Analyzing complex vowel articulations from acoustic data

D. Sky Onosson
University of Victoria
Presentation

- **Topic of study**: Canadian Raising in Manitoba
  - research conducted as part of MA thesis (2010) at University of Manitoba
  - associated with Rob Hagiwara’s (2006) *Winnipeg Vowels Project*

- **Research question**: how to extrapolate from acoustic data to articulation *(among others!)*

- **Goal**: develop a method which extracts crucial information on diphthong articulation — including *position, duration and velocity* — without being overly complex
CR in Manitoba

• Subjects and data:

  • 8 speakers: Manitoba-born, female, native English speakers, ages 24-34

  • wordlist task: elicitation of 200+ tokens per speaker of /aj/ in a variety of phonetic contexts, with contrasting voiced/voiceless codas

  • 1,600+ tokens of /aj/ in total

• Study of /aw/ remains for future research
Transcription of diphthongs

• Various sources differ in representation of CR diphthongs: [aj] ~ [aɪ], [aw] ~ [aʊ]

• Either choice might be justified on phonetic or phonological grounds, depending on how define the term “diphthong”

• Miret (1998): diphthong is not a well defined term:
  • Catford (1977): “a sequence of two perceptually different vowel sounds in one and the same syllable”
  • Ladefoged (1982): “single vowels with continuously changing qualities”
Monophthongs vs. diphthongs

• Single or multiple articulations in sequence

• Absence or presence of articulatory motion and trajectory

  • Trajectory: “The path of any body moving under the action of given forces” (OED)

• Trajectories are physical paths — in this case, of speech articulators, e.g. the tongue — which can be extrapolated from acoustic data
Canadian Raising

- Martin Joos first described Canadian Raising in 1942:

  “The Canadian diphthongs /aj, aw/ have higher initial tongue-position in pre-fortis context than elsewhere, while for all other syllabics there is only a difference in length in the two kinds of context.”

- Joos suggests CR arose from “a shift from a difference essentially of length to a difference essentially of quality”

- My research suggests that length is still highly significant — more so than quality; Canadian Raising is more a process of shortening than it a process of raising
Developing a method

- CR is process of variation involving diphthongs

- Diphthongs involve an articulatory trajectory

- Given a different initial position (nucleus), each CR allophone/variant must have a different articulatory trajectory

- What is the best method to compare varying articulatory trajectories?
Monophthong

Formants typically measured at a single point selected to represent vocalic nucleus

Image: Ladefoged (2001)
Diphthongs

Where is the vocalic nucleus?
How many positions at which to measure formants?
Diphthongs

Measuring at one position is clearly inadequate
Diphthongs

Two timepoints: details of formant trajectory not preserved
Diphthongs

Adding a third point is a better representation — but some detail still missing
Multiple timepoints

Graph of data from nine measurement points overlaid on spectrogram (i.e. formants measured every 10%)
Comparison of CR formants

- Formant values averaged across all speakers, tokens
- Use of percentile scale fails to indicate durational differences between allophones
- Formant values and trajectories appear visually very distinct
Durational differences

• CR allophones exhibit large differences in duration:
  • before voiced segment (non-raised) [aj]: 293 ms
  • before voiceless segment (raised) [ʌj]: 159 ms
  • 184% difference in duration (alternatively, [ʌj] is 54% shorter)
• Non-raising varieties of English: vowels before a voiceless coda shorter than in other contexts (Peterson & Lehiste 1960, Chen 1970, Umeda 1975…)
• Question: How to incorporate duration into comparison of CR variants?
Incorporating duration, non-raised allophone

- Non-raised allophone is the longer variant, forms baseline for comparison

- Percentile timepoints recalculated as percentage of mean duration, 293 ms
Incorporating duration, raised allophone

- Raised allophone percentile time scale recalculated to mean duration = 159 ms

- Time axis scaled to match non-raised duration baseline = 293 ms
Incorporating duration, both allophones

- Comparison of both allophones aligned at left edge — articulatory onset
Incorporating duration, both allophones

- Raised allophone aligned to articulatory offset by shifting rightwards by the durational difference between two variants (136ms)
Incorporating duration

- Details of both articulatory trajectory and duration are well indicated.
- Similarities can be observed between the two allophones which were not readily apparent prior to including durational information.
- Non-obvious differences can also now be observed, i.e. presence/absence of nuclear steady state.
Canadian Raising
Redefining Canadian Raising

- **Raising** is slight, and evenly distributed *throughout* the articulation rather than occurring only at the nucleus.

- **Fronting** appears to be even more significant; F2 is higher (*fronted*) in the raised allophone, and the difference *increases* over time.

- **Shortening** of the raised allophone.

- **Steady-state phase** comprises half of the *non-raised* allophone, almost entirely absent in *raised* allophone.
Questions…

• How best to describe CR allophonic differences in terms of an articulatory model (e.g. *Articulatory Phonology*)? Are the duration and steady-state differences best accounted for as a single process, or multiple processes?

• What’s going on with /aw/?
  - Chambers (1989) has suggested the two diphthongs are not necessarily part of a single phonological process as indicated by the occurrence of raising with only one of the diphthongs in some American dialects, e.g. Roberts (2007) in Vermont

• What’s going on in other English dialects, which *don’t* have raising but *do* have pre-voiceless shortening?
  - e.g. Thomas (2000) looked at /ai/ production in Ohio and Texas, pre-/d/ and pre-/t/ exhibit truncation at different edges of the articulation — pre-/d/ truncates the glide, pre-/t/ truncates the nuclear steady state, overall duration much less divergent

• Is this method applicable to studies of diphthongs in other languages?
Figure 2. Spectrogram of *tide*...*tight* uttered by a female speaker from Johnstown, with a schematic diagram of the first three formants.
References


