THE PHONETICALLY DETERMINED DISTRIBUTION OF ATR HIGH VOWELS IN CANADIAN FRENCH

Gabriel Poliquin

Department of Linguistics
Harvard University

1. INTRODUCTION

The paper presented here is a fragment of a larger work on the vowel system of Canadian French\(^1\) (CF), which gives special attention to the distribution of tense, lax and ATR vowels. The paper is especially concerned with three categories of vowels: high, front mid unrounded and front mid rounded vowels. Because of space limitations, I will only be concerned with the first case here; readers interested in the two latter categories are referred to Poliquin (2002). The focus of the paper is even narrower, in that I will only be treating one case of high vowels, namely, those that exhibit tongue root advancement. The reason for further narrowing of the focus is not space, but interest in this case. ATR high vowels are the most interesting of the system in that their distribution is motivated by phonetic factors; that is the hypothesis put forward by this paper. The paper will therefore be concerned with the three main pillars of evidence that support this hypothesis: knowledge concerning the interaction of valves of the vocal tract, acoustic evidence, and, most importantly, preliminary results of studies using magnetic resonance imaging (MRI). Firstly though, I will give a brief description CF high vowels, and their distribution.

2. CF HIGH VOWELS

CF has three high vowels, /i, y, u/, each with three allophones: tense, [i, y, u]; lax, [I, y, u], and ATR, [Ii, yY, uu]\(^2\). Distribution of the allophones (in a phrase final syllable) is essentially contingent on the absence or presence of a coda consonant, and in the case of ATR allophones, on the presence of a particular class of coda consonants. Tense allophones are only allowed in open syllables, while lax allophones are only allowed in closed syllables. The pattern is illustrated in (1):

\begin{align*}
\text{(1) } & \text{ri} = \text{rice} & \text{t}'y = \text{you} & \text{tu} = \text{cough, n.} \\
& \text{Rid} = \text{wrinkle} & \text{t}'yk = \text{tuque} & \text{tus} = \text{cough, 3rd pers. sg. ind. pres.} \\
& \text{*Rid} & \text{*t}'yk & \text{*tus} \\
& \text{*RI} & \text{*t}'Y & \text{*tu}
\end{align*}

ATR allophones are found only in syllables closed by a voiced fricative and nowhere else in native CF words\(^3\). Tense and lax high vowels are disallowed before voiced fricatives:

\begin{align*}
\text{(2) } & \text{t}'z = \text{stem} & \text{eklyz} = \text{lock} & \text{luv} = \text{she-wolf} \\
& \text{*t}'z & \text{*eklyz} & \text{*luv} \\
& \text{*t}'z & \text{*eklyz} & \text{*luv}
\end{align*}

Now it is not immediately clear why ATR allophones should be differentiated from tense allophones. In many dialects of CF, vowels that surface before coda voiced fricatives have the same quality as tense vowels, but are much longer. Could this be a differentiation in length only caused by 'lengthening effect' of the fricative? Some

---

1. The dialect under consideration here is that of Montreal.
2. In many dialects of CF, these are not diphthongs, but long 'tense' vowels. The representation chosen here reflects the pronunciation of the speaker used for this study.
3. Many English words with high vowels have been borrowed, interestingly, the high vowels surface as ATR-like. We will not be concerned with loan words here.
previous literature on the topic has considered that to be case: vowels before voiced fricatives are simply tense vowels that have undergone a lengthening rule (Gendron 1966, Dumas 1978). Such hypotheses are not explanatory however, as the question remains: though length may be an effect of the following voiced fricative, why does the vowel have a tense quality? The hypothesis put forward here is explanatory in that the quality and length of the vowel will be explained by considering articulatory factors. Essentially, it will be proposed that the combination of a high vowel and a voiced fricative results in the vowel having ATR quality, as a result of pharyngeal expansion necessary for proper voicing.

