This thesis examines the acquisition of the English lax-tense contrast in front vowels (/i/-/u/, /e/-/e/) by adult Catalan learners of English as a second language. The English lax-tense distinction is based on differences in tongue position, duration and phonotactic distribution, where lax vowels are restricted to closed syllables. Importantly, Catalan has a similar vowel height distinction (/i/, /e/, /e/), but no duration contrast or phonotactic restrictions for vowels. Further, English, but not Catalan, has ambisyllabic consonants. These specific differences raise the following two questions: How do second language learners acquire a novel segmental contrast, and how does the acquisition of nonsegmental information interact with the acquisition of segmental structure? These questions are addressed in a series of perception, production and psycholinguistic experiments.

The perceptual similarity of English and Catalan vowels is assessed to predict relative success in vowel acquisition. Results from perception and production experiments illustrate that perceived similarity alone does not predict the likelihood of accurate perception and production. Specifically, duration, a non-L1 feature, was used to establish a contrast for a vowel that had no close match in Catalan. This in turn influenced the cues used in the
categorization of the other vowel in the pair despite having a near identical match in the L1. Further, differences in L1 and L2 cue weighting for highly similar vowels may affect perception accuracy. These findings demonstrate that factors such as relative weighting of phonetic cues and the need to establish oppositions must be taken into account when predicting success in acquisition.

The role of nonsegmental information is investigated using syllable manipulation and picture matching tasks. Catalan learners of English demonstrate knowledge of English vowel phonotactics. However, use of this constraint is limited by interference from L1 syllabification, which in turn may affect production ability. Additional evidence for the interaction of segmental and nonsegmental structure in L2 acquisition comes from correlated results across experiments.

Taken together, the findings underscore the need to regard vowels as members of a contrastive system rather than individually acquired categories and to consider nonsegmental information in the acquisition of L2 phonological structure.
Acknowledgements

I would like to express my gratitude to many different people without whose help and support this thesis could not have been possible.

First and foremost, I am grateful to my supervisor, Ron Smyth. He has been a consistent source of encouragement, guidance and support. His genuine interest and enthusiasm kept me focussed and helped me to successfully complete my Ph.D. I am also thankful to Keren Rice, who also has been a constant source of encouragement and advice throughout my UofT years. Thanks also to Hank Rogers for his time, effort and advice, to Pascal van Lieshout, for his insightful comments and generosity in letting me use the facilities at the Speech Pathology department, and to Murray Munro, for kindly agreeing to act as an external appraiser and for giving invaluable advice and encouragement for future work. All these people deserve my deepest appreciation for their assistance in making the defence date possible when the timing was so tight.

My sincere thanks go too to all the members of the Linguistics department who have made my graduate student life a time of intellectual and personal growth. In particular, I thank my fellow graduate students who have shared with me the ups and downs of graduate life. I am indebted to many, but I especially want to express my gratitude to Abdul-Khalig Ali, David Bennett and Trisha Causley for their friendship, Susana Bejar, Luba Butska and Naomi Cull for their support in the last stages of the thesis, Nila Friedberg, Alexei Kochetov and Hoa Pham for their inspiration as forerunners of PhD completion, and to Carolyn Smallwood, my eternal graduate school peer and intermittent housemate. Finally, special thanks go to Craig Chambers for making this crazy last year of the thesis a truly memorable one. Heartfelt thanks to all of you.
I would also like to express appreciation to all the participants and consultants in my experiments, both in Canada and in Barcelona. Special thanks go to the members of the Casal Català de Toronto, the students at the Universitat Autònoma de Barcelona and my colleagues at UofT who participated in the experiments. These experiments would not have been possible without the help of Rick Dale and Michael Kennedy, who lent me their programming knowledge and native speaker voice (respectively). Thanks also go to my colleagues at UAB who covered my teaching duties during recent trips to Toronto.

Finally, I would like to thank my friends and family back home for their long lasting support, encouragement, and patience.
Table of Contents

Chapter One. A study of second language vowel acquisition ............................................. 1
1.0. Introduction .......................................................................................................................... 1
1.1. Main issues in the acquisition of second language phonology ......................................... 1
1.2. Learning L2 segments: The effect of phonetic similarity .................................................. 2
1.3. Empirical research on L2 vowel acquisition ........................................................................ 5
1.4. The relative weighting of different acoustic cues .............................................................. 11
1.5. The acquisition of non-segmental phonological characteristics ........................................ 14
1.6. The current study: Goals, assumptions and hypotheses ...................................................... 17
1.7. Overview of the thesis ........................................................................................................ 25

Chapter Two. Phonological characteristics of English and Catalan ....................................... 28
2.0. Introduction .......................................................................................................................... 28
2.1. The target language: English ............................................................................................... 28
2.1.1. Tense and lax vowels in English .................................................................................... 28
2.1.2. Spectral and temporal differences .................................................................................. 31
2.1.3. The Lax Vowel Constraint ............................................................................................. 34
2.1.3.1. The status of phonotactics in the phonology ............................................................. 35
2.1.3.2. Productivity and violability of the Lax Vowel Constraint ............................................. 38
2.1.4. The syllabification of intervocalic consonants .............................................................. 42
2.1.4.1. Psycholinguistic evidence for CVC-VC syllabification ............................................ 44
2.2. The L1: Catalan .................................................................................................................... 48
2.2.1. The Catalan vowel system .............................................................................................. 48
2.2.2. Catalan syllable structure ............................................................................................. 50
2.3. Acoustic comparison of Catalan and English vowels ......................................................... 52
2.4. Conclusion .......................................................................................................................... 54

Chapter Three. Perceived similarity between Catalan and English vowels ............................... 56
3.0. Introduction ......................................................................................................................... 56
3.1. Measuring phonetic similarity ............................................................................................ 56
3.2. Perceptual assimilation of English vowels to Catalan vowels ........................................59
  3.2.1. Stimulus preparation ..................................................................................................59
  3.2.2. Procedure ..............................................................................................................60
3.3. Experiment 1. Catalan native speakers in Ontario .......................................................62
  3.3.1. Subjects ..................................................................................................................62
  3.3.2. Results ...................................................................................................................62
3.4. Experiment 2. Catalan native speakers in Catalonia ....................................................65
  3.4.1. Subjects ..................................................................................................................65
  3.4.2. Results ...................................................................................................................66
3.5. Experiment 3. Perceptual assimilation of Catalan vowels to English vowels ............71
  3.5.1. Subjects and procedure ........................................................................................71
  3.5.2. Results ...................................................................................................................71
3.6. Discussion and conclusions ..........................................................................................73

Chapter Four. Perception of L2 vowels ...........................................................................77
4.0. Introduction ..................................................................................................................77
4.1. Vowel identification involving natural stimuli .............................................................79
  4.1.1. Subjects ..................................................................................................................80
  4.1.2. Stimuli ....................................................................................................................81
  4.1.3. Procedure ..............................................................................................................85
  4.1.4. Experiment 4. Vowels in isolation ........................................................................85
  4.1.5. Experiment 5. Vowels in monosyllables .................................................................90
  4.1.6. Experiment 6. Vowels in bisyllabic environments ..................................................93
  4.1.7. Conclusions .........................................................................................................98
4.2. Experiment 7. Vowel identification involving a synthetic vowel continuum ..........102
  4.2.1. Subjects ................................................................................................................105
  4.2.2. Stimuli ..................................................................................................................105
  4.2.3. Procedure .............................................................................................................106
  4.2.4. Results ..................................................................................................................107
  4.2.5. Conclusions .........................................................................................................118
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>F1 and F2 values of English high and mid front vowels for male speakers of British English and American English.</td>
<td>32</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>F1 and F2 values of English high and mid front vowels for female speakers of British English and American English.</td>
<td>32</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>Duration ratio /i/-/ɪ/ as reported in different studies.</td>
<td>33</td>
</tr>
<tr>
<td>Table 2.4</td>
<td>Formant values of Catalan high and mid front vowels.</td>
<td>49</td>
</tr>
<tr>
<td>Table 2.5</td>
<td>Mean durations in milliseconds for Catalan high and mid front vowels.</td>
<td>49</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>PAT results with the Catalan group in Canada. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings.</td>
<td>62</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>PAT results with the Catalan group in Canada. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings across talkers.</td>
<td>63</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>PAT results with the Catalan group in Catalonia. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings.</td>
<td>66</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>PAT results with the Catalan group in Catalonia. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings across talkers.</td>
<td>67</td>
</tr>
<tr>
<td>Table 3.5</td>
<td>Identification and goodness ratings of English vowels and perceptual assimilation of Catalan vowels to English vowels by English native speakers.</td>
<td>72</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Formant values for vowel /ɛ/ in Experiment 4 (isolated vowels).</td>
<td>83</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Formant values for vowel /ɛ/ in Experiment 5 (CVC words).</td>
<td>83</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Formant values for vowel /ɛ/ in Experiment 6 (CVCVC words).</td>
<td>83</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Vowel durations in ms by talker for each experiment.</td>
<td>84</td>
</tr>
</tbody>
</table>
Table 4.5  English native speakers. Identification scores for English target vowels in isolation and means across talkers.

Table 4.6  Catalans in Toronto (TC group). Identification scores for English target vowels in isolation and means across talkers.

Table 4.7  Catalans in Barcelona (BC group). Identification scores for English target vowels in isolation and means across talkers.

Table 4.8  Catalans in Toronto (TC Group). Identification scores for English target vowels in the CVC context and means across talkers.

