Native Dialect Effects in Non-native Production and Perception of Vowels

by

Irina Marinescu

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Spanish and Portuguese
University of Toronto

© Copyright by Irina Marinescu 2012
Native Dialect Effects in Non-native Production and Perception

Irina Marinescu

Doctor of Philosophy

Department of Spanish and Portuguese

University of Toronto

2012

Abstract

This dissertation examines the role of the native dialect in non-native perception and production in the specific case of Cuban and Peninsular Spanish as native varieties and of English vowels /Q, √, A/ as the target. In most second language studies, the learners’ native variety is assumed to be homogenous, regardless of their regional variety. Nevertheless, regional varieties differ in non-trivial ways and such differences need to be considered when analyzing second language acquisition. This dissertation takes shape around the main research question of whether potentially systematic differences between vowels in the native dialects of Cuban and Peninsular Spanish would produce matching phonetic differences in non-native perception and production of English vowels. This question was addressed in three experiments that compared native vowels in these Spanish varieties, as well as the perception and production of the target English /Q, √, A/ by Cuban and Peninsular Spanish learners.

Significant cross-dialectal differences were identified in the production of native vowels, namely, locations of /i, o/, durations of /e, a, o/ as well as different variability patterns, which were predicted to influence mapping of sounds in L2 perception and production. In L2 perception, discrimination for advanced and naïve listeners from each dialect was tested with English /Q - √,/.
/\ - A/, and /Q- A/. No clear native dialect effect could be identified; however, Cuban advanced
listeners obtained high error rates with /\ - A/. In L2 production, though, dialect-specific patterns
were found: Peninsular learners produced /Q, \, A/ with significant spectral but no durational
differences whereas Cuban learners produced /\, A/ with no significant spectral but with duration
differences.

I concluded that the native dialect was one of several factors generating the differences between
Peninsular and Cuban participants’ perception and production of English /Q, \, A/. The conjoint
effect of the native dialect, input and learning experience were shown to have contributed to the
distinctions. This finding contributes to second language acquisition research because it stresses
the need to control for learners’ native dialect. This research contributed new acoustic data on
Cuban Spanish and on L2 English. It uncovered specific patterns and interlanguage strategies of
Spanish learners of English.
Acknowledgments

At the long-awaited and much-anticipated moment of finishing this dissertation, my best thoughts and deepest gratitude go to the people who have helped me along the intricate journey of my doctoral studies. I am more than grateful to Professor Laura Colantoni who had advised and guided me in my academic endeavours long before she became my thesis supervisor. With her readiness to discuss big questions and small details, theoretical and practical; her amazing ability to make complex things look clear; her warmth and humour and, above all, extensive knowledge and critical sense, she has been the driving force and a role model in my quest to become a scholar. I thank Laura for trusting, encouraging and helping me to carry on when I had found myself at an impasse and for training and mentoring me throughout the years.

I am also very grateful to my committee members Professors Jeffrey Steele and Ana Teresa Pérez-Leroux who teamed up to lead me effectively through the process of producing my dissertation. I wish to thank Jeff for his sweeping generosity and enthusiasm in discussing, listening and offering constructive feedback at all the stages of my research. Our discussions would always offer food for thought and be sources of positive energy for weeks on end. Jeff has urged me to approach my work with an objective eye, always to keep in mind the big picture and challenged me to find alternative solutions and explanations, thus teaching me to become a better editor and critical thinker. I express my warmest and most sincere thank you to Ana Teresa for being a constant source of encouragement and support in the program. Her insightful comments, precision, and ability to pursue and elicit the essence have helped me to keep on track and sift through unnecessary details. Her metaphors and practical advice about academic pursuits are landmarks in my training. I consider myself very fortunate to have benefited from such a remarkable team of mentors and scholars.

I would like to thank Professor María Cristina Cuervo for being a member of my oral examination committee and for helping and training me as an instructor throughout the years. I am grateful to Professor Jorge Guitart for accepting to act as an external appraiser of my dissertation; it has been a great honour. Many other scholars have volunteered their time and expertise to help me with various aspects of my dissertation. I thank Professors Alexei Kochetov and Chandan Narayan for their insight on perceptual testing protocols, Professor Juli Cebrian for kindly accepting to discuss with me in detail various aspects of my dissertation, Professor Michol
Hoffman for sharing her expertise on Canadian English and Professor Steven Rupp for his support in the program. I thank Dr. Monique Herbert for her interest in my research and help with the statistics software. I greatly appreciate the generous help of Professor Noelia Ramón García and Dr. Ana Ferrando Carretero in recruiting participants at the University of León, Spain and providing a space for testing.

I also wish to express my warmest thanks to all the members of the Department of Spanish and Portuguese, faculty, staff, and graduate students. In particular, I owe thanks to my colleagues, Yasaman Rafat, Natalia Mazzaro, Anna Limanni, Danielle Thomas, Yadira Álvarez, and Tanya Battersby for making my graduate experience a better, funnier, more meaningful, and more tolerable experience.

Last but not least, my family deserves the most heartfelt thank you for supporting me emotionally, financially and in many other ways throughout the years, for trusting and encouraging me, for making me laugh and uprooting me from my desk in order to enjoy life beyond the thesis. I thank my husband Florin for his readiness to solve any of my technical problems, from software misbehaviour to computer and recording device updates, for putting his programming skills at work to write the perceptual testing software I used here and for lending his ear to my hypotheses and discussions. Dedicating this dissertation to him comes as a natural and legitimate gesture of thank you.
# Table of Contents

Acknowledgments ..................................................................................................................... iv

Table of Contents ...................................................................................................................... vi

Chapter 1 Introduction ................................................................................................................ 1

1 The starting point ................................................................................................................... 1

2 Motivation for the linguistic variables selected ................................................................. 2

3 Theoretical support ................................................................................................................ 4

3.1 Phonetic theories .......................................................................................................... 4

3.2 Theoretical models of SLA ............................................................................................ 5

4 Dissertation outline ............................................................................................................. 6

Chapter 2 Methodology .............................................................................................................. 8

1 Introduction ....................................................................................................................... 8

2 Experimental design ......................................................................................................... 9

2.1 Experiment on L2 production ....................................................................................... 10

2.1.1 Motivation for task selection ............................................................................. 10

2.1.2 L2 production Experiment. Oral sentence-completion task ................................ 12

2.2 Experiment on L2 perception ..................................................................................... 16

2.2.1 Motivation for the task selection ........................................................................ 16

2.2.2 L2 perception – AX discrimination task ............................................................ 18

2.3 Experiment on L1 production ..................................................................................... 21

2.3.1 Motivation for the task selection ........................................................................ 21

2.3.2 L1 production – Reading task ............................................................................ 22

3 Summary ............................................................................................................................. 23

Chapter 3 L1 production ........................................................................................................... 25

1 Introduction ....................................................................................................................... 25
2 Phonetic theories and cross-linguistic comparisons .............................................................. 25
  2.1 Phonetic theories of speech production ................................................................. 26
      2.1.1 The Theory of Adaptive Dispersion .............................................................. 26
      2.1.2 The Quantal Theory ...................................................................................... 28
      2.1.3 The Articulatory Settings or the Base of Articulation Theory ....................... 28
  2.2 Acoustic measurement of vowel formants and within-category variability .......... 30
      2.2.1 Vowel location ............................................................................................. 31
      2.2.2 Within-category variability ......................................................................... 32

3 Previous research on dialect differences ............................................................................... 33
  3.1 Cross-dialectal vowel differences in languages other than Spanish ..................... 33
  3.2 Cross-dialectal vowel differences in Spanish varieties ........................................... 35

4 Normalization Procedures .................................................................................................. 38

5 The current study ............................................................................................................... 40
  5.1 Hypotheses and predictions ...................................................................................... 40
  5.2 Vowel locations ......................................................................................................... 42
      5.2.1 Qualitative results for vowel locations ......................................................... 42
      5.2.2 Data normalization ....................................................................................... 43
      5.2.3 Quantitative results for vowel locations ....................................................... 44
  5.3 Dispersion and intervocalic distances .................................................................... 45
      5.3.1 Qualitative results for dispersion and intervocalic distances ......................... 45
      5.3.2 Quantitative results for intervocalic distances and dispersion ..................... 47
  5.4 Variability ................................................................................................................ 48
      5.4.1 Cross-dialectal patterns of variability ............................................................ 48
      5.4.2 Within-category variability .......................................................................... 50
  5.5 Vowel duration – qualitative and quantitative analysis ........................................... 52

6 Hypothesis evaluation and discussion ............................................................................... 56
7 Conclusions ......................................................................................................................... 58
Chapter 4 L2 perception ........................................................................................................ 59
1 Introduction ......................................................................................................................... 59
2 General issues in vowel perception .................................................................................. 60
3 Low and back vowels in native and non-native perception ............................................. 62
   3.1 Low and back vowels in identification tasks .............................................................. 62
   3.2 Contrasts with low and back vowels in discrimination tasks ...................................... 64
   3.3 Vowel inventory mismatches and experience in L2 perception ................................. 65
4 Vowel perception in cross-variety studies ...................................................................... 67
   4.1 Cross-linguistic and L2 vowel perception ................................................................. 67
   4.2 Cross-dialectal vowel perception ............................................................................. 68
   4.3 Native dialect effects in non-native perception ......................................................... 69
5 The current study ............................................................................................................. 71
   5.1 Hypotheses ............................................................................................................... 71
   5.2 Results ...................................................................................................................... 73
      5.2.1 Contrast difficulty ............................................................................................ 74
      5.2.2 Experience with the L2 .................................................................................. 75
      5.2.3 L1 dialect effect in L2 perception .................................................................... 76
   5.3 Hypothesis evaluation and discussion ..................................................................... 78
   5.4 Conclusions ............................................................................................................. 80
Chapter 5 L2 production ..................................................................................................... 82
1 Introduction ....................................................................................................................... 82
2 Theoretical approaches to SLA: three models ............................................................... 83
   2.1 The Speech Learning Model ................................................................................... 84
   2.2 The formal linguistic model of categorization ......................................................... 85
   2.3 The Perceptual Assimilation Model for L2 .............................................................. 86
3 Methods for assessing L2 production ................................................................. 88
   3.1 Acoustic comparisons .................................................................................. 88
   3.2 Accentedness and intelligibility judgments ................................................. 90
4 Previous studies in second dialect production ............................................... 91
5 Canadian English as spoken in Toronto .......................................................... 94
6 The current study ............................................................................................... 96
   6.1 Hypotheses and predictions ....................................................................... 97
   6.2 Results ......................................................................................................... 99
      6.2.1 Individual variability ......................................................................... 105
   6.3 Hypothesis evaluation and discussion ....................................................... 106
7 Conclusions ........................................................................................................ 109
Chapter 6 Discussion .......................................................................................... 111
1 Introduction ....................................................................................................... 111
2 Vowel comparisons in Peninsular and Cuban Spanish .................................... 111
   2.1 How real are the cross-dialectal differences? .............................................. 112
   2.2 Some findings in the light of different phonetic theories ......................... 114
      2.2.1 TAD – dispersion and speech rate .................................................... 114
      2.2.2 QT and variability of point vowels ................................................. 115
      2.2.3 AS/BA theory and cross-dialectal preferences ............................... 115
3 Non-native perception and native dialect influence in L2 ............................... 117
   3.1 The use of spectral and temporal cues in L2 perception ......................... 117
   3.2 The relationship between L2 perception and L2 production .................. 122
4 Non-native production and native dialect influence in L2 ............................. 124
   4.1 Arguments in favour and against dialect effects in L2 production ........... 125
   4.2 The role of input in the interpretation of L2 production results .............. 127
5 Final conclusions, relevance and contribution .............................................. 129
Chapter 1
Introduction

1 The starting point

The fact that the native language influences non-native production and perception is common knowledge in the field of second language acquisition and cross-linguistic research. First (L1) and second language (L2) phonological systems interact in various ways and many hypotheses have been tested in order to explain how L2 sounds relate to those of the native language. In most second language studies the learners’ native variety is treated monolithically as it is assumed to be homogeneous, that is, all native speakers of a given language are assumed to behave in the same way, regardless of their regional variety. Nevertheless, regional varieties do differ in non-trivial ways at the allophonic and sometimes even at the phonemic level. Not much is known about the role of the learners’ native dialect in the production and perception of second language sounds. If the starting point in acquisition is different for L2 learners from distinct L1 dialects and assuming there are sizeable differences between the dialects themselves, then the end points in L2, or at least the point where comparisons are made, might be different too. This dissertation aims to inform this area as it explores whether and how the native dialect influences non-native production and perception by focusing on two widely spoken languages, Spanish and English.

The present investigation on these languages was sparked by the empirical observation that there is a great amount of variation among Spanish learners of English in how specific English vowels are produced. For instance, the word-internal vowel in study, /stʌdɪ/, is produced with various qualities ranging from [stadi] to [stodi], with the latter realization more frequent for learners of English with a Caribbean Spanish background (Cuban, Dominican, Puerto Rican). Similarly, the vowels in stop (/stʌp/) and back (/bɒk/) are produced with a great amount of variability by Spanish learners of English. With this empirical observation as a starting point, and a small corpus of previous studies (Guitart, 1985, 1988, 1996) probing into the role of the native dialect in L2, this dissertation takes shape around the main research question of whether potentially systematic differences between vowels in the native dialects of Spanish would produce matching
differences in non-native perception and production of English vowels. As such, the first set of research questions explores the differences between vowels in two Spanish dialects, which, despite having identical phonemic inventories, are predicted to have different phonetic realizations. The second set of research questions investigates how listeners having different native dialects perceive specific target language (TL) vowels and whether the possible differences found can be attributed to the native variety. The third line of questioning addresses non-native production in two different groups of learners with different native dialects. Particular patterns of non-native perception and production are predicted to parallel the differences found between the native dialects.

2 Motivation for the linguistic variables selected

A rationale for the selection of the native dialects and the target language vowels to be studied is warranted. With respect to the native varieties of Spanish, because a specific pattern was observed for learners of English with a Caribbean background (/u/ being produced as [o], (Guitart (1996)), Cuban Spanish was thus chosen as one of the regional dialects of the Caribbean. Peninsular Spanish spoken in central and Northern Spain was selected as the other dialect for comparison because this variety has been analyzed in previous studies (Godínez, 1978; Guitart, 1996; Morrison & Escudero, 2007; Chládková, Escudero, & Boersma, 2011) and can thus serve as a reference and control for the data obtained in the present investigation of cross-dialectal differences. While there are some notable differences in the consonant inventories of these two varieties, their phonemic vowel inventories are identical and any potential cross-dialectal differences could be detected only in the phonetic realization of vowels. In order to reduce idiosyncratic variation, homogeneous and localized varieties\(^1\) that would reflect real cross-dialectal differences were targeted for the comparison and the particular local varieties of Cuban and Peninsular Spanish satisfied well this criterion. Thus, the present research adds to the growing body of dialectal comparisons across Spanish varieties new acoustic data on vowels in the understudied Cuban Spanish and corroborates already existing data on central and Northern Peninsular Spanish.

---

\(^1\) The localized varieties of León, Spain and Holguín, Cuba that are specifically investigated in this dissertation will be referred to as Peninsular (PS) and, respectively, Cuban Spanish (CS) to facilitate the discussion.
With respect to the TL vowels, the low and mid back English vowels /Q, √, A/ are of particular interest for this research for several reasons. First, as compared to studies focusing on the L2 English front vowels /i, l, e/, /E, Q/ (Bohn & Flege, 1990, 1992; Bohn, 1995; Cebrian, 2002, 2006; Escudero & Boersma 2004; Flege, Bohn, & Jang, 1997; Morrison, 2008a, b, 2009), far less attention has been given to the low and back vowels. The vowels /√, A/ were explored specifically in few studies (Højen & Flege, 2006; Flege & MacKay, 2004) or along with all the other vowels in the English inventory (e.g. Bohn & Bungaard Nielsen, 2010; Bohn & Steinlen, 2002; Cutler, Smits, & Cooper, 2005; Flege, 1995). A possible reason for this is that a great amount of regional and sociolinguistic variability is reported for these vowels (e.g. Rogers, 2000), as most of them are used to characterize specific patterns of variation across the Englishes of the world. These include the Northern Cities Vowel Shift (/A/ > /Q/ > /E/ > /√/) in American English or the /A, √/ low back merger in Canadian English (e.g. Labov, Ash, & Boberg, 2006). It is beyond the scope of this SLA research to investigate these variation patterns in English; however, the data on Canadian English obtained for the present research will be discussed in relation to recent findings in the field of variation in English. As such, the present research aims to expand the existent corpus of data on the English vowels /Q, √, A/ with an emphasis on how they are rendered in L2 perception and production.

Another reason for investigating English /Q, √, A/ in the L2 of Spanish learners is that the cross-linguistic inventory differences can shed light on specific interlanguage phenomena like assimilation patterns and perceptual strategies such as cue-weighting. Whereas the Spanish vocalic inventory has only five vowels of which only /a/ is low and /o/ is mid back, the English vocalic inventory is more than double in size and has several vowels occupying the low and mid back acoustic/perceptual space, namely /Q, √, A, √/. From an SLA standpoint, studying the discrepancies between L1 Spanish and L2 English in the low and back acoustic/perceptual space is highly informative with respect to one-to-multiple-categories assimilation patterns as well as the perceptual strategies of spectral and temporal cue weighting that may differentiate groups of learners with distinct native dialectal backgrounds. In relation to the perceptual strategies that L2 listeners adopt, the other reason for studying the contrasts among /Q, √, A/ is to explore whether L2 listeners favour the use of temporal cues when spectral cues are not readily available. This is a strategy regarded as being universal but experimentally demonstrated only for the TL English /i
- /I/ - /Q - /E/ and /E - e/ contrasts (Bohn, 1995; Cebrian, 2002, 2006). The weight that listeners attach to spectral and temporal cues may be different when perceiving contrasts among /Q, √, ∫/ than when perceiving those among /i, I, Q, E/.

3 Theoretical support

3.1 Phonetic theories

Having laid out the motivation for selecting the linguistic variables, namely language varieties and vowels to be investigated in this study, I present next the theoretical approaches that I will use throughout in developing the argumentation. In order to describe and compare vowel production in L1 and L2, this research relies on acoustic data, which are interpreted and discussed from the perspective of three phonetic theories, namely, the Theory of Adaptive Dispersion (TAD), the Quantal Theory (QT), and the Articulatory Settings or Base of Articulation theory (AS/BS). According to the TAD (e.g. Liljencrans & Lindblom, 1972), vowels (speech sounds, in general) in a language are distributed in the acoustic/perceptual space in such a way as to ensure an optimal balance between minimal articulatory effort and sufficient perceptual contrast between categories in a given speech situation. The Quantal Theory (Stevens, 1972, 1989) proposes that sounds across languages seek stable regions in which the acoustic parameters remain steady despite great variations in the articulatory configurations and avoid unstable regions in which the acoustic parameters vary greatly even for small variations in the articulatory configurations. The Articulatory Settings or Base of Articulation theory (e.g. Honikman, 1964; Bradlow, 1995) speculates that languages differ in that they have specific default articulatory settings, which are given by the most frequent consonant place of articulation in a particular variety. Thus, in cross-linguistic comparisons of vowels, acoustic spaces may appear shifted along a specific dimension (height, frontness) in one variety as compared to another, hence, a base-of-articulation effect. The phonetic theories presented here will be discussed in detail and referred to throughout the dissertation, in terms of predictions, types of data required, relevance and implications for the present research. At this point, I note that it is the TAD that best serves the purposes of the cross-dialectal and cross-linguistic experiments pursued in this dissertation because it facilitates such comparisons through acoustic measures of vowel space dispersion and contrast distinctiveness. Thus, the quantitative assessment of the differences between dialects is possible based on measurements of formants, vowel locations in
the acoustic space as well as distances between vowels. Moreover, these acoustic measures are also employed to quantify the differences between the TL vowels produced by the two groups of L2 learners that have different native dialects and the degree to which their L2 vowels approximate the TL vowels. As such, this dissertation has theoretical and empirical implications in that it discusses these phonetic theories as applied to cross-dialectal and cross-linguistic research and highlights how their predictions/assumptions can be tested and the types of data for which they are the most adequate.

3.2 Theoretical models of SLA

Whereas the phonetic theories mentioned above, in particular the TAD, are instrumental in the implementation of the cross-dialectal and cross-linguistic comparisons, theoretical assumptions and models of second language acquisition are also called upon in order to articulate the hypotheses regarding the L2 production and perception. Three theoretical approaches, the Speech Learning Model (Flege, 1995, 2003), the formal linguistic model of categorization (Escudero & Boersma, 2004) and the Perceptual Assimilation Model in L2 (Best, 1995; Best & Tyler, 2007) that will be discussed throughout this study are briefly presented next.

The Speech Learning Model (SLM) predicts the perceptual behaviour of L2 learners based on similarities between L1 and L2 phonetic categories. If an L2 sound is perceived as dissimilar, or new, the learner is likely to create a new phonetic category in the interlanguage. If an L2 sound is perceived as similar to some extent to a native category the learner will likely assimilate it to that L1 category, ignoring what s/he might perceive as fine-grained differences from the target. As such, L2 learners perceive and produce L2 sounds that are greatly dissimilar from the TL better than L2 sounds that are somewhat similar to the TL. New phonetic categories are created in the first case, whereas assimilation occurs in the latter case.

Escudero and Boersma (2004) propose a formal linguistic model of categorization and demonstrate via a particular case how the L2 acquisition process can be simulated. The model postulates a perception grammar with acoustic-perceptual constraints that adjust in successive stages in the learner’s interlanguage until they optimally match the environmental input. While using phonetically formulated constraints, the model taps into the more abstract, phonological level of categorization and explains concepts of category creation and category reuse that parallel
those of new phonetic category creation and assimilation of an L2 to a similar L1 sound set forth by the SLM.

The Perceptual Assimilation Model in L2 (PAM-L2) predicts the perceptual behaviour of L2 learners and describes several types of possible assimilations of L2 to L1 sounds as well as what accurate discrimination of such non-native contrasts is expected to be. This model is particularly illustrative in the case of great inventory discrepancies between L1 and L2, as L2 sounds may be assimilated to different L1 categories (two-category assimilation), to one L1 category with various degrees of accuracy (category-goodness and single-category assimilations), or go uncategorized. PAM-L2 considers both the phonetic and the phonological levels in explaining how L2 perceptual learning occurs and tackles the role of the lexicon in the creation of a new phonological category.

Like the phonetic theories presented in the previous section, these SLA models will be contextualized in relation to the objectives and the hypotheses proposed in the present study. Their predictions and operating concepts will be used in discussing and interpreting the results of the experiments undertaken here. The findings of this study are compatible with each of these SLA models in explaining the patterns in the interlanguage of L2 learners residing in their native regional environment rather than immersed in the TL environment.

4 Dissertation outline

Having briefly presented the object of study, the research questions and the linguistic variables that will be investigated in this study as well as the theoretical support sought, the remainder of this dissertation is organized in five chapters. Chapter 2 presents the methodological protocol, Chapters 3, 4, and 5 outlines three experiments on L1 production, L2 perception and L2 production, and Chapter 6 the general discussion, conclusions and contributions. Detailed literature reviews are included in each of Chapters 2 through 5. As such, Chapter 2 starts by motivating the selection of the testing protocol for each of the three experiments by comparing and weighting the relevance of previous findings to the type of data and results expected in this study. Descriptions of the tasks, materials, participants and measurements are subsequently laid out. Chapter 3 presents the first experiment, which analyzes L1 production data from native speakers having the Cuban (CS) and Peninsular Spanish (PS) dialects in order to establish the extent of the phonetic differences between vowels in these two varieties. The three phonetic
theories mentioned above will be discussed in detail in order to select the theoretical and experimental framework that best matches the purposes of the cross-dialectal comparison. An up-to-date overview of the research on dialects precedes the acoustic comparison of the PS and CS dialects. The second experiment (Chapter 4) focuses on L2 perception and aims to establish a hierarchy of difficulty for the contrasts between the English low and mid back vowels /AU, and to highlight the role of linguistic experience with these L2 contrasts as well as the influence that the native dialect exerts on non-native perception. General issues regarding vowel perception, the nature of the vowels investigated as well as previous studies that are relevant to native dialect effects in perception are addressed in Chapter 4. In order to further investigate the role of the native dialect in SLA, an L2 production experiment in Chapter 5 will provide acoustic data on how learners with different native dialects produce the L2 vowels and how well their vowels approximate the TL vowels. Several concepts set forth in the three models of SLA introduced above will be discussed in relation to the present study along with the sparse research on the effects of the native variety in the realization of L2 sounds. Finally, in Chapter 6 the results of these three experiments will be discussed in relation to each other and interpreted in the light of the phonetic theories and SLA models in order to draw conclusions and highlight the contributions and directions for future work.
Chapter 2
Methodology

1 Introduction

As outlined in Chapter 1, this research aims to identify the effects that different native dialects may have on the realization of non-native vowels. Having selected two Spanish dialects from Northern Spain and Cuba and focusing on the target English /Q, √, A/ vowels, I develop this study from the assumption that differences in the phonetic realization of vowels in these two Spanish dialects produce matching differences in how these specific TL vowels are produced and perceived. To obtain data on L1 production, L2 production and L2 perception, I designed three experiments whose specific objectives are briefly outlined next. First, the L1 production experiment is intended to assess the acoustic differences in the realization of vowels in Peninsular Spanish (PS) and Cuban Spanish (CS). Based on existing literature on various Spanish dialects, which will be presented in detail in Chapter 3 along with specific hypotheses, vowels are predicted to be more fronted and shorter in PS as compared to CS. Second, the L2 perception experiment is designed to investigate the existence of distinct patterns of perceptual errors for the two groups of English learners with different native dialects. In order to isolate the dialect effects from the L2 experience effect, L2 learners and naïve L2 listeners in each dialectal group are tested. Drawing on reports that the TL /√/ tends to be perceived as /a/ by Peninsular listeners, but as /o/ by Caribbean listeners (e.g. Guitart, 1996), more errors are predicted for the back vowels /√, A/ in CS group of L2 listeners than in the corresponding PS group. Finally, the goal of the L2 production experiment is to test whether native dialect factors determine systematic differences in the realization of non-native /Q, √, A/ as measured acoustically. The effects of the native variety on L2 production are likely to be found in the phonetic implementation of the TL vowels. For example, if vowels in the PS dialect tend to be realized as more fronted than those in the CS dialect, then L2 learners from PS should possibly produce L2 vowels that are more fronted as compared to those produced by learners from CS. Similarly, shorter durations of the native vowels in one dialect as compared to the other may result in different durations or use of duration in producing specific TL vowels. Specific interlanguage
strategies and assimilation patterns are predicted, with L2 learners in the PS group tending to assimilate the TL /ɒ/, /ʌ/ to a front category and the L2 learners in the CS group merging the TL /ʌ/, /ɒ/ into a back category.

This chapter is structured around the description of the methodology used to design the three experiments undertaken in this dissertation: an L1 reading task, an L2 perceptual discrimination task, and an L2 sentence completion task. Each of the three experiments is described in a separate section including a discussion motivating the selection of a particular experimental paradigm from a set of possible tasks, as well as detailed descriptions of the procedures, stimuli, and participants. The rationale of the sequence in which tasks are to be performed accompanies the presentation of the elicitation methods in the following sections.

2 Experimental design

Several sets of data and participant groups are necessary for the exploration of the native dialect effects in non-native production and perception of vowels. First, production data from PS and CS advanced learners of English will be compared to each other and in relation to the TL /ɒ/, /ʌ/, /ɒ/ vowels (Experiment on L2 production). Second, TL speech samples from native speakers of English are required to set up an L2 perceptual task as well to serve as control and reference for the L2 production task. Third, perceptual judgments from PS and CS listeners of English, both naïve L2 listeners and advanced L2 learners, are necessary to map the perceptual behaviour of all these four groups and assess the influence of the native dialect on L2 vowel perception (Experiment on L2 perception). A perceptual task tapping the phonetic rather the phonemic level of processing needs to be selected in order to be suitable for both L2 learners and non-learners of English. Finally, acoustic data on PS and CS vowels will be elicited in order to demonstrate the extent of the cross-dialectal phonetic differences (Experiment on L1 production). Informal speech, mimicking conversational style is likely to yield speech samples rich in dialect specific characteristics and thus suitable for cross-dialectal comparisons. In the testing sequence, the L2 production task comes first, followed by the L2 perception task and the L1 production task at the end of the session. Specific reasons for this testing sequence will be adduced in the description of each of the three experiments, which are presented in the following sections.
2.1 Experiment on L2 production

2.1.1 Motivation for task selection

The aim of the L2 production task is to elicit the English low and mid back vowels /Q, V, A/ from advanced PS and CS learners of English and, based on acoustic data, to compare how the two dialectal groups of learners differ from each other in the realization of these TL vowels and whether the native dialect shapes L2 vowel realizations. A wide range of elicitation methods of various degrees of formality are currently used to test L2 speech, including not exhaustively, reading of word lists or short texts, picture-naming tasks, sentence completion or cloze tasks, map tasks, conversation on pre-established topics. As previously mentioned in Chapter 1, there is variability in the phonetic realization of /Q, V, A/ in the speech of Spanish learners of English. In order to show that specific patterns of variability exist and may be attributed to the L2 learners’ native variety, I selected an oral sentence completion task. Advanced L2 learners listened to short definitions or questions in English and had to complete the sentences with the appropriate word and then repeat that word in a carrier phrase. The decision to use a sentence completion task is motivated, on the one hand, by the need to obtain from all participants a balanced set of stimuli with a particular structure, namely CVC real English words, with C an obstruent and V one of the /Q, V, A/ vowels. On the other hand, the L2 speech samples need to have some degree of spontaneity and informality in order to reflect possible transfer from the participants’ L1 varieties, and additionally to limit the orthography effects. By selecting this task, I assume that effects that could bias the results, such as those of task formality, orthographic input and the subject’s attentional focus on pronunciation, are minimized. In the next section I elaborate on how these effects are offset in the elicitation procedure chosen for the L2 production experiment designed for the present study.

Reading tasks have many advantages, including the possibility of controlling the structure and number of stimuli elicited, requiring little or no training, imposing a minimal memory load on the subject, and being quick and easy to implement. However, in terms of task formality, reading
tasks have the disadvantage of yielding ‘lab data’, which may differ from casual speech.\(^2\) Several studies (e.g. Holden & Nearey, 1986; Eckman & Iverson, 1993) have shown that the task formality interferes with the linguistic variety of the speaker and may level out particular cross-dialectal phonetic differences found in casual speech.\(^3\) This implies that reading tasks introduce a certain degree of formality that reduces the speech sample of key features found in casual speech. Non-native speakers performing a reading task in their L2 (e.g. words with similar structures in a predictable carrier phrase) tend to pay more attention to pronunciation than to content, thus producing hyperarticulated tokens, different from informal speech. As opposed to reading tasks, conversation tasks are more likely to produce peculiar forms found in casual speech that are representative of a particular linguistic variety. For example, Eckman and Iverson (1993) comparing L2 data from a reading task versus data from a conversation task, report higher pronunciation error rates for specific variables (consonant clusters) in the latter case. The main disadvantage, though, is that the number of stimuli obtained for analysis from conversation tasks is difficult to control and to balance across participants and linguistic variables. A middle ground between these two elicitation methods are the sentence completion, picture-naming and map tasks. With such tasks, there is more control of the structure and number of stimuli while the advantage of obtaining less formal speech samples is preserved.

Another effect that I have taken into consideration when deciding on the elicitation procedure is the possible effect of orthographic input, which was shown to influence non-native production (Young-Scholten, 1995) because it favours transfer from the native language. When reading aloud, L2 learners access specific mechanisms of processing written words and may use the pre-established relations between graphemes and their phonological representations in their native language, different from those in the target language. Pronunciation of written tokens in a reading task in L2 may differ from pronunciation of tokens produced without orthographic input. As such, in the oral sentence completion task designed for this particular study, participants receive aural rather than orthographic input.

\(^2\) For example, hyperarticulated forms can be identified in lab speech as speakers may resort to the formal registers whereas in casual speech less careful forms are present, like segment reduction, lenition, which may reflect dialectal idiosyncrasies.

\(^3\) In L1, “dialect differences which are clear in varieties of ‘casual’ speech can be greatly muted in reading word lists.” (Labov, 1980, as cited in Holden & Nearey, 1986, p. 17).
The sentence-completion task also has the advantage of redirecting the participants’ attentional focus from pronunciation to vocabulary retrieval, as my purpose here is to elicit L2 speech that is as casual as possible. If participants focus on pronunciation, they are likely to produce more careful and formal speech, whereas if focusing on something other than pronunciation (i.e. vocabulary), their speech samples are likely to contain more phonetic details attributable to the native language dialect. Previous studies on L2 production have successfully obtained informal speech samples by disguising an L2 production task as a vocabulary-learning task (Broselow & Finer, 1991), or a grammaticality judgment task (Hancin-Bhatt & Bhatt, 1997). Listeners’ attention has thus been redirected towards processing meaning and grammaticality and away from pronunciation. Taking into account these considerations, the L2 production elicitation method has been designed as an oral sentence-completion task, which is described next.

2.1.2 L2 production Experiment. Oral sentence-completion task

The task was administered individually to 19 advanced learners of English and 4 English native speakers who listened to short sentences in the TL played back on a portable computer. Participants were requested to finish each incomplete utterance that they heard with the best fitting word and then repeat that word in a carrier phrase as indicated in the instructions given at the beginning of the test. For example, the L2 listeners heard the incomplete utterance ‘The opposite of good is ____’ and they have to answer: ‘Bad. Say bad once more’. Their answers were recorded. Prior to testing, two female English native speakers from Toronto pre-recorded the instructions, sample questions for familiarization and fifty-eight test questions, the answers and the repetition of the target word embedded in the carrier sentence (which were edited out and used to build the English corpus). The materials used for testing included speech samples from both English native speakers, as questions uttered by one speaker alternate with questions produced by the other speaker. The sound files corresponding to the instructions, sample questions and test questions were assembled into a Power Point Presentation on a portable computer. There was no orthographic presentation of the questions; each presentation slide includes only the number of the question and the shortcut icon to the corresponding audio file. After being debriefed in English on the nature and duration of the experiments and signing the consent form, participants proceeded with the experiment. A short training session in English helped subjects familiarize themselves with how to use the computer for this particular task, as well as with the test instructions and pronunciation of the two English native speakers.
2.1.2.1 Stimuli

The set of stimuli shown in Table 2.1 included a total of 48 target words and 10 distracters evenly distributed among the English vowels /Q, √, A/.

