Variation of rhythm metrics in regional varieties of Acadian French

Władysław Cichocki*, Sid-Ahmed Selouani, Alaidine B. Ayed, Catherine Boudreau and Yves Perreault

*Corresponding author's address: French, University of New Brunswick, Fredericton, E3B5A3, New Brunswick, Canada, cicho@unb.ca

This preliminary investigation studies variation of rhythm metrics in dialects of Acadian French spoken in New Brunswick (Canada). The aim is to determine whether regional and social factors are significant sources of this variation. Data are recordings of 140 speakers who represent five geographic regions, both genders and two age groups (20-30 and 40-55 years of age). Sound files were segmented manually; eight interval-based metrics were calculated. Mean metric scores indicate that regional varieties of Acadian French pattern with other dialects of French; these scores are similar to those of other so-called syllable-timed languages. Regional differences are found for several metrics (%V, Delta C, VarcoC, nPVI-C, rPVI-C), although the five geographic regions are not all clearly distinguished by these metrics. A major pattern that emerges is regional variation in high-vowel devoicing and/or deletion. Analyses also show that social factors are significant sources of interspeaker variability: gender (VarcoV, nPVI-C, rPVI-C) and age (DeltaV, VarcoV). These results suggest a certain amount of complementarity between regional and social factors in their effects on rhythm metrics.
1. INTRODUCTION

The development of rhythm metrics has motivated a significant body of research into linguistic rhythm. While it may be that these metrics provide only a partial picture of the general notion of rhythm in language (Arvaniti 2009), they are providing important insights into how speech timing functions. Many studies examine cross-language variation with a view to clarifying the classification of languages into syllable-timed and stress-timed categories.

There is also a growing literature that applies these metrics to differences within a single language. Research is showing that there can be considerable variation between dialects of the same language. In a study on fifteen regional varieties of Italian, Giordani and D’Anna (2010) note “wide-ranging variation”. Using several metrics (%V, ΔV, ΔC, nPVI and rPVI) the authors find that certain varieties resemble mora-timed and stress-timed languages, even though Italian is generally considered to be syllable-timed. Work on nine dialects of French (Obin et al. 2012) finds gradient differences among referential (or standard), regional and contact varieties. The authors note that rhythm metrics partially distinguish among these dialects. Among other examples of research on cross-dialect differences are studies of Arabic (Ghazali et al. 2002), British English (Ferragne and Pellegrino 2004), Greek (Topintzi et al. 2010), Italian (White et al. 2009), Portuguese (Frota and Vigario 2001) and Spanish (O’Rourke 2008). In addition, sociophonetic research is showing that social factors – such as gender, age and class – may also affect variation within a regional variety. Nokes and Hay (2012) examine variation in vocalic nPVI measures to document a change in the rhythm structure of New Zealand English. However, Fagyal (2010) finds that rhythm metrics do not confirm perceptual impressions of rhythmic differences in the French spoken by two ethnic groups who live in a suburb of Paris.

The aim of our research program is to describe the rhythm structure of the varieties of French spoken in the province of New Brunswick (Canada). These varieties belong to a dialect known as Acadian French, which is distinct from its well-known neighbour Québec French. The map in Figure 1 identifies the geographic area under consideration. Our study does not include varieties of Acadian French that are spoken in the neighboring provinces of Nova Scotia, Prince Edward Island and Newfoundland.

![Map of New Brunswick](image)

**FIGURE 1.** Map of New Brunswick showing the five regions surveyed in the RACAD corpus.

This paper is a preliminary report of our research; we present a listing of the scores on eight rhythm metrics and a description of some of the variation in these rhythm scores. We address two main research questions. We ask, first, whether New Brunswick Acadian French, like many varieties of French, belongs to the class of syllable-timed languages. Our second research question asks whether regional and social factors are significant sources of variation in the rhythm scores. Because this is a preliminary study, we are interested in identifying some of the next steps in our research.
2. METHOD

2.1 Speakers and Materials

Data are taken from the RACAD (Reconnaissance automatique de l’acadien) speech corpus, which was designed for research on automatic speech recognition (Cichocki et al. 2008). This database contains recordings of sentences read by 140 speakers who represent the five main francophone regions in New Brunswick: Northwest, North, Northeast, Southeast and Moncton-Dieppe (see Figure 1). To include social variation, the corpus has equal numbers of males and females, and there are two age groups, younger (between 18 and 24) and older (30 to 55). Table 1 summarizes information about the speakers in the sample.