2. ARTICULATORY AND ACOUSTIC JUSTIFICATIONS FOR THE HYPOTHESIS

The presence of +ATR vowels before voiced fricatives is the occurrence of a well-known phenomenon, which, in auto-segmental terms, is seen as the spreading of ATR values from obstruents to vowels. The phenomenon is described in Vaux (1992, 1994) for several languages including some Armenian dialects, Mon Khmer Languages, Akan, and Madurese. I will illustrate the phenomenon through examples from Buchan Scots English, which also exhibits similar behaviour. As described in Dieth (1932), the vowel [ɪ] only ‘occurs before nasals + cons. [...], the voiced fricatives v, ð, z, and partly before the voiced stops b, d, g,’ while [ɨ] ‘occurs in all other positions, [...] and all voiceless consonants,’ [ɛ] on the other hand, ‘occurs for [ɨ] in ‘broad’ speech, specially under full stress’ (39):

(3) live ‘lif’ vs. left ‘lift’
\textbf{mrid} ‘mother’ vs. \textbf{meθ} ‘might’
\textbf{mrdn} ‘dung-heap’ vs. \textbf{miten} ‘gloves’
\textbf{brd} ‘bid’ vs. \textbf{bɛt} ‘bit’

The distribution of the more advanced allophone [ɪ] (as opposed to [ɨ]/[ɛ]) is contingent on the following obstruent having a positive value for voice. In fact, the relevant feature for the obstruent is [ATR], as advancement of the tongue root is necessary for proper voicing of segments under certain conditions. In co-articulation with the following obstruent, the [ATR] feature manifests itself on the vowel as well.

2.1. Inferential Articulatory Evidence

Spreading of [ATR] values is the phonological expression of phonetic processes that are involved in the production of voicing. These are very well known, and their role in phonology has been discussed in literature dating back to Halle & Stevens (1971). As described in Shadle (1997) and Stevens (1998), production of voicing necessitates a pressure differential between the sub- and supra-glottal areas. Lowering of pressure in the supra-glottal area is done by means of expansion of the pharyngeal airway by forward movement of the tongue root. Advancement of the tongue root in such a way is responsible for the production of the preceding +ATR vowel.

Thinking only in terms of sub- and supra-glottal pressure, the hypothesis is problematic. Though all voiced obstruents involve a build-up of supra-glottal pressure, the pressure build-up involved in the production of plosives is presumably greater, given that their production requires complete closure of the vocal tract and total obstruction of airflow. There is thus an implicational hierarchy. The production of both plosives and fricatives involves the reduction of a pressure rise in the supra-glottal airway. Reduction may be implemented either passively, by flexible expansion of the cavity’s walls, or actively, by articulatory movement (advancement of the tongue root, or lowering of the larynx). If the latter strategy is implemented by a given language to reduce pressure rise in the production of fricatives, we would expect the language to do the same for the production of plosives, as they involve a greater pressure build-up. The case of CF contradicts the hierarchy. If the implicational hierarchy were to be verified by CF, we would expect that tongue root advancement in the production of voiced stops, and the consequent realisation of the preceding vowel as a diphthong. That is not so. The case of CF is surprising in this regard. Among the languages discussed in Vaux (1994) that exhibit this phenomenon, all verify the hierarchy. If [ATR] vowels are found adjacent to fricatives, they are also found adjacent to stops. What is more, most languages only show this behaviour for vowels adjacent to stops, not for fricatives. CF is therefore a contradictory case.
However, the case is only contradictory and surprising in this regard. Kohler (1984) provides finer-grained considerations of articulatory mechanisms, which offer insights into the matter. In the phonological literature, consonant classes such as /p, t, k/ and /b, d, g/ are traditionally differentiated with respect to [voice]. Though the distinction may suit the needs of most analyses, Kohler argues that phonological inquiry would gain much by considering the distinction between fortis and lenis consonants, and that the feature [+voice] should be reserved only for those consonants which exhibit actual glottal periodicity. The fact is that scores of languages distinguish pairs of consonantal arrays such as /p, t, k/ and /b, d, g/, but neither array need show any glottal periodicity at all. The standard example that comes to mind is Chinese, where the first array is made of aspirated consonants and the second of devoiced consonants. Kohler argues that these pairs should be distinguished using the feature [+fortis], where the former array is characterised by the positive value of the feature and the latter by the negative. The positive value of the feature characterises those segments that show a greater auditory signal, which, in turn, is correlated with greater articulatory power.