Table 2.1
Set of Stimuli for L2 Production Task.

<table>
<thead>
<tr>
<th>Minimal sets</th>
<th>/Q/</th>
<th>√/</th>
<th>/A/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triads</td>
<td>cat</td>
<td>cut</td>
<td>caught</td>
</tr>
<tr>
<td></td>
<td>hat</td>
<td>hut</td>
<td>hot</td>
</tr>
<tr>
<td></td>
<td>bat</td>
<td>but</td>
<td>bought</td>
</tr>
<tr>
<td></td>
<td>cap</td>
<td>cup</td>
<td>cop</td>
</tr>
<tr>
<td></td>
<td>bag</td>
<td>bug</td>
<td>bog</td>
</tr>
<tr>
<td></td>
<td>hag</td>
<td>hug</td>
<td>hog</td>
</tr>
<tr>
<td></td>
<td>stack</td>
<td>stuck</td>
<td>stock</td>
</tr>
<tr>
<td>Pairs</td>
<td>back</td>
<td>buck</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>pack</td>
<td>puck</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>bad</td>
<td>bud</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>cab</td>
<td>cub</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>black</td>
<td>-</td>
<td>block</td>
</tr>
<tr>
<td></td>
<td>tap</td>
<td>-</td>
<td>top</td>
</tr>
<tr>
<td></td>
<td>pad</td>
<td>-</td>
<td>pod</td>
</tr>
<tr>
<td></td>
<td>pat</td>
<td>-</td>
<td>pot</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>duck</td>
<td>doc</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>pup</td>
<td>pop</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>gut</td>
<td>got</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>dug</td>
<td>dog</td>
</tr>
<tr>
<td>Singletons</td>
<td>catch</td>
<td>pub</td>
<td>Bob</td>
</tr>
<tr>
<td></td>
<td>dad</td>
<td>tub</td>
<td>God</td>
</tr>
<tr>
<td></td>
<td>flag</td>
<td>shut</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>gap</td>
<td>come</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>tag</td>
<td>run</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. Shaded cells correspond to the subset of stimuli analyzed in this study.

The stimuli consisted of real English words with a C1VC2 structure, in which V was one of the /Q, √, A/ vowels, C1 a stop or a fricative /S h/ and C2 a stop or an affricate /tS/. There were six minimal triads (e.g. cat – cut – caught) and several minimal pairs (e.g. buck – back, duck – doc, tap – top). The distracters were also monosyllabic words including one of the target vowels but having a different structure with a consonant cluster and/or some other consonant (nasal, rhotic, lateral). Several types of incomplete utterances formed the testing set of sentences were presented to the participants, such as:
short definitions (e.g. A feline pet is a ____ (cat)),
synonyms or antonyms (e.g. A synonym of to return is to come____ (back)),
abbreviated forms or informal words (e.g. An informal word for policeman is ____ (cop) or a short form for doctor is ____ (doc)),
irregular Past Tense verb paradigms (e.g. to sit - sat - sat and to get - ___ - ____ (got - got)).

The full list of questions is given in Appendix 1. As the purpose was to obtain casual speech, only tokens that were produced spontaneously, with no hesitation and no prompting from the researcher were selected for analysis. Based on this criterion and the L2 learners’ performance, the subset of tokens chosen for analysis is shown in shaded cells in Table 2.1. Given that vowel duration varies depending on the voicing of the following consonant (vowels are shorter if followed by a voiced consonant) and the strength of this effect varies cross-linguistically (being stronger in English than in Romance languages (Keating, 1990)), a subset of words ending in a voiceless consonant was selected for analysis.

2.1.2.2 Testing equipment

The recording equipment used included an M-Audio Microtrack 24/96 professional 2-channel mobile digital recorder and a lavaliere unidirectional microphone. The recordings were made at a sampling rate of 44.2 kHz and a quantization of 16 bits. The audio files containing the extracted tokens were downsampled at 22.1 kHz and saved in wave format using Adobe Audition 1.5. The experiments took place in sound-attenuated or quiet rooms with carpeted floors, desks and draped windows in order to obtain high quality speech samples.

2.1.2.3 Participants

Forty-six participants were recorded at three different locations in Toronto, Canada; León, Spain; and Holguín, Cuba. The English native speaker group consists of six female participants from Toronto, 20-39 years of age, mean 29.6, of which two contributed speech samples for the setup of the L2 production and perception experiments and, along with the remaining four, provided speech samples for the Canadian English corpus. Each L2 group, with Peninsular and, respectively, Cuban backgrounds, consists of two subgroups, one of advanced learners of
English and one of non-learners. The criteria used to assign participants to the experienced group were the formal training in English and the use of English in everyday activities. Participants in both advanced groups from the PS and CS dialects had obtained or were pursuing a university degree in English language and literature and used English more than 10 hours/week. The non-learner groups had little or no exposure to English. Table 2.2 summarizes the distribution of participants with various linguistic backgrounds and experience with L2 in the three experiments that will be described in this chapter. The same participants in the corresponding groups provided data for all experiments (two experiments in the case of non-learners, and three experiments in the case of the advanced learners).

### Table 2.2
**Distribution of participants in the three studies**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>PS non-learners</th>
<th>PS advanced learners</th>
<th>CS non-learners</th>
<th>CS advanced learners</th>
<th>Native English</th>
<th>Total speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 production</td>
<td>--</td>
<td>10</td>
<td>--</td>
<td>9</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>L2 perception</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>43</td>
</tr>
<tr>
<td>L1 production</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>--</td>
<td>40</td>
</tr>
</tbody>
</table>

For the L2 production experiment, speech samples from 19 L2 learners of English, 10 from the PS variety and 9 from the CS variety and from the two female English native speakers mentioned above were used. The advanced group from PS included ten subjects (7 female, 3 male), mean age 36.8, sd 7.4. They had extensive formal training in English as they were students beyond the 3rd year (8) or had graduated (2) and used English on a daily basis for work, travel, or conferences.

The advanced group from CS consisted of nine participants (7 female, 2 male) mean age 29.2, sd 5.6, with university degrees in English language and literature (6) and other (3). They worked in the Public Relations Office in a tourist resort and reported using English on a daily basis with foreign tourists. Detailed background information on the participants is given in Appendix 2.

### 2.1.2.4 Measurements

The set of stimuli that have been selected to reflect spontaneous answers to the questions of the test included approximately 20 tokens (minimum 13, maximum 26) from each participant and vowel category, thus yielding a total of 583 vowels for the L2 learners from PS and 541 vowels for the L2 learners from CS. A similar set of words has also been selected for analysis from the
six English native speakers and consists of 366 vowels. Several acoustic parameters were necessary for the analysis: vowel duration, the F1, F2 and F3 formants and the fundamental frequency F0. All measurements were obtained with Praat 4.2.34. Vowel duration was measured between the release of the previous stop or the reduction of the frication noise (in the case of /h/ and /S/) and the onset of the following voiceless stop, signaled by formant transitions and abrupt changes (increase at the beginning and decrease at the end of the vowel) in the intensity curve. Once the vowel’s boundaries were identified, they were further verified by playing back the sound within and beyond the selection to ensure the selection is accurate. For formant measurements, a 50 ms window was placed in the middle of the vowel or where formants showed steady values and the average values of F1, F2, F3, F0 for the 50 ms selection were obtained using Praat’s functions ‘Formant’ and ‘Pitch’. The L2 perception task is the second experiment in the testing sequence and is presented in the next section.

2.2 Experiment on L2 perception

2.2.1 Motivation for the task selection

As outlined in § 1, the aims of this experiment are to identify native dialect effects in non-native perception of the English /Q, Ñ, Ą/ vowels and to compare how perception of these vowels changes from naïve to experienced learners in the two dialectal groups. As such, four groups of listeners are tested: L2 advanced learners and non-learners having PS and, respectively, CS as native dialects. A native dialect effect can be posited if both L2 learners and non-learners from the PS dialect differ in the same way from L2 learners and non-learners from the CS dialect. The potential preference of CS for more posterior phonetic realization of vowels in their L1 may be reflected in more perceptual errors with back vowels in L2 and may determine systematic differences vis-à-vis L2 listeners from a different L1 dialect, i.e. PS. The selection of a particular perceptual paradigm, identification versus discrimination, has to be carefully considered in order to be able to highlight different patterns in perceiving L2 vowels and reliably link these patterns to the native dialect. Next I discuss the characteristics of each of these perceptual tests and bring arguments in favour of the AX discrimination task to be used for this particular experiment.
In the identification task, subjects listen to one item at a time and have to identify and label it by selecting a tag from a list, either an orthographic symbol, a key word or a picture.\textsuperscript{4} In L2, this method gives insight into the perceptual assimilation and categorization patterns, as the listener assesses the similarity of the physical token to a mental representation. Moreover, if response time is measured and goodness of fit ratings accompany the identification task, we gain information on the type of assimilation and on how well the physical sound approximates the listener’s mental representation of that sound. Precise quantitative information on perceptual assimilations, crossover between adjacent categories, and cue weighting can be obtained through identification tasks using synthetic stimuli (e.g. Holden & Nearey, 1986; Bohn, 1995; Rochet, 1995; Escudero & Boersma, 2004). However, in the particular case of this study, L2 advanced learners of English and naïve listeners would have to identify L2 sounds using a label corresponding to one of the English vowels /\textipa{Q}, \textipa{V}, \textipa{A}/. This would be problematic for naïve listeners, as they cannot be asked to choose from a set of unfamiliar labels. Moreover, creating a continuum of stimuli varying from /\textipa{Q}/ to /\textipa{V}/ to /\textipa{A}/ requires the use of more than one or two parameters (e.g. F2 for the /i, l, y/ vowels as in Rochet, 1995; F1 and duration for /i, l/ as in Escudero and Boersma, 2004) as the /\textipa{Q}, \textipa{V}, \textipa{A}/ vowels differ considerably from each other along F1, F2 and duration.

Using a discrimination paradigm mitigates these problems. Rather than selecting an identification label for a particular sound, in discrimination tasks, subjects have to compare two stimuli and decide whether they are identical or not. Within the discrimination paradigm there are several possibilities: the AX task, the extensively used ABX or AXB tasks as well as others (4IAX or 4-interval forced-choice) using sets of two pairs of tokens. Since data for all three experiments have to be elicited in one session, considerations regarding the duration and memory load of each task are important. The perceptual task is demanding per se because participants have to listen to a great number of similar tokens, make fast decisions and concentrate steadily during the whole duration of the test, approximately 20 minutes. Therefore, I employ the AX

\textsuperscript{4} For example, in an L2 perception task testing identification of the English /i, l/ vowels, Spanish listeners hear the L2 English word bit and have to identify the vowel using either a symbol provided or by choosing between 2 images of words containing the corresponding to one of the /l, i/ vowels, like ship and sheep (cf. Escudero & Boersma, 2004).
task, as it is reported to be the least demanding in terms of memory load and stimulus uncertainty (Strange & Shafer, 2008). The subject listens to the item A (one of the English vowels in a C1VC2 word) and shortly after to the item X and decides whether they are the same or not. Thus, the memory load is minimal, as the listener assesses only one pair of sounds in each trial, as opposed to triads, as in ABX or AXB or sets of two pairs in the more complex discrimination tasks. The main drawback of the AX paradigm is the response bias (Brannen, 2002), that is, listeners tend to respond with all the same or all different and may decide on a ‘different’ answer based on criteria that are irrelevant to the task. For instance, subjects may base their decision on purely physical differences between tokens (tokens with different intensities, different speakers producing each of the items of the pair, etc.). However, a careful selection and pairing of stimuli is likely to eliminate, at least in part, the response bias effects in the AX task.

2.2.2 L2 perception – AX discrimination task

Having presented the motivation for the task selection, in this section I describe the AX discrimination task that tests how the English /O-/A/, /O-/Q/ and /A-/Q/ contrasts are perceived in the four groups of L2 listeners with different native dialectal backgrounds and different levels of linguistic experience with English. As it will be discussed in Chapter 4, in discrimination tasks, specific levels of processing of the differences between the vowels of a contrast (auditory, phonetic or phonological) can be exploited by setting the interval between the presentations of stimuli of a pair at particular values. According to Werker and Logan (1985), long interstimulus intervals, ISI, (> 1500 ms) favour phonological processing; intermediate ISI’s (500 – 1000 ms) favour phonetic processing, whereas short ISI’s (< 250 ms) allow for auditory processing. For the task undertaken here, the phonetic level of processing is targeted because it suits both L2 advanced learners and naïve listeners who do not have L2 phonological categories. Thus, by setting the ISI at 1 second, the task provides insight into the perception of phonetic differences between the L2 /O-/A/, /O-/Q/, and /A-/Q/ contrasts. Participants, seated at a portable computer were instructed to listen carefully to pairs of words with the target vowels and to answer as fast as possible whether the items were the same or different, by pressing the corresponding decision button.

---

5 In procedures with high memory loads, like ABX, subjects have access only to a phonetic code whereas if the task poses lower memory demands, subjects access the acoustic code that facilitates within-category discrimination (Werker & Logan, 1985, p. 36).
key. They were advised that they could not re-listen to any of the files, but they could pause the experiment between trials. To reduce the possibility of the participants’ pressing only one of the decision keys, they were informed that approximately half of the pairs are the same and half are different. Before proceeding to the actual test, a trial test that includes 6 pairs of words is presented to the subject.

2.2.2.1 Stimuli

The two English native speakers whose speech samples were used for the creation of the sentence-completion task described in §2.1.2., also provided the stimuli for the current experiment, a two-talker AX discrimination task. Tokens shown in Table 2.1 were used to create a block of 72 pairs by pasting together various words separated by a one-second ISI. The set included: an equal number of ‘same’ and ‘different’ pairs (2 x 36); an equal number of target vowel pairs /Q - A/, /A - √/, and /√ - Q/ (3 x 12); within the ‘different’ pairs, tokens produced by one speaker combined with an equal number of tokens produced by the other speaker (2 x 18); within the ‘same’ pairs, an equal number of identical tokens, same name – same speaker (physically identical) and same name – different speaker (2 x 18). Within each category, the order of presentation was counterbalanced for vowels and speakers. Sample stimuli are cat1_CE2_caught1_CE2 or cub2_CE1_cab2_CE2, in which the number next to the word indicates the round in which it was elicited, and CE1, CE2 represent the speakers’ codes. Nine distracter pairs were also added; the items in these pairs share the vowel but differ in voicing of the coda stop (e.g. but - bud). The 81 pairs were randomized and used to create two blocks and a total of 162 audio files of token pairs. A short trial test was set up to familiarize participants with the software and testing material. A full list of stimuli is given in Appendix 3.

2.2.2.2 Testing equipment

The perceptual testing software was developed in LabView 7.1 and runs on a portable computer. The screen (Figure 2.1) shows two large decision buttons ‘same’ and ‘different’ that are activated by pressing one of the computer function keys (F1 or F12) and a Pause button. The response time window displays the reaction time in milliseconds. The response timer is activated automatically at the end of the sound file and stops as soon as the subject has pressed one of the decision buttons. The delay between presentations of consecutive pairs (the intertrial interval ITI) can be varied between 1 and 5 seconds and was set at 3 seconds at the beginning of each
test. The researcher selects the trial or the testing session and the audio files are played back automatically after the subject has pressed one of the decision buttons. Participants’ answers (‘same’ or ‘different’) and their response times are stored automatically next to each pair of stimuli in Excel format.

Figure 2.1
Interactive Screen in the AX Discrimination Task

2.2.2.3  Participants

For the L2 perception test, 39 non-native listeners grouped by dialect and experience with L2 and 4 English native listeners were tested. The four English natives belong to the group of six described earlier in § 2.1.2.3, the other two who provided the speech samples did not take the L2 perception task. The L2 advanced learner groups with PS and CS backgrounds were the same as in the L2 production experiment described above. The non-learner group from Spain included ten subjects (9 female, 1 male), mean age 28.9, sd 8.9 with university education, who reported having minimal or no exposure to English. The non-learner group from Cuba consisted of ten participants (7 female, 3 male) mean age 32.6, sd 5.6. They worked in the same tourist resort as the L2 advanced group, but they had little or no exposure to English, nor did their jobs require any interaction with foreign tourists.

2.2.2.4  Measurements

Participants provide ‘same’ or ‘different’ answers to the 162 pairs equally distributed among the /\ - A/, /\ - Q/ and /A - Q/ contrasts. Their incorrect answers for each contrast are transformed
into error rates by reporting them to the total number of pairs in the corresponding category. The error rates are used to report perceptual performance and are further analyzed statistically.

2.3 Experiment on L1 production

2.3.1 Motivation for the task selection

The main purpose of this experiment is to compare vowels in two Spanish dialects, Peninsular and Cuban, and test the hypothesis that there are specific differences between vowels in these two regional varieties. Most of the cross-dialectal studies, whose findings will be reviewed in Chapter 3, elicited data on vowels through reading tasks (e.g. Recasens & Espinosa, 2006; Morrison & Escudero, 2007; Escudero, Boersma, Schurt Rauber, & Bion, 2009). In those experiments, the target vowel in a word with a specific consonantal and prosodic configuration was embedded in a carrier phrase, which was read in several rounds by each participant. In order to be able to compare and integrate the findings of the present study with previous comparisons across Spanish dialects, the same elicitation procedure has been adopted for this L1 production experiment, namely, a reading task. Bearing in mind the discussion in § 2.1.1 about the factors likely to interfere with production, namely, the task formality, the effects of orthographic input and the attentional focus, I argue next that such factors have a lesser weight for this particular experiment.

The L1 production task is the final experiment in a 45-minute testing session that includes the L2 production and L2 perception experiments described in § 2.1 and § 2.2. It is widely accepted that the beginning of the interview or test is more formal than other moments of the task, and it is characterized by clear hyperarticulated speech. This is the reason for which in some studies the initial part of an interview is not analyzed or is used to elicit the subject’s background information as the participant gets more comfortable with the test setting. By placing this reading task in L1 at the end of the testing session, after two demanding and uncommon tasks focusing on L2, I assumed that the subject was already comfortable with the test setting and likely relieved to be asked to perform a reading task in his/her native language. Whereas during experiments on L2 production and L2 perception, English is used exclusively for instructions and clarifications in order to prime the participants’ English mode, at this point of the interview, the L1 production task, the researcher switches the language of addressing participants from English to Spanish. The background questionnaire is filled at this point and the corresponding questions
are asked and answered in Spanish. The reading task in Spanish was designed as a very short one, requiring less than 5 minutes to complete and posing minimal memory load and attention requirements. To reduce task formality, the target words, which are real Spanish words, were presented embedded in a carrier phrase. Participants were explicitly instructed to speak at a normal to fast pace, as if talking to friends and, prior to testing, they were invited to familiarize themselves with the target sentences printed on flash cards. The effects of orthographic input are irrelevant for this task since subjects were asked to read sentences in their native language that has an almost one-to-one grapheme-to-sound orthography.

2.3.2 L1 production – Reading task

As already shown, participants were asked to read the target words embedded in a carrier phrase at a normal to fast pace that parallels informal speech. The researcher controlled the speech rate by presenting the flash cards with the target sentences quickly and steadily. This strategy aimed to reduce the effects of task formality and ensured that the vowels elicited were not hyperarticulated but rather similar to those produced in informal speech. For this experiment, instructions and clarifications were given in Spanish. The flash cards were randomized and presented in two rounds.

2.3.2.1 Stimuli

The L1 reading task includes 49 target words with a CVCa structure with the stress on the first syllable, as shown in Table 2.3. V is one of the Spanish vowels /i, e, a, o, u/ and C is a consonant (stop, approximant, fricative, affricate, rhotic, nasal). The tokens are embedded in the carrier phrase Digo___ otra vez. (I say___ again). A subset of tokens in labial context has been selected for analysis and appear in the shaded cells in Table 2.3.
Table 2.3

List of Stimuli in L1 Reading Task

<table>
<thead>
<tr>
<th>/i/</th>
<th>/e/</th>
<th>/a/</th>
<th>/o/</th>
<th>/u/</th>
</tr>
</thead>
<tbody>
<tr>
<td>viva</td>
<td>beba</td>
<td>baba</td>
<td>boba</td>
<td>buba</td>
</tr>
<tr>
<td>pipa</td>
<td>pepa</td>
<td>papa</td>
<td>popa</td>
<td>pupa</td>
</tr>
<tr>
<td>quita</td>
<td>quema</td>
<td>cada</td>
<td>coca</td>
<td>guita</td>
</tr>
<tr>
<td>guisa</td>
<td>guerra</td>
<td>gata</td>
<td>coba</td>
<td>cuca</td>
</tr>
<tr>
<td>dicha</td>
<td>beta</td>
<td>capa</td>
<td>copa</td>
<td>curro</td>
</tr>
<tr>
<td>cita</td>
<td>teta</td>
<td>tacha</td>
<td>dota</td>
<td>duda</td>
</tr>
<tr>
<td>tira</td>
<td>seda</td>
<td>dada</td>
<td>bota</td>
<td>tuba</td>
</tr>
<tr>
<td>tiza</td>
<td>seta</td>
<td>vaca</td>
<td>boga</td>
<td>zumo</td>
</tr>
<tr>
<td>diva</td>
<td>beca</td>
<td>paga</td>
<td>poca</td>
<td>dura</td>
</tr>
</tbody>
</table>

Note. Shaded cells correspond to items used in the analysis

2.3.2.2 Participants

Forty Spanish native speakers, learners and non-learners, 21 from PS and 19 from CS participate in the L1 reading task. Learners in the PS and CS groups are the same participants who provide data in the experiments on L2 production and L2 perception, whose profiles were described in § 2.1.2.3 and § 2.2.2.3.

2.3.2.3 Measurements

The data on Spanish dialects consist of vowels in bilabial context obtained from 21 speakers from PS and 19 speakers from CS. Each speaker contributed three or more vowels for each category, thus yielding 442 vowels for PS and 381 vowels for CS. The parameters measured were vowel duration, the F1, F2 and F3 formants and the fundamental frequency F0. The vowel segmentation and measurement procedure was the same as for the L2 vowels described in § 2.1.2.4. The acoustic data were used to draw vowel plots in the F1 x F2 space and to compare patterns of within-category variability.

3 Summary

This chapter has provided a detailed description of the methodology used in this study, which aims to obtain evidence concerning the role of the native dialect in non-native production and perception of the English /Q, V, A/ vowels. Three experiments were designed to compare vowels in two Spanish dialects and to elicit data on L2 production and perception from groups of
advanced learners of English and non-learners having different native dialects. Given the varied and complex nature of the data that needed to be collected from a large number of participants in one session, a careful selection of the methods was required. Thus, in selecting the elicitation procedures, I took into account factors like task formality, orthographic input, attentional focus and obtain speech samples that were as casual as possible both in L1 and L2 production. As such, an oral sentence-completion task for L2 production, an AX discrimination task for L2 perception and a reading task for L1 production was chosen. In order to further control formality in the L2 production task, only a subset consisting of the stimuli that were produced spontaneously, without any prompting, were selected for further analysis. The L2 perceptual task was designed so as to tap into the phonetic processing level. Finally, in the L1 production task, participants were encouraged to produce the stimuli at a quick pace, thus mimicking the informal register. The next three chapters of this dissertation are dedicated to the experiments and present the relevant literature review, hypotheses, results and their interpretation and conclusions.
Chapter 3
L1 production

1 Introduction

Having outlined the general objectives and main hypotheses of this study and the methodology of the experiments, this chapter presents a dialect comparison of two Spanish varieties, focusing on the theoretical framework suitable for such comparisons, previous findings on other cross-dialectal analyses and on vowels across Spanish varieties, as well as on specific predictions for Cuban and Peninsular Spanish, results and their interpretation. Regional varieties of a language may differ in the size and dispersion of their vowel spaces, distances between vowel categories, the degree of within-category variability as well as inherent duration. These acoustic parameters, which are based on formant and duration measurements, are commonly used to characterize differences between linguistic varieties. This chapter focuses on providing an acoustic description of vowels in Cuban and Peninsular Spanish based on data obtained through the reading task described in Chapter 2. Although Spanish is considered to be a well-investigated language, there are few studies that address vowel variation across its regional varieties and they all converge on the fact that vowels in several Spanish varieties differ indeed but not to a great extent. The aim of the present chapter is to identify the extent of such cross-dialectal differences, in order to subsequently explore the consequences of these differences for L2 production and perception. Specifically, differences in the realization of native vowels may be carried over into L2 production and determine different realizations for learners of English with different native backgrounds.

2 Phonetic theories and cross-linguistic comparisons

In the following section I present three phonetic theories relevant for cross-linguistic comparisons, along with the types of data needed when testing their predictions. By contrasting these theories I build a case in favour of the one that best suits the aims of the present cross-dialectal comparison.
2.1 Phonetic theories of speech production

Several theories of speech production have been formulated in an attempt to capture cross-linguistic sound patterns in vowel systems and provide a phonetic explanation to language universals. Among them, the Theory of Adaptive Dispersion (Liljencrants & Lindblom 1972; Lindblom, 1986), the Quantal Theory (Stevens, 1972, 1989) and the Articulatory Settings or Base of Articulation Theory (Honikman, 1964; Bradlow, 1995; Gick, Wilson, Koch, & Cook, 2004) are outlined in the following section and implications for the present study are discussed. In particular, I set forth arguments in favour of using the Theory of Adaptive Dispersion for the present study.

2.1.1 The Theory of Adaptive Dispersion

The Theory of Adaptive Dispersion (TAD) provides a framework suitable for cross-linguistic comparisons, not only for languages with different vocalic systems (Disner, 1984; Bradlow 1995) but also for dialects with similar, if not identical, phonemic vowel inventories (Recasens & Espinosa, 2006, 2009). According to TAD, vowels are distributed in the vocalic space in such a way as to ensure an optimal balance between minimizing the articulatory effort and creating sufficient perceptual contrast between categories. Thus, in a particular speech setting, speakers and listeners adapt their production and perception to the demands of the communication situation (Lindblom, 1990). This theoretical framework allows the direct evaluation of the differences between linguistic varieties through two specific measures, namely vowel space dispersion and contrast distinctiveness, both based on acoustic measurements of vowel formants. Thus, vowel space dispersion refers to how far apart are the point vowels (vowels with extreme F1, F2 values) in the F1 x F2 acoustic space and contrast distinctiveness refers to how much distance exists between adjacent vowels represented in the F1 x F2 space.

Conventionally, vowels are described in terms of height (whose acoustic correlate is F1; the lower the F1, the higher the vowel) and frontness (whose acoustic correlate is F2; the higher the F2, the more fronted the vowel) and represented in an F1 x F2 plot whose axes are arranged in such a way so that front vowels appear to the left, back vowels to the right, high vowels at the top and low vowels at the bottom of the plot (e.g. Figure 3.1). This kind of representation exploits the relationship that exists between the acoustic and articulatory properties of vowels. In addition to F1 and F2, higher formants, particularly F3, a correlate of roundness, the fundamental
frequency, F0, as well as the intrinsic duration (at a normal speech rate), the dynamic properties of vowels (formant movement or the vowel inherent spectral change) are parameters that can help characterize vowels.

The first measure, the vowel space dispersion, is directly related to the inventory size, as TAD predicts that the larger the vowel inventory of a language, the more dispersed the vowels in the F1 x F2 space. The degree of dispersion is given by the acoustic distance between the point vowels /i, a, u/. Significant vowel space expansion is reported to occur in languages with more than 14 vowels (Livjin, 2000), with greater distances between the point vowels /i-a/, /u-a/, /i-u/. However, even comparisons between varieties with similar (and small) inventories may show differences in vowel dispersion (e.g. Jongman, Fourakis, & Serrano, 1989).

The second measure is the contrast distinctiveness, which is given by the Euclidian distance between two adjacent vowels. The distance between two points in a bi-dimensional plot (in this case F1 x F2) is calculated as the square root of the sum of the squared projections of the distance on the axes. The projection on the F1 axis of the distance between two vowels is given by the difference between the F1’s corresponding to these vowels and, similarly, the projection on the F2 axis is given by the difference between the F2’s of the vowels. If this distance is very small, the contrast becomes difficult to discriminate and is likely to become neutralized. Thus, the contrast distinctiveness measure can signal a near-merger in a specific variety as compared to another (Recasens & Espinosa, 2009). Intervocalic distances may differ across linguistic varieties with identical inventories as a result of specific phonetic or phonological processes (e.g. assimilation, vowel harmony, aspiration)⁶ that may be present in one variety but not in the other (see § 2.1.1 in Chapter 6).

---

⁶ For example, more prominent nasal assimilation in a particular dialect causes adjacent vowels to become nasalized, which affects formant bandwidth, which may further differentiate vowels across dialects. Another example is /s/-aspiration (present, for instance, in Andalusian Spanish). In this case, word-final aspiration causes a more open articulation of preceding vowel, necessary for maintaining a morpho-phonological contrast (sg-pl). This more open articulation may propagate from word-final position to other vowels in the word (vowel harmony), which may generate differences from vowels in dialects without /s/-aspiration.
2.1.2 The Quantal Theory

In contrast to TAD that claims that vowels tend to disperse in the acoustic space according to principles of minimal articulatory effort and sufficient perceptual contrast, Quantal Theory (QT) formulated by Stevens (1972, 1989) proposes the existence of stable regions\(^7\) that vowels tend to occupy more frequently across languages. There is a non-linear relationship between acoustic and articulatory parameters characterized by stable regions in which configurations of articulatory gestures may vary greatly with little effect on the acoustic parameters in contrast to unstable regions in which small manipulations of the articulatory parameters result in important changes in the acoustic parameters. QT proposes that such quantal relations between articulatory and acoustic parameters represent universal principles by which vowel systems are organized and that vowels ‘seek’ such regions of stability. In particular, the point vowels /i, A, u/ tend to appear at the same acoustic locations across languages (Stevens, 1972). Therefore, less cross-linguistic variability is expected for these vowels as opposed to others. However, this prediction has been disconfirmed (Bradlow, 1995; Pisoni, 1980). As the articulatory parameters in the so-called preferred regions of stability may vary without affecting the acoustic output, linguistic varieties can implement the same acoustic parameter using different articulatory configurations. The implication for the present study is that, although vowels prove to have similar acoustic realizations in the Peninsular and Cuban varieties, they may still differ in the realization of their articulatory configurations. However, without an adequate articulatory experiment, one can only speculate on the articulatory differences between dialects.

2.1.3 The Articulatory Settings or the Base of Articulation Theory

Whereas TAD and QT make predictions regarding universal properties of vowel systems, the Articulatory Settings (AS) or Base of Articulation (BA) Theory (Honikman 1964; Bradlow 1995; Gick et al., 2004) addresses language-specific characteristics that differentiate the default articulatory configurations across linguistic varieties. Without actually performing an articulatory or acoustic experiment, but rather through direct observation and native speakers’ intuitions, Honikman reports that English and French articulatory settings differ in many respects. The

---

\(^7\) In these stable regions, also known as hot spots (Livjin, 2000), the acoustic/auditory parameters are stable, insensitive to variation in the articulatory parameters. They are separated by areas of instability in which the acoustic/auditory parameters vary dramatically even for small changes in the articulatory adjustments.
articulatory settings used to compare these languages are represented by a series of parameters, namely the jaw position, the configuration of the lips, the muscular activity of the oral cavity, the most frequent consonant place of articulation (or base of articulation, in Bradlow’s terminology), as well as the involvement of several sections of the tongue in articulation. According to Honikman, cross-linguistic differences between articulatory settings may be the cause of foreign accents reported even for advanced learners who may be transferring their L1 configurations into their L2 production.

Gick et al. (2004) found empirical support for language-dependent articulatory settings through two experiments using X-ray data from English and French native speakers. They measured several articulatory parameters (pharynx width, velopharyngeal width, tongue body distance from hard palate and, respectively, from alveolar ridge, lower-to-upper jaw distance, lip protrusion) corresponding to the inter-utterance rest position. Significant differences were found for the inter-utterance rest position in English versus French. In order to determine whether the settings corresponding to the inter-utterance rest position function as targets, their variability (magnitude of standard deviations) was compared to that of a specified articulatory target, the vowel /i/. No significant differences were found between the standard deviations of the inter-utterance rest position and /i/, suggesting that the former is indeed specified as a speech target and, consequently, is language-specific.

Bradlow (1995) tested several predictions of TAD, QT and the Base of Articulation theory with acoustic data on the vowels that English and Spanish share, namely /i, e, o, u/ in CVC words, and attributed the differences between English and Spanish vowels in part to a base of articulation effect. Specifically, the English vowels were systematically shifted towards higher F2 values, indicating a more fronted tongue position as compared to Spanish vowels. This shift of vowels in the F1 x F2 space in one language as compared to their counterparts in the other language represents the so-called base of articulation effect (Bradlow, 1995; Recasens & Espinosa, 2006). Bradlow claims that speakers across linguistic varieties show systematic preference for particular articulations (e.g. ‘cardinal alveolar tongue position’ for English speakers, versus back articulations reported for Arabic and other languages).

In the same vein, though analyzing articulatory (electropalatographic) data on coronal consonants, Kochetov and Colantoni (2011) found a tendency towards a more posterior
articulation in Cuban Spanish as compared to Argentinean Spanish. If this tendency represents a base of articulation effect, it is likely to be matched in the vocalic domain as well, with the Cuban variety having vowels articulated at more posterior locations. Finally, an acoustic study on two Spanish varieties, Peninsular and Puerto Rican, (Valle 1996, as cited in Guitart, 1996) that measured F1 and F2 formants of /a, o/, also points to a higher [a] and a more fronted [o] in the Peninsular variety as compared to the Puerto Rican variety, which potentially represents a base of articulation effect.

The three phonetic theories presented above make testable predictions regarding potential differences between linguistic varieties as concerns the phonetic realization of their vowels. In particular, TAD provides two measures – vowel dispersion and contrast distinctiveness – readily available from direct acoustic measurements. QT predicts less variability (i.e. relatively stable regions, (Bradlow, 1995)) for point vowels across languages. Relevant to the present study, the QT acknowledges the possibility that articulatory configurations differ cross-linguistically despite similar acoustic outputs. However, to test this prediction, articulatory data need to be corroborated by acoustic measurements. With respect to the AS/BA theory, it predicts that differences among languages may be attributed to different articulatory settings, namely to a systematic preference in speakers of a variety for particular places of articulation (e.g. more anterior versus more posterior articulations).

