Vowel production in Winnipeg

Robert Hagiwara

University of Manitoba

Abstract

General properties of the Canadian English vowel space are derived from an experimental-acoustic study of vowel production underway in Winnipeg, Manitoba. Comparing the preliminary Winnipeg results with similar data from General American English confirm previously described generalizations for Canadian English: the merger of low-back vowels, the relative retraction of /æ/ and the relative advancement of /u/ and /ʊ/. However, a similar comparison of the Winnipeg sample with comparable southern California data dispute the accuracy of the so-called Canadian Shift (Clarke et al. 1995) as features of ‘general’ Canadian and Californian English. The utility of acoustic analysis in uncovering subtle phonetic distinctions is further revealed in a discussion of Canadian Raising. The Winnipeg speakers produce a directional shift in both the nucleus and offglide of the diphthongs /ai, ao/, rather than just adjusting the height of the nucleus. This process applies to all three diphthongs (including /oi/).

1. Introduction

‘General’ Canadian English (assuming that such a thing exists) does not seem to have the same kind or degree of regional variation that some other ‘national’ varieties of English have. Nonetheless, the perception that Canadian English has no significant variation at the phonetic/phonological, or ‘accent’, level is diminishing. Several previous studies have identified variables which have some geographic or ethno-social ‘tuning’, e.g. the quality of the low vowel nuclei of the /ao/ and /ai/ diphthongs (Boberg 2005b; Hung et al. 1993). However, it remains difficult to compare the whole vowel system across existing studies, due in part to the wide range of methodologies employed in studying specific aspects of the vowel system by different researchers in various contexts. What is needed is some common frame of reference.

In part to address this problem, several colleagues and I have been working on using standardized acoustic experimental methods to facilitate comparisons across speakers and dialects, with the intention of using such findings as ‘baselines’ or points of reference to describe other kinds of variation—whether due to prosody, segmental context, register, rate or dialect. Acoustic analysis is now almost commonplace as a companion to traditional methods of reporting vowel production, such as impressionistic transcription. While using acoustic measurements may increase the precision of the analysis, it remains difficult to compare results derived from different studies, due to the differences in data collection, speech styles and materials collected, and even some of the analytical techniques employed. Thus a standard method is crucial to achieving maximum comparability.

This paper presents a first look at results from such an acoustic-experimental study of vowel production in Canadian English. First in Section 2, I will describe the goals and method employed in the larger project and the current subset of speakers. Then in Section 3, I will discuss the characteristics of the vowel space used by these speakers, comparing them to the presumed ‘General American’ standard (Peterson and Barney 1952) and to a comparable
dataset from southern California English. The results from these comparisons will be discussed with reference to patterns reported for Canadian English, especially the proposed “Canadian Shift” (Clarke et al. 1995). In Section 4, I will use the positions of the simplex vowels in the vowel space as a backdrop for describing the time course of the diphthongs /ai, ao, oi/. This will allow me to offer an acoustic characterization of the “raising” pattern seen in the Winnipeg sample, and discuss some of the implications of the results for our understanding of the process.

2. The Winnipeg vowels project

The data presented here are derived from an ongoing study of English and French in and around Winnipeg, Manitoba. It is designed as an experimental study, along the lines of classic studies of vowel acoustics (Hillenbrand et al. 1995; Peterson and Barney 1952), and generally following the recommendations put forward in Hagiwara et al. (1999) for the creation of a comprehensive database of such results, but the general approach can be appended to any study as an adjunct to whatever other methods are used.

The goal of this project is to produce quantifiable acoustic baselines for the description of vowels in Winnipeg, both for their own sake and to provide a standard of comparison for the quantification of phonetic variation. For this paper, however, I will confine my discussion to general properties of vowel dispersion in the acoustic space, leaving the issue of quantification aside for future work.

The data presented in this paper are derived from the speech of ten monolingual English speakers (five women and five men), 18-25 years of age. They are natives of Winnipeg, and children of natives of Winnipeg. No attempt has been made at this time to control for possible Winnipeg-internal ethnic, geographic, or cultural dialectal variants. While these may turn out to be interesting, the goal here is to provide an overview of ‘general’ Winnipeg English vowel production rather than specific variants.

Table 1. Target words used in the Winnipeg Vowel Project script.