Importantly, ' [+fortis] is not proposed as an abstract feature, but it will be argued that degrees of articulatory power can provide its phonetic base' (Kohler, 1984:152). In abstraction, CF plosives and fricatives do belong to the same class, as they both involve obstructions of the vocal tract; in theory, they should behave similarly as a result of the pressure build-up that is involved in their articulation. Though the phonological prediction is phonetically grounded, we may gain something by grounding it further, so as to capture what sort of phonological generalisation is made by a child acquiring CF. It seems plausible that the child, in this case at least, may organise his system in terms of [+fortis], rather than [+voice].

How is [+fortis] characterised, articulatorily? In the articulation of any consonant, Kohler argues that one should consider the coordinated actions of the oral cavity’s three valves: the oral, the velopharyngeal and the glottal. The articulation of a ‘quintessential’ [+fortis] consonant, such as a voiceless fricative, involves the greater stricture of the oral and velopharyngeal valves, as well as the tightening or abduction of the vocal folds. In short, a [+fortis] consonant involves stricture, or tightening of at least one of the three valves. If a [+fortis] consonant is articulated with abduction of the vocal folds rather than tightening, greater stricture of the oral valve is necessary to ensure greater acoustic salience relative to lenis counterparts. This results in more energy being involved in the articulation of the release. Lenis consonants, on the other hand, involve a general slacking of a particular point in the oral tract. For instance, the articulation of a devoiced stop will involve abduction or tightening of the glottal valve, but slacking of the oral valve. The system entails that, if there is a general slacking of the valves (especially the glottal), voicing ensues. As Kohler himself points out: ‘owing to these physiological conditions, voicing is an extreme manifestation of the lenis feature in stops and fricatives, since it greatly reduces the air-stream power and the tension associated with powerful articulatory movements’ (Kohler, 1984: 163).

Not only is the force of stricture important, duration of stricture is also to be considered. The production of fortis stops involves more extensive movements with greater average velocity. This is confirmed by studies presented in Fujimura & Miller (1979). Fricatives on the other hand (fortis and lenis) involve greater muscle coordination that is sustained for a greater period of time ‘leading to longer vowels and consonants’ (Kohler, 1984:155). This would explain the greater average length of CF vowels when preceding fricatives. What was accounted for by means of a phonological rule in past studies of this phenomenon can be reduced to a simple phonetic effect.

2.2. Inferential Acoustic Evidence

Kohler’s dynamic perspective not only considers degrees of stricture in terms of location and duration, but also takes into account the timing of different articulations one relative to others. The articulatory movements involved in the production of a particular consonant invariably affect the articulations of segments adjacent to it. This is, of course, the cause of co-articulation phenomena. It has been observed in Kohler (1981) that the articulatory movements necessary to ensure proper voicing of a coda obstruct are initiated during production of the preceding vowel (termed ‘preparation of voicing’ Kohler, 1984: 163)). Kohler (1981) shows how this surfaces in

4 [+fortis] should not be correlated with aspiration, it is a happenstance of Chinese that only its [+fortis] consonants are aspirated. As pointed out by Kohler, Hindi has both aspirated and unaspirated [+fortis] stops that each have [-fortis] equivalents.
In Kohler’s study of French in this respect, greater or lesser differences between $F_2$ and $F_1$ can be observed from the very beginning of the vowel. The results were established based on the comparison of French voiced (lenis) and voiceless (fortis) stops.

I have performed the same experiment on my corpus of data, comparing the differences between the second and first formants at the beginnings, and ends of vowels. All vowels were followed by a voiced or voiceless obstruent. The goal of the experiment is to find which obstruents pattern like lenis consonants, and which pattern like fortis consonants. Given the results of Kohler (1981), one would expect that vowels showing an increase in the difference of $F_2$ and $F_1$ are those that precede a lenis consonant. The increase in difference is explained by a rise in $F_2$, which is correlated with an advancement of the tongue root, necessary for proper voicing of the consonant.