To conclude, in order to test various predictions set forth by the phonetic theories outlined here, specific types of data are necessary. Whereas TAD requires only acoustic data, QT and AS/BS require both acoustic and articulatory data. Given that the present study involved three separate experiments as well as a great number of participants who were tested in their home countries, adding an articulatory experiment would have been difficult to implement. As such, using TAD as a theoretical framework for cross-dialectal comparisons and its easy-to-compute measures appears to be the optimal choice.

2.2 Acoustic measurement of vowel formants and within-category variability

Having chosen TAD as the theoretical framework for testing cross-dialectal differences, it is necessary to briefly review previous literature on its practical implementation, in particular, measures of formants, duration and variability.
2.2.1 Vowel location

The first step in assessing the differences between two linguistic varieties is to compare their inventories and vowel distribution in the acoustic space. It has been shown that linguistic varieties with identical phonemic vowel inventories may differ, however, with respect to the location of their vowels in the F1 x F2 space. Based on acoustic measurements from separate studies, I show in Figure 3.1 that the five-vowel systems /i, e, a, o, u/ of Spanish, Greek, and Hebrew do not actually coincide in the F1 x F2 space. Spanish and Greek vowels were elicited in similar conditions and contexts (p-vowel-alveolar-a) (Bradlow, 1995; Jongman et al., 1989), whereas Hebrew vowels were produced in a p-vowel-p context (Most, Amir, & Tobin, 2000). Vowels in Spanish are reported to be systematically more fronted than vowels in Greek (Bradlow, 1995). The visual inspection of vowel locations in Figure 3.1 indicates that whereas the mean formant values for /u/ tend to be similar for the three languages, there is more difference among the mean formant values for /i, a/.8

---

8 Figure 3.1 and the arguments in this section are provided only for comparison. There are several differences in the data sets of each of these studies, including stimuli and materials, number of speakers, and more importantly, non-normalized formant values are reported.
Figure 3.1
Vowel locations in three languages with identical phonemic inventories

Note. The sources of acoustic data are: Bradlow (1995) for Spanish, Most et al. (2000) for Hebrew and Jongman et al. (1989) for Greek.

2.2.2 Within-category variability

Although comparisons between similar linguistic varieties sometimes produce non-significant differences between vowel locations, i.e., vowel spaces overlap to a great extent, within-category variability may be what really differentiates these linguistic varieties. For instance, a comparison of two Dutch dialects (Koopmans-van Beinum, 1973) yielded much more variability for the Utrecht dialect as compared to the standard Dutch variety. Variability was visually assessed in F1 x F2 plots by comparing the size of ellipses enclosing 50% of the realizations of vowels in each dialect investigated. Since the data elicitation method was the same for both varieties and other factors were carefully controlled (speech rate, register, gender, education), variability
within the category is considered a dialectal effect. As such, the visual inspection of the category distributions across varieties, as well as comparisons of standard deviations of F1 and F2 and maximal ranges for F1 and F2 within a vocalic category (Recasens & Espinosa, 2006) are measures of variability commonly used in cross-dialectal comparisons.

This section outlined the main acoustic measures that can be used to highlight phonetic differences among linguistic varieties with identical phonemic systems like Peninsular and Cuban Spanish. These acoustic measures are vowel dispersion, intervocalic distances, and within-category variability (standard deviations and maximal ranges). Dialects with identical phonemic inventories may differ from each other in some, if not all, of these acoustic parameters, as illustrated in the next section.

3 Previous research on dialect differences

Dialect comparisons are undertaken for a variety of reasons, such as to test or amend predictions of specific phonetic theories, like the Theory of Adaptive Dispersion or Quantal Theory (Recasens & Espinosa, 2006); to highlight specific language variation patterns for vowels (shifts and mergers) with sociolinguistic relevance (Recasens & Espinosa, 2009); to provide a qualitative and quantitative map of how language varies geographically (Godínez, 1978; Recasens & Espinosa, 2006; Escudero et al., 2009); or to evaluate whether and how learners’ native background affects L2 learners’ phonetic realization of target language categories (Holden & Nearey, 1986; Guitart, 1988; Morrison & Escudero, 2007). Regardless of the research goals, vowels have been found to differ across regional varieties of a language. The following sections present some of these findings, which are organized in cross-dialectal vocalic differences for regional varieties other than Spanish (Catalan, Portuguese, and Russian) as well as for Spanish varieties (Argentinean, Caribbean, Mexican, Peninsular, and Peruvian).

3.1 Cross-dialectal vowel differences in languages other than Spanish

For regional varieties of a language that have the same phonemic vowel inventories, like Spanish, dialectal differences, if any, are likely to be found in the phonetic (allophonic) realization of vowels. Specifically, in one variety vowels may be systematically realized as higher (lower F1), and/or more fronted (higher F2), and/or longer than in other variety. For
example, a comparison between Brazilian and European Portuguese (Escudero et al., 2009) determined that the dialects’ seven vowels occupy similar acoustic spaces. However, significant differences were found in the location of the mid-low front and back vowels /E/ and /ζ/ and the central vowel /a/, which are lower in Brazilian Portuguese than in European Portuguese. The different locations of /E/ and /ζ/ affect the intervocalic distances to the neighbouring vowels /e/ and /o/, which are thus greater in Brazilian Portuguese. A dialect effect was also found for vowel duration, with shorter vowels in Brazilian than in European Portuguese.

A very detailed cross-dialectal study (Recasens & Espinosa, 2006) analyzed acoustic and articulatory data from four varieties of Catalan: Eastern (EC), Western (WC), Majorcan (MC), and Valencian (VC). Several consonantal contexts were included. With respect to position in the acoustic space, significant differences were found along the F1 dimension for the four vowels /i, a, ζ, u/, whereas differences along the F2 dimension were non-significant. Differences in height (F1) were greater for low and mid-low vowels particularly for the Majorcan dialect that had a closer /a/ and a more open /ζ/. This dialect, which possesses the extra vowel /’/ as compared to the other three Catalan dialects, also tended to have a more dispersed vocalic space. However, this tendency proved to be only marginally significant. The intervocalic distance measure helped distinguish a dialectal effect along the F1 dimension particularly for the Eastern Catalan dialect, which exhibited a more equidistant distribution of vowels in the vowel space. Within-category variability was also measured taking into account vowels produced in several contexts (labial, dento-alveolar, palatal, velar, r/l). Against TAD-based predictions, there was more variability for the point vowel /i/ and less for /o, ζ/ in Majorcan Catalan as compared to the other dialects.

Holden and Nearey (1986) compared vowels in three Russian dialects (Minsk, Moscow and Kiev) in production and perception. Each dialect has the six vowels /i, ^, e, a, o, u/. Four of these six vowels proved to have similar locations. The main cross-dialectal differences were found for the high central vowel /’/ which tended to be diphthongized and more fronted, as well as for the vowel /u/ that was more posterior in the Kiev dialect compared to the other two dialects. No measures of vowel space dispersion, intervocalic distances, or duration (vowels were uttered in isolation) were reported; however, vowel location itself along the front-back dimension (F2) characterized the differences among these three dialects.
The dialect comparisons reviewed above reveal that some differences in vowel formant frequencies do exist. These differences are apparent mostly along the vowel height (F1) dimension as well as in the degree of within-category variability and affect the front vowels to a greater extent. Cross-dialectal duration differences are also reported. However, unless specifically controlling for the speech rate, it is difficult to tease apart speech rate effects from inherent duration differences.

3.2 Cross-dialectal vowel differences in Spanish varieties

The literature on dialectal differences for Spanish vowels is not particularly rich, mainly because the most salient dialectal differences among Spanish varieties are found in consonants rather than vowels. Indeed, regional varieties of Spanish are most easily differentiated with respect to various phonetic and phonological processes affecting consonants (/s/-aspiration and deletion, stop lenition, lleísmo-yeísmo-žeísmo, neutralization of liquids in coda, realization of /x/ and /T/, rhotic assimilation). Hence, vocalic differences tend to be overlooked.

Acoustic data for individual Spanish varieties do exist (e.g. Argentinean - Guirao & Borzone de Manrique, 1971; Peninsular – Quilis, 1981), but important methodological differences including data elicitation, stimuli, subjects render any comparison difficult. Early comparative studies on Spanish dialects (Godínez, 1978; Quilis & Esgueva, 1983) uncovered some cross-dialectal differences for vowels. For instance, Godínez (1978) examined Mexican, Argentinean, and Peninsular varieties. His analysis found that the Peninsular vowel space was smaller than that of the other two varieties; front vowels in the Mexican and Argentinean varieties were higher than in the Peninsular variety and [a] had a posterior realization in all three dialects rather than a central position as it is traditionally reported. The least variability was observed for front vowels in the Argentinean dialect. Much variability was found for [a] in Peninsular Spanish and for [o] in Mexican Spanish. The author acknowledges, however, that the report was based on a very small number of participants (6-4-6) and stimuli and more data are necessary for an accurate characterization of cross-dialectal differences. Other methodological inconsistencies can be identified, too. For instance, whereas isolated vowels were elicited from the Mexican group (6

---

9 Hualde and Prieto (2002) acknowledged some differences in vocalic sequences. Specifically, some vowel sequences (dúeto, riáda) are realized as diphthongs in some varieties and as hiatuses in others.
speakers), vowels in the pVpa context inserted in a carrier phrase were obtained from the Argentinean and Peninsular groups (4 and 6 speakers, respectively). Moreover, the participants in the Mexican and Argentinean groups were from the same city (Tijuana and, respectively, Buenos Aires), whereas the Peninsular speakers came from various regions of Spain.

Quilis and Esgueva (1983) measured Spanish vowels in stressed and unstressed position in CVCa words, where C was a bilabial (/p, b, m/). The main goal of the study was to compare the Spanish vowels to the cardinal vowels of the IPA. Twenty-two participants (11 male, 11 female) were tested. The majority of the speakers were from Spain, but some were from Latin America. Formant averages were presented together, regardless of the participants’ dialectal background but separated for men and women. Men produced stressed and unstressed vowels that were shorter than those of women. With respect to duration, however, vowels produced by Peninsular speakers were found to be slightly shorter as compared to those produced by Latin American speakers.

A previous attempt at comparing the Caribbean and Peninsular Spanish (Valle 1996, as cited in Guitart, 1996) suggested that Peninsular (Asturian) Spanish speakers produced a higher [a] and a more fronted [o] as compared to Caribbean (Puerto Rican) speakers. The analysis, based on a small group of speakers (4 per dialect) and stimuli, did not include however, all vowels in the phonemic inventory of Spanish, or any duration measures. This makes it impossible to compare vowel spaces and the exact location of vowels.

These early studies contribute valuable data on the Spanish vowel system by showing that there is cross-dialectal variation in the realization of vowels and thus setting the baseline for further investigation. However, they contain methodological inconsistencies with data elicitation and analysis (different materials used to elicit data from several dialectal groups, collapsing data from participants with different dialectal backgrounds). These early studies also lack data normalization procedures and the statistic evaluation of the results. More recent studies, which are presented next, rectified these inconsistencies through a careful experimental design and analysis, and by including larger groups of participants and stimuli, data normalization procedures and statistic analyses.

Morrison and Escudero (2007) conducted a thorough analysis of vowels (14 speakers per dialect, 60 stimuli per speaker) in sentence-final position in Peruvian and Peninsular Spanish. Each of
the Spanish vowels appeared several times in the following carrier sentence ‘En CVCe y CVCo tenemos V’ (In CVCe and CVCo we have V) and only the sentence-final isolated vowels were analyzed, probably to avoid co-articulation effects. Few cross-dialectal differences were found. Specifically, only /o/ was found to differ significantly along the F2 dimension (more than 10%), having a more fronted realization in Peninsular Spanish. The other differences, although statistically significant, were deemed to be too small to have linguistic relevance (i.e. to be perceptually detectable). Duration was significantly shorter (more than 30%) for Peninsular Spanish, however, the authors were cautious in characterizing the Peninsular dialect as having intrinsically shorter vowels. An alternative explanation based on different speech rate effects is offered as a possibility.

O’Rourke (2010) examined two Peruvian dialects, Lima and Cuzco, as part of a larger study on bilingualism and language contact. Vowels were extracted from a read corpus obtained from 6 male speakers for each variety. A larger vowel space, shifted towards the front (a base-of-articulation effect), is reported for the Cuzco variety as compared to the Lima variety. Vowel locations were significantly different along the F1 dimension for /i, a, o/ but not for /e, u/ and not for all speakers. Similarly, along the F2 dimension, /o, u/ did not reach significance for all speakers.

In a recent study, Chládková, Escudero, and Boersma (2011) compared Peninsular and Peruvian vowels in controlled consonantal and phrasal contexts with the purpose of exploring whether such contexts affected vowels differently across dialects. Twenty speakers from each dialect produced each of the five Spanish vowels in several consonantal contexts, /p, t, k, f, s/, in CVCe or CVCo isolated words followed by the carrier sentence ‘En CVCe y CVCo tenemos V’ (the same as in Morrison & Escudero, 2007). The first analysis, in which consonantal contexts were collapsed, revealed that [a] was higher, [e, o] were more centralized and shorter in Peninsular Spanish as compared to Peruvian Spanish. The median sentence duration was also shorter for the Peninsular speakers. The second analysis, which explored the context effect on vowel duration, F0, F1, and F2, detected cross-dialectal differences in the realization of [o, u, i] in /s/-context. Specifically, in /s/-context [o, u] were more fronted and [i] was more posterior in Peninsular than

---

10 The authors do not motivate their choice; my assumption is that they intended to avoid co-articulatory effects.
in Peruvian Spanish. In /p/- and /t/-contexts, [e] was also found to differ cross-linguistically, being realized as more posterior in Peninsular Spanish. The differences found between vowels in the /s/-context were attributed to the different realization of /s/ in the two dialects (apico-alveolar in Madrid Spanish versus dental in Lima Spanish) as well as to the presence of the additional interdental fricative /T/ in Peninsular Spanish.

In summary, the studies conducted on Spanish varieties to date indicate that dialects differ indeed with respect to vowel location, either along the F1 or the F2 dimension, or both, although the differences (in percentages, roughly corresponding to tens up to 100 Hz) are not very large. Peninsular Spanish appears in most of the comparative studies and its vowel space seems to be more fronted as compared to other varieties. This section also touched several aspects regarding data analysis, in particular, data normalization which is discussed in detail in the next section in order to determine a normalization technique that is adequate for cross-dialectal comparisons.

4 Normalization Procedures

A thorough cross-dialectal comparison requires that raw acoustic data be normalized in order to eliminate inherent individual and gender variation attributed to differences in vocal tract length. Nevertheless, any regional characteristics need to be preserved. Several normalization procedures are currently used and they can be grouped into vowel-intrinsic and vowel-extrinsic techniques (Adank, Smits, & van Hout, 2004). Vowel-intrinsic techniques assume that all information regarding a particular vowel of a talker is found in the vowel itself and thus use direct formant measurements of that particular vowel in Hertz or transformations of formant values on an auditory scale, such as logs, mels, or barks. In contrast, vowel-extrinsic techniques assume that all vowels of a speaker contribute to the identification of a particular vowel and, thus, such techniques include formants of all vowels in order to calculate the normalized formants of a particular vowel. Adank et al. (2004) evaluated eleven vowel-intrinsic and extrinsic normalization techniques to determine which method was best at preserving phonemic information and sociolinguistic variation, while eliminating anatomic variation. The analysis was carried out on nine Dutch monophthongs produced by a large group of participants (160) divided into men and women and stratified in eight regions. A series of discriminant analyses, which recognize patterns in the data, as well as multivariate analyses, which highlight the contribution of specific factors to the variation found in normalized data, were conducted.
With respect to normalization methods that best preserve phonemic information, Adank et al.’s results show that vowel-extrinsic techniques are better than vowel-intrinsic techniques. In particular, in Syrdal and Gopal’s (1986) vowel-intrinsic procedure, which takes into account formant ratios F0-F1, F2-F3, a great amount of phonemic information was lost. On the contrary, among the vowel-extrinsic procedures, Lobanov’s (1971) z-score transformation and Nearey’s (1978) single log-mean transformation were those normalization procedures that preserved most of the phonemic information.

The normalization methods were also assessed for their ability to reduce gender variation. According to the logic of discriminant analyses, a good normalization method should produce results close to chance level, meaning that a normalized vowel has a 50% chance of being attributed either to a man or to a woman. From this perspective, Lobanov’s and Nearey’s methods proved once again to be the best procedures.

Finally, with respect to preserving regional variation, the eleven methods were assessed separately for each of the nine Dutch monophthongs. As with the other criteria, Nearey’s and Lobanov’s normalization procedures were the best, producing percentages higher than chance level (12.5%, as there were 8 regions vowels could be assigned to). These percentages indicated the probability for a vowel to be attributed to the correct region. Separate scores calculated for each vowel indicated some variation in the chances that a vowel received the correct classification. For instance, /A/ had a 26% probability of being attributed to the correct region using Lobanov’s method and 23% using Nearey’s method whereas /Ç/ had a 27% probability with Lobanov’s method and 29% with Nearey’s method.

In summary, Adank et al.’s study demonstrated that Lobanov’s (1971) normalization method was among the best at preserving regional variation and eliminating most of the gender-related variation in the normalized data. A similar comparison of normalization procedures tested on Catalan dialects is reported by Recasens (2008); with Lobanov’s method, the shape of the normalized versus non-normalized vowel spaces was better preserved for all dialects investigated while individual variation was eliminated. Given the excellent performance of normalized data with respect to regional variation and the straightforward computation formula, shown in 5.2.2, the present study uses Lobanov’s normalization procedure.
5 The current study

In the previous sections I have motivated the choice of TAD as a theoretical framework for the present cross-dialectal study and reviewed the most important findings and methods focusing on the acoustic measures that are likely to highlight differences between linguistic varieties. The current study aims to contribute qualitative and quantitative acoustic data on two Spanish varieties, to corroborate production data on Peninsular Spanish with existing results and provide acoustic data on Cuban Spanish and to answer the question of whether the native variety affects non-native production.

5.1 Hypotheses and predictions

The studies reviewed in § 3 found cross-dialectal differences by comparing acoustic measures of formants and duration. For some vowels, parameters (height - F1, backness - F2), or speakers, the cross-dialectal differences were either below the level of significance or were too small to have linguistic relevance. Nevertheless, in some cases, vowels did differ phonetically. Based on what it is known so far, I predict several cross-dialectal differences, outlined below:

With respect to vowel location and dispersion in the acoustic space, the vowel spaces of Cuban and Peninsular Spanish show similar degrees of dispersion (measures of intervocalic distances between point vowels, /a-i/, /a-u/, /i-u/, do not differ). According to Godínez (1978), Morrison and Escudero (2007), and Chládková et al. (2011), the Peninsular vowel space is somewhat smaller than that of other Spanish dialects, however, the differences are not statistically significant for all cross-dialectal pairs of vowels. With respect to location, back vowels [o, u] in Peninsular Spanish are more fronted than Cuban vowels. Thus, some significant differences for these back vowels in F2 (higher F2 for PS than for CS) are predicted and possibly for [a] along the F1 (higher [a], i.e. lower F1 for PS as compared to CS).

In relation to vowel location prediction, the second hypothesis addresses intervocalic distances. Studies that compared PS against other dialects (Morrison & Escudero, 2007; Godínez, 1978) have shown that the PS vowel triangle tends to be smaller than that in other varieties. In addition to the visual inspection of the vowel plots in these two dialects, two sets of intervocalic distances
are used to compare dispersion, namely, distances from point vowels /i, a, u/ to centroids (d(i, C), d(a, C), d(u, C))\textsuperscript{11} and relative distances between point vowels (d(i-a), d(i-u), d(a-u)). I predict different degrees of dispersion of the vowel spaces in PS and CS, which are reflected in smaller intervocalic distances in PS as compared to CS.

With respect to \textit{variability}, dialects may differ in the range within which F1 and F2 vary (see § 2.2.2). Differences in standard deviations of F1 and F2 are predicted cross-dialectally, as one dialect may articulate more precisely, i.e. show less variation, along the height dimension (F1) whereas the other may be more accurate along the backness dimension (F2). Specifically, since a higher and more fronted vowel space is reported for PS, less variation is predicted for its front vowels and more variation for its back vowels (i.e. greater standard deviations for PS back vowels /o, u/). Likewise, if the Cuban vowel space tends to be more posterior, less variability is predicted for its back vowels and more variability for front vowels /i, e/ in CS.

Finally, with respect to \textit{duration}, there are several mentions in the literature that vowels are shorter in Peninsular Spanish as compared to other varieties (e.g. Morrison & Escudero, 2007; Godínez, 1978). In line with previous reports, differences in inherent vowel duration between the two dialects are predicted, specifically, with PS having shorter vowels than CS. In the present study the elicitation method was the same for both dialectal groups and differences attributable to speech rate effect are thus considered to be minimal. Duration was shown to affect vowel space dispersion (Moon & Lindblom, 1994; Bradlow et al., 1996). Specifically, in clear speech, characterized by longer (hyper-articulated) vowels, the vocalic space tends to be larger than in fast or informal speech. Therefore, the relationship between intervocalic distances and duration needs to be explored before attributing inherently longer durations to one variety compared to the other. A strong correlation between duration and intervocalic distances (vowel to centroid) would signal a speech-rate effect.

Summarizing, the following hypotheses are tested:

- \textbf{H1 - vowel locations:} the PS vowels are more fronted than CS vowels, particularly the back vowels /o, u/.

\textsuperscript{11} This measure yielded better results than the intervocalic distances (Recasens & Espinosa, 2006).
• H2 - vowel space dispersion and intervocalic distances: distances from the point vowels /i, a, u/ to the corresponding centroids are different in the two dialects investigated, pointing to a smaller space for PS than for CS.

• H3 - within-category variability: different patterns of variability are expected: more variability with back vowels for PS and more variability with front vowels for CS are predicted.

• H4 - duration: shorter vowels are expected for PS as compared to CS

5.2 Vowel locations

5.2.1 Qualitative results for vowel locations

The present study compared production data from Cuban and Peninsular Spanish, obtained from 40 speakers (21 for PS and 19 for CS) and 4 stimuli per speaker for each of the five Spanish vowels /i, e, a, o, u/, for a total of more than 800 stimuli (see Chapter 2, § 2.3.2.3). Although vowels were obtained in all contexts (labial, alveolar, velar) and context combinations (labial-alveolar, labial-velar, alveolar-velar), only tokens in the labial contexts (i.e. pVpa, bVba) were analyzed because vowels in labial context have been shown to have the least amount of coarticulation as compared to other contexts (Recasens, 2007). Mean formant values and standard deviations for each dialect and vowel are given in Appendix 4.

As groups included both men and women, differences in vocal tract lengths affect the distribution of vowels in the F1 x F2 space. Moreover, given that the proportion of male and female participants in the experiment was different for the two dialectal groups (in PS there were 4 men and 17 women and in the CS there were 5 men and 14 women), the acoustic vowel spaces were plotted separately for male and female speakers (Figure 3.2). Each data point represents the mean F1 and F2 values obtained from measurements of 80 vowels per category produced by women and 20 vowels per category produced by men. From the visual evaluation of the distributions in Figure 3.2, the high front vowels in the Peninsular group appear more fronted than in the Cuban group, both for men and for women and the acoustic vowel space is slightly more dispersed in the Peninsular variety. Men’s vowel spaces are smaller and concentrated towards the lower frequencies in the F1 x F2 domain as compared to women’s vowel spaces, which is the usual gender effect reported on vowel formants.
Figure 3.2
Cross-dialectal vowel spaces for men and women

5.2.2 Data normalization

To obtain normalized formant values, I used Lobanov’s (1971) z-score transformation and calculated normalized formants F0 through F3 for each of the 40 participants with equation (1):

\[ F_{\text{normalized}} = \frac{F_i - m_i}{sd_i} \]

In this equation, \( F_i \) represents the raw formant value (F0 through F3) for talker \( t \), \( m_i \) is the average formant frequency across the five Spanish vowels for talker \( t \) and \( sd_i \) is the standard deviation for average \( m_i \). The table containing mean normalized formant values for each dialect and vowel is given in Appendix 5.

As discussed in § 4, the normalization procedure reduced gender variation and maintained phonemic and cross-dialectal differences. For a better visual evaluation of the cross-dialectal difference, formant averages for each vowel category are presented in Figure 3.3, in which each
data point is given by normalized F1 and F2 means across speakers for each vowel category and dialect (80 points per vowel category and dialect).

Figure 3.3
Cross-dialectal vowel spaces with normalized formant values

Just like with raw data, the normalized vowel space of Peninsular Spanish appears slightly higher and more expanded in the area corresponding to the high front vowel /i/. Moreover, in PS dialect, /o/ is more fronted and /a, e/ are higher as compared to Cuban Spanish. These tendencies apparent in Figure 3.3 are evaluated quantitatively next.

5.2.3 Quantitative results for vowel locations

To compare cross-dialectal differences in vowel locations, a series of one-way ANOVAs was run on normalized formant values with the independent variable ‘dialect’ (PS and CS). The number of ANOVAs was ten (2 formants x 5 vowels). Of the ten analyses performed, only three reached the level of significance, as follows. For the high front vowel /i/, F1 is significantly lower in CS than in PS, \( F(1, 160) = 4.319, p = .039 \), F2 is significantly higher in PS than in CS, \( F(1, 160) = 8.763, p = .004 \). For the mid back vowel /o/, F2 is significantly higher in PS than in CS \( F(1, 160) = 45.879, p < .001 \). The differences between the mean normalized F1 and F2 values for the /i/’s in PS and CS are 4.97% (F1) and 8.93% (F2). A 10.29 % difference in the mean normalized F2 values and a large effect size were found for the /o/’s in PS and CS.
5.3 Dispersion and intervocalic distances

5.3.1 Qualitative results for dispersion and intervocalic distances

In this section, vowel space dispersion in PS and CS is evaluated through comparisons between distances among point vowels /a-i/, /a-u/, /u-i/; the greater the distances, the more dispersed the vowel space. Distances involving the mid vowels /e, o/ are also compared. Intervocalic distances are calculated using two separate methods: (i) the Euclidean distance between pairs of vowels and (ii) the distance from each vowel to the centroid - a central point in the vowel space corresponding to the average across formants of all vowels produced by a speaker.

(i) The Euclidean distance in the F1 x F2 space between two points V1(F1, F2) and V2(F1, F2) is given by the equation (2):

\[ d(V1, V2) = \sqrt{(F1_{V1} - F1_{V2})^2 + (F2_{V1} - F2_{V2})^2} \]

In equation (2) F1_{V1} and F2_{V1} correspond to the mean values of normalized F1 and F2 across all repetitions of the vowel V1 produced by all speakers of each dialect.\(^\text{12}\) The intervocalic distances between the point vowels /a-i/, /a-u/, and /i-u/ shown in Figure 3.4, are measures of the degree of dispersion. In particular, /a-i/ and /i-u/ distances are larger, while /a-u/ is shorter in PS than in CS.\(^\text{13}\) This suggests that while the PS vowel space tends to expand more towards the front, the CS vowel space expands more towards the back. Inspecting the intervocalic distances between mid and high point vowels (/e-i/ and /o-u/), it is apparent that these distances are greater in PS than in CS. With respect to distances between mid and low vowels (/a-e/ and /a-o/), /a-o/ is greater in CS than in PS, suggesting different relative locations for [o] (i.e., [o] is closer to [a] in PS and is closer to [u] in CS).

---

\(^\text{12}\) Only normalized data are used. If using non-normalized data, F2 would contribute too much weight to the intervocalic distance (Recasens & Espinosa, 2006) thus obscuring the effect that F1 may have on intervocalic distance.

\(^\text{13}\) How much of a difference must there be for it to be real? The perceptual threshold differs across languages (intuitively, languages with more vowels, i.e. more crowded vowel spaces have lower thresholds) and across listeners of a specific variety. In Chapter 6 I further discuss this threshold and propose that the threshold corresponds to the smallest intervocalic distance transformed onto a relevant psychoacoustic scale in a particular L1 variety.
(ii) The Euclidean distances calculated with equation (2) provide information on the relative distance between vowels in the F1 x F2 space. A different measure of evaluating dispersion (based on Recasens and Espinosa, 2006; Bradlow, Torretta, & Pisoni, 1996) is to determine for each dialect a reference point in the vowel space that takes into account the grand mean of all vowels in the system and then calculate distances from point vowels to this reference value or centroid across all vowels in a variety. Thus, for each dialect I calculated the centroid F1 and F2 values by averaging normalized F1 and F2 values for all vowels pooled together. Then, I calculated averages of normalized formant values for each of the five vowels in each dialect. Euclidean distances were computed from these average values per vowel to the centroids. Normalized distances between point vowels and centroids in PS and CS are presented in Figure 3.5. Each dot represents the average normalized F1 and F2 values across approximately 80 vowels per category and each centroid corresponds to the grand mean of normalized F1 and F2 values across approximately 400 vowels per dialect.
Comparing the distances from point vowels to the corresponding centroid, Peninsular point vowel /i/ is slightly farther apart from the centroid than in the case of Cuban Spanish, suggesting a more dispersed vowel space in the Peninsular variety. The centroid in the Peninsular variety is higher and more fronted than in the Cuban variety with consequences on the distances to the mid front and back vowels as well. The detailed statistic analyses are reported in the next section.

5.3.2 Quantitative results for intervocalic distances and dispersion

Normalized distances to centroids for all speakers and dialects were submitted to one-way ANOVAs with ‘dialect’ as independent variable. Of the five analyses, only the distance from /i/ to the centroid was significantly different across the two dialects ($F(1, 160) = 12.364, p = .001$), with the /i – C/ distance greater in PS than in CS. In other words, the PS vowel space is only slightly more dispersed in the high front region than the CS vowel space.
5.4 Variability

5.4.1 Cross-dialectal patterns of variability

For each vowel category and dialectal group, tokens whose F1 or F2 values were within two standard deviations from the corresponding mean were identified. To obtain a clearer picture of variability, non-normalized data were used.\(^{14}\) Outliers, i.e. tokens whose F1 or F2 were beyond two standard deviations were manually removed, then the vowel distributions in both dialects were replotted. Ellipses were drawn around each vowel category so as to optimally include all dots (Figures 3.6 and 3.7). Thus, the visual comparison of the size of ellipses incorporating only data points within 2 standard deviations offers a picture of the cross-dialectal differences in terms of variability. Along the F1 dimension, there is more variability for PS vowels than for CS (ellipses are higher). However, along the F2 dimension, there is less variability in PS than in CS for /i, e/ (ellipses in CS are elongated) and more variability for /a, o/ in PS than in CS. In both dialects /u/ seems to vary little. There is also more category overlap between /i/ and /e/ in CS as compared to PS.\(^{15}\)

\(^{14}\) Normalization levels out differences by removing/highlighting deviations in the data. Removing points that are beyond 2 standard deviations in non-normalized data should have the same effect in normalized data. Moreover, outliers should be the same, be they in normalized or non-normalized data.

\(^{15}\) Further variability patterns could be highlighted if analyzing more contexts. Potentially, dialects differ in co-articulatory patterns (QT).
Figure 3.6
Vowel variability patterns in PS
5.4.2 Within-category variability

As shown in the previous section, vowels display different cross-dialectal patterns of variability. To illustrate more accurately how each of the five vowels of Spanish vary across PS and CS, standard deviations of the normalized F1 and F2 values per vowel (80 entries for each dot) were plotted in Figures 3.8 and 3.9.
Figure 3.8
Standard deviations of F1 of vowels in PS and CS

Figure 3.9
Standard deviations of F2 of vowels in PS and CS
There is more variation along F1 (height) for all PS vowels as compared to the CS vowels and the greatest amount of within-variety variation is found for the low vowel /a/. In PS, /i/ too varies in height greatly. With respect to the F2 dimension (front-back) PS shows less variation with front vowels /i, e/ as compared to CS, whereas CS shows less variation with low /a/ and mid back /o/ as compared to PS. Both dialects show similar degrees of variability with /u/. In summary, there is more overall variability along F1 (height), particularly for /i/ in PS than in CS, whereas CS displays more variability along F2 (front-back) for front vowels, especially /i/. Thus, the pattern of variability, quantified in normalized standard deviations, confirms the tendencies in non-normalized data that can be visually assessed from Figures 3.6 and 3.7.

5.5 Vowel duration – qualitative and quantitative analysis

Overall, duration averages were higher in the CS group as compared to the PS group (Table 3.1).

<table>
<thead>
<tr>
<th>Dialect</th>
<th>/i/</th>
<th>/e/</th>
<th>/a/</th>
<th>/o/</th>
<th>/u/</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS</td>
<td>84</td>
<td>84</td>
<td>90</td>
<td>88</td>
<td>85</td>
</tr>
<tr>
<td>CS</td>
<td>91</td>
<td>99</td>
<td>109</td>
<td>105</td>
<td>93</td>
</tr>
</tbody>
</table>

Average durations were computed for each talker and vowel (Figure 3.10, 3.11, 3.12, 3.13, and 3.14) and plotted separately with increasing values for PS and CS.
Figure 3.10
Average duration per talker and dialect for /i/

Figure 3.11
Average duration per talker and dialect for /e/
Figure 3.12
Average duration per talker and dialect for /a/

Figure 3.13
Average duration per talker and dialect for /o/
As shown in the plots above, Cuban vowels are longer than the corresponding Peninsular vowels. A series of five one-way ANOVA’s was performed on average vowel durations with ‘dialect’ as independent variable. There were no significant cross-dialectal duration differences for the high vowels [i, u]. For the remaining vowel categories, the PS vowels were significantly shorter than CS vowels, as follows. For [e], \( F(1, 38) = 9.932, p = .003 \), the PS [e] was 16.3% (13.8 ms) shorter than the CS [e]. For [a], \( F(1, 38) = 10.811, p = .002 \), the difference was 19.8% (18 ms). Finally, for [o], \( F(1, 38) = 20.183, p < .001 \), the difference was 19.9% (17.5 ms). The effect size is small and the differences in duration are smaller than 20 ms, which may be not significant from a phonetic and perceptual standpoint, i.e., listeners from other varieties may not detect the differences perceptually. However, as Morrison and Escudero (2007) point out, “if the observed vowel duration difference between two dialects is intrinsic, rather than simply a speaking rate effect, then this could have an effect on L2 speech studies where vowel duration is a relevant factor” (p. 1508).