Table 4.9  Catalans in Barcelona (BC Group). Identification scores for English target vowels in the CVC context and means across talkers.

Table 4.10 English native speakers. Identification scores for English target vowels in the CVCVC context and means across talkers.

Table 4.11 Catalans in Toronto (TC Group). Identification scores for English target vowels in the CVCVC context and means across talkers.

Table 4.12 Catalans in Barcelona (BC Group). Identification scores for English target vowels in the CVCVC context and means across talkers.

Table 4.13 Correct identification by vowel and group for each experiment.

Table 4.14 Formant values for English high and mid front vowels.

Table 4.15 Percentages of ‘beat’ responses for English native speakers.

Table 4.16 Percentages of ‘beat’ responses for Catalan speakers in Barcelona.

Table 4.17 Percentages of ‘beat’ responses for Catalan subjects in Toronto.

Table 4.18 Percentage of ‘bit’ responses for English native speakers.

Table 4.19 Percentage of ‘bit’ responses for Catalan subjects in Barcelona.

Table 4.20 Percentages of ‘bit’ responses for Catalan subjects in Toronto.

Table 4.21 Percentages of ‘bet’ responses for English native speakers.

Table 4.22 Percentages of ‘bet’ responses for Catalans in Barcelona.

Table 4.23 Percentages of ‘bet’ responses for Catalan subjects in Toronto.
Table 4.24  Correct identifications of each prototypical vowel in each length condition.

Table 5.1  F1 and F2 frequencies in Hertz for the female L2 and native speakers in this study, plus reported data.

Table 5.2  F1 and F2 frequencies in Hertz for the male L2 and native speakers in this study, plus reported data.

Table 5.3  Mean vowel duration and tense/lax vowel duration ratios in the L2 data.

Table 5.4  Mean vowel duration and tense/lax vowel duration ratios for the English native speakers.

Table 5.5  Duration ratio /i/-/ə/ as reported in different studies.

Table 5.6  Percentages of correct identification of the target vowels produced by the L2 learners, listed by word type.

Table 5.7  Mean goodness ratings per vowel and environment for Catalan subjects.

Table 5.8  Listening tests results and correlations.

Table 6.1  Percentages of CVC-VC syllabifications in the Repetition task.

Table 6.2  Percentages of CVC-VC syllabifications in the Reverse Order task.

Table 6.3  Percentages of LVC-motivated responses obtained in the picture matching task by both groups.

Table 6.4  Percentage of final unmodified responses in the singular derivation task by both groups.

Table 6.5  Percentages of LVC-obeying responses per group and task.

Table 7.1  Summary of results for each vowel by experiment.

Table 7.2  Percentages of LVC-obeying responses per group and task.
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Catalan vowel inventory.</td>
<td>48</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Distribution of English and Catalan vowels in the F1-F2 vowel space according to the values presented in Tables 2.1 and 2.4</td>
<td>53</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Perceptual assimilation of English vowels to Catalan vowels by native Catalan speakers in Canada</td>
<td>64</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Perceptual assimilation of English vowels to Catalan vowels by native Catalan speakers in Catalonia.</td>
<td>67</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Percentages of ‘beat’ responses for English native speakers.</td>
<td>109</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Percentages of ‘beat’ responses for Catalan subjects in Barcelona.</td>
<td>109</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Percentages of ‘beat’ responses for Catalan subjects in Toronto</td>
<td>110</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Percentage of ‘bit’ responses for English native speakers.</td>
<td>111</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Percentage of ‘bit’ responses for Catalan subjects in Barcelona.</td>
<td>111</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Percentages of ‘bit’ responses for Catalan subjects in Toronto.</td>
<td>112</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Percentages of ‘bet’ responses for English native speakers.</td>
<td>113</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Percentages of ‘bet’ responses for Catalan subjects in Barcelona.</td>
<td>113</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>Percentages of ‘bet’ responses for Catalan subjects in Toronto.</td>
<td>114</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>F1 and F2 values for vowels /i/, /u/ and /e/ in the h_d condition.</td>
<td>132</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>F1 and F2 values for vowels /i/, /u/ and /e/ in the h_b condition.</td>
<td>132</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>F1 and F2 values for vowels /i/, /u/ and /e/ in the h_pus condition</td>
<td>133</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Mean vowel durations per vowel and environment conditions for the Catalan learners of English</td>
<td>137</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>Mean vowel durations per vowel and environment conditions for the English native speakers.</td>
<td>137</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Correct identification of the target vowels produced by the L2 learners in each environment.</td>
<td>143</td>
</tr>
<tr>
<td>Figure 5.7</td>
<td>Mean goodness ratings per vowel and environment for Catalan subjects.</td>
<td>148</td>
</tr>
</tbody>
</table>
# List of Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Descriptive statistics for the Catalan native speakers' results (the BC group).</td>
<td>215</td>
</tr>
<tr>
<td>B</td>
<td>Syllable manipulation tasks: List of test words.</td>
<td>217</td>
</tr>
<tr>
<td>C</td>
<td>Picture labelling tasks: List of test words.</td>
<td>219</td>
</tr>
<tr>
<td>D</td>
<td>Picture labelling tasks: Pictures.</td>
<td>220</td>
</tr>
</tbody>
</table>
Chapter One. A study of second language vowel acquisition

1.0. Introduction

It is well attested that adult second language (L2) speech rarely achieves the degree of proficiency (native-like pronunciation) that is characteristic of first language (L1) speech. Research in L2 acquisition attempts to explain this difference. This thesis addresses this question by examining the acquisition of a second language vowel contrast and by evaluating the effect of non-segmental phonological characteristics such as syllabification and phonotactics on the acquisition of the contrast. This chapter reviews the claims and findings of the most relevant theories on the acquisition of L2 vowels, outlines the assumptions of the current thesis, and presents the hypotheses and predictions to be tested in subsequent chapters.

1.1. Main issues in the acquisition of second language phonology

L2 speech is commonly characterized by the failure to sound like native speech, particularly when L2 learning starts after childhood (Ellis, 1994; Long, 1990; Patkowski, 1990; Flege, Munro and MacKay, 1995). This has often been associated with neurological change arising as the result of normal maturation (Lenneberg, 1967; Penfield and Roberts, 1959; Scovel, 1988). In addition to the age of learning the L2, other factors, including the amount of exposure to the L2 and the relative use of the L2 and L1, have been proposed as responsible for the difficulty in achieving native-like speech (Best and Strange, 1992; Flege, 1995, 1997a; Flege, Schmidt and Wharton, 1996; Flege, Bohn and Jang, 1997). In addition, psychological factors such as motivation or
attitude towards the target language speaking community, and social factors such as age, 
gender, and social class, among others, also affect the process of second language 
acquisition (Ellis, 1994).

One important linguistic source of difficulty in acquiring a second language is 
interference from the first language (L1). L1 interference refers to the way the learner's 
existing linguistic knowledge influences the course of acquisition of the L2. For example, 
the phonology of the L1 may cause learners to filter out acoustic differences that are not 
phonemically relevant in the L1 (Trubetzkoy, 1939). Thus, errors that characterize non-
native speech may be due to incorrect representations of L2 sounds (Rochet, 1995; Flege, 
1995). The most commonly discussed cases of L1 interference in interlanguage 
phonology involve the segmental level, but interference can also affect other aspects of 
the phonology, including phonological rules and processes, syllable structure, 
phonotactics, stress patterns and intonation (Ellis, 1994; Tarone, 1987; Sato, 1984; 

1.2. Learning L2 segments: The effect of phonetic similarity

The awareness of the important role of L1 interference gave rise to the 
Contrastive Analysis Hypothesis, proposed by Lado (1957). Lado hypothesized that all 
nonnative deviations are due to interference or negative transfer, such that those elements 
in the L2 that are similar to the learner's L1 will be simple for the learner to acquire 
whereas those that are different will be difficult. Empirical evidence soon cast doubt on 
the strong claim made by this hypothesis and research turned to a more careful
investigation of the factors that interact with L1 interference, including the degree of crosslinguistic similarity (Ellis, 1994).

The notion of perceived similarity between L1 and L2 sounds is crucial in the development of current theories of L2 category formation. Research has shown that learners are language-specific perceivers and that previous linguistic experience determines the way L2 sounds are perceived and categorized. For example, Rochet (1995) found that native Portuguese subjects tended to misidentify French /y/ tokens as /i/ whereas native English subjects tended to misidentify the same vowel tokens as /u/. Further, in a repetition task, native Portuguese subjects produced /i/-quality vowels when they heard French /y/ tokens whereas native English subjects tended to produce /u/. The difference could be explained by the more fronted articulation of English /u/, which is thus closer to /y/ than Portuguese /u/.

A general account of this effect is provided by Kuhl's Native Language Magnet model (Kuhl, 1993, 1991). According to this theory, language specific vowel prototypes interfere with the adult learners' abilities to perceive some L2 contrasts by acting as a perceptual magnet that pulls L2 vowels towards L1 prototypes. L2 vowels that are located near an L1 vowel prototype are discriminated less readily than vowels that are not located near an L1 prototype.

Two models that further investigate the patterns of perceptual assimilation of L1 and L2 sounds are the Perceptual Assimilation Model (Best and Strange, 1992; Best, 1994, 1995) and the Speech Learning Model (Flege, 1995). Both models generate predictions of relative perceptual ability on a range of the possible perceptual relations between native and non-native sounds. The models differ in that the former supports a
gestural basis for perception whereas the latter favours an auditory or psychoacoustic basis.