<table>
<thead>
<tr>
<th>Region</th>
<th>younger females</th>
<th>older females</th>
<th>younger males</th>
<th>older males</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northwest</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>North</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>Northeast</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>39</td>
</tr>
<tr>
<td>Southeast</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>Moncton-Dieppe</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>140</td>
</tr>
</tbody>
</table>

The sentences in the RACAD database were selected from lists of phonetically balanced sentences of French. In this study, we examine the two sentences in the database, given in (1) and (2), that were read by all 140 speakers. None of the other sentences in this database was read by all speakers, and using them would have introduced an additional source of variation in rhythm scores (see Wiget et al. 2010), which we chose to avoid in this initial study.

(1) Je viens de lire dans l’Acadie Nouvelle qu’un pêcheur de Caraquet va monter une petite agence de voyage.
   “I just read in Acadie Nouvelle that a fisherman from Caraquet is going to open a small travel agency.”
(2) C’est le même gars qui l’année passée a vendu sa maison à cinq Français d’Europe.
   “He is the same guy who sold his house last year to five Frenchmen from Europe.”

2.2 Analysis

Segmentation and labeling of the speech samples were done manually at the LARIHS Laboratory at the Université de Moncton using Praat. All segmentation was verified by two of the co-authors. Vocalic and intervocalic (consonantal) intervals were extracted with a Praat script, and rhythm metrics (following Grabe and Low 2002, and Ramus et al. 1999) were calculated with routines written with Matlab.

The following is a list of the rhythm metrics calculated:

- %V (percentage of total utterance duration composed of vocalic intervals)
- ∆V (standard deviation of vocalic interval duration)
- VarcoV (coefficient of variation of vocalic interval duration)
- nPVI-V (normalized pairwise variability index for vocalic intervals)
- AC (standard deviation of consonantal interval duration)
- VarcoC (coefficient of variation of consonantal interval duration)
- nPVI-C (normalized pairwise variability indices for consonantal intervals)
- rPVI-C (raw pairwise variability indices for consonantal intervals)

A series of univariate ANOVAs was carried out on the subject scores with SPSS.
3. RESULTS AND DISCUSSION

3.1 Situating Acadian French among the Rhythm Classes

Table 2 gives the mean scores and standard errors for the eight rhythm metrics based on speech data from all 140 speakers in the corpus.

<table>
<thead>
<tr>
<th>Rhythm metrics</th>
<th>%V</th>
<th>49.03 (0.24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔV</td>
<td>55.99</td>
<td>0.89</td>
</tr>
<tr>
<td>VarcoV</td>
<td>50.78</td>
<td>0.52</td>
</tr>
<tr>
<td>VarcoC</td>
<td>57.39</td>
<td>0.95</td>
</tr>
<tr>
<td>nPVI-V</td>
<td>55.61</td>
<td>0.56</td>
</tr>
<tr>
<td>nPVI-C</td>
<td>47.36</td>
<td>0.44</td>
</tr>
<tr>
<td>rPVI-C</td>
<td>56.97</td>
<td>0.47</td>
</tr>
</tbody>
</table>

To provide a context in which to interpret these scores, Figures 2a and 2b plot Acadian French timing scores along with those of three languages that represent the so-called main rhythm classes: English (a stress-timed language), Spanish (a syllable-timed language) and Catalan (sometimes referred to as a mixed or intermediate language). Comparison data are from Payne et al. (2011). Figure 2a, which associates %V and ΔC, shows that Acadian French is fairly close to Spanish, a representative of syllable-timed languages. Acadian French has a slightly higher value of %V than Spanish; it also has a higher value of ΔC, although this value is certainly not as high as the ΔC score for English. In the (nPVI-V, rPVI-C) plane in Figure 2b, Acadian French is once again close to Spanish, and it has a slightly greater degree of consonantal interval variability than Spanish. These results suggest that New Brunswick varieties of Acadian French tend to resemble syllable-timed languages.

FIGURE 2a. Distribution of Acadian French, Catalan, English and Spanish over the (%V, ΔC) plane.