---

\[16\] Given that all speakers adhered to the instruction of producing speech samples at a normal to fast pace, I considered that normalizing for duration was unnecessary.
In order to check the possibility of speech rate and, implicitly, vowel duration affecting the vowel space dispersion, a series of correlations were performed between durations and distances to centroids (5 distances x 2 dialects). Three correlations came out as significant, as follows. In CS, for /e/, the Pearson correlation coefficient was .350, \( p = .002, \) \( N = 77 \) and for /a/ it was .283, \( p = .003, \) \( N = 73 \). In PS, for /a/ the Pearson correlation coefficient was .432, \( p < .001, \) \( N = 82 \). The effects are nevertheless small (CS /a/) to moderate (CS /e/ and PS /a/), indicating that the speech rate marginally influenced vowel dispersion. Worth noticing is that the effect occurs with the open vowel /a/ which is intrinsically a long vowel.

6 Hypothesis evaluation and discussion

As similar studies on cross-dialectal differences in vowels have shown (Godínez, 1978; Morrison & Escudero, 2007, Chládkova et al., 2011) and in line with the predictions of the current experiment, differences between vowels in Peninsular versus Cuban varieties do exist. Several acoustic parameters were compared. Hypothesis 1 - vowel locations predicted, based on previous findings on Spanish dialects, that the PS vowels are more fronted than the CS vowels, in particular [o, u]. Significant differences along F2 (front-back dimension) were found for the back vowel [o] (higher F2 for PS than for CS) but no difference for [u]. Despite previous reports of PS [a] being higher than in the Caribbean variety (Valle 1996, as cited in Guitart, 1996; Chládková et al., 2011), the [a]’s in PS and CS did not differ significantly here. With respect to front vowels [i, e], significant differences in vowel location were found for [i] along both F1 and F2, but not for [e]. The differences are of approximately 10% along the F2 dimension, which correspond to a few hundred Hertz on the non-normalized scale and potentially produce cross-dialectal differences in the quality of the perceived vowel. In accordance with Chládková et al. [e, a, o, u], but not [i], appear as more centralized in PS, although only [o] is statistically more centralized as compared with the [o] in CS. As such, H1 received only partial confirmation: from the back vowels, only [o] is more fronted and from the front vowels, only [i] is more fronted in PS as compared to CS.

Hypothesis 2, vowel space dispersion and intervocalic distances, predicted a smaller space for PS than for CS distances, which should be reflected into different distances from the point vowels /i, a, u/ to the corresponding centroids in the two dialects investigated. Except for the distance \( d(i, C) \), which is greater in PS, all the other distances were found to differ non-
significantly across dialects. Based on the point vowel-to-centroid measure, PS vowel space is slightly more dispersed in the front high region corresponding to \([i]\). The alternate measure of intervocalic distances, calculated relatively to other vowels (not to the grand mean across all vowels), confirmed the tendency for the distance \(d(i, a)\), between the front point vowel \([i]\) to \([a]\) to be larger in PS than in CS. However, the distance from the back point vowel \([u]\) to \([a]\) tends to be greater in CS than in PS. In other words, the PS space is significantly more dispersed in the high-front domain, whereas the CS tends to be slightly more dispersed in the back high domain. The intervocalic measures also indicate different proportions between \(/a-o/\) and \(/o-u/\) distances for the two dialects suggesting differences in the category boundary. Based on these two measures of intervocalic distances, I conclude that, overall, vowel spaces are similarly dispersed in the two dialects despite a tendency for the PS vowel space to expand more in the front-high region.

According to Hypothesis 3, within-category variability, different patterns of variability were expected: more variability along \(F1\) (height) for PS and more variability along \(F2\) (backness) for CS are predicted. Overall, PS vowels vary more along the \(F1\) dimension but show less variability than CS vowels for \([i, e]\) along the \(F2\) dimension (Figure 3.8 and 3.9). More variability was found for CS with front vowels along the \(F2\) dimension. A great amount of variability for \([a]\) along \(F1\) is present in both varieties, which is in line with a universal pattern of greater variability in the realization of open vowels (Recasens & Espinosa, 2006). Variability of the back point vowel \([u]\) is the lowest along \(F1\) and \(F2\) in both dialects but the opposite is true for the front point vowel \([i]\), which varies largely along \(F1\) in PS and along \(F2\) in CS. As compared to other studies, Godínez (1978) found a different pattern of variability in the three dialects he explored. Overall, front vowels showed less variability than back vowels whereas \([a, o]\) were the most variable. In his study, Peninsular Spanish displayed moderate variability as compared to Argentinean and Mexican Spanish, which were the least and, respectively, the most variable. Recasens and Espinosa (2006) also reported front vowels as varying less than back vowels in all four Catalan dialects analyzed, thus supporting the fact that within-category variation patterns are dialect- and language-specific.

Finally, with respect to Hypothesis 4 concerning vowel duration differences, the analysis yielded statistically shorter \([e, a, o]\) vowels in PS as compared to CS, whereas high vowels \([i, u]\) have similar lengths. This is not surprising as high vowels tend to be shorter as compared to low
vowels and thus, there is less variability across varieties with respect to duration in high vowels. Just like in the case of spectral differences between the [i, o] vowels in PS and CS, the question is whether a statistically significant cross-dialectal duration difference for [e, a, o] vowels is phonetically significant. I discuss this issue at length in Chapter 6.

7 Conclusions

The acoustic exploration of the cross-dialectal differences between PS and CS yielded the expected results. These dialects with identical phonemic vowel inventories differ to some extent in the phonetic realization of their vowels. The visual inspection of the normalized vowel spaces in each dialect (Figure 3.3) suggests that [i, o, u] are more fronted in PS that in CS, although statistically only [i, o] are actually more fronted, with [i] also being higher in PS than in CS. Duration of [e, a, o] was also confirmed to be shorter in PS than in CS. While the cross-dialectal differences are not very large, they might indicate a different organization of the native perceptual vowel space as found for the dialects of Russian in Holden and Nearey’s (1986) study (see § 3.1) with direct influence on the realization of specific non-native sounds. It is also worth reminding the reader that only vowels in bilabial context were analyzed, although vowels in other contexts were also elicited. These other contexts could provide additional information on co-articulatory patterns and potentially greater cross-dialectal differences could be identified. As it will be discussed in Chapter 6, there are significant differences between the consonant inventories of these two dialects as well as different phonological processes involving consonants (/s/-aspiration and /s/-deletion in coda position, nasalization, different stages of stop lenition), which are possibly affecting the adjacent vowels. The analysis included several acoustic parameters (formants and durations) that helped to describe the differences in location, intervocalic distances and variability patterns. Given the differences in duration, an additional parameter, the vowel inherent spectral change, would add further information to the cross-dialectal comparisons, namely more formant movement in the dialect characterized by longer vowels (CS). While these are avenues worth exploring in the future, the data analyzed in this experiment show that the differences between the vowels of the PS and CS dialects do exist and suggest that such phonetic differences are large enough to produce a different organization of sounds in L2.
Chapter 4
L2 perception

1 Introduction

This chapter examines several factors in the reportedly different perception of the English vowels /Q, √, A/ by Spanish native listeners with different dialectal backgrounds. Previous studies with inexperienced Spanish learners of English (Guitart, 1985, 1996) found that each of the /Q, √, A/ vowels may be heard differently and, more importantly, that the learner’s native variety might influence L2 perception. Specifically, English /√/ tends to be perceived as /a/ by Peninsular listeners, but as /o/ by Caribbean listeners (see the detailed discussion of Guitart’s studies in § 4.3). The present research reconsiders the hypothesis of the native dialect effects in non-native perception proposed by Guitart using a different experimental design and examines the effects of experience with L2 and the type of contrast. Thus, the current study aims to answer two questions: (i) Does non-native perception differ for listeners with distinct native varieties and what dialect-specific perceptual strategies are responsible for the differences? and (ii) How does the type of contrast and experience with non-native contrasts affect L2 perception? Given the reduced number of studies exploring the effects that the native variety has on non-native perception, this research aims to contribute information on such effects. The nature of the L2 contrasts analyzed here and the role of experience are also relevant in language acquisition and language teaching involving two widely spread linguistic varieties, English and Spanish.

The chapter is structured as follows: in section 2 general issues regarding the nature of vowel perception, processing levels and between- and within category discrimination are presented. In section 3, I explore the perceptual characteristics of the low and back vowels in identification and discrimination tasks as well as related issues on inventory size, L2 experience and inherent contrast difficulty. Section 4 reviews the literature on non-native perception, in particular on cross-linguistic, cross-dialectal differences, and native dialect effects in non-native perception. The remaining sections present a perceptual experiment, hypotheses and predictions (5.1), results (5.2), hypothesis evaluation and discussion (5.3) and conclusions (5.4).
2 General issues in vowel perception

This section presents several concepts that are essential to understanding how sounds are perceived and processed, what types of tests are used to ‘measure’ perception, how vowel perception differs from consonant perception and what is their relevance for the present study. Speech perception refers to the “internal mental (and physiological) process by which the perceiver recognizes incoming stimulus events as instances of mental categories” (Strange & Shafer, 2000, p. 159), a process allowing the listener to capture the acoustic signal, code and interpret it. The continuous flow of speech is coded at different levels of processing (auditory, phonetic) and interpreted as discrete categories at the phonemic level. In speech perception all these levels of analysis are abstract (Strange, 1995) as opposed to speech production, for which the phonetic (acoustic and articulatory) level is physically measurable. To obtain concrete information on how speech is processed perceptually, identification and discrimination tasks are used (see also Chapter 2). In identification tasks subjects listen to one stimulus at a time and have to identify a particular sound by choosing one label from a set usually provided by the researcher. In discrimination tasks, subjects listen to several (2, 3, or 4) stimuli at a time and have to discriminate between them, i.e. decide whether the stimulus in a specified position is identical or different from the other(s) provided in each specific trial. The purpose of these tasks is to obtain accuracy and response time measurements that give insight into how speech is processed. Identification tasks may be accompanied by category goodness ratings that quantify the perceptual distance between the real stimulus and its mental representation. In discrimination tasks, specific levels of analysis can be tapped into by a careful selection of the interstimulus interval (very long ISI > 1500 ms for phonemic processing, ISI of 500 - 1000 ms for phonetic processing and very short ISI < 250 ms for auditory processing, Werker & Logan, 1985). Long ISI’s allow listeners to code the incoming stimulus, compare it with a mental representation and thus categorize it. By the time the next stimulus is presented, the memory trace of the previous one has decayed and only the result of the categorization is available for comparison. It is assumed that this kind of categorization that accesses a mental code represents the phonemic level of processing. Perception occurs at the phonetic level for shorter ISI’s, in which listeners do not have enough time to access any mental representation. During the shorter interval between presentation of stimuli, memory traces of sounds are still available for comparison and thus listeners “show a sensitivity to phonetic distinctions that are used in some other (not their own)
languages” (Werker & Logan, 1985, p. 36). Finally, at the auditory level, for which stimuli are separated by a very short ISI, only acoustic differences can be processed, as no judgments about their phonetic and/or phonological nature can be made.

Having briefly presented two types of tests that are used to explore speech perception and processing levels in speech perception, in the following section I describe specific characteristics of vowel perception in order to define concepts of categorical versus continuous perception, within- and between-categories and category boundaries that will be further used in the argumentation of this chapter. Based on the overall shapes of the identification and discrimination functions obtained experimentally for vowels and consonants, it has been shown that stop consonants are perceived categorically whereas long vowels are perceived more continuously (Strange, 1995). Stated otherwise, when presented with a continuum of stimuli in which relevant acoustic parameters vary in equal steps from one to another (e.g. from /ba/ to /pa/ or from /bæ/ to /dæ/ as in Pisoni, 1973), (English) listeners do not perceive small changes within a category (e.g. a change from a /pa/ with a VOT = 50 ms to a /pa/ with a VOT = 70 ms) but detect well a category change (e.g. from /ba/ to /pa/ around a VOT of 30 ms). This category change detection is reflected into an abrupt slope of the identification function in the vicinity of the between-category change. The discrimination function shows a peak in the vicinity of the category change, and a drop to chance level for between-category changes. In other words, categorical perception of stop consonants means that within-category (i.e. phonetic) differences are not well perceived whereas between-category (i.e. phonemic) differences are well perceived.

In contrast, vowel perception is less categorical and more continuous. The fact that vowels are perceived more continuously means that there is a less abrupt slope of the identification function and a less prominent peak in the discrimination function for between-category changes whereas within-category changes are discriminated well above chance. As such, vowel discrimination is fairly good for both between- and within-category changes as compared to consonants.17 This implies that vowel perception relies more on phonetic differences rather than on the phonemic categories of the native language (i.e. experience), as is the case of consonant perception (Strange, 1995, p. 18). However, the influence of experience in discriminating phonetic vowel

17 Pisoni (1973) and others cited within argue that the reason for that is the nature of the cues present in the two types of sounds: longer, more salient cues for vowels versus shorter weaker cues for consonants (e.g. voicing for the entire duration of the vowel, 100-300 ms versus the aspiration burst of a stop, 30-70 ms).
The continuous nature of vowel perception also suggests that boundaries between vowel categories are less well defined than for consonants (Strange, 2007). Perception is more sensitive at detecting changes in sounds in areas near the category boundaries and less sensitive for within-category changes in areas far from category boundaries (Strange, 1992). If a vowel contrast receives incorrect answers systematically and requires longer response times, it may be the case that one or both vowels of the contrast are far from the boundary that the listener has established for a particular vowel contrast. This is particularly true in cross-linguistic perception when different vowel inventories in L1 and L2, and thus different category boundaries, cause a non-native between-category contrast to be perceived as a within-category contrast, and consequently to be discriminated less accurately.

The general issues on vowel perception presented in this section have methodological implications for the current cross-linguistic perceptual experiment. I argued in favour of an experimental design based on discrimination that elicits the phonetic interpretation of non-native vowel contrasts by setting the ISI at an intermediate value. Thus, the results are likely to show the participants’ sensitivity to phonetic distinctions that may not be present in their L1.

3 Low and back vowels in native and non-native perception

In this section I discuss the nature of the English low, mid and back vowels /Q, \( \hat{v} \), A/ and the fact that they appear to be inherently difficult to distinguish from each other. Findings from native and non-native perception studies are presented to show that these vowels are often misclassified in identification tasks and that contrasts between them tend to be poorly discriminated. Arguments regarding the weakness of the acoustic and articulatory cues, the small perceptual distances between these vowels, inventory size mismatches between L1 and L2, and the slow improvement with perceptual training in L2 are set forth.

3.1 Low and back vowels in identification tasks

The domain of low and back vowels represents a source of perceptual confusion even for native listeners. Several perceptual identification and classification studies (Syrdal & Gopal 1986; Cutler, Smits, & Cooper, 2005) converge on the finding that among the American English
vowels, /Q, √, A, ü, U/ are the most often misidentified by native and non-native listeners. For example, Cutler et al. (2005) presented 15 AE vowels in CV and VC syllables\(^\text{18}\) to three groups of listeners (American, Australian, and Dutch) to assess how vowel identification is influenced by dialect and language. According to the classification matrix (Table 4.1), the low and back vowels /Q, √, A, ü/ had lower identification scores as compared to the other vowels in the American English inventory despite the fact that the listeners belonged to the same variety as the speaker who uttered the test syllables.

Table 4.1
Classification matrix of four American English vowels as identified by American listeners (Cutler et al., 2005)

<table>
<thead>
<tr>
<th>Stimulus /V/</th>
<th>Response /Q/</th>
<th>Response /A/</th>
<th>Response /√/</th>
<th>Response /ü/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/Q/</td>
<td>89.1</td>
<td>1.6</td>
<td>-</td>
<td>1.6</td>
</tr>
<tr>
<td>/A/</td>
<td>14.1</td>
<td>35.9</td>
<td>21.9</td>
<td>20.3</td>
</tr>
<tr>
<td>/√/</td>
<td>6.3</td>
<td>14.1</td>
<td>67.2</td>
<td>7.8</td>
</tr>
<tr>
<td>/ü/</td>
<td>1.6</td>
<td>29.7</td>
<td>4.7</td>
<td>59.4</td>
</tr>
</tbody>
</table>

For example, /A/ was mistaken for /Q/, /√/ or /ü/ and was identified correctly only in 35.9% of the cases. Moreover, the neighboring vowels /√/ and /ü/ were incorrectly classified as /A/. While identification rates were fairly good for /Q/ (percentages in the high 80’s), identification of /√, A, ü/ was much lower and showed more variability, i.e. more categories were identified as the target. Cutler and colleagues found a similarly high misidentification scores for listeners from a different dialect (Australian English) and from a different language (Dutch). On the whole, results showed there was more confusion for back vowels, which suggests that the acoustic cues that signal back articulations (F2, essentially) are less robust and more likely to produce perceptual ambiguity. Lindblom (1986) attributed this effect to the reduced mobility of articulators and sensory control at the back of the oral cavity that correlated with less salient acoustic-perceptual phenomena (as compared to the front). Robustness of perceptual cues has been assessed through perceptual tasks in conditions of increasing signal-to-noise ratios (e.g. Cutler et al. 2005; Narayan, 2008). As such, these tasks have provided information on which cues are easily masked in noise and which manage to be transmitted even in noisy conditions.

---

\(^{18}\) These were embedded in multispeaker babble in three conditions of sound to noise ratios.
sum, even without considering factors like transfer from L1 and inventory size, contrasts involving back vowels are inherently more difficult.

### 3.2 Contrasts with low and back vowels in discrimination tasks

In addition to identification tasks suggesting that some back vowels are often misidentified, discrimination tasks bring further support that contrasts involving back vowels may be difficult for native and non-native listeners. A measure of the inherent difficulty of a contrast is the similarity/dissimilarity score that correlates with the perceptual distance between the elements of the contrast. Low dissimilarity scores (i.e. smaller distances) indicate that the contrast is difficult to perceive. Native or non-native judges are asked to assign similarity/dissimilarity ratings to vowel pairs.¹⁹ For example, in a study that investigated cross-linguistic discrimination of the English vowel contrasts, Flege (1995) found that English /A-Ã/ and /Q-E/ contrasts were those that generated the highest error rates among non-native listeners with various L1s (Spanish, Portuguese, Dutch, German, Korean, and Arabic). The /A-Ã/ contrast was the most problematic even for those L2 listeners (the L1 Dutch and German groups) whose overall performance was otherwise similar to that of English natives. Performance on the /Q-A/ contrast was also poor for the L1 Spanish group.

In a separate study that assessed the perceptual similarity between English and Spanish vowels, Flege, Munro, and Fox (1994) found that the /A-Ã/ contrast received low dissimilarity scores from English native speakers confirming that this contrast is difficult. Moreover, Flege, Munro and Fox (1994) also measured the differences between the Spanish and English vowels. Specifically, dissimilarity ratings for the Spanish-English pairs /a-Q/, /a-A/ and /a-Ã/, among other vowels, were elicited from both English and Spanish native speakers. Each of the English /A, Q/ paired with the Spanish /a/ produced high dissimilarity ratings from English native speakers but low ratings from Spanish listeners. The /a-Ã/ pair, however, generated low ratings from all groups suggesting that the small acoustic/auditory distance between these two vowels

---

¹⁹ Alternatively, the perceptual distance can be calculated from acoustic measurements transformed on a psychoacoustic scale (e.g. from Hertz to mels or Barks). Perceived dissimilarity increases as the distance between vowels in F1 x F2 acoustic space increases (Flege et al., 1994). However, it has been shown that acoustic similarity may not always reflect perceptual similarity (e.g. Bohn & Flege, 1992).
made them sound similar. Such measurements also give an indication of the possible perceptual assimilation patterns suggesting that the Spanish /a/ might be the L1 category used in assimilating the target language English /\( /A, Q /\) that obtained low scores (1.9, 2.1, and 3.1, respectively, as compared to the highest 7.8). In sum, the studies presented in this section showed that the contrasts between English /\( /Q, \sqrt{}, A /\) pose perceptual difficulties to native and non-native speakers and, that, in the particular case of native Spanish listeners of English, the proximity of these vowels to the Spanish /a/ contributes to their difficulty. This also brings about another important factor in non-native perception, the inventory size effect, and in relation to it, the experience with L2 contrasts. These factors are discussed in the next section with the purpose of explaining why some contrasts are difficult for non-native listeners and to what extent such difficulties may be overcome with experience and training.

3.3 Vowel inventory mismatches and experience in L2 perception

Vowel inventory size plays an important role in perception (Flege, 1995; Wagner & Ernestus, 2008). The greater the vowel inventory of a language, the greater the number of dimensions/features/cues required and used to perceive contrasts. However, the more crowded a particular area is in the perceptual vowel space, the more attuned and sensitive perception is to capturing fine-grained differences among phones in that vowel space. Thus, perception is shaped by the native inventory. This is true for native listeners’ perception of L1 contrasts. However, the task of non-native listeners is different and more difficult particularly if their native language inventory is smaller than that of the TL. For instance in the Spanish inventory, the only low vowel is /a/, whereas in English there are two low vowels /\( /Q, A /\) and the mid back /\( /\sqrt{} /\) (reportedly similar to /a/, see above). With fewer phonetic categories to attend to in L1, non-native listeners have to learn to sharpen their perception to the specific contrasts in L2.

Experience with L2 is expected to influence the listeners’ perceptual reattunement to TL sounds and contrasts and to help improve their identification or discrimination scores. However, experience with L2 was shown to affect unevenly L2 perception. Thus, some contrasts that are problematic for inexperienced non-native listeners (and even for native listeners) remain difficult for experienced non-native listeners and performance with such difficult contrasts improves slowly with experience or training. For example, Lambacher, Martens, Kakehi, Marasinghe, and
Molholt (2005) tested the effects of training on Japanese learners of English, and found that, after a perceptual training experiment with the American English vowels /Q, √, A, ç, ´/, Japanese learners showed much less improvement with the identification of /√, A/ than with the other vowels. In (TL) English, /√, A/ differ both spectrally and in duration but they are spectrally similar to (L1) Japanese /a/. Thus, given that vowel length is phonemic in Japanese and duration represents the primary perceptual cue, the Japanese listeners in this study may have only attended to the duration cues while interpreting the spectral differences as mere variations in vowel quality. Bohn and Flege (1990, as cited in Lambacher et al., 2005, p. 244) note that “not all nonnative sounds are equally identified and produced well. Some can be readily improved, whereas others are inherently more difficult and require more time, input, and training.”

In the same vein, Levy and Strange (2008) examined how experience with L2 and the phonetic context influenced non-native perception of the French contrasts, /i-y/, /u-ø/ and /y-ø/ of inexperienced and experienced American listeners. Not all of these vowel contrasts were equally easy to discriminate. Among them, /u-y/ was difficult because in the alveolar context it sounded like a between-category contrast to American English listeners. As predicted, perception improved with experience but unevenly as, occasionally, inexperienced listeners performed better than experienced listeners on the difficult contrast /u-y/. For this specific contrast, the difference in error rates between inexperienced and experienced listeners was only 4% in alveolar contexts, while in bilabial contexts the inexperienced actually had lower error rates than experienced listeners (8% versus 24.5%). The authors attribute this effect to the influence of the L1 coarticulatory patterns that “continue to exert their influence on some highly experienced L2-learners’ perception of non-native vowels” (p. 154).

The aim of this section was to show that many factors influence L2 perception (e.g. L1-L2 inventory differences, experience with L2) and that the low and back vowels and contrasts among them, which are analyzed in the current study, tend to be inherently difficult. Previous studies found that English /√, A/ are often misidentified; the /√-A/ contrast poses problems in discrimination even for native listeners and more so for non-native listeners (Spanish included).

---

20 For native AE listeners, the alveolar context causes /u/-fronting and thus further reduces the /u-y/ perceptual distance
In the particular case of the Spanish vowel inventory, its low /a/ is perceptually similar to English /æ/, /ɒ/, /ʌ/ with potential consequences on discrimination in both inexperienced and experienced groups. Vowel inventory mismatches in L1-L2 as well as TL specific coarticulatory patterns influence L2 perception.

4 Vowel perception in cross-variety studies

While in § 2 and § 3 general aspects of vowel perception were presented, this section addresses specific findings in cross-variety studies focusing on particular aspects of L2 perception for Spanish listeners as well as on how the native dialect influences L2 perception. The studies reviewed next are grouped under cross-linguistic, cross-dialectal perception and L1 dialect effects in L2 perception.

4.1 Cross-linguistic and L2 vowel perception

Perception is language specific, that is, native speakers have refined their perception to recognize automatically the contrasts in their language. Since phones in languages differ in many ways (e.g. spectrally, temporally, phonemic or allophonic status, coarticulatory patterns), perceptual strategies that listeners adopt also vary cross-linguistically and these native language strategies influence L2 and cross-linguistic perception. Studies using multiple groups of L2 learners that have various L1’s often show very different perceptual patterns with L2 that are linked to L1 (e.g. Cutler et al. 2005; Flege, 1995; Flege, Bohn, & Jang, 1997; McAllister, Flege, & Piske, 2002).

For example, in McAllister et al. (2002), advanced learners with different L1s (Estonian, English, and Spanish) were tested on how they perceived the L2 Swedish quantity distinctions. The hypothesis investigated was that perceptual performance with L2 quantity is influenced by the presence of temporal distinctions in L1. Three L1s were selected based on the importance of quantity (vowel and consonant duration) given in the phonetic and phonological system of each of these languages. L1 Estonian has such distinctions; thus Estonian listeners used temporal cues extensively and showed native-like performance for L2 Swedish vowels. L1 English has some long-short (tense-lax) distinctions, so native English listeners performed well, but not as well as those with L1 Estonian. In L1 Spanish, however, duration has no phonological function, and L1 Spanish listeners made little use of the durational cues when identifying L2 Swedish vowels. As
such, the presence versus the absence of the duration feature in the native language influences how L2 duration contrasts are perceived. McAllister et al.’s study isolated the L1 influence in L2 perception by contrasting the perceptual behaviour of learners with native languages that exploit vowel duration differently.

As an alternative to using several groups of learners with different linguistic backgrounds, Fox, Flege, and Munro (1995) adopted a statistical technique, namely, a multidimensional scaling analysis, to highlight specific perceptual strategies that L2 listeners use. Based on dissimilarity ratings for pairs of Spanish /i, e, a/ and English /i, I, e, E, O, V, A/ vowels, they showed that English and L2 Spanish listeners perceive English vowels differently. Specifically, whereas English monolinguals use 3 underlying dimensions to categorize vowels (height, backness and central/non central distinctions), Spanish listeners use only 2 dimensions (height and proximity to a native prototype vowel). Moreover, vowel height tends to be strongly correlated with duration for (American) English monolinguals but not for Spanish native speakers. The multidimensional scaling analysis also provides information on the weight each dimension (cue) is given. Namely, it seems that height is weighted more than front-back (Fox, Flege, & Munro, 1995, p. 2545), and duration is weighted more when spectral cues are not readily processed (also see Bohn, 1995). Thus, L2 perception is influenced not only by differences in the perceptual spaces but also by different strategies of weighting acoustic cues. Such strategies of cue weighting vary not only for listeners with different native languages but also for listeners from distinct dialects of a language. This issue is addressed in the next section.

4.2 Cross-dialectal vowel perception

Several studies showed that the importance listeners from different L1 varieties attach to cues may also vary cross dialectally. Escudero and Boersma (2004) found dialectal differences between Scottish and Southern English in the perception of the contrast between high front vowels. Whereas Scottish English listeners favored spectral cues, Southern English listeners perceived the contrast based on a combination of spectral and temporal cues. A similar finding is reported for the French vowels /o-ʃ/ and /A-ə/ in two dialects, Standard and Swiss French (Miller & Grosjean, 1997). In contrast with Standard French that uses mainly spectral cues, in Swiss French duration is given a more important weight in vowel identification. Some distinctions in perception were also reported between American and Australian English (Cutler et al., 2005) for
/A-Ç/ and /Q-E/ in the use of duration and tenseness cues. Thus, cross-dialectal differences are reflected in different perceptual strategies of weighting spectral and temporal cues. These cross-dialectal differences in the perceptual strategies may have repercussions on the processing of non-native phones, which is the question at the centre of the present study.

4.3 Native dialect effects in non-native perception

A small number of studies have shown that learners with different native dialects perceive L2 sounds differently. Given the relevance for the present research, these studies are discussed next, emphasizing findings on Spanish dialects and English /Q, √, A/ vowels.

In order to highlight native dialect effects in non-native perception, Holden and Nearey (1986) first compared L1 perception and production in three Russian varieties that have identical phonemic inventories (/i, Æ, e, a, o, u/). Using a forced identification task with synthetic stimuli, the authors mapped the perceptual vowel spaces in each variety and found that there are significant differences in the distribution of /Æe, o/ across the three dialects. Then they further tested how English /√/ was perceived by listeners from each dialect with the purpose of identifying its assimilation patterns to Russian vowels. A forced identification task with real English stimuli showed that L2 listeners perceive the non-native vowel /√/ as /a, o/, or /e/ depending on their native variety and on the extension of area occupied by the native vowel to which /√/ was assimilated. As such, native perception was shown to influence how non-native sounds are mapped.21

Morrison (2008a) compared non-native perception in Mexican and Peninsular Spanish and found differences in the identification of the Canadian English front high and mid vowels /i, l, E/ and showed that listeners from these two dialects map vowels differently. Specifically, a synthetic continuum which allowed for English /i, l, e, E/ and Spanish /i, e, ei/ responses was created. The main difference between the performance of monolingual listeners with distinct dialectal

21 Acoustically L1 vowels do not differ much across the dialects, however, the authors cite previous articulatory data that corroborate the perceptual differences. In other words, there might be a dialectal preference for specific articulatory settings, which does not show acoustically but is reflected in perception.
background was found for English /l/ which was identified as /i/ or /e/ by Mexican Spanish listeners but mostly as /e/ by Peninsular Spanish listeners. Different assimilation patterns for English /l/ were thus proposed, namely, a category-goodness assimilation for the Mexican Spanish group and a two-category assimilation for the Peninsular Spanish group. As such, the perception of the /i - l/ contrast may be more problematic for Mexican Spanish learners than for Peninsular Spanish learners for which there is a better spectral separation.

Finally, Guitart (1996, citing data from unavailable conference presentations by Valle 1995, 1996) discusses native dialect effects in non-native perception of English /ɪ/, ɐ, ɔ/. Specifically, he reports Valle’s perceptual experiment in which two groups of Peninsular and Puerto Ricans listeners were asked to identify American English /ɪ/, ɐ, ɔ/ presented in CVC nonce words as one of the Spanish vowels /a, o/. The English /ɪ/ was identified as [a] in 83% of cases by Peninsular and as [o] in 71% of the cases by Puerto Ricans listeners. Guitart (1985, 1988, 1996) also found a similar differential substitution for inexperienced groups of learners with heterogeneous dialectal backgrounds (1985) and with Puerto Rican background (1988, 1996). It is not clear whether dialect was controlled for in these studies. Nevertheless, this differential substitution of /ɪ/ is given a phonological explanation based on contrastive features. Whereas English vowels are specified for four features, L2 listeners use only three features and ignore the fourth one in perceiving the L2 vowel. According to Guitart (1985), L2 listeners create a feature hierarchy depending on the perceived acoustic saliency of each feature. Thus, when perceiving TL /ɪ/,

some listeners ignore its [-low] feature and perceive it as an /a/ while others ignore the [-round] feature and perceive it as an /o/. Although no explanation is given for the preference of one set of features over another, nor for the dialect role in such preference, Guitart’s studies corroborate findings on the categorization of the L2 /ɪ/ as /a, o/ (Holden & Nearey, 1986) and on the possible L1 dialect effects in L2 perception (Morrison, 2008). Moreover, Guitart (1996) cites Valle’s findings on L1 varieties (Peninsular and Caribbean) and L2 vowels that are analyzed in the present research.

To sum up, several approaches that illustrate how perception works for listeners whose language or dialect is different from the TL were discussed in this section, from studies that contrast L2 performance of multiple L1 groups of learners (McAllister et al., 2002), statistical modeling of
dissimilarity ratings (Fox et al., 1995), and analyses of cue weighting based on identification
tasks (Escudero & Boersma, 2004; Miller & Grosjean, 1997), to plotting of perceptual maps
(Holden & Nearey, 1986; Morrison, 2008) and phonological analyses (Guitart, 1985, 1988,
1996). Whereas the studies reviewed here do not present exhaustively issues in cross-variety
perception, their findings are relevant for the present research in terms of varieties analyzed,
methods used and, moreover, in making predictions about native listeners from different Spanish
dialects and their perception of the English low and back vowels.

5 The current study

The main objective of the current study is to show whether the native dialect influences how
non-native sounds are perceived. Thus, two groups of Spanish native speakers from Spain (PS)
and Cuba (CS), each including two subgroups of learners and non-learners of English, were
tested in an AX discrimination task on the Canadian English contrasts /\text{A}-\text{\AA}/, /\text{Q}-\text{\AA}/ and /\text{Q}-\text{A}/ (see
Chapter 2, § 2.2.2). Based on the research reviewed in § 3 and § 4, the present study makes
specific predictions regarding the influence that the contrast inherent difficulty, the experience
with L2 contrasts and the native variety have in non-native perception of the TL /\text{Q}, \text{\AA/}, \text{\AA/} vowels.
The next sections present in detail the hypotheses, the results, hypothesis evaluation and
conclusion.

