ACOUSTIC MEASUREMENT OF CANADIAN ENGLISH-ACCENTED FRENCH

by

Nicole Gavac
B.A., Simon Fraser University 2002

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
Department of French

©Nicole Gavac 2006

SIMON FRASER UNIVERSITY

Fall 2006

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

APPROVAL

Name: **Nicole Gavac**

Degree: **Master of Arts**

Title of Thesis: **Acoustic Measurement of Canadian English-Accented French**

Supervisory Committee:

Chair: **Dr. Réjean Canac-Marquis**
Associate Professor, Department of French

____

Dr. Christian Guilbault
Senior Supervisor
Assistant Professor, Department of French

____

Dr. Phyllis Wrenn
Supervisor
Associate Professor, Department of French

____

Dr. Murray Munro
External Examiner
Professor, Department of Linguistics

Date Defended/Approved: November 20, 2006

SIMON FRASER UNIVERSITY library

DECLARATION OF PARTIAL COPYRIGHT LICENCE

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Revised: Fall 2006

SIMON FRASER UNIVERSITY library

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, BC, Canada

ABSTRACT

This study examined the speech of female Canadian English learners of French at three proficiency levels in an effort to quantify Canadian English-accented French. Five segmental and two suprasegmental cues were measured on a corpus of delayed-repetition and spontaneous speech: voiceless stop VOT; F2 – F1 acoustic space for / u/ and /y/; F2-shifting for /u/, /o/, and /i/; F2 – F1 acoustic space for /o/ and /ɔ/; F1 of unstressed /a/; temporal reduction of /u/, /y/, and /a/; and a rhythmic Variability Index. Hypotheses predicted more experienced learners would perform more comparably with native speakers, and display less variation, than less-experienced learners. F2 – F1 acoustic space of /u/ – /y/ and /o/ – /ɔ/, and rhythmic variability index values patterned most closely with hypotheses. Results were discussed in terms of acoustic measurement, cue variability, and speaker task performance. Findings provide preliminary data on Canadian English-accented French linking spontaneous and elicited speech.

Keywords: foreign accent; French; female speakers; segmental analysis; suprasegmental analysis; acoustic quantification; rhythm

Subject Terms: second language acquisition; French language – Phonetics

DEDICATION

To my parents and their friends, who inspired me to pursue my curiosity about the phenomenon of foreign accents… and especially to Mr. Stepan Cervinek, who could inspire a whole research program on the subject!

ACKNOWLEDGEMENTS

I would like to thank my senior thesis supervisor, Dr. Christian Guilbault, for guiding me through my Master's program, and for providing endless hours of support throughout the thesis writing process. Many thanks also to Dr. Phyllis Wrenn and Dr. Murray Munro, for their expertise in shaping the final versions of this work. The Department of French at SFU cultivated my academic curiosity, and afforded me a nurturing learning environment… thank you all for your commitment and enthusiasm!

My acknowledgements would not be complete without recognizing the unwavering support of my fellow graduate students. I consider them an inspiration to my professional life, and wish them the best in all of their endeavours. Specifically, I would like to extend my indebtedness to Andrea Williams, whose friendship played an integral role in maintaining my focus and sanity! I would also like to thank my friends and family – especially my partner Jesse, whose passion and commitment serves as an inspiration to any mortal pursuing academic excellence.

Finally, I am thankful to many sources of funding that helped support my research, such as the Social Sciences Humanities Research Council, the C.D. Nelson Memorial Graduate Scholarship, the Dr. Tai Whan Kim Memorial Graduate Scholarship, and the SFU Graduate Fellowship. These financial contributions made it possible to focus on my studies, and I appreciate these institutions' dedication to the promotion of research in the Social Sciences.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background

Since the nineteen-sixties, the study of foreign accent in North America has seen significant growth. Examining foreign accent has become an important issue, given that research in this field could further our understanding of certain aspects of second language acquisition, as well as our understanding of the link between speech perception and production. The study of foreign accents in English has dominated the field, while foreign accent as manifested in other target languages has been relatively less researched. Although English-accented French has been the object of substantial investigation (Brown & Cichocki, 1995; César-Lee, 1999; Champagne-Muzar, Schneiderman & Bourdages, 1993; Flege, 1987b; Flege & Hillenbrand, 1984; Greasley, 1989; Grover, Jamieson & Dobrovolsky, 1987; Guilbault, 2002; Le Clézio, 1986; Neufeld, 2001; Rochet, 1995; Walz, 1979; among others), these studies have not examined this foreign accent from the same perspective. Instead, they have approached the topic from diverse perspectives, such as speech perception, speech production, sociolinguistic and/or pedagogical implications. Therefore, cross-comparison of acoustic data across these studies is difficult. Furthermore, detailed speaker data from these studies is often unavailable, which renders speaker comparisons difficult as well. Specifically, the need remains for rigorous acoustic quantification of English-accented French speech productions of a relatively well-defined group of speakers, preceding perceptual judgments of such productions.

1.2 Statement of the Problem

Previous characterizations of Anglophone accent in French have relied heavily on contrastive phonological analyses in order to identify the salient cues responsible for perception of such a foreign accent (Casagrande, 1984; Dansereau, 1990; Garant Viau, 1994; Valdman, 1993; Walker, 2001). Moreover, most such descriptions of the features of English-accented French do not rely on experimental quantification. These characterizations have predominantly made use of impressionistic perceptual information on the part of native-speaking educators for the purpose of improving the pronunciation of foreign-language students. While this approach can present a clear orientation for additional research on foreign accent, it cannot explain the nature of foreign accent, such as the articulatory, acoustic, or perceptual aspects of the phenomenon.

Relatively few studies have used acoustic measures to quantify Canadian English-accented French, and those that have either studied only a small number of selected parameters or did not treat the parameters systematically. Of the previous studies of English-accented French, only César-Lee (1999) identified certain segmental and suprasegmental cues as potential carriers of foreign accent correlating with perceived degrees of foreign accentedness. The cues identified are:

- The difference between F2 – F1 frequency of /y/ and /u/
- Total duration of utterance
- Voice-onset time (VOT) of syllable initial /p/
- Syllable duration

Unfortunately, these cues have been identified only in that one study. In addition, other cues that may contribute to the perception of a Canadian English-accent in French may not have been studied systematically. Although other studies may have examined a specific segmental or suprasegmental feature of English-accented French (e.g., intonation patterns, by Grover, Jamieson and Dobrovolsky, 1987; the perception of /y/, by Rochet, 1995; or VOT values of /t/, by Flege and Hillenbrand, 1984), a detailed acoustic account of segmental and suprasegmental cues for the same speakers could provide a more comprehensive understanding of a speaker's accented productions. In addition, no previous acoustic production studies examining English-accented French have utilized both delayed-repetition and spontaneous speech samples, in an effort to draw further accented-speech production comparisons.

1.3 Purpose of the Present Study and Research Questions

The present research offers a more global consideration of the nature of Canadian English-accented French as substantiated by empirical acoustic data. A physical-acoustic correlate to the theoretical notion of foreign accent would not only expand our understanding of the nature of phonetic acquisition of a second language (L2) but could also provide some of the much-needed groundwork in speech production necessary for determining sources for the perception of a foreign accent.

The primary purpose of the present study is to identify and quantify some of the most salient acoustic cues of Canadian English-accent in French as proposed in previous research. How does non-native speech compare to native speech in French? Is acoustic variation in segmental and suprasegmental production readily identifiable between native and non-native speech samples? The present acoustic experimentation can then serve as a foundation for subsequent perceptual studies, which could substantiate arguments for the relative influence of specific segmental and prosodic features of speech on the perception of a foreign accent.

Theoretical issues in the field of foreign accent abound, from difficulty in defining the notion of foreign accent to the idiosyncratic experience of producing or perceiving foreign accent (Markham, 1997). Many studies have examined foreign accent through perceptual tasks (see Piske, MacKay & Flege, 2001, for a review), and yet many others have sought the acoustic quantification of foreign-accented productions (Caramazza, Yeni-Komshian, Zurif & Carbone, 1973; Flege, 1987b; Flege, Munro & Skelton, 1992; Shah, 2002). A relatively few have actually sought to correlate production values with the perception of foreign-accented speech (César-Lee, 1999; Jonasson & McAllister, 1972; Magen, 1998). In part because of the paucity of such acoustic-perceptual studies, the relationship between the production and the perception of foreign accent still merits academic consideration.

In order to address some of the empirical issues involved in understanding the phonetic acquisition of an L2, such as the need for more acoustic quantification of foreign-accented speech, a speech production experiment was conducted examining the utterances of adult female Anglophone learners of French at three stages of language acquisition (beginner, intermediate and advanced). The hypotheses motivating this study stem from previous phonological and acoustic research on other foreign-accented speech (Jonasson & McAllister, 1972; Magen, 1998; Shah, 2002), where researchers found that those L2 learners' speech productions which have greater acoustic variability or which differ considerably from native speakers' productions are perceived as more accented than L2 learners' speech productions which pattern more closely to native speaker productions.

In the present study, Anglophone learners of French and native-French speakers were recorded and their utterances analyzed acoustically. Specifically, the segmental cues which were quantified include measurements of 1) VOT of initial voiceless stops, 2) F2 – F1 acoustic space for /u/ and /y/, 3) diphthongization of /u/, /o/, and /i/, 4) F2 – F1 acoustic space for /o/ and /ɔ/, and 5) F1 spectral quality in unstressed environments for /a/. The suprasegmental elements of speech the present study examined are 1) temporal vowel reduction for /u/, /y/ and /a/ in unstressed environments, and 2) rhythmic variation in spontaneous speech (measured by the Variability Index). These acoustic parameters were chosen because of their alleged importance in previous research for influencing the perception of a foreign accent (César-Lee, 1999; Dansereau, 1990; Flege & Hillenbrand, 1984; Guilbault, 2002; Le Clézio, 1989; Nearey & Rochet, 1994; Rochet, 1995; Valdman, 1993; Walz, 1980).

This study takes an experimental approach in attempting to determine how the acoustic cues under analysis could prove relevant to the perception of a foreign accent in French. Theoretical frameworks and models of second-language acquisition, such as the contrastive analysis hypothesis, markedness theory, Flege's Speech Learning Model (1995), and Best's Perceptual Assimilation Model (1995), are also presented, because of their influence in shaping hypotheses on accented pronunciation.

1.4 Organization of the Present Study

This research presents the results of two experimental tasks (a delayed-repetition sentence-production task and a spontaneous-speech elicitation task) dealing with the acoustic measurement of Canadian English-accented French. Prior to an assessment of these experiments in chapters 4 and 5, Chapter 2 will provide a review of the literature, which discusses foreign accent theory, factors affecting the degree of foreign accent, methods used in the study of foreign accent, models predicting pronunciation difficulties, and an overview of Canadian English-accented French. The chapter concludes with a detailed presentation of the hypotheses motivating the current study.

Chapter 3 presents the methodology used in the development and undertaking of the two experimental tasks conducted in the scope of the present research. These tasks include recording speakers of French at three stages of language acquisition (beginner, intermediate and advanced), as well as native speakers of French. Specific methodological considerations, such as subject selection, stimuli development, and speech material measurement and analysis, are discussed in detail.

Chapter 4 reports on the results from the two experimental tasks listed above. Chapter 5 provides a discussion of the results, as framed by the treatment of theoretical and methodological issues pertaining to foreign accent in the literature review. A comparison between the current data analysis and previous empirical results obtained in the study of analogous segmental and suprasegmental cues is presented. The chapter concludes with suggestions of future directions for research.

1.5 Significance of the Study

The present study provides a systematic evaluation of several acoustic segments as well as spontaneous speech production of non-native and native speakers of French, in order to address the empirical need for quantification of Canadian English-accented French. Specifically, the research presents VOT and vowel formant values across multiple segments, allowing greater intra-speaker and intra-segmental comparison. Furthermore, the study examines both elicited and spontaneous speech, allowing for additional levels of data assessment within the same speaker set. Finally, the study provides acoustic values for female speech in French, much needed in the field due to a paucity of such values.

While its conclusions may not be definitive, this study provides a point of comparison for future endeavours dealing with the aforementioned topics. A perceptual correlate of degree of foreign accent to the collected speech samples, such as auditory-perceptual ratings, would substantiate the effect of acoustic variational patterns found in speakers' segmental and suprasegmental productions (Markham, 1997; Munro, 1993; Wayland, 1997).

2 LITERATURE REVIEW

2.1 A Theoretical Look at Foreign Accent

The study of foreign accent involves an analysis of the complex interactions between the production and the perception of speech. Although formally studied in North America for decades, the notion of foreign accent remains difficult to characterize due to the many factors implicated in its treatment (Burgess, 2001; Flege, 1995; Markham, 1997; Munro, 1998). Definitions of foreign accent abound, but are often of a problematic nature (Burgess, 2001). In this section, several definitions of foreign accent will be presented, as well as the many theoretical implications raised by these multifaceted approaches to this phenomenon.

When discussing the notion of foreign accent, it is necessary to specify the linguistic elements involved in its treatment. Munro (1998) defines foreign-accented speech as "non-pathological speech produced by second language learners that differs in partially systematic ways from the speech characteristics of native speakers of a given dialect" (p. 139). César-Lee (1999) speaks of the "faulty production of a target language (L2), due to faulty perception, faulty articulation, or a combination of both" (p. 1) when referring to foreign accent. Flege (1995) identifies foreign accent in English from the listener's perspective: "listeners hear foreign accents when they detect divergences from English phonetic norms along a wide range of segmental and suprasegmental (i.e., prosodic) dimensions" (p. 233). Shah (2002, p. 3) echoes Flege's remarks by emphasizing the intrinsically perceptual nature of foreign accent. These definitions expose the complicated nature of the concept of foreign accent, as they recognize the many linguistic elements involved in its treatment, whether considering the phonetic characteristics of a non-native utterance, the perception of the native listener, or the potential underlying causes of a foreign accent.

Markham (1997) provides us with an especially involved and detailed conceptualization of foreign accent. While conceding that the only objective judgment in isolating a foreign accent is the identification of a deviant phonetic characteristic in non-native speech, Markham underscores the essentially subjective nature of the concept of foreign accent (p. 85). According to this account, each listener experiences a foreign accent in a wholly idiosyncratic manner, determined by the listener's concepts of *native* and *local*[1]. These notions are affected by:

1) The listener's beliefs and knowledge about himself within a group,
2) How the listener regards his own group in relation to other groups,
3) Experience and knowledge of other groups.

Markham's views highlight the sociolinguistic underpinnings of the phonetically salient foreign accent phenomenon. Moreover, he clearly distinguishes many types of foreign accents, such as listener-subjective foreign accent and acoustic accent, which are the most relevant to the present study. Markham states that the two most commonly-used approaches to studying foreign accent – physical-objective quantification and acoustic measurement of foreign accent, and listener-subjective judgments of foreign accent – each serve an important, specific purpose in furthering research, but ultimately provide

[1] Here, the terms *native* and *local* are used to make a distinction between *native local accent, native non-local accent*, and *non-native accent* (Markham, 1997, p. 85).

the most valuable analysis when used jointly. Munro (1998) and Wayland (1997) echo the need for perceptual correlates to the acoustic dimensions of foreign accent.

Other theoretical concepts important to the study of foreign accent include the *intelligibility* and *comprehensibility* of non-native utterances. Derwing and Munro (1997; Munro & Derwing, 1995) maintain that these concepts are related to (and sometimes confused with) foreign accent, but that the relationship between these three notions is partially independent. In more recent research, the authors have clearly distinguished the three dimensions (Munro & Derwing, 2001, p. 454):

Accentedness	The degree to which the listener believes an utterance differs phonetically from native-speaker utterances.
Comprehensibility	The degree of difficulty the listener reports in attempting to understand the utterance.
Intelligibility	The extent to which a particular utterance is actually understood.

Although the three notions are related and the significance of each is noted, only the concept of *accentedness* will be focused on due to its relevance to the present study of the acoustic nature of foreign accent.

For the purposes of the present study, Munro's (1998) definition of foreign accent will be used, given that the focus of the current project is on production, rather than perception. Munro's claim that foreign accented speech "differs in partially systematic ways from the speech characteristics of native speakers of a given dialect" (p. 139) is particularly noteworthy because the aim of the present study is, essentially, to isolate how Anglophone accent in French differs "in partially systematic ways" from the utterances of native French speakers.

2.2 Factors Affecting Degree of Foreign Accent

Having established a working definition of foreign accent, an examination of its causes is relevant, and could potentially highlight some of the acoustic phenomena possibly identified or associated with the manifestation of a foreign accent in a non-native speaker. Although all of the commonly identified factors contributing to the manifestation of a foreign accent in the non-native speaker's speech will be presented, only those most pertinent to the foreign-language student will be explored more seriously.

Piske, MacKay and Flege (2001) provide a comprehensive review of the factors that are claimed to affect the degree of foreign accent in an L2. The authors identify the eight most commonly researched variables believed to influence the degree of perceived foreign accent in an L2:

1) Age of L2 learning (AOL)
2) Length of residence (LOR)
3) Gender
4) Formal Instruction
5) Motivation
6) Language learning aptitude
7) Amount of L1 use
8) Amount of L2 use

The four variables that will be considered in the following discussion include *AOL*, *formal instruction*, *motivation* and *amount of L2 use*. The other variables will not be discussed due to their lack of relevance or application to the proposed study. Clearly, the variable *length of residence* does not apply in the foreign language classroom, neither does *amount of L1 use*[2]. *Gender* will not be a variable in this study as a result of experimental protocol, which includes the examination of female speech only. The

[2] Within the scope of the present study, the amount of L1 use is assumed to be equivalent across learners, as L2 use is relatively restricted to the foreign language classroom.

language learning aptitude variable will not be taken into account because of the inconclusive nature of limited research on the topic (Piske, MacKay & Flege, 2001).

Although the variable *age of L2 learning* has frequently been identified as the best single predictor of degree of foreign accent in non-native speakers (Asher & Garcia, 1969; Tahta, Wood & Loewenthal, 1981; Thompson, 1991), the role that other factors play in the perception a foreign accent is not negligible. The relationship between a younger age of acquisition and a less perceptible foreign accent is not fixed (Purcell & Suter, 1980), as some young learners may retain a foreign accent, while certain older learners attain accent-free speech in their L2 (Flege, Munro & MacKay, 1995; Moyer, 1999; Thompson, 1991). The well-known effect of AOL on L2 acquisition has been drawn on as evidence of a critical period in language learning. However, the hypothesis favouring neurological maturation as the main cause of restricted pronunciation performance – the critical period hypothesis – has been contested, and alternate explanations suggested (see Archibald & Libben, 1995, for a review). For example, the interaction or interference hypothesis (Yeni-Komshian, Flege & Liu, 2000) and the sensitive period hypothesis (Long, 1990) have been proposed as alternatives to the critical period hypothesis, to account for the more general pronunciation tendencies of non-native speakers.

The AOL variable is relevant to the present study because the speakers comprising the experimental groups of learners have varying AOL for French. However, the circumstances of their language learning are incongruent with the standard L2 acquisition setting since it is uniquely the formal classroom that serves as the principal learning environment, and not the society at large.

The other three variables – *motivation, amount of L2 use*, and *formal instruction* – have varying implications for the present research. While motivation to speak the L2 without a foreign accent has been shown to correlate with the degree of foreign accent – as has an increased amount of L2 use – an increased amount of formal instruction has not been shown to affect the degree of foreign accent (Flege & Fletcher, 1992; Moyer, 1999; Piske, MacKay & Flege, 2001; Thompson, 1991). Piske, MacKay and Flege (2001) propose that a lack of attention to pronunciation teaching in the foreign language classroom may explain the insignificant effect of instructional variables. However, as the present study is evaluating the speech of foreign language students who only have access to the L2 in the foreign language classroom, the amount of formal instruction has become an independent variable, in effect defining the three experimental groups.

While it is necessary to make a relevant and concise characterization of foreign accent within the present framework, Burgess (2001) correctly underscores the actual state of the theoretical discussion surrounding the causes of foreign accent: "Suffice it to say that in our current state of knowledge it is impossible to assign foreign accent to a single cause or even a definitive congeries of causes that will cover all instances" (p. 27). Evidently, the domain of foreign accent remains complex and continues to require the critical attention of researchers.

2.3 Frameworks for Studying Accented Speech

Research on foreign accent has exploited various methods of studying accented speech, such as acoustical analysis, perceptual analysis, and integrated acoustic-perceptual analysis of foreign accent. Because of the proliferation of their use, these methods prove to be most pertinent to the present goal of quantifying Canadian English-accented French.

Furthermore, researchers have developed language-learning models intending to account for various parameters of foreign-accented speech. The following section will introduce these major frameworks used in the study of accented speech.

2.3.1 Acoustical Analysis

In order to characterize a particular foreign accent, researchers have often measured the acoustic values of segmental and suprasegmental elements of speech. The actual method used in the acoustic analysis depends on the element under investigation. Acoustical properties of vocalic segments have been quantified through the measurement of formant patterns and durational quantification. Examination of stop consonantal segments has included the measurement of voice-onset time (VOT). The measurement of suprasegmental features of non-native speech has included intensity, duration, inter-syllabic variation, and fundamental frequency contours of the utterance. This acoustic analysis of non-native, accented speech has also often consisted of a subsequent comparison to the acoustic measurements of native speaker productions. While this method has been used in order to study or quantify non-native speech, this approach does not propose any predictions for L2 speaker pronunciation difficulties.

Much previous cross-language and foreign accent research has relied on acoustic analysis. The following studies illustrate the utility of such an approach in quantifying variation in pronunciation. Mack (1982) measured voicing-dependent effects on vowel durations of native English speakers, native French speakers and French-English bilinguals, and discovered that the acoustic values of French monolingual (and bilingual) utterances differed significantly from those of the Anglophones' utterances. Flege (1984) found that significant differences between the VOT values of French and American speakers in the production of /t/ may be related to the perception of a foreign accent. Gottfried and Beddor (1988) manipulated vowel duration as part of a synthetic continuum in a /k/ V /t/ context in order to determine its perceptual saliency to French and to English speakers. Magen (1998) made several acoustic measurements of Spanish speakers' English utterances (both segmental and suprasegmental) with the intention of isolating the acoustic parameters of foreign-accented speech. These studies demonstrate that it is possible to differentiate speaker proficiency levels based on the acoustic measurements of specific segments, because speakers display noticeably different values on the analyzed speech parameters.

The acoustic analysis of foreign-accented speech has been used frequently in an attempt to quantify an accented speech signal (Wayland, 1997). In theory, by measuring and comparing native and non-native utterances, the acoustic properties of foreign-accented speech can be revealed. Unfortunately, not all speech properties are easily defined and/or attributable to a single acoustic measurement. Because of the inherently variable nature of speech, it is often hard to extract relevant acoustic properties of segments under investigation. Therefore, it is often difficult to attribute the presence of a foreign accent to differences in the speech signal. While some aspects of speech lend themselves well to measurement (VOT, formant frequency patterns, duration), others may not (voice quality, /ʀ/ quality). Because certain phonological properties of speech are inherently difficult to quantify, their effect on the perception of a foreign accent, as well as their role in foreign-accented speech in general, has not been adequately assessed.

As such, the explanatory power of such an approach has been questioned. Munro (1993, p. 40) clearly indicates some of the shortcomings of acoustic analysis:

> One of the drawbacks of comparing sets of acoustic data in L2 studies is that such an approach does not necessarily reveal which characteristics of the non-native productions cause native listeners to hear them as accented. Microscopic acoustic analyses may reveal differences between two sets of data on several parameters. However, these differences may have varying degrees of importance in *perceived* accentedness.

Nonetheless, such reservations do not render the use of acoustic analyses invalid when studying foreign accent. More exactly, they underscore the need for a balanced approach to foreign accent research. Acoustic analyses are vital to the conception of foreign-accented speech but, without perceptual correlates, cannot account for a complete understanding of the phenomenon.

Notwithstanding this *caveat*, the present study recognizes the acoustic approach as an integral first step in the understanding of Canadian English-accented French. The main objective of the project is to contribute to the body of research that quantifies many aspects of Canadian English-accented French by providing a detailed and consistent acoustical data set of multiple parameters of L2 speech.