FIGURE 2b. Distribution of Acadian French, Catalan, English and Spanish over the (nPVI-V, rPVI-C) plane.
Comparisons with data published for French indicate that rhythm scores for Acadian French fall within the general range of scores found in other dialects of French. The mean nPVI-V score for Acadian French (47.4) is very close to scores for three other regional varieties of French spoken in Canada: Windsor (45.5) and Hearst (43.9) in Ontario, and Québec City (44.5) in Québec (based on Tennant 2012). White and Mattys (2007) report an nPVI-V score of 50 for “français neutre” (presumably a variety spoken in France), which is also close to the Canadian scores. Obin et al. (2012) calculate four metrics (%V, ΔC, nPVI-V, rPVI-V) in their comparison of nine different varieties of French spoken in Europe and Africa. For all four metrics, the Acadian French scores are very similar to those reported for regional dialects in Switzerland; the Acadian scores are nevertheless distinct from those of urban varieties found in Paris and Lyon in France.

### 3.2 Sources of Variation in the Rhythm Metric Scores

#### 3.2.1 Variation with Respect to Region, Gender and Age

There is considerable variation in the rhythm scores. Our initial analysis studies the effects of the three external factors – region, gender and age – that were used in designing the RACAD database. Table 3 summarizes the results of a series of 3-way ANOVAs that test the effects of these factors on each of the rhythm metrics.

<table>
<thead>
<tr>
<th>Rhythm metrics</th>
<th>Region</th>
<th>Gender</th>
<th>Age</th>
<th>Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>%V</td>
<td>.024</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ΔV</td>
<td>–</td>
<td>–</td>
<td>.001</td>
<td>–</td>
</tr>
<tr>
<td>VarcoV</td>
<td>–</td>
<td>.022</td>
<td>.003</td>
<td>–</td>
</tr>
<tr>
<td>nPVI-V</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>gender*age (.051)</td>
</tr>
<tr>
<td>ΔC</td>
<td>.021</td>
<td>–</td>
<td>–</td>
<td>gender*age (.054)</td>
</tr>
<tr>
<td>VarcoC</td>
<td>.018</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>nPVI-C</td>
<td>.027</td>
<td>.001</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>rPVI-C</td>
<td>.027</td>
<td>.049</td>
<td>–</td>
<td>gender*age (.027)</td>
</tr>
</tbody>
</table>

Region is the factor with the greatest number of significant effects: five metrics. Gender differences are significant on three metrics, and age on two. There are very few two-way interactions and no three-way interactions.

The rhythm metrics appear to be capturing different aspects of regional and social variation. On the interval-based measures (%V, ΔV, VarcoV, ΔC, VarcoC), there is a complementarity among effects: region and the two social factors have significant effects on different metrics. However, on the two consonantal pairwise variability indices (nPVI-C, rPVI-C), both region and gender have significant effects. The nature of these effects on the rhythm metrics and the relation between the regional and social factors will be developed in follow-up research.

Another issue for further research is the relation of the metrics to phonetic realizations. The clustering of significant region effects on the four consonantal variability measures (ΔC, VarcoC, nPVI-C, rPVI-C) suggests that these metrics are tapping into phonetic processes that can affect non-vocalic (consonantal) intervals. One of these is the shortening and occasional deletion of the high vowels /i, y/ in the context of the assimilation of /t, d/. Examples of these deletions in the corpus are une petite agence [ynpætlstæjøs], l’Acadie Nouvelle [lkædznuvel], a vendu sa maison [avɔdзsamezɔ]. The result is longer non-vocalic intervals that contribute to durational variability. This and other phonetic processes that affect durational variability will be examined more closely in future research.

#### 3.2.2 Some Patterns of Regional Variation

Some regional patterns emerge from variation in the rhythm scores. Effects for region are found on %V and on the four metrics that are associated with consonantal intervals. To illustrate these effects, Figure 3 plots the five regions of New Brunswick on the (%V, VarcoC) plane. The %V metric arranges the regions along a north-south axis; the VarcoV metric separates the Northwest from the other four regions. Briefly, the Northwest has the highest VarcoC score; the Southeast has the highest %V score. The interesting result is that rhythm metrics provide a fairly regular, albeit partial, distinction among the geographic regions.
4. CONCLUSIONS AND FUTURE DIRECTIONS

This study presents a preliminary description of speech timing in New Brunswick Acadian French. Rhythm metric scores are similar to those of other dialects of French, and they classify this variety among syllable-timed languages. Both regional and social factors have effects on variation in the rhythm scores. Distinctions among geographic regions are partially characterized by timing differences. Follow-up work will examine the relation between region, gender and age as sources of variation in these scores, as well as phonetic processes that can affect durational variability.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the New Brunswick Innovation Foundation (NBIF), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC).

REFERENCES