5.1 Hypotheses

In § 3.1 and § 3.2, I laid out evidence showing that among the English vowels, low and mid back
/\text{A}, \text{\AA/}, \text{\AA/} receive low identification scores from both native and non-native listeners. The same
vowels occur in contrasts that non-native speakers discriminate poorly. Moreover, contrasts
involving English /\text{Q}/ and /\text{A}, \text{\AA}/ are also reported to be difficult for Spanish listeners because of
the perceived similarity of these vowels to Spanish /\text{a}/ (Flege, Fox, & Munro, 1994). To test the
contrast difficulty of the English low and mid back vowels, I hypothesized that these contrasts
can be hierarchically ordered, with /\text{A}-\text{\AA}/ as the most difficult, followed by /\text{\AA/}-\text{\AA/} and /\text{\AA/}-\text{\AA/}
(hypothesis 1). This hierarchical pattern, reflected in discrimination error rates, can be observed

22 The following sections reproduce in part Marinescu’s (2011) study.
for all groups of listeners tested: Cuban (CS) and Peninsular (PS) Spanish learners and non-learners as well as for English native speakers (the control group).

In § 3.3, it was pointed out that the discrimination of L2 contrasts improves with experience in L2, and that improvement may be slower with some difficult L2 contrasts. Based on the hierarchy of difficulty proposed in hypothesis 1, the second hypothesis addresses the role of experience with the L2 contrasts in discrimination performance. It is expected that experience with L2 contrasts will determine lower error rates for the learner groups especially with the easy contrast /Q-A/ whereas for /V-Q/ and /A-V/ contrasts the error rates will drop less dramatically in the advanced group as compared to the monolingual group (hypothesis 2). The control group of English native speakers is expected to have excellent discrimination scores.

Given that English /V/ is reportedly substituted differently by listeners from distinct L1 dialects (see § 4.3) and that back vowels are perceptually confusable (see § 3.2), I specifically investigate whether the /A-V/ contrast is perceptually more difficult for one dialectal group than for the other. I assume that in processing this contrast, listeners are likely to use categories situated in the low and back perceptual space of L1, that is, /a/ and /o/. Given the phonetic differences between the [o]'s in these dialects ([o] is 10% more fronted in PS than in CS, see Chapter 3), the location of the /a-o/ perceptual boundary is likely to be more fronted in PS than in CS. The target /A-V/ vowels are thus closer to the PS /a-o/ boundary than to the CS /a-o/ boundary. Since perception is more sensitive at detecting differences in sounds in areas near a category boundary (Strange, 1995), target vowels that are closer to the L1 boundary are better discriminated. Thus, hypothesis 3 states that there will be different cross-dialectal error patterns in the discrimination of the L2 /A-V/ contrast. Specifically, Cuban listeners are more likely to err with the /A-V/ contrast, as for them the boundary of this contrast is far from their (posterior) L1 /a-o/ boundary. Conversely, Peninsular listeners tend to have a lower error rate with the /A-V/ contrast because its boundary is closer to the /a-o/ boundary their perception is attuned to.

In sum, the following hypotheses are tested in this perceptual discrimination task:
• H1 - contrast difficulty: the hierarchy of contrast difficulty is, from difficult to easy, /A-/√/, /√-Q/, /Q-A/ and is reflected in error rates. Qualitatively, all groups obtain the same hierarchy of difficulty.

• H2 - experience with L2 contrasts: non-learners have higher error rates than L2 learners. English native listeners have the lowest error rates.

• H3 - L1 dialect effect in L2 perception: different discrimination scores for the /A-/√/ contrast indicate that L2 perception is different for listeners with distinct L1 dialects, specifically, CS will show higher error rates than PS on this particular L2 contrast.

5.2 Results

The full presentation of the methodology, motivation for the task selection, stimuli and participants is detailed in Chapter 2, § 2.2. In brief, a two-speaker AX discrimination task with 162 trials was presented to 39 native Spanish listeners and 4 native English listeners. The stimuli were pairs of real English C1VC2 words including one of the vowels of interest, /Q, √, A/, and were separated by an ISI of 1 second. The independent variables were the contrast (/A-/√/, /√-Q/, /Q-A/), the L2 experience (advanced learners, non-learners), and the native dialect (PS, CS) and the dependent variable was the error rate.

The number of incorrect answers for the ‘different’ pairs was counted for each contrast and subgroup and converted into error rates. Participants’ mean error rates and standard deviations are reported for each contrast and subgroup in Table 4.2.

23 There were 24 ‘different’ pairs for each contrast. Error rates were calculated as percentages from the number of errors on ‘different’ pairs out of a maximum of 24.
Table 4.2

Mean error rates and standard deviations for each contrast and group

<table>
<thead>
<tr>
<th>Participant groups</th>
<th>n</th>
<th>/A-Ã/</th>
<th></th>
<th>/Ã-Q/</th>
<th></th>
<th>/Q-A/</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>error rate</td>
<td>SD</td>
<td>error rate</td>
<td>SD</td>
<td>error rate</td>
<td>SD</td>
</tr>
<tr>
<td>PS non-learners</td>
<td>10</td>
<td>44.5</td>
<td>18.8</td>
<td>42.0</td>
<td>19.1</td>
<td>11.2</td>
<td>10.0</td>
</tr>
<tr>
<td>PS advanced</td>
<td>10</td>
<td>13.7</td>
<td>10.5</td>
<td>13.3</td>
<td>11</td>
<td>2.9</td>
<td>4.8</td>
</tr>
<tr>
<td>Peninsular total</td>
<td>20</td>
<td>29.1</td>
<td>21.7</td>
<td>27.7</td>
<td>21.2</td>
<td>7.0</td>
<td>8.7</td>
</tr>
<tr>
<td>CS non-learners</td>
<td>10</td>
<td>35.8</td>
<td>18.2</td>
<td>42.9</td>
<td>16.4</td>
<td>12</td>
<td>16.9</td>
</tr>
<tr>
<td>CS advanced</td>
<td>9</td>
<td>37.9</td>
<td>21.7</td>
<td>10.1</td>
<td>9.5</td>
<td>1.8</td>
<td>3.0</td>
</tr>
<tr>
<td>Cuban total</td>
<td>19</td>
<td>36.8</td>
<td>19.8</td>
<td>27.4</td>
<td>21.4</td>
<td>7.2</td>
<td>13.2</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>32.9</td>
<td>20.8</td>
<td>27.5</td>
<td>21.0</td>
<td>7.1</td>
<td>11.0</td>
</tr>
</tbody>
</table>

Overall, the /A-Ã/ contrast had the highest error rates (32.9%), followed by /Ã-Q/ (27.5%) and /Q-A/ (7.1%). Examining discrimination scores for the experience factor, learners have better scores than non-learners for all contrasts, with the notable exception of the CS advanced groups whose mean error rate (37.9%) surpasses that of the non-learners (35.8%). With respect to the dialect factor, error rates are very similar for /Ã-Q/ (27.7 versus 27.4 %) and /Q-A/ (7 versus 7.2 %) but they differ much more for /A-Ã/ (29.1 versus 36.8 %). The four English native listeners (controls) were very accurate in discriminating the contrasts and they were not included in Table 4.2 or in the quantitative analysis. Pooled together, they made only 17 errors out of 648 trials, and all but one errors was either with the distracters or with the ‘same’ pairs. The fact that their error rate was very low (0.3%) confirmed that the stimuli and the testing procedure were properly designed. In the following sections the results of quantitative analyses are presented.

5.2.1 Contrast difficulty

While overall mean error rates shown in Table 4.1 suggest that /A-Ã/ was the most difficult contrast followed by /Ã-Q/ and /Q-A/ in both dialectal groups, a closer look at the scores for each dialect and experience level (Figure 4.1) shows that in the PS groups (learners and non-learners) the error rates for both /A-Ã/ and /Ã-Q/ contrasts are similarly high. In the CS non-learner group, the differences between these two contrasts are greater, /Ã-Q/ has the highest error rate in the CS monolingual and /A-Ã/ is notably high in the CS advanced group. A one-way repeated measures ANOVA with these three contrasts returned significant values ($F(2, 37) = 58.9, p < .001$). Pairwise comparisons show that error rates with the /Q-A/ contrast are significantly different
from each of the /A-/ and /Q-/ contrasts. The hierarchy of difficulty is thus partially supported, as only the /Q-A/ contrast, separated by a larger perceptual distance, is confirmed to be the easiest of all, whereas the /A-/ and /Q-/ contrasts are both similarly difficult.

Figure 4.1
Mean error rates in the discrimination task per groups and contrasts

5.2.2 Experience with the L2
To test the effects of experience with the L2 contrasts, repeated-measures ANOVA with contrasts (between) and experience (within) was performed. As expected, the main effect of experience is significant ($F(3, 35) = 8.596, p < .001$), advanced learners discriminate the English contrasts better than non-learners (Figure 4.1). Post hoc tests revealed that there were significant differences in discrimination between advanced and non-learners within the PS group ($p < .001$) but not within the CS group. The interaction between experience and dialect was not significant ($F(1, 35) = 1.46, p = .235$).

Given the high error rates for the /A-/ across subgroups, a 2x2 ANOVA with dialect and experience was further performed for this difficult contrast only (on arcsine-transformed data).
There were significant differences between PS advanced and PS non-learners for the /ʌ-ʌ/ contrast \((F(1, 35) = 6.537, p = .015)\) but no significant differences between the CS subgroups, for which the mean error rate is higher in the advanced group (37.9%) than in the non-learner group (35.8%). Apparently, experience with the L2 /ʌ-ʌ/ contrast caused no discrimination improvement, as CS advanced learners failed to follow the pattern found for the other contrasts, for which the error rates decreased considerably as compared to the non-learners’ group. A significant interaction between dialect and experience was also found \((F(1, 35) = 6.932, p = .013)\) and is further discussed in the next section.

5.2.3 L1 dialect effect in L2 perception

A repeated-measures ANOVA with contrasts (within) and dialect (between) showed that, overall, there are non-significant differences between dialectal groups. However, a separate 2x2 analysis of variance performed only on the /ʌ-ʌ/ contrast returned the above-mentioned significant interaction between dialect and experience \((F(1, 35) = 6.932, p = .013)\); the advanced PS group performed better than the advanced CS group \((F(1, 35) = 7.407, p = .01)\). If native dialect shaped indeed non-native perception, similar behaviour would be expected both for learners and non-learners. However, this was not the case in the non-learner groups, for which there was no significant main effect of dialect, as their mean error rates and variability were comparable (cf. Table 4.2). Therefore, the differences between dialectal groups can be pinpointed to different L2 learning experiences in the learner groups rather than to the dialect solely. The significant interaction between dialect and experience for /ʌ-ʌ/ provides further support to the fact.

An aspect worth noticing is the great amount of individual variability within the CS advanced group for /ʌ-ʌ/ (Figure 4.2). In this group, performance is comparable to those in the non-learner groups, as error rates range from 0 to more than 60%.
According to the background questionnaire elicited at the time of testing (see Appendix 4), almost all advanced learners in both dialectal groups had or were pursuing a degree in English. Nevertheless, different levels of proficiency may have been the source of the great variability in performance. To test these effects, proficiency scores were compiled from a vocabulary test elicited as part of an L2 production experiment (see Chapter 5) that accompanied the current perceptual task. Participants had to answer 58 vocabulary questions and, based on the numbers of correct answers without any prompting from the researcher, a score was assigned to each one. The proficiency scores were 66.6% (sd 10.5) for PS and 65.7% (sd 10.6) for CS. For both dialectal groups of advanced learners, Pearson correlation coefficients were computed to assess the relationship between proficiency scores in the vocabulary test and accuracy in discriminating the L2 contrasts. The correlation between the two variables was not significant in either group: PS ($r = -.327, n = 10, p = .357$), CS ($r = -.144, n = 9, p = .711$) suggesting that proficiency alone was not responsible for the great variability in the perceptual performance.
5.3 Hypothesis evaluation and discussion

The results presented in the previous sections addressed L2 perception in relation to three factors, contrast difficulty, experience with L2, and the native dialect of the participants. Hypothesis 1 proposed a hierarchy of difficulty in which error rates decrease from /A-/ to /Q-Ã/ and /Q-A/. The rationale for this hierarchy was the inherent difficulty of contrasts involving the English low and mid back vowels as well as the perceptual behaviour reported in previous studies with L2 Spanish listeners. Although data pooled together for all participants show that discrimination errors decreased as predicted for these three contrasts (32.9, 27.5, 7.1 %), a statistically significant difference was found only between /Q-A/ and /A-/ and between /Q-A/ and /Q-/ contrasts. Judging by the relative position of these vowels in the acoustic-perceptual space, it is likely that the perceptual distance between vowels affected contrast discrimination. The distance in the /Q-A/ contrast involving a front and a back vowel is large; therefore this contrast has the best discrimination scores. On the contrary, the distance between the back vowels in /A-/ is small, so this contrast is difficult. The /Q-/ contrast also generated high error rates. Given that mid (and high) back vowels tend to be centralized in alveolar contexts (Strange et al., 2007), the distance between [Ã] and the low front /Q/ is reduced, rendering the contrast difficult. The mid back (or centralized) vowel /Ã/ received poor scores in the confusion matrix even from native listeners (Syrdal & Gopal, 1986; Cutler et al., 2005). In the present study the vowel /Ã/ appears in contrasts that rank high in the difficulty hierarchy, indicating that L2 listeners confuse it easily. The fact that /A-/ and /Q-Ã/ contrasts were difficult to discriminate corroborates Flege’s (1995) study in which Spanish learners of English had high error rates with these same contrasts. It is also worth noting that while vowels in these two pairs are spectrally close to each other, they differ in tenseness (duration), which, nevertheless, does not make them perceptually easier to discriminate. The reason is likely to be found in the cross-linguistic differences of weighting spectral and duration cues in English and Spanish. A detailed discussion of this issue can be found in Chapter 6, once all the necessary data will have been presented.

---

24 At this point, I am referring to the canonical description reported for these English vowels in the literature: /Q/ is low front unrounded lax, /Ã/ is mid back, possibly centralized, unrounded lax and /A/ is low back rounded tense. Acoustic measurements (and their transformation into perceptual distances) will be provided in Chapter 5.
With respect to Hypothesis 2, experience with L2 contrasts benefited the L2 learners in both dialectal groups, except for the /\text{A-}\text{Ã}/ contrast in the CS advanced learner group. Considerable improvement in L2 perception can be observed in Table 4.3, which presents how much discrimination error rates decreased with L2 experience in each dialectal group. The easy /\text{Q-}\text{A}/ contrast shows the greatest improvement in both dialectal groups; the error rates drop more in CS than in PS for the /\text{\text{-Q/ and Q-}}\text{A/} contrasts, but for /\text{A-}\text{\text{V/}} there is actually an increase in the error rates from CS non-learners to learners. This is a surprising effect. To explain it, I speculate that L2 learners have either adopted an interlanguage perceptual strategy that does not help them distinguish between two TL vowels that are spectrally close and differ in duration as well, or that transfer from L1 is still in effect, i.e., L2 learners use only spectral cues as in their L1 not both spectral and duration cues as required in the TL. Furthermore, the boundaries between these vowels may fall in an area of the perceptual space for which the learners’ perception is less sensitive, as it is discussed in the next paragraph dedicated to L1 dialect effects in L2 perception. To conclude, results on experience with L2 confirm previous findings (Lambacher et al., 2005; Levy & Strange, 2008) that some perceptual contrasts may take longer to master and may require specific training.

Table 4.3

<table>
<thead>
<tr>
<th>Decrease (percentages) in the discrimination error rates with L2 experience</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vowel contrasts</strong></td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>PS</td>
</tr>
<tr>
<td>CS</td>
</tr>
</tbody>
</table>

Hypothesis 3 set out to test whether the native dialect influences non-native perception and proposed that such influence will show in specific patterns of discrimination errors across dialects, specifically more errors with back vowels for CS than for PS participants.

Advanced Cubans had significantly more errors than advanced Peninsulars with the /\text{A-}\text{\text{V/}} contrast, but performed comparably with the other two contrasts /\text{\text{-Q/ and Q-}}\text{A/}. It is known that perception is sharper at category boundaries and less accurate far from these boundaries (Strange, 1995). In other words, between-category contrasts are better discriminated than within-category contrasts. The fact that the advanced Cuban group obtained a higher error rate with the L2 /\text{A-}\text{\text{V/}} suggests that it represents a within-category contrast, which is farther from the
interlanguage boundary they may have established for this contrast (L1 /a-o/ boundary). On the other hand, the good performance that the Peninsular learner group has with this contrast indicates that their interlanguage boundary of the /A-Ã/ contrast is closer to the L1 /a-o/ boundary or it may even be the case that /A-Ã/ is already a between-category contrast, that is, they have created separate categories for these L2 vowels. Thus, learners from distinct native dialects map the L2 /A, Ñ/ vowels differently. Additionally, greater variability in discrimination accuracy points to a fuzzier boundary of the contrast for the Cuban advanced group. Learners rely on their native /a-o/ boundary to discriminate the /A-Ã/ contrast, but since their native vowels differ cross-dialectally, the outcome of their perceptual behaviour is different. From the perspective of the Perceptual Assimilation Model in L2 (Best & Tyler, 2007), the /A-Ã/ contrast is assimilated as a two-category contrast by the PS group but as a category-goodness contrast by the CS group. The fact that PS learners adopt a two-category assimilation strategy is consistent with Morrison’s (2008) findings on the /i-I/ contrast for a group of Peninsular listeners suggesting that L2 learners from PS tend to create perceptual categories that have good spectral separation.

5.4 Conclusions

This study proposed a hierarchy of contrast difficulty and predicted that contrasts involving two back vowels would be more difficult to discriminate than contrast involving a front and a back vowel. Results showed that discrimination was more accurate only for the ‘easy’ contrast /Q-A/, but there were no significant differences in discrimination for the ‘difficult’ contrasts /A-Ã/ and /Ã-Q/. The overall discrimination accuracy was above chance, and similar error patterns emerged for all groups, except for CS advanced. The hierarchy of difficulty obtained for L2 matches the pattern of difficulty reported for L1 (Syrdal & Gopal, 1986) and supports the finding that pairs involving front and back vowels are inherently easier to perceive as compared to contrasts involving two back vowels.

Experience with L2 contrasts had an effect on the listeners’ performance as advanced learners discriminated contrasts better than the non-learners. The contrast difficulty was signaled by higher error rates in the monolingual groups as compared to the advanced groups. For the
‘difficult’ contrast /A-\textipa{v}/ the error rates dropped 69% in non-learners vs. advanced learners in the PS group; for the /\textipa{v}-Q/ contrast, the drop was 68% in the PS group and 76% in the CS group. Finally, the drop was 74% in the PS and 85% in the CS group for the ‘easy’ contrast /Q-A/, which almost half of all listeners discriminated correctly. Thus, L2 experience generally, but not always, determines a lower error rate with L2 contrasts.

Overall, there is little evidence for L1 dialect effect in L2. However, an unexpected finding was the poor performance on the /A-\textipa{v}/ contrast in the CS advanced group. Perception is the same for non-learners in the PS and CS groups, but it diverges for the learner groups as they adopt different paths when processing this non-native contrast. Since the native dialect and proficiency effects are inconclusive, other factors like the learning experience, the input, the amount of L2 use, knowledge of other languages need to be taken into consideration. Based on the background questionnaire, participants in this study differed with respect to the other languages they knew, to the L2 input and immersion. The Peninsular group had more exposure to other languages (French, Catalan, German), had spent various amounts of time (1-8 months) in English-speaking environments, including Canada, and had received native input at some point during their studies. The Cuban group had less experience with other languages and received native input during their studies only via recorded materials. At the time of the experiment, however, the Cubans had more exposure to the TL due to their everyday contact with native and non-native English speakers, which is likely to have contributed to their highly variable performance.
Chapter 5
L2 production

1 Introduction

As I mentioned in the opening of Chapter 1, the exploration of TL English /Q, \, / vowels as produced and perceived by Spanish native speakers was sparked by the empirical observation of a case of differential substitution of the English /\ / vowel as [a] or [o] by Spanish native speakers and a native dialect effect was suggested to be the cause (Valle, 1996, as cited in Guitart, 1996). Variable realizations in L2 have also been observed for English /Q, / vowels. It has been demonstrated that speakers’ native variety influences their L2 production (e.g. Flege, 2002; Ingram & Park, 1997; Hammarberg, 1997; Munro, Flege, & MacKay, 1995). L2 learners use their native categories as a reference when processing and producing TL sounds, as transfer of the native articulatory settings in the interlanguage may be present even at advanced stages of acquisition. Since L2 learners from PS and CS are likely to rely on their L1 categories when perceiving and producing L2 English vowels, I hypothesize that there will be differences between the L2 vowels produced by PS learners and those produced by CS learners, reflecting the differences between the native vowels in PS and CS. As such, if in the PS dialect vowels tend to be shorter and more fronted and there is more variability in the realization of the front vowels, then (all or specific) L2 vowels are realized as more fronted, shorter and possibly there is more confusion with the L2 front vowels as compared to the TL. Similar reasoning for L2 vowels produced by learners from the CS group predicts more posterior L2 vowels and more confusion with the back vowels, given that native vowels in CS are more posterior and there is more variability for the back vowels.

So far, explorations of the cross-dialectal differences between L1 vowels in PS and CS (Chapter 3) and native dialect effects in L2 perception (Chapter 4) have revealed that there are some differences in the phonetic realization of L1 vowels ([i, o] are more fronted in PS versus CS, front vowels are more variable in PS, back vowels are more variable in CS, [e, a, o] are shorter in PS versus CS). The perceptual experiment described in Chapter 4 showed no clear native dialect effects in L2, but confirmed that non-native listeners have perceptual difficulties especially with
the /Q-/A/ and /Q-/ contrast, and that L2 advanced learners from CS persist in their high error rates with the /Q-/ contrast. The present chapter further explores the role of the native dialect in L2 production. Specifically, the aim here is to obtain TL /Q, \( \sqrt{ } \), \( \check{} \) vowels from advanced PS and CS learners of English and, based on acoustic data, to compare how the two dialectal groups of learners differ from each other in the realization of these TL vowels and whether the native dialect shapes L2 vowel realizations. If the differences between L2 vowels produced by PS and CS learners show the same tendencies found for the L1 vowels with respect to frontness, duration and variability, then the native dialect effect in L2 production will be supported.

The following sections of this chapter present three theoretical models of acquisition (§ 2), methods used to assess L2 production (§ 3), previous studies with relevance to dialect acquisition (§ 4), and the experiment on L2 production, including a section on Canadian English vowels (§ 5), hypotheses, results, and hypothesis evaluation (§ 6).

2 Theoretical approaches to SLA: three models

The native dialect effect in non-native vowel production is an understudied research area. Although some studies have addressed native dialect effects in L2 perception (Holden & Nearey, 1986; Guitart, 1996; Morrison, 2008), next to nothing has been said about potential differences in L2 production. To my knowledge, there are no studies that analyze native dialects as a source of differences in L2 vowel production. However, since many studies analyzed differences in L2 production of groups with various native languages (e.g. Ingram & Park, 1997; Flege, Bohn, & Jang, 1997; Bohn & Flege, 1992), their methodology and findings (e.g. measures used in comparisons of L2 and TL sounds, assimilation patterns, relation between acoustic and perceptual similarity) may inform the current research. Thus, the present study contributes an analysis of the influence of the native dialect in the L2 vowel production with a focus on the English /Q, \( \sqrt{ } \), \( \check{} \) vowels as produced by advanced learners having different native dialects. The

25 Branner’s (2002) study, though, looks at differential substitution of TL English interdental fricative by learners from different native French dialects.

26 The languages investigated were L1 Japanese and Korean, L2 Australian English.

27 The languages investigated were L1 German, Spanish, Mandarin, Korean, L2 American English.

28 The languages investigated were L1 German and L2 American English.
analysis of potential native dialect effects in L2 production uses concepts and considerations set forth in three well-known models of acquisition, namely, Flege’s Speech Learning Model (1987, 1995, 2003), the formal linguistic model proposed by Escudero and Boersma (2004) and the Perceptual Assimilation Model for L2 articulated by Best and Tyler (2007). These models are presented in the following sections.

2.1 The Speech Learning Model

The Speech Learning Model predicts the perceptual behaviour of L2 learners based on differences between L1 and L2 phonetic categories. The comparison unit is the diaphone, a pair of an L1 sound and the corresponding L2 sound. As Flege acknowledges, this is a working model developed from generalizations of experimental data that address specific stages only, not the whole acquisition process, as the SLM accounts principally for the later stages of acquisition. Although no theoretical formalization has been formulated for the underlying mechanisms of the categorization processes, as Escudero and Boersma (2004) point out, the model provides valuable proposals concerning the mapping of L1 onto L2 sounds.

The basic assumptions of this model are that a new phonetic category is likely to be created if the L2 sound is perceived as new/different from the already existing L1 categories (that are transferred to the interlanguage) and that an already existing phonetic category is going to be reused if the L2 sound is perceived as similar to an L1 sound. On the other hand, similar sounds pose more perceptual problems: although they share features with the L1 sounds, they are not completely identical to those L1 sounds. L2 learners may misperceive or ignore some of the differentiating features of the L2 sounds and assimilate them to L1 phonetic categories. In Flege’s terminology, this is the mechanism of equivalence classification that explains the inaccurate acquisition of similar L2 categories. Apparently, one drawback of the model is that no production factors are considered, as the premise is that accurate perception implies accurate production, as well. The other drawback is the difficulty of quantifying the construct of similar categories. However, in an overview of the research on cross-language phonetic similarity of vowels, Strange (2007) discusses several empirical approaches to assessing similarity in L2 production like acoustic comparisons, statistical techniques of modeling interlanguage categorization, as well as perceptual assimilation tasks accompanied by accentedness and/or intelligibility ratings. Thus, such measurements provide qualitative and quantitative information.
on the degree of acoustic and/or perceptual similarity between diaphones and the amount of overlap between cross-language categories, making the Speech Learning Model a readily testable one.

2.2 The formal linguistic model of categorization

Like Flege’s Speech Learning Model, the formal linguistic model of categorization proposed by Escudero and Boersma (2004) is based on perception and articulated within the framework of the probabilistic version of the Optimality Theory (Boersma, 1998, as cited in Escudero & Boersma, 2004). The input, i.e. an auditory event, is processed by a perception grammar, which determines an optimal output, i.e. a phonological category. Continuous families of phonetic constraints regulate the mapping of an auditory event onto a discrete phonological category. The authors illustrate this model with the specific case of L2 acquisition of the /i/-l/ contrast in two target dialects of English (from Southern England and Scotland) by Spanish native speakers. Thus, the study aims to model the acquisition process and to show that the environmental input shapes the perceptual learning, as L2 learners adapt their perceptual strategies to optimally match specific traits of the input. In this particular case, the input is characterized by reliance on durational and spectral cues in the Southern English dialect as opposed to reliance on mainly spectral cues in Scottish English. In order to accurately perceive TL /i, l/, L2 learners immersed in the corresponding target dialect, may reuse L1 categories (this representing an instance of L1 transfer), shift category boundaries or create new categories. For example, L2 learners in Southern England initially assimilate both TL /i, l/ to their native /i/, and later they create a duration category /i^\prime/, i/ to approximate the TL /i-l/ contrast. In the case of L2 learners in Scotland, initially they assimilate the TL /i, l/ to their native categories /i, e/ and later they shift the boundary between these two vowels to F1 values that approximate the target boundary between /i, l/. In particular, creating a short-long category distinction is considered to reflect a universal preference for learners’ use of duration as a means of implementing an L2 contrast rather than creating a spectral contrast (i.e. splitting an L1 category into two distinct categories

29 So, in perception, Scottish listeners rely mainly on spectral cues whereas Southern English listeners rely on spectral and duration cues. This perceptual pattern is apparent in production, too. Native English speakers produce the /i-l/ contrast with a large F1 difference (142 Hz) and a small duration difference (9 ms) in the Scottish dialect and with some F1 difference (45 Hz) and a large duration difference (45 ms) in the Southern English dialect.
that differ in F1/height). Under this approach, long /a/ would map to the English /A/ versus short /a/ to the English /Æ, Q/.

Although this model focuses mainly on perceptual learning, it is relevant to the present experiment because it gives insight into a series of possible interlanguage processes of category reuse, category formation, and category split, which are likely to motivate and explain the patterns found in L2 production. Such constructs are particularly informative for this study for which mappings from fewer L1 categories (Spanish /a, o/) to multiple TL categories (English /A, Æ, Q/) are analyzed. Moreover, Escudero and Boersma’s study has shown that differences in the environmental input (i.e. two distinct English dialects) cause L2 learners to adopt different paths when acquiring the /i, I/ contrast leading to different endpoints (i.e. phonological categories). Thus, one can further speculate that differences in the starting points of the acquisition for learners with different L1 dialects (who start mapping TL vowels based on phonetically different L1 vowels) also result in differences in the endpoint.

2.3 The Perceptual Assimilation Model for L2

According to Best and Tyler (2007) the goal of the Perceptual Assimilation Model in L2 (PAM-L2) is to predict the development of the L2 learners’ interlanguage system addressing both the phonetic and phonological levels. The comparison unit is not the SLM’s diaphone (an L1-L2 pair of sounds) but rather an L2 minimal pair of sounds whose discrimination accuracy and assimilation to L1 sounds are assessed. The authors discuss four possible situations of L2 minimal contrasts that illustrate L2 learners’ perceptual behaviour and formulate predictions regarding discrimination accuracy.

The first case is a two-category (TC) assimilation, for which both phonological categories of an L2 contrast are perceived as equivalent to distinct L1 phonological categories. Alternately, one member of the L2 contrast is a good match for an L1 category but the other is uncategorized (no match in L1), and thus it represents a case of categorized-uncategorized assimilation. Crucially,

---

30 L2 production is not discussed in this study. My understanding is that it is assumed that accurate perception entails accurate L2 production as well.
the L1 and L2 phonetic and phonological categories involved in this type of assimilation are perceived as equivalent and discrimination is predicted to be very good.

The second case, category-goodness (CG) assimilation, occurs when both L2 phonological categories are perceived as equivalent to the same L1 phonological category but one is perceived as a better match than the other. Discrimination is predicted to be good as L2 learners are able to distinguish between minimal lexical pairs. Phonetically, the member of the L2 pair that matches better the L1 category is assimilated to a common L1/L2 phonetic category whereas for the member of the L2 pair that is perceived as more distant from the L1 category, a new phonetic category will be created.

The third case, single category (SC) assimilation, corresponds to the situation when both phonological categories of an L2 contrast are perceived as equivalent to the same L1 phonological category but are perceived as equally close or equally far from that L1 phoneme. In other words, such L2 phones are treated as homophones and thus they are difficult to discriminate. The authors hypothesize that a new phonetic category needs to be learned at least for one of the members of the pair before a phonological category can be created. Moreover, establishing a new phonetic category depends on detecting differences between minimally contrasting words. Thus, the richer a particular TL contrast in minimal pairs and the higher the frequency of such words, the higher the chances for the L2 learner to implement a new L2 category.

The fourth possibility is when the members of the TL contrast are perceived as not clearly belonging to any L1 category (Uncategorizable). Depending on how close to each other the members of the L2 pair are and how close to the L1 categories, one or two new L2 phonological categories may be learned. Discrimination is predicted to range from poor to very good. These predictions capture several instances of L2 learners’ perceptual behaviour and could account for patterns in L2 vowel production. In particular, specific perceptual assimilation patterns for L2 vowels may be paralleled in L2 production.

These three models discussed above differ from each other in several respects as they tap into different levels of representation. On the one hand, Flege’s discussion of phonetic category formation draws on the acoustic-phonetic properties of sounds and on the acoustic and perceptual similarities existing between sounds in learners’ L1 and the TL, as the SLM targets
acquisition for advanced learners as well as bilinguals. On the other hand, Escudero and Boersma’s discussion of categorization implies a more abstract level of L2 processing, uncovering specific interlanguage phenomena and being capable of modeling the entire acquisition process, not only specific stages. Finally, Best and Tyler’s PAM-L2 discusses perceptual learning of both phonetic and phonological categories in L2 and address the role of lexical factors in shaping L2 perception. The present exploration of L1 dialect effects in L2 production seeks to explain its findings from the perspective of each of these acquisition models.

3 Methods for assessing L2 production

The following sections discuss types of data commonly used in the evaluation of L2 production as well as previous work on acquisition involving dialects as either the native or the target variety. As summarized by Strange (2007), there are several methods of assessing cross-linguistic similarity of vowels, namely, acoustic comparisons, perceptual assimilation tasks accompanied by accentedness ratings, and/or intelligibility labeling as well as statistical techniques that compare vowel distributions in L1 and L2. Some of the studies that used these approaches are discussed next.

3.1 Acoustic comparisons

L2 vowel production accuracy can be assessed through direct acoustic measurements of formants and durations, which may then be compared against native vowel production; conclusions regarding the formation of phonetic categories can then be drawn. For example, in Bohn and Flege (1992), the authors’ aim was to test the effects of experience on the production of the new vowel /Q/ and of the similar vowels /E, I, i/ by German learners of English. The first experiment aimed to obtain and compare the native English /Q, E, I, i/ and native German /E, I, i/. Data from three groups of participants (10 native German experienced, 10 native German inexperienced and 10 native English controls) were elicited through a reading task. Formants F0 through F3 and durations were measured and subsequently normalized into Bark dimensions of vowel height (B1-B0) and frontness-backness (B2-B1) to eliminate gender differences. Ellipses enclosing data points obtained for each of the L1 vowels show that there was a considerable amount of overlap between the corresponding /E, I, i/ vowels between German and English, confirming the cross-linguistic similarity hypothesized for these vowels. In contrast, the ellipsis enclosing the English
/Q/ overlapped with no German category, as it only touched the ellipsis corresponding to 
German /a/ and /E/, thus suggesting that /Q/ represents a new category for the German learners.

In the second experiment, L2 English vowels produced by German experienced and 
inexperienced learners were compared to the same vowels as produced by native speakers. As 
expected, the L2 vowels /l, i/ produced by the experienced learners were closer to the target 
language vowels both spectrally and in duration as compared to those produced by inexperienced 
learners. With respect to the similar vowel /E/ and the new vowel /Q/, experienced, but not 
inexperienced, learners produced them accurately.

Acoustic comparisons of vowels in the F1 x F2 space across languages are quite useful and 
illustrative. Based on comparisons between realizations of vowels by native speakers of the 
languages being analyzed, similarity relationships can be highlighted and predictions concerning 
the difficulty of producing a vowel in a speaker’s L2 can be articulated. Moreover, comparisons 
between the target vowels and what L2 learners actually produce may corroborate predictions on 
similarity.