Like the Native Language Magnet, the Perceptual Assimilation Model proposes that an L2 learner's ability to perceive the L2 contrast will depend on the degree to which an L2 contrast can be assimilated to L1 categories. The model describes several ways in which L2 vowel categories can be assimilated to L1 vowels and predicts L2 discrimination ability accordingly. An L2 sound may be assimilated to an L1 category (i.e. categorized as an L1 sound) and may be perceived as a good, an acceptable or a deviant exemplar of that L1 category. Alternatively, the sound may be uncategorizable in terms of L1 sounds (i.e., falling within the L1’s phonological space but between L1 categories). A third possibility is that the sound falls outside the L1’s phonological system, that is, it is not heard as a speech sound. The model predicts that category formation for L2 sounds is more likely to occur in the case of L2 sounds that are moderately similar to L1 sounds than in the case of sounds that are very similar to L1 sounds or else are too dissimilar to be assimilated to any L1 category (Best and Strange, 1992). The model also proposes that discrimination between L2 sounds may improve as a function of L2 experience.

According to the Speech Learning Model (e.g., Flege, 1995), successful acquisition of L2 sounds depends on the ability to create target-like perceptual categories for them. For sounds that are not found in the L1, this is possible if the learner is able to discern the phonetic differences between the L2 and the closest L1 sounds. This ability becomes less likely as age of learning increases. The mechanism of equivalence classification (Flege, 1987) may prevent formation of a new category. Thus, the model
claims that adult learners, with sufficient experience and exposure to the L2, will be more likely to establish new and more authentic phonetic categories for sounds that do not have an easily identifiable counterpart in the L2, i.e., "new" sounds, than for "similar" sounds, which will be readily substituted by their L1 counterparts (Bohn and Flege, 1997). One important characteristic of the Speech Learning Model is that the L1 and the L2 share a common phonological space. Thus a bilingual’s L2 category may be deflected away from an L1 category to maintain contrast within the system. In addition, it is possible that an L2 category may be based on different features or feature weights than the corresponding category in a monolingual speaker (Flege, 1995). Finally, the Speech Learning Model proposes a close link between perception and production, claiming that the accuracy with which L2 sounds are produced is limited by the accuracy of the perceptual categories for those sounds.

1.3. Empirical research on L2 vowel acquisition

The relationship between the age at which L2 learning commences and the ability to establish native-like categories has been examined in a number of studies, which have also addressed the effect of experience in the L2. Flege, MacKay and Meador (1999) evaluated the production and perception of English vowels by three groups of Italian native speakers who differed in the age of arrival in Canada. They found that the results for the two groups of subjects who had arrived at an earlier age (at 7 and 14 years of age) did not differ significantly from the results obtained with native English speakers. However, the late-arriving subjects (around 19 years of age) were significantly less successful. This supports the prediction of the Speech Learning Model that the likelihood
of establishing native-like categories decreases as the age of learning increases. Similar results were obtained by Munro, Flege and MacKay (1996) also with Italian learners of English with respect to goodness ratings of English vowel productions. However, different results were obtained by Pallier, Bosh and Sebastián-Gallés (1997), who found that Spanish-Catalan bilinguals whose first language was Spanish but who had grown up in a bilingual environment since childhood were not able to perceive the Catalan /e/-/ɛ/ contrast in a native-like manner, thus showing no effect of experience in the perception of a new vowel contrast. Flege, MacKay and Meador point out that the discrepancy between the two studies could be accounted for by the amount of L1 use, given that the Spanish/Catalan bilinguals may have had more opportunity to use their L1 (Spanish) than the early Italian/English bilinguals.

Research has shown that L2 learning can heighten learners' awareness of crosslinguistic phonetic differences. In a multidimensional scaling study, Fox, Flege and Munro (1995) found that more advanced Spanish learners of English resembled native English speakers more closely in their use of certain perceptual dimensions in vowel identification than less advanced Spanish learners of English. In a study by Flege (1991), a group of Spanish monolinguals and a group of Spanish learners of English used Spanish letters to identify English /i, ɪ, ɛ, æ/ as one of the five Spanish vowels /i, e, a, o, u/ or selected a ‘none’ response if deemed different from the L1 vowels. In general, Spanish learners responded "none" significantly more often than Spanish monolinguals, and experienced learners chose more "none" responses than inexperienced learners, suggesting that sensitivity to phonetic differences between L2 and L1 sounds may increase with L2 experience.
With respect to the effect of perceived similarity on the formation of accurate representations, studies yield different results. Flege, Munro and Fox (1994) examined the perceived dissimilarity of pairs of vowels drawn from Spanish and/or English by Spanish learners of English and English monolingual speakers. Vowel pairs consisting of a Spanish /i/ token and an English /i/ token obtained comparable dissimilarity ratings by experienced and inexperienced Spanish subjects. Thus, English /i/ did not emerge as a "new" vowel and experience had little effect on the perception of this vowel. However, both Spanish and English speakers could apparently distinguish the two /i/'s because they rated pairs consisting of a Spanish /a/ and an English /i/ as being more dissimilar than Spanish /a/ and Spanish /i/. With respect to English /u/, both native English speakers and Spanish speakers rated the vowels in /a/-/u/ pairs as less dissimilar than vowels in /a/-/i/ pairs. However, Spanish subjects rated the vowel pair /u/-/i/ as significantly more similar than did the English speakers. Thus, learners may be able to discriminate between two target vowels but still fail to categorize them as separate vowels.

Flege's (1997b) study of Dutch learners of English illustrates a case of identical or near identical sounds. Following acoustic measurements, Dutch /u/ was classified as identical to English /u/ in F1-F2 space, and Dutch /i/ was classified as similar to English /i/ because it is lower in the acoustic space. In vowel identification tasks performed by English speakers, the ratings for English /i/ and /u/ produced by experienced and inexperienced Dutch learners were as high as those produced by native English speakers (97-93%). This was predicted for the identical vowel /u/ but not for the similar vowel /i/.

The author claims that, even if not as identical as /u/, Dutch /i/ may be so close to the
English counterpart that the substitution of the Dutch vowel for the similar L2 vowel does not reduce intelligibility. Similar cases are the high identification rates for Spanish production of English /e/ and German production of English /i/ and /ɛ/ (Flege, 1993), indicating that near identical vowel substitution goes unnoticed by native listeners.

Turning now to studies examining the acquisition of new sounds, Busà (1992) found that nearly all Italian learners tested tended to produce English [u], classified as phonetically similar to Italian /u/, with F2 values lower (i.e. more Italian-like) than those of native English speakers. However, in the case of the new L2 vowel [ʊ], although the F1 and F2 values tended to be lower than for native English speakers, a larger number of Italians produced [ʊ] tokens with native-like spectral and temporal properties than was the case for [u].

The predicted effect of experience on the acquisition of new vowels is further examined in studies by Bohn and Flege. Bohn and Flege (1990) suggest that experienced German learners of English are eventually more likely to establish a new phonetic category for English /æ/, a new vowel, than for English /i, ɪ, ɛ/, which are perceptually similar, if not identical, to their German counterparts. Results of vowel identification tasks indicated that L2 experience did not affect perception of /i, ɪ, ɛ/; both experienced and inexperienced learners distinguished them. But with the new vowel /æ/, experienced learners resembled the English native speakers more closely, indicating that new vowels can be acquired by adult L2 learners with sufficient exposure. Bohn and Flege (1992) tested the production of these English vowels by German learners both acoustically and by intelligibility assessment in a labelling experiment. Both experienced and
inexperienced L2 speakers produced the similar English vowels /i, ɪ, ɛ/ with equal accuracy. As for /æ/, acoustic measurements (but not intelligibility results) showed that the experienced but not the inexperienced learners produced the new vowel /æ/ in much the same way as the English speakers. Thus, experience had some positive effect on the creation of new vowel categories.

Results from other studies run counter to the predicted effect of experience. Flege and Bohn (1992) examined Spanish learners' production of English vowels as assessed by the number of productions rated as correct by native listeners. The ratings for /i/ and /i/, considered on the basis of spectral measurements to be similar and new, respectively, were 57% for /i/ and 61% for /i/ for experienced late learners and 69% /i/ and 51% /i/ for inexperienced late learners. Similarly, both experienced and inexperienced learners obtained higher identification scores for the similar vowel /ɛ/ (99% and 91%, respectively) than for the new vowel /æ/ (73% and 70%, respectively). These results fail to support the claim that experienced late learners will master new L2 vowels (such as /i/ and /æ/) or that they would acquire the similar vowel /i/ and the new vowel /i/ differently.

Similarly, in the study of perceptual similarity discussed above, Flege (1991) found that English /i/ was identified with Spanish /i/ by both monolingual subjects and L2 learners. In the case of English /æ/, Spanish monolinguals yielded the highest identifications as Spanish /i/, and experienced learners provided the most "none" identifications. Nevertheless, only a few subjects used the "none" label consistently for /æ/ (and /æ/, another new vowel) lending little support to the expected effect of experience on the
acquisition of new vowels. Finally, in their study of Italian speakers’ production of English vowels, Munro, Flege and MacKay (1996) found no clear relationship between phonetic distance between L1 and L2 vowels and success in production as measured by goodness ratings. For example, neither /ɛ/ (as in err) nor /æ/ occur in Italian but the latter was much more accurately produced than the former. On the other hand, the production of English /i/, which is quite close to Italian /i/, was more often judged to sound English-like. These results do not lend support to the hypothesis that more similar sounds will be produced more accurately than more distant sounds.