Measurements in (4) (in Hertz) each correspond to the arithmetic mean of frequencies (either $F_1$ or $F_2$) measured at a given point in the vowel (starting ($P_1$) or ending ($P_2$)) over four tokens of the same word in the same environment. The last row of each table corresponds to the increase or decrease in the differences between $F_2$ and $F_1$ between $P_1$ and $P_2$:

(4a) $\begin{array}{|l|l|l|l|l|}
\hline
& it & id & is & iz \\
\hline
F_1 & 408 & 474 & 370 & 415 \\
F_2 & 2054 & 1760 & 1895 & 1682 \\
F_2-F_1 & 1646 & 1286 & 1525 & 1267 \\
\pm & -360 & -258 & -200 & +183 \\
\hline
\end{array}$

(4b) $\begin{array}{|l|l|l|l|l|}
\hline
& ip & ib & if & iv \\
\hline
F_1 & 404 & 450 & 399 & 411 \\
F_2 & 1936 & 1695 & 2038 & 1825 \\
F_2-F_1 & 1532 & 1245 & 1639 & 1414 \\
\pm & -287 & -225 & -106 & +125 \\
\hline
\end{array}$

(4c) $\begin{array}{|l|l|l|l|l|}
\hline
& rt & rd & rs & r\textsuperscript{2}z \\
\hline
F_1 & 369 & 391 & 394 & 405 \\
F_2 & 1845 & 1641 & 1577 & 1516 \\
F_2-F_1 & 1496 & 1250 & 1183 & 1111 \\
\pm & -246 & -72 & -61 & +69 \\
\hline
\end{array}$

Vowels followed by palatal and velar obstruents were not measured as the CF segment inventory does not provide the necessary environments for a proper controlled study. Stops are meant to be compared to fricatives so as to see which show the acoustic signal characteristic of a lenis or fortis consonant. However, CF has no palatal stops one can compare with palatal fricatives, and no velar fricatives one can compare with velar stops.
3. BEYOND INFERENTIAL EVIDENCE: PRELIMINARY RESULTS FROM MRI STUDIES

Given what we know concerning the coordination of the vocal tract's different valves and given the study of relevant CF acoustic signals, we can infer that the succinct production of a high vowel and a voiced fricative necessitates pharyngeal compensation. Such an articulatory strategy, necessary for proper voicing of the fricative,
may require anticipatory advancement of the tongue root, which, in turn, spurs advancement of the tongue body, causing the high vowel to be produced at a higher frequency. The hypothesis is strongly motivated, but not verified as of yet.

If one considers that the evidence gathered so far is only inferential, the question begs to be asked: what sort of evidence could confirm or disprove the hypothesis? Moving mages of the interior of the vocal tract would of course be best suited to this purpose, but how does one obtain them?

In a perfect world, the choice of which technology to use depends only on the needs of the study. In a less perfect world, the choice is dictated by what is available. Though I operate in a less than perfect world, I was lucky enough, that, what happened to be available was perfect. The need of the study is this: to find whether maximal pharyngeal expansion in the articulation of a [high V] + [vee, fric] sequence coincides with the articulatory maneuver that ensures partial obstruction of the vocal tract. For example, in the syllable [i'v'], the experiment would attempt to verify if maximal pharyngeal expansion coincides with retraction of the lower lip toward the teeth. A positive result would indicate two things: first, that tongue root advancement is an articulatory maneuver aimed, not at producing a particular allophone of /l/, but aimed at ensuring proper voicing of the obstruent; second, because maximal advancement is timed with retraction of the lower lip, the tongue root is actually moved forward before retraction of the lip, that is, during production of the high vowel, thus leading the vowel to have a higher F₂.