<table>
<thead>
<tr>
<th>Vowel category</th>
<th>/hVd/</th>
<th>/hVt/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>heed</td>
<td>heat</td>
</tr>
<tr>
<td>/ɪ/</td>
<td>hid</td>
<td>hit</td>
</tr>
<tr>
<td>/e/</td>
<td>aid</td>
<td>ate</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>head</td>
<td>pet</td>
</tr>
<tr>
<td>/æ/</td>
<td>had</td>
<td>hat</td>
</tr>
<tr>
<td>/ɑ/</td>
<td>odd</td>
<td>hot</td>
</tr>
<tr>
<td>/ɔ/</td>
<td>hawed</td>
<td>ought</td>
</tr>
<tr>
<td>/o/</td>
<td>ode</td>
<td>oat</td>
</tr>
<tr>
<td>/u/</td>
<td>hood</td>
<td>put</td>
</tr>
<tr>
<td>/u/</td>
<td>who'd</td>
<td>hoot</td>
</tr>
<tr>
<td>/ʌ/</td>
<td>Hudd</td>
<td>hut</td>
</tr>
<tr>
<td>/ɑ/</td>
<td>herd</td>
<td>hurt</td>
</tr>
<tr>
<td>/ai/</td>
<td>hide</td>
<td>height</td>
</tr>
<tr>
<td>/ao/</td>
<td>how'd</td>
<td>out</td>
</tr>
<tr>
<td>/oi/</td>
<td>Boyd</td>
<td>Hoyt</td>
</tr>
</tbody>
</table>
Speakers for the study were audio-recorded reading from a script containing target words that illustrate the 15 potentially contrastive English vowel phonemes in a standard context. Where possible, real monosyllabic words of the form /hVd/ and /hVt/ were used. Where real (or easily spelled) words of the appropriate shape were not available, some substitutions were made. The words used in the Winnipeg script are listed in Table 1. Each target word is presented in a frame (“Say ___ once”), and appears in the script five times in random order.

Recordings were made in the Linguistics Laboratory at the University of Manitoba on high-quality digital recording equipment in a sound-attenuated room. The data were digitized at 44.1 kHz, and transferred to the computer. All analysis was done using Kay Elemetrics MultiSpeech (Model 3700 Version 2.3) software. Vowel durations were determined by taking the difference between the vowel start point (the time at which periodic energy in F2 begins) and the vowel end point (the time of the onset of closure, if visible, or the last regular period of voicing in F2, if not). Timepoints for further measurement were established at 25%, 50% and 75% of vowel duration (hereafter referred to as timepoints 1, 2, and 3). At each of these timepoints, the frequencies of the first four formants were measured by simultaneous evaluation of wide-band spectrograms, LPC formant histories, and LPC and narrow-band FFT slices taken at the measurement point.

While the ultimate goal of the project is to quantify inter- and intra-category variation, the data from the ten speakers under discussion here are probably not enough to provide any conclusive results. I prefer at this time to offer a more impressionistic account of these speakers’ vowel systems derived from the acoustic analysis.

To do this, I will present diagrams, similar in form to the standard vowel space diagrams but derived from the acoustic measurements rather than impressionistic placement of symbols in the vowel trapezoid. To further the approximation of the perceptual distribution of vowels in the acoustic space, I have used a technique I call “coarse auto-normalization” to derive auditory distances from the underlying values in Hz. In short, the F1 and F2 frequencies of each token are converted into an auditory distance from ‘neutral’ first and second resonance frequencies. The neutral resonances for each speaker are estimated by pooling that speaker’s data from all timepoints for the plain vowels (excluding the diphthongs and syllabic /r/), and deriving a linear slope-intercept formula by regressing formant number by frequency. The resulting slope-intercept formula should approach Hz = 1000r – 500 (for adult male speakers) or Hz = 1200r – 600 (for adult female speakers), where r stands for resonance (formant) number. These reproduce the neutral resonances of vocal-tract length tubes of uniform diameter (Chiba and Kajiyama 1941; Fant 1960). It is beyond the scope of the present paper to compare this technique with the many other possible normalization techniques. However, I believe coarse auto-normalization may have some advantages. By using all the available formant frequency information across a large subset of vowels, upper formants can normalize for lower formants and individual speakers can normalize for themselves.

Using this technique, the F1xF2 position of a vowel in the space can be expressed for each individual in terms of its distance from his or her own calculated neutral, and these can be plotted in a form which visually mimics the familiar vowel space diagram. For reference purposes, I have provided in Table 2 the underlying first and second formant averages, in Hertz, from the present study as well as those from the two other studies discussed below.
3. The Winnipeg vowel space

Figure 1 represents the men’s and women’s vowel averages for the current subset of the Winnipeg English study. These are derived from the F1 and F2 measurements at timepoint 2, using coarsely auto-normalized distances as the coordinates in the space. The crosshairs in the figure represent the coordinates of the neutral F1 and F2 (i.e. the origin point of the coordinate system). The large symbols (in squares at the top of the figure for the men and in circles at the
bottom of the figure for the women) are located at the coordinates of the average F1 and F2 for each vowel category in a coarsely auto-normalized space.