In sum, the results from different studies are not entirely consistent on how perceived dissimilarity affects the ability to create new vowel categories. One possible source of these discrepancies is the accurate classification of L2 sounds as new or similar, a classification that is not always straightforward. For example, Flege (1987) observed that productions of a new vowel, the French [y], by native English learners of French showed more French-like F2 frequencies than did their productions of a similar vowel, the French [u]. These results were interpreted to support the claim that more accurate representations will be formed for new sounds than for similar sounds. However, Rochet (1995) observed that neither French /y/ nor /u/ has the acoustic properties of the English prototype /u/. In addition, both French vowels were perceptually categorized as /u/ by English listeners, casting doubt on Flege’s the classification of French /u/ and /y/ as a similar and new vowel respectively. This points to the need for a reliable way of measuring crosslinguistic phonetic similarity. As seen above, Bohn and Flege's (1992) production study reveals a discrepancy between acoustic measurements and intelligibility ratings of L2 vowel production. Similarly, Flege (1997b) assumes first
that English /i/ is a similar vowel to Dutch /i/ on the basis of acoustic comparisons but results from listening tests indicate that Dutch /i/ may go unnoticed in English, which would mean that it is a near identical sound rather than a similar sound. One strategy for addressing this problem (which is adopted in the current study) is to supplement acoustic measurements with perceptual assimilation tasks, as advocated in recent L2 speech research (Strange, 1999; Rochet, 1995; Flege, Bohn and Jang, 1997).

1.4. The relative weighting of different acoustic cues

Another important issue in vowel acquisition is the use and relative weighting of acoustic cues. For example, it has been hypothesized that vowel length in the L1 influences the perceptual salience of the distinction between tense and lax vowels (Flege, Munro and Fox, 1994). Some studies lend support to this view. In a study of L2 vowel perception in which vowel duration and spectrum were varied, Bohn and Flege (1990) found that inexperienced German learners of English, but not experienced ones, mainly used duration cues in differentiating the English /e/-/æ/ pair. The authors conclude that when spectral cues are insufficient to differentiate an L2 vowel contrast, duration will be used. This factor was not found with the English /i/-/ɪ/ contrast, which was distinguished by all learners on the basis of vowel quality. This indicates that duration cues are exploited in differentiating /e/ from the new vowel /æ/, whereas there is no need to resort to such cues with /i/-/ɪ/, a vowel contrast that has a similar counterpart in an L1 vowel contrast based on vowel quality.
Minnick-Fox and Maeda (1999) found that native speakers of Japanese, a language with contrastive length distinctions for vowels of nearly identical quality (e.g., /i/ and /i:/), relied on duration as a primary cue to identify English high front vowels (/i/-/i/). Similarly, Altenberg and Vago (1987) found that highly experienced Hungarian speakers of English substituted Hungarian short /i/ for English /i/. Hungarians tend to regard the contrast between English /i/ and /i/ as analogous to that of Hungarian /i:/ and /i/, leading them to equate Hungarian short /i/ with English /i/. Finally, in a study involving Arabic native speakers, Munro (1993) found that the Arabic speakers exaggerated duration differences between English tense and lax vowels in the fashion of the long vs. short vowel contrast in Arabic, even after more than fifteen years in the United States. The results in Minnick-Fox and Maeda (1999), Altenberg and Vago (1987), and Munro (1993) thus lend support to the hypothesis that the phonological status of vowel duration in the L1 influences the perceptual salience of temporal variations in the L2 (Flege, Munro and Fox, 1994).

McAllister, Flege and Piske (1999) tested the hypothesis that an L2 contrastive category will be difficult to acquire if it is based on a phonetic feature not exploited in the L1. They evaluated the production and perception of the Swedish phonemic vowel length contrast by native speakers of four languages that differ in the use they make of duration as a vocalic cue. The subjects were speakers of Swedish, Estonian, a language with an overall prominence of duration in the prosody, English, which makes use of both spectral and durational cues in distinguishing vowels, and Spanish, a language with no length distinction. The results of both production and perception tests showed that Estonians
ranked very close to the native speakers, followed by English speakers, and Spanish
speakers were the least successful. These results indicate that success in learning the
length contrast is related to the role of the duration feature in the L1.

Nevertheless, Bohn (1995) and Flege and Bohn (1992) claim that late learners can
detect temporal differences between a pair of unfamiliar L2 vowels more readily than
spectral differences, a claim that is supported by research showing that temporal
variations also seem salient for speakers whose L1 has no length contrast. In particular,
studies have shown that L2 learners rely heavily on durational cues for the English /i/-/ɪ/
contrast, regardless of whether length is contrastive in the L1. Flege, Bohn and Jang
(1997) compared the production and perception of English vowels by German, Spanish,
Korean, Mandarin and native English speakers. They found that native speakers
produced larger spectral differences but small temporal differences between /i/ and /ɪ/,
and /ɛ/ and /æ/. On the other hand, many non-native speakers failed to produce spectral
differences but instead produced exaggerated length differences. With respect to
perception, subjects were tested on synthetic stimuli with varying temporal and spectral
cues. Whereas English speakers relied mainly on spectral differences, many non-natives
relied on temporal cues.

In addition, Bohn and Flege (1990; Flege and Bohn, 1989) found that not only
inexperienced German learners of English but also Spanish learners, whose L1 has no
length contrast, relied on duration when asked to identify the members of synthetic vowel
continua whose endpoints did not contrast in their L1. Flege and Bohn (1992) compared
the production of English /i/ and /ɪ/ by early and late Spanish learners. They found that
all learners were able to produce temporal contrasts, but only early learners produced a significant spectral contrast. Flege and Bohn (1989) found that Spanish learners exploited temporal cues to a greater extent than spectral cues in their identification of English /i/ and /u/, with no two clear separate L1 counterparts, but overreliance on duration was not evident with /e/ and /æ/, which may be identified in terms of Spanish /e/ and /a/. Thus duration was used in the absence of L1-like spectral contrast. The predominance of temporal cues with Spanish learners of English was also found by García-Lecumberri and Cenoz (1997) for perception. The use of temporal cues in both perception and production was also found with Mandarin learners of English (Wang, 1997; Wang and Munro, 1999).

To summarize, evidence from a range of studies indicates that, at least initially, L2 learners may exploit temporal cues to a greater extent than spectral cues to identify unfamiliar vowels in a second language. In addition, even when new (non-L1) categories are established for an L2 sound, these categories may be based on a different cue or cue weighting than the target language's categories. Consequently, production of the L2 sound may not be accurate.

1.5. The acquisition of non-segmental phonological characteristics

In addition to learning which phonetic segments are phonologically distinctive and which phonetic variants are appropriate in which contexts (allophonic constraints), L1 language learners come to know the syllabic and prosodic structure and the phonotactic constraints of their language (Lindblom, 1990; Strange, 1999). A limited
number of studies have evaluated the suprasegmental factors in the acquisition of L2 phonology, focusing primarily on interference effects from the L1.

L1 interference can cause phonotactic modifications, that is, constraints on syllable structure in the L1 may affect the acquisition of L2 structures (e.g., cases of consonant cluster simplification or schwa epenthesis). Prosodic effects may also occur (e.g., using syllable timing for a stress-timed language, or transferring intonational patterns). Studies have shown that the L1 syllable and stress patterns are indeed carried over to the L2 (Tarone, 1987; Chela-Flores, 1996; Trammell, 1993; Sato, 1984; Broselow, 1987, among others). In addition, universal tendencies to avoid codas and onsetless syllables are also present in L2 acquisition (Eckman, 1977; Tarone, 1987; Youssef and Mazurkevich, 1998).

With respect to syllable structure, Trammell (1993) found that the transfer of L1 syllabification patterns results in allophonic and even phonemic errors in L2 English. As examples, Trammell mentions errors involving consonant length, vowel length, nasalization, aspiration, alveolar stop flapping, or the pronunciation of dark [l]. Trammell also explains that speakers of syllable-timed languages have difficulty with the lack of clear syllable boundaries in languages with ambisyllabic such as English. Similarly, Delattre (1965) found evidence that the different syllabification patterns of American English speakers and French speakers resulted in differences in pronunciation of English sequences. More specifically, when producing the phrase take it, native English speakers showed the effect of ambisyllabic since the word final consonant /k/ in take was anticipated and had a reduced burst and a short closure. French learners of English, on the other hand, showed little or no anticipation, long closure and a stronger release burst,
probably due to the preference for resyllabification found in French. Of relevance to the current study, Catalan learners have also been found to transfer syllabification patterns into their L2 English. Cebrian (1996, 2000) observed that Catalan speakers resyllabified a word-final consonant with a following onsetless word-initial vowel. Cebrian (1999) tested the syllabification of middle intervocalic consonants by native English speakers and Catalan learners of English and found that the two groups differed in their syllabification strategies. The English native speakers, but not the Catalan learners of English, showed a tendency to syllabify singleton intervocalic consonants with the preceding vowel, especially when the vowel was lax. The effect of this difference in syllabification on the perception and production of the lax-tense vowel contrast will be examined in the current study. As we have seen then, L1 prosodic structure is likely to cause interference in L2 acquisition. The transfer of an L1 syllabification pattern that is different from that of the L2 may be responsible for segmental mispronunciation in the L2.