The objective is then to see if maximal pharyngeal expansion is timed with retraction of the lower lip. Now, magnetic resonance imaging (MRI) may be the most promising technique to verify such a thing. MRI can provide the precise and total vocal tract imaging necessary for our objectives. However, the ideal piece of evidence would be moving images, so that we may compare expansion at the point of lip retraction with expansion at further and anterior points in time. Traditionally though, MRI would not be suited for this purpose as image capture speed is much too slow. Where a film camera has a shutter speed of 24 frames per second (fps), an MRI magnet has a 'shutter speed' of approximately 1 fps. In order to capture articulator transitions that last 20-40 ms, MRI must be able to capture images at a much higher speed. The challenge of speeding up MRI scans has been overcome by the work of researchers at the Massachusetts Eye & Ear Infirmary, which was presented in Nissenbaum et al. (2002). The thing to understand is that an MRI magnet does not capture images, but data, and that data is stored in multi-dimensional arrays. A still MRI of a patient's brain say is formed by means of a 2-D Fourier transformation of a three-dimensional data matrix. The matrix should be pictured as a loaf of bread, each slice being a two-dimensional array of data. The two-dimensional array (a slice of the bread) should be pictured as a rectangular table of data with 256 rows. With each scan of the machine, data is collected 256 times, once every 7 ms. Each data collecting event corresponds to one row of the array being filled. In the case of a still MRI, each two-dimensional array of the matrix corresponds to one cut of the object being filled. By combining different cuts together, we get a complete image of the object under investigation. Now, to make moving images, we proceed entirely differently. Where each two-dimensional array corresponded to a cut of an object for a still MRI, in the case of moving images, each array corresponds to a point in time during an utterance (there are 128 slices). In the case of a still image, with each scan, one array was filled. In the case of a moving image, a particular row of each slice is filled. So, with the first scan, we fill the first row of 128 arrays, with the second scan, we fill the second row, and so on. To fill the 256 rows of data, the subject is asked to repeat the same utterance 256 times. The first slice of the matrix will then correspond to a mid-sagittal view of the subject at the beginning of the utterance, while the last slice will correspond to a view of the subject at the end of the utterance. Slices in between correspond to intermediary points. Each slice of the matrix is then converted into images that are minimally different as positions of the various articulators shift over time. Producing an MRI 'movie' is then much like making a cartoon.

I have applied the technique to the study of tongue root advancement in CF. At the time of writing the experiments have not been fully controlled, and only provide partial results. However, the results that were obtained were controlled to a degree such that they are already convincing. These will be presented here. The hypothesis to test is this: production of a high vowel requires greater stricture of the oral valve of the vocal tract leading to an increase of pressure in the supra-glottal cavity; voicing of a fricative requires reduction of supra-glottal pressure; given the observations of Kohler (1984), we would expect pharyngeal expansion by means of tongue root advancement to occur so as to insure pressure reduction. In the case of a mid vowel however, there is much less stricture of the vocal tract; because less stricture leads to less pressure, no pharyngeal expansion will be necessary.
Because there will be no tongue root advancement in the production of a fricative following a mid vowel, the vowel will not be affected.

The results were based on two MRI scans of myself, each scan corresponding to a different utterance. In the first utterance, a word containing a high vowel and a voiced fricative was placed in a carrier sentence in which it had phonological prominence. For the second scan, the same carrier sentence was used, but the target word was changed to one containing a mid unrounded vowel and a voiced fricative. During the collection of data, each carrier sentence was uttered 256 times, each of which had to be synchronised with the scanner. To do this, the following strategy was used. A very short question-answer dialog was recorded, where the question prompted the carrier sentence as an answer. The dialog was repeated 256 into headphones worn by the subject, who uttered the carrier sentence in synchrony with the answer to the question he was hearing. The two dialogs are shown below:

(5)a. Q: C'est Claire ou Yves qui attend?
   sê klɛʁ o jɛv kjas tah
   is-it Claire or Yves that is waiting?

   A: Yves attend.
   jɛv aš tah
   Yves (is) waiting.

b. Q: C'est Adam ou Eve qui attend?
   sê adɔ u ev kjas tah
   is-it Adam or Eve that is waiting?

   A: Eve attend.
   ev aš tah
   Eve (is) waiting.