2.3.2 Perceptual Analysis

The perceptual approach to the evaluation of foreign accent has been used extensively in previous research. As foreign accent is predominantly defined by a perceptual construct (as per section 2.1 above), subjective native-listener judgments have often been included in characterizations of L2 learners' foreign accent (studies such as Asher & Garcia, 1969; Brennan & Brennan, 1981; Jonasson & McAllister, 1972; Magen, 1998; Piske, Flege & MacKay, 2001; Tahta et al., 1981; Thompson, 1991, among others).

Essentially, a perceptual approach to the study of foreign accent consists of native speakers rating speech productions for accentedness according to subjective, internal standards of native speech (Markham, 1997). These ratings are usually conducted using a fixed interval scale (7-, 9-, or 11-point scale) where one end of the scale represents heavily foreign-accented speech production, while the other represents foreign accent-free speech production. Southwood and Flege (1999) studied foreign accent rating techniques in order to determine whether linear partitioning is a valid scale to use for the perceptual quantification of accentedness. The authors found that a 9- or 11-point interval scale was most amenable for accentedness judgment of Italian ESL speakers, in order to avoid the ceiling effects of smaller interval scales. The researchers also found a high degree of intra-judge reliability, while inter-judge reliability was lower. These reliability discrepancies were potentially attributed to differences between judges' internal standards of foreign accent, scaling artifact, and the lack of a physical referent for foreign accent. The authors highlight the need for future studies determining how familiarity with foreign accent may influence perceptual ratings, as well as foreign-accent research "that relates acoustic variables to global judgements of degree of perceived foreign accent" (p. 348), in order to provide definable physical units to the notion of accentedness.

Native-speaker judgment of foreign accent has been correlated with several aspects of the L2 learner's linguistic experience, such as age of acquisition, and length of residence (Piske, Mackay & Flege, 2001). However, one of the disadvantages of using exclusively a perceptual approach to foreign accent is that while speaker variables can be accounted for (such as linguistic experience), it is not possible to identify acoustically which parts of the speech samples contribute to the perception of a foreign accent. Therefore, Markham's (1997) call for an integrated approach, using both acoustic analysis and perceptual correlates to foreign accented speech, takes advantage of the benefits of both methods, while addressing their weaknesses.

2.3.3 An Integrated Approach

Many studies have used the integration of both acoustic and perceptual approaches to studying foreign accent (Jonasson & McAllister, 1972; Magen, 1998; Munro, 1993). Jonasson and McAllister (1972) provide an excellent illustration of this approach. The goal of their study is twofold: 1) to identify consistencies in the relationship between the acoustic signal of American-accented Swedish speech and the perception of this signal by a native speaker, and 2) to establish which of the acoustic deviations had a more pronounced effect on the perception of foreign accent.

In their first experiment, the authors recorded one American speaker (AS) and one Swedish speaker (SS) producing two lists of nonsense words (one following the phonotactic constraints of English, the other, those of Swedish). The utterances were then analyzed for consonant, vowel, and syllable durations. The authors found that while SS made a consistent distinction between long and short vowels in the Swedish words, AS's utterances displayed higher acoustic variability and inconsistent durational distinctions. The second experiment consisted of a perceptual task for 20 native Swedish speakers to determine the perceptual boundary of vowel consonant ratios (*/ett/* vs. */e:t/*), as well as the degree of naturalness of the utterance. It was found that the listeners indicated a well-defined perceptual boundary between the two stimuli. When the acoustic values of AS's utterances in experiment 1 were compared with the perceptual values considered *native* in experiment 2, most of the speaker's productions did not fall within the category, and would therefore never have been considered natural. Based on these findings, the authors concluded that it is in fact possible to correlate acoustic measurements with perceptual judgments of foreign accent. They further argue for the existence of an *accent component* accounting for faulty L2 learner speech production, composed of both incorrect pronunciation rules based on L1 interference, as well as learner hypotheses about the target language, which are incorrect.

Magen's (1998) study of Spanish-accented English correlated judgments of degree of foreign accent with acoustic measurements of several segmental and suprasegmental characteristics of L2 speech. She further manipulated the speech signals synthetically, in an effort to remove certain non-target-like acoustic features, in order to identify the relative influence of different types of pronunciation errors in a subsequent judgment of degree of foreign accent. The manipulated acoustic features of phonological elements include syllable structure, vowel, consonant, and stress quality. Upon acoustic manipulation, the listeners were perceptive of syllable structure, final /s/ deletion, consonant articulation, and lexical and sentence stress, but not of voicing differences. The importance of Magen's contribution lies 1) in her manipulation of natural speech (as opposed to synthetic speech), which may preserve certain natural acoustic dimensions in the stimulus set, and 2) in her comparison of acoustic measurements with judgments of foreign accent, thereby permitting an evaluation of relative influence of segmental and suprasegmental features of the accented speech. However, the author justly notes the challenges of using synthetic manipulation of natural speech: certain acoustic parameters are more easily manipulated synthetically, while others are not possible to modify, thus potentially skewing the data set; and the resulting manipulated speech does not necessarily sound natural, which could thereby reverse the benefits of using modified natural speech stimuli in the first place.

While this study recognizes the benefits of an integrated approach to the study of foreign accent, only acoustic measurements will be undertaken, because of the need to obtain a larger data pool composed of multiple segmental and suprasegmental acoustic parameters (as opposed to a more limited sample), and because of the aim to compare elicited speech with spontaneous speech performance. However, a subsequent perceptual analysis of the dataset obtained in the present research would not only be prudent, but desirable, due to the possible explanatory findings resulting from such an additional analysis.

2.3.4 Language-Learning Models

Several language-learning models have been proposed to account for foreign-accented speech, including the Contrastive Analysis Hypothesis (CAH) (Weinreich, 1953), Markedness Differential Hypothesis (MDH) (Eckman, 1977), the Perceptual Assimilation Model (PAM) (Best, 1995), and the Speech Learning Model (SLM) (Flege, 1995), which are outlined below. The goal of these models has been to predict pronunciation difficulties, thereby characterizing the foreign accent of L2 learners.

Contrastive Analysis Hypothesis (CAH). Making comparisons between languages seems almost instinctive when dealing with the learning of an L2. The use of a contrastive approach to the teaching of an L2 has been well documented. Formal contrastive phonological analyses were first developed in the mid 1950s (Weinreich, 1953). The contrastive phonological approach postulates that pronunciation difficulties arise from a discrepancy between the phonemic inventories of the native and target languages. This approach is based on developing a detailed phonological inventory of both languages and studying those phonemic differences. The prediction, identification, and isolation of pronunciation problems can then be realized. However, the formalization of the contrastive phonological analysis approach has led to the identification of limitations and faulty predictions inherent in this method, such as a lack of procedural objectivity (Jackson, 1971), and inadequate phonological or phonetic description (Brière, 1966; Rochet, 1995).

Currently, it is acknowledged that this method fails to account for all pronunciation problems of various native speakers of a target language (for instance, Wardhaugh, 1970). For example, in the case where a target phoneme of the L2 does not exist in the L1, native speakers of various L1s may not make the same phoneme substitutions. Rochet (1995) studied the production and perception of the French /y/ by English and by Portuguese speakers. The study consisted of two tasks: 1) an imitation task in which subjects were asked to repeat a list of monosyllables containing the vowels /i/, /u/, /y/, and /a/, and 2) a perceptual task in which subjects were asked to identify synthetic stimuli as /i/ or /u/ (on a high vowel continuum). While neither of the phonological inventories of the two languages contains /y/, English speakers substituted /u/ and Portuguese speakers substituted /i/ for the target vowel /y/. Rochet's perceptual analyses indicated that the subjects' substitutions corresponded to their perceptual boundaries of the phonemes in the L1. Jonasson and McAllister (1972) further assert that the contrastive method has little explanatory power when attempting to predict L2 speech errors, which contribute to the perception of a foreign accent. In their study of American

English-accented Swedish, the L2 speaker's pronunciation deviations in Swedish occurred with the short/long vowel distinction, a phonological opposition present in both languages' vowel systems. These examples clearly illustrate the limitations of the predictive power of the contrastive phonological hypothesis.

In effect, although influence of the L1 on the acquisition of an L2 cannot be ignored (Zybert, 1997), simple first-language transfer does not adequately explain all instances of the manifestation of a foreign accent (Garnica & Herbert, 1979). Because contrastive phonological analysis is based on abstract linguistic notions such as the phoneme, the resulting predictions of pronunciation difficulties are not necessarily phonetically salient to the learner. For example, this approach does not account for how a learner will deal with allophonic variation in a target language. Therefore, current research has made use of more detailed acoustic information thought to be relevant to the learner in the elaboration of complex L2 learning models capable of more accurate predictions of pronunciation difficulties, based on the notion of cross-linguistic influence. Nonetheless, contrastive phonological analysis is one of the most developed approaches for predicting the pronunciation difficulties of L2 speakers.

While the predictive power of contrastive analysis is limited, the application of a contrastive approach based on the notion of cross-linguistic influence serves as a useful first step in establishing the parameters of comparison between two languages. For example, Shah (2002) examines the temporal characteristics of Spanish-accented English related to the perception of accentedness by native speakers of American English. In developing her hypotheses, the author makes extensive use of a contrastive analysis of the phonological rules governing English and Spanish. She subsequently predicts that

Spanish speakers will have difficulties with the aspects of English pronunciation that differ from those of Spanish. Undertaking acoustic and perceptual analyses on Spanish L2 learners of English, Shah concludes that the pronunciation deviations of Spanish speakers could be predicted by phonological analyses, and that the perceived accentedness of Spanish speakers' productions of multisyllabic words in English may in fact be related to the temporal deviations found in their speech.

In the foreign language classroom, contrastive phonological analysis is the principal method used in most pronunciation teaching materials. In the case of teaching French to Anglophones, the identification of particular pronunciation problems specific to these learners is based on a comparison of the two phonological systems (Casagrande, 1984; Dansereau, 1990; Green & Poulin, 1971; Valdman, 1993). Consistent with this precedent, this study will make use of comparative analysis between English and French as an initial step towards elaborating hypotheses predicting the acoustic deviations of Canadian English-accented French from native French speech.

Markedness Differential Hypothesis (MDH). This approach to language-learning proposes that the universal markedness of phonological characteristics in the target language can serve to predict the relative difficulty in the pronunciation of these forms (Eckman, 1977). Therefore, phonological forms in the target language that differ from those in the L1, will be more difficult to acquire if they are marked[3], while those forms that differ but are unmarked will be relatively easier to learn. The notion of markedness is

[3] Within the theory of markedness, Crystal (2003) defines the terms *marked* and *unmarked* as follows, "an unmarked property is one which accords with the general tendencies found in all languages; a marked property is one which goes against these general tendencies – in other words it is exceptional (…). Markedness in this sense can be represented as a continuum along which language-universal and language-specific properties can be related." (p. 284).

generally attributed to typological markedness, which is based on tendencies found across languages in phonological inventories. However, the MDH presents certain theoretical and practical limitations, such as when learners are more challenged by an unmarked form, or determining the level of markedness of a form.

Cichocki, House, and Lister's (1997) investigation of Cantonese-French interlanguage variants of French nasal vowels provides a clear example of the limitations of the MDH. The authors conclude that while the MDH was able to provide some useful predictions for identifying the level of difficulty for the acquisition of certain French nasal vowels by Cantonese speakers, it was unable to account for marked pronunciations transferring from L1 to L2. For example, the Cantonese learners retained marked diphthongal pronunciations, even though the target monophthongal French nasal vowels are unmarked in relation to vocalic diphthongization.

Perceptual Assimilation Model (PAM). Assuming a direct realist view of cross-language speech perception, Best (1994, 1995) outlines a framework of perceiving non-native speech known as the Perceptual Assimilation Model (PAM), which predicts the L2 speaker's patterns of perceptual assimilation or categorization patterns of L2 segments. The author claims that L2 speakers compare non-native segments with native segmental constellations, often integrating L2 segments to the nearest native segment in the perceived phonological space. Non-native segments can be assimilated to a native category as an uncategorizable speech sound, or as a non-speech sound. Moreover, non-native contrast assimilation follows from these segmental assimilation patterns. Finally, Best (1994) posits that language perception and production are informationally compatible, meaning that speakers produce L2 sounds according to their perception of the target sound. This production-perception correlation suggests that L2 learners accurately perceiving L2 sounds will produce the segments more accurately than segments they miscategorize. Best (1995) demonstrates the various assimilation patterns used by American listeners in their perception of several non-native vowel contrasts. However, no explanations are given as to why listeners draw on different perceptual strategies in the various contrast contexts.

While this approach provides important direction for cross-language research, the issue of quantifying perceptual distance between L1 and L2 speech sounds, or 'gestural similarities' between the two sounds, remains problematic. Furthermore, while the model assumes a direct relationship between production and perception, it does not provide any means to confirm such an assumption.

Speech Learning Model (SLM). This model "is concerned primarily with the ultimate attainment of L2 pronunciation, so work carried out within its framework focuses on bilinguals who have spoken their L2 for many years, not beginners" (Flege, 1995, p. 238). Flege claims that many L2 pronunciation errors are perceptually motivated, but not exclusively (other factors include the phonotactic constraints of the L1, among others). The SLM makes explicit predictions for the production and perception of L2 segments based on specific postulates and hypotheses. The SLM is the most robust speech-learning model available for foreign accent speech production research. While other models propose more abstract levels of organization or conceptualisation for the prediction of accented speech, the SLM not only deals with these abstract levels of conceptualisation, but it also provides pronunciation predictions on a phonetic level of representation, which is most relevant to acoustic measurement studies.

While innovative, the model has been critiqued for certain shortcomings (Guilbault, 2002; Rochet, 1995, among others). Flege makes use of the 'new category' concept in his second and third hypotheses (H2 and H3[4]), which predict that learners pronounce 'new' phones better than 'similar' phones. However, Rochet's main concern is the validity of the notion of 'new category' for L2 speech sounds. The difficulty in determining acoustic distance between two sounds in L1 and L2 lies in the criteria used to quantify such an acoustic space. Conditions governing the perception of a truly 'new' category are relatively rare, because the very existence of "uncommitted space" is questionable. Nonetheless, this model provides the most comprehensive framework for predicting accented speech available. The SLM will be used to shape the hypotheses predicting non-native pronunciation in the scope of this study.

2.4 Canadian English-Accented French

In describing Canadian English-accented French it is important to include both segmental and suprasegmental properties of speech. Flege (1987a) affirms the influence of segmental articulation in foreign-accented speech, which supports this study's objective of examining vowels and consonants. Furthermore, previous research also identifies suprasegmental production in the L2 as an even more important factor in the perception of a foreign accent (Anderson-Hsieh, Johnson & Koehler, 1992). Specific segmental and

[4] H2: a new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.
H3: the greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned.

suprasegmental elements commonly identified as problematic between French and English are outlined below, in order to validate the present research decisions of analyzing particular aspects of French target language production.

2.4.1 Segmental Characteristics

The phonological characteristics of English and French have been well-documented throughout the literature (Casagrande, 1984; César-Lee, 1999; Dansereau, 1990; Ostiguy & Tousignant, 1993; Valdman, 1993; Walker, 2001), as well as the many phonetic contrasts found between the two languages (César-Lee, 1999; Delattre, 1981; Flege, 1984; Mack, 1982; Nearey & Rochet, 1994; Walz, 1980). This section will examine the most relevant differences between the English and French systems through an exploration of the consonant and vowel inventories, which will lay the framework for the research and methodological considerations undertaken in the present research. Segments not considered to pose difficulties for Anglophone learners of French will not be treated. For example, neither the production of voiceless fricatives, nor the production of /m/ and /n/, are assumed to contribute to the perception of an Anglophone accent in French because English speakers seem to produce these segments in a target-like manner (Le Clézio, 1986).

2.4.1.1 Consonant Inventories

Table 2.1 presents commonly accepted phonetic differences between the consonant inventories of English and French, which pose pronunciation difficulties for Anglophone learners of French (Casagrande, 1984; César-Lee, 1999; Dansereau, 1990; Flege, 1984; Valdman, 1993; Walz, 1980) relevant to this study.

Table 2.1 Highlights of Comparison for English and French Consonants.

Segment	*French*	*English*
/p/, /t/, /k/	Unaspirated in syllable-initial position; released in syllable-final position	Aspirated in syllable-initial position; often unreleased in syllable-final position
/b/, /d/, /g/	Voicing lead	Unaspirated, voiceless
/ʀ/	Uvular trill, lip position dependent on following segment; considerable dialectal differences	Retroflex consonant with lip rounding; considerable dialectal differences

While the production of French /ʀ/ is considered to contribute noticeably to the perception of a foreign accent (Dansereau, 1990; Walz, 1979; 1980), much work still needs to be done for the systematic acoustic quantification of the corresponding English and French segments (/ɹ/ and /ʀ/ respectively), along with the many approximant productions spanning the two target segments (i.e., the apical /r/ discussed in Walz, 1980). Other phonetic variations not included in the table may also play a role in a deviant L2 production of French. Place of articulation of consonants may vary slightly between English and French; for instance, coronal stops are commonly described as alveolar in English, whereas in French they are described as dental (Dansereau, 1990). Further examples of variation include the contextually-specific voicing effects of fricatives, distinctive for each language (Valdman, 1976). Dansereau (1990) also identifies a general tendency in consonant pronunciation for French and English, noting that French consonants often require more articulatory effort and tension than their English counterparts.

While each one of the identified differences between the consonant inventories of English and French could benefit from further examination due to their potential influence on the perception of an Anglophone accent in French productions, the present study will focus on the production of syllable-initial voiceless stops (/p/, /t/, and /k/): the measurement of these consonants is included because of the relatively established role VOT is thought to play in the perception of an Anglophone accent in French, and the existence of some previous research available for data comparison (Caramazza et al., 1973; César-Lee, 1999; Flege, 1987b; Flege & Hillenbrand, 1984; Nearey & Rochet, 1994). These earlier studies are discussed below.

The importance of this distinguishing temporal cue, voice-onset time, is well-documented in cross-language studies (Caramazza et al., 1973), and the production and perception differences of the voiceless stops /p/, /t/, and /k/ in syllable-initial position – present in both languages – have frequently been studied. Specifically, Shriberg and Kent (2003) have defined VOT as the "time difference between the release of the stop closure and the beginning of the vocal fold vibrations" (p. 79) of the subsequent vowel. The acoustic quantification of VOT is relevant when comparing English and French syllable-initial stops due to differences in their degrees of aspiration (see Table 2.1). VOT durational measurements of voiceless stop consonants in English and French have been obtained for monolingual speech of both languages, as well as for L2 learners (bi-directionally) (Caramazza et al., 1973; Flege, 1987b; Nearey & Rochet, 1994).

Flege's (1987b) investigation into the effect of equivalence classification[5] for 'new' and 'similar' L2 phones includes VOT durational measurements of /t/ in the production of the English word *two* and the French *tout*, by 42 female speakers. One of the merits of Flege's study is that it is based on the evaluation of French learner

[5] Flege (1987b) defines the term *equivalence classification* as "A basic cognitive mechanism which permits humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category." (p. 49).

productions at various levels of proficiency. The speaker groups included monolingual English speakers, three proficiency levels of Anglophone learners of French (from less to most experienced), Francophone learners of English, and monolingual French speakers[6]. Flege hypothesized that these learners would not distinguish 'similar' L2 phones such as /t/ acoustically, thereby providing evidence for an effect of equivalence classification between L1 and L2 for Anglophone learners of French. Table 2.2 summarizes his durational findings for the VOT of /t/ across all six speaker groups.

Table 2.2 Flege's (1987b) Mean VOT Values for /t/ (in ms).

		French "tout"	*English "two"*
Group A	Monolingual English Speakers	-	77
Group B	Less Experienced English Learners of French	72	77
Group C	More Experienced English Learners of French	46	72
Group D	Most Experienced English Learners of French	43	56
Group E	Most Experienced French Learners of English	51	49
Group F	Monolingual French Speakers	33	-

These findings indicate that distinct VOT values for /t/ exist between monolingual English productions (77 ms) and monolingual French productions (33 ms). These results also demonstrate that L2 learner VOT values in L2 can pattern closely with L1 VOT values in less experienced learners, or take on intermediate VOT values between L1 and L2 for more experienced learners. For example, the less experienced English learners of

[6] The native English-speaking subjects were from Chicago, Illinois. The target French was considered a standardized variety, while the native French speakers were from Paris, France.

French produced comparable VOT values for /t/ in both English and French (77 ms and 72 ms respectively), while more experienced learners produced VOT values (around 44 ms) in between the monolingual VOT values of both languages. Furthermore, these findings highlight potential effects of L2 learning on the production of /t/ in L1 with more experienced L2 learning subjects (Groups D and E) not producing comparable monolingual VOT values in their L1s, suggesting evidence of bi-directional linguistic influence on speech production. Altogether, these findings indicate that observable differences exist in the VOT productions between non-native speakers of French of varying language proficiency levels.

Caramazza et al. (1973) provide VOT values for syllable-initial /p/, /t/, and /k/ produced by 40 subjects divided into the following groups: monolingual Canadian French speakers, monolingual Canadian English speakers, native Canadian French bilinguals speaking in French, and native Canadian French bilinguals speaking in English. The VOT measurements were taken from three tokens for each consonant in a stop + [a] context. Table 2.3 outlines the authors' findings for the production values.

Table 2.3 Caramazza et al.'s (1973) Mean VOT Values for /p/, /t/, and /k/ (in ms).

	/p/	*/t/*	*/k/*
Monolingual French Speakers	18	23	32
Bilingual French Speakers in French	20	28	35
Bilingual French Speakers in English	39	48	67
Monolingual English Speakers	62	70	90

These production values provide distinct VOT values for voiceless stops in both native English and native French productions. These VOT values for /t/ are comparable to Flege's (1987b) VOT values insofar as they display similar patterning between monolingual speakers. The authors found also that monolingual English speaker VOT values did not overlap between the three phonemic contrasts, whereas there was considerable VOT durational overlap for the monolingual French speakers, which suggests that English speakers may produce a clearer categorical VOT distinction than French speakers.

When examining the learner VOT distributions for /p/, /t/, and /k/, an interesting pattern emerged: the Canadian French bilingual speakers' VOT values patterned closely with the monolingual French speakers in their French productions, but shifted towards native English VOT values in their English productions. However, analysis of variance revealed significant differences between learner production VOT values in English and native English production VOT values. The authors therefore made a case for unidirectional interlanguage interference, suggesting that the phonological properties of the first language influence L2 productions, but not vice versa. While the Canadian French bilinguals' productions were not assessed for degree of foreign accent, these VOT results could indicate a possibility of the presence of a foreign accent in the L2 (English), because of the variation in acoustic values of the segments.

Nearey and Rochet (1994) provide a comprehensive data set of VOT production values for voiced and voiceless stops across ten vowel contexts for native speakers of English and native speakers of French; in other words, the authors controlled for vowel context for each elicitation of VOT production. Concerning the voiceless stops /p/, /t/, and /k/ produced by native English speakers, the authors found that the two vowels associated with the longest VOT values are /i/ and /u/ for all places of articulation. For native speakers of French, /u/, /y/, and /i/ are associated with the longest VOT values (see Appendix H for full list of values).

The authors found significant effects for place of articulation and vowel context in both languages. However, a significant two-way interaction is present only in the French VOT findings. The authors conclude "there appears to be greater variation in French VOT as a function of both vowel context and place of articulation." (p. 7): place of articulation effects in French range from about 3 to 35 ms, while vowel-related effects range from about 24 to 28 ms. These secondary influences on VOT values (the place of articulation of the stop and the following vowel) are important to bear in mind when considering VOT stimulus data. Furthermore, this greater variation found in French VOT values seem to corroborate Caramazza et al.'s (1973) data, according to which French VOT values display more variation than English VOT values.

