5aSC22 — Monophthongs and formant movement in North American English

Overview

This poster represents a summary of several exploratory analyses of databases of acoustic measurements of vowel production in two studies of similar dialects of North American English: Southern California and Winnipeg, Manitoba. These dialects are phonologically similar, exhibiting features of Labov's (1991) 'third dialect' group (complete or near-complete merger of /a/-/ɔ/, lack of general tensing of /æ/ and some centralization of /u/) characteristic of most of Canadian and western American English).

In investigating these and other features, however, systematic differences between the output of these historical processes (e.g. different realizations of the merged low-back vowel) suggests that these dialects may not share the sociophonetic history suggested by their phonological similarity (Hagiwara 2005, in press; cf. Clarke et al. 1995).

The data analyzed here represent background data for other work--the Californian data was collected original to provide a backdrop for the analysis of the formants of /r/. The Canadian data is part of a larger project on vowels in Canadian English and French.

This poster's primary goal is to investigate patterns of movement in the formants of the phonologically simplex vowels, both for descriptive/comparative completeness, and for investigating the degree to which these vowels can properly be regarded as phonetic monophthongs, ultimately with the intention of shedding light on the proper phonetic representation of formant movement in vowel production.

Not surprisingly, none of the vowels in the database can be described as 'flat', i.e. with very long, very steady formants throughout most of their duration. However, some are flatter and some more dynamic at different points of the vowel. Trends in the patterns of dynamism seen in the two datasets which will suggest avenues of approach for the representation of vowels, and the relation between presumed steady-state targets, controlled vs. simply transitional movement, and related issues which should be tested further in less-formal speech contexts.

Speakers and method

The present speakers were all young adults (18-26 years) at the time of recording. Californian speakers were recorded between 1992 and 1994 in Los Angeles. Manitoban speakers were recorded between 2002 and 2004 in Winnipeg. There are fourteen Californians (nine women, five men) and ten Canadians (five women and five men). The Californians have been described elsewhere (Hagiwara 1995, 1997), but those reports were based on single-point (steady state of F2) measurements, and on a combination of three phonetic contexts.

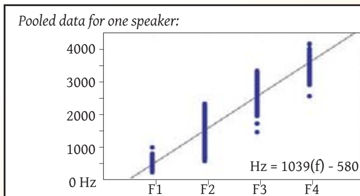
The Californian and Canadian data for the present work represents only /hVd/ (Californian) and /hVd/-/hVt/ (Canadian) tokens. Multiple repetitions of each vowel in standard context were recorded by each speaker from a randomized script.

F1 through F4 frequencies were measured for each token at three points--25%, 50%, and 75% of vowel duration (see Hagiwara 2005a, 2005b, for further discussion).

The raw data for each speaker was subjected to Coarse Auto-normalization, in which a given formant frequency is expressed as a distance (in Bark) from a neutral formant frequency calculated independently for each speaker. (Sidebar adapted from Hagiwara, 2005b.)

This poster is based on analyses of coarsely auto-normalized data only.

Pooled data 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
F2 [i] = 2132 Hz
= 13.5 Bark
= -2.3 (11.2-13.5) ΔBark

Y coordinate
F1 [i] = 213 Hz
= 2.1 Bark
= 2.3 (4.4-2.1) ΔBark

General properties of the Californian and Canadian vowel spaces

In keeping with general description, both populations show general properties of Labov's (1991) 'third dialect' group, but with differences in phonetic realization (see Hagiwara 2005a, in press, for further discussion).

- Both dialects have non-tensed /æ/ in the /hVd/ context. However, the Californian /æ/ is significantly lower than its Canadian counterpart.
- Both dialects have merged low-back vowels, but the result is a much lower (and slightly more central) vowel in Californian than Canadian.
- Both dialects centralize /u/. The Californians tend to unround /u/, and the centralization trend seems to be general to the back vowels. However, the Canadians seem to centralize without unrounding /u/ (at least to the same degree), and the centralization trend seems to exclude /o/.

Both dialects share a gender-effect whereby women use a greater range of the coarsely auto-normalized F1 space than the men. This expanded F1 range is distinct from the F2 range, which is not significantly different for gender. (ANOVA of pooled data, Pop x Sex x Timepoint: F1 (Sex) F=49.3 p<.0001.) There is a main effect for Population in F2, with the Canadians using an expanded F2 range relative to the Californians. (F2 (Pop) F=41.7 p=.0003).

