Predicting Vowel Duration in Spontaneous Canadian French Speech

Darcie Williams, François Poiré

Department of French, University of Western Ontario, London, Canada
darciewilliams13@gmail.com, fpoire@uwo.ca

Abstract

This study examines variables influencing vowel duration of French spoken in Windsor, Ontario, in order to see whether their respective effects on vowel duration are organised hierarchically. We first consider the data distribution of four female speakers before carrying out a statistical principal components analysis. Our results show that the variables are classified into three underlying factors: syllable structure, syllable position and vowel properties. This last factor group includes the factors phonological vowel class and diphthong status, and always explains the majority of the variability in vowel duration. Syllable position also accounts for some of this variation in certain cases. The consistent hierarchy of these factors across the statistical analyses confirms that a vowel’s properties are the most important in determining its duration, followed first by the syllable’s position in the utterance, and second by the syllabic structure.

Index Terms: vowel duration, variation, Ontario French

1. Introduction

We seek to establish a hierarchy of variables identified in the literature as influencing vowel duration by statistically regrouping these variables into underlying factors. In order to achieve this goal, we conduct an acoustic and distributional analysis on recordings from four female speakers from the Phonologie du français contemporain (PFC) project, and carry out a factorial (principal components) analysis. We analyse the results of these calculations in order to discover what variables play the most important roles in determining vowel duration, and in what contexts. These observations will allow certain preliminary conclusions regarding the subjacent system controlling vowel duration.

1.1. Problem

Vowel duration is a well-known and studied topic; research comparing phonetic and phonological traits normally states that a vowel’s duration is the result of its intrinsic length combined with identified phonetic conditions. However, these variables are almost never classified based on empirical data. A study of English voice synthesis is an exception: the author uses mathematical formulas to determine vowel properties [1]. He notes that variables controlling vowel duration are not constant, but have a combinatory effect of either limitation or augmentation; he proposes a sum-of-products model to account for this cumulative effect.

Although this research analyses English, it presents several relevant questions for our study. Is there a difference, as the study suggests, between perceived and actual vowel duration? To what degree are phonological classifications, often based on perception, reliable? Can we arrange variables influencing vowel duration into a consistent hierarchy?

1.2. Variables influencing vowel length in French

French vowel articulation implies four basic parameters: tongue height, rounding, retraction and nasality [2]. These properties are linked to vowel duration [3]: nasal vowels are longer than oral vowels, and low vowels are longer than high vowels [4]. In Canadian French (CF), duration and intensity are identified as intrinsic characteristics of the vocalic system [4]; thus, CF contrasts phonologically long and short vowels [5]. This gives CF a distinct rhythm from what is called “standard French” (SF) [6], where this contrast is disappearing [5]. Some researchers suggest that this distinction applies to all vowels except /a/, which is always short, and nasal vowels, which are always long [2]. Others submit that there are four oral and four nasal phonologically long vowels: the back “a” [a], the closed “o” [o], the nucleus of “peu” little [œ], and the long “e” [œ], [œ], [œ], and [œ] [5]. These eight vowels always maintain their inherent length when accented or in a closed syllable, and often do when in an open or closed penultimate syllable [6].

The four long oral vowels have two possible allophones: a stable long vowel realised in open syllables, or a diphthong realised in final position closed syllables [2]. The other vowels do not normally diphthongize, but vowels such as [a, e, œ, ε], when lengthened by a following consonant, can undergo diphthongization, as can lengthened nasal vowels [5]. These cases all depend on stress [5]. In French, stress is not lexical, as in English [7], but rather prosodic; the final syllable in an accentual phrase receives the primary accent [3]. Stress manifests mainly in duration [5].

Thus, the final position is ideal for maximizing vowel duration [5]. Results of certain studies [6] suggest that in spontaneous spoken CF, phonological vowel classification is often overcome by syllabic position in determining vowel duration. One exception to this structural model is penultimate syllable lengthening in CF [8]. This syllable is normally the weakest of the word or the accentual phrase [5]. However, two types of vowel lengthening in this position are identified. The first is simply an increase in the penultimate syllable’s duration without displacement of the final stress; the second involves moving this final stress to the penultimate syllable [8].

Syllable structure also influences vowel duration, since closed syllables favour vowel lengthening. In fact, any stressed vowel can lengthen if the syllable’s coda begins with one of the lengthening consonants [R, z, ð, v, vR] [5]. When the nucleus is one of the four nasal vowels or one of the three oral vowels [a, o, ə], any coda can increase that vowel’s duration [7]. Even short vowels can sometimes lengthen, both in closed penultimate syllables and in open syllables, as long as the lengthening consonant and the vowel in question form part of the same morpheme [6]. We see that in a given
syllable of CF, the coda exerts more of an influence over the nucleus’ duration than does the onset [3, 4]. In addition to the lengthening consonants we have discussed, there are ten “shortening” consonants that, as codas, limit the duration of the syllable’s nucleus: \{p, t, k, b, d, g, m, n, ŋ, l\} [5].