With respect to the effect of L1 phonotactic constraints, Rochet and Putnam Rochet (1999) found that the phonotactic constraint that governs the distribution of the high and mid front unrounded vowels in French affects their L2 English. French /i/, /e/ and /ɛ/ are contrastive in open syllables but the contrast is neutralized in closed syllables, where only /i/ and /ɛ/ are contrastive. This L1 phonotactic constraint accounts for the French learners' substitution patterns in the case of the final vowels in happy and presented (Rochet and Putnam Rochet consider these to be realizations of an unstressed high front vowel /ɪ/). This L2 vowel is perceived and produced as /e/ in an open syllable
(e.g., happy), and as /i/ in a closed syllable (e.g., presented). Thus, an L2 vowel category may be perceived as belonging to two different categories depending on the context the L2 vowel appears in and whether this context constrains segmental contrasts in the L1.

The examples discussed so far show that, along with the L1 sound inventory, other characteristics of the L1 phonology can have an effect on the shaping of the L2 system. Importantly, L1 syllabification patterns are likely to interfere with L2 syllabification and may affect the way L2 sounds are acquired. Similarly, L1 phonotactic constraints may have a negative effect on L2 perception and production abilities.

1.6. The current study: Goals, assumptions and hypotheses

Based on the models and studies reviewed in the previous sections, the most relevant findings and conclusions about the acquisition of L2 sounds by adult learners can be summarized as follows:

- Adult learners evaluate L2 vowels sounds partly from the perspective of their L1 phonology. Thus, L2 sounds will initially be assimilated to L1 sounds, i.e., perceived and categorized as belonging to some L1 sound category.
- L2 sounds will be assimilated to the phonetically closest L1 sounds. This mapping will be based on the vowel properties that are contrastive in the L1.
- The similarity of L1 and L2 sounds cannot be determined on the basis of acoustic measurements (e.g., F1-F2 distances) alone. Similarity must be assessed from the perspective of the perceiver by means of perceptual assimilation tasks.
- Subsequent experience in the L2 may allow learners to discern differences between L2 and L1 vowels, a prerequisite for new vowel category formation.
• L2 vowel categories may differ from the target categories in that they may be based on different features or feature weights than the target categories.

• Non-segmental phonological knowledge from the L1, such as syllabification processes, may cause interference and affect L2 segmental acquisition.

Although these statements provide the starting points for an understanding of how L2 categories are acquired, a number of important questions remain.

Question 1. How do second language learners implement a contrast that is largely based on a distinction that is not found in the L1? The available literature on this topic is often unclear or contradictory. One goal of this thesis is to explore this problem in greater detail.

As we have seen, the notion of phonetic similarity has been a central consideration in predicting the ability to form accurate L2 categories. This study follows current trends in L2 research in assessing phonetic similarity using perceptual assimilation tasks. In these tasks, monolingual listeners are presented with L2 speech stimuli, and asked to indicate to which L1 phonetic category each L2 token is most similar, and to rate its "goodness" as an exemplar of that category (Ingram and Park, 1997; Flege, Bohn and Jang, 1997; Strange, 1999). Three different results may be obtained:

a) An L2 sound is consistently categorized as a good instance of one L1 vowel (that is, consistently assimilated to one L1 vowel) and obtains good ratings,
b) An L2 sound is consistently assimilated to one L1 vowel but with poor goodness ratings, or
c) An L2 sound is identified as instances of multiple L1 vowels with poor goodness ratings.

These three alternatives correspond to the classification of L2 sounds as ‘identical,’ ‘similar’ and ‘new’, respectively, in Flege’s (1987) Equivalence Classification Hypothesis.

As discussed above, different models make different predictions based on this classification. The traditional contrastive analysis (Lado, 1957) predicts better acquisition of more similar sounds than of new sounds. Conversely, Flege’s (1995) Speech Learning Model predicts that the greater the perceived dissimilarity between L1 and L2 sounds, the greater the likelihood that a new and authentic (i.e., target-like) category is established, given sufficient exposure and experience. Finally, a third possibility is outlined by the Perceptual Assimilation Model, which agrees with the Speech Learning Model in that learners will be more likely to establish new categories for L2 sounds that are moderately dissimilar than for those that are very similar to L1 sounds. However, this model differs in its prediction that learners are unlikely to form a new category for L1 sounds that are so dissimilar that they cannot be related to any L1 category (Best and Strange, 1992).

As discussed above, empirical evidence yields inconclusive results that do not fully support any alternative. However, an important distinction needs to be made between the ability to create a new (i.e., non-L1) representation for an L2 sound and the accuracy of that representation. As we have seen, learners may establish new categories
to represent L2 sounds but these categories may differ from target language categories. For instance, Munro, Flege and MacKay (1996) found that few of the English vowels produced by the adult Italian speakers were rated as good productions. However, an intelligibility test showed that most L2 productions were correctly identified, indicating that Italians were producing a larger number of vowels than the total number of vowels in their L1 inventory. Thus, the Italian speakers had established new L2 vowel categories which were identifiable as intended. However, the inaccuracy of most of these L2 categories resulted in low goodness ratings.

This fact illustrates that perceived (dis)similarity may allow us to predict whether or not L2 learners will establish new categories. In that respect, I follow the Speech Learning Model and the Perceptual Assimilation Model in their claim that learners will be more likely to establish a new vowel category if they are able to discern phonetic differences between the target sound and the closest L1 sound. However, accuracy of L2 categories will depend on the ability to establish those new categories based on the same vowel features and feature weights as in the target language.

It follows then that learners will be unlikely to establish new categories for strongly assimilated vowels because the perceived similarity between the L1 and L2 sounds will prevent the detection of phonetic differences between the two sounds. Thus, the L1 category will be used to perceive and produce the L2 vowel. The use of the L1 category, however, may result in successful perception and production in the case where the sounds are practically identical, as measured by the fact that substitution of the L1 sound for the target sound may go unnoticed by target language listeners, a possibility that is also contemplated in the Speech Learning Model. With respect to non-identical
sounds, success in the perception and production of the L2 vowels will depend on the accuracy of the new vowel representation in terms of target-like cues and cue weighting.

As discussed above, there are different views about the role of various phonetic cues in the formation of L2 vowel categories, in particular with respect to vowel duration and vowel quality. According to Bohn (1995) and Flege and Bohn (1992), learners will note temporal differences more readily than spectral differences regardless of whether length is distinctive in the L1. This view runs counter to McAllister, Flege and Piske (1999) and Flege (1995), who predict that L2 features that are not contrastive in the L1 will be difficult for late learners to perceive and produce. These incompatible outcomes suggest there is a need to reevaluate the particular conditions under which non-L1 features are attended to.

As we have seen, the starting point for adult learners is to assimilate L2 vowels to L1 vowels according to vowel properties or features that are contrastive in the L1. When categories must be established for relatively new L2 vowels, it is proposed that features that are not contrastive in the L1 may become more noticeable. For example, the fact that vowels in the L1 are contrasted spectrally will facilitate learners' initial detection of spectral contrasts among L2 sounds. However, if the L1 spectral distinctions cannot be directly used to establish new L2 categories, learners may become more attentive to other non-L1 features and exploit them in their categorization of the L2 sounds. Thus, L2 learners may become attuned to duration contrasts that are absent in the L1 when an L2 contrast is based on both temporal and spectral differences but the spectral differences find no match in the L1. Conversely, when duration is the sole distinction between two target vowels and it is not distinctive in the L1, learners may not readily detect the
difference between the L2 sounds, which in turn eliminates the need to use duration cues to construct new categories. There is in fact some evidence for this prediction in McAllister, Flege and Piske (1999), who observed a greater use of duration by Spanish speakers in their perception and production of Swedish high and low vowels, for which there is both a spectral and temporal distinction, than for mid vowels, which only differ in duration. A third case would be that of a spectral and temporal vowel contrast in the L2 for which the L1 has a similar spectral contrast. In this case the prediction is that the temporal difference would not be readily noticed.

In the case under study, the target vowel contrasts in the L2 (English tense and lax pairs) involve both spectral and temporal differences. The L1, Catalan, has vowel contrasts that are based on the spectral dimension only. The prediction is that if each vowel in an English tense and lax pair is strongly assimilated to a Catalan vowel (based on spectral characteristics), temporal differences will not be prominent. However, if no Catalan spectral match can be found for one (or both) of the English vowels, duration may come to play a more prominent role, in accordance with the view in Bohn (1995).

In summary, L2 vowels will be assimilated to L1 vowels on the basis of L1 distinctive features. Perceived dissimilarity between L1 and L2 sounds is necessary for new vowel category formation. Successful perception and production of L2 vowels will depend on whether the L2 vowel categories are established on the basis of the same feature and feature weights as in the target language. Finally, inventory building is influenced by the need to maintain sound oppositions between the existing and the new vowels. This leads to an additional hypothesis, namely that in the process of constructing an L2 vowel system, the need to accommodate new vowels and establish new vowel
oppositions may affect the categorization of neighbouring vowels, even that of L2 categories that are a good match with L1 categories.

Question 2. How does the acquisition of nonsegmental information interact with the acquisition of segmental categories?