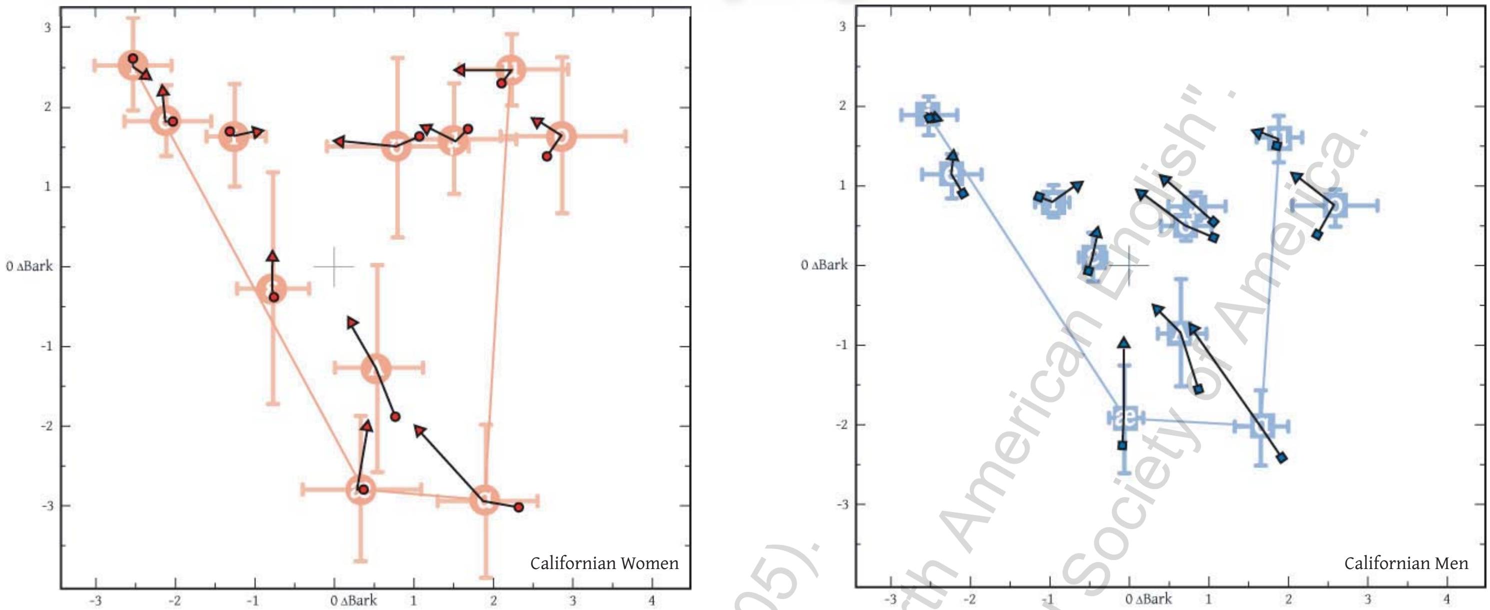
Both formants show a main effect for Timepoint, there were no significant intearactions (at the .001 level). Scheffe's post-hoc tests show that for F1, the 75% timepoint is significantly different from either of the first two. For F2, the only significant difference is the 25%-75% comparison.

I interpret this as suggesting that the vowel space between the first two timepoints is very similar, but overall the second half of the vowel (any vowel) involves a degree of 'reduction' in the space.

This must be confirmed for all vowels separately, and obviously needs to be explored in a phonologically less-restricted dataset. However, taking the vowel space and vowel production as a whole, it looks like some kind of targeted steady-state (or directed movement) must be located in the first half of the vowel. In the second half, vowel-specific control may ease. I surmise that the resulting movement in the second half of the vowel is transitional, perhaps interpolated between the vowel target and the upcoming consonant transition.