However, this approach is not always unproblematic. According to some studies (Strange, 2007; 
Flege, Bohn, & Jang, 1997), methodological problems like speaker normalization, phonetic 
context effects, speech rate effects, and cross-linguistic prosodic and phonotactic differences can 
be avoided through careful experimental design. For example, when diaphones are similar, 
researchers must ensure that vocal tract differences are filtered out while preserving other 
differences, including regional differences in the particular case of the present research. Thus, an 
adequate normalization procedure must be selected. Moreover, the phonetic context in which 
vowels appear is important in cross-linguistic comparisons, as it has been shown that, for 
instance, the alveolar context tends to produce a greater degree of vowel fronting, in American 
English than in Parisian French or Northern German (Strange, 2007). Therefore, comparisons of 
vowels in more neutral contexts (e.g. bilabial) are likely to avoid this problem. Speech rate also 
has to be kept constant across groups, as fast speech affects not only vowel duration but also its 
spectral properties, including target undershoot. Syllable structure differences between languages 
should also be taken into account. Although Bradlow (1995) demonstrated that comparing 
certain English vowels (/i, e, o, u/) in CVC contexts with corresponding Spanish vowels in 
CVCV contexts that are very frequent in Spanish but infrequent in English led to no apparent
differences, it is worth keeping in mind that this might not be true for the other vowels not investigated in her study or for the same vowels in other linguistic varieties. More subtle problems like transfer of L1 coarticulation patterns or specific strategies of using vowel quality and quantity into the interlanguage or language-specific formant movements or F0 characteristics may not be easily controlled and explained. Despite these problems that can be controlled, direct acoustic comparisons between L1 and L2 vowels are informative and readily obtained. As an alternative, accentedness and intelligibility judgments are used in assessing L2 production, which is discussed next.

3.2 Accentedness and intelligibility judgments

In order to corroborate the acoustic findings, accentedness and/or intelligibility judgments may be elicited from native listeners of the target language, who assign labels and category goodness scores to the vowels produced by L2 learners (e.g., Bohn & Flege, 1992; Flege, Bohn, and Jang, 1997). In many studies, such accentedness and/or intelligibility tasks have completely replaced acoustic comparisons (e.g. Flege, Munro, & MacKay, 1995, on L1 Italian L2 Canadian English). For instance, Flege, Bohn, and Jang (1997) had a group of native listeners of the target language (American English) label the /l, i, Q, E/ vowels as produced by learners of various L1’s (German, Spanish, Mandarin, and Korean). A different pattern of native speaker identification errors was highlighted for each group of learners and conclusions regarding the perceptual similarities between L1 and target language vowels were consequently drawn. Of interest here, the vowels produced by the Spanish group were misidentified as follows: /l/ was heard as /i/ and vice versa; /Q/ was heard as /A/ or as a vowel posterior to the target; whereas most of the /E/’s were identified accurately by the judges. Further exploration of the acoustic data through Euclidean distances showed that the Spanish learners produced a smaller spectral contrast for /l-i/ and a larger one for /Q-E/ as compared to English native speakers. These findings suggest that targets /l, i/ are mapped to Spanish native speakers’ /i/, target /Q/ is assimilated to native /a/, and /E/ might be similar to Spanish /e/. These patterns of assimilation are different from the other languages investigated in this study, supporting once again the effect of the native variety in L2 acquisition.
A problem with accentedness ratings is that the rater’s decision may not be determined solely by the realization of a single segment but rather by the combination of several cues that lead to the perception of foreign accent. To minimize such effects, category goodness ratings are elicited, which involve raters’ attention being directed towards a specific L2 segment and a comparison with its similarity to the hearer’s native counterpart (Bohn & Flege, 1992).

Many of the studies that have used both acoustic measurements and intelligibility ratings (e.g., Flege, Bohn, & Jang, 1997; Bohn & Flege, 1992) found a correlation between the results obtained through these two methods, thus suggesting that either approach is sufficient to confirm the similarity patterns predicted based on comparisons between diapophones as produced by native speakers of each variety involved in the study. Transformations of acoustic data into auditory and perceptual relevant data (like Barks, mels or log Hertz) add further information regarding the cross-linguistic patterns of perceptual similarity. In the present study, I adopted the method of directly comparing the acoustic realizations of the L2 vowels as produced by Cuban and Peninsular learners of English. These would provide an objective picture of the cross-dialectal effects in L2, unmediated by potentially subjective judge ratings. Having presented these approaches to analyzing L2 production, I focus next on reviewing studies that explored dialect acquisition and native dialect effects in L2 production.

4 Previous studies in second dialect production

Among the few studies that have addressed dialect acquisition and L2 and native dialect effects in L2, three are relevant to the current research and will be presented in detail in this section. Despite the fact that their goals, hypotheses, and linguistic varieties investigated are different from those of the current study, these studies are discussed here because of the particular relevance of the theoretical constructs presented (Escudero & Boersma, 2004), and of the experimental results on dialect acquisition of specific phonetic differences between vowels (Fox & Tevis McGory, 2007; Munro, Derwing, & Flege, 1999).

In Escudero and Boersma (2004; see also § 2.2), Spanish learners of two different English varieties, Scottish and Southern English, were tested via an identification task to demonstrate how different target dialect in the input influences L2 perception. It was shown that L2 learners adopt the perceptual strategies of cue weighting of the ambient speech community, namely L2 learners of Southern English rely both on spectral and duration cues just like the native speakers.
in that region whereas L2 learners of Scottish English match the native Scots’ reliance on spectral cues only. Only the perception of the /l-l/ contrast was investigated. This study also provides a theoretical framework for L2 categorization, bridging the gap between perception research and phonological theory.

A similar study is that of Fox and Tevis McGory (2007) that investigated the production and perception of Japanese learners of American English immersed for more than 2 years in two distinct regional varieties characterized by different vowels. Ten monophthongs were analyzed, namely /i, ɪ, e, ɛ, o, ɒ, ʌ, ʊ, ə, ʌ/. Standard American English (Ohio) and Southern American English (Alabama) vowels differ both spectrally and temporally. To compare them spectrally, F1 and F2 formants were measured at the 25%, 50% and 75% points through each vowel. The measure calculated for the group as a whole was the vector length (or spectral change, in other terminology), which was calculated as the median distance in Hertz between the 25% and 75% points in each vowel. All Southern American vowels (except /i/) had longer trajectories as compared to Standard American vowels and 6 out of 10 vowels were diphthongized. To compare vowel durations, ratios were calculated for tense-lax vowel pairs (/i-ɪ/, /e-ɛ/, /u-ʌ/ and /ɒ-ʌ/); smaller ratios were found between pairs in Southern American English indicating that the tense-lax distinction was weaker (i.e., tense and lax vowels were produced with similar durations in this variety) than in the Standard dialect. Based on these measurements, L2 production was then assessed. It was hypothesized that L2 learners would tend to follow the norm in their dialect of immersion, that is, Japanese-speaking learners in Alabama would produce vowels that match the Southern American target whereas learners in Ohio would produce vowels similar to those of Standard American. Interestingly, this was not the case, as both learner groups produced vowels more similar to those in the Standard American dialect. The authors concluded that learners used the Standard variety they heard in the media, which was likely the same they were initially exposed to in their ESL classes. Moreover, in a follow up perceptual identification task, both groups of learners obtained higher identification scores for vowels of the Standard variety than for the Southern American variety and the advantage of the learners immersed in the Southern variety in the identification of the Southern American vowels was minimal. Attitudes towards a non-standard variety, which may restrict its use, as well as the limited time participants in this group spent actually talking to speakers using this variety are put forward as an explanation.
To summarize the two studies discussed so far, Escudero and Boersma analyzed only the /I-i/ contrast through a perceptual identification task and found that Spanish learners adopted the perceptual strategies in the ambient speech variety, namely, spectral and temporal cue weighting in Southern English and spectral cue weighting in Scottish English. Fox and Tevis McGory analyzed four vowel contrasts through production and perception tasks and, in contrast, determined that Japanese learners favoured Standard American English, regardless of the ambient speech variety. The contradictory findings of these studies are likely attributable to different analyses of the perceptual data (reliance versus intelligibility scores) as well as to sociolinguistic issues regarding the varieties investigated (learners perceived the Southern American dialect as non-standard and thus resisted adopting its peculiar characteristics which was not the case with the learners of the European English varieties). Other factors that may be contributing to these divergent findings are the learning experience in the home country and differences in typological proximity between Japanese and Spanish versus English.

Another study involving acquisition and various dialects is that of Munro, Derwing, and Flege (1999), which analyzed the extent to which Canadian speakers living in Alabama acquired this American English variety. The analysis relies on accent judgments from native speakers of the Canadian and Alabaman varieties who listened to speech samples from three groups of 10 speakers each: Canadians in Alabama (CA), Canadians in Canada (CC) and Americans in Alabama (AA). Speech samples obtained from a picture description task were rated on a nine-step-scale ranging from ‘very Canadian’ (1) to ‘very American’ (9). The overall degree of accentedness (Experiments 1, 2) as well as particular phonetic characteristics (Experiment 3) were assessed. Results of Experiment 1 showed that 22 Canadian judges accurately assigned the CC and AA groups to the corresponding dialects, while rating the CA group at intermediate scores (4.0). This suggests that Canadians living in Alabama had adopted “an intermediate degree of American accent” (p. 385). In Experiment 2, the same speech samples were also rated by a group of 27 judges from the Alabaman variety. Once again, the CC and AA groups received scores towards the Canadian and, respectively, the American end on the scale, whereas the migrant group of CA received intermediate scores (5.8). However, further analysis of the ratings that the two groups of judges assigned to CA speakers showed that this group was perceived as more Canadian than Alabaman. This suggests that, although dialect acquisition did occur, native-dialect peculiarities were still present in the participants’ speech samples. In Experiment 3, two
specific phonetic characteristics, the mean articulation rate (syllables/second) and the realization 
of the /aj/ diphthong in isolated words were analyzed as dialect markers. Two phonetically 
trained judges listened to multiple sets of one-word tokens from each participant extracted from 
the speech samples used in Experiments 1 and 2. Their ratings on words targeting specific dialect 
differences matched those that untrained judges assigned to the longer speech samples. This 
suggests that crucial information regarding the degree of dialectal change in the CA group can be 
pinpointed to specific phonetic realizations of sounds that are known to distinguish between 
given dialects. Overall, this study on dialect acquisition is relevant to the L2 production 
experiment that will be described in this chapter because it shows that native dialect (D1) 
peculiarities are carried over into the interlanguage and that small phonetic deviations from the 
norm can be perceived even by naïve listeners from a different dialect (D2).

In summary, in this section I have reviewed a few studies that examined dialects as target 
varieties in acquisition. Although these studies do not discuss the influence of the native dialect 
in L2 production, which is the focus of the present chapter, they provided valuable information 
on the interlanguage of learners from the same L1 acquiring different target dialects of L2 and of 
learners acquiring a different dialect of their L1. The interlanguage is thus shaped by both the 
native dialect and by the ambient linguistic variety.

5 Canadian English as spoken in Toronto

Before proceeding to the presentation of the L2 production experiment, a few considerations on 
the native English variety used to set up the experiment are warranted. The materials used in 
testing L2 production in the current experiment and in the L2 perception experiment (Chapter 4) 
were obtained from native Canadian speakers who also form the control group. Given that their 
data on Toronto English serve as a reference in the present study, the next section describes the 
main findings and tendencies reported in the literature for this variety.

Canadian English used to be presented as a relatively homogenous and stable dialect of North 
America until the seminal study of Clarke, Elms and Youssef (1995) signaled a change in

---

31 According to references cited in Munro, Derwing, and Flege (1999), in Canadian English /aj/ is raised to [vä] 
(Canadian Raising) whereas in Alabaman English /aj/ is almost monophthongal [a].
progress, namely the lowering of the lax front vowels /I, E, Q/ as a result of the merger of /ç/ and /A/. Although the change matches to some extent the Northern Cities Vowel Shift (i.e. lax front vowel lowering), specific trends were found exclusively in the Canadian variety, particularly with the low vowels. Specifically, /E/ only lowers but does not retract towards /Ä/. Mid back /Ä/ lowers and/or becomes centralized so that it “occupies anywhere from a mid to low central to low back position” (Clarke et al., 1995, p. 221). Front low /Q/ retracts and lowers towards a central /a/, apparently because mid back /ç/ and low back /A/ merge to a low back category. This /A/ category presents a variable degree of rounding, at least in the 1995 corpus. Some fronting of the high back vowels /u, U/ is also reported. It is worth adding that the citation-form speech samples were obtained from 16 speakers in their mid 20s living in Toronto and that the analysis is based on both transcription and acoustic analysis. All possible phonetic contexts grouped under the manner of articulation, the place of articulation and voicing of the following consonant as well as the gender and age of the speakers were analyzed thus yielding a full picture of the Canadian Shift.

The Atlas of North American English (Labov, Ash, & Boberg, 2006) reports the current state of affairs of the vowel system and the mergers and shifts English vowels have undergone. The tendency of variation of the vowels of Canadian English is characterized by two trends: (i) the Canadian Shift, namely the low back merger of /ç/ in low back position as /A/ that triggers the above-mentioned lowering of the lax front vowels and (ii) the Canadian raising of the nucleus of /aj/ and /aw/ to /Äj/ and /Äw/ before voiceless consonants. Data on mergers and chain shifts as well as on the phonemic inventory of Canadian English were obtained from judgments on minimal pairs and rhymes in different phonetic contexts. As such, the Atlas (2006) contains acoustic data and maps for Canadian English from 22 informants across Canada, of which eight (7 female, 1 male) were from Ontario.

A recent study (Roeder & Jarmasz, 2010) provides further information on the Canadian vowel shift in Toronto from sociolinguistic interviews, analyzing a larger population sample (33 speakers) distributed across three age groups (as opposed to the above-mentioned studies in which only the young group is represented). Their findings focusing exclusively on the front vowels /I, E, Q/ are similar to the previously cited reports in that /Q/ is lowering and retracting
but, contrary to Clarke et al. (1995), in that /E/ is retracting and /I/ is retracting and lowering. The authors acknowledge that the comparison of these two studies is problematic due to important methodological differences (e.g., different elicitation methods, stimuli, groups). Results indicate that the trajectory of the Canadian shift is approaching the end in Toronto, that /I/ has stopped changing, /E/ is only lowering but no longer retracting while only /Q/ is still changing. Given the different stages of change affecting these vowels, the authors offer an explanation based on Dispersion Theory as an alternative to the pull-chain shift triggered by the /ç/-/Ay/ merger set forth in the previous studies.

Although these three studies differ non-trivially in the data elicitation and analysis, they are relevant to the present research as a reference point for the variability patterns that are likely to be found in the data on Toronto English. Thus, the corpus containing the /Q, v, Ay/ vowels obtained from 6 female speakers with ages between 20-30 (the control group in this study) is predicted to display considerable variability for /Q/, centralized /v/’s and high /Ay/’s. Given the reduced number of speakers, vowels, and factors (age and gender) taken into account, the analysis aims to serve as a reference for comparisons between L2 and target vowels.

6 The current study

The same 19 participants (10 from the Peninsular variety, 9 from the Cuban variety) that formed the advanced learner groups in the perceptual task presented in Chapter 4 participated in the L2 production experiment reported next. Acoustic data were obtained by means of a sentence-completion task presented in detail in Chapter 2. The same testing task was used to obtain data on the target vowels from a group of 6 English native speakers from Toronto. The main objective was to highlight potential differences in the realization of the target English vowels /Q, v, Ay/ as produced by advanced learners with different native dialectal backgrounds. Given the discrepancy in the inventory sizes of Spanish and English, non-native learners have to adjust their perception and production to accommodate three English vowels /Q, v, Ay/ in a similar space to that occupied by Spanish /a/ alone (e.g., Boersma & Escudero, 2008; Bohn & Flege, 1992). Several questions emerge, regarding the strategies learners employ to produce these vowels: do they reuse their native categories /a, o, e/ to approximate several target language categories, or do they rather create one or more new categories? In what way does this new category exploit
spectral and duration cues? Do the attested differences in the cross-dialectal vowel inventories reported in Chapter 3 affect the categorization of target vowels? In other words, does the more anterior vowel space in Peninsular Spanish influence their production of front L2 vowels and, respectively, does the more posterior vowel space in Cuban Spanish influence their production of back L2 vowels? Does L2 vowel production indicate different category assimilations of L2 vowels for participants depending on their native dialect? In order to answer these questions, the corpus of Canadian English (CE) vowels obtained from native speakers in Toronto as well as measurements of the non-native vowels obtained from Peninsular (PS) and Cuban (CS) vowels are compared as follows: (i) non-native and target vowel comparisons (CE, PS, CS) to check how well L2 learners produced the target vowels, (ii) comparisons between non-native vowels to assess how learners from PS differ from learners from CS; and (iii) the non-native vowels /Q, √, A/ within each group (PS, CS) to check whether there are specific patterns of vowel overlap that differ cross-dialectally.

6.1 Hypotheses and predictions

According to the studies reviewed in § 3 and § 4, advanced learners are likely to produce vowels that approximate well those in the target language, particularly if these vowels are perceived as new. However, the vowels /Q, √, A/ in the current study are similar to the native categories (mainly to /a/)33, so some or all of these vowels will not be accurately produced. Thus, the first hypothesis (H1 - attainment) addresses L2 production and states that there are differences in the realization of each of the L2 vowels /Q, √, A/ as compared to those of the target language, as follows:

- Differences between the target /Q/ as produced by English native speakers and the L2 /Q/ as produced by participants in either dialect group are predicted in the F1 dimension with

32 This question cannot be answered for now, because only three rather than all English vowels were elicited from participants. The comparisons will look into the relative position of the L2 vowels produced by learners from distinct native dialects as compared to TL and will seek to highlight any possible shift towards the front of /Q, √, A/ as produced by PS and towards the back of /Q, √, A/ as produced by learners from CS.

33 Based on dissimilarity judgments for Spanish-English pairs /a-Q/, /a-√/, /a-A/ elicited from Spanish native learners of English (Flege, Munro, & Fox, 1994). These pairs obtained low dissimilarity scores suggesting that L2 /Q, √, A/ may be assimilated to /a/, see discussion in Chapter 4, § 3.2.
both groups of learners producing a higher, more centralized vowel (i.e. lower F1 as compared to the target).

- Differences between the target /\i/ as produced by English native speakers and L2 /\i/ as produced by L2 learners are predicted, in particular for the L2 /\i/ produced by the CS group, with lower F2 and F3 as compared to the target.

- Differences between the target /A/ as produced by English native speakers and L2 /A/ are predicted for both groups of learners.

The second hypothesis (H2 - native dialect effects in L2) addresses differences between L2 vowels as produced by learners from distinct native dialects. Given the phonetic differences between the native vowels of PS and CS and cross-linguistic transfer attested even with advanced learners, the L2 vowels produced by PS speakers are predicted to be more fronted and shorter than those produced by CS speakers, most notably for the L2 /\i/ (higher F2 for PS than for CS). Different intervocalic distances are predicted for the pairs of L2 vowels /Q-\i/, /A-\i/ as produced by learners from PS versus CS, namely smaller /Q-\i/ distance in PS, smaller /A-\i/ distance in CS and similar /Q-A/ distances reflecting specific preferences for more fronted in PS and respectively more posterior articulations in CS.

The third hypothesis (H3 - L2 vowel overlap patterns) targets the distribution of L2 vowel means for individual learners from each dialectal group. Based on the intervocalic distance evaluation in Hypothesis 2, specific patterns of L2 vowel overlap are expected. Since perceptually the CS group of advanced learners has more discrimination errors with the /\i-A/ contrast (according to findings in Chapter 4), the learners’ production of these vowels is likely to mirror their perception. Thus, more overlap is predicted for /A i/ as produced by the nine L2 learners from CS (i.e., smaller, non-significant differences between /A, i/ along F1, F2, F3) and, similarly, more overlap is predicted for /Q, i/ as produced by the ten learners from PS (i.e., smaller, non-significant differences between /Q, i/ along F1, F2, F3).

To sum up, the hypotheses to be tested here are as follows:
- **H1 - attainment** - there are differences in the realization of each of the L2 vowels /ʊ, ʌ, æ/ as compared to those of the target language.

- **H2 - native dialect effects in L2** - the English vowels produced by learners from PS are more fronted and shorter than those produced by learners from CS.

- **H3 - L2 vowel overlap patterns** – L2 learners from PS tend to produce the English /ʊ, ʌ/ overlapping into a front category while L2 learners from CS to produce the English /æ, ʌ/ overlapping into a back category.

### 6.2 Results

More than 1,500 vowels produced by L2 learners and English native speakers were analyzed (F1, F2, F3, and durations were measured) as presented in detail in Chapter 2. Formants were normalized using Lobanov’s normalization procedure (described in Chapter 3) that was shown to be the best at eliminating gender differences while preserving regional variation (Adank et al., 2004). Figure 5.1 provides means for normalized vowels in each of the /ʊ, ʌ, ʌ/ categories produced by all speakers in each of the PS, CS, and CE groups.
Figure 5.1
Mean normalized formant values for L2 (as produced by L2 learners) and TL (as produced by English native speakers) /Q, ñ, A/ vowels

The quantitative analysis, which is presented next, included two additional parameters: F3, a correlate for vowel roundness, and vowel duration. Indeed, most of the tendencies observed in Figure 5.1 proved to be statistically significant. The first series of tests compared how closely L2 vowels produced by learners from PS and, respectively, CS approximate the TL vowels produced by English native speakers (Hypothesis 1 - attainment) and to what extent L2 vowels produced by learners from PS differ from those produced by learners from CS (Hypothesis 2 - native dialectal effects in L2). Four one-way ANOVA’s were conducted for each of the /Q, ñ, A/ vowels to compare the effect of the linguistic variety (Canadian English, PS, CS) on normalized F1, F2, F3, and duration of the vowels. Several significant effects at the $p < .05$ level were found, as presented in Tables 5.1, 5.2, and 5.3. In the case of statistically significant results, further exploration through post hoc Tukey HSD tests indicated the source of the differences.
Table 5.1
Vowel /ɒ/: results of comparisons among CE, PS, and CS for normalized F1, F2, F3, and duration. Significant differences are further explored through pairwise comparisons of CE-PS, CE-CS, PS-CS.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overall F, p</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CE-PS</td>
<td>CE-CS</td>
</tr>
<tr>
<td>F1 norm</td>
<td>F(2, 22) = 5.515, p = .011</td>
<td>p = .023</td>
</tr>
<tr>
<td>F2 norm</td>
<td>F(2, 22) = 1.91, p = .172 n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>F3 norm</td>
<td>F(2, 22) = 1.205, p = .319, n.s</td>
<td>n.s.</td>
</tr>
<tr>
<td>Duration</td>
<td>F(2, 22) = 1.244, p = .308, n.s</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

With respect to the low front vowel /ɒ/, its realization in both learner groups differs significantly from the target particularly along the height (F1) dimension. There are no statistical differences along the other parameters investigated, nor between the realizations of /ɒ/ in the two groups of learners.

Table 5.2
Vowel /ʌ/: results of comparison among CE, PS, and CS for normalized F1, F2, F3, and duration. Significant differences are further explored through pairwise comparisons of CE-PS, CE-CS, PS-CS.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overall F, p</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CE-PS</td>
<td>CE-CS</td>
</tr>
<tr>
<td>F1 norm</td>
<td>F(2, 22) = 4.791, p = .019</td>
<td>p = .067</td>
</tr>
<tr>
<td>F2 norm</td>
<td>F(2, 22) = 18.381, p = .000</td>
<td>n.s.</td>
</tr>
<tr>
<td>F3 norm</td>
<td>F(2, 22) = 4.798, p = .019</td>
<td>p = .014</td>
</tr>
<tr>
<td>Duration</td>
<td>F(2, 22) = 3.668, p = .042</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

There are more substantial differences in the realization of /ʌ/: as compared to the target, /ʌ/ is significantly more open (lower F1) for L2 learners from PS and it is significantly more posterior (lower F2) and longer for L2 learners from CS. There are also statistically significant differences between the realization of /ʌ/ in the two groups of learners, as it is lower and more fronted for L2 learners from PS as compared to L2 learners from CS.
Table 5.3
Vowel /A/: results of comparison among CE, PS, and CS for normalized F1, F2, F3, and duration. Significant differences are further explored through pairwise comparisons of CE-PS, CE-CS, PS-CS.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overall F, p</th>
<th>CE-PS</th>
<th>CE-CS</th>
<th>PS-CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 norm</td>
<td>F(2, 22) = 1.981, p = .162 n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>F2 norm</td>
<td>F(2, 22) = 5.841, p = .009 n.s.</td>
<td>n.s.</td>
<td>p = .007</td>
<td>n.s.</td>
</tr>
<tr>
<td>F3 norm</td>
<td>F(2, 22) = .881, p = .449 n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Duration</td>
<td>F(2, 22) = 3.710, p = .041 n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .037</td>
</tr>
</tbody>
</table>

Finally, with respect to /A/, this vowel is more fronted (higher F2) for L2 learners from CS, as compared to that of English native speakers. The only statistically significant difference between the realization of /A/ in the two groups of learners is the significantly shorter duration for L2 learners from PS.

The next step of the analysis included tests regarding Hypothesis 3 - L2 vowel overlaps, by which /Q-/Ã/ are predicted to overlap in PS, and /Ã-/A/ in CS, i.e., there are no significant differences between the vowels of each pair. In this case, within each dialect, L2 vowels (averages of F1, F2, F3, and duration for each participant in each dialectal groups) were compared through a linear mixed model with repeated measures for vowels. The output of the analysis is shown in Tables 5.4 for vowels produced by L2 learners from PS and in Table 5.5 for vowels produced by L2 learners from CS.

Table 5.4
L2 vowel overlap patterns for L2 learners from PS

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overall F, p</th>
<th>/Ã-/A/</th>
<th>/Q-/Ã/</th>
<th>A-Q/</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 norm</td>
<td>F(2, 9) = 25.718, p &lt; .000</td>
<td>p = .01</td>
<td>p = .013</td>
<td>p &lt; .000</td>
</tr>
<tr>
<td>F2 norm</td>
<td>F(2, 9) = 78.803, p &lt; .000</td>
<td>p &lt; .000</td>
<td>p = .023</td>
<td>p &lt; .000</td>
</tr>
<tr>
<td>F3 norm</td>
<td>F(2, 9) = 12.195, p = .003</td>
<td>p = .003</td>
<td>n.s.</td>
<td>p = .002</td>
</tr>
<tr>
<td>Duration</td>
<td>F(2, 9) = 2.965, p = .103 n.s.</td>
<td>p = .04</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

The PS group’s realizations of English /Q, √, A/ constitute spectrally different categories, as all differences are significant. It is worth noting, however, that for the /Q-/√/ comparison along F3, the differences are statistically non-significant. Interestingly, there is no effect of vowel type on duration, suggesting that PS learners hardly produce any difference between tense (long) and lax
(short) vowels. Duration means across the three vowel categories are similar: 112 ms for /Q/, 98 ms for /ι/, and 107 ms for /A/.

Table 5.5

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overall F, p</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>/ι/-A/</td>
</tr>
<tr>
<td>F1 norm</td>
<td>F(2, 8) = 16.896, p = .001</td>
<td>n.s.</td>
</tr>
<tr>
<td>F2 norm</td>
<td>F(2, 8) = 152.383, p &lt; .000</td>
<td>p = .039</td>
</tr>
<tr>
<td>F3 norm</td>
<td>F(2, 8) = 2.611, p = .134 n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Duration</td>
<td>F(2, 8) = 9.952, p = .007</td>
<td>p = .005</td>
</tr>
</tbody>
</table>

In the CS group, however, /ι/, /A/ are produced with similar F1 and F3 and only marginally significant differences along F2. Contrary to what was found in the PS variety, duration is used to differentiate between tense and lax vowels. The average durations for /Q/, /ι/, /A/ are, respectively, 136 ms, 118 ms and 144 ms.

Finally, in order to obtain a more precise picture of the vowel locations in each dialect and support findings for Hypothesis 3, Euclidean distances were calculated using normalized formant values and the formula (2) introduced in Chapter 3 and repeated below:

\[ d(V_1, V_2) = \sqrt{(F1_{V_1} - F1_{V_2})^2 + (F2_{V_1} - F2_{V_2})^2} \]

Distances between pairs of vowels were calculated for each speaker and mean distances for each pair of vowels were further obtained for each of the 19 speakers. Three one-way ANOVAs were performed on these means in order to assess the effect of dialect on intervocalic distances. All mean distances between pairs of L2 vowels were found to differ significantly across groups of learners from PS and CS. The /Q-ι/ distance is smaller in PS than in CS \( F(1, 17) = 8.575, p = .009)\), adding evidence that these categories tend to overlap more towards the front of the F1 x F2 space in the former variety. Conversely, the /ι-A/ distance is smaller in the CS variety \( F(1, 17) = 17.032, p = .001)\) pointing to more category overlap towards the back of the vowel space. The /Q-ι/ distances are slightly larger in the PS variety \( F(1, 17) = 5.527, p = .031)\) indicating a more expanded vowel space. To illustrate differences in intervocalic distances and dispersion of the vowel space, formant means for each speaker, dialect, and vowel category are given in Figures 5.2 (for PS) and 5.3 (for CS). One data point in the plot corresponds to mean F1 and F2.
values measured for one speaker and vowel category. Formant means across CE speakers are also represented (large empty shapes) for comparison purposes.

Figure 5.2
Mean normalized formant values for each speaker and vowel category in PS
Indeed, these plots indicate a more dispersed L2 vowel space of the PS learners, with smaller distances between /Q-∫/ and greater for /√-A/ and vice versa for CS learners. Although L2 vowels tend to be realized as separate categories by PS learners, a preference for a fronted realization of the /Q-∫/ is apparent, while the /A/ vowel spans more along both F1 and F2. In the CS group of learners, the /√-A/ category overlap is more prominent, as is the separation from the /Q/ category. Figures 5.2 and 5.3 also show that there is some individual variability particularly for L2 learners from PS, whose realizations of /√/, /A/ seem to spread out more in the F1 x F2 space. In the next section, I analyze the individual variation in relation to the learner’s background.

6.2.1 Individual variability

A closer look at the individual patterns of vowel categorization may explain the greater variability found in the PS group. Six out of 10 participants in this group showed a clear-cut
overlap of the /Q, √/ categories, which are separated from /A/. For the remaining 4 participants, two (PS6 and PS9) had no specific pattern, as they produced all three English vowels with a great degree of overlap. The two remaining participants (PS3 and PS5) showed overlap between /√, A/, a pattern observed with the CS group as a whole. In the CS group, 7 out of 9 participants produced the /√, A/ overlap while for the other two (CS4 and CS9), all vowels tended to overlap. Quantitatively, participants in CS group were more consistent in their realization of the L2 vowels according to the pattern proposed for their group (77% of learners in CS had the /√, A/ overlap) as compared to the PS group (60% of learners in PS had the /Q, √/ overlap). A closer examination of the subjects’ background questionnaire (Appendix 2) suggests that Peninsular participants PS3 and PS5, whose production involves the /√, A/ overlap, a pattern that is predominant in the CS group, had some exposure to American English. Specifically, PS3 had spent approximately 2 months in the US, whereas PS5 had an American English native speaking partner, with whom, nevertheless, she interacts predominantly in Spanish, as her reported amount of English use per week (15 hours/week), is in line with the other participants. The CS group, in which most of the participants showed the /√, A/ overlap, had also more exposure to Englishes of North America, especially Canadian.

6.3 Hypothesis evaluation and discussion

Several predictions were explored in this study. With respect to attainment (Hypothesis 1), which predicted differences between the L2 (as produced by L2 learners) and target vowels (as produced by English native speakers), the L2 learners’ /A/ was acoustically realized as the closest to native Toronto English along all four parameters investigated (the only significant difference was found with the CS group who produced more fronted /A/s). The differences predicted for this L2 vowel as compared to the native target vowel proved to be actually very small. Thus, the prediction regarding this vowel was not supported. The next most accurately produced vowel is the low front vowel /Q/, which is similar to the target in F2, F3 and duration parameters, but is higher (lower F1) than in English. The differences between the L2 /Q/ and the target as produced by English native speakers proved to be as predicted. Finally, the learners’ production of the mid
back/centralized /\}/ diverged the most from that of the controls, as this vowel was lower in PS and more posterior and longer in CS. Although the acoustic similarity of L2 learners’ production of a vowel vis-à-vis the target does not necessarily reflect the perceptual similarity that exists between the vowels of a diaphone (L1-L2 pair), in this case, the inaccuracies in the realization of L2 /\}/’s parallel those obtained in the discrimination task presented in Chapter 4 in which the /\}/ appeared in the most confusable contrasts.

To summarize, in the non-native production of native PS speakers, vowels were spectrally similar to those of the native English speakers. Specifically, /\}/ was the closest to the target, followed by /\}/ and /Q/. No significant differences in duration were found between the English vowels of the PS and control groups. On the other hand, in non-native CS speakers’ production, all three vowels differed from the target along at least one parameter, (/Q/ in F1; /\}/ in F2 and duration; and /\}/ in F2).

An assumption that this study made was that the target vowels /Q, \, \}/ are perceived as similar to the native category /a/ (and maybe /o/) and thus assimilated to it, following the SLM. Given the similarity between diaphones, no new categories would be created, but rather existing L1 categories would be reused with variable degrees of success (e.g., the Spanish /a/ for the English /Q/). A three-to-one assimilation would, however, fall short of realizing all contrasts between /Q, \, \}/, a problem that most of the L2 learners managed to avoid. According to the L2 production data summarized in Figures 5.2 and 5.3, the three target vowels are apparently mapped to two interlanguage phonetic categories, a fronted low category and a posterior back category. Additionally, some L2 learners, those from CS, used duration distinctively. In the PS group, there is a tendency (not significant, though) to assimilate the English /Q, \}/ to one interlanguage front low category, separately from a higher and more posterior category used to realize target /\}/. However, as opposed to the target language, these L2 contrasts do not exploit duration. With the CS group, target /Q/ was assimilated to a front category while targets /\}, \}/ were realized in the Spanish speakers’ interlanguage through a back and slightly higher category. Moreover, within the spectrally merged L2 category (/\}, \}/), a duration contrast was also employed. It is worth reminding the reader that the L2 vowels were elicited through a sentence-completion task.
participants listened to short definitions and then had to complete each sentence with the appropriate word. As discussed in Chapter 2, participants were told that the task’s objective was to assess their vocabulary. No orthographic input was provided, as the definitions were played back with no written subtitles, as the aim was to obtain semi-spontaneous speech. As such, participants focused on vocabulary retrieval rather than pronunciation and their production is likely to reflect closely the categories they have created for English.