The acquisition of the L2 phonological system is not limited to learning an inventory of sound contrasts. Knowledge of prosodic and phonotactic constraints is also necessary, as discussed above. The current study addresses this issue by focusing on the acquisition of L2 phonotactic constraints and syllabification processes and how this knowledge influences the acquisition of L2 sounds. The acquisition of the English tense-lax contrast provides an opportunity for investigating this issue. This is because, as discussed in greater detail in the next chapter, the contrast between English lax and tense vowels is partly based on nonsegmental information. The lax vowels, but not the tense vowels, are restricted to closed syllables. No stressed lax vowels are found in open syllables. This L2 phonotactic constraint, known as the Lax Vowel Constraint, is a novel nonsegmental feature for the Catalan learner of English. Moreover, English and Catalan differ also in syllable structure in that the former, but not the latter, is characterized by less defined syllable boundaries and ambisyllabicity.

This study will thus investigate the acquisition of segmental as well as nonsegmental structure. The importance of both L1 and L2 syllabification in the shaping of the L2 phonological system has been discussed above. For instance, recall that failure to acquire L2 syllabification, and English ambisyllabicity in particular, has been found to affect segmental inaccuracy (Trammell, 1993; Delattre, 1965). An issue that needs to be
addressed is whether phonotactics of the L2 play any role in the process of acquiring an L2, that is, whether L2 phonotactics are actually learned by L2 learners. On the one hand, learners may not actually need to learn phonotactics because a faithful representation of the segmental sequence for a given word should be sufficient to allow for accurate perception and production. On the other hand, there is evidence that L1 phonotactics are learned and are represented as a distinct component of the grammar, as will be discussed in the next chapter. Thus, it is at least plausible that L2 phonotactics are also acquired as part of a speaker's phonological knowledge. If so, it should be possible to demonstrate the effect of L2 phonotactic constraints on the course of L2 acquisition. Thus, this thesis will examine the acquisition of non-segmental structure and its interaction with the acquisition of segmental structure.

Some preliminary evidence relevant to this issue is provided by Cebrian (1996), who found that adult Catalan learners of English were more successful in producing a lax-tense contrast between /i/ and /i/ in CVC words (e.g., heat–hit, leave–live) than in bisyllables with one intervocalic consonant (heating–hitting, leaving–living). It was hypothesized that this was due to the joint effect of knowledge of the English Lax Vowel Constraint and Catalan-like syllabification of the middle consonant. As discussed in the next chapter, native English speakers often syllabify a middle consonant with a preceding stressed vowel if that vowel is lax, thus preserving the closed syllable environment. However, if the Catalan learners of English apply their L1 onset maximization syllabification (as expected from the high transferability of L1 prosodic structure in L2 acquisition), the lax vowel is left in an open syllable and will consequently violate the lax vowel constraint. The Catalan speakers' solution is to produce the vowel as the tense
counterpart. This explains the lesser success in maintaining the contrast between lax and tense vowels in bisyllables with one intervocalic consonant. Additional support for this view was found in Cebrian (1999), who found that Catalan learners failed to syllabify middle consonants with preceding lax vowels whereas native English speakers showed the expected tendency for coda syllabification. As a result, some of the Catalans’ productions of /i/ were produced like the tense counterpart /i/ when followed by a consonant-vowel sequence. However, those findings were based on limited data and rested upon the assumptions that learners know the lax vowel constraint and that L1 syllabification is difficult to overcome. The current study will test a larger number of subjects, and will more carefully assess knowledge and productivity of the phonotactic and syllabic constraints, as well as their interaction with vowel acquisition. A working hypothesis of this study is that the acquisition of a second language phonology involves the acquisition of both segmental and non-segmental structure, including syllable structure and phonotactics, and that these will interact such that the acquisition of non-segmental structure (or the lack of it) may have an effect on segmental acquisition.

1.7. Overview of the thesis

This thesis will attempt to answer the questions outlined above by examining Catalan speakers’ acquisition of two L2 vowel pairs that illustrate the lax-tense vowel contrast in English: /i/-/i/ and /e/-/e/. The relevant phonetic and phonological characteristics of English and Catalan segmental and non-segmental structure are described in Chapter 2. Chapter 3 assesses the phonetic similarity between the English sounds and the closest Catalan sounds by means of a perceptual assimilation task in
which native Catalan speakers identify English vowels in terms of Catalan vowel categories and provide goodness ratings. In addition, the reverse technique is also used, namely testing the perceptual assimilation of Catalan vowels to English with a group of native English speakers. The idea is that English perception of Catalan vowels in terms of English vowels will help determine the degree of identity between L1 and L2 vowels by illustrating which L1 vowels can go unnoticed if substituted for their L2 vowel counterpart. The predictions resulting from the perceptual assimilation tasks are then evaluated in subsequent experiments addressing the perception (Chapter 4) and production (Chapter 5) of the vowels. Vowel identification tasks are used to test the ability to perceive the L2 vowels and the relative weighting of acoustic cues, duration and vowel quality. Acoustic cue reliance is measured by means of an identification task involving a synthetic continuum in which stimuli vary in spectral and temporal properties. Two groups of L2 learners participated in the perception experiments: a group of Catalan immigrants in Canada and a group of Catalan undergraduate students in English Philology at a Catalan university. By testing these two groups, the effect of differences in experience and L2 input on L2 vowel perception and categorization will also be examined. The production of the English vowels by the same group of undergraduate students will be analyzed by means of acoustic measurements as well as intelligibility and goodness rating tests (Chapter 5). After assessing vowel perception and production, acquisition of L2 non-segmental structure will be evaluated in Chapter 6. The same group of Catalan undergraduate students was tested in addition to a group of native English speakers. Knowledge of the Lax Vowel Constraint will be examined by means of singular-plural picture labelling tasks which involve matching nonsense words with
pictures according to whether the test words are best suited for a singular meaning or a plural meaning. (e.g., /tez/, as opposed to /teIz/). Syllabification strategies are tested using syllable manipulation tasks involving bisyllabic words with singleton intervocalic consonants. Finally, a summary of the main conclusions of the thesis including a global analysis of the results for the different experiments will be presented in the last chapter, Chapter 7.
Chapter Two. Phonological characteristics of English and Catalan

2.0. Introduction

This chapter describes the characteristics of the L1 and the L2 focusing on the features that are relevant to the current study. The English vowel system is described, concentrating on the distinction between the so-called tense and lax vowels in English, a distinction based on both spectral and durational differences. A crucial property of lax vowels is their phonotactic restriction to closed syllables, the Lax Vowel Constraint (LVC). This chapter discusses the validity of the LVC as an active constraint in English in light of empirical evidence from psycholinguistic studies. One source of support for the LVC is the syllabification of medial intervocalic consonants in English. The strategies of syllabification in English are thus presented. This is followed by a description of the Catalan vowel system as well as the basic Catalan syllable structure. The chapter finishes with an acoustic comparison of the two languages.

2.1. The target language. English

2.1.1. Tense and lax vowels in English.

English has a larger vowel inventory than Catalan, including more distinctions in the central and low vowel space and a distinction between lax and tense vowels. The target feature in this second language acquisition study is the lax/tense vowel distinction.
The division of vowels into tense and lax in British English and American English is illustrated in (1), based on Rogers (2000).

(1) General American British English (RP)

Lax vowels
---
\(i \varepsilon \æ u \a\)  \(i \varepsilon \æ u \a\)

Tense vowels
---
\(i \varepsilon \æ u \varepsilon \u\)  \(i \varepsilon \æ u \varepsilon \u\)

The terms lax and tense are commonly used to refer to a vowel opposition based on acoustic characteristics and phonotactics, but they do not imply a difference in muscular activity (Rogers, 2000). In terms of acoustic differences, the tense/lax opposition is associated with variations in height and backness and a difference in vowel length (Ladefoged and Maddieson 1996). The tense vowels are longer than the lax vowels in all comparable contexts, thus they are also known as long and short vowels. Further, tense vowels are more peripheral in the vowel space than their lax counterparts. For instance, /i/ is longer, higher and more front than /u/.

In addition to the phonetic difference, a crucial difference between lax and tense vowels lies in their phonotactics, which is perhaps the main motivation for the phonological distinction: (stressed) lax vowels only appear in closed syllables; they cannot appear in open syllables. Importantly, this restriction does not affect tense vowels. Some examples of tense/lax vowel oppositions are given in (2), which also illustrates the

---

1 The English high and mid tense vowels /i/ and /e/ may also be represented as /iy/ and /ey/ in figures and graphs.
syllabic restriction on lax vowels (the Lax Vowel Constraint, discussed in more detail in the next section).²

<table>
<thead>
<tr>
<th></th>
<th>Tense/Long</th>
<th>Examples</th>
<th>Lax/Short</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>High front</td>
<td>/i/</td>
<td>beat</td>
<td>/u/</td>
<td>bit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bee</td>
<td></td>
<td>*b[ɪ]</td>
</tr>
<tr>
<td>Mid front</td>
<td>/ɛ:/</td>
<td>bait</td>
<td>/ɛ/</td>
<td>bet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bay</td>
<td></td>
<td>*b[ɛ]</td>
</tr>
<tr>
<td>High back</td>
<td>/ʊ/</td>
<td>boot</td>
<td>/u/</td>
<td>put</td>
</tr>
<tr>
<td></td>
<td></td>
<td>blue</td>
<td></td>
<td>*b[ʊ]</td>
</tr>
</tbody>
</table>

In addition to the vowel pairs in (2), Giegerich (1992) lists other vowel pairs in similar oppositions, namely, /æ/-/æ/, /oʊ/-/ʌ/ and /ɔ/-/ɒ/ (the latter in British English, not in General American). This set of vowel pairs illustrates the system of vowel oppositions that underlie the English vowel system. The oppositions involve long/short or tense/lax contrasts between vowels that are similar in quality and differ in syllabic restrictions. Although the validity and comparability of each vowel opposition is debatable and its status as such may vary from one English variety to another (e.g., /æ/ in American English varieties), the crucial point in such an analysis is that vowels are organized into a system of oppositions and that L2 learners have to acquire this system in shaping their L2 phonology. The opposition that is of interest to this study is the lax-tense distinction in high and mid front vowels specifically, which holds across most dialects of English. The

² In fact, the low front lax vowel can be pronounced word finally in the onomatopoeic word ‘baa’, the bleating cry of the sheep. Nevertheless, the lax/tense distinction seems to hold more strongly of high and mid lax vowels than low vowels, especially given the variability in the low vowel inventory across English varieties. The issue of the violability of the Lax Vowel Constraint is discussed in the following section.
following sections provide a more detailed discussion on the acoustic and phonotactic differences between these English vowels.