The formality of the situation also influences vowel duration, because the more formal the situation, the more intact the articulatory and durational properties of unstressed syllables [5]. As a result, vowel traits differ slightly between reading and conversational speaking styles.

2. Corpus

Our Windsor, Ontario corpus of four female speakers is part of the international PFC project. Of four available speaking styles, we analyse “reading” and “controlled conversation”; each speaker has read a designated news article, followed by an informal interview with a PFC researcher. We seek to contrast the formality of these two speech events. We also consider the sociolinguistic variables of age and audible influence of English. Speakers one and two are 88 and 60 years old, respectively; neither speaker exhibits an audible influence of English when speaking French. Speakers three and four are 41 and 16 years old, respectively; both show a noticeable influence of English when speaking French.

Our analysis of three minutes of recordings for each speaker in each style, with all hesitations, interruptions, pauses, and stutters eliminated, yields more than 600 syllables in “reading”, and between 800 and 1000 syllables in “controlled conversation” for each speaker. In our corpus, 80% of the syllables in “reading” are open (2019 of 2509 syllables); in “controlled conversation”, these syllables make up 85% of the total (3112 of 3713 syllables). Also, four syllabic structures are more common than others in our corpus, varying by speaker and style: CV (from 53% to 61% of the data); CCV (11% to 17%); CVC (11% to 16%); and V (7% to 12%).

3. Method

3.1. Acoustic analysis

Selected recordings are segmented using PRAAT software. Our segmentation identifies the duration of the onset, nucleus and coda of every syllable of each recording. Segmenting is conducted using the spectrogram; voicing and formant structure determines vowel boundaries. We follow Cedergren and Perreault’s [9] prosodic analysis of CF by dividing speech events into syllables, rhythmic groups and intonational groups. The highest level, “accentual phrase”, combines phonological words based on the intonational curve of the spectral analysis and changes in fundamental frequency [10]. “Phonological word” [9] is next; these groupings are determined by the functional status of each lexical word. “Full” words, such as nouns, terminate the grouping; “clitics”, such as prepositions, do not. The final three levels are “lexical word”, “syllable”, and “segment”.

Segmentation at the lexical and syllabic levels presents the issue of resyllabification. For example, in the words “petite amie”, girlfriend, the graphic ending of “petite” is the phonetic onset of the first syllable of “amie”:

\[po.tiː + [a.mi] \rightarrow [po.ti.ta.mi]\] (1)

In such cases, we follow the Maximum Onset Principle set out by Kahn [11]. Sometimes, the application of this principle creates normally impossible sequences; one such example from our corpus is the sequence [dR] from “beaucoup de français”, a lot of French:

\[bo.ku] + [d] + [tRə.se] \rightarrow [bo.ku.d.tRə.se]\] (2)

Since no French syllables begin or end with [dR], this onset violates Hooper’s Legality Principle [12]: segmented onsets and codas must exist elsewhere in the segmented language. We obey both principles in our segmentation: we do not create non-existent onsets; we displace a maximum of two consonants from the coda of one syllable to the onset of the next; and we never displace the phoneme [s] if this movement creates an onset of more than three consonants.

3.2. Principal components analysis

In order to transfer our data into SPSS, we code durational, structural, phonological and prosodic classifications using increasing whole numbers (\{0, 1, 2, 3, ...\}) that express binary (yes/no) or ordered (tiny/small/large/enormous) relationships. Due to the arbitrary nature of the ordering for the external (sociolinguistic) independent variables, we exclude these variables from the factorial analysis. We include the dependent variable (vowel duration) without distinction in the list of variables to analyse. We can thus directly compare its behaviour to that of the other variables by examining the factor groupings of the entire system.

We perform principal components analyses for each speaking style in SPSS; this test clarifies the latent structure of the system controlling vowel duration by grouping known variables into underlying factors. The test finds the relative weight of each variable that allows the factor to account for the maximum possible variability in the corpus. After the first factor has been determined in this way, the test passes on to the second, taking what is left of the variability and assigning the maximum weight to this factor, and so on. These factors are displayed in a Scree Plot according to their eigenvalue – their underlying value in terms of the amount of variability for which they can account. Factors with an eigenvalue above 1.0 are extracted for use in the rotated components matrix.