César-Lee (1999) examined syllable-initial /p/ in her study on the correlation between acoustic production and the perception of a foreign accent in French with adult female learners. These syllable-initial /p/ values were found to correlate with the perception of a foreign accent in the productions of non-native American English speakers of French: speakers judged as having a 'native' accent produced mean VOT values of 8 ms, those judged as having a 'mild' foreign accent produced mean VOT values of 13 ms, speakers judged as having a 'moderate' foreign accent produced mean VOT values of 17 ms, and speakers judges as having a 'strong' foreign accent produced mean VOT values of 35 ms. The VOT values for syllable-initial /p/ rated as native-like are lower than those reported by Caramazza et al. (1973) and Nearey and Rochet (1994) for native speakers of French. However, César-Lee's VOT values are based on a single token of the target consonant in stressed position, extracted from a short passage that was read by the subjects. Furthermore, the author excluded the analysis of /t/ and /k/ in comparable phonetic contexts from the study, thereby excluding the possibility of cross-consonantal comparison of VOT values.

VOT values varying across speaker groups in the context of an experimental study could constitute a valid acoustic measurement of foreign-accented speech. However, the nature of individual speaker differences must also be taken into consideration. Allen, Miller and DeSteno (2003) suggest that even when speaking rate is controlled, significant differences in VOT production can still persist across subjects. They propose that native listeners may be able to use these speaker-specific VOT productions in order to identify the speaker, or indeed to recognize words produced by familiar speakers. Finally, they emphasize the need to control for speech rate in evaluating the production of VOT because of its potential influence on findings, especially when undertaking comparative analyses. These factors must be taken into consideration when discussing the results of the present study.

In summary, while many studies have examined the VOT productions of syllable-initial voiceless stops in both English and French – for native speakers as well as Anglophone learners of French, no study has looked at the VOT production of all three voiceless stops in the speech of female Anglophone learners of French for three language proficiency levels. Therefore, the need for enhancing the current data pool by including multiple stop tokens of /p/, /t/, and /k/ across comparable stress positions for the same subject sample (at varying degrees of language proficiency) remains.

2.4.1.2 Vowel Inventories

The standardized vowel systems of Canadian English and French have also been the subject of much research. Certain phonological and phonetic differences between the two languages have commonly been identified (Casagrande, 1984; César-Lee, 1999; Dansereau, 1990; Delattre, 1981; Garant Viau, 1994; Le Clézio, 1986; 1989; Mack, 1982; Valdman, 1993; Walz, 1980). Table 2.4 below summarizes key findings and observations on the nature of the two vowel systems, as they relate to Anglophone learners of French.

Table 2.4 Highlights of Comparison for English and French Vowels.

Segments	*French*	*English*
Number of Vowels	16 vowels (incl. nasals)	14 vowels (incl. diphthongs)
/ɥ/[7] and /y/	Present	Absent
Nasal vowels	/ɑ̃/, /ɛ̃/, /ɔ̃/, (/œ̃/)	Absent
Diphthongs	Absent	Present

Just as articulatory characteristics were identified for the consonant inventories of French and English, other general tendencies in vowel production distinguish the two languages as well. For example, lip positioning, vowel height, and vowel tension differ between French and English. César-Lee (1999) and Dansereau (1990) both associate the following properties to the production of French vowels: higher articulatory effort, relatively higher 'high' vowels, and a clear tendency toward fronting and rounding. The subsequent resonance frequency measurements for nine female French speakers presented in Table 2.5, as measured in Dowd, Smith and Wolfe (1997), provide a point of comparison for the acoustic values that will be obtained in the present study.

[7] While /ɥ/ is included in the vowel inventory, it is actually a semi-vowel, or glide.

Table 2.5 R1 and R2 Values for Native French-speaking Women (in Hz) (SD) (Dowd, Smith & Wolfe, 1997).

Vowel	*R1*		*R2*	
/y/	350	(70)	1960	(160)
/u/	390	(60)	869	(110)
/o/	520	(40)	920	(100)
/ɔ/	690	(100)	1090	(110)
/a/	870	(90)	1420	(180)

Martin's (1998) results for female French speakers provide somewhat lower formant values for /o/ (380 Hz [F1], 803 Hz [F2]) and /ɔ/ (592 Hz [F1], 1420 Hz [F2]). In addition, Martin (2004) provides a complete set of average female formant values for Canadian French for all Standard French Vowels (see Appendix I) as an Internet reference, which serves as an important point of comparison for the data collected in this study.

The contrast /u/ – /y/ has been examined in language acquisition studies of Anglophone learners of French (César-Lee, 1999; Flege, 1987b; Flege & Hillenbrand, 1984), because such a phonological contrast does not exist in the English vowel inventory. While conventionally /y/ is thought to be mispronounced more systematically than /u/ by learners of French (especially by beginners), previous acoustic analysis has demonstrated that /y/ can in fact be pronounced in a more native-like manner by more experienced learners of French than /u/ (Flege, 1987b; Flege, 1995; Flege & Hillenbrand, 1984; Schweyer, 1996). For example, Flege (1987) found that native English learners of French did not produce F2 values for /u/ equalling those of monolingual subjects, whereas their F2 values for /y/ were more comparable to those of monolingual subjects. These results suggest that Anglophone learners of French actually produce a more native-like /y/ than /u/ in terms of F2 frequency. The author proceeds to plot the mean F2 frequency findings for /u/ and /y/, revealing "important differences between groups in the magnitude of the F2 difference between French /u/ and /y/" (p. 58). Whereas native French/English bilinguals and French monolinguals demonstrated marked F2 differences between /u/ and /y/ (785 Hz and 866 Hz, respectively), less experienced and more experienced English learners of French produced smaller /u/ – /y/ differences (61 Hz and 379 Hz, respectively). Flege uses these findings to support his claim for the equivalence classification of similar sounds in an L2, such as /u/, as is found in both French and English.

César-Lee (1999) also examines the /u/ – /y/ production contrast in the speech of American learners of French. In her investigation of the pronunciation of /u/ and /y/ by Anglophone learners of French, the author posits that the ratio between the formants (F1 and F2) is most important in measuring the vowel quality of the two segments because their F1s are relatively similar, while their F2 values differ greatly. The author thereby supports an F2 – F1 acoustic space representation for findings relating to the /u/ – /y/ contrast. The author's findings seem to parallel those found by Flege (1987b), in that learner values for /y/ seem to pattern closer with native /y/ productions than learner values for /u/ pattern with native /u/ productions.

While the distinction /o/ – /ɔ/ has not been studied acoustically in L2 French studies of foreign accent the opposition will be examined here, since the failure to correctly produce this contrast has anecdotally been attributed to the presence of a foreign accent in French speech production (Dansereau, 1990). Moreover, in his study on possible factors contributing to an English accent in French, Le Clézio (1986) found that French learners produced /o/ less satisfactorily than /ɔ/. In fact, in their non-target-like production of /o/, learners either used a vowel closer to /ɔ/, or they used a diphthong of the [əʊ] type. The present study seeks to quantify production differences acoustically across learners and native speakers for this target opposition.

Martin (1998) provides an exhaustive acoustic study of the French mid-vowel system by university-aged male and female Québécois speakers, which can serve as a reference point for the present investigation into this contrast. His speech data is based on single word tokens produced by four female and four male speakers. His findings indicate that the contrast /o/ vs. /ɔ/ remains the most stable of the French mid-vowel system in the Québécois population. Expressed in terms of an F2 – F1 acoustic space, the author found values of 423 Hz for /o/, and 828 Hz for /ɔ/ produced by female speakers. The difference between these values was found to be significant using *t*-tests.

Finally, the present research also includes acoustic measurement of the segment /a/ in non-native and native French productions in unstressed environments, because spectral centralization (or F1 lowering) of /a/ in unstressed environments has been associated with the perception of a foreign accent in French in the productions of Anglophone learners of French (César-Lee, 1999). Delattre (1981) reports F1 values of 750 Hz for /a/ in stressed position, and 650 Hz for /a/ in unstressed position in the speech of native male speakers of French. César-Lee (1999) found that spectral centralization (corresponding to F1 lowering) of /a/ in an unstressed environment is associated with the perception of a foreign accent in French: as F1 frequencies for /a/ decrease, the degree of perceived foreign accent increases. Female speakers rated as native-like produced mean F1 values of /a/ at 681 Hz in an unstressed environment, whereas speakers judged as having a strong foreign accent produced mean F1 values of 530 Hz. Speakers judged as having a mild or moderate foreign accent clustered closely with native speakers, with mean F1 values for /a/ at 717 Hz and 692 Hz, respectively. While this vocalic reduction was examined from a spectral centralization perspective, there may also be a corresponding temporal reduction phenomenon, which could also influence the perception of a foreign accent. Such potential temporal reduction will be discussed in section 2.4.2, which deals with the suprasegmental characteristics of foreign-accented speech.

In summary, while many differences exist between the vowel inventories of English and French, the following segments will be examined due to their potential influence on the perception of an Anglophone accent in French: the F2 – F1 acoustic space of /u/ and /y/; the F2 – F1 acoustic space of /o/ and /ɔ/; and the spectral centralization (F1) of /a/. No previous study has provided formant values for all of these vowels across multiple stimulus tokens for the same speaker groups. Therefore, the acoustic analysis of these vowels collectively will provide a more comprehensive consideration of French learner speech at multiple proficiency-levels.

2.4.1.2.1 Diphthongization

Conventionally, Standard French is thought of as having 'pure', or monophthongal vowels, with relatively stable steady-states (César-Lee, 1999; Casagrande, 1984; Dansereau, 1990; Dowd, Smith & Wolfe, 1997; Le Clézio, 1986; 1989; Walz, 1979; 1980), whereas the vowel inventory of English includes several diphthongs (César-Lee, 1999; O'Grady & Archibald, 2000). While some varieties of French, such as Canadian French, do display certain dipthongized vowels (Martin, 2002; Santerre, Dufour & McDuff, 1985; Walker, 1984), these regionalized features tend to be muted in the foreign language classroom setting (Valdman, 1993)[8]. Furthermore, the diphthongized phenomena of Canadian French follow a certain structure, whereas diphthongization in the non-native student's productions of French does not adhere to the same set of constraints. For example, according to Walz (1980), Anglophone learners of French at the beginner level consistently substitute diphthongs for word final [e] and [o] in a variety of target words. Le Clézio (1989) also discusses diphthongal productions of /e/ and /o/ by learners of French, as well as inappropriate diphthongization of /i/. Therefore, although it has not been acoustically quantified, inappropriate diphthongization in French by Anglophone speakers has been associated with the perception of a foreign accent (Casagrande, 1984; Dansereau, 1990; Le Clézio, 1986; Tranel, 1987; Walz, 1980).

Santerre, Dufour and McDuff (1985) discuss the perception of dipthongization in Canadian French varieties, and the sociolinguistic impacts thereof. The authors claim that diphthong perception is not exclusively a phenomenon of perceiving spectral variation, but rather a complex interaction between acoustic, psychological and sociolinguistic factors in decoding a speech sound in context. The authors propose that vowel duration, intonation, and intensity variation must also be taken into consideration in addition to a simple formant measurement of vowel quality because their study of the correlation between spectral variations in naturally produced vowels and the degree of perceived diphthongisation was inconclusive. In order to control for durational, intonational, and intensity variables, Santerre et al. develop synthetic stimuli varying along spectral, durational and intensity dimensions. The authors were then able to conclude that

[8] Valdman (1993) discusses the pedagogical language norm, which allows for multiple 'correct' variations, but advances a mostly unmarked pronunciation by language instructors.

diphthong perception is in fact linked to spectral variation, and that longer vowel duration promotes diphthong perception somewhat.

Le Clézio (1989) discusses the potential impact of inappropriate diphthongization by Anglophone learners of French. He suggests that native perception of inappropriate dipthongization depends on the width of the diphthong and the direction of variation. For example, an /i/ production which begins at the "margin of the area of dispersal of French /i/ and ends well inside it" (p. 68) could escape detection, while /e/, /o/, or /i/ tokens produced as [eɪ], [əu] and [ɪə], respectively, are considered foreign-accented. This is due to the fact that the latter set of phonetic realizations move away from the centers of gravity of French /e/, /o/, and /i/. Unfortunately, Le Clézio does not quantify the level of acoustic variation needed for diphthongized Anglophone pronunciations to be considered foreign-accented. Due to the current paucity of acoustic measurements of diphthongizational phenomena in foreign-accented speech, it is necessary to quantify such potential factors, which may affect the degree of perceived foreign accent in Anglophone productions of French.

2.4.2 Suprasegmental Characteristics

In the scope of the present study, the terms *suprasegmental* and *prosody* are used interchangeably, and refer to speech properties such as stress, rhythm, speaking rate and intonation[9]. The differences between the two prosodic systems of French and English have been the subject of discussion in much previous research (César-Lee, 1999; Guilbault, 2002; Vaissière, 1991; Wenk & Wioland, 1982). The prosodic systems of the two languages display variation in intonation, stress, rhythm and rate, but only the two

[9] See the entry for *prosody* in Crystal (2003, p. 378) for a more nuanced discussion of the two terms.

aspects relevant to the present project will be presented: rhythm and speech rate. While all parts of the prosodic system could benefit from further investigation, when examining spontaneous speech samples – which is one of the aims of the current project – the study of rhythm and speech rate are relatively well established, and because of the existence of previous research results, other data is available for comparison (César-Lee, 1999; Deterding, 2001; Guilbault, 2002).

Traditionally, English and French have been grouped into two separate language rhythmic classes: English has been classified as a stress-timed language, while French is considered a syllable-timed language (Dauer, 1983; Vaissière, 1991, among others). Pike (1945) originally proposed the two rhythmic classifications, where stress-timed languages seem to demonstrate relatively equal timing between stressed syllables, and syllable-timed languages demonstrate relatively equal timing across all syllables. Dauer (1983) suggests that stress-timed and syllable-timed languages display distinctive phonological and phonetic properties. Among the most important properties are the following 1) stress-timed languages display a different syllable structure to that of syllable-timed languages, and 2) they also exhibit vowel reduction phenomena absent in syllable-timed languages. For example, syllables in stress-timed languages display more complex structures (e.g., CCV and CCCV) than the syllables of syllable-timed languages.

Miller (1984) develops this rhythmic classification model further by proposing a continuum along which languages are categorized, tending either toward a more stress- or syllable-timed constraint. Thus, instead of an absolute categorization, languages occupy positions relative to one another (e.g., English is more stress-timed than French). Ramus, Nespor and Mehler (1999) confirm that the categories of this relatively established rhythm classification system for various languages are in fact meaningful, as opposed to just serving as theoretical constructs. In their study of the perception of rhythm in eight languages (including English and French), the authors found that segmenting the speech signal into consonants and vowels accounted for the perception of rhythmic classes. Thus, the rhythmic categorization of languages seems to reflect actual properties of the speech signal of different languages, as well as perceptual rhythmic boundaries. While the authors support the rhythmic categorization of languages into more syllable- or stress-timed classifications[10], they argue that it is conceivable that other categories could exist as well. Until the rhythmic structures of more languages are examined, it is impossible to conclude that the existing categories can account for every language rhythm.

Based on the similarities and differences between the English and French prosodic systems, such as varied stress assignment and rhythmic properties of the two languages, César-Lee (1999) hypothesizes that English-speaking learners of French could produce non-target rhythm- and stress-patterns which would contribute to the perception of a foreign accent. Specifically, she found that the suprasegmental cues correlated with perceived levels of foreign accentedness include the total duration of utterance, VOT[11] of syllable initial [p], and syllable duration. The syllable duration feature was measured by comparing the syllable length of a target syllable and its adjacent syllables (one or two preceding syllables, and one after the target syllable). However, this method of establishing inter-syllabic variability does not account for differences in speech rate across subjects, or the larger rhythmic environment (i.e., a whole phrase vs. a single

[10] The authors also discuss Japanese as being the only known mora-timed language.
[11] VOT has alternately been considered a segmental cue, or a suprasegmental cue, as its variation has temporal ramifications. For the purposes of this study, it has been included in the segmental portion of the analysis.

word). Furthermore, this method does not take French final-syllable lengthening into consideration, which could interfere with the durational measurements of inter-syllabic variation. A more robust measurement of rhythm has been proposed in the form of a Variability Index (Deterding, 2001; Low, 1998), which includes normalization procedures in order to control for speech rate and makes use of spontaneous speech, which is more representative of actual speaker productions in real-life situations.

Previous studies on rhythmic variation across languages have made use of the Variability Index (VI) to facilitate valid quantification of such variation (Deterding, 2001; Guilbault, 2002). The VI, as proposed by Deterding (2001), is intended for use on conversational or spontaneous speech, rather than read data. It also

> measures the whole syllable rather than just the vowel, excludes consideration of the final syllable in the analysis, and uses normalization based on the whole utterance (excluding the final syllable) rather than a localized "pairwise" normalization. (p. 219)

The VI also provides an average syllable length value, calculated after pauses have been removed, which can be used as a speech rate measurement. Deterding (2001) justifies the use of a Variability Index over other methods of measuring rhythmic variability in the following

> It would be possible to calculate the difference of each syllable duration from the average syllable duration of the utterance, and then use the average of all these differences as a measure of variability. However, if a speaker produces a number of syllables quickly followed by a number of syllables slowly, the deviation of each syllable from the average would be high, even if the speaker had been using syllable-timed rhythm with a change of pace in the middle. (p. 223)

Guilbault (2002) applied the Variability Index to the spontaneous productions of Canadian English learners of French. Specifically, he compared the utterances of less experienced learners of French (EL1), and more experienced learners of French (EL2), with the utterances of native Canadian French (CF) and native European French (EF) speakers, and hypothesizes that English L2 learners of French will exhibit more inter-syllabic variability than the two groups of French native speakers. This hypothesis is based on the assumption that English displays higher inter-syllabic variability than French. Table 2.6 below provides the author's results.

Table 2.6 Guilbault's (2002) Variability Index Values (p. 119).

Speaker Group	*Variability Index*	*Standard Deviation*
EL1	0.4056	0.2486
EL2	0.4298	0.2399
CF	0.3508	0.2042
EF	0.2866	0.1547

Higher values, such as 0.4056 and 0.4298, obtained for spontaneous-speech samples of less experienced French learners (EL1) and the experienced French learners (EL2), demonstrate greater inter-syllabic variation. Lower values, such as 0.2866 obtained for spontaneous-speech samples of native European French speakers, denote lower inter-syllabic variation because they approach the 0 score, which indicates greater rhythmic uniformity (no inter-syllabic variation). While Guilbault's aim was not to quantify a global degree of perceived foreign accent, his findings do indicate a statistically significant difference between the inter-syllabic variability scores of less experienced French learners (EL1) and of native European French speakers (EF). No significant effects for level of L2 proficiency, for age of L2 learning, or for time spent in a French-dominant environment were found. Because Guilbault did not perform any segmental analysis on the same speaker groups, comparison across segmental and suprasegmental phenomena for the same speaker groups is impossible. However, his study provides an important foundation for the understanding of the rhythmic patterns produced by Canadian English learners of French.

In order to further examine suprasegmental cues of Canadian English-accented French, the present study considers both the temporal aspects of vowel reduction in unstressed environments for /a/, /u/, and /y/; as well as more general inter-syllabic temporal variation, as quantified by the VI. First, an average durational percentage of the unstressed vowel will be calculated. While it is evident that vowel reduction is a feature of French prosody, it is argued that non-native speakers would demonstrate greater vowel reduction than native speakers of French. Second, the use of a VI is proposed in order to quantify the rhythmic variability of L2 learner speech, instead of using a simple "pairwise" syllable comparison. The VI will not only give a more robust measurement of inter-syllabic variability, but will also provide speech rate values, useful for inter-speaker comparison.

2.5 Current Study

The literature review has revealed that previous research on foreign accent has generated numerous studies dealing with the acoustical and perceptual features of accented speech, several dealing particularly with the nature of English-accented French. These studies have provided some acoustic measurements of particular segments in French, as produced by Anglophone speakers, as well as some preliminary examinations of the rhythm of English-speaking learners of French. However, much specific quantification of English-accented French speech requires further investigation, such as speaker performance on multiple parameters of speech (segmental and suprasegmental) with corresponding acoustic production values, as well as spontaneous and non-spontaneous speech samples for the same speaker set.

Many models and approaches have been used in the characterization of foreign accented speech, such as contrastive analysis (CA), acoustic analysis, and perceptual analysis. The contrastive analysis method of predicting pronunciation difficulties based on L1 interference has been refuted, as negative transfer from L1 to L2 provides insufficient explanation of accented non-native speech (Garnica & Herbert, 1979; Jonasson & McAllister, 1972; Wardhaugh, 1970). Instead, the notion of cross-linguistic influence has replaced CA, where there is a two-way interaction between the influences of L1 and L2, in order to encompass a variety of pronunciation constraints wider than merely the L1's influence on L2. Moreover, while the use of acoustic and perceptual analysis to define the parameters of accented speech has been relatively successful in quantifying foreign accented speech, often only a very restricted number of phonological or phonetic features of such speech have been examined using the same sample population. Finally, the need for more acoustic quantification of female speech remains at issue, as most studies establishing pronunciation norms and tendencies have based their findings on male speech.

The research justification for the present study lies in the current lack of foreign-accent data linking the segmental speech analysis of female Anglophone learners of French with a suprasegmental analysis of their spontaneous productions. One of the objectives of the present study is therefore to identify and quantify some of the more relevant segmental and suprasegmental acoustic cues of Canadian-accented French in female speech. Based on previously discussed a) cross-linguistic interference between

English and French, b) established research findings in the field, c) research inconsistencies identified in other studies, and d) the desire to contribute to the larger body of research dealing with the nature of foreign accent, the present experiment aims to answer the following general research questions:

1) How does non-native speech compare acoustically to native speech in French?
2) Is acoustic variation in segmental production readily identifiable between native and non-native speech samples?
3) Is variation in suprasegmental production readily identifiable between native and non-native speech samples?

In order to address these research questions, three hypotheses have been postulated. The predictions of these hypotheses stem from previous phonological and acoustic research where researchers found that those L2 learner speech productions displaying greater acoustic variability, and/or differing considerably from native speaker productions, are perceived as more accented than L2 learner speech productions patterning more closely to native speaker productions (César-Lee, 1999; Jonasson & McAllister, 1972; Magen, 1998; Shah, 2002; Walz, 1980). More specifically, the present experiment tests the following general hypotheses:

H1 The three experimental groups (female Anglophone learners of French) would demonstrate noticeably different acoustic values in their speech productions when compared to the control group (female native speakers of French).

H2 The three experimental groups would demonstrate greater acoustic variability (obtained in the measurement of acoustic cues) than the control group.

H3 As speaker language proficiency increased, acoustic values would become increasingly target-like and display less acoustic variability.

Recognizing the acoustic approach as an integral first step in the understanding of Canadian English-accented French, the study proposes to contribute to the body of research intending to quantify Canadian English-accented French. In order to address the above research questions, the study seeks to quantify many aspects of Canadian English-accented French by providing a detailed and consistent acoustical data set of the following parameters, which were identified in the literature review (see section 2.4) as potential factors influencing the degree of foreign accent in Canadian English-accented French productions: 1) the VOT of syllable-initial /p/, /t/, and /k/; 2) the F2-F1 acoustic space for /u/ and /y/; 3) the spectral values over time for /u/, /o/, and /i/ to identify potential diphthongization; 4) the F2-F1 acoustic space for /ɔ/ and /o/ to determine target-like vowel provision; 5) the spectral F1 values for /a/ in unstressed environments; 6) the temporal values for vowel reduction in unstressed environments for /u/, /y/ and /a/; and 7) the rhythmic variation values as □nalysed by the Variability Index. The analysis of these parameters will provide not only a more comprehensive data set for the speech of Canadian English learners of French in the foreign language classroom setting, but it would also establish a link between the segmental and suprasegmental production values of Canadian English learners of French across the same speaker groups.

3 METHODOLOGY

The primary objective of this study is to obtain acoustic measures of speech in an attempt to quantify the selected acoustic cues of Canadian English-accented French. With this objective in mind, the study includes two experimental tasks eliciting production material from non-native learners and native speakers of French. The first task is a delayed-repetition sentence-production task, where carrier sentences contain the target segments under analysis, while the second task includes the elicitation of spontaneous speech for the analysis of suprasegmental features. This section of the study will present participating subjects, task stimuli, experimental protocol and procedures, and subsequent data analysis.