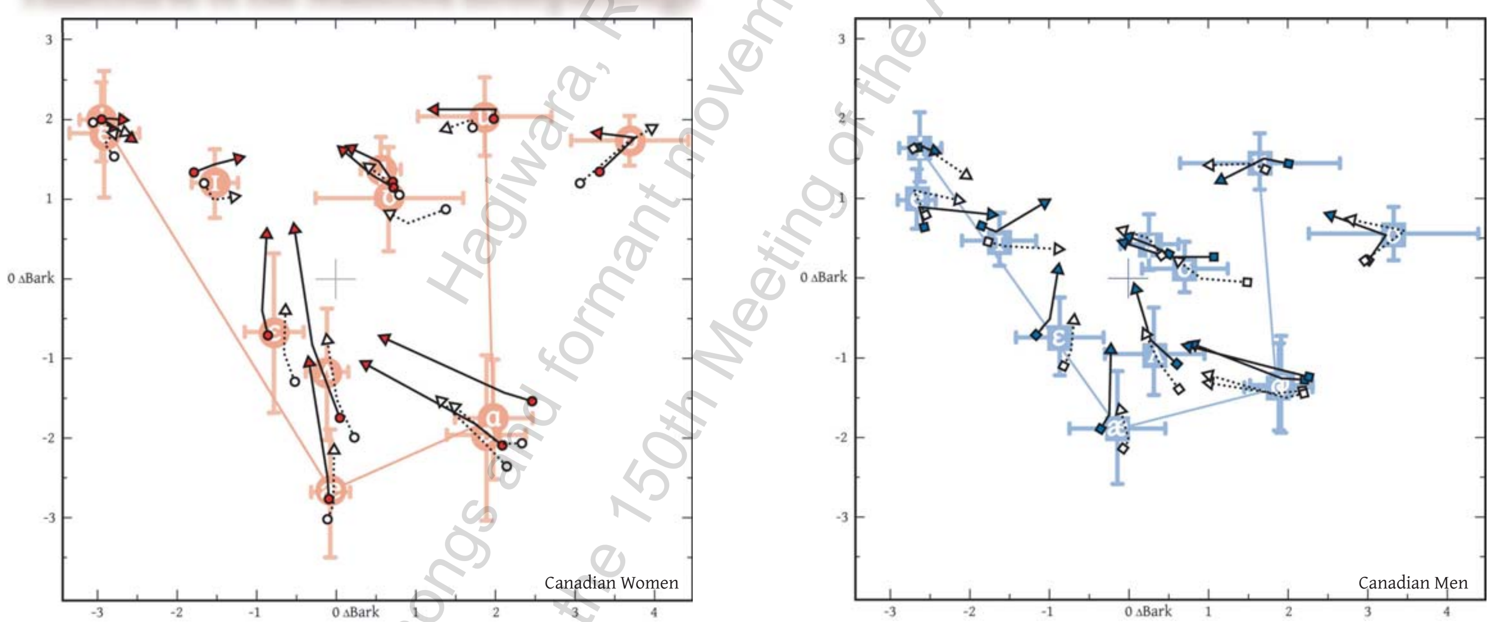
The 'actual' consonant transition, involving a very sharp change in formant slope, does not occur until very late in the vowel. The 'reduction' seen at the 75% timepoint may thus be at least partially independent of the following consonant.

Timecourse of the Southern Californian monophthongs



Women's data (left, in red) and men's data (right, in blue). Large symbols represent the coarsely auto-normalized positions of the formants at the vowel midpoint. Bars indicate ±1SD of vowel midpoint. Arrows plot the course of the vowels. The path of the arrow begins at the coordinates for the 25% timepoint, bends at the vowel midpoint, and ends at the 75% timepoint.

Timecourse of the Manitoba monophthongs



Women's data (left, in red) and men's data (right, in blue). Large phonetic symbols represent the coarsely auto-normalized positions of the formants at the vowel midpoint averaged across voicing contexts (/hVd/ /hVt/). Bars indicate ±1SD of vowel midpoint. Dotted/open arrows plot the course of the vowels in the /hVt/ context, and solid/filled arrows the /hVd/ context. The path of the arrow begins at the coordinates for the 25% timepoint, bends at the vowel midpoint, and ends at the 75% timepoint.

Reduced-format image of this poster (in PDF format) available from the ASA conference website, or from the author's personal website: http://home.cc.umanitoba.ca/~robh/files/asa150.pdf

Additional information from the author by e-mail:
robh@cc.umanitoba.ca

Robert Hagiwara
University of Manitoba

Are the tense vowels diphthongs?

Not in the conventional sense of having /ɪ/ or /ʊ/ off-glides.

Considering first the mid vowels, the movement between the 25% and 50% timepoints in these two vowels is consistent with the direction of an appropiate offglide. However, this movement does not seem to be sustained in the second half of the vowel (i.e. to the 75% timepoint). The movement seen at the end of these vowels (with the possible exception of the Canadian women's "oat" productions) appears to follow the general trend of transitioning toward the following consonant in the second half.

The /u/ vowel doesn't even move in the expected direction in the first half, and the distance moved in the first half of the front tense vowels /i, e/ is not significant.

Taken as a group, most of the vowels' paths seem to point toward a close F1 value (consistent with following closure) and an F2 value just higher (slightly front) of neutral. This is consistent with these data's following coronal plosive and its 'loci' of transition.

