2pSC23 Revisiting the Canadian English Vowel Space

Robert Hagiwara
University of Manitoba

The Winnipeg Vowels Project and method

This presentation draws from a subset of data gathered as part of the Winnipeg Vowels Project, currently under way at the University of Manitoba. The goal of the project is to study vowel production and variation in local English and French. As a first step, I am trying to establish 'baseline' frequencies for the formants of the vowels to serve as a standard of comparison with other dialects, languages, or variants.

The present subset of speakers are 10 monolingual English speakers (5 men and 5 women), 18-25 years of age, all natives of Winnipeg, and children of monolingual, anglophone natives of Manitoba. Possible ethno-social variation within the heterogeneous Winnipeg community will be left for future research. These ten speakers represent a typical subset of the general population.

These speakers were digitally recorded in a laboratory setting, using methods based on recommendations from Hagiwara et al (1999, "An interactive atlas of English vowels: Design considerations, JASA 106(4.2): 2243). Each read a script containing target words of /hVd/ or /hVt/ shape, in the frame "Say ___ once.". (Real words were used, some substitutions to the target shape were necessary.) Each target word appeared five times in the randomly ordered script.

Digital audio recordings were made at 44.1 kHz (16-bit) on professional quality equipment. Once transferred to the computer, working files were downsampled to 22 kHz. All analysis was done using Kay Elemetrics MultiSpeech. Durations were determined from vowel start point (onset of periodic energy in F2) and end point (last glottal pulse before consonant closure). From these, three timepoints (25%, 50%, and 75% of vowel duration) were located for formant measurement.

The frequencies of the first four formants were measured using simultaneous evaluation of wide band spectrograms, LPC formant histories, and FFT and LPC slices taken at the timepoint under consideration. (Analysis parameters were adjusted to produce the best results for each speaker.) The resulting formant measurements were subjected to the Coarse Auto-normalizaton procedure (see below), and the results plotted as the diagrams seen here.

This poster is concerned primarily with overall properties of the space at vowel midpoint, and some observations about the timecourse of the vowels. Comparison of the vowel midpoint data with similar data from General American and Southern California English are currently in press at *Canadian Journal of Linguistics*. Statistical and quantitative analyses are in preparation.

Coarse auto-normalization and these vowel diagrams

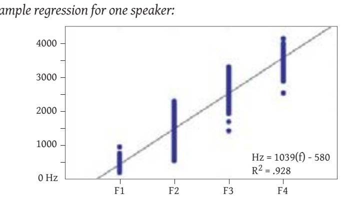
For each speaker, a regression (formant number x frequency) is calculated from all available plain-vowel formants and timepoints. The slope-intercept formula is taken to represent that speaker's neutral resonance curve. These should approximate *Hz = 1000(f) - 500* for men and *Hz = 1200(f) - 600* for women.

All measurements are converted to an auditory scale (Bark), and the value of any formant frequency is expressed as an auditory distance (ΔBark) from that speaker's presumed neutral. Since this procedure allows each speaker's data to normalize for itself, I call this **coarse auto-normalization**.

Coarse auto-normalization is suitable for relatively balanced datasets such as are gathered experimentally, and does not rely on the accuracy or precision of any one formant measurement of any particular vowel category.

In this poster, the coarsely auto-normalized results for F1 and F2 are used as coordinates in plotting the vowel diagrams. The result resembles the placement of vowels in a traditional vowel diagram with height and backness/rounding dimensions.

Sample regression for one speaker:



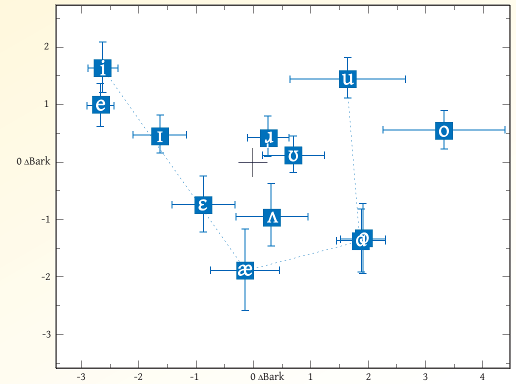
Resulting calculated neutrals:	F1	F2	F3	F4
	459	1498	2537	3576 Hz
	4.4	11.2	14.6	16.6 Bark