The target words are the first word of each answer. Their position is one of phonological prominence as they constitute the information awaited by the question; they thus carry stress conveying focus. In a simple declarative sentence, the voiced fricative of the target word would resyllabify as the onset of the following syllable. Special stress on the target word prevents such resyllabification from occurring, insuring that the fricative is indeed in coda position; a glottal stop separates the fricative from the following nucleus.

For each carrier sentence, two frames of the scan were isolated. The first of each corresponds to the first frame of the film, at a point in the utterance when there is no tongue root advancement, corresponding to a beginning point in the vowel. The second frame corresponds to the frame where there is maximal lower lip retraction:
(6)a. *Yves attend* (high vowel + vce fricative).

\[\alpha\]

\[\beta\]

b. *Eve attend.* (mid unrounded vowel + vce fricative)

\[\alpha\]

\[\beta\]

Image analysis using Matlab software proceeded as follows. A script was written so as to allow a point to be chosen on the pharyngeal wall, and a line parallel to the abscisse drawn through it. Another script was then written for a point with identical co-ordinates to be chosen in all three remaining images (point A). Then, for each image, another point on the line was chosen, one corresponding approximately to the line and the outermost edge of the tongue root (point B). Distance between point A and B was then measured for each image. Essentially, measurement is done by counting the number of pixels between each point with the help of the software, which then converts the number of pixels into measurements in millimeters; each pixel is equal to 0.8 mm. Results are shown in the following table:

<table>
<thead>
<tr>
<th>Images</th>
<th>Distance between points A and B (in mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)a)(\alpha) - (\varepsilon)' (frame 1/126)</td>
<td>16.34</td>
</tr>
<tr>
<td>(6)a)(\beta) - (\varepsilon)' (frame 70/126)</td>
<td>21.5</td>
</tr>
<tr>
<td>(6)b)(\alpha) - (\varepsilon) (frame 1/126)</td>
<td>13.76</td>
</tr>
<tr>
<td>(6)b)(\beta) - (\varepsilon) (frame 69/126)</td>
<td>13.76</td>
</tr>
</tbody>
</table>

Measurements show that when *Eve* was the target word, the distance between point A (constant on the pharyngeal wall) and point B (point on outer edge of the tongue root) remains the same before and during lower lip retraction (13.76 mm). When *Yves* is the target word however, there is a lengthening of the distance between A and B by 5.16 mm. One can see very clearly in (6)a)\(\alpha\) that the tongue root is indeed moved forward, causing the body of the
tongue to jut forward. Measurements are accurate, but not thorough. It remains to be seen if maximal advancement occurs during lip retraction. This will involve measuring the distance between A and B for every frame between the beginning and end of lip retraction. My initial assessment indicates that the hypothesis will be verified; what is certain, and what the approximations tell us, is that we can be certain that tongue root advancement occurs at some point during lip retraction, and it would appear that maximal tongue root advancement co-occides with it.

The results need to be controlled further however, as we have to eliminate all possibilities. There is the unlikely possibility that there exists tongue root advancement in the production of other segments, and that there are no co-articulatory effects on the preceding vowel. If that is the case, it would be difficult to uphold the view that it is tongue root advancement which changes the quality of the preceding vowel. We must therefore control for these kinds of cases. Also, we must see what happens in syllables that are closed by a voiced fricative, but where there is no impact on the preceding vowel. Such cases exist in the inventory of English loan words in CF. The word *business* for instance was borrowed without any major changes; the first syllable contains the target environment: *[biz. nis]*. There are one of two possibilities. According to the hypothesis because the voice fricative is preceded by a high vowel, we should see tongue root advancement, which is necessary for its proper voicing. However, there is no co-articulation on the preceding effect in this case. There are two possibilities: either, tongue root advancement is simply timed in such a way that maximal pharyngeal expansion does not correspond with stricture, or, full voicing is ensured by another articulatory strategy, namely, passive expansion of the pharyngeal airway.

At the very least, the hypothesis is far from being disproved, and preliminary results seem encouraging. At this point in time, we can predict that it is unlikely that further results will completely disprove the hypothesis.

REFERENCES