For comparison, the vowel centres from Peterson and Barney (1952) representing the “General American” vowel system are included as small IPA symbols. The underlying data from the Peterson & Barney (1952) study were not available. Neutral frequencies were calculated from the averages of three formants of the vowels reported in that study, and coarsely auto-normalized. This should not make a difference at the present level of visual comparison of the vowel spaces.

The ‘point’ vowels (/i, æ, ɑ, u/) are joined by a line in all four cases (solid for the Canadian vowels and dotted for the American vowels) to provide additional visual reference.

The Canadian and American vowels, on average, seem to occupy similar ranges in the auditory space, although the highest vowels seem to be slightly higher in the General
American space (indicating greater deviation from neutral). Obviously, as closely related dialects of English, a great deal of similarity is expected. However, there are obvious differences.

First, the General American vowel system clearly distinguishes the /ɑ/ and /ɔ/ categories, where these are merged, or at least greatly overlap, in the Canadian space. The merger of low-back vowels is regarded as characteristic of Canadian English, as well as a broad dialect range of American dialects (Labov 1991). I will return to this point with reference to the Californian space below.

Another difference between the Canadian and General American vowels is the relative advancement of the /u, ʊ/ and /ʌ/ categories for the Canadians. Interestingly, Canadian /o/ remains a back, round vowel, of similar backness to the General American /ɔ/ category. Peterson and Barney (1952) did not include the higher-mid vowels /e/ and /o/, in their study, so direct comparison is not possible.

Retraction of /æ/ has been reported in Canadian English, both as part of the Canadian Shift (Clarke 1993), and independently as a phenomenon observed in Vancouver English (Esling and Warkentyne 1993). This is also seen in the Winnipeg sample. It appears also that /æ/ is slightly lowered in the Winnipeg sample, although this appears to apply to the other front vowels as well, at least for the men, and thus seems to confirm the presence of Canadian Shift (Clarke et al. 1995).

Canadian Shift was specifically proposed to involve a drag-chain: triggered by the merger of the low-back vowels, the /æ/ category begins to retract into the available low-central space, with lowering (and retraction) of the front lax vowels resulting from that. However, the lowering seen in the /æ/ category in the Winnipeg sample seems to be something occurring to the front vowels generally, if at all, including tense /i/. What appears to be happening instead is that the increased distance in the backness dimension between /i/ and /æ/ results in a redistribution of backness values for the front lax vowels. Note that for the men in both samples, the front lax vowels fall on the line between the /i/ and /æ/ centres. There is some indication of additional retraction of /i/ and possibly /ɛ/ among the Winnipeg women, which might be expected if the Canadian Shift were incipient in this sample. However, it also appears that /ɪ/ is, if anything, more retracted than /ɛ/, which is not explained by the proposed chain-shift.

If Canadian Shift is to be interpreted as involving lowering of the front lax vowels relative to the height of /i/, it does not appear that such lowering is taking place in the Winnipeg sample. The implications of this with respect to the Canadian Shift will again be addressed below in comparison with the California vowel system.

To summarize the comparison with General American, the Canadian vowels appear to exhibit the following differences from the American ones:

- Merger of the low-back vowels, resulting in a higher-low vowel of similar backness to /a/<br>
- Relative advancement (centralization) of /u, ʊ/ and /ʌ/<br>
- Some retraction and lowering of /æ/, and some redistribution of the front lax vowels
In the original proposal regarding the Canadian Shift (Clarke et al. 1995), striking similarities with the Californian system of vowels were observed. However, that comparison was based on impressionistic reports of Californian vowels made by a different set of researchers (Hinton et al. 1987; Luthin 1987), I am fortunate in that I have access to data previously reported for southern California (Hagiwara 1995, 1997), collected using similar experimental-acoustic methodology as the Winnipeg sample. In order to facilitate comparison, I have re-digitized the Californian data from the original recordings and re-measured the /hVd/ tokens from the Californian study using the current techniques. Figure 2 shows the Winnipeg vowels again, but this time compared with coarsely auto-normalized F1 and F2 of the Californian vowels at timepoint 2. The similarities are indeed striking, but there are important differences as well.