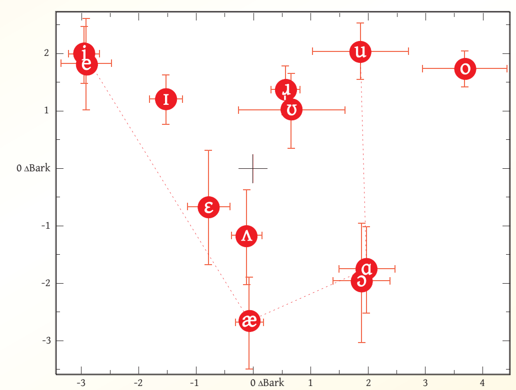
Coarsely auto-normalized coordinates of [i]:

X coordinate	Y coordinate
F2 [i] = 2132 Hz	F1 [i] = 213 Hz
= 13.5 Bark	= 2.1 Bark
= -2.3 (11.2-13.5) ΔBark	= 2.3 (4.4-2.1) ΔBark

The Canadian English vowel space

These figures represent the coarsely auto-normalized F1xF2 averages for the men (in blue) and the women (in red) at vowel midpoint (50% of vowel duration). Error bars indicate ±1 SD from the mean.

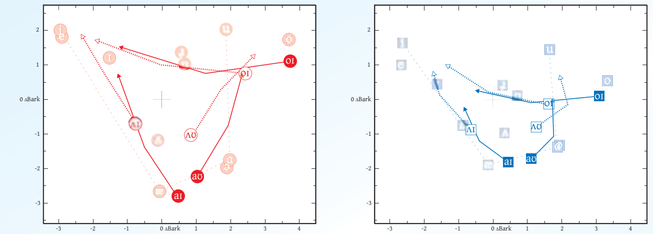




General properties of the space:

- ▲ Functionally complete merger of low-back vowels /ɑ-ɔ/
- ▲ Advancement of /u, ʊ, ʌ/ relative to /o/
- ▲ Some retraction of /æ/ and re-adjustment of front vowels
- ▲ Few non-trivial gender differences
  - ◆ women's vowels distributed more distantly in the space (i.e. are more peripheral)
  - ◆ in particular, women seem to use more of the height dimension than men

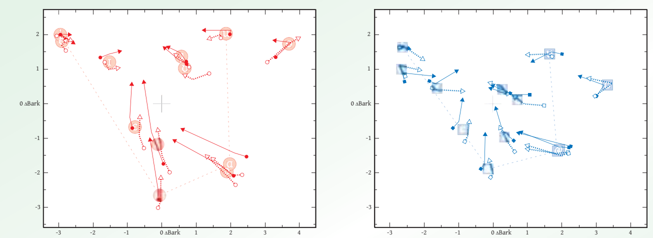
The truth about Canadian "Raising"



Solid arrows represent the longer /hVd/ context, while dotted arrows represent the shorter /hVt/ context (i.e. the environment of Canadian Raising). Three timepoints are represented for each diphthong. Phonetic symbols are located at the F1xF2 position at timepoint 1 (25%), the 'bend' at the vowel midpoint, and the arrowheads at the position at timepoint 3 (75%). The women's data are in red, the men's in blue. as in the main figures. For reference, the positions of the monophthongs are included in the background.

- ▲ Similar auditory distances covered, despite average 30% difference in duration
- ▲ 'Visible' portion of vowel 'advances' in the direction of movement, as a function of environment
  - ◆ Non-categorical shift of nucleus to the height of /ʌ/ (but not its backness)
  - ◆ Similar shift in midpoint and endpoint positions
- ▲ Same process applies to /oɪ/ diphthong, resulting in 'fronting' rather than 'raising'
- ▲ The above don't generally describe the behaviour of the monophthongs; this process is distinct from shortening/hiding typical of voiceless environment

A closer look at the monophthongs



These views show the dynamic information for the monophthongs. Similar key as for the dipthongs (above). Solid arrows represent the path of the vowel in the longer /hVd/ context; dotted arrows represent the shorter /hVt/ context.

- ▲ In general, monophthongs have very little movement in the first half, suggesting a relatively steady state followed by long, slow transition
- ▲ Different 'strategies' for duration adjustment (cf. diphthongs, above)
  - ◆ Lower vowels show much shorter movement in shorter environment, suggesting gestural hiding
  - ◆ Higher vowels show similar auditory distances traveled in the two environment, suggesting stiffness/velocity adjustment
- ▲ Except for women's /o/ in the voiceless environment, none of the 'tense' vowels moves in the 'expected' direction

Contact information:
Rob Hagiwara
Linguistics Department, University of Manitoba, Winnipeg, MB, CANADA R3T 5V5
robh@cc.umanitoba.ca
www.umanitoba.ca/linguistics/robh/

ASA 149, Vancouver BC