With respect to potential native dialect effects in L2 production (Hypothesis 2), I hypothesized that a more fronted vowel space in native PS would be carried over to L2 production and similarly, a preference for more posterior articulations in CS would result in more posterior L2 vowels. However, this hypothesis was not supported for the peripheral L2 vowels /Q, /A/. Qualitatively (Figure 5.1), these vowels are actually more posterior in the L2 English of the PS group than in that of the native CS speakers, although statistically their locations are similar. The apparently more expanded L2 vowel space of the PS learners (which does not actually reach the significance level, but can be appreciated in Figure 5.2) may be attributed to the fact that some of the learners in this group have been exposed to an input variety in which the /ɛ-ʌ/ merger is absent, namely British or Irish English (varieties most of the PS participants were familiar with, according to the background questionnaire). In other words, some of the tokens were realized with higher, more (/o/-like) posterior articulations, leading to more spread along these dimensions. With respect to the L2 vowel /ʌ/, the Hypothesis 2 was indeed confirmed, as this vowel was significantly more fronted for L2 learners in the PS group as compared to those in the CS group.

This difference was reflected in distinct patterns of vowel overlap (Hypothesis 3). Some learners in the PS group managed to create three distinct categories well spread in the vowel space while others showed a tendency towards mapping /Q, /ʌ/ to one low front category and did not use duration distinctively. Consequently, Hypothesis 3 received only partial confirmation, as category overlap was supported statistically only for the CS group of learners, with learners using a mid-low back spectral category for both /ʌ/, /A/ but distinguishing between them by realizing shorter /ʌ/-s. Thus, from the perspective of Escudero and Boersma’s formal linguistic model of categorization, learners in the two dialectal groups diverged in their strategies of realizing English /Q, /ʌ, /A/: while the PS group exploited the spectral properties of these vowels
to a greater extent (slightly greater expansion of the vowel space, better separation of categories with some speakers, greater spectral similarity to the English targets), they hardly used duration, a cue relevant to tense-lax distinctions. The CS group showed a clear three-to-two assimilation pattern that resulted in vowels that did not match the English targets, particularly in the case of /ɜ/, but apparently created a duration category to realize the non-native contrasts. As such, the L2 production results presented here are compatible with predictions on perceptual learning set forth by Escudero and Boersma (2004) in that, depending on specific conditions in the interlanguage, like different ambient input or distinct native varieties, L2 learners reuse L1 categories and shift L1 boundaries (L2 learners from PS), or reuse L1 categories and create new categories (duration) to approximate TL categories (L2 learners from CS).

From the perspective of PAM-L2 (Best and Tyler, 2007), L2 learners from PS assimilate the English vowels according to a two-category assimilation type (the categorized-uncategorized subset), in which English /Q, ʌ/ clearly match L1 phonological categories (likely /a, o/) and /ɜ/ is uncategorized and thus is variably produced towards a front category. In perception, L2 learners from CS have high error rates for the /A-ɜ/ contrast, but discriminate well /ɜ-Q/ and /Q-A/. In L2 production, /A, ɜ/ overlap. Based on the PAM-L2 considerations, L2 learners from CS adopt a category-goodness assimilation pattern for /A-ɜ/, as both these L2 categories are perceived as the same L1 phonological category, with /ʌ/ being the better match and a new phonetic category (duration) is created for /ɜ/, probably due to the necessity of distinguishing minimal lexical pairs containing /A-ɜ/.

7 Conclusions

The L2 production experiment presented in this chapter refined the findings of the L2 perceptual experiment in Chapter 4. Sharper differences between the two groups of L2 learners with distinct dialectal backgrounds were highlighted, suggesting different interlanguage strategies of producing the English /Q, ɜ, ʌ/ vowels. Whereas the hypothesis regarding native dialect effects in L2 production (H2), which predicted that L2 learners from PS would produce more fronted L2 vowels, was supported only for /ɜ/ but not for /Q, ʌ/, the hypothesis regarding specific vowel overlap patterns (H3) shed light on particular strategies of mapping L1 into L2 sounds suggesting
that the native variety along with the ambient input play a role in how the English vowels are produced by L2 learners from different Spanish dialects.
Chapter 6
Discussion

1 Introduction

This study was carried out with the purpose of exploring the role of the native dialect in the production and perception of non-native sounds. The main research question was whether systematic acoustic differences between vowels in the native dialects of Spanish would produce matching differences in non-native perception and production of English vowels. Specific questions about the phonetic realization of native and non-native vowels and the perceptual discrimination of the non-native contrasts /ɪ-/ʌ/, /ʌ-/ɒ/, and /ʌ-/ɒ/ (as in hut-hot, hut-hat and hot-hat) were addressed. Most questions received a definitive answer, as the hypotheses that articulated them experimentally were either partially supported or disconfirmed. Hypothesis evaluations were presented at the end of each of the Chapters 3, 4, and 5. I present next an in-depth discussion of the relationships across all sets of results and their implications for the big-picture phenomena. As such, in this chapter the main findings of the experiments on native production (§ 2) and non-native perception (§ 3) and production (§ 4) are reiterated and discussed in relation to each other along with several other questions that emerged post-hoc from the analyses. The interpretation of these findings within three L2 acquisition models is discussed as well, focusing on contribution, conclusion and future research.

2 Vowel comparisons in Peninsular and Cuban Spanish

Apart from verifying existing data on PS and contributing new data on CS, the aim of the cross-dialectal comparison was to identify the extent of the phonetic differences between vowels in these varieties and the potential consequences for L2 vowel perception and production. As such, comparisons based on acoustic measures of vowel locations in the acoustic space (spectral parameters), durations, vowel space dispersion (intervocalic distances), and within-category variability yielded the following statistically significant differences: [i] is higher (4.9%) and more fronted (9.8%) in PS than in CS, [o] is more fronted (10.3%) in PS than in CS, [e, a, o] are shorter (20%) in PS than in CS; the PS vowel space is more dispersed towards the high front
region while the CS is more dispersed (tendency) towards the high back region; there is more variability along F1 in PS and along F2 for the CS front vowels.

2.1 How real are the cross-dialectal differences?

The main question that these findings raise is whether these cross-dialectal differences that are statistically significant are also phonetically relevant. Would native speakers of the Cuban dialect characterize PS vowels as different from their own and vice versa? An identification task with category goodness ratings performed by listeners from one dialect on vowels of the other dialect are likely to produce highly accurate identification scores and only category goodness ratings might provide some indication of cross-dialectal spectral differences. However, with respect to the shorter vowels in the PS dialect, duration might be what really distinguishes its vowels from other varieties. Morrison and Escudero (2007, p. 1508) noted that “In the Spanish-speaking world, Spaniards have a reputation for speaking quickly” cautioning, nevertheless, that shorter vowels may be thus a speech-rate effect. Speech rate was not controlled in the experiments reported here, however, given the elicitation procedure and the instructions that participants received (see Chapter 2 for details), such effects were likely minimized.\(^{34}\)

The fine-grained phonetic differences between vowels in PS and CS may pass unnoticed to speakers/listeners of the opposite dialect\(^ {35}\) or may be interpreted as idiosyncratic individual variation. However, such differences between native vowels can be essential in L2 production and perception. Different boundaries between L1 categories across dialects and/or different intrinsic durations in the L1 dialects may cause L2 learners/listeners to organize and implement non-native sounds differently in their interlanguage, particularly those L2 sounds whose spectral or temporal characteristics fall into the critical/sensitive domain of cross-dialectal differences. A detailed discussion of the pros and cons of the dialect effects in L2 production is presented in §4.1.

---

\(^{34}\) Had I used speech samples taken from conversations (i.e., spontaneous speech), I would have obtained more individual variability for vowel duration. Instead, semi-spontaneous speech elicited from participants who were instructed to read stimuli in carrier sentences from cards presented at a controlled pace are likely to reflect duration the way it is for each variety.

\(^{35}\) These phonetic differences are at most interpreted as allophonic variation or free variation but most native speakers are not aware of the differences.
2.1.1. Phonological processes and phonetic realization of vowels

In the previous section I have speculated that the extent of the cross-dialectal differences between vowels in one dialect may not be large enough to be readily noticed by listeners from other dialects. However, dialects differ with respect to other more salient phenomena that, in turn, generate significant differences in vowels. Specifically, in the Andalusian dialect in Southern Spain, the /s/-deletion in coda and word final position causes stressed and final vowels to be realized as more open (with a lower tongue position) than in the standard Peninsular dialect. The vowels involved most frequently in this morpho-phonological process are /e, a, o/. In the context of final /s/-deletion, the functional load of the noun’s plural or the verb’s second person singular morpheme -s is transferred to the vowel that receives a lower articulation, in the case of the mid vowels /e, o/ and a more advanced or palatalized for /a/ (Moreno Fernández, 2004, p. 974). As various authors (e.g. Penny, 2000) indicate, the contrast between the noun’s singular and plural or the verb’s second and third person is rendered only on the basis of a tongue height distinction. For example, chic[o] (boy) ~ chic[œ] (boys) differ only in vowel quality between [o] and [œ]. Given that these pairs of vowels /e-E/, /o-Œ/ and /a-Q/ are used to differentiate words, they actually function as phonemes, and thus the phonemic inventory of the Andalusian dialect contains eight vowels, three more than the standard Peninsular Spanish. Therefore, an instance of phonetic variation triggered by /s/-deletion leads to a phonological split in the vowel system. This phenomenon has been explored only in the Andalusian Spanish but not in relation to any other variety. Since /s/-aspiration and deletion is a widespread characteristic of Cuban Spanish (Terrell, 1979) too, it may be the case that it affects the vowel quality in the same way as in the Andalusian dialect. In contrast, in the Northern Peninsular Spanish dialect, /s/-deletion is hardly present, if present at all. In other words, the presence versus absence of /s/-deletion processes found in Cuban Spanish as opposed to Peninsular Spanish may influence the vowel quality, namely, vowels in an /s/-deletion context are more open as compared to vowels in a similar context but in dialects without /s/-deletion.

In the present study, the analysis focused only on vowels in bilabial context. It is worth noticing that even in this neutral context, [i, e, a, u] vowels tend to be lower/more open in CS than in PS as in Figure 3.3 (Chapter 3), which shows cross-dialectal vowel spaces. I speculate that in the context in which /s/-deletion operates, the differences in height between vowels in PS and CS are
even more substantial than in the bilabial context. Phonetic evidence from an acoustic comparison of Peninsular and Peruvian vowels (Chládková et al., 2011) confirmed that the dialectal differences are more pronounced and affecting more vowels in /s/-context as compared to other contexts (/p, t, k, f/). Specifically, in /s/-context, /i, a, o, u/ have more centralized values along F2 in Peninsular as compared to Peruvian Spanish. Further exploration of the centers of gravity of the /s/’s pointed to different articulation across the Peninsular and Peruvian dialects.\footnote{According to Chládková et al. (2011, p. 425) and other studies cited within, different places of articulation of /s/ affect the centers of gravity, the more posterior the articulation, the lower the centre of gravity. Moreover, articulatory differences in the consonants tend to be reflected in the neighboring vowels.} Based on these findings, a follow-up study on PS and CS would explore whether processes specific to one dialect but not to the other generate indeed phonetic differences between vowels and furthermore, whether such differences can be detected by listeners in other dialects.

### 2.2 Some findings in the light of different phonetic theories

The following section discusses how the findings of the cross-dialectal comparisons inform the phonetic theories presented in Chapter 3, in particular, speech rate effects in relation with the vowel space dispersion, within-category variability and its relevance to the acoustic-articulatory predictions of the Quantal Theory and vowel location differences within the framework of the Articulatory Settings/Base of Articulation theory.

#### 2.2.1 TAD – dispersion and speech rate

The Theory of Adaptive Dispersion provided the acoustic measures, namely vowel space dispersion and contrast distinctiveness, which were used in the cross-dialectal comparisons of vowels. Although the experiments were not designed with the purpose of testing any of the assumptions of TAD, it is worth discussing here the relationship between vowel space dispersion and speech rate and the findings of the present research. It has been shown (e.g. Moon & Lindblom, 1994) that the vowel space is more dispersed in clear speech,\footnote{Clear speech is produced when speakers adapt their style according to the situation, as in talking in a noisy environment, to foreign interlocutors or to infants (Moon & Lindblom, 1994). Clear speech also supposes better intelligibility and is acoustically characterized by an expanded vowel space and less within-category variability, longer vowels, and longer transitions.} which is also characterized by longer vowels. Assuming that there is a relationship between vowel space...
dispersion and vowel duration at speech rates faster than clear speech, as targeted in the present experiment (in Chapter 3), I speculate that the dialect with a more expanded vowel space, namely PS, would also have longer vowels. However, the cross-dialectal comparison showed that the PS vowel space is more expanded towards the high front region but its [e, a, o] vowels are actually shorter than in the CS dialect. This can be interpreted as a confirmation that PS low and mid vowels are indeed intrinsically shorter than in the other dialect.

2.2.2 QT and variability of point vowels

The findings of the present cross-dialect research have also relevance for the Quantal Theory. Two testable predictions of the QT are that the /i, a, u/ point vowels, considered stable reference points in the articulatory and acoustic space, are characterized by less variability than other non-point vowels (Pisoni, 1980) and that they can be found at similar locations cross-linguistically (Bradlow, 1995). These predictions were not supported in either study. In other words, there was no clear low variability pattern for the point vowels, nor were they situated at similar locations cross-linguistically. Recasens and Espinosa (2006) reported a more consistent variability pattern for point vowels across the four Catalan dialects analyzed in their study, namely little variability for /i/ and a great amount of variability for /u, a/.

In line with these studies, the present research also disproved the QT’s predictions in that the /i/’s were at different locations cross-dialectally and there was a high degree of within-category variability for /i, a/ and low variability for /u/. This suggests that within-category variability of point vowels is not always low, as predicted by QT, and that point vowels are not universally found at the same (specific) locations. Given the various patterns reported in the literature, within-category variability should be thus considered a language- or even a dialect-specific characteristic.

2.2.3 AS/BA theory and cross-dialectal preferences

Finally, the results of the present research are discussed from the perspective of the AS/BA theory. The question is whether the acoustic differences between PS and CS reflect indeed a base of articulation effect. First, there is indication in the literature that the PS back vowels are more fronted than in other varieties (e.g. Godínez, 1978; Morrison & Escudero, 2007). Second, based on intervocalic distances between point vowels and the centroids identified in the present study,
the PS vowel space expands more towards the front whereas the CS tends to disperse more towards the back. Third, there are different variability patterns with PS speakers having greater standard deviations along the height dimension (F1) and CS speakers along the front-back dimension (F2) for the front vowels. Finally, Kochetov and Colantoni (2011) identified a preference for more posterior articulations of coronals in the Cuban dialect as compared to other varieties, which is likely to extend to the vocalic domain.

Previous findings concur with those of the current study in suggesting that indeed PS speakers have a preference for more fronted articulations whereas CS speakers tend to articulate sounds at more posterior articulatory settings. However, when comparing vowel locations acoustically, only [i, o] are more fronted in PS than in CS. A clear-cut base of articulation effect would be reflected in a systematic shift of the entire vowel space in one variety versus the other, as in Bradlow (1995). Without articulatory data to confirm the preference for different articulatory settings across dialects, the acoustic data do not support the AS/BA theory. Moreover, the duration differences found between vowels cross-dialectally cannot be readily interpreted within this framework, unless specific patterns of gestural overlaps and timings are incorporated to the analysis.

To conclude, the present acoustic exploration of cross-dialectal vowel differences is compatible with the current phonetic theories and previous findings in that three out of five PS vowels are shorter and more fronted than in other dialects and that point vowels are not necessarily found at the same locations across linguistic varieties, nor do they show less within-category variability. Further acoustic and articulatory exploration might establish whether consonantal processes specific to one variety influence vowel quality too and whether PS and CS have different articulatory settings.

---

38 All of the /i, e, o, u/ vowels investigated in Bradlow’s study were found to be shifted along F2 in English as compared to Spanish.
3  Non-native perception and native dialect influence in L2

Having identified the extent of the cross-dialectal differences in L1, in Chapter 4 I explored whether PS and CS listeners perceive L2 vowels differently and whether the differences in L2 perception can be related to the acoustic differences found between their native vowels. In order to separate dialect effects from learning experience effects, advanced learners and non-learners of English were tested on three English contrasts, /æ-ʌ/, /ʌ-Q/ and /Q-ʌ/, whose perceptual difficulty was assessed as well.\(^{39}\) Whereas no clear native dialect effect in non-native perception could be identified in the non-learner groups, as their perceptual behaviour is similar, differences in the discrimination of the /æ-ʌ/ contrast were found between the advanced learner groups (13.7\% versus 37.9\% error rate), which suggest that CS listeners have more difficulties discriminating between these back vowels than PS listeners. The high error rates with the /æ-ʌ/ contrast indicate that CS listeners perceived the differences between these vowels as a within-category contrast rather than belonging to distinct categories. This behaviour is mirrored in the realization of the /æ, ʌ/ vowels that tend to overlap in the L2 production of the CS learners.

Drawing on the results of the L2 perceptual task, in the next two sections I will discuss two issues: the use of spectral and temporal cues and the relationship between L2 perception and L2 production.

3.1  The use of spectral and temporal cues in L2 perception

An issue often discussed in L2 perception studies is how learners use spectral and duration cues for vowel discrimination and specifically whether their use of cues has become more similar to that of the listeners in the target language. As discussed in Chapter 2, discrimination tasks do not provide direct information on the use of spectral and duration cues as identification tasks do (those in which synthetic stimuli varying along frequency and duration in controlled steps are used). However, the perceptual behaviour regarding the use of cues in L2 can still be qualitatively assessed employing the results of the discrimination task, which is discussed next.

\(^{39}\) Given the great amount of variance in the learner/non-learner groups, reporting discrimination scores per dialectal groups with learners and non-learners pooled together is unwarranted.
The use of cues varies cross-linguistically, as summarized in Chapter 3: English listeners use both spectral and temporal cues whereas Spanish listeners use mainly spectral cues. Although vowel duration is a salient cue, L2 learners may not readily process temporal cues if these are not used at all in L1 and if L2 requires the use of temporal cues in combination with spectral cues. In other words, Spanish learners of English have to discriminate vowels that differ both in quality and duration although duration is not phonologically contrastive in Spanish. Activating the duration feature is not always successful, as suggested by McAllister et al.’s (2002) study (see Chapter 4, § 4.1) in which Spanish advanced listeners showed low sensitivity to short-long vowel contrasts in Swedish. If duration only distinguishes an L2 contrast, then learners whose L1 has no duration contrast perform poorly. However, Bohn (1995) and Cebrian (2006) found, respectively, that Spanish and Catalan listeners do use temporal cues when identifying English vowels. Bohn argues that the use of duration cues is a universal strategy when temporal and spectral cues in L2 do not match those in L1 and thus are not readily accessible to the learner.40

Another possibility is when an L2 contrast exploits both temporal and spectral cues and the spectral cues match those in L1 (Cebrian, 2006). In this case the L2 learner does not activate the duration feature in his interlanguage grammar because the more-or-less familiar spectral cues are sufficient and the temporal cues are ignored. A similar situation is reflected in the present study and is described next.

In order to analyze cue weighting, I compared the availability of cues in the /A-/Ã/, /Q-/Ã/, and /A-/Q/ contrasts (Table 6.1) in relation to the L2 listeners’ perceptual performance with these contrasts (Table 6.2). The availability of specific cues was quantified through two measures: the perceptual distance, calculated as the Euclidean distance between vowels and the duration difference between pairs of vowels. The measures were computed from mean values of the acoustic parameters (F1, F2, durations) of the /Q, Ñ, A/ vowels produced by Canadian English native speakers (see Chapter 5) and transformed onto an auditorily/perceptually relevant scale (mel). As such, a small perceptual distance implies that spectral cues are insufficient (as for /A-/V/.

40 Alternately, Bohn speculates that “non-native listeners prefer to attend to the least varying dimension in perceptually differentiating non-native contrasts” (p. 300). As such, the Spanish listeners in his study used duration because they perceived it as varying less than the spectral information. While this is the case for the contrasts in Bohn’s study (beat-bit, bet-bat), L2 listeners may have a different preference for other contrasts, as in hat-hut-hot.
and /Q-Ã/ and a small duration difference implies that temporal cues are insufficient (as for /A-Q/). Labels like ‘insufficient spectral cues’ were formally operationalized as follows. I considered that spectral cues were insufficient when the intervocalic distance that characterizes a TL contrast fell below the smallest intervocalic distance in L1, that being d(o, u) (see Figure 3.4). In other words, this smallest distance is assumed to be a perceptual threshold so that, to hear contrasts separated by distances smaller than the threshold, listeners would have to attune their perception to L2 contrasts. In PS the smallest distance is 268 mel and in CS is 194 mel. As such, both /A-Ã/ and /Q-Ã/ contrasts fall below this threshold and thus are more difficult to perceive, whereas /A-Q/ is above the threshold, thus easy to perceive. Duration cues were considered insufficient if the duration difference between the vowels in the pair was smaller than 20 ms, known to be the perceptual threshold (Noteboom, 1999). In this case, only the /A-Q/ contrast has insufficient temporal cues.

Table 6.1

<table>
<thead>
<tr>
<th>Contrasts</th>
<th>Spectral cues / Euclidean distance (mel)</th>
<th>Temporal cues / Duration difference (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/A-Ã/</td>
<td>136.5</td>
<td>71</td>
</tr>
<tr>
<td>/Q-Ã/</td>
<td>146.6</td>
<td>58</td>
</tr>
<tr>
<td>/A-Q/</td>
<td>307.8</td>
<td>13</td>
</tr>
</tbody>
</table>

Note. A contrast’s spectral cues are considered ‘insufficient’ if intervocalic distances are smaller than 268 mel. A contrast’s temporal cues are considered ‘insufficient’ if duration difference is smaller than 20 ms. Contrasts having sufficient cues are in the shaded cells.

The L2 listeners’ discrimination performance with the /A-Ã/, /Q-Ã/, and /A-Q/ contrasts is shown in Table 6.2, which reproduces the error rates obtained in the four groups of PS and CS advanced learners and non-learners of English, previously reported in Chapter 4 (Table 4.2).

Table 6.2

<table>
<thead>
<tr>
<th>Contrasts</th>
<th>L2 listeners’ discrimination performance/ Error rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PS non-learners</td>
</tr>
<tr>
<td>/A-Ã/</td>
<td>44.5</td>
</tr>
<tr>
<td>/Q-Ã/</td>
<td>42.0</td>
</tr>
<tr>
<td>/A-Q/</td>
<td>11.2</td>
</tr>
</tbody>
</table>
Note. Error rates higher than 35% are considered to indicate ‘poor’ discrimination, whereas error rates lower than 14% are considered to indicate ‘good’ discrimination of the contrast.

Comparing, thus, the availability of specific cues and listeners’ performance (error rates), it appears that contrasts characterized by large spectral differences (/A-Q/) are better discriminated than contrasts characterized by large duration differences (/A-√/ and /Q-√/). In line with previous findings (McAllister et al., 2002; Morrison 2008b, 2009), spectral cues are weighted more than temporal cues in the monolingual groups whose L1 does not exploit duration, and the use of duration is not necessarily the initial strategy that inexperienced listeners use when processing a contrast that encodes both spectral and temporal cues. In the advanced groups, however, discrimination accuracy improves for most of the learners and contrasts, suggesting that learners’ perception has become attuned to the target language spectral and, to some extent, to temporal differences between vowels. The only exception is the advanced CS group for /A-√/ in (37.9% error rate). Based on the poor performance in the advanced CS group for the /A-√/ contrast, I speculate that, although available, temporal cues are still underexploited, even when spectral cues are insufficient. As such, the inability to perceive well the spectral cues that differentiate /A/ from /√/, which are separated by a small perceptual distance, combined with the ongoing development of the ability to use temporal cues, causes the high error rate for the /A-√/ contrast.

In order to further discuss the unusually high error rates in the CS advanced group, I compared the discrimination of /Q-√/ versus /A-√/ with experience and dialect. Although both contrasts are similar in terms of availability of spectral and temporal cues, only discrimination of the /Q-√/ contrast improves with experience in both dialectal groups. Discrimination of /A-√/ improves in the PS advanced group, but not in the CS advanced group. If the /A-√/ contrast is inherently difficult, why do PS advanced learners manage to discriminate it? Could it be because their native dialect allows them to map them as different categories or is it simply because of the learning experience? A strong argument in favour of the learning experience is the fact that the CS inexperienced group does not perform worse than the PS inexperienced group.

It is worth noting that results for L2 perception and L2 production suggest divergent strategies of using spectral and temporal cues in the CS group, for the /A-√/ contrast. Whereas in L2
perception, spectral cues are given more weight than temporal cues (again, the evidence is the poor discrimination of the /A √/ that have good temporal cues but insufficient spectral cues), in L2 production CS learners produce the /A √/ vowels with no significant spectral differences but with different durations. A possible explanation is that in L2 perception, CS advanced learners still rely heavily on spectral cues (a strategy transferred from L1) to discriminate vowels that differ both spectrally and temporally. In L2 production, however, when acquiring individual lexical items, CS learners may have trained themselves to produce short and long vowels to distinguish between, for example, hut - hot or but - bought. Although not specifically controlled for the frequency of occurrence, the target tokens were high frequency words (e.g. cut, cat, dog, cup, cap, shut) for which learners are likely to have strong lexical representations established early in the acquisition. Moreover, the nature of the sentence-completion task used to elicit L2 production and the instructions given at the time of testing, which focused on vocabulary retrieval, may have caused learners to produce the /A - √/ contrast the way they have stored it in the lexicon, i.e., with no spectral differences but with significantly different durations. As a follow up study, it would be interesting to compare how such high frequency words versus nonce or low frequency words with similar structures are produced and perceived by CS learners. An imitation task or a word learning task would likely show whether the /A- √/ contrast in such nonce words is perceived more like in the TL or whether the cue weighting strategies favour spectral cues against temporal cues.

Apart from the problematic case of the /A- √/ contrast in the advanced CS group, learners’ discrimination of the other contrasts, /Q- √/ and /A-Q/, shows that they have successfully established some interlanguage strategy to deal with contrasts that differ both spectrally and temporally. Discrimination scores for /Q-√/ and /A-Q/ are good in both groups of learners. While this section discussed the relationship between L2 perception and production from a qualitative perspective of the differential use of spectral and temporal cues, the next section presents a quantitative approach to the possible link between L2 perception and L2 production, by exploring correlations between perceptual and acoustic measures.
3.2 The relationship between L2 perception and L2 production

A question that emerges from this perceptual experiment is how discrimination of non-native contrasts relates to the production of vowels involved in these contrasts. Specifically, are contrasts that are difficult to discriminate (like the /A-/ or /Q-/) also produced less distinctively, i.e., separated by smaller distances? Moreover, do subjects who perceive contrasts well also produce them with a greater intervocalic distance? Findings for L1 tend to support the hypothesis that better perceivers also create more distinct contrasts and less variable categories (e.g. Perkell et al., 2004; but see Ainsworth and Paliwal, 1984, as cited in Perkell et al., 2004 that argue for a lack of a production-perception correlation). In L2 studies, Cebrian (2002) found several significant correlations on data pooled for all L2 learners between perception and production, namely, vowels that are better perceived, like /i, ɪ, ɐ/, but not /eɪ/, are also more accurately produced along the F1, F2, and duration dimensions. McAllister et al. (2004), however, did not find significant correlations between perception and production. In their exploration of the individual variation through multiple regression analyses, there was no clear relationship between perception and production of the duration contrast between Swedish vowels. Specifically, Spanish learners of Swedish with good lexical knowledge of Swedish (quantified as the number of times they had to be prompted to produce the target word in a sentence-completion task) were not necessarily either good perceivers or good producers of the target duration contrast. Similarly, in Peperkamp and Bouchon (2011), French-English bilinguals’ discrimination performance with the English /i-/ vowels correlated neither with how accurately they pronounced these vowels nor with how native-like they sounded. In sum, whereas in L1 good perceivers tend to have good pronunciation as well, in L2, findings are contradictory, possibly due to different L1-L2 segmental inventories, contrasts investigated, elicitation tasks and/or levels of proficiency in L2.

41 Peperkamp and Bouchon noted that when off-line tasks (identification or goodness rating) were used in data elicitation, correlations between perception and production were found. On the contrary, when online tasks were used (as their speeded ABX discrimination task (ISI of 500ms)), no such correlations were found. They argue that this happens because subjects have insufficient time to access the phonological loop during perception.
In order to assess the relation between perception and production in L2 for learners in the present study, I computed two series of Pearson product-moment correlation coefficients, first between the accurate discrimination of each of the contrasts (/A-\n/, /Q-\n/ and /A-Q/) (percentages of correct answers) and contrast distinctiveness (intervocalic distances between pairs of the /Q, \n, A/ vowels). The second set of correlations was computed between the accurate discrimination of the /A-\n/ and /Q-\n/ contrasts in perception and tense-lax duration ratios A/\n/ and Q/\n/ in production. There were no significant correlations to indicate that learners who discriminate better the L2 contrasts also produce them better. According to Peperkamp and Bouchon (2011, p. 164), a reason for the lack of correlation is that discrimination tasks with short ISI’s leave listeners little time “to access the phonological loop during perception and subvocally rehearse the stimuli”, hence their variable and imprecise production. This is also indirect evidence that phonetic, not phonological, processing was indeed obtained by selecting an ISI of 1 second in the AX discrimination task used in the present study. Corroborating these findings with the results on the correlations between accurate discrimination and vocabulary knowledge (see Chapter 4), I conclude that the learners in this study have varying abilities in L2 with respect to perception, production and vocabulary size. In other words, a good perceiver is not necessarily a good producer and someone with good vocabulary knowledge is not necessarily a good perceiver.

To summarize the discussion in this section, first, inexperienced L2 listeners use mainly spectral cues in discrimination, but not duration cues, because their error rates are high when spectral cues are poor, despite good duration cues. As such, it appears that, contrary to Bohn (1995), the use of duration is not a universally preferred strategy in the interlanguage of learners whose L1 has no duration contrasts, or at least it does not apply for the particular contrasts analyzed here. Second, L2 experienced learners have learned to discriminate most of the contrasts that exploit both spectral and duration cues, with the notable exception of the /\n-A/ in the CS advanced group. Whether the improved discrimination is due to the fact that listeners have mastered both spectral and temporal cues or only spectral cues, it still remains to be determined through an identification task in which spectral and temporal cues are varied orthogonally. Third, there is no

42 Arguments regarding the use of duration versus spectral cues are based on specific contrasts (/i-I/) but they may not hold for other vowel contrasts (like /Q-E/) as in Bohn and Flege (1992).
clear relationship between L2 perception and L2 production, but this is not surprising in L2, as it has been previously reported in the literature. Although advanced learners were tested here, their performance in perception and production has yet to become native-like.

4 Non-native production and native dialect influence in L2

The L2 production experiment aimed to highlight specific characteristics of the learners’ interlanguage vowels and a native dialect effect in L2. An additional goal was to corroborate L2 perceptual data. The influence of the native dialect on non-native production can be posited if specific differences between native vowels are matched in the realization of the non-native vowels. Specifically, in the PS dialect, some vowels (/i, o/) are more fronted and shorter and the vowel space is more dispersed as compared to the CS dialect. Consequently, I hypothesized that the TL vowels as produced by PS advanced learners would also be more fronted, shorter and more dispersed as compared to those produced by learners in the CS group. Such findings can be held as evidence of both dialect effect and transfer present at the advanced stages of acquisition. The results, however, reflect a much more complex picture, although in part they point towards what had been predicted. Namely, PS learners produce the /Q, V, A/ vowels with significant spectral differences and no duration difference whereas CS learners produce the /V, A/ vowels with no significant spectral difference but with duration differences, L2 vowels produced by PS are shorter than those of CS, and only L2 /V/ is more fronted for PS learners. The complexity of the findings originates from the great discrepancies between Spanish and English in terms of inventory size, dimensions/cues/features used to produce and perceive vowels. These discrepancies allow for various possibilities of mapping fewer native categories into multiple L2 categories. Moreover, several acquisition factors like TL input that learners have received, and learner variables that are difficult to control, like exposure to the TL, language use, other languages subjects have learned, contribute to the complex picture, making the task of isolating the native dialect effect in L2 production a difficult one. The following sections of the discussion present the arguments in favour and against native dialect effects in L2 production and discuss the role of the input in relation to the results obtained in L2 production.
4.1 Arguments in favour and against dialect effects in L2 production

As previously reported in the literature on Spanish varieties (see Chapter 3), vowels differ to some extent along acoustic parameters (formants, duration, formant movements, fundamental frequency and laryngeal adjustments) and possibly along articulatory parameters. In § 2, I discussed that the phonetic differences found for PS and CS vowels are not large enough for native speakers to detect. However, such differences play a much more important role in L2 as they may cause learners to adopt different strategies of mapping L1 sounds to L2. I discuss next how each of the differences found between PS and CS vowels influence the L2 vowels that they produce as arguments in favour of dialect effects.

The PS vowel space was found to be more dispersed towards the front (the distance between /i/ and the centroid is larger than in CS) and this seems to influence dispersion in L2, as the L2 vowels of the PS learners tend to spread more in the F1 x F2 space and there is a better category separation for the /Q, √, A/ vowels (i.e., greater intervocalic distances for the L2 vowels). Unfortunately, no other vowels than /Q, √, A/ were obtained in L2. However, given that the Spanish /i/ is significantly higher and more fronted in PS than in CS, I speculate that a sizeable native dialect effect in L2 would show in the realization of the TL /i, I/ vowels. As such, advanced PS learners would map the TL /i, I/ to two different spectral categories, possibly ignoring the duration differences whereas the CS learners would use duration to differentiate between spectrally similar vowels. A follow up study would look into the realization of the other TL vowels by PS and CS learners.