3.1 Subjects

Non-native speakers. A total of fifteen non-native speakers, all female, were recruited for participation in this production experiment. The non-native speakers comprising the three experimental groups were students recruited from the Department of French at Simon Fraser University[12]. Each speaker was enrolled in a French class at the time of data collection. A Language Background Questionnaire (see Appendix B) determined the suitability of each potential participant by providing demographic information and information on linguistic experience. All subjects reported having English as their first language. None reported having fluency in any language other than English or French (Appendix D). The age of these subjects ranged from 19 to 27, with a mean age of 22.

[12] One subject, P11, was a student recruited from a beginner French class at UBC.

The speakers were grouped into proficiency levels according to their French instructional levels:

1) *Beginner* 100 level French language classes
2) *Intermediate* 200 level French language classes
3) *Advanced* 300 level French language classes

For each proficiency level, five speakers were recruited, from which two were retained, resulting in six experimental subjects in total. The sample size was determined by the consideration of speech sample quality, a speaker's ability to complete the tasks appropriately, and consistency in linguistic experience across experimental groups. Two criteria governed the final selection of experimental subjects: no prolonged exposure to French in francophone regions or French Immersion experience, and limited exposure to additional languages.

All of the beginner and advanced subjects, as well as one of the intermediate participants satisfy all of these requirements (P11, P13, P22, P32, and P37). One intermediate subject, P24, had French Immersion and limited experience with Portuguese. However, this subject was deemed most suitable from the available intermediate subjects. These possible confounding factors on pronunciation ability will be discussed in the results section.

Native speakers. Two female native speakers were recorded for the present study in order to provide a control group for the experiment. Both are French language teachers at Simon Fraser University; subject C11 comes from France, while subject C12 is from Québec. Both are older than the experimental subjects, with subject C11 in her 50s, and subject C12 in her 40s. They completed a Language Background Questionnaire in

French, in order to determine their eligibility for participation in the study (Appendix C). Both grew up in francophone regions with limited exposure to other languages until adulthood. They report being French-dominant in their language use, with the majority of their professional life, cultural activities, and communication being conducted in French. Appendix D provides a summary of the linguistic differences between all subjects.

3.2 Stimuli

3.2.1 Task 1: Sentence-Production Task

The first task consisted of a delayed-repetition sentence-production task, with carrier sentences containing the target segments under analysis. Target words were embedded in sixty-three carrier sentences, and randomized when presented to the subjects (five different versions of stimulus presentation). The target words were chosen according to the following criteria:

1) Inclusion of the target cue in the appropriate stress environment (see Table 3.1 below: Appendix A lists the final inventory of sentences containing the target segments),
2) Learner familiarity[13] (most words were taken from beginner textbooks familiar to the students).

[13] Flege, Takagi, and Mann (1996) note the potential effects of subjective lexical familiarity on experimental tasks. Therefore, in the present study, it was important that all words included in the stimuli set be familiar to the learners.

Table 3.1 Task 1: Acoustic Measurements of Target Cues for Analysis.

1. VOT
VOT duration of /p/ in stressed position
VOT duration of /t/ in stressed position
VOT duration of /k/ in stressed position
2. F2-F1 acoustic space for /u/ and /y/
F1 and F2 of /u/ in stressed position (at the ½ point of vowel steady-state)
F1 and F2 of /y/ in stressed position (at the ½ point of vowel steady-state)
3. Diphthongization of /u/, /o/, and /i/
F1 and F2 of /u/ at the ¼ point of vowel steady-state
F1 and F2 of /u/ at the ¾ point of vowel steady-state
F1 and F2 of /o/ at the ¼ point of vowel steady-state
F1 and F2 of /o/ at the ¾ point of vowel steady-state
F1 and F2 of /i/ at the ¼ point of vowel steady-state
F1 and F2 of /i/ at the ¾ point of vowel steady-state
4. F2-F1 acoustic space for /ɔ/ and /o/
F1 and F2 of /o/ in stressed position (at the ½ point of vowel steady-state)
F1 and F2 of /ɔ/ in stressed position (at the ½ point of vowel steady-state)
5. Spectral centralization in unstressed environments
F1 of /a/ in unstressed position (at the ½ point of vowel steady-state)
6. Temporal reduction in unstressed environments
Duration of /u/ in stressed and unstressed positions
Duration of /y/ in stressed and unstressed positions
Duration of /a/ in stressed and unstressed positions

After the production task, a total of sixty-seven cues were analyzed for each subject, typically five tokens of each cue, except in the case of /y/ in unstressed position and /a/ in stressed position, where only four tokens of each cue were □nalysed (the omission of the fifth token is explained in section 3.4.2). Acoustic measurement of these tokens included the acoustic segmentation of each phone from adjacent phones, in order to obtain vowel steady-state duration or VOT duration, as well as spectral measurement determining formant values for the vocalic cues.

3.2.2 Task 2: Elicitation of Spontaneous Speech

After the sentence-production task, speakers were asked to converse with the researcher in French for approximately eight minutes. This spontaneous-speech production task was included for the analysis of suprasegmental features, such as rhythmicity. The researcher asked straightforward, open-ended questions that were relevant to the speaker, in order to encourage the production of spontaneous utterances in French of at least eight syllables duration (see Appendices E and F for question lists), to be used in the Variability Index calculations. For example, the researcher asked subjects to talk about a favourite movie, their reasons for studying French, or common weekend activities.

3.3 Procedure

All speakers who replied to the research subject advertisement poster were invited to a recording session as long as they met subject criteria, as determined through a mini-language background questionnaire via e-mail. The subjects were each paid a small honorarium upon completion of all experimental tasks. The recording sessions took place in a quiet, private room, without any other equipment operating (computers, fans, air conditioning, etc.). The speech samples of both experimental tasks were recorded on a BurnIT CDR-830 Compact Disc Recorder unit, through a Focusrite trak master pre-amplifier, using a condenser AKG Acoustics microphone (MicroMic Series III). The researcher monitored noise and recording levels via headphones throughout the session.

Subjects began the recording session by completing the full language-background questionnaire and a consent form[14](Appendices B and C). The subjects were told that

[14] The experimental protocol was reviewed and approved by Simon Fraser University's Research Ethics Board, Ref. #36286. A copy of the approval is filed with the university library.

they were free to leave at any point if they felt so inclined. The researcher then briefly explained the procedure (in English for non-native speakers, in French for native-speakers) of the two tasks they would be asked to complete.

The stimuli for the sentence-production task were displayed on a laptop computer via a PowerPoint presentation, which included written instructions corresponding to the researcher's verbal instructions. The first slide thanked subjects and familiarized them with the navigation of the presentation, while the second slide provided detailed experimental instructions (Table 3.2).

Table 3.2 Experimental Instructions.

1	The researcher will read out the slide number.
2	You will read the sentence once out loud.
3	Move to the next blank slide (by pressing the arrow key or clicking the mouse).
4	As naturally as possible, say the sentence twice, with a momentary pause between each repetition.
5	If you make a mistake, or cough, etc., just say the sentence again.
6	Move on to the next sentence and repeat the procedure.

The subjects then had the opportunity to practice on three trial sentences before reviewing the instructions again and beginning the production task. Some subjects found it hard to remember to advance the slide after the initial repetition, so the researcher advanced the slides for them. The subjects were reminded to speak naturally when uttering the sentences, as if conversing with someone, in order to ensure a more genuine speech sample. This task took them between 30 and 45 minutes each to complete.

Prior to recording the spontaneous component of the session, the researcher conversed with subjects in French, reiterating the parameters of the task. Subjects were asked to respond naturally and expand on their ideas as much as possible. The ensuing interview was recorded, including some preliminary conversation (greetings, activities of the day, time of year, etc.) in order to assuage any anxiety over the task. Because of the limited language-proficiency of some of the subjects, certain subjects were not able to provide enough spontaneous sentences of appropriate length. Therefore, the researcher employed some compensatory strategies. First, she asked a handful of more situationally-specific questions within the conversation, and asked for complete sentence answers. Furthermore, she gave the option of studying a simple dialogue at the end of the task, to inspire the speakers with longer answers (Appendix G). The subjects then improvised their own answers in a conversational situation, based on the dialogue. Overall, speakers reported feeling comfortable during both tasks.

3.4 Analysis

For each of the eight speakers all of the materials were recorded, and then converted to a *.wav* format at a 44 kHz sampling rate using Adobe Audition software. The files containing target information were then filtered, downsampled to 11 kHz (as required by CSL analysis parameters), and saved as an *.nsp* file in Kay Elemetrics' MultiSpeech (v.2.7.0, Model 3700), using the hardware provided with the DOS version (Model 4300B). All sound files were identified by a code to preserve subject anonymity.

3.4.1 Stop Consonants

For the sentence-production data, the VOT duration of the voiceless stop consonants /p/, /t/, and /k/ was measured. Using an amplitude waveform linked to a spectrographic representation, VOT was measured from the onset of the initial burst of aspiration after stop closure (corresponding to an aperiodic oscillation on the waveform or vertical striation on the spectrogram) to the onset of periodicity/vowel onset (first positive value after the zero point on the upslope of the sound wave) (César-Lee, 1999). In the case of a discrepancy between the two representations (waveform and spectrogram), the waveform measurement was retained, as waveform-based measurement has been considered a more accurate form of measurement (Francis, Ciocca & Yu, 2003).

3.4.2 Vowels

For the sentence-production data, the vocalic cues were tagged for vowel onset, steady-state onset, steady-state offset, and vowel offset. The vowel onset and offset criteria correspond to the onset and offset of periodicity in the sound wave, as viewed on a waveform representation (Mack, 1982). The steady-state onset and offset criteria correspond to the relative consistency of the shape, the presence of upper formants, and periodicity of the sound wave (Flege, 1984; Shah, 2002). All tags were placed at the first positive value after the zero point on the upslope of the sound wave. Figure 3.1 below illustrates an example of the tagging of the vowel /ɔ/ in the word 'bol' for native speaker C11.

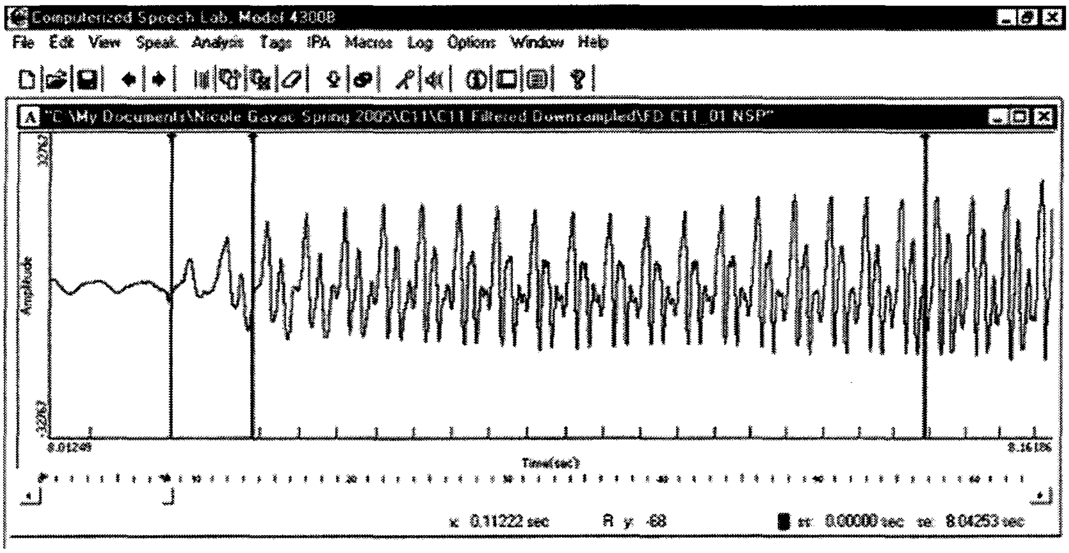

Figure 3.1 *Print Screen Snapshot of the Waveform 'bol' for Speaker C11,* the bar at the first tag corresponds to the vowel onset of [ɔ]; the bar at the second tag corresponds to the steady-state onset; and the third bar corresponds to the steady-state offset.

Formant frequency values for vocalic cues were obtained using spectrograms and formant history overlays, with the use of linear predictive coding (LPC) analyses for verification. Vowel formant measurement of F1, F2, and F3 took place at ¼, ½, and ¾ of the duration of the steady-state. These points of the vowel steady-state are coded to as T1, T2, and T3, respectively. F1 and F2 formant measurements at the mid-vowel point (T2) were used for subsequent analysis and comparison for the vowels. This methodological consideration was made in order to minimize consonantal co-articulatory effects on formants, as neighboring consonants – or more generally, phonetic context – influence vowel quality (Shriberg & Kent, 2003). Furthermore, vowel mid-point measurement could allow for possible comparison with formant values obtained in previous research (Martin, 2002). Vowel analysis for the pairs /u/ – /y/ and /o/ – /ɔ/ includes the calculation of an F2 – F1 acoustic space, which is considered a valid method of comparison for these minimal pairs by César-Lee (1999; for the segments /u/ – /y/), and Martin (1998; for the segments /o/ and /ɔ/).

F1 and F2 formant measurements of /u/, /o/, and /i/ at ¼ and ¾ of the vowel steady-state (T1 and T3, respectively) were used for analysis of diphthong-related properties (Martin, 2002). The presence of diphthongization was determined by the occurrence of a second formant sweep, or F2-shifting, between 200 and 400 Hz, as put forward by Schouten and Peeters (2000).

Occasional segmentation difficulties occurred in defining vowel onset and offset, depending on the phonetic environment surrounding the vowel (adjacent liquids, glides or nasals) or extreme vowel reduction (e.g., the word /tɔ'mat/ being realized as /t'mat/). Such segments were omitted from the final analysis.

Finally, durational measurements of /u/, /y/, and /a/ were compared between stressed and unstressed environments. Average durational values of each segment were calculated for each environment, providing an *average durational percentage*. This percentage is the unstressed vowel duration expressed as a percentage of stressed vowel duration: the higher the percentage, the less temporal vowel reduction taking place, the lower the percentage, the more vowel reduction taking place.

3.4.3 Spontaneous Speech

The spontaneous-speech samples were obtained in order to perform Variability Index (VI) calculations (Deterding, 2001; Guilbault, 2002). First, the eight-minute speech samples were transcribed in order to identify utterances longer than eight syllables, appropriate for use in the calculation of a VI analysis. After determining the representative nature of the utterance (natural rhythm relative to speaker, no use of English words, relative lack of hesitations, redirections, and non-speech sounds), five sentences were segmented and tagged for syllable duration for each of the eight speakers. All remaining hesitations, repetitions and pauses (defined as an audible silence between segments) were removed from the sentence.

The calculation of a VI requires the segmentation of speech samples into syllables, in order to compute the normalized durational value of consecutive syllables, excluding the final syllable of the rhythmic group. The final syllable of the rhythmic group is excluded because final-syllable lengthening, which is common in French, could interfere with the measurement of syllable-timing (Deterding, 2001). In the present study, the syllabification rules for French were used to guide the syllable segmentation procedure. Guilbault (2002, p. 82 – 83) presents a summary of the French syllabification rules proposed by Delattre (1940) and experimentally confirmed by Beaudoin (1996):

- A single intervocalic consonant is syllabified as the onset of the following syllable, leaving the preceding syllable without a coda (open syllable).
- Preference is given to splitting consonant clusters. However, the retention of clustering patterns according to the following rules:
    - The sonority of the first consonant is clearly lower than that of the second consonant.
    - The pronunciation of the first consonant is more fronted than the second.

Deterding also identifies some potential syllable segmentation issues, as a consequence of the merging of words in spontaneous-speech samples. An illustration of such a vowel or segment deletion/merging phenomenon is relatively common in colloquial French, especially in the case of the schwa: e.g., *je suis* becomes *j'suis* (Walker, 2001). In this study, the duration of the remaining consonant (in this case 'j') was included in the calculation of the following syllable (e.g., *je suis* pronounced as *j'suis* → *[ʃɥI]* was considered one syllable rather than two).

Subsequent analysis of the VI included calculating syllable duration, normalized syllable duration, and the durational difference between syllables. The following algorithmic formula is used in the calculation of the VI (Deterding, 2001; Guilbault, 2002), where the normalized duration of the *k*th syllable is d_k, and the number of syllables in the utterance is n.

$$VarIndex = \left(\sum_{k=1}^{n-2} \left| \mathrm{d}_{k+1} - \mathrm{d}_{k} \right| \right) / (n-1)$$

Table 3.3 below provides an illustration of VI calculations done in this study; the sentence was produced by a native speaker of French (C11). When calculating the VI, syllables found at the end of syntactically disjunctive (left-dislocated) adverbial phrases were removed, because they are prosodically-marked (lengthened). Inclusion of such syllables would skew the VI. Because of the relative durational consistency across syllables in the utterance presented in Table 3.3, the VI value is very low, at 0.1415. Lower VI scores are typical of syllable-timed languages, while higher VI scores are more typical of stress-timed languages.

Table 3.3 *Calculation of a VI of 0.1415 for the Native Speaker Utterance "Cet été je suis partie pendant deux mois en".*

Syllable number		*Syllable duration (s)*	*Normalized syllable duration*	*Difference from next syllable*
1	sɛ	0.1554	1.0629	0.0166
2	te	0.1530	1.0463	0.0560
	te	-	-	-
3	ʃɥI	0.1612	1.1023	0.0058
4	paʀ	0.1620	1.1082	0.1656
5	ti	0.1378	0.9426	0.1488
6	pɑ̃	0.1160	0.7937	0.3115
7	dɑ̃	0.1616	1.1052	0.2664
8	dø	0.1226	0.83879	0.1608
9	mwa	0.0991	0.6780	
10	ɑ̃			
	AVERAGE	*0.1462*		***0.1415***

3.5 Summary

The study examined the speech of Canadian English learners of French at three proficiency levels (beginner, intermediate, and advanced), along with the speech of two native French speakers. The primary objective was to obtain acoustic measures of speech in order to quantify the most salient acoustic cues of Canadian English-accented French, and to link segmental and suprasegmental speech production values. The study was composed of two experimental tasks: a delayed-repetition sentence-production task and the elicitation of spontaneous speech. The first task provided segmental tokens of the vowels and stop consonants under investigation. The second task provided spontaneous-speech samples for the suprasegmental analysis of the rhythm of Anglophone utterances of French. The following chapter will discuss the results obtained from the analysis of the data obtained in the two experimental tasks.

4 RESULTS

This study obtained acoustic measures of speech in an attempt to identify and quantify the most relevant acoustic cues of Canadian English-accented French. Briefly, the three main hypotheses were, H1) the experimental group, or non-native, productions would differ acoustically from native production values, H2) the experimental group productions would demonstrate greater variability, and H3) as language proficiency level increased, acoustic values would become increasingly target-like. The two experimental tasks elaborated in the previous chapter elicited production material from Canadian English-speaking learners and native speakers of French for the chosen five segmental and two suprasegmental cues: 1) VOT of /p/, /t/, and /k/; 2) F2 – F1 acoustic space for /y/ and /u/; 3) diphthongization of /u/, /o/, and /i/; 4) F2 – F1 acoustic space for /o/ and /ɔ/; 5) spectral quality (F1 values) for /a/ in unstressed environments; 6) temporal vowel reduction of /a/, /u/, and /y/; and 7) rhythmic variation measured by the VI. This section successively examines the results obtained for each of the cues under analysis. Results are expressed as group averages with standard deviation values across conditions.

4.1 VOT of /p/, /t/, and /k/

The VOT values for /p/, /t/, and /k/, were obtained in the first delayed-repetition speech-production task. The specific hypotheses predicted that 1) non-native speakers would produce longer VOT values than native speakers, 2) non-native speaker productions would display higher variability, and 3) as language proficiency increased, VOT values would decrease and become less variable (in other words, become more native-like). Five voiceless stop tokens in stressed environments were extracted from target words embedded in carrier sentences. The following table, Table 4.1, summarizes the basic VOT values obtained for each speaker for the consonants /p/, t/, and /k/, as well as average group values.

Table 4.1 Average VOT Values (in ms) for /p/, /t/, and /k/, across Speaker Groups (Standard Deviations).

		/p/	/t/	/k/
Beginner	P11	55.3 (33.4)	65.3 (24.2)	74.4 (27.9)
	P13	64.1 (18.7)	58.0 (9.8)	74.1 (16.8)
	Average	*59.7 (26.0)*	*61.6 (17.0)*	*74.3 (22.3)*
Intermediate	P22	64.7 (29.2)	72.9 (21.0)	85.9 (16.1)
	P24	20.1 (6.5)	30.4 (23.5)	52.4 (16.1)
	Average	*42.5 (17.9)*	*51.7 (22.3)*	*69.2 (16.1)*
Advanced	P32	37.0 (13.8)	39.5 (24.1)	67.3 (8.6)
	P37	30.1 (22.9)	24.2 (9.7)	34.5 (7.8)
	Average	*33.6 (18.3)*	*31.9 (16.9)*	*50.9 (8.2)*
Native	C11	50.4 (32.9)	46.9 (26.8)	69.2 (30.9)
	C12	39.5 (18.4)	35.3 (18.2)	60.9 (18.8)
	Average	*44.9 (25.6)*	*41.1 (22.5)*	*65.1 (24.8)*

The following graphs, Figures 4.1, 4.2, and 4.3, illustrate average VOT production and standard deviation values for each of the voiceless stops under investigation.

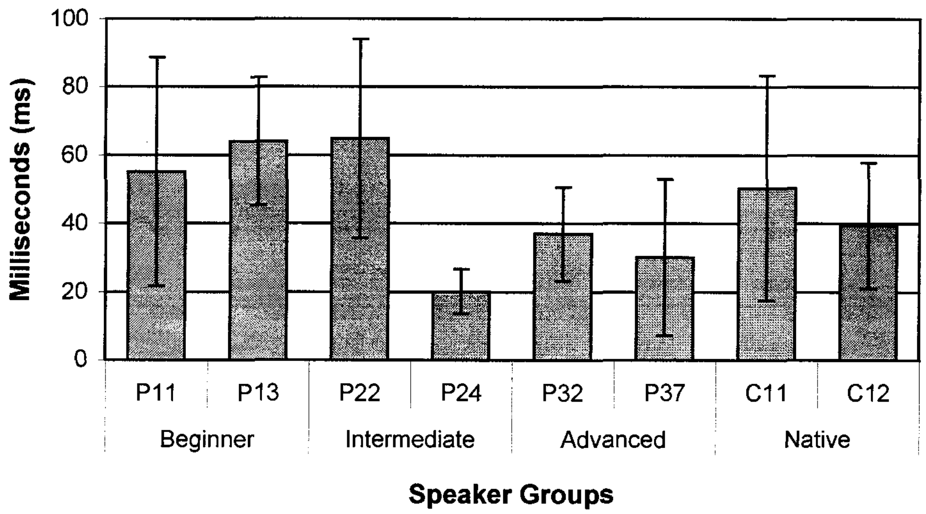

Figure 4.1 *Average VOT of /p/*, with standard deviation represented by error bars.

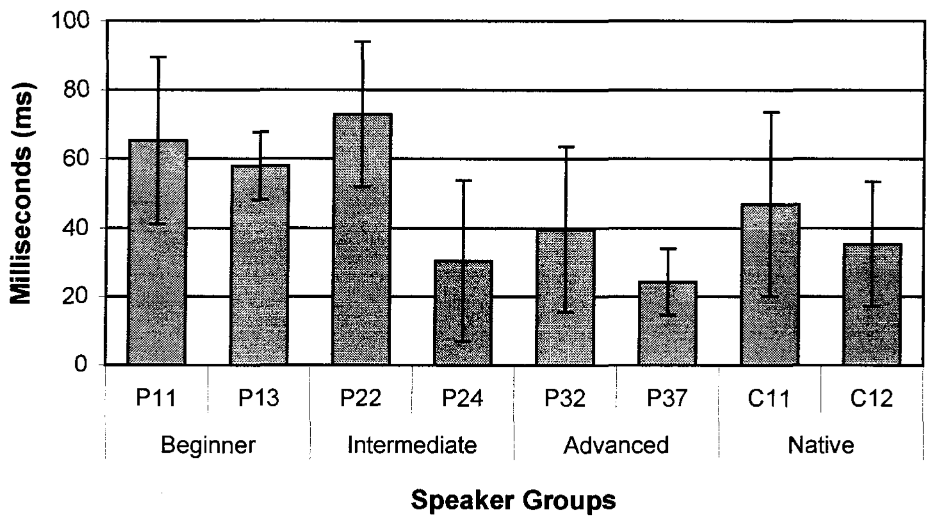

Figure 4.2 *Average VOT of /t/*, with standard deviation represented by error bars.