2.1.2. Spectral and temporal differences

This thesis focuses on the acquisition of the English high and mid front vowels, namely, /i e ’ e/, the vowels in ‘beat’, ‘bit’, ‘bait’ and ‘bet’, respectively, by Catalan speakers. The spectral characteristics of these English vowels are given in Tables 2.1 and 2.2 below, which provide the F1 and F2 values for male and female speakers’ high and mid front vowels. Given that the two dialects have been found to differ in the use of the two kinds of acoustic cues (vowel quality and vowel duration), as discussed below, the values for both American English and British English are given. In all three studies vowels were elicited in /h/+vowel+/d/ words in citation form (i.e., heed, hid, hayed, and head). Values for the vowel /e/ are not reported in two of the studies, which analyze only monophthongal vowels with longer steady states. Indeed, together with the height and length difference, the mid vowel contrast is also distinguished by greater transitions from the vowel’s steady state to the high off-glide in /e/, at least in some environments and in some English dialects.

Apart from the exclusion of /e/, there are notable differences between the data in Peterson and Barney (1952) and those in Hillenbrand et al. (1995) despite the fact that the latter study was an attempt to replicate former. Hillenbrand et al. suggest that the discrepancy may result from a dialect difference, since Peterson and Barney do not provide much detail about their speakers, whereas the speakers in Hillenbrand et al. were a dialectally homogeneous group of General American speakers, mostly from Michigan’s
lower peninsula and other areas of the upper midwestern United States. Another explanation can be found in diachronic change, given that the data in the two studies were collected several decades apart. In any case, listening tests in both studies showed very high rates of correct vowel identification. Thus, despite the discrepancy in formant frequencies, listeners had no problem identifying all vowels in each study. A comparison of the results of Deterding’s study of British English (Standard Southern British English, similar to RP) to the two studies of American English shows that both dialects exhibit similar spectral differences among the vowels under study.

<table>
<thead>
<tr>
<th>British English</th>
<th>American English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1^a</td>
</tr>
<tr>
<td>/i/</td>
<td>270</td>
</tr>
<tr>
<td>/ɪ/</td>
<td>390</td>
</tr>
<tr>
<td>/eɪ/</td>
<td>-</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>530</td>
</tr>
</tbody>
</table>

Table 2.1. F1 and F2 values of English high and mid front vowels for male speakers of British English (Deterding, 1997) and American English (a = Peterson and Barney (1952), b = Hillenbrand et al. (1995)).

<table>
<thead>
<tr>
<th>British English</th>
<th>American English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1^a</td>
</tr>
<tr>
<td>/i/</td>
<td>310</td>
</tr>
<tr>
<td>/ɪ/</td>
<td>430</td>
</tr>
<tr>
<td>/eɪ/</td>
<td>-</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>610</td>
</tr>
</tbody>
</table>

Table 2.2. F1 and F2 values of English high and mid front vowels for female speakers of British English vowels (Deterding, 1997) and American English vowels (a = Peterson and Barney (1952), b = Hillenbrand et al. (1995)).
The temporal contrast between tense and lax vowels is illustrated in Table 2.3, which provides the duration ratio between the two vowels as reported in a number of studies. The tense vowel is on the whole between 20% and 90% longer than the lax vowel. The percentage varies with regard to the following sound and also to the variety of English involved. The ratios show means across different contexts, hence the range of variation in some studies. Although British English tends to have a greater durational difference, length is still important in American English and the difference in duration between lax and tense vowels is still significant (e.g., Munro, 1993). Ratios for /ɛ/-/ɛ/ are also provided when available.

<table>
<thead>
<tr>
<th>/i/-/ɪ/</th>
<th>/ɛ/-/ɛ/</th>
<th>English variety</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.57 – 1.94</td>
<td>-</td>
<td>British English</td>
<td>Giegerich (1992)</td>
</tr>
<tr>
<td>1.50 – 1.85</td>
<td>-</td>
<td>British English</td>
<td>Lindsey (1990)</td>
</tr>
<tr>
<td>1.30 – 1.35</td>
<td>-</td>
<td>American English</td>
<td>Lindsey (1990)</td>
</tr>
<tr>
<td>1.60</td>
<td>1.43</td>
<td>American English</td>
<td>Flege, Munro &amp; Fox (1994)</td>
</tr>
<tr>
<td>1.20</td>
<td>1.20</td>
<td>American English</td>
<td>Munro (1993)</td>
</tr>
<tr>
<td>1.28</td>
<td>1.34</td>
<td>American English</td>
<td>Hillenbrand et al. (1995)</td>
</tr>
</tbody>
</table>

Table 2.3. Duration ratio /i/-/ɪ/ as reported in different studies.

Despite variability in the duration ratio due to contextual and dialectal differences, the studies show a clear difference in duration between lax and tense vowels. Although the contrast between lax and tense vowels is described in terms of both temporal and spectral differences, there is evidence that native speakers, at least of some English varieties, rely mostly on spectral cues in discriminating the two vowel types, especially in
the case of high vowels. Hillenbrand, Clark and Houde (2000) conducted a study to
determine the role of duration in vowel recognition and found that the lax-tense vowel
pairs in high vowels (/i/-/ɪ/ and /u/-/ʊ/) are in fact minimally affected by duration,
supporting a stronger relative weighting of spectral over temporal cues. Native English
speakers’ reliance on spectral cues was also reported in a number of works evaluating the
relative weighting of cues in native and L2 English, as discussed in Chapter 1. These
studies showed that temporal differences seem to be more salient for L2 learners. This
issue will be investigated further in this study by examining Catalan learners of English.

The other issue under study in the current work is the effect of the syllabic context
in the implementation of the contrast and its relation to the phonotactic restriction on lax
vowels. We turn now to a more detailed analysis of the Lax Vowel Constraint in English.

2.1.3. The Lax Vowel Constraint

The LVC is a phonotactic constraint that disallows lax vowels in open stressed
syllables. Before addressing its influence on the acquisition of a second phonology it is
important to determine first the status of phonotactics in the phonology. One important
question is whether phonotactic constraints constitute a distinct part of a speaker’s
phonological knowledge or if the patterns produced by speakers simply reflect the
segmental sequences that are stored in their lexicon. If it is a genuine body of linguistic
knowledge, then we have to ask how it is acquired, in particular in the case of the LVC,
both by L1 and L2 learners. The cognitive status of phonotactic knowledge in relation to
other parts of the grammar and its contribution to the shaping of the phonology are
discussed in the next section.
2.1.3.1. The status of phonotactics in the phonology

Studies on speech segmentation provide empirical evidence for the status of phonotactic constraints. The segmentation process appears to be highly sensitive to the typicality or wordlikeness of sound sequences (Bailey and Hahn, 2001). Bailey and Hahn studied whether this effect of typicality was the result of language phonotactics or similarity to existing words in the lexicon (lexical influences). The former would imply a separate component for phonotactic knowledge whereas the latter would emphasize lexical neighbourhoods as the mechanism at work in facilitating speech segmentation. On the basis of results from wordlikeness judgement experiments, the authors conclude that phonotactic and lexical influences are distinct, indicating that phonotactic knowledge is not merely implicit in the lexicon but constitutes a component of its own. These results are in accordance with Vitevich and Luce (1998), who also concluded that lexical influences and phonotactics are two separate bodies of knowledge. More evidence comes from studies of on-line word recognition with adults, indicating that phonotactic constraints are used in segmenting speech (McQueen, 1998; Seguí, Frauenfelder and Halle, 2001).

With respect to the role of phonotactics in acquisition, there is growing evidence that infants become sensitive to the phonotactics of their native language sometime during the second half of their first year. Children learning English are able to segment speech into words initially on the basis of stress patterns (Jusczyk, 1999). Thus, 7.5-month-old children display adult-like abilities in speech segmentation only with words that conform to the predominant stress patterns in English words. Three months later,
children make use of other information such as allophonic cues and phonotactic patterns in order to establish word boundaries. In a study that focused on consonant cluster sequences, Mattys and Jusczyk (2001) found that 9-month-old infants used probabilistic phonotactics to segment speech into words. Knowledge of phonotactics is thus incorporated into the children’s grammar early on in the acquisition process and this knowledge is actively used in the process of word recognition.