A mathematical function synthesises significant factors into the rotated components matrix; this function manipulates the relative orientation of the factors to maximize correlations between variables. We add a varimax rotation with Kaiser normalisation to this function in order to fully accentuate these correlations. The relative weight of each variable is expressed as a number between -1 and 1; this value is significant if it is greater than 0.30 or less than -0.30. A factor is significant if it contains three significant variables. The sign (positive or negative) of a variable’s weight within a given factor grouping indicates its relationship with the other variables. Suppose X and Y are the values of two variables for a given factor. If they share the same sign (+ or -), they are in direct correlation, so that if X increases, so will Y; if they have contrary signs, they are in indirect correlation. The relationships expressed in the matrix depend wholly on the whole number classification system; therefore, a thorough understanding of its organisation is essential to the accurate interpretation of these correlations.

Each factor can explain a certain percentage of the variability presented by each variable. We can calculate this amount by
using the following formula:

\[ 100X^2 \]  

where \( X \) is the weight of the variable for the factor in question. Since vowel duration is one of the variables in the analysis, this calculation is very useful in determining the influence that each factor grouping has on the variability of vowel duration. The other variables are as follows: the nucleus’ phonological class; the vowel’s status as a diphthong; the word’s status as a clitic; the syllable’s position in the lexical word, phonological word, and accentual phrase, respectively; the number of elements in the syllable; the coda’s phonological class; the syllable’s structure (open or closed); and the coda’s lengthening properties.

4. Results

4.1. All speakers

Each principal components matrix contains factors rotated to best fit the data set; the arbitrary nature of this rotation can potentially compromise the reliability of comparing different analyses to one another. We will therefore only consider the two analyses that directly compare our four speakers.

4.1.1. Reading

The Scree Plot for the reading corpus isolates four underlying factors (Figure 1); in the rotated components matrix (Table 1), only three of these factors are significant. The first factor groups the structural properties of the syllable: the number of elements in the syllable, the coda’s phonological class, and whether the syllable is open or closed all indirectly correlate with the coda’s lengthening properties. This underlying factor explains the majority of the variation in the corpus as a whole, but is not at all significant for vowel duration (-0.11).

The second factor generally addresses the data’s position: the word’s status as a clitic, the syllable’s position in the phonological word and its position in the accentual phrase. The syllable’s position in the lexical word is almost significant in its direct correlation with these variables. Their indirect correlation with the number of elements in the syllable and vowel duration indicates that the final position in prosodic or phonological groups, non-clitics, and more complex syllabic structures favour longer vowel durations. This factor explains 12.6% of the corpus’ variability of vowel duration.

The third factor explains 56.7% of the variability of vowel duration in the corpus, and groups variables describing the intrinsic properties of the nucleus. The nucleus’ class is indirectly correlated with the vowel’s duration and its status as a diphthong. Not surprisingly, these correlations confirm that phonologically long vowels and diphthongs tend to have systematically longer surface durations than those of phonologically neutral or short vowels.

4.1.2. Conversation

The Scree Plot (Figure 2) pinpoints five significant underlying factors based on their Eigen values; three are further isolated in the rotated components matrix (Table 2). The groupings are more coherent than those of the reading corpus, and strongly correspond with those of the individual speakers’ results for this style.

The first factor shares many similarities with that of the reading corpus: it shows that while most of the variability of the corpus has to do with the syllable’s structural properties, they do not effectively explain variability of vowel duration. The coda’s status as a lengthening consonant does not figure in this factor as significant; if the fourth and fifth underlying factors were eliminated from the matrix, this variable’s significance would very likely increase.

The second latent factor explains 11.6% of the variability in vowel duration, and clearly assembles variables describing the lexical and prosodic nature of the data. The syllable’s position in the lexical word, phonological word, and accentual phrase, and the word’s status as a clitic, are significant, and are indirectly correlated with vowel duration.

The matrix’s most important subjacent factor for vowel duration is, again, the third and last to consider. This component explains 44.2% of the variability in vowel duration in conversation. The three variables characterising the vowel (vowel duration, nucleus class, diphthong status) are in the same correlative relationship as they are in the reading analysis, solidifying its results.
We saw that syllable distribution remains fairly constant across speaking styles, even though the constituents of the reading style are predetermined by the choice of reading material. Our factorial analyses demonstrated that three ordered, underlying factors explain the variability of our corpus. Of these, we noted that “structure”, the strongest factor in explaining this general variation, is always the weakest in accounting for the variability of vowel duration. The strongest factor in this last respect is “vowel”, and it is more stable in reading than in conversation. We observed that between these two subjacent factors lies the third, “position”; half of the variables of this factor present a high degree of instability. Future studies can analyse the links between these unstable variables, and how these connections may confound results. The overall cohesion of our empirical results suggests that the observed patterns and hierarchy (vowel > position > structure) of factors influencing vowel duration are exact and reliable, and can be tested in future studies on this topic.

7. References