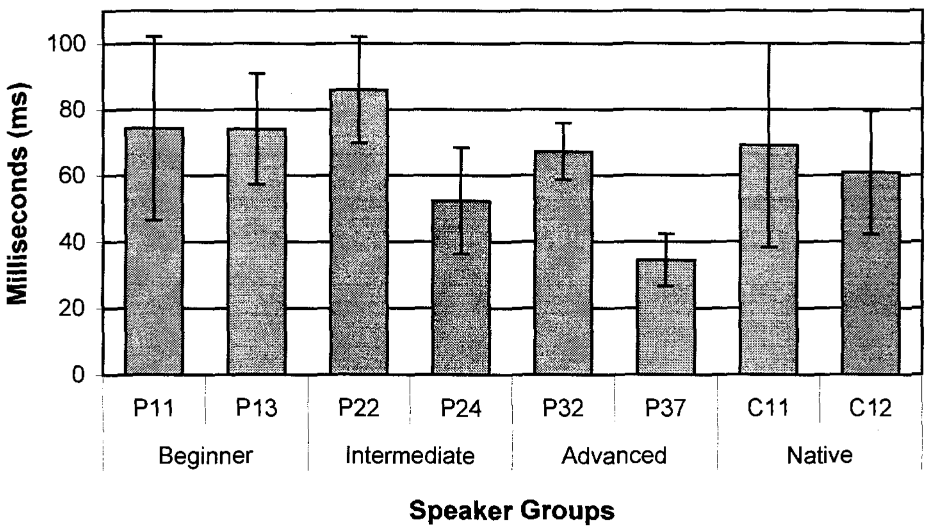

Figure 4.3 *Average VOT of /k/,* with standard deviation represented by error bars.

As predicted, the beginner group produced the longest VOT values across all three stop consonants in stressed position, with average VOT values between 59.7 ms (for /p/) and 74.3 ms (for /k/). However, instead of producing the shortest VOT values, the native speaker group consistently displayed longer VOT values than the advanced speaker group. The native speakers produced average VOT values between 41.1 ms (for /t/) and 65.1 ms (for /k/). Advanced speakers produced average VOT values between 31.9 ms (for /t/) and 50.9 ms (for /k/). In fact, the native speaker group behaved more comparably with the intermediate speaker group, which produced average VOT values between 42.5 ms (for /p/) and 69.2 ms (for /k/).

When examining intra-group variability for VOT performance, the two beginners performed comparably, as did the two native speakers, across all three consonant stops. Speakers within these groups did not differ from each other more than about 10 ms for any of the given stops. Speakers within the advanced speaker group differed more substantially from each other, with a difference of between 7 ms (for average VOT of /p/

values) to 33 ms (for average VOT of /k/ values). Finally, the intermediate speakers differed from each other most substantially, with a difference of between 34 ms (for average VOT of /k/ values) to 45 ms (for average VOT of /p/ values).

The results presented in the table and graphs illustrate that the variability in the production of stop-consonant VOTs (as represented by standard deviation values) is inconsistent with the hypothesis predicting high beginner variability, decreasing towards a lower native speaker variability of VOT production. Instead, the advanced speaker group displayed the lowest average standard deviation across all three stop-consonant production conditions (14.5 ms). The intermediate group displayed the second-lowest average standard deviation values (18.7 ms). Even the beginner speaker group displayed smaller standard deviation values than the native speaker group (beginners: 21.8 ms; natives: 24.3 ms). An examination of individual standard deviations for the VOT results reveals that both native speakers displayed higher standard deviation than four of the six learners, both advanced speakers, as well as one of the intermediate speakers and one of the beginners. Native speaker C11 displayed the highest standard deviation of all speakers, averaging 30.1 ms across all three stop consonants, with beginner speaker P11 displaying comparable standard deviation patterns, at 28.5 ms.

When compared to VOT values found by Nearey and Rochet (1994), these VOT results fit within the range of previously documented VOT values for /p/, /t/, and /k/. The current native speaker productions fall within the upper range of Nearey and Rochet's (1994) documented VOT values for /p/ (18 to 48 ms) and /t/ (21 to 49 ms), while their VOT of /k/ productions are 3 ms longer (32 to 62 ms). When compared to Flege's (1987b) findings for /t/, the current native speaker productions fall between his native bilingual and native monolingual speaker groups, who produced 33 and 51 ms VOT for /t/, respectively. Our native speaker VOT values are much higher than those recorded by Caramazza et al. (1973) for all three stops, and for /p/ recorded by César-Lee (1999).

The non-native speaker VOT productions are considerably higher than those recorded by César-Lee (1999) for /p/, who recorded values between 13 and 35 ms, but somewhat lower than those reported by Flege (1987b) for /t/, who reported values between 42 and 72 ms. No non-native speaker VOT of French /k/ production data was found for comparison in the literature. However, when compared to VOT of /k/ productions reported by Nearey and Rochet (1994) for English monolinguals (VOT between 80 and 87 ms) and French bilinguals (VOT between 32 and 62 ms), the VOT of /k/ values produced by the non-native speakers of French examined in the scope of this study span the two groups, with advanced speakers falling within the French native-speaker VOT range, and beginner and intermediate speakers patterning closer with the native English VOT of /k/ production values.

4.2 F2 – F1 Acoustic Space for /y/ and /u/

Previous research has found that examining the F2 – F1 acoustic space in the production of /y/ and /u/ could be related to the perception of a foreign accent (see Section 3.4.1). The hypotheses motivating the current research predicted the following: H1) non-native speakers would display a larger F2 – F1 acoustic space for /u/ than native speakers, and a smaller F2 – F1 acoustic space for /y/ than native speakers, H2) non-native speaker productions would display higher acoustic variability, and H3) as language proficiency increased, F2 – F1 acoustic space values for both segments would become increasingly target-like and become less variable. Table 4.2 presents the values obtained for each speaker group.

Table 4.2 *Average F2 – F1 Acoustic Space (in Hertz) for /u/ and /y/, across Speaker Groups (Standard Deviations).*

			/u/		*StDev*		/y/		*StDev*
		F1	F2	F2 – F1		F1	F2	F2 – F1	
Beginner	P11	415	1972	1557	(318)	433	2321	1888	(99)
	P13	442	1479	1037	(339)	449	1533	1084	(425)
	Average	*429*	*1726*	*1297*	*(329)*	*441*	*1927*	*1486*	*(262)*
Intermediate	P22	449	1611	1162	(304)	448	1462	1014	(341)
	P24	433	1249	816	(209)	423	2271	1848	(121)
	Average	*441*	*1430*	*989*	*(247)*	*436*	*1867*	*1431*	*(231)*
Advanced	P32	381	1509	1128	(175)	369	1952	1583	(188)
	P37	440	970	530	(75)	429	2215	1786	(198)
	Average	*411*	*1240*	*829*	*(125)*	*399*	*2084*	*1685*	*(193)*
Native	C11	330	1096	766	(104)	321	1934	1613	(129)
	C12	395	1100	705	(101)	415	2056	1641	(139)
	Average	*363*	*1098*	*766*	*(103)*	*368*	*1995*	*1627*	*(134)*

Figures 4.4 and 4.5 below represent the average difference in Hertz in F2 – F1 values (in other words, the F2 – F1 acoustic space) for each of the two segments /u/ and /y/ by speaker.

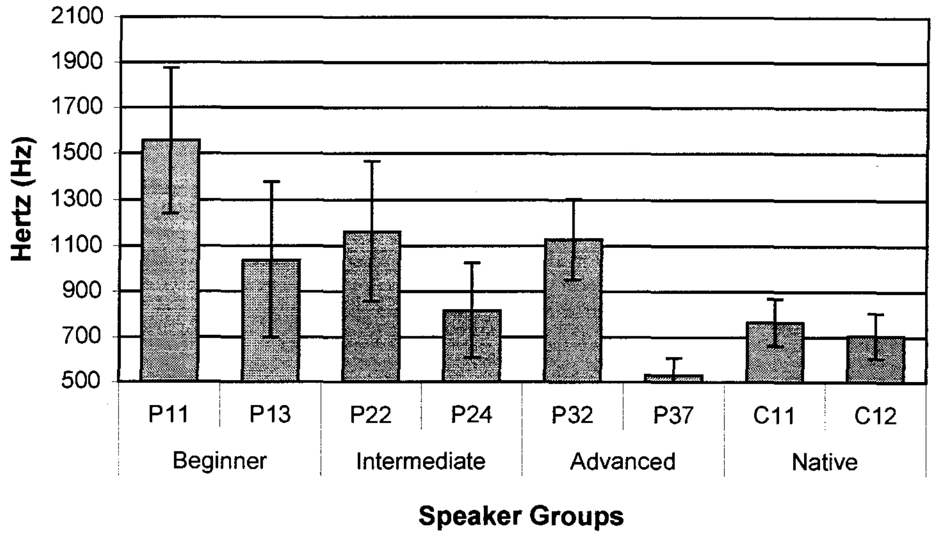

Figure 4.4 *Average F2 – F1 Acoustic Space of /u/,* in stressed position, with standard deviation represented by error bars.

As represented in Table 4.2, the beginner group displayed the highest average F2 – F1 difference for /u/, at 1297 Hz, followed by the intermediate group at 989 Hz; the native speaker group displayed the lowest average F2 – F1 difference at 766 Hz, with the advanced group displaying the second lowest values at 829 Hz. Consequently, an inverse relationship is observable between learner proficiency level and F2 – F1 difference in the production of /u/. In other words, the figure indicates that as learner proficiency level increases, the F2 – F1 acoustic space decreases. Furthermore, standard deviation values also steadily decrease as learner proficiency increases (beginner: 329 Hz; intermediate: 257 Hz; advanced: 125 Hz; native: 103 Hz). These findings would seem to support the predictions of all three of the hypotheses of the current study. However, as illustrated in Figure 4.4, when taking individual variation into account, none of the non-native speaking groups performs consistently, with up to a 600 Hz difference in the F2 – F1 acoustic space for /u/ between the two speakers.

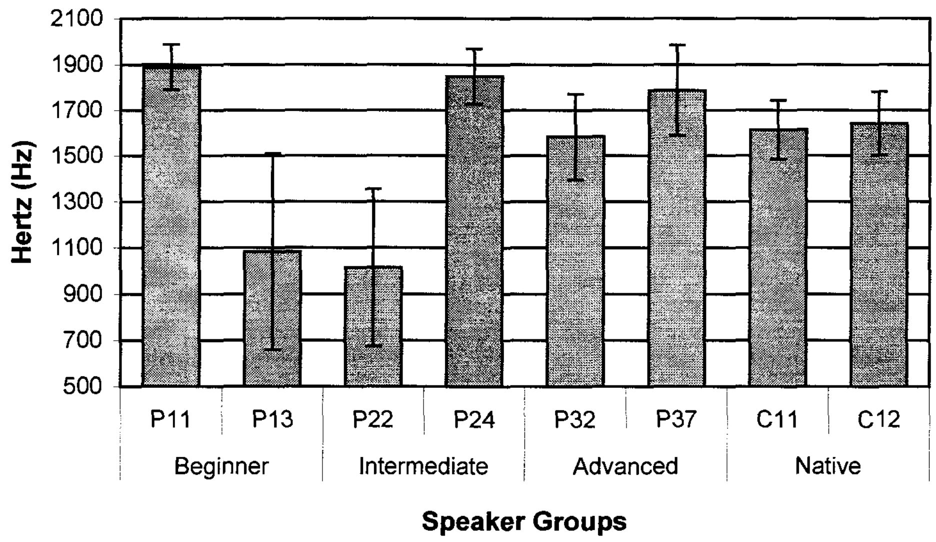

Figure 4.5 *Average F2 – F1 Acoustic Space of /y/,* in stressed position, with standard deviation represented by error bars.

The F2 –F1 acoustic space of /y/ values do not demonstrate the same kind of inverse relationship found with the F2 – F1 acoustic space values of /u/. As shown in Table 4.2, the beginner and intermediate groups displayed a somewhat smaller average F2 – F1 acoustic space for /y/ (beginner: 1486 Hz; intermediate: 1431 Hz) than the advanced and native speaker groups (advanced: 1685 Hz; native: 1627 Hz). However, a slight decrease in variation, as captured by average standard deviation values, is observable as learner proficiency increases (beginner: 262 Hz; intermediate: 231 Hz; advanced: 193 Hz; native: 134 Hz). Consequently, while the H2 prediction of acoustic variation seems to be supported by these results, the predictions of the other hypotheses, H1 and H3, are not clearly substantiated.

In addition, it is evident from individual results illustrated in Figure 4.5 that two speakers (beginner speaker P13 and intermediate speaker P22) produced a considerably smaller F2 – F1 acoustic space for /y/, at around 1040 Hz, than the other speakers, who cluster between 1600 Hz and 1900 Hz. The two speakers produced an F2 – F1 acoustic space for /y/ (approximately 1050 Hz) almost equivalent to their F2 – F1 acoustic space for /u/ (approximately 1100 Hz). These values would indicate that virtually no distinction is being made between the two segments, /u/ and /y/.

In fact, it may be more valuable to consider the F2 – F1 acoustic space results for /u/ and /y/ simultaneously for all of the speakers, in order to gain a more comprehensive understanding of the phenomena under investigation. The values presented in Table 4.2 would seem to indicate that beginners were not producing a relevant acoustic distinction between /u/ and /y/, whereas more advanced learners were able to produce an acoustic distinction between /u/ and /y/. Figure 4.6 below plots the values for the two F2 – F1 acoustic spaces for each speaker. Whereas beginners are not demonstrating a stable /u/ – /y/ distinction, with an average of only 189 Hz between the two F2 – F1 acoustic spaces, this phonological distinction develops an acoustic/spectral correlate as learner proficiency level increases (intermediate: 441 Hz; advanced: 854 Hz; native: 893 Hz).

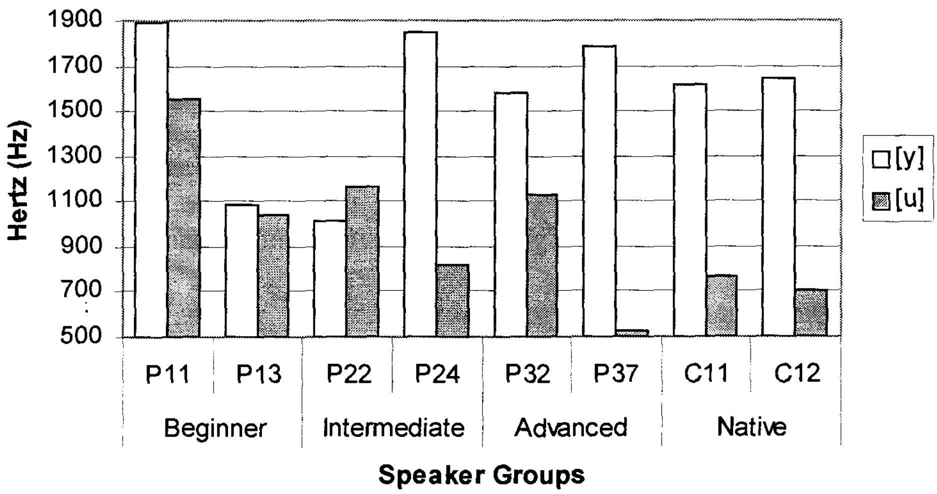

Figure 4.6 Average F2 – F1 Acoustic Space for both /u/ and/y/.

The above figure demonstrates a gradual increase towards more native-like production values as learner proficiency increases, which corresponds to the predictions of H3.

When compared to the F2 – F1 acoustic space results for /u/ and /y/ reported by César-Lee (1999), the current values differ, but follow a similar pattern. César-Lee reports much higher values and increased range for F2 – F1 acoustic space for /y/ across speaker groups (from 1398 to 2026 Hz) than the present study (1431 to 1685 Hz). She also reports higher values and increased range for F2 – F1 acoustic space for /u/ across speaker groups (from 1196 to 2033 Hz) than the present study (736 to 1297 Hz). However, while the absolute values may differ, the same /u/ and /y/ acoustic production patterns emerge, with early learners not making a clear distinction in their F2 – F1 acoustic spaces for /u/ and /y/, whereas more advanced and native speakers produce a more pronounced distinction between the two segments, as captured by discrete acoustic values for the two F2 – F1 acoustic spaces.

4.3 Diphthongization of /u/, /o/ and /i/

This study examined the production of /u/, /o/, and /i/, with the intention of capturing potential diphthongizational phenomena. The hypotheses predicted the following for this speech parameter: H1) non-native speakers would demonstrate greater spectral variation throughout the vowel steady-state, H2) non-native speakers would demonstrate greater acoustic variability in their productions, and H3) as language proficiency increased, production values would become increasingly target-like.

Diphthongs are often characterized by F1- and/or F2-shifting within the vowel nucleus (Martin, 2002; Schouten & Peeters, 2000). Schouten and Peeters (2000) characterize diphthongs as containing an F2 formant shift between 200 and 400 Hz. While the significance of the absolute value of the F2 shift depends on its relation to the speaker's fundamental frequency, the purpose of the present study was to provide a preliminary quantification of any potential intra-speaker F2-shifting. Table 4.3 presents the difference (in Hertz) of the formant values for each target vowel at two intervals in the vowel steady-state: T1 (¼ point of the vowel steady-state) and T3 (¾ point of the vowel steady-state). The table presents absolute values in order to capture formant variation, as opposed to the directionality of F2-shifting.

Table 4.3 Formant Shifting, as Represented by the Average Difference (in Hertz) between T1 and T3, for /u/, /o/, and /i/ (Standard Deviations).

			/u/		/o/		/i/	
Beginner	P11	F1:	38	(27)	63	(36)	65	(56)
		F2:	148	(119)	252	(110)	57	(56)
	P13	F1:	36	(27)	36	(26)	58	(39)
		F2:	245	(218)	64	(90)	72	(63)
	Average	*F1:*	*37*	*(27)*	*49*	*(31)*	*62*	*(48)*
		F2:	*197*	*(169)*	*158*	*(100)*	*65*	*(60)*
Intermediate	P22	F1:	47	(27)	96	(87)	60	(27)
		F2:	203	(118)	171	(155)	96	(45)
	P24	F1:	23	(11)	98	(23)	35	(11)
		F2:	225	(160)	221	(95)	96	(61)
	Average	*F1:*	*35*	*(19)*	*97*	*(55)*	*48*	*(19)*
		F2:	*214*	*(139)*	*196*	*(125)*	*96*	*(53)*
Advanced	P32	F1:	14	(15)	38	(16)	24	(25)
		F2:	195	(102)	221	(175)	112	(90)
	P37	F1:	29	(28)	38	(35)	38	(36)
		F2:	50	(50)	135	(135)	52	(47)
	Average	*F1:*	*21*	*(22)*	*38*	*(26)*	*31*	*(31)*
		F2:	*123*	*(76)*	*178*	*(155)*	*82*	*(69)*
Native	C11	F1:	12	(18)	16	(13)	6	(5)
		F2:	145	(93)	100	(97)	60	(58)
	C12	F1:	20	(20)	24	(11)	4	(6)
		F2:	122	(67)	151	(47)	70	(55)
	Average	*F1:*	*16*	*(19)*	*20*	*(12)*	*5*	*(6)*
		F2:	*134*	*(80)*	*126*	*(72)*	*65*	*(57)*

The values presented in Table 4.3 illustrate that while some spectral variation exists within the vowel steady-state, this acoustic variation is limited and inconsistent.

Although Schouten and Peeters (2000) do not quantify or characterize F1-shifting within diphthongization phenomena, F1-shifting results are briefly presented here. When considering group averages, spectral variation corresponding to F1-shifting remains relatively minor: the highest variation is displayed by the intermediate speaker group for the segment /o/, at 97 Hz, with the beginner group displaying the next highest F1-shifting, at 62 Hz for /i/. When considering individual performance, beginner speaker P11 displays relatively higher F1-shifting for /o/ and /i/, with T1 – T3 variation around 64 Hz. Intermediate speakers P22 and P24 display the highest F1-shifting, with T1 – T3 variation for /o/ at 96 Hz for speaker P22, and 98 Hz for speaker P24.

Spectral variation of F2 also remains relatively minor, whether considering group or individual averages. For example, only once do group averages of F2-shifting exceed 200 Hz – the necessary diphthong threshold mentioned by Schouten and Peeters (2000) – which the intermediate group displays for the segment /u/, at 214 Hz. When considering individual performance, beginner speaker P13 and both intermediate speakers (P22 and P24) display F2-shifting greater than 200 Hz for the segment /u/; beginner speaker P11, intermediate speaker P24, and advanced speaker P32, display F2-shifting greater than 200 Hz for the segment /o/. On the other hand, no speakers produced F2-shifting greater than 200 Hz for the segment /i/. Figures 4.7, 4.8, and 4.9 below present absolute F2-shifting values in graph form for individual and group averages for /u/, /o/, and /i/, respectively. As with Table 4.3, rising and falling F2-shifting results are treated in their absolute values, in order to illustrate the amount of formant variation, as opposed to directionality of F2-shifting.

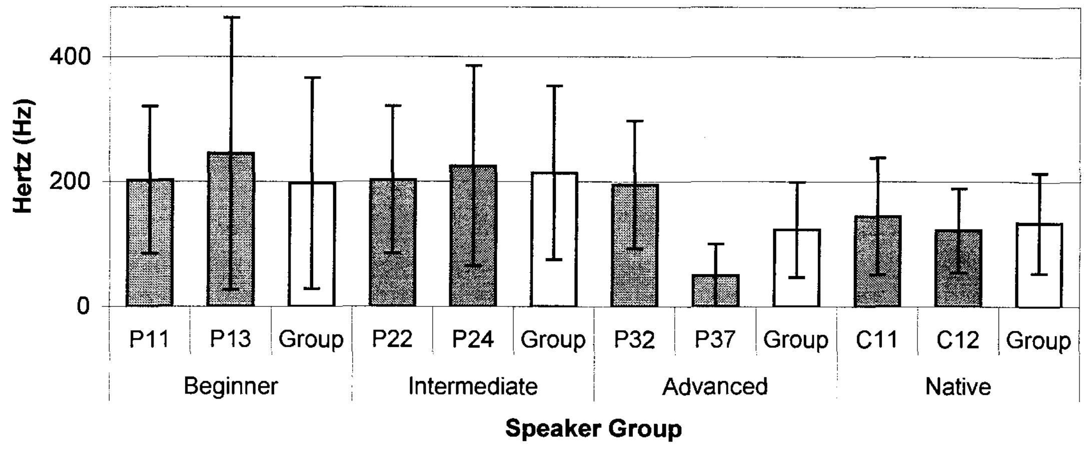

Figure 4.7 *Average F2-shifting within Vowel Steady-state for /u/,* absolute F2 difference between T1 and T3, with standard deviation represented by error bars.

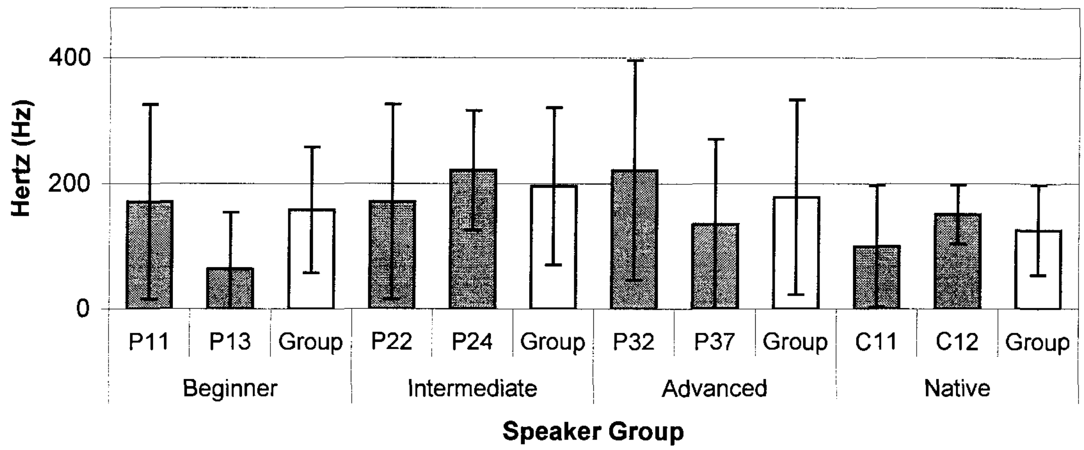

Figure 4.8 *Average F2-shifting within Vowel Steady-state for /o/,* absolute F2 difference between T1 and T3, with standard deviation represented by error bars.