Recent work has demonstrated that adult speakers are able to acquire new phonotactic restrictions. Onishi, Chambers and Fisher (2002) tested the ability to extract a probabilistic phonotactic constraint while listening to a list of English-like words with a novel restriction on the distribution of consonants with respect to syllable position and adjacent vowels. They found that knowledge of the constraint was reflected in the latency of subjects' cued productions after the listening experience.

There is also evidence that phonotactic constraints influence the perception and production of segmental structure. Empirical studies have revealed significant differences in perceptual difficulty of non-native sound contrasts as a function of the phonotactic and phonetic contexts in which the contrasts are presented (Strange, 1999). Thus, identification of phonological segments may be influenced and overridden by phonotactic constraints. Solé (1989) argues that phonotactic information acts as a filter for the acoustic signal and the subsequent articulatory gestures. Evidence comes in fact from the effect of the English Lax Vowel Constraint on French borrowings into English and from English accent in French and Spanish, and from a perception experiment in which the same segmental structure was presented in different syllabic structures to speakers of different languages. Results showed that perception varied according to the phonotactic
characteristics of the listeners’ languages. With respect to the treatment of foreign words, French borrowings into English (where the original French words ended in a front mid vowel, close to English lax vowel /e/, such as ‘ballet’ or ‘gourmet’) are often pronounced with the tense vowel /e/ in English. Similarly, tensing final lax vowels is a common trait of English-accented French. Tensing affects the vowel /e/ in open syllables but not in closed syllables, as illustrated by the native English speakers’ pronunciation of the French example elle est très belle [ɛl e tre bel] (‘she is very beautiful’) as [ɛl e tre bel]. Similarly, the substitution of English mid tense vowels for Spanish monophthongal mid vowels is also typical of English-accented Spanish, as in [no se] for no sé ([no se]), ‘I don’t know’ (Solé, 1989).

Further empirical support comes from Rochet and Putnam Rochet (1999), who found that the phonotactic constraint that governs the production of the high and mid front unrounded vowels in French also governs their perception. French has two front mid unrounded vowels (/e/ and /ɛ/) that are contrastive in open syllables (e.g., fée /fɛ/ ‘fairy’ and fait /fɛ/, ‘done’). However, the contrast is neutralized in closed syllables. Rochet and Putnam Rochet investigated if the syllable-based neutralization constrained perception using an identification task involving a synthetic continuum from vowel /i/ to /e/ to /ɛ/. The vowels in the continuum were presented in either CV or CVC frames. The results showed that listeners categorized the stimuli into three vowels (/i/, /e/ and /ɛ/) when presented in open syllables, but only two vowel categories emerged with CVC stimuli (/i/ and /ɛ/). In fact, the vowel category space occupied by /ɛ/ in open syllables
was mostly occupied by /i/ in CVC context, relating /i/ and /e/ rather than the traditional /e/-/e/ neutralization account. The results of this study clearly indicate an effect of phonotactics on both the production and perception of segmental structure. In addition, as discussed in the previous chapter, Rochet and Putnam Rochet found that this phonotactic constraint affects French speakers’ L2 English: it accounts for the French learners' substitution patterns in the case of the final vowels in happy and presented (Rochet and Putnam Rochet consider these to be realizations of an unstressed high front vowel /ɪ/). This L2 vowel is perceived and produced as [e] in an open syllable (e.g., happy), and as /i/ in a closed syllable (e.g., presented).

To conclude, evidence from a range of studies indicates that phonotactic constraints are a distinct component of phonological knowledge. All of the findings reported above would be unexpected if the distributional patterns found in the wordforms of a given language had not been somehow extracted and internally represented by the language speaker. Moreover, the studies on infants, children, and even adults illustrate a predisposition to extract these patterns from speech. It seems reasonable to assume that phonotactics are learned concomitantly with the other components of the speaker’s L1 grammar and thus it is possible that they are acquired in the course of the acquisition of a second language. Let us consider now the status and productivity of the English Lax Vowel Constraint more specifically.

2.1.3.2. Productivity and violability of the Lax Vowel Constraint

Arguments for the productivity of the LVC in English go back to Sapir (1933), who listed three kinds of possible syllable endings in English, illustrated in (3) below: a
tense vowel or diphthong, a tense vowel or a diphthong followed by one or more consonants, and a lax vowel followed by one or more consonants. A (stressed) syllable ending in a lax vowel is not a possible alternative. Sapir observed that his students (native English speakers) tended to hear and transcribe word-final lax vowels in non-English words with a following glottal stop. This was not the case with word-final tense vowels. In addition, they produced the final lax vowels with diphthongization (vowel ‘drawling’). This would indicate that the LVC is an active phonotactic constraint in English.

(3) Possible word endings

| /bi:/ | /baj/ |
| /bjt/ | /bet/ |
| /bi:t/ | /bajt/ |

Impossible endings

| */bi/ | */be/ |

However, the LVC is not as inviolable a constraint as some others, including as the distribution of aspirated stops in English, where syllable initial pretonic voiceless stops are aspirated without exception. In repetition tasks, monolingual native English speakers find difficulty in repeating unaspirated initial stops (rendering them either voiced stops or aspirated stops) but they do not have an equal difficulty in repeating final lax vowels. The results from a nonsense word repetition task involving eight speakers in Westby (1984) showed an overall greater difficulty in repeating the words ending in a lax vowel than those ending in a tense vowel. Final tense vowels were correctly repeated 88-100% of the time. The percentages for lax vowels were: 50% for /æ/ and /u/, 87.5% for /t/ and /ʌ/ and 100% for /ε/. The distributional restriction on lax vowels is thus not all-or-nothing. Nonetheless, this task may involve repetition of sound-based information stored
in echoic or short-term memory and may not tap into the subjects’ phonological system. In addition, correct repetition of unaspirated stops involves altering physiological or articulatory conditions such as voice onset time, something which is not the case with the repetition of lax vowels.

Other sources of evidence for the LVC in English include borrowings, English accent in foreign languages, and dictionary syllable division conventions (Westby, 1984; Solé, 1989). As discussed in the previous section, the LVC is responsible for the tensing of final lax vowels in French borrowings into English. Similarly, tensing final lax vowels is a common trait of English-accented French. Dictionaries often indicate syllable breaks according to this constraint so that an intervocalic consonant following a tense vowel is syllabified with the following vowel whereas a consonant preceded by a lax vowel is grouped with the preceding vowel (de-mon vs. lem-on). The actual relation between the LVC and syllabification of medial consonants is discussed in more detail below. Finally, when we find a sequence of two vowels the first vowel is in an open syllable and, accordingly, is tense rather than lax (e.g., pious, variety, theatre). These examples illustrate that the LVC is valid for word-internal as well as word-final open syllables. More evidence of this will be presented from syllabification of medial consonants.

An experimental methodology designed to test subjects’ knowledge of the LVC involves the application of the rule of plural formation and the use of picture labelling tasks. This kind of technique has been used to test children’s knowledge of the plural formation rule and its application with respect to surface phonetic constraints (Anisfield and Tucker, 1967; Berko, 1958; Messer, 1967). In these tasks, children had to learn a number of non-sense words and derive their plural form (or singular form) with the aid of
visual prompts. Westby (1984) and Wallner (1986) applied this method to the study of native speakers’ knowledge of the LVC. The studies consisted of a series of picture labelling tasks in which subjects had to match a given nonsense word with a picture. Words varied in shape so that some ended in a lax vowel plus /z/ (e.g. /driz/) and others ended in a tense vowel plus /z/ (e.g., /skeiz/). Pictures showed either one item or two equal items so that a plural word would be preferable to designate a two-item picture and a singular word would be best to refer to a single-item picture. The tasks thus are intended to test subjects’ recognition of allowable and unallowable singular forms based on their knowledge and productivity of the LVC. The tasks varied in the number of pictures and words per trial so that they involved choosing one of two words for one picture, one of two pictures for one word, or matching two words with two pictures. Support for the LVC comes from choosing words ending in a tense vowel plus /z/ as a possible plural form and words ending in a lax vowel plus /z/ as a singular (given that choosing them as a plural form would imply a violation of the LVC in their singular form).

The results in Wallner (1986) show high productivity of the LVC. Wallner tested 25 subjects and found that they used the LVC in a variety of picture labelling tasks (1 word 2 pictures, 2 words 1 picture, 2 words 2 pictures). She also tested children aged between four and eleven and found that the age of acquisition of the LVC was around six. In the case of Westby (1984) the results showed no effect of the LVC except in the forced double matching technique (2 words 2 pictures). Westby concludes that the constraint’s productivity is high with borrowings only. However, the lack of a stronger effect in Westby’s experiments may be due to the small number of subjects tested, 8 or 4
depending on the task. Wallner (1986) in fact conducted a version of Westby’s study with a larger number of subjects and, as we have seen, found an effect of the LVC in all tasks.

Evidence from borrowings, English accent in foreign languages, dictionary syllable division conventions, and psycholinguistic experiments converge on the conclusion that the LVC is a productive constraint. Further evidence comes from English syllabification, specifically from the syllabification of intervocalic consonants. This is discussed next.

2.1.4. The syllabification of intervocalic consonants

How syllable boundaries are established in English has been a central focus of study. Some theories use syllable weight requirements as one argument for ambisyllabicity or coda syllabification of medial consonants. Thus, in (C)V(C)V sequences, the intervocalic consonant is often analyzed as part of the first syllable, either as ambisyllabic (Kahn, 1976; Kreidler, 1995; Pulgram 1970; Gussenhoven and Jacob, 1998), or as a coda (Selkirk, 1982