Like the Winnipeg vowels, the California vowels show advancement of the back vowels. However, it is clear that /o/ is participating in this advancement in California in a way that is not seen in the Canadian sample.

Figure 2. Men’s (top) and women’s (bottom) vowel centres, as in Figure 1. Small symbols represent the centres from Southern California.
While the merger of low-back vowels is complete in both dialects, the resulting vowel is different. In the Canadian space, the merged vowel is usually rounded (or at least acoustically further back), and higher in the space than the /æ/ category. The analogous vowel in Californian English is definitely the lowest vowel in the system, lower than /æ/ and not at all rounded. There is little evidence of lowering (relative to /i/) of the front lax vowels in either dialect. Interestingly, the front lax vowels for the Californian men seem to be significantly more retracted, relative to the other front vowels, than for either the Californian or Canadian women.

As originally proposed, the Canadian Shift was suggested to be common to both Canadian and California vowel systems, triggered by the merger of the low back vowels /ɑ/ and /ɔ/. The acoustic comparison provided here casts doubt on the generality of Canadian Shift in two ways. As discussed earlier, the proposed shift does not seem to describe the Winnipeg vowels, and thus is not a general feature of varieties of general Canadian English (I leave aside the interesting question whether the Winnipeg or Ontario samples should better represent ‘general’ or ‘standard’ Canadian English). Secondly, while retraction (if not lowering) of the front vowels seems to be occurring in the Californian sample, this cannot have been triggered by the merger of low-back vowels in the fashion of a traditional drag-chain. The merged low-back vowel in Californian does not leave the void in the low-back space it does in the Canadian space.

Contra Clarke et al. (1995), other researchers have suggested that lowering of the front lax vowels is a less conspicuous feature of Canadian dialects than retraction (Boberg 2005a; De Decker 2002; Hoffman 1999). The present data suggest that even this retraction is at best incipient in the Winnipeg sample, especially compared to the Californian sample. The present subset does not allow us to test whether the apparent retraction of /i/ among the Winnipeg women represents the beginning of a Montreal-style retraction (Boberg 2005a). An older cohort of speakers (and more speakers in general), planned as part of the larger Winnipeg project, will allow for some refinement of this point in the future.

There is, however, an additional similarity observed in Clarke et al.’s data and the Winnipeg and California samples that may involve a chain shift. I surmise that the low-back merger may be the trigger of a push-chain, involving the advancement of the /ʌ/ category. This might be crowding the lower-front space, resulting in lowering of the /æ/ category. Retraction of /æ/ may simply follow as a secondary result of lowering a front vowel. A similar proposal was made by Hoffman (1999).

In sum, the Canadian Shift, as described by Clarke et al (1995), does not seem to characterize either the Winnipeg or the Southern California samples, and the present data call into question the explanation offered for the similarity between the two dialects. However, some of the patterns seen in Clarke et al’s study, the lowering and retraction of /æ/, and the advancement of /ʌ/ along with some of the back vowels, appear to characterize both dialects. From the present data, it cannot yet be determined whether or not there is a ‘Canadian Shift’ of one sort or another going on in Winnipeg English, or whether the Canadian and Californian systems in fact involve similar processes of change.
4. About Canadian Raising

No discussion of vowel production in Canadian English would be complete without some mention of Canadian Raising (Chambers 1973; Joos 1942), arguably the most salient phonetic marker of Canadian English.

Figure 3 represents the production of diphthongs in the Canadian sample, compared with the monophthongs. As these sounds must be modeled dynamically, I have included data from all three timepoints. The large vowel symbols are located at the coordinates of the coarsely auto-normalized F1 and F2 at timepoint 1, the arrowhead represents the coordinates at timepoint 3, and the angle in between the coordinates of the vowel midpoint. The path of each diphthong in the voiced /d/ context is represented by a solid line, and the path with the following voiceless /t/ (the result of Canadian Raising) by a dotted line.

Looking first at the /aɪ/ and /aʊ/ diphthongs in the longer, voiced consonant-environment (solid lines), we can observe that on average, the diphthongs start low and central. The /aɪ/
diphthong then proceeds forward and up in the space, the /ao/ diphthong up and back. The nuclei of the shorter, ‘raised’ diphthongs are in fact raised in the space to approximately the height, but not the backness, of /ʌ/. However, it is not simply raising of the nuclei that has taken place, but of the entire vowel trajectory. The entire diphthong appears to have ‘advanced’ along the path of the transition in the ‘raising’ context. Whether this is merely the product of vowel shortening or indicates a global shift in the featural ‘targets’ cannot be determined at this time. It is perhaps worth noting that a similar auditory distance seems to be covered in both cases, even though the ‘raised’ diphthongs are on average 30% shorter than their unraised counterparts.