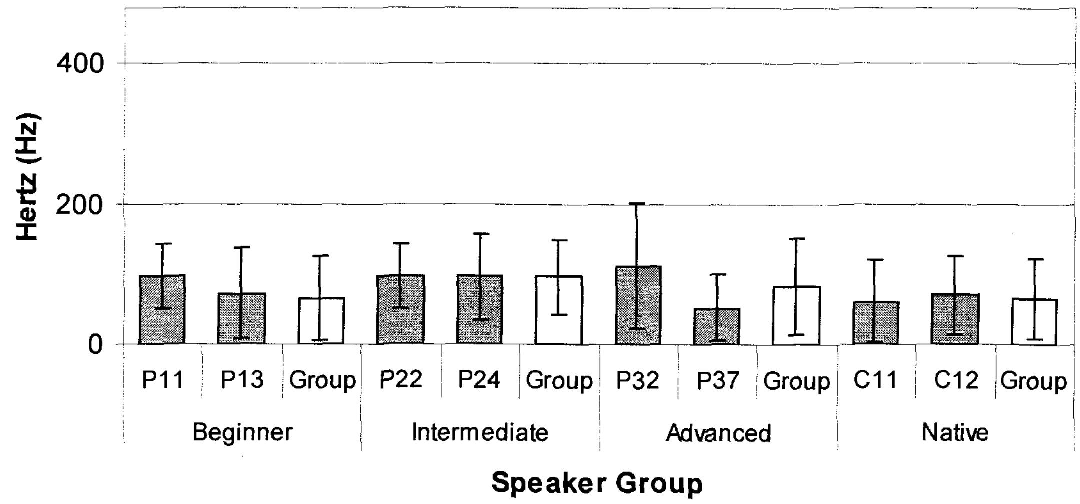

Figure 4.9 *Average F2-shifting within Vowel Steady-state for /i/,* absolute F2 difference between T1 and T3, with standard deviation represented by error bars.

As illustrated in the above figures, potential diphthongization effects for /u/, /o/, and /i/ differ somewhat according to target cue. First, /u/ steady-state F2-shifting demonstrates the highest values of all of the three segments under evaluation. The group average for intermediate speakers just attains the requisite 200 Hz threshold for diphthong characterization, as proposed by Schouten and Peeters (2000). Furthermore, when standard deviations are considered, some average beginner and intermediate group productions will also fall within the 200 to 400 Hz steady-state variation range needed for diphthong characterization. In other words, the standard deviation indicates that there are raw F2-shifting values within the 200 to 400 Hz range. In fact, some raw values for beginner speaker P13's F2-shifting are above 400 Hz. Even the native speaker group standard deviation values reach just above the 200 Hz variation range. The advanced speaker group average values do not reach the 200 Hz threshold, even when taking standard deviation into account.

Second, /o/ steady-state F2-shifting also suggests the potential presence of diphthongization, especially when considering non-native speakers' values of formant variation. While none of the group average values falls above the 200 Hz steady-state F2 shifting threshold, standard deviations for all non-native speaker groups extend well into the 200 to 400 Hz steady-state variation range needed for diphthong characterization. When taking individual results into account, beginner speaker P11, intermediate speaker P24, and advanced speaker P32 all produce mean F2-shifting values within the 200 to 400 Hz range.

Third, /i/ steady-state F2-shifting is quite low across all group average results, never reaching the requisite 200 Hz, even when taking standard deviations into consideration. Only the raw values of advanced speaker P32 demonstrate F2-shifting to within the 200 Hz range. The beginner speaker and native speaker group average F2-shifting values are quite low, at only 65 Hz, suggesting a fairly consistent formant steady-state production across the duration of the vowel.

Overall, these mean value results do not support the hypotheses predicting the presence of inappropriate diphthongizational phenomena in non-native French productions, or a decrease in acoustic variation as learner proficiency increases. The standard deviation values representing acoustic variation sometimes almost match the production values, indicating large deviations from the average. Unfortunately, due to a paucity of acoustic data on this speech dimension for vocalic productions of Anglophone learners of French, no other values are available for comparison to the current dataset.

4.4 F2 – F1 Acoustic Space for /o/ and /ɔ/

Foreign-accentedness has been anecdotally attributed to non-target-like vowel production in /o/ and /ɔ/ environments (Dansereau, 1990; Le Clézio, 1986). As documented in Martin (1998), an acoustic dimension to the phonological distinction /o/ – /ɔ/ remains relevant for native French speakers. Therefore, this study examined speaker production of /o/ and /ɔ/ across ten tokens and the hypotheses predicted the following: H1) non-native speakers would display a larger F2 – F1 acoustic space for /o/ than native speakers, and a smaller F2 – F1 acoustic space for /ɔ/ than native speakers, H2) non-native speaker productions would display higher acoustic variability, and H3) as language proficiency increased, F2 – F1 acoustic space values for both segments would become increasingly target-like and become less variable. Table 4.4 presents the values representing the F2 – F1 acoustic space for /o/ and /ɔ/ across all four speaker groups.

Table 4.4 *Average F2 – F1 Acoustic Space (in Hertz) for /o/ and /ɔ/, across Speaker Groups (Standard Deviations).*

			/o/		*StDev*		/ɔ/		*StDev*
		F1	F2	F2 – F1		F1	F2	F2 – F1	
Beginner	P11	513	1252	739	(109)	684	1305	621	(120)
	P13	555	1174	619	(57)	676	1183	507	(142)
	Average	*534*	*1213*	*678*	*(83)*	*680*	*1244*	*564*	*(131)*
Intermediate	P22	683	1271	588	(174)	674	1141	467	(99)
	P24	575	1249	674	(134)	749	1643	894	(159)
	Average	*629*	*1260*	*631*	*(154)*	*712*	*1392*	*681*	*(129)*
Advanced	P32	569	1242	673	(169)	685	1300	615	(115)
	P37	484	984	500	(189)	644	1415	771	(240)
	Average	*527*	*1113*	*587*	*(179)*	*665*	*1358*	*693*	*(178)*
Native	C11	414	906	492	(70)	610	1284	674	(90)
	C12	473	984	511	(138)	598	1385	787	(126)
	Average	*444*	*945*	*502*	*(104)*	*604*	*1335*	*731*	*(108)*

As detailed in Table 4.4 above, beginners demonstrated larger F2 – F1 acoustic space for /o/, at 678 Hz, than for /ɔ/, at 564 Hz, which is contrary to pronunciation trends found in native speaker productions. Intermediate learners displayed a difference of only 50 Hz between the F2 – F1 acoustic space for /o/ (631 Hz) and /ɔ/ (681 Hz), but the appropriate native-like compact/diffuse trend in the relation between F2 – F1 acoustic spaces began to emerge with this speaker group. Advanced learners also demonstrated an appropriate compact/diffuse contrast between F2 – F1 acoustic spaces for /o/ (587 Hz) and /ɔ/ (693 Hz), which was somewhat more pronounced than that of the intermediate learners. Finally, the native speakers demonstrated the most pronounced compact/diffuse relation between F2 – F1 acoustic spaces for /o/ (502 Hz) and /ɔ/ (731 Hz). Figures 4.10, 4.11, and 4.12 below illustrate the F2 – F1 acoustic space values for /o/ and /ɔ/, respectively, across individual speakers.

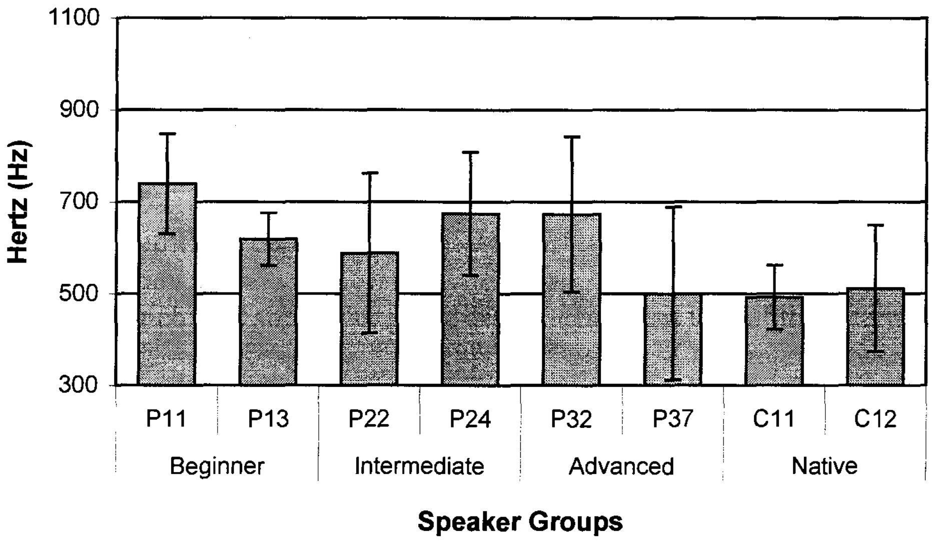

Figure 4.10 *Average F2 – F1 Acoustic Space of /o/,* in stressed position, with standard deviation represented by error bars.

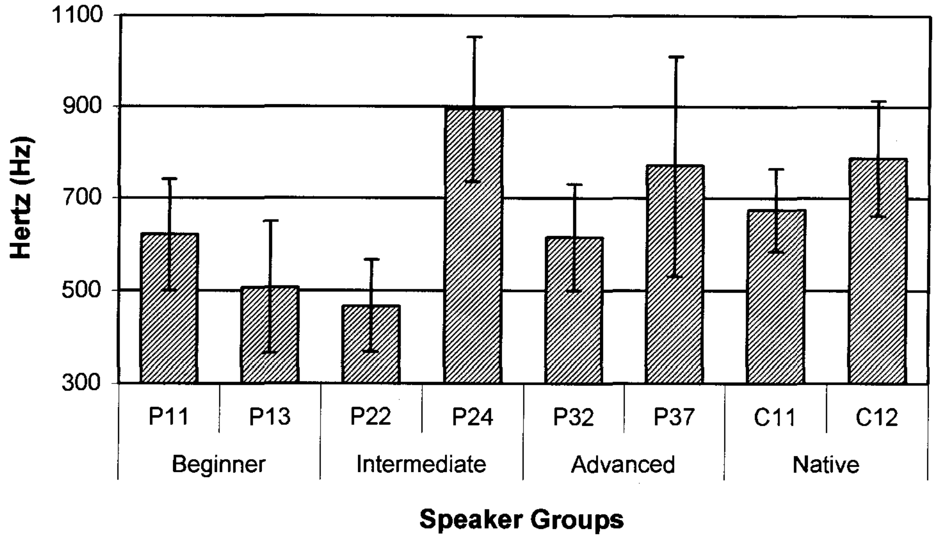

Figure 4.11 *Average F2 – F1 Acoustic Space of /ɔ/,* in stressed position, with standard deviation represented by error bars.

The mean production values for the F2 – F1 acoustic space for /o/ and /ɔ/ display a chiasmus-like relationship, as illustrated in Figure 4.12 below. In other words, F2 – F1 of /o/ decreases as language proficiency increases (inverse relationship), and F2 – F1 of /ɔ/ increases as language proficiency also increases (direct relationship).

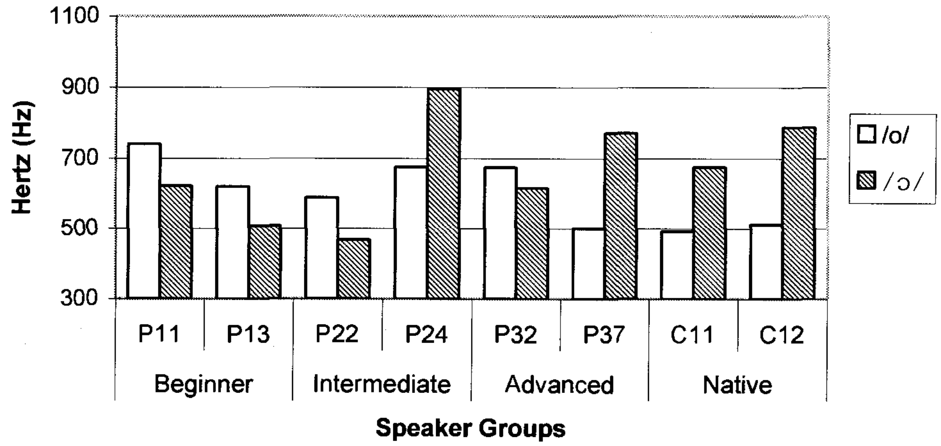

Figure 4.12 *Average F2 – F1 Acoustic Space for both /o/ and /ɔ/,* values display a chiasmus-like relationship.

In light of these results, this study's hypotheses (H1 and H3) seem to be supported as follows: H1) non-native speakers are producing a larger F2 – F1 acoustic space for /o/ than native speakers, and a smaller F2 – F1 acoustic space for /ɔ/ than native speakers; and H3) as language proficiency increases, F2 – F1 acoustic space values for both segments become increasingly target-like.

On the other hand, some important intra-group variation is present – most notably the 427 Hz difference in F2 – F1 acoustic space between intermediate speakers P22 and P24, for /ɔ/. In addition, the learner proficiency variability predictions of H2 are not supported by these results, as the beginner group and the native group display the lowest standard deviation values, while the intermediate and advanced speakers display the highest standard deviation values for both vowels.

Martin's (1998) acoustic findings for the production of /o/ and /ɔ/ suggest that female native speakers make a distinction between the F2 – F1 acoustic spaces of /o/ and /ɔ/. He reports that the F2 – F1 acoustic space for /o/ is 423 Hz, and 828 Hz for /ɔ/. However, in a formant value chart for female French vowels, Martin (2004) presents very similar F2 – F1 acoustic space values for both segments: 430 Hz for /o/, and 490 Hz for /ɔ/[15]. The current study suggests that native speakers of French do make a distinction between the two segments with a 229 Hz difference between the two F2 – F1 acoustic spaces (502 Hz for /o/, and 731 Hz for /ɔ/), which is consistent with Martin's (1998) findings. Regrettably, no previous learner data is available for comparison with the present non-native speaker results on this dimension of French learner speech.

[15] No explanation of the different results is available, as the Martin (2004) material is a stand-alone vowel chart posted on the internet, without any methodological account.

4.5 F1 of Unstressed /a/

This study examined the F1 production values of /a/, with the intention of capturing potential centralization phenomena in unstressed syllables. The hypotheses predicted the following: H1) non-native speakers would display lower F1 values for unstressed /a/ than native speakers, H2) the non-native speakers would display higher acoustic variability in their F1 values than native speakers, and H3) as language proficiency increased, F1 values would also increase. Table 4.5 presents average F1 and standard deviation values of unstressed /a/ for each speaker group, with Figure 4.13 providing graphic illustration.

Table 4.5 Average F1 Values of Unstressed /a/ (Standard Deviations).

		/a/	
Beginner	P11	761	(86)
	P13	715	(83)
	Average	*738*	*(85)*
Intermediate	P22	871	(72)
	P24	975	(45)
	Average	*923*	*(59)*
Advanced	P32	783	(33)
	P37	819	(75)
	Average	*801*	*(54)*
Native	C11	696	(82)
	C12	787	(49)
	Average	*742*	*(66)*

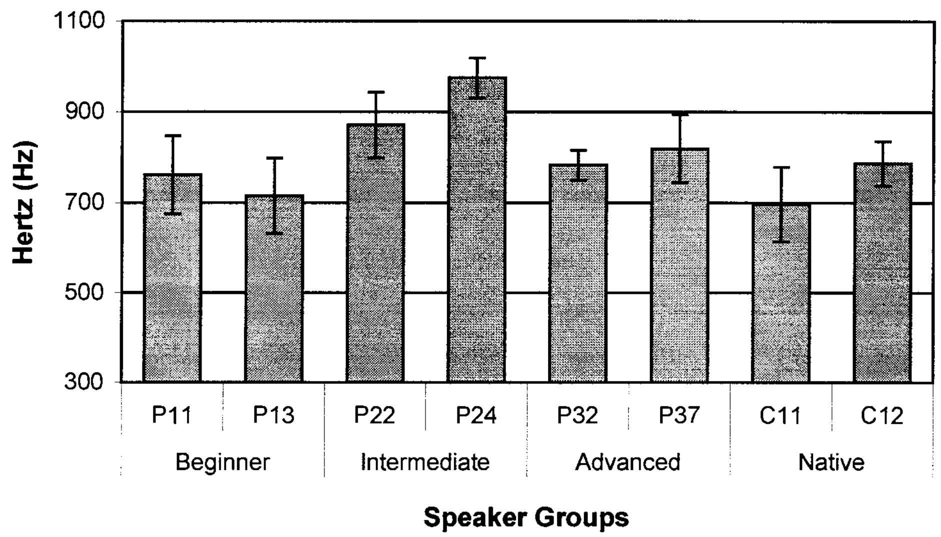

Figure 4.13 *Average F1 of Unstressed /a/,* with standard deviation represented by error bars.

The results do not support any of the study's hypotheses predicting pronunciation performance for unstressed /a/. The values do not indicate any patterning according to language proficiency. Moreover, standard deviation values are fairly consistent across all speaker groups (between 33 and 86 Hz). Finally, none of the F1 values for unstressed /a/ falls below Delattre's (1981) /a/ value threshold of 650 Hz for native speakers of French. Therefore, these results would suggest that the subjects are not centralizing their vowels in unstressed syllables, contrary to the study's hypothesis.

4.6 Temporal Vowel Reduction in Unstressed Environments for /u/, /y/, and /a/

A temporal variation measurement was conducted to quantify temporal reduction in unstressed environments for /u/, /y/, and /a/. Vowel measurements were taken from two sets of word stimuli: first they were measured in a stressed environment, and second, in an unstressed environment (penultimate syllable) (see Appendix A). The hypotheses predicted the following: H1) non-native speakers would display more temporal vowel reduction in unstressed environments than native speakers, H2) non-native speakers would demonstrate higher acoustic variability, and H3) as language proficiency increased, temporal vowel reduction values would decrease towards native-like values and become less acoustically variable.

Table 4.6 below presents average durational percentages for each speaker group. The *average duration ratio* is the unstressed vowel duration expressed as a percentage of stressed vowel duration. For example, the duration of a token of unstressed /u/ would be expressed as a percentage of the duration of a token of stressed /u/. Therefore, there is an inverse relationship between the percentage value and vowel reduction phenomena: the higher the percentage, the less temporal vowel reduction is taking place, since the unstressed vowel is closer in duration to the stressed vowel. Conversely, the lower the percentage value, the more temporal vowel reduction is taking place, as the unstressed vowel duration decreases compared to stressed vowel duration values. Figure 4.14 illustrates the data values presented in Table 4.6.

Table 4.6 Average Durational Ratio for /u/, /y/, and /a/, across Speaker Groups (Standard Deviations).

	/u/		/y/		/a/	
Beginner	65%	(51)	24%	(9)	42%	(22)
Intermediate	38%	(14)	32%	(13)	43%	(14)
Advanced	60%	(14)	53%	(15)	84%	(13)
Native	41%	(17)	55%	(9)	65%	(13)

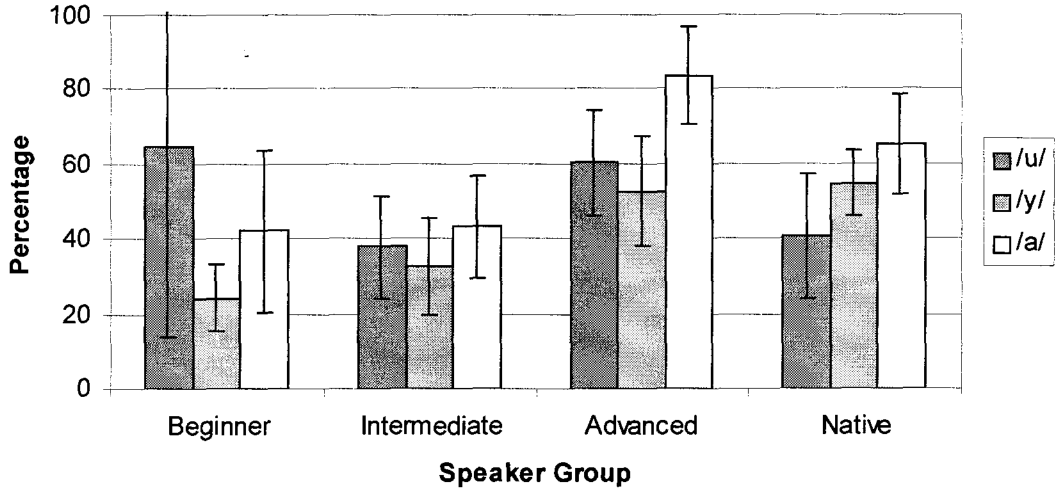

Figure 4.14 *Average Durational Ratio for /u/, /y/, and /a/,* unstressed vowel duration expressed as a percentage of stressed vowel duration, with standard deviation represented by error bars.

As illustrated in the figure above, the hypotheses did not accurately predict non-native and native speaker performance in the temporal reduction of /u/, /y/, and /a/. For /u/, the native speaker group and intermediate group demonstrated the most temporal vowel reduction, expressed as the lowest average durational percentages for unstressed /u/. Contrary to the hypothesis, beginners actually produced the least amount of temporal vowel reduction for /u/, with unstressed /u/ production duration at 65% of stressed /u/ duration. Advanced speaker productions fell 5% behind the beginner speaker group, with unstressed /u/ duration at 60% of stressed /u/ duration. While insufficient temporal reduction could be a sign of inadequate mastery of the language or the result of an unnatural speech elicitation situation (e.g., the delayed-repetition of a constructed sentence), the fact is that both the beginner speaker group, as well as the advanced speaker group display the least amount of temporal reduction. It would therefore be difficult to attribute inadequate language proficiency as the cause of decreased temporal vowel reduction.

On the other hand, temporal vowel reduction values for /y/ and /a/ did not pattern the same as for /u/. In this case, beginners display the greatest temporal reduction, with unstressed /y/ duration at 24% of stressed /y/ duration. Intermediate speakers also demonstrate high temporal reduction, with unstressed /y/ duration at 32% of stressed /y/. The advanced and native speaker groups performed comparably, producing unstressed /y/ at 54% and 55% of stressed /y/, respectively.

Unstressed /a/ duration values range from about 43% of stressed /a/ duration for the beginner and intermediate groups, to 65% of stressed /a/ duration for the native speaker group, and 85% of stressed /a/ duration for the advanced speaker group. The temporal vowel reduction results for /y/ and /a/ would seem to suggest some effect for language proficiency level in the temporal reduction of unstressed vowels.

Furthermore, the standard deviation values presented in Table 4.6 and illustrated in Figure 4.12 seem to corroborate the predictions of H2 and H3, which predicted that non-native speaker productions would be more temporally-variable, and that as language proficiency increased, variability would decrease. For /u/, beginners display extremely high variability at 51%, with the rest of the groups clustering between 14 and 17%. For /y/, all of the groups displayed relatively low variation, with beginners and the native speakers displaying only 9% variability, and the intermediate and advanced speaker groups around 14% variability. For /a/, beginners once again display the highest variability, with a standard deviation score of 22%, while the rest of the speaker groups perform similarly with 13 to 14% standard deviation scores. Overall, these results suggest that beginners are demonstrating the most variability, whereas the rest of the speaker groups are performing relatively comparably. Unfortunately, no previous data on the temporal reduction of /u/, /y/, and /a/ by French learners is available for comparison with these results.