The difference between the realization of /o/ across the PS and CS dialects bears more relevance for the L1 effect on L2 production because of its proximity to the low and mid back TL vowels analyzed in this study. With the /a/’s at similar locations across dialects and with a more posterior /o/ in CS than in PS, the /a-o/ category boundary is also more posterior in CS than in PS. Moreover, /a, o/ vary less in CS than in PS along both F1 and F2 and there is no category overlap between /a/ and /o/ in CS, as opposed to PS (see Chapter 3, Figures 3.6 and 3.7). The fact that PS speakers allow more variation and category overlap for /a, o/ in L1 suggests that they may be using the entire /a-o/ continuum when mapping the TL /Q, √, A/ vowels, whereas CS
speakers are more conservative particularly with using the more distant /o/ category. This explains why their realizations of /ɜ, ʌ/ are merged into a single back category whereas those of the PS learners spread more towards the mid back /o/.

Another significant difference between PS and CS with potential influence in L2 production is the shorter /e, a, o/ vowels found for PS (see Chapter 3, § 5.5). L2 vowels produced by PS learners were also shorter than those of the CS learners (see Chapter 5, § 6.3). As such, the PS learners, accustomed to producing short vowels in L1, do not use duration at all as a cue to the TL tense-lax contrasts. On the contrary, CS learners, whose L1 vowels are longer and thus encode more perceptually relevant duration cues, also produce longer L2 vowels and moreover, use duration to create contrasts between spectrally similar L2 vowels.

Different variability patterns found in L1, namely, more variability along F1 in PS and along F2 for the CS front vowels, in addition to more dispersion towards the front in PS and a tendency to more dispersion towards the back in CS are likely reflected in the different patterns of vowel overlap in L2. Specifically, the preference for more posterior articulations in CS can be related to the /ɜ, ʌ/ overlap into a back category for the CS learners whereas the preference for the more fronted articulations in PS causes 6 out of 10 PS learners to produce the /ʊ, ɨ/ vowels as overlapping into a front category. Overall, PS learners produce the /ʊ, ɨ/ vowels as spectrally distinct, which can be attributed to the fact that there are more articulatory-acoustic resources towards the front of the oral cavity. Conversely, the /ɜ, ʌ/ are not spectrally distinct in the CS group because of the reduced possibilities of articulators to create the required acoustic output. There is also more perceptual confusion at the back, particularly for the /ɜ, ʌ/ contrast, reflected in the high error rates in discrimination for the CS learner group.

Having discussed several arguments in favour of native dialect effects in L2, I move next to the arguments against such effects. The differences found between the native vowels of PS and CS

---

43 An alternate explanation for the /o/-like vowels in L2-PS (in words like hot, caught) is the input they received, namely varieties of English in which there is no /A-ɔ/ merger.
are quantitatively small, although statistically significant and with large size effects.\textsuperscript{44} Specifically, the differences between the /i/'s are 9.8%, between the /o/'s are 10.3% along F2, and between durations 20% (see § 5.2.3). Translated into psychoacoustic measures, the /i/'s differ by approximately 283 mel, the /o/'s by approximately 150 mel and durations by 20 ms. I believe that such differences are likely irrelevant in L1,\textsuperscript{45} but may be contributing to the different organization of L2 sounds in the learners’ interlanguage, possibly in close relation to other factors like input, which is discussed in the next section.

4.2 The role of input in the interpretation of L2 production results

The interpretation of results of this cross-dialectal and cross-linguistic study must take into account the role of the input that learners tested here have received. Several factors that were shown to be related to input (Flege, 2010; Bohn & Bundgaard-Nielsen, 2009), like the age of acquisition (AOA), the language use, the length of residence in English-speaking countries (LOR), were controlled at the time of testing and both groups of learners had similar profiles. Proficiency scores computed from a vocabulary test were also found to be similar for the two groups of learners (66.6 for PS and 65.7 for CS, presented in Chapter 4, § 5.2.3). Both groups have been exposed to mixed native and non-native input, but the main difference is that the native input came from different varieties of English and at different stages of acquisition. Specifically, most of the PS learners have had an Irish English native teacher at some point in their studies whereas CS learners have been exposed to native speech via recorded materials, some of Canadian English during instruction. Moreover, learners in both groups also have heterogeneous exposure to various Englishes, including foreign-accented English, through traveling (PS group) and through contact with native and non-native English speakers (CS group). While some of these input-related factors can be and were controlled for in the present study, the TL variety that learners hear is difficult to control, particularly because of the

\textsuperscript{44} e.g. F(1, 160) = 45.8, p <.001, in Chapter 3, § 5.2.3
\textsuperscript{45} Going back to the minimal threshold calculations in § 3.1, the difference between the /o/'s across dialects are smaller than the thresholds, 268 mel in PS and 194 mel in CS. However, the differences between /i/'s are greater, so listeners may actually hear the cross-dialectal differences.
proximity of the testing locations to distinct English-speaking countries. The influence of the different target variety in the input on L2 production is discussed next.

According to Escudero and Boersma (2004), input is crucial as L2 learners acquire the perceptual strategies of the ambient language. Similarly, Munro, Derwing, and Flege (1999) found that Canadian English native speakers immersed in a dialect different from their own (Alabama English) adopt - to some extent - the speech characteristics of their environment. On the contrary, in Fox and Tevis McGory’s (2007) study, L2 learners resist adopting the characteristics of the ambient language if they regard it as a non-standard variety. All these studies draw conclusions on the influence of input based on L2 learners living for some time in the country of the TL and thus being in an immersion situation, characterized by rich input from native speakers of the TL and many opportunities to use the TL. However, participants in the current study are in a different setting. The L2 learners in the PS group are university students who use the TL exclusively for their English classes and receive both native and non-native input from their Irish English native and non-native instructors and peers. Similarly, the L2 learners in the CS group receive native and non-native input from the foreign tourists they interact with at work. Moreover, both groups of learners were interviewed in their native countries, Spain and Cuba, not immersed in the TL environment. Therefore, beyond their language classes and, respectively, their work hours, they do not hear nor use English. The reason for this detailed presentation of the learners’ background is to show that they have exposure to mainly heterogeneous target language input, despite some possible influences of the European English on PS learners and of North American English on CS learners. The possible influences of the TL varieties are discussed next in further detail.

In Chapter 5, I speculated that exposure to the English varieties that maintain the contrast /ʌ-/ (European English) may be the cause of the good spectral separation of L2 categories in the PS group. The extra category existing in their input led them to produce some of the /ʌ/-tokens with a more /o/-like quality and thus with a better /ʌ-ʌ/ separation. On the contrary, the two PS

---

46 This situation is classified as FLA (classroom foreign language acquisition) and contrasted to SLA, i.e. a more constrained learning context vs. a natural communicative context. Best and Tyler (2007) caution against conflating perception results from FLA listeners with SLA listeners.
learners who had more exposure to North American English produced the /\AA/-/\AE/ less distinctly. This is comparable to the overlap pattern that the majority of CS learners display and it may be originating from more exposure to North American English.

While the preference for a particular target variety (European English versus North American English) may be having a long-term impact on the L2 categories that learners have established, it is also possible that, at the time of testing, learners from both groups tried to imitate the variety of the two Canadian English native speakers, who gave the instructions and asked the questions in the L2 production elicitation task that lasted approximately 30 minutes. If present indeed, this phonetic convergence would have had the effect of reducing the effects of the different inputs that learners received, all remaining differences being attributable to the native varieties of the learners.

To summarize the discussion above, the effects of different target varieties in the input (European English versus North American English) on the L2 production of the /\AA/, /\AE/, /\AE/ vowels are moderated by several factors. First, such specific target variety input is restricted mainly to classroom setting for PS learners, and to the workplace for CS learners, respectively. Second, both groups of learners receive great amounts of non-native input from non-native instructors, peers, and tourists. Third, L2 learners in this study were tested in their home countries rather than in the TL country in which case the input would have arguably had a more significant effect. Fourth, participants may have converged to the Canadian English characteristics of the speakers in the recorded testing materials. All in all, the differences in the organization of sounds found in L2 production can be attributed in part to native dialect differences.

5 Final conclusions, relevance and contribution

The findings of the three experiments presented and discussed in this thesis contribute to a growing body of cross-dialectal and cross-linguistic research focusing on whether and how differences between native dialects are reflected in non-native production and perception of vowels. In the particular case of PS and CS participants listening to and producing the English /\AA/, /\AE/, /\AE/ vowels, the dialect itself is one of the several factors that generate the differences found between groups. The fact that the native dialect has only a moderate effect on L2 is readily explained by the small differences between the native vowels across the PS and CS dialects,
which may pass unnoticed by listeners from the opposite dialect. Greater differences between vowels across dialects, which can be easily perceived by listeners from other dialects (e.g. American English perceiving Australian English (Cutler et al., 2005), Canadian English judging Southern British English, American English and Australian English vowels (Bohn & Bungaard-Nielsen, 2009)), are more likely to generate stronger dialect effects in L2. Likewise, cross-dialectal differences might also be greater for varieties with larger vowel inventories than Spanish and thus be carried over to a greater extent in L2. The differences between the native vowels across dialects represent thus different starting points in the acquisition of the TL /Q, V, A/ vowels from which learners develop their L2 categories by adopting specific strategies of mapping L1 sounds to L2 in accordance to other modeling factors like heterogeneous input, mixed native and non-native input, different learning experiences, L3 and other foreign languages, individual variability. As such, the final conclusion is that the differences found in L2 production and perception are the result of the conjoint effect of several factors among which the native dialect, the input and the learning experience play an important role.

This finding bears relevance for second language acquisition studies because it shows the need to control the L2 learner’s native dialect and that mixing L2 learners from different varieties of a language may lead to erroneous findings. This caveat is particularly relevant when the L2 learners speak a language with a great amount of geographical variation because in such a case, the native dialect influence may have a strong effect in the realization of L2 sounds. This research contributed new data on Cuban Spanish and comparisons across dialects of Spanish as well as a whole corpus of acoustic measurements (as opposed to accentedness and intelligibility ratings) on L2 English. It also revealed specific patterns and interlanguage strategies that Spanish learners of English adopt and added to the sparse body of data on how the English /Q, V, A/ vowels are dealt with in L2.
References


Guitart, J. M. (1985). The resolution of phonological ambiguity in a simulated English-Spanish borrowing situation”. In L.D. King & C. A. Malley (Eds.) Selected Papers from the XLth linguistic symposium on Romance languages, University of North Carolina, Chapel Hill. (pp. 117-125) Amsterdam, Philadelphia: John Benjamins.


Appendix 1

Appendix 1 – L2 production – sentence-completion task

You will hear a set of short definitions and you will have to finish the sentence with one word that best suits the definition. After that, repeat the word in the sentence Say (word) once more.

Example 1:

You hear the sentence: a synonym of aircraft is _____

Your answer should be plane. Say plane once more.

Example 2:

You hear the sentence: a short form for prime minister is _____

Your answer should be PM. Say PM once more.

Example 3:

You hear the sentence: to creep - crept - crept and to sleep - _____ - _____

Your answer should be slept - slept. Say slept once more.

The test begins now.

A feline pet is a ____ (cat). Say (cat) once more.

The short form for doctor is ____ (doc). Say (doc) once more.

Mickey Mouse and Donald ____ (duck). Say (duck) once more.

A synonym of to return is to come ____ (back). Say (back) once more.
The name of a popular pizza restaurant is “Pizza ____” (hut). Say (hut) once more.

A canine pet is a ___(dog). Say (dog) once more.

An informal word for insect is ____ (bug). Say (bug) once more.

The short form for popular music is _____ (pop). Say (pop) once more.

A flower that has not opened yet is a ___ (bud). Say (bud) once more.

In summer when the temperature is 35 degrees, it is ____ (hot). Say (hot) once more.

An informal word for dollar is ____ (buck). Say (buck) once more.

You can take a bath in a bath ____ (tub). Say (tub) once more.

An informal way to say that someone has courage is to say that he has ____ guts. Say (guts) once more.

A knife is used to ____ (cut). Say (cut) once more.

You can write a note on a note- _____ (pad). Say (pad) once more.

A synonym of to embrace someone is to give him/her a ____ (hug). Say (hug) once more.

The opposite of good is ___ (bad). Say (bad) once more.

A synonym of however is ____ (but). Say (but) once more.

A synonym of marijuana is _____ (pot). Say (pot) once more.

To stick – stuck – stuck and to dig - ____ - ____ (dug – dug). Say (dug) once more.

A small animal that looks like a mouse that has wings and can fly is a ____ (bat). Say (bat) once more.

The rubber disk that hockey players skate after is called a ____ (puck). Say (puck) once more.

A brand of portable media players is the i-____(pod). Say (pod) once more.
A public place where you can have a drink is called a ____ (pub). Say (pub) once more.

A short form for Robert is ____Bob. Say (Bob) once more.

To catch – caught – caught and to buy - ____ - ____(bought - bought). Say (bought) once more.

A baby lion or baby bear is called a ____ (cub). Say (cub) once more.


A synonym of to close is to ____ (shut). Say (shut) once more.

A synonym for taxi is ___ (cab). Say (cab) once more.

An exclamation of surprise is: Oh, my ___! (God). Say (God) once more.

A baby dog is called a ___(pup). Say (pup) once more.

A label that shows the price is called a price-___ (tag). Say (tag) once more.

The opposite of bottom is ___ (top). Say (top) once more.

An informal word for policeman is ___ (cop). Say (cop) once more.

The short form for Patrick is _____ (Pat). Say (pat) once more.

You drink coffee from a ____ (cup). Say (cup) once more.


When you step in the subway train you have to mind the ____ (gap). Say (gap) once more.

An informal word for father is ___(dad). Say (dad) once more.

A piece of cloth with the nation’s colours and symbols is called a ____ (flag). Say (flag) once more.

To put one’s belongings into luggage for travel is to _____ (pack). Say (pack) once more.

The opposite of white is ____ (black). Say (black) once more.
To come – came – come and to run – ran - _____ (run). Say (run) once more.

Something to protect your head from sun is a _____ (hat). Say (hat) once more.

To run – ran – run and to come – came - ____ (come). Say (come) once more.

A small animal with spines on the back is a hedge-_____ (hog). Say (hog) once more.

You put your groceries in a plastic ____ (bag). Say (bag) once more.

When a player passes the ball to you, you have to _____ (catch) it. Say (catch) once more.

A pile of objects each on top of the last is called a ____ (stack). Say (stack) once more.

In your kitchen, water comes out of the ____ (tap). Say (tap) once more.

To dig – dug – dug and to stick - ____ - ____ (stuck – stuck). Say (stuck) once more.

The area between two streets is called a ___ (block). Say (block) once more.

An ugly old woman is called a ____ (hag). Say (hag) once more.

A synonym of swamp is ____ (bog). Say (bog) once more.

Broth made from meat or vegetables is called ___ (stock). Say (stock) once more.

The short form for capital letter is ____ (cap). Say (cap) once more.

The short form for ice-cappuccino is ice- ____ (cap). Say (cap) once more.
## Appendix 2

### Background information for participants in the study

<table>
<thead>
<tr>
<th>Dialect Group</th>
<th>Dialect Code</th>
<th>Sex</th>
<th>Age</th>
<th>AOO</th>
<th>Education/occupation</th>
<th>Current use h/wk</th>
<th>Other languages</th>
<th>Immersion in mths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peninsular Spanish Monolinguals</td>
<td>PS1_M</td>
<td>F</td>
<td>45</td>
<td>---</td>
<td>Psychology</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS2_M</td>
<td>F</td>
<td>48</td>
<td>---</td>
<td>Psychology</td>
<td>---</td>
<td>Cat, Fr</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS3_M</td>
<td>F</td>
<td>27</td>
<td>---</td>
<td>Teachers’ College</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS4_M</td>
<td>F</td>
<td>39</td>
<td>---</td>
<td>Tourism</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS5_M</td>
<td>F</td>
<td>36</td>
<td>---</td>
<td>Social work</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS6_M</td>
<td>F</td>
<td>27</td>
<td>---</td>
<td>Biology</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS7_M</td>
<td>F</td>
<td>28</td>
<td>---</td>
<td>Physical education</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS8_M</td>
<td>M</td>
<td>36</td>
<td>---</td>
<td>Teachers’ College</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS9_M</td>
<td>F</td>
<td>41</td>
<td>---</td>
<td>EGB</td>
<td>---</td>
<td>French</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>PS10_M</td>
<td>F</td>
<td>41</td>
<td>---</td>
<td>Biology</td>
<td>---</td>
<td>French</td>
<td>---</td>
</tr>
<tr>
<td>Advanced learners</td>
<td>PS1_B</td>
<td>F</td>
<td>42</td>
<td>6</td>
<td>Psychology</td>
<td>5</td>
<td>Fr, Germ</td>
<td>2 Canada</td>
</tr>
<tr>
<td></td>
<td>PS2_B</td>
<td>F</td>
<td>23</td>
<td>8</td>
<td>English</td>
<td>7</td>
<td>French</td>
<td>5 Ireland</td>
</tr>
<tr>
<td></td>
<td>PS3_B</td>
<td>F</td>
<td>23</td>
<td>8</td>
<td>English</td>
<td>20</td>
<td>Fr, Germ</td>
<td>2 Irl USA</td>
</tr>
<tr>
<td></td>
<td>PS4_B</td>
<td>F</td>
<td>23</td>
<td>6</td>
<td>English</td>
<td>20</td>
<td>Fr, Germ</td>
<td>9 UK</td>
</tr>
<tr>
<td></td>
<td>PS5_B</td>
<td>F</td>
<td>43</td>
<td>11</td>
<td>English</td>
<td>15</td>
<td>French</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>PS6_B</td>
<td>M</td>
<td>26</td>
<td>14</td>
<td>English</td>
<td>15</td>
<td>French</td>
<td>1 UK</td>
</tr>
<tr>
<td></td>
<td>PS7_B</td>
<td>F</td>
<td>40</td>
<td>8</td>
<td>English</td>
<td>5</td>
<td>---</td>
<td>4 UK</td>
</tr>
<tr>
<td></td>
<td>PS8_B</td>
<td>M</td>
<td>22</td>
<td>9</td>
<td>English</td>
<td>25</td>
<td>French</td>
<td>8 UK</td>
</tr>
<tr>
<td></td>
<td>PS9_B</td>
<td>M</td>
<td>25</td>
<td>6</td>
<td>English</td>
<td>25</td>
<td>Fr, Jap, Pol</td>
<td>10 UK</td>
</tr>
<tr>
<td></td>
<td>PS10_B</td>
<td>F</td>
<td>22</td>
<td>8</td>
<td>English</td>
<td>20</td>
<td>Fr, Russian</td>
<td>9 Irl</td>
</tr>
<tr>
<td>Cuban Spanish Monolinguals</td>
<td>CS1_M</td>
<td>F</td>
<td>28</td>
<td>---</td>
<td>Law/paralegal</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS2_M</td>
<td>M</td>
<td>26</td>
<td>---</td>
<td>Engineering</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS3_M</td>
<td>F</td>
<td>41</td>
<td>---</td>
<td>IT</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS4_M</td>
<td>F</td>
<td>38</td>
<td>---</td>
<td>Engineering</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS5_M</td>
<td>F</td>
<td>33</td>
<td>---</td>
<td>Chefs school</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS6_M</td>
<td>F</td>
<td>26</td>
<td>---</td>
<td>Preuniversitaria</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS7_M</td>
<td>F</td>
<td>35</td>
<td>---</td>
<td>Preuniversitaria</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS8_M</td>
<td>M</td>
<td>32</td>
<td>---</td>
<td>Engineering</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS9_M</td>
<td>F</td>
<td>36</td>
<td>---</td>
<td>Preuniversitaria</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS10_M</td>
<td>M</td>
<td>31</td>
<td>---</td>
<td>Engineering</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Advanced learners</td>
<td>CS1_B</td>
<td>F</td>
<td>24</td>
<td>6</td>
<td>English</td>
<td>20</td>
<td>French</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS2_B</td>
<td>F</td>
<td>25</td>
<td>6</td>
<td>English/sales</td>
<td>20</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS3_B</td>
<td>F</td>
<td>33</td>
<td>6</td>
<td>English</td>
<td>20</td>
<td>French</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS4_B</td>
<td>M</td>
<td>30</td>
<td>7</td>
<td>English</td>
<td>20</td>
<td>French</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS5_B</td>
<td>F</td>
<td>33</td>
<td>6</td>
<td>English</td>
<td>20</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS6_B</td>
<td>F</td>
<td>38</td>
<td>7</td>
<td>Teachers’ College</td>
<td>20</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>CS7_B</td>
<td>F</td>
<td>34</td>
<td>6</td>
<td>English</td>
<td>20</td>
<td>---</td>
<td>30 Canada</td>
</tr>
<tr>
<td></td>
<td>CS8_B</td>
<td>F</td>
<td>23</td>
<td>6</td>
<td>Computer science</td>
<td>&gt;20</td>
<td>French</td>
<td>26 Canada</td>
</tr>
<tr>
<td></td>
<td>CS9_B</td>
<td>M</td>
<td>23</td>
<td>11</td>
<td>Medical school</td>
<td>&gt;20</td>
<td>---</td>
<td>5 Canada</td>
</tr>
</tbody>
</table>
## Appendix 3

### Stimuli used in the AX discrimination task (experiment 2)

<table>
<thead>
<tr>
<th>Wave File Name</th>
<th>Expected Answer</th>
<th>Response Time</th>
<th>Subject Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>cat1_CE1_caught3_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cup1_CE1_cap3_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cup1_CE1_cop1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug2_CE1_bag1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dug2_CE2_dog1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gut1_CE1_got1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>back1_CE1_back1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bad1_CE1_bat1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bag2_CE2_bag1_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stuck1_CE2_stack1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>block1_CE2_block1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bog1_CE2_bog2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bought1_CE1_bought3_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stuck3_CE1_stock1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>buck2_CE2_puck2_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bud1_CE1_bat2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug1_CE1_bug2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cap1_CE2_cap3_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug2_CE2_bog1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but1_CE1_bought_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duck2_CE2_duck2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but2_CE1_bat3_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cab2_CE1_cab3_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cap1_CE1_cap4_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hag2_CE2_hag1_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but1_CE2_bat2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cap3_CE1_cop2_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat1_CE2_caught1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cop1_CE1_cop2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>buck2_CE1_back1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cop1_CE2_cop2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cub1_CE2_cab2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bag2_CE2_bog2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bat1_CE1_bat2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hut1_CE1_hat1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bat1_CE1_bought1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>black1_CE2_black1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cub1_CE2_cub2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hag3_CE2_hog1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>back1_CE1_pack1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hut1_CE2_hat1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cup1_CE1_cop1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug2_CE2_bug2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cut1_CE1_cat1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hag1_CE2_hog2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 1</td>
<td>Wave 2</td>
<td>Comparison</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>cap1_CE2_cap3_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug2_CE2_bog1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but1_CE1_bought_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duck2_CE2_duck2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but2_CE1_but3_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cab2_CE1_cab3_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cap1_CE1_cap4_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hag2_CE2_hag1_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but1_CE2_but2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cap3_CE1_cop2_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat1_CE2_caught1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cop1_CE1_cop2_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>buck2_CE1_back1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cop1_CE2_cop2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cub1_CE1_cab2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bag2_CE2_bog2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bat1_CE1_bat2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hut1_CE1_hat1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bat1_CE1_bought1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>black1_CE2_black1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cub1_CE2_cub2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hug3_CE2_hog1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>back1_CE1_pack1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hut1_CE2_hat1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cup1_CE1_cop1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug2_CE2_bug2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cut1_CE1_cat1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hag1_CE2_hog2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>but1_CE1_bat1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cut2_CE2_caught1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>doc1_CE2_doc1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hog1_CE1_hog1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bog2_CE1_bog1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bug2_CE2_bag1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hug2_CE2_hug1_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duck2_CE2_dug2_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dug1_CE1_dog1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>god1_CE1_got1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cup2_CE1_cup1_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gut1_CE1_got1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duck1_CE1_doc1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hag1_CE1_hag2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cup1_CE1_cap3_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hat1_CE2_hot2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bat1_CE2_bat2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hog1_CE2_hog1_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hug1_CE1_hug2_CE1.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stuck3_CE1_stuck2_CE2.wav</td>
<td>same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cut1_CE1_caught3_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hug3_CE2_hag1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cut2_CE2_cat1_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>doc2_CE1_dog2_CE2.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hut1_CE1_hot1_CE1.wav</td>
<td>different</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
bought1_CE2_bought2_CE2.wav  same
pack1_CE1_pack1_CE2.wav  same
cat1_CE1Caught3_CE1.wav  different
pad1_CE1_pad1_CE2.wav  different
pat2_CE1_pot1_CE2.wav  different
hat1_CE1_hot1_CE1.wav  different
pod1_CE2_pot1_CE1.wav  different
pop1_CE2_pop1_CE1.wav  same
puck1_CE1_pack1_CE2.wav  different
hut1_CE2_hot2_CE2.wav  different
puck2_CE2_puck2_CE1.wav  same
pup1_CE2_pup1_CE1.wav  same
stack1_CE2_stock1_CE1.wav  different
stock1_CE1_stock1_CE2.wav  same
pod1_CE2_pad1_CE1.wav  different
stuck3_CE1_stuck1_CE2.wav  same
top1_CE1_tap1_CE2.wav  different

Distracters in Experiment 2:

bud1_CE1_but2_CE2
duck2_CE2_dug2_CE1
god1_CE1_got1_CE2
doc2_CE1_dog2_CE2
pad1_CE1_pad1_CE2
pod1_CE2_pot1_CE1
bad1_CE1_bat1_CE1
back1_CE1_pack1_CE1
buck2_CE2_puck2_CE1
Appendix 4

Mean formant values (Hz) of Spanish vowels and standard deviations for the PS and CS dialects

<table>
<thead>
<tr>
<th>Vowels</th>
<th>PS</th>
<th>CS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>sd</td>
</tr>
<tr>
<td>/i/</td>
<td>394</td>
<td>61</td>
</tr>
<tr>
<td>/e/</td>
<td>519</td>
<td>62</td>
</tr>
<tr>
<td>/a/</td>
<td>755</td>
<td>85</td>
</tr>
<tr>
<td>/o/</td>
<td>517</td>
<td>59</td>
</tr>
<tr>
<td>/u/</td>
<td>412</td>
<td>43</td>
</tr>
</tbody>
</table>
### Appendix 5

Mean F1, F2, F3 (Hz) and duration (ms) values for L2 vowels as produced by L2 learners from PS and CS and Canadian English vowels

<table>
<thead>
<tr>
<th>Subject code</th>
<th>/Q/ F1</th>
<th>/Q/ F2</th>
<th>/Q/ F3</th>
<th>/Q/ dur</th>
<th>/I/ F1</th>
<th>/I/ F2</th>
<th>/I/ F3</th>
<th>/I/ dur</th>
<th>/A/ F1</th>
<th>/A/ F2</th>
<th>/A/ F3</th>
<th>/A/ dur</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS1</td>
<td>887</td>
<td>1751</td>
<td>3040</td>
<td>226</td>
<td>868</td>
<td>1688</td>
<td>3011</td>
<td>140</td>
<td>634</td>
<td>1069</td>
<td>3442</td>
<td>174</td>
</tr>
<tr>
<td>PS2</td>
<td>1008</td>
<td>1633</td>
<td>2717</td>
<td>93</td>
<td>912</td>
<td>1597</td>
<td>2657</td>
<td>91</td>
<td>740</td>
<td>1225</td>
<td>2969</td>
<td>81</td>
</tr>
<tr>
<td>PS3</td>
<td>789</td>
<td>1831</td>
<td>2540</td>
<td>120</td>
<td>645</td>
<td>1323</td>
<td>2384</td>
<td>106</td>
<td>649</td>
<td>1101</td>
<td>2562</td>
<td>120</td>
</tr>
<tr>
<td>PS4</td>
<td>930</td>
<td>1546</td>
<td>2393</td>
<td>110</td>
<td>899</td>
<td>1533</td>
<td>2491</td>
<td>109</td>
<td>766</td>
<td>1281</td>
<td>2558</td>
<td>123</td>
</tr>
<tr>
<td>PS5</td>
<td>933</td>
<td>1780</td>
<td>2824</td>
<td>110</td>
<td>746</td>
<td>1382</td>
<td>3376</td>
<td>99</td>
<td>761</td>
<td>1214</td>
<td>3433</td>
<td>99</td>
</tr>
<tr>
<td>PS6</td>
<td>691</td>
<td>1551</td>
<td>2383</td>
<td>114</td>
<td>629</td>
<td>1392</td>
<td>2356</td>
<td>100</td>
<td>609</td>
<td>1176</td>
<td>2402</td>
<td>111</td>
</tr>
<tr>
<td>PS7</td>
<td>859</td>
<td>1523</td>
<td>2580</td>
<td>97</td>
<td>829</td>
<td>1569</td>
<td>2681</td>
<td>112</td>
<td>651</td>
<td>1081</td>
<td>2863</td>
<td>113</td>
</tr>
<tr>
<td>PS8</td>
<td>708</td>
<td>1455</td>
<td>2726</td>
<td>90</td>
<td>709</td>
<td>1395</td>
<td>2756</td>
<td>80</td>
<td>597</td>
<td>1058</td>
<td>2868</td>
<td>91</td>
</tr>
<tr>
<td>PS9</td>
<td>639</td>
<td>1714</td>
<td>2464</td>
<td>87</td>
<td>638</td>
<td>1313</td>
<td>2384</td>
<td>77</td>
<td>643</td>
<td>1138</td>
<td>2431</td>
<td>87</td>
</tr>
<tr>
<td>PS10</td>
<td>855</td>
<td>1606</td>
<td>2488</td>
<td>83</td>
<td>842</td>
<td>1591</td>
<td>2480</td>
<td>78</td>
<td>722</td>
<td>1268</td>
<td>2708</td>
<td>81</td>
</tr>
<tr>
<td>CS1</td>
<td>777</td>
<td>2013</td>
<td>2806</td>
<td>146</td>
<td>770</td>
<td>1262</td>
<td>2608</td>
<td>126</td>
<td>786</td>
<td>1196</td>
<td>2611</td>
<td>135</td>
</tr>
<tr>
<td>CS2</td>
<td>934</td>
<td>1968</td>
<td>3092</td>
<td>145</td>
<td>809</td>
<td>1509</td>
<td>3079</td>
<td>137</td>
<td>818</td>
<td>1401</td>
<td>3060</td>
<td>151</td>
</tr>
<tr>
<td>CS3</td>
<td>955</td>
<td>1623</td>
<td>2317</td>
<td>181</td>
<td>921</td>
<td>1442</td>
<td>2706</td>
<td>135</td>
<td>923</td>
<td>1421</td>
<td>2683</td>
<td>204</td>
</tr>
<tr>
<td>CS4</td>
<td>895</td>
<td>1549</td>
<td>2537</td>
<td>200</td>
<td>825</td>
<td>1307</td>
<td>2604</td>
<td>180</td>
<td>795</td>
<td>1255</td>
<td>2552</td>
<td>204</td>
</tr>
<tr>
<td>CS5</td>
<td>857</td>
<td>1480</td>
<td>2502</td>
<td>112</td>
<td>777</td>
<td>1190</td>
<td>2585</td>
<td>113</td>
<td>771</td>
<td>1200</td>
<td>2681</td>
<td>137</td>
</tr>
<tr>
<td>CS6</td>
<td>806</td>
<td>1883</td>
<td>2598</td>
<td>143</td>
<td>772</td>
<td>1356</td>
<td>2415</td>
<td>114</td>
<td>778</td>
<td>1364</td>
<td>2512</td>
<td>122</td>
</tr>
<tr>
<td>CS7</td>
<td>906</td>
<td>1646</td>
<td>2547</td>
<td>110</td>
<td>792</td>
<td>1393</td>
<td>2662</td>
<td>97</td>
<td>821</td>
<td>1391</td>
<td>2872</td>
<td>114</td>
</tr>
<tr>
<td>CS8</td>
<td>898</td>
<td>1645</td>
<td>2505</td>
<td>120</td>
<td>790</td>
<td>1397</td>
<td>2585</td>
<td>93</td>
<td>784</td>
<td>1264</td>
<td>2663</td>
<td>121</td>
</tr>
<tr>
<td>CS9</td>
<td>678</td>
<td>1244</td>
<td>2035</td>
<td>85</td>
<td>623</td>
<td>1121</td>
<td>2061</td>
<td>80</td>
<td>558</td>
<td>1011</td>
<td>2137</td>
<td>91</td>
</tr>
<tr>
<td>CE1</td>
<td>972</td>
<td>1642</td>
<td>2298</td>
<td>146</td>
<td>870</td>
<td>1480</td>
<td>2577</td>
<td>111</td>
<td>801</td>
<td>1160</td>
<td>2371</td>
<td>164</td>
</tr>
<tr>
<td>CE2</td>
<td>893</td>
<td>1644</td>
<td>2420</td>
<td>141</td>
<td>819</td>
<td>1527</td>
<td>2425</td>
<td>96</td>
<td>780</td>
<td>1221</td>
<td>2290</td>
<td>140</td>
</tr>
<tr>
<td>CE3</td>
<td>948</td>
<td>1711</td>
<td>2185</td>
<td>140</td>
<td>826</td>
<td>1511</td>
<td>2475</td>
<td>89</td>
<td>831</td>
<td>1247</td>
<td>2625</td>
<td>133</td>
</tr>
<tr>
<td>CE4</td>
<td>928</td>
<td>1631</td>
<td>2387</td>
<td>108</td>
<td>829</td>
<td>1555</td>
<td>2484</td>
<td>76</td>
<td>838</td>
<td>1253</td>
<td>2591</td>
<td>115</td>
</tr>
<tr>
<td>CE5</td>
<td>979</td>
<td>1623</td>
<td>2511</td>
<td>122</td>
<td>885</td>
<td>1403</td>
<td>2675</td>
<td>86</td>
<td>823</td>
<td>1077</td>
<td>2702</td>
<td>129</td>
</tr>
<tr>
<td>CE6</td>
<td>915</td>
<td>1739</td>
<td>2865</td>
<td>118</td>
<td>812</td>
<td>1700</td>
<td>2967</td>
<td>77</td>
<td>799</td>
<td>1280</td>
<td>3236</td>
<td>111</td>
</tr>
</tbody>
</table>
Copyright Acknowledgements