The result is that the initial portion of the ‘raised’ diphthong has the height (but not the backness) of the /ʌ/ category, but not because of a phonological process acting specifically on its height. The backness of the ‘nucleus’ seems to depend on the general trajectory of the movement. Thus I suggest that Canadian Raising should not be seen as a change in the categorical specification of the nucleus, but as a ‘shift’ in the portion of the trajectory being realized (and the speed at which it moves).

This new characterization of Canadian “Raising” as advancement along the path of movement also describes the variants of the /oi/ category seen in Figure 3. Both the men’s and women’s data show that in the non-‘raising’ environment (with following /d/), the /oi/ diphthong begins in region of [o], and then moves forward towards the /i/ region. In the ‘raising’ environment (with voiceless /t/), the entire trajectory of the diphthong has moved forward in the space, apparently preserving auditory distance covered and the path of movement. Instead of ‘raising’, /oi/ ‘fronts’ or ‘advances’.

These findings are consistent with recent proposals regarding the voicing/length correlation and the production of diphthongs (Moreton and Thomas 2004; Thomas 2000). According to these proposals, shortening a diphthong tends to increase the excursion of the offglide relative to the nucleus (probably due to increasing the speed of the transitional movement), while lengthening tends to reduce the distance of the nucleus relative to the glide. While proposed primarily with respect to variation in reflexes of historical /ai/, the general scenario holds here for all three diphthongs. It will be interesting to see whether there are degrees of variation in Canadian English that can speak to the historicity of Moreton and Thomas’s proposal in Canadian English.

One of the advantages of acoustic measurements is that it allows us to detect more subtle distinctions than can easily or reliably be detected by ear. I am not convinced this difference in the realization of the /oi/ diphthong is the sort of thing that would have been noticed simply by careful listening, or that the general similarity with the output of “Canadian Raising” would be apparent if it were not for the direct acoustic characterization of the process. These findings also suggest different avenues of approach for experimental study of production and perception of diphthongs and other sounds in Canadian English: Can these differences be perceived? How do they relate to the speed and distance articulators must travel? How do raising and non-raising dialects differ with respect to the effect of the voicing/lengthening correlation?
5. Conclusions

To summarize, this subset of the Winnipeg sample appears to confirm the similarity between Canadian English vowels and the system found in Southern California (Clarke et al. 1995). In spite of these superficial similarities, however, significant differences exist between the two systems, differences which might only have been clearly revealed by acoustic analysis of directly comparable (experimentally controlled) datasets. Similarly, acoustic analysis of diphthong production in the Canadian sample reveals a very different conception of Canadian Raising than the traditional phonological rule would have suggested.

Obviously this work is in its infancy and much further research is required. The study of the dynamic aspects of all of the vowels requires expansion, and the details of the proposed view of Canadian “Raising” need to be worked out. The distribution of vowel categories in the space and the distribution of vowel productions within category must be explored more fully. Of course, it is critical to determine the degree to which the norms of a specific population (such as the young Winnipeggers discussed here) represent ‘general’ Canadian English.

As the Winnipeg Vowels Project is expanded, these and other features of the Canadian vowel space can be explored and quantified. As more speakers are added, I hope that we will be able to look at these features in different age cohorts and quantify the effect of age or other socio-demographic factors on vowel production. Similarly, I hope these data will allow for the quantification of other variables in vowel production, such as occur at different rates, in different registers, and of course other phonological contexts. Finally, I hope that this work will facilitate the quantification of comparisons with other languages and varieties of (especially Canadian) English, and help bridge the gap between traditional sociolinguistic studies and experimental linguistics.

Acknowledgments

I wish to thank the organizers and participants of the Canadian English in the Global Context conference for providing the perfect forum for the first presentation of this work. I have benefited greatly from the discussions at and following the conference. I want to thank Charles Boberg, Paul DeDecker, Michol Hoffman, Jennifer Ingle, and Eric Thomas for helping to clarify my thinking on a number of points addressed here. For their insightful and meticulous comments on various drafts of this manuscript, I wish to thank Daniel Bérubé, Keren Rice, Kevin Russell, and two anonymous reviewers. Any remaining errors or misapprehensions are entirely my own responsibility.

References