4.7 Rhythmic Variation

Using the spontaneous speech elicited from speakers across all groups, a Variability Index was calculated based on five utterances for each speaker. The VI provides two measurements: 1) a rhythmic variability rating, with higher scores reflecting a higher inter-syllabic variability; and 2) average syllable duration. The hypotheses predicted the following: H1) non-native speakers would obtain higher VI scores than native speakers, H2) non-native speakers would demonstrate greater variability, and H3) as language proficiency increased, VI scores would decrease, and become less variable. Table 4.7 presents the data obtained in the Variability Index calculations across all speakers, and Figure 4.15 illustrates the VI values graphically.

Table 4.7 *Variability Index Calculations, based on 8-syllable Spontaneous Utterances (Standard Deviation).*

		Avg. Syl. Duration (ms)		*Variability Index*	
Beginner	P11	397.6	(92.0)	0.6869	(0.1146)
	P13	272.4	(75.8)	0.6170	(0.2460)
	Average	*335.0*	*(83.9)*	*0.6519*	*(0.1803)*
Intermediate	P22	334.9	(90.6)	0.9689	(0.0507)
	P24	238.1	(67.0)	0.5086	(0.1124)
	Average	*286.5*	*(78.8)*	*0.7388*	*(0.0815)*
Advanced	P32	213.0	(63.3)	0.5475	(0.1920)
	P37	183.4	(13.2)	0.4196	(0.0943)
	Average	*198.2*	*(38.2)*	*0.4836*	*(0.1432)*
Native	C11	149.6	(20.9)	0.2447	(0.1266)
	C12	192.5	(29.4)	0.4623	(0.0993)
	Average	*171.1*	*(25.2)*	*0.3535*	*(0.1129)*

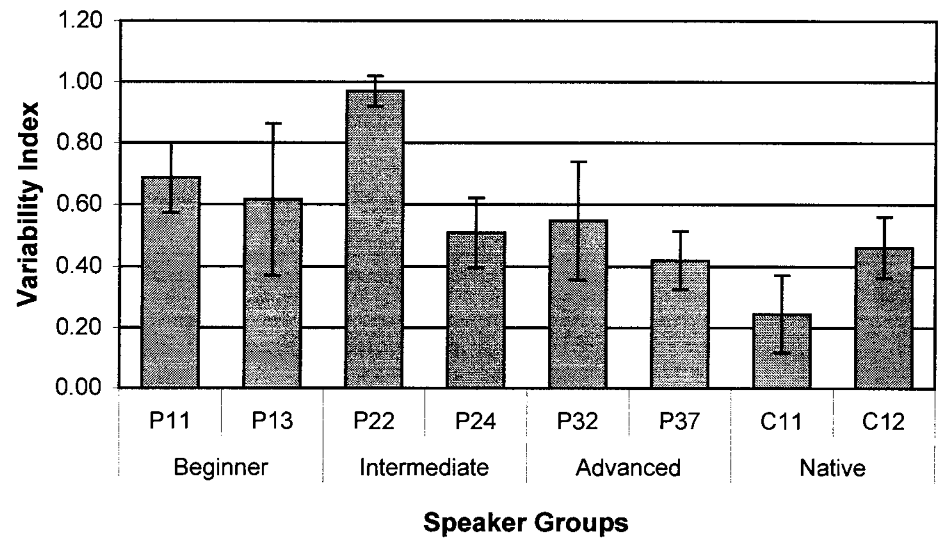

Figure 4.15 *Average Variability Indices*, with standard deviation represented by error bars.

As detailed in Table 4.7, native speakers displayed the least amount of inter-syllabic variability, with an average score of 0.3535, followed closely by the advanced speaker group, which scored 0.4836. Both the beginner and intermediate speaker groups displayed high VI ratings, with scores of 0.6519 and 0.7388, respectively. While the beginner group scored slightly closer to the native speaker group than the intermediate group (due to the extremely high variability score of intermediate subject P22, at 0.9689), these results could still suggest a trend toward native speaker inter-syllabic variability as language proficiency increases, as hypothesized in the study.

These results seem to confirm the study's hypotheses (H1 and H3) insofar as non-native speakers are demonstrating higher inter-syllabic variability, and as language proficiency increases, inter-syllabic variability decreases. Guilbault's (2002) Variability Index analyses found that intermediate learners of French scored 0.4056; advanced learners scored 0.4298; native Canadian French speakers scored 0.3508; and native European French speakers scored 0.2860 (p. 119). While the results of this study are considerably higher in value than Guilbault's findings for native speakers and learners of French, the same trend is apparent, with the least experienced learners of French obtaining the highest rhythmic variability scores, and native French speakers obtaining the lowest rhythmic variability scores.

On the other hand, hypothesis H2 is not supported by the present results as all speaker groups demonstrate consistently low variability, with individual standard deviation values ranging between 0.0507 and 0.2460. In fact, these standard deviation values are lower than previously recorded values for inter-syllabic variability, as presented by Guilbault (2002, p. 119), who reports standard deviation values between 0.1547 and 0.2486 for native speakers and learners of French. Finally, Deterding (2001) explains that Variability Index measurements do not reflect speech rate, but are, rather, an independent measurement of inter-syllabic variability; therefore, variability results should not be biased by differing speaking rates.

However, the average syllable duration values obtained in the present Variability Index analysis suggest that speaking rate may correlate with language proficiency, as illustrated in Figure 4.16 below. In other words, average syllable duration, which is representing speech rate here, seems to decrease as language proficiency increases. Beginners display the slowest speaking rate, with average syllable durations of 335 ms, while intermediate learners produce syllables of somewhat shorter duration, at 287 ms. Advanced speakers produce syllables at an average duration of 198 ms, with native speakers producing the shortest syllable durations, at 171 ms each.

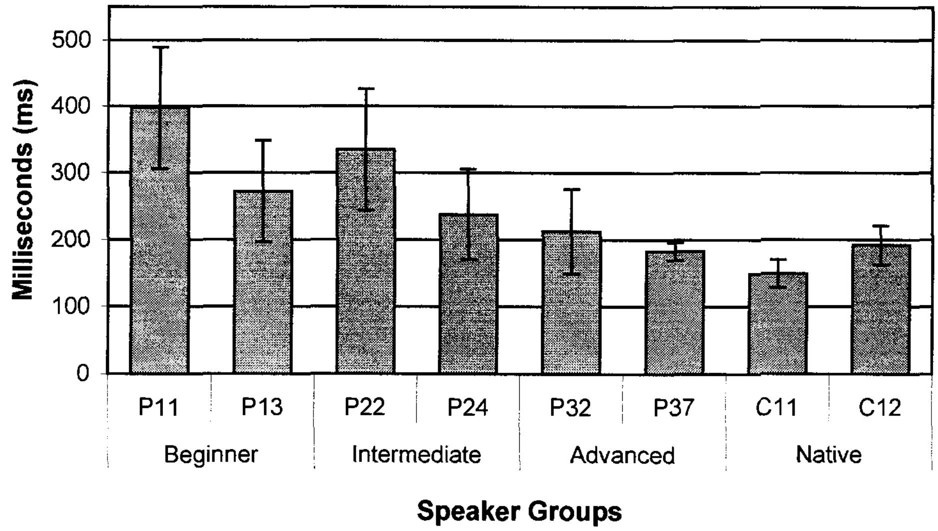

Figure 4.16 *Average Syllable durations,* with standard deviation represented by error bars.

If taken as an indicator of speech rate, these average syllable duration values suggest that as language proficiency increases, speech rate increases towards more native-like values. These results therefore seem to support the hypotheses predicting speaker performance varying according to language proficiency. Unfortunately, no previous VI syllable duration results for learners and native speakers of French were found for comparison. In addition, standard deviation values also decrease as learner proficiency increases, supporting the predictions of hypothesis H2. The beginner and intermediate groups demonstrate the highest variability at 84 ms average standard deviation and 79 ms average standard deviation, respectively. The advanced and native speaker groups display considerably lower average standard deviation values at 38 ms and 25 ms, respectively.

4.8 Summary

The values presented in this chapter are the result of the measurement and analysis of seven segmental and suprasegmental cues found in the elicited and the spontaneous speech of learners and native speakers of French. These cues include the VOT of syllable-initial /p/, /t/, and /k/; the F2 – F1 acoustic space of /y/ and /u/; the spectral quality of /u/, /o/ and /i/ for the presence of possible diphthongizational properties; the F2 – F1 acoustic space of /o/ and /ɔ/; the possible F1 spectral centralization of /a/ in unstressed syllables; the possible temporal reduction of /u/, /y/ and /a/ in unstressed syllables; and the calculation of inter-syllabic variability according to the Variability Index.

Some evidence of values and measurements trending towards native-like production along the language proficiency variable (beginner, intermediate, and advanced) exists for three of the seven cues: the F2 – F1 acoustic space of /y/ and /u/, the F2 – F1 acoustic space of /o/ and /ɔ/ (target-like provision), and inter-syllabic variability according to the calculations of the Variability Index. Findings are inconclusive for the other four cues under analysis, which could be due to various factors, and will be discussed in the following chapter.

5 DISCUSSION

The aim of this research was to provide an acoustic dimension to the Canadian-English accent in French. In previous research, L2 learners' speech productions displaying greater acoustic variability than native speaker productions, or differing considerably from native speakers' productions, were perceived as more accented than L2 learners' speech productions patterning more closely to native speaker productions (César-Lee, 1999; Jonasson & McAllister, 1972; Magen, 1998; Shah, 2002). The present study investigated three hypotheses predicting that H1) the experimental groups would demonstrate different acoustic values than native speakers of French for all acoustic cues under analysis; H2) non-native speakers would demonstrate greater acoustic variability, and H3) as the learners' language proficiency increased, acoustic values would become increasingly target-like and display less acoustic variability.

In examining the speech of female learners of French for seven potential acoustic cues of foreign accent, the findings of the study suggest that while the current hypotheses predicting pronunciation trends according to learner proficiency may hold true for certain cues, inconclusive results obtained in the analysis of many of the cues may either cast doubt on the predictive power of the hypotheses, indicate that the cues in question are not related to the acoustic realization of a Canadian English-accented French, or reveal that other methodologies should be explored. The results obtained in the scope of this study will be discussed for the specific acoustic cues measured, followed by a more general discussion.

5.1 Cues under Investigation

5.1.1 VOT of /p/, /t/, and /k/

The present study found that while beginners consistently produced the longest VOT values across all three consonant stops, the native speakers clustered with intermediate speakers with the second longest VOT values, and the advanced speakers produced the shortest VOT values. Furthermore, although standard deviation scores seemed to decrease with language proficiency, native speakers displayed the highest standard deviation in their VOT productions. These results do not confirm the hypotheses because, although there seems to be a decrease in VOT values according to learner proficiency, native speakers do not provide the anticipated, lowest VOT production values. While VOT values obtained by César-Lee (1999) and Flege (1987b) for speakers of increasing language proficiency and native speakers have demonstrated inverse correlations (higher proficiency or more native-like accent = lower VOT values), there could be some potential reasons for the unexpected native speaker VOT production values, such as vowel context, general subject variability trends, and speech rate issues.

First, the present study did not control for vowel context following the target voiceless stop. As discussed in the literature review, Nearey and Rochet (1994) found that VOT values consistently varied across place of articulation of the target stop, as well as across vowel context (vowel following the target stop). Since the vowel contexts for the present /p/, /t/, and /k/ stimuli included front, mid and back vowels, it is possible that this could increase VOT production variability, as opposed to focusing solely on one vowel context (i.e., just /pa/, /ta/, and /ka/). However, because every speaker produced the same tokens, this potential parameter of variability (vowel context following the target stop) was controlled for across all speaker groups.

Second, high subject variability in VOT productions has been identified in other studies (Allen, Miller & DeSteno, 2003; Flege, 1987a; Nearey & Rochet, 1994). Flege (1987a) found that a single speaker can easily display a 5 – 15 ms standard deviation in the production of /t/ in a single phonetic context (based on an average VOT value of 80 ms). Nearey and Rochet (1994) found very high subject variability in their study of English and French VOT production values. In their study, these authors examined the utterances of ten Canadian English speakers, and eight native French speakers. They concluded that larger samples are needed in order to draw conclusions on population norms. As this study only examined eight speakers in total, with only two subjects representing each language proficiency level, the data sample is too small to make any population norm generalizations. It may also be too small for general comparison across subjects.

Third, Allen, Miller and DeSteno (2003) emphasize the importance of controlling for speech rate, as this parameter can greatly influence VOT value results. Furthermore, these authors indicate that even when speech rate is controlled for, VOT values can still vary dramatically between native speakers for the same stop tokens. While this study did not control for speech rate across the elicited stimuli, speech rate was measured in the form of average syllable length in the calculation of the Variability Index for the spontaneous-speech productions. In those calculations, beginners demonstrated the slowest speech rate, with a rate increase towards native speakers, who had the fastest speech rate. If one infers that speech rates would be comparable for the elicited data (although perhaps the gap between speaker groups lessens, as the language material provided decreases cognitive load), one would expect that VOT values still fall within the same hypothesized categories (with beginners displaying the longest VOT values, and native speakers producing the shortest VOT values). However, in interpreting the results obtained in this study, these expectations are not supported. Consequently, it is possible that speech rate is not a contributing factor to the inconsistent results obtained in the present study.

Another possible factor accounting for the unpredictable VOT productions of the native speakers could be their length of residence in a predominantly English-speaking environment. Although both subjects grew up in francophone regions as monolinguals, they have since become bilingual, and have been residing in British Columbia for most of their adult lives. Previous research has indicated that prolonged exposure to a second language can affect a speaker's native productions (Flege, 1987b). This may be the case for the native speakers used in the scope of this study. Perceptual evaluations of accentedness of these speakers' utterances by monolingual French speakers could either a) substantiate the presence of a potential pronunciation shift towards a more *bilingual* representation, or b) could in fact support the claim that these two speakers are representative of native, monolingual speakers of French.

5.1.2 F2 – F1 difference for /y/ and /u/

The results obtained for determining the existence of an adequate distinction between the acoustic productions of /u/ and /y/ support the hypotheses that learners progress towards more native-like production values as their language proficiency levels increase. These results support previous findings on French learners' ability (or inability) to pronounce the two segments appropriately (César-Lee, 1999; Rochet, 1995). The acoustic findings of this study also corroborate previous hypotheses that French learners actually produce /y/ more appropriately than /u/ (Flege, 1987b; Flege & Hillenbrand, 1984; Schweyer, 1996). Moreover, speaker production variability of the two segments also decreased as language proficiency increased. This finding would confirm conclusions drawn from other research, where researchers found that L2 learners' speech productions displaying greater acoustic variability are perceived as more accented (César-Lee, 1999; Jonasson & McAllister, 1972; Magen, 1998; Shah, 2002, Walz, 1980).

5.1.3 Diphthongization of /u/, /o/ and /i/

The diphthongization results obtained in the scope of this study are inconclusive about the nature or the presence of the potential diphthongization of /u/, /o/, and /i/ by Anglophone learners of French, and therefore do not confirm the hypotheses. Overall, these mean value results could suggest that either French learners are not diphthongizing their /u/, /o/, and /i/ vowels inappropriately in French, (thereby refuting H1), or the current methods used to quantify these acoustic phenomena are not adequately capturing the speech productions. While some formant shifting was evident for certain vowels and speakers (/u/: 1 beginner and 1 intermediate speaker; /o/: 1 speaker each for beginner, intermediate and advanced), no clear patterns emerged.

Though it is possible that learners are not inappropriately diphthongizing these vowels, impressionistic findings would suggest that the issue might profitably be treated from other perspectives, in order to determine its significance in the perception of foreign-accented speech (Le Clézio, 1986; 1989). The present methodology, which consisted of measuring only F2-shifting, may not have been sufficient to capture the multifaceted differences between diphthongized pronunciations which may be considered foreign-accented and native speaker vowel productions. In light of these assertions, it would be desirable to develop an adequate set of production and related perceptual studies, in order to effectively identify L2 French learners' potential diphthongization, and how these productions are recognized (on a foreign-accented/non-accented continuum) by native speakers.

5.1.4 F2 – F1 Acoustic Space for /o/ and /ɔ/

The results of this study suggest that an adequate /o/ – /ɔ/ distinction in French may emerge as language proficiency increases. These results also confirm that the two native French speakers under investigation are both still producing a stable distinction between the two segments, as represented by the F2 – F1 acoustic space, even though the status of this mid-vowel contrast in varieties of French from France has been put into question, meaning that there may no longer be a relevant contrast between these two segments in those varieties (Walker, 1984). While it is not the intention of the present study to resolve this issue, the results would suggest that the opposition is still relevant, as the native speakers produce the two segments discretely.

The study's results seem to support Le Clézio's (1986) impressionistic findings that learners produce /ɔ/ in a more target-like manner (and therefore potentially less foreign-accented manner) than /o/. However, Le Clézio's findings were only based on the investigation of French productions of two learner proficiency levels: intermediate and advanced[16]; while this study examined three learner proficiency levels: beginner, intermediate, and advanced. Therefore, while similar acoustic trends for the two segments are found in intermediate and advanced learner productions for both the present study and Le Clézio's (1986) study, the values obtained for beginner productions in this study do not seem to follow the same pattern. In fact, beginner productions of both /o/ and /ɔ/ display F2 – F1 acoustic spaces which contradict target productions (meaning that beginners are producing /o/ as [ɔ], and /ɔ/ as [o]). Although these findings are of some interest, the limited sample does not allow generalization to the larger Canadian English-speaking, beginner French-learner population.

5.1.5 Spectral Quality (F1) of /a/ in Unstressed Environments

The hypotheses that learners would produce more centralized /a/ in unstressed syllables was not supported by the results of the analysis, which indicate that while some F1 variation exists between speaker groups, all groups are still producing French unstressed /a/ with appropriate F1 values. Furthermore, standard deviation values are relatively consistent across speaker groups, thereby refuting H2.

Certain procedural methods may have encouraged the learners to produce less centralized pronunciations of the target vowel in this study than in natural speech. The production task (of delayed-repetition) may have elicited a more careful sentence production than would be the case in spontaneous speech. However, César-Lee (1999), using the same methodology, was able to isolate learner centralization of /a/. In order to

[16] Although Le Clézio never states explicitly the proficiency level of his subjects (who were university students learning French in England), the text the subjects were asked to read in order to participate in the study is relatively advanced, and therefore serves as an indicator of their level of language capacity in French.

measure conclusively actual speaker performance, the issue merits further consideration. Increasing the sample size, for example, with more target vowels, or more environments, may be sufficient to shed more light on this issue.

5.1.6 Temporal Vowel Reduction in Unstressed Environments for /u/, /y/, and /a/

This study also examined potential temporal variation between stressed and unstressed tokens of /u/, /y/, and /a/. Temporal reduction of /u/ did not pattern as predicted. In fact, beginners displayed the least amount of temporal reduction, followed closely by the advanced speaker group, with the intermediate and native speakers exhibiting temporal reduction of about 60%. The production of /y/ and /a/ tokens in stressed and unstressed environments does seem to reinforce the hypothesis that French learners temporally reduce their vowels excessively in unstressed environments. While beginner and intermediate speakers frequently reduce these vowels more than 50%, the advanced and the native speakers never do.

The inconsistent results across the three segments could be due to the data analysis methods, which may be inappropriately capturing production values. As speech rate was not controlled for in the elicitation task (task 1, which provided the production data sample), even small variations in speaking rate could have seriously influenced the temporal measurements. Therefore, controlling for speech rate could provide more consistent results. In addition, exploring other methods of quantifying vocalic reduction could also prove useful, for example, quantifying vocalic reduction through syllable length measurement, as César-Lee (1999) has done.

5.1.7 Rhythmic Variation using the Variability Index

The hypothesis predicting that French learners display greater rhythmic variation than native speakers seems to be substantiated by the results of this analysis. The results could also suggest a gradual trend towards a more native-like rhythm as language proficiency increases. This trend in rhythmic variability is mirrored by the speech rate results also obtained in the execution of the VI analysis. As learner proficiency increases, speech rate also increases towards more native-like rates. On the other hand, the hypothesis predicting that French learners would exhibit greater standard deviation than native speakers in their VI values is not substantiated. Overall, the Variability Index results obtained in the scope of this research support previous claims that French learners demonstrate greater inter-syllabic variability, or a higher rhythmic variability, than native speakers of French (Guilbault, 2002).

Of particular interest in the current study are the rhythmic VI values of intermediate speaker P22. First, the speaker displays extremely high VI values (0.9689), which is much higher than other recorded VI values for English speakers, whether in English (0.448 for British English, 0.543 for Singapore English: Deterding, 2001, p. 225; between 0.1817 and 0.6104 for Canadian English: Guilbault, 2002, p. 116), or in L2 French (0.4056 for intermediate learners, 0.4298 for advanced learners: Guilbault, 2002, p. 119). However, the speaker's standard deviation in her productions is extremely low (0.0507), demonstrating a stable VI measurement. In fact, the standard deviation values of P22's spontaneous productions are lower than any other recorded VI standard deviation values[17].

[17] Deterding (2001, p. 225): Native British English STDev: 0.164, Native Singapore English STDev: 0.172.

Further examination reveals that the spontaneous sentences used in the calculation of P22's VI included a scripted speech sentence, in which the researcher described a very simple scenario and suggested answers, and then the subject was able to use those answers as her own afterwards. This sort of 'spontaneous' elicitation could have lessened subject anxiety as well as the cognitive load of the speaking task, which would in turn have promoted more native-like L2 performance. However, it is clear that the scripted speech sentence did not have this effect on the speaker's rhythmic variation at all; the scripted speech sentence scored 0.9647 on the VI, which is very close to the speaker VI average. Therefore, it seems to be unlikely that the inclusion of this scripted speech sentence skewed the speaker's VI results.

In order to confirm the potential trends found in the VI values, perceptual evaluations of the subjects' spontaneous speech would need to be performed. These evaluations could then determine the level of accentedness attributed to each spontaneous utterance and allow for a more substantiated correlation between acoustic values and degree of foreign accent. While other factors could play a role in the foreign accent ratings (segmental cues), the perceptual tests could provide a general indicator. Similarly, resynthesis of the subjects' spontaneous sentences to preserve rhythmic properties while degrading segmental information could be used in order to determine the *nativeness*, or accentedness, of the sentences. Comparable methods have been used in previous studies of L1 perception of rhythm (Ramus & Mehler, 1999; Ramus, Nespor & Mehler, 1999).

Guilbault (2002, p. 119) Intermediate French learner STDev: 0.2486, Advanced French learner STDev: 0.2399, Native Canadian French STDev: 0.2042, Native European French STDev: 0.1547.

5.2 General Discussion and Directions for Future Research

The main objectives of this study were to identify and quantify some of the more salient segmental and suprasegmental acoustic cues of Canadian English-accented French in female speech, as well as to establish a link between spontaneous and elicited speech materials in the productions of native and non-native French speakers. When considered globally, the results of this study could suggest that there are in fact acoustic measurements of certain segmental and suprasegmental cues that seem to vary consistently with language proficiency, between native and non-native speakers of French, which in turn may contribute to the perception of a Canadian-English accent in French. Specifically, the research questions will be revisited, in order to provide a more integrated interpretation of the findings.

How does non-native female speech compare acoustically to native female speech?

Based on the results of the two experiments conducted in the scope of this study, it is possible to conclude that there are in fact some measurable acoustic differences between speech productions of non-native and native female speakers of French. In particular, segmental production differences between native and non-native speakers were most apparent for the vowel sounds /u/ and /y/, and /o/ and /ɔ/. The measurement of these segments revealed non-target-like acoustic realizations of these segments, most notably by the beginner and the intermediate speakers. Furthermore, VI results based on spontaneous-speech samples also suggest the existence of measurable differences between native and non-native French productions. In fact, the native and non-native speaker values obtained in the measurement of spontaneous-speech samples patterned closely with these native and non-native speaker values obtained in the measurement of elicited speech productions. This similar patterning could indicate that speaker performance on spontaneous and on elicited speech material is comparable.