While the paths seen here are derived from the timepoints measured (25%, 50% and 75% of vowel duration), the last timepoint (for most tokens) was well before the beginning of what I would have identified as the closure transition (i.e. the point at which the 'slope' of the formant changed sharply heading into the closure. So while not strictly 'just' transition to the consonant, I'd suggest that these values do not indicate a late 'target' in the vowel but 'relaxation' and interpolation between the vowel target (realized somewhat earlier) and the upcoming consonant transiton.

Obviously this will need to be explored further with a more varied dataset.

Do the lax vowels /ɪ, ɛ, ʊ/ transition towards schwa?

Not if we take 'schwa' to mean the neutral, unshaped vowel. The Coarse Auto-normalization algorithm starts with the assumption that the neutral resonances of a uniform tube (i.e. schwa) can be inferred from the vowel production of an individual speaker. 'True' schwa should have the neutral resonances of the speaker's vocal tract, and thus become the origin point (⟨0,0⟩) of the vowel plot in the coarsely auto-normalized space. On the charts at left, this point is indicated by the small crosshair.

So the idea that lax vowels transition towards schwa (while the tense vowels transition toward the periphery) is open to interpretation here. It seems to me all the vowels generally tend toward formant values at the 75% timepoint which are consistent with movement towards the upcoming consonant closure. Taken as a group, the vowel paths seem to point towards some ultimate target which is very high/close and just front of center. I have interpreted this as consistent with the 'locus' of the upcoming consonant transition. It is equally consistent with some kind of 'reduced' vowel (such as [ɨ]), but this this vowel is not 'true' schwa, and this pattern is not characteristic of the lax vowels in particular.

What about those long movements in the lower vowels?

I'm not sure. No one (as far as I know) has suggested that either dialect as distinctive schwa-like offglides in the low vowels (except in /æ/ for some speakers).

In general, the vowels show very little movement between the 25% and 50% timepoints. This is especially true for most of the lower vowels. The second half of the vowels (between the 50% and 75% timepoints) shows much longer travelling distance, and for the most part this movement is consistent with an ultimate 'target' predicted by the upcoming consonant (here, /d/ or /t/).

The asymmetrical pattern of relatively steady state (between 25% and vowel midpoint) and long transition is typical of the lower vowels in both populations, and quite distinct from the higher vowels (which even in the coarsely auto-normalized space show lesser overall distances traveled, and at least for some a more symmetrical patterns of movement distances. The lower vowels seem to hit their targets fairly early (by the 25% timepoint) and maintain those frequencies fairly steadily up to the vowel midpoint, at which time some 'relaxation' or 'reduction' may begin to take place.

So is any of this movement diphthongal?

At this time, I'd say no. But of course this depends on your definitions (and must all be confirmed in a more naturalistic dataset).

It seems to me that the true diphthongs /aɪ, oɪ, aʊ/ have two clearly defined targets which are specific to the underlying vowel category and independent of surrounding context. (Being agnostic about whether we should properly represent vowels as static targets or dynamic movements, I hesitate to continue to talk about 'targets'. I'd rather talk about movement as 'directed' or under category-specific control of the speaker, and 'undirected', being transitional, or interpolated, or otherwise predictable from context. But for simplicity, I will continue to talk about 'targets', with caveats).

The movement seen in these data overall suggest that these vowels have only a single target, typically achieved in the first half of the vowel at least by vowel midpoint. Movement seen in the second half of the vowel is consistent with general transitional movement. Whether this should be represented as a schwa-like target or I'm right in surmising a relationship with the upcoming consonant transition, it seems to be a general property of vowel production for these speakers, and should not be regarded a vowel-category specific target.

References

Clarke, Sandra, Ford Elms & Amani Youssef (1995). "The third dialect of English: Some Canadian evidence". *Language Variation and Change* 7, pp. 209-228.
Hagiwara, Robert (in press). "Vowel production in Winnipeg". *Canadian Journal of Linguistics/La revue canadienne de linguistique.*
Hagiwara, Robert (1995). "Acoustic realizations of American /r/ produced by women and men". *UCLA Working Papers in Phonetics* 90, pp. 1-187.
Hagiwara, Robert (1997). "Dialect variation and formant frequency: The American English vowels revisited". *Journal of the Acoustical Society of America* 102, pp. 655-658.
Hagiwara, Robert (2005a). "How 'general' is General Canadian? Vowel production in Winnipeg". Presented at *Canadian English in the Global Context*, Toronto, ON.
Hagiwara, Robert (2005b). "Revisiting the Canadian English vowel space". Presented at the *149th Meeting of the Acoustical Society of America*, Vancouver, BC.
Labov, William (1991). "The three dialects of English." In P. Eckert (ed.), *New Ways of Analyzing Sound Change* (Academic Press), pp. 1-44.