On the other hand, some of the segmental and suprasegmental cues under investigation revealed no consistent patterning of native and non-native speech productions. Language proficiency levels did not seem to impact acoustic performance in the production of diphthongized vowels /u/, /y/, and /a/, unstressed /a/, and temporal vowel reduction of /u/, /y/, and /a/. Furthermore, while VOT values patterned along language proficiency levels for the non-native speakers, native speakers did not perform as expected.

Does acoustic variation in segmental and suprasegmental production readily distinguish native and non-native speech samples? While some evidence, such as the production values of /u/ and /y/, suggests that variation in segmental and suprasegmental production distinguishes native and non-native speech samples, most of the cues under investigation did not yield distinct values according to language proficiency level. These findings would indicate that, at this time, it is not possible to distinguish the acoustic variation patterns of Canadian English learners of French and native speakers of French.

The present research has contributed to the existing literature on foreign-accented French in general, and more particularly to the study of Canadian English-accented French. This study has provided acoustic quantification for some of the more relevant acoustic cues associated with the perception of an English accent in French. Furthermore, it has established a preliminary link between spontaneous and delayed-repetition speech productions of non-native female speakers of French, along varying language proficiency levels.

This research could also prove useful in helping establish stronger ties between linguistic analysis of learner speech and French language instruction. This study has provided a broad range of production values for female beginner-level English learners of French, which had not been documented as extensively before. Furthermore, these values are contrasted with more advanced and native speakers, providing a point of comparison for the beginner level of speech performance. In addition, the non-native subjects of this study were only exposed to French in the language classroom, meaning that their pronunciation performance was most likely influenced predominantly by instructional methods.

While the acoustic measurement results obtained in the scope of this study provide an initial step towards the quantification of a Canadian-English accent in French, perceptual correlates to the production data would further substantiate the inferences made on the basis of these findings. Such evaluations could determine the level of accentedness attributed to each acoustic representation, and allow for a substantiated correlation between acoustic values and degree of foreign accent. In other words, perceptual correlates could corroborate the saliency of these cues for foreign accent perception. As put forth by Munro (1993), and discussed by Wayland (1997) "detailed acoustic analysis may reveal differences between native and non-native speech on various parameters. However, these differences may influence native speakers' perception of accentedness to a varying degree. Thus, acoustic data should be related to perceptual data." (Wayland, 1997, p. 247)

In conclusion, acoustic measurements of certain segmental and suprasegmental cues seem to vary consistently with language proficiency, which in turn may contribute to the perception of a Canadian-English accent in French. Canadian English-speaking learners of French and native speakers of French also seemed to perform consistently across the spontaneous and elicited samples. Future research could provide perceptual correlates to the physical-acoustic dimensions of Canadian English-accented speech presented in this study.

APPENDICES

Appendix A: Final List of Target Words

Target words in carrier sentences according to acoustic cue

1. VOT of syllable-initial voiceless stops /p/, /t/, and /k/

   i. /p/ in stressed position

   Le mot ***poux*** est très joli.
   Le nom ***tapis*** semble très joli.
   L'adverbe ***pas*** n'est pas joli.
   Le nom ***empire*** est très beau.
   Le mot ***poupée*** est très joli.

   ii. /t/ in stressed position

   L'adjectif ***tout*** semble très joli.
   Le nom ***couteau*** est très joli.
   Le pronom ***tu*** semble très beau.
   Le nom ***pâté*** est très joli.
   Le mot ***bateau*** est très joli.

   iii. /k/ in stressed position

   Le mot ***avocat*** semble très joli.
   Le nom ***cou*** semble très beau.
   Le nom ***coucou*** est très beau.
   Le mot ***cœur*** est très beau.
   Le mot ***court*** est très joli.

2. Production of /u/ and /y/

   i. /u/ in stressed position

   Le mot ***toutou*** est très joli
   L'adjectif ***tout*** semble très joli.
   Le nom ***coucou*** est très beau.
   L'adverbe ***partout*** est très beau.
   L'adjectif ***doux*** semble très beau.

   ii. /y/ in stressed position

   Le mot ***pu*** est très laid.
   L'adjectif ***aigu*** est très joli.
   Le pronom ***tu*** semble très beau.
   La préposition ***du*** est très jolie.

3. Diphthongization of /u/, /o/ and /i/ in stressed position

/u/ Le mot ***toutou*** est très joli.
L'adjectif ***tout*** semble très joli.
Le nom ***coucou*** est très beau.
L'adverbe ***partout*** est très beau.
L'adjectif ***doux*** semble très beau.

/o/ Le mot ***chapeau*** est très beau.
Le mot ***bateau*** est très joli.
Le mot ***boulot*** semble très beau.
L'adjectif ***beau*** semble très beau.
Le nom ***couteau*** est très joli.

/i/ Le mot ***lit*** semble très joli.
Le mot ***bougie*** est très joli.
Le mot ***ami*** semble très beau.
Le nom ***tapis*** semble très joli.
Le nom ***partie*** est très joli.

4. Production of /ɔ/ and /o/

/ɔ/ Le mot ***bol*** est très beau.
Le nom ***botte*** est très joli.
Le mot ***porte*** est très joli.
L'adjectif ***bonne*** est très joli.
Le mot ***donne*** est très joli.

/o/ Le mot ***chapeau*** est très beau.
Le mot ***bateau*** est très joli.
Le mot ***boulot*** semble très beau.
L'adjectif ***beau*** semble très beau.
Le nom ***couteau*** est très joli.

5. Spectral centralization in unstressed environments for /a/.

Le nom ***tapis*** semble très joli.
Le nom ***patate*** est très joli.
Le mot ***café*** semble très joli.
Le mot ***bateau*** est très joli.
Faire ***attention*** c'est très bon.

6. Temporal vowel reduction in unstressed environments for /u/, /y/, and /a/.

   i. /u/ in stressed position

      Le mot ***toutou*** est très joli
      L'adjectif ***tout*** semble très joli.
      Le nom ***coucou*** est très beau.
      L'adverbe ***partout*** est très beau.
      L'adjectif ***doux*** semble très beau.

   ii. /u/ in unstressed position

      Le mot ***toutou*** est très joli.
      Le nom ***couteau*** est très joli.
      Le mot ***poupée*** est très joli.
      Le nom ***coucou*** est très beau.
      Le mot ***boulot*** semble très beau.

   iii. /y/ in stressed position

      Le mot ***pu*** est très laid.
      L'adjectif ***aigu*** est très joli.
      Le pronom ***tu*** semble très beau.
      La préposition ***du*** est très jolie.

   iv. /y/ in unstressed position

      Le pronom ***duquel*** est très joli.
      Le mot ***tulipe*** est très beau.
      Le mot ***dupé*** semble très joli.
      Le mot ***publique*** est très joli.

   v. /a/ in stressed position

      Le mot ***avocat*** est très joli.
      L'adverbe ***pas*** n'est pas joli.
      Le nom ***patate*** est très joli.
      Le nom ***patte*** est très joli.

   vi. /a/ in unstressed position

      Le nom ***tapis*** semble très joli.
      Le nom ***patate*** est très joli.
      Le mot ***café*** semble très joli.
      Le mot ***bateau*** est très joli.
      Faire ***attention*** c'est très bon.

Appendix B: Language Background Questionnaire (Non-native French Speakers)

Name

Tel/e-mail

Gender

Age

1. What is your first language?
2. What are your parents' first languages?
   Mother:
   Father:
3. Do you speak any languages other than English and French? If so, please indicate which one(s), in what capacity you use them, and the age at which you started learning them.
4. How old were you when you started studying/learning French?
5. How long have you been studying French?
6. Where have you studied French? (location and educational level)
7. Which French classes have you been or are currently enrolled in (at the High School and/or University level)?
8. Have you ever been enrolled in French immersion? If so, where and for how long?
9. Do you use French outside of classes? If so, under what circumstances and how often?
10. Have you spent time in a francophone region? If so, where and for how long? Did you speak French while you were there?
11. Please indicate, to the best of your knowledge, where your French instructors came from and which classes (or levels) of French they taught you.
12. To the best of your knowledge, are you aware of any hearing problem you may have? If so, which one?
13. If you are a student, which program are you enrolled in at Simon Fraser University? If you are working, what is your job title?

Appendix C: Language Background Questionnaire (Native French Speakers)

Nom

Tél/courriel

Sex

Âge

1. Avez-vous le français comme langue maternelle?
2. Quelle est la langue maternelle de vos parents?
   Mère:
   Père:
3. Quelle langue utilisez-vous majoritairement à la maison?
4. Avez-vous grandi dans un milieu bilingue? Si oui, specifiez où et quelles langues vous parliez à la maison.
5. Parlez-vous une (ou plusieurs) autre(s) langue(s) que l'anglais et le français?
6. Si oui, laquelle (lesquelles), à quelle fréquence et dans quelles circonstances?
7. Depuis combien de temps estimez-vous être dans un milieu majoritairement anglophone?
8. Est-ce que la majorité de vos activités journalières (études, travail) se déroulent en français? Si non, dans quelle langue?
9. Savez-vous si vous avez un quelconque problème d'audition? Si oui, lequel?

Appendix D: Demographic Information

Table 6.1 Demographic Information

Subj.	Age	AOL	LOL	Group	Immersion	Other languages (AOL)	Time in Francophone Region
P11	27	10	8	Beg.	No	None	None
P13	25	13	5	Beg.	No	None	None
P22	19	11	8	Int.	No	None	None
P24	20	5	13	Int.	13 yrs (K-12)	Int. Portuguese (12+)	None
P32	22	8	9	Adv.	No	Some Spanish (19+)	5 wk Bursary Prog. (QU 2002)
P37	20	13	7	Adv.	No	Tagalog (Barely)	None
C11	40+	N/A	N/A	Nat.	N/A	Some Spanish (teens)	N/A
C12	40+	N/A	N/A	Nat.	N/A	None	N/A

AOL	Age of Acquisition
LOL	Length of Learning (in years)
Immersion	Language program offered in BC public schools with subject instruction in French

Appendix E: Question List for Spontaneous-speech Elicitation (Non-native French Speakers)

Beginner Students

Qu'est-ce que tu aimes faire le weekend?
Qu'est-ce que tu aimes porter?
Aimes-tu voyager? Où as-tu voyager?
Quel cours prends-tu?
Pourquoi tu étudies le français?

Est-ce que tu as des frères ou des sœurs? Peux-tu décrire ce membre de ta famille?
Quel temps fait-il dehors?

Intermediate and Advanced Students

Quel cours prends-tu?
Pourquoi tu étudies le français?
Est-ce que tu aimes étudier à SFU? Pourquoi?
Qu'est ce que tu veux faire quand tu finis à SFU?
Penses-tu pouvoir trouver un emploi dans ton domaine?
Que penses-tu des relations entre les étudiants et les professeurs à l'université?

Qu'est-ce que tu aimes faire le weekend?
Qu'est-ce que tu fais dans tes temps-libre?
As-tu des passions?
Raconte-moi l'histoire du meilleur film (du plus beau spectacle/du meilleur livre) que tu as vu.

Aimes-tu voyager? Où as-tu voyager?
Raconte-moi ton plus beau voyage.

Pourrais-tu décrire un peu ta famille?

Tu gagnes 10 millions de dollars demain. Qu'est ce que tu fais?

Appendix F: Question list for Spontaneous-speech Elicitation (Native French Speakers)

Language Instructors

Aimez-vous enseigner le français à SFU?
Quelle importance donnez-vous à l'enseignement de la prononciation dans vos cours de langue?

Aimez-vous voyager?
Racontez-moi un de vos voyages préférés.

Racontez-moi l'histoire du meilleur film (du plus beau spectacle/du meilleur livre) que vous avez vu.

Appendix G: Sample Dialogue for Spontaneous-speech Elicitation (used for the beginners and one intermediate subject)

Q. *Bonjour. Qu'est-ce que tu fais aujourd'hui?*

R Je vais à la bibliothèque maintenant.

Q. *Pourquoi?*

R. Parce que j'ai besoin d'un livre de mathématiques.

Q. *Est-ce que tu as des cours de français aujourd'hui?*

R. Non, mes cours de français sont le mardi et le jeudi.

Appendix H: Nearey and Rochet's (1994) English and French VOT Findings

Table 6.2 Nearey & Rochet's (1994) Mean VOT Values (in ms) for English Voiceless Stops, /p/, /t/, and /k/ (for 10 subjects)

Vowel	/p/		/t/		/k/	
/i/	71.9	(15.7)	85.1	(17.6)	86.8	(11.5)
/e/	65.2	(13.9)	67.9	(12.5)	80.5	(11.0)
/o/	59.8	(12.9)	74.7	(19.7)	79.5	(10.7)
/u/	73.2	(8.3)	74.0	(19.8)	80.4	(13.2)

Table 6.3 Nearey & Rochet's (1994) Mean VOT Values (in ms) for French Voiceless Stops, /p/, /t/, and /k/ (for 8 subjects)

Vowel	/p/		/t/		/k/	
/i/	24.0	(12.7)	48.0	(10.0)	59.4	(8.0)
/a/	17.9	(9.2)	33.7	(25.1)	32.2	(7.4)
/ɔ/	22.2	(7.1)	20.8	(4.7)	40.1	(9.1)
/o/	33.4	(12.2)	28.9	(9.2)	51.1	(9.9)
/u/	40.4	(15.3)	42.9	(19.9)	62.2	(13.4)
/y/	47.9	(19.5)	49.3	(10.7)	51.2	(8.3)
/œ/	25.2	(16.2)	24.0	(6.3)	36.2	(6.3)

Appendix I: Martin's (2004) Female French Canadian Formant Values

Table 6.4 Martin's (2004) Female French Canadian Formant Values

	F1	*F2*	*F3*	F4
i	360	2680	3670	4280
e	420	2470	3180	4580
ɛ	700	2280	3090	4630
a	950	1570	3150	4370
u	400	800	2900	4030
o	420	850	3040	4160
ɔ	650	1140	3150	4090
ɑ	800	1370	3040	4200
y	350	2200	2480	3840
ø	420	1710	2700	4130
œ	530	1630	2750	4070
ə	630	1710	3060	4180

REFERENCE LIST

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America (JASA), 113, 1,* 544 – 552.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42, 4,* 529 – 555.

Archibald, J., & Libben, G. (1995). The effects of age. *Research Perspectives on Second Language Acquisition*, Ontario: Copp Clark Ltd, 288 – 309.

Asher, J. J., & Garcia, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal, 53,* 334 – 341.

Beaudoin, M. (1996). *The Syllable Structure of French as Perceived by First and Second-Language Speakers*. Unpublished doctoral dissertation, University of Alberta, Alberta.

Best, C.T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman, & H. C. Nusbaum (Eds.), *The Development in Speech Perception: The Transition from Speech Sounds to Spoken Words* (pp. 167 – 224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 171 – 205). Timonium, MD: York Press.

Brennan, E. M., & Brennan, J. S. (1981). Measurements of accent and attitude toward Mexican-American speech. *Journal of Psycholinguistic Research, 10, 5,* 487 – 501.

Brière, E. J. (1966). An investigation of phonological interference. *Language, 42,* 768 – 796.

Brown, C., & Cichocki, W. (1995). Listeners' reactions to four French accents: A study of gender as a variable in linguistic attitudes. *Linguistica Atlantica, 17,* 45 – 62.

Burgess, C. S. (2001). *Speaking Rate, Fluency, and Accentedness in the Speech of Second Language Learners*. Unpublished doctoral dissertation, Simon Fraser University, British Columbia.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America (JASA), 54, 2,* 421 – 428.

Casagrande, J. (1984). *The Sound System of French.* Washington, DC: Georgetown University Press.

César-Lee, B. V. (1999). *Quantification of Accented Pronunciation by American-English Speakers in French-as-a-Foreign-Language Setting.* Unpublished doctoral dissertation, University of Florida, Florida.

Champagne-Muzar, C., Schneiderman, E. I., & Bourdages, J. S. (1993). Second language accent: The role of the pedagogical environment. *International Review of Applied Linguistics (IRAL), 31, 2,* 143 – 160.

Cichocki, W., House, A. B., & Lister, A. C. (1997). Cantonese speakers and the acquisition of French nasal vowels. *Revue Parole, 1,* 1 – 15.

Crystal, D. (2003). *A Dictionary of Linguistics & Phonetics.* Malden, MA: Blackwell Publishing.

Dansereau, D. (1990). *Savoir Dire: Cours de phonétique et de prononciation.* Lexington, MA: D. C. Heath and Company.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11, 1,* 51 – 62.

Delattre, P. (1940). Tendances de coupe syllabique en français. *Publications of the Modern Language Association, 55, 2,* 579 – 595.

Delattre, P. (1981). *Studies in Comparative Phonetics: English, German, Spanish and French.* Heidelberg: Groos.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition (SSLA), 20,* 1 – 16.

Deterding, D. (2001). Letter to the Editor, The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics, 29,* 217 – 230.

Dowd, A., Smith, J., & Wolfe, J. (1997). Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time. *Language and Speech, 41, 1,* 1 – 20.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27, 2,* 315 – 330.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America (JASA), 76, 3,* 692 – 707.

Flege, J. E. (1987a). The instrumental study of L2 speech production: Some methodological considerations. *Language Learning, 37, 2,* 285 – 296.

Flege, J. E. (1987b). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15, 1,* 47 – 65.

Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 232 – 277). Timonium, MD: York Press.

Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America (JASA), 91, 1,* 370 – 389.

Flege, J. E., & Hillenbrand, J. (1984). Limits of phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America (JASA), 76, 3,* 708 – 721.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America (JASA), 97, 5,* 3125 – 3134.

Flege, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final English /t/ – /d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America (JASA), 92, 1,* 128 – 143.

Francis, A. L., Ciocca, V., & Yu, J. M. C. (2003). Accuracy and variability of acoustic measures of voicing onset. *Journal of the Acoustical Society of America (JASA), 113, 2,* 1025 – 1032.

Garant Viau, C. (1994). *La portée des sons.* (Les cahiers de français, N° 12). Québec: Laval University.

Garnica, O. K., & Herbert R. K. (1979). Some phonological errors in second language learning: Interference doesn't tell it all. *International Journal of Psycholinguistics, 6,* 5 – 19.

Gottfried, T., & Beddor P. S. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech, 31, 1,* 57 – 75.

Greasley, V. (1989). French Pronunciation Difficult? *Modern Languages, 70, 1,* 44 – 48.

Green, J. R., & Poulin N. A. (1971). *French Phonology for Teachers: A Programmed Introduction.* Philadelphia: Center for Curriculum Development.

Grover, C., Jamieson, D. G., & Dobrovolsky, M. B. (1987). Intonation in English, French and German: Perception and production. *Language and Speech, 30, 3,* 277 – 295.

Guilbault, C. (2002). *The Acquisition of French Rhythm by English Second Language Learners.* Unpublished doctoral dissertation, University of Alberta, Alberta.

Jackson, K. H. (1971). The verification and comparing of contrastive analysis. PCCLLU Papers, *Working Papers in Linguistics, 3,* 201 – 209.

Jonasson, J., & McAllister, R. H. (1972). Foreign accent and timing: An instrumental phonetic study. *Papers from the Institute of Linguistics University of Stockholm (PILUS), 14,* 11 – 40.

Le Clézio, Y. (1986). Possible factors of the English accent in French. *Speech, Hearing and Language, 2,* 237 – 253.

Le Clézio, Y. (1989). The English accent in French – further considerations. *Speech, Hearing and Language, 3,* 63 – 72.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition (SSLA), 12,* 251 – 285.

Low, E. L. (1998). *Prosodic Prominence in Singapore English.* Unpublished doctoral dissertation, Cambridge University.

Mack, M. (1982). Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *Journal of the Acoustical Society of America (JASA), 71, 1,* 173 – 178.

Magen, H. S. (1998). The perception of foreign-accented speech, *Journal of Phonetics, 26,* 381 – 400.

Markham, D. (1997). *Phonetic Imitation, Accent and the Learner*. (Travaux de l'institut de linguistique de Lund, Vol. 33). Lund: Lund University Press.

Martin, P. (1998). Les voyelles d'aperture moyenne en français du Québec. *Cahiers de l'ILSL, 11,* 215 – 242.

Martin, P. (2002). The vocal system in Quebec French: From acoustics to phonology. *La Linguistique, 38, 2,* 71 – 88.

Martin, P. (2004). *Identification des sons du français.* Retrieved January 25, 2006, from http://www.lli.ulaval.ca/labo2256/illust.html.

Miller, M. (1984). On the perception of rhythm. *Journal of Phonetics, 12,* 75 – 83.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition (SSLA), 21,* 81 – 108.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech, 36, 1,* 39 – 66.

Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition (SSLA), 20,* 139 – 154.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45, 1,* 73 – 97.

Munro, M. J., & Derwing, T. M. (2001). Modelling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition (SSLA), 23,* 451 – 468.

Nearey, T. M., & Rochet, B. L. (1994). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association, 24, 1,* 1 – 19.

Neufeld, G. (2001). Non-foreign-accented speech in adult second language learners: Does it exist and what does it signify? *ITL, Review of Applied Linguistics, 133 – 134,* 185 – 206.

O'Grady, W. D., & Archibald, J. (2000). *Contemporary Linguistic Analysis, An Introduction*, 4th Ed. Toronto; Addison Wesley Longman.

Ostiguy, L., & Tousignant, C. (1993). *Le français québécois: normes et usage*. Montréal: Guérin Universitaire.

Pike, K. L. (1945). *The Intonation of American English.* Ann Arbor, Michigan: University of Michigan Publications.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29,* 191 – 215.

Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning, 30,* 271 – 287.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America (JASA), 105, 1,* 512 – 521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73,* 265 – 292.

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 379 – 410). Timonium, MD: York Press.

Santerre, L., Dufour, S-P., & McDuff, S. (1985). La perception de la diphtongaison: son importance dans les grands corpus. *Revue québécoise de linguistique, 4, 4(1),* 33 – 53.

Schouten, M. E. H., & Peeters, W. J. M. (2000). Searching for an explanation for diphthong perception: Dynamic tones and dynamic spectral profiles. *Phonetica, 57,* 17 – 39.

Schweyer, D. (1996). The effect of consonantal environment on French listeners' perception of English [u]. *Canadian Acoustics, 24, (3),* 37.

Shah, A. P. (2002). *Temporal Characteristics of Spanish-Accented English: Acoustic Measures and Their Correlation with Accentedness Ratings*, Unpublished doctoral dissertation, City University of New York (CUNY), New York.

Shriberg, L. D. & Kent, R. D. (2003). *Clinical Phonetics*, 3rd Edition. Boston, MA: Allyn and Bacon.

Southwood, M. H., & Flege, J. E. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics & Phonetics, 13, 5,* 335 – 349.

Tahta, S., Wood, M., & Loewenthal, K. (1981). Foreign accents: factors relating to transfer of accent form the first language to a second language. *Language and Speech, 24, 3,* 265 – 272.

Thompson, I. (1991). Foreign Accents Revisited: The English pronunciation of Russian immigrants. *Language Learning, 41, 2,* 177 – 204.

Tranel, B. (1987). *The Sounds of French: An Introduction.* Cambridge: Cambridge University Press.

Vaissière, J. (1991). Rhythm, accentuation and final lengthening in French. In J. Sundberg, J. Nord & R. Carlson (Eds.), *Music, Language, Speech and Brain* (pp. 108 – 120).

Valdman, A. (1976). *Introduction to French Phonology and Morphology*. Rowley, MA: Newbury House.

Valdman, A. (1993). *Bien entendu! Introduction à la prononciation française*. New Jersey: Prentice Hall.

Walker, D. C. (1984). *The Pronunciation of Canadian French*. Ottawa: University of Ottawa Press.

Walker, D. C. (2001). *French Sound Structure*. Calgary: University of Calgary Press.

Walz, J. (1979). *The Early Acquisition of a Second Language Phonology*. Hamburg: Buske.

Walz, J. (1980). An empirical study of pronunciation errors in French. *French Review, 53, 3,* 424 – 432.

Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly, 4,* 123 – 130.

Weinreich, U. (1953). *Languages in Contact*. The Hague: Mouton.

Wenk, B. J., & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics, 10,* 193 – 216.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition, 3, 2,* 131 – 149.

Zybert, J. (1997). Acquisition of L2 phonetic features. *Papers and Studies in Contrastive Linguistics, 32,* 103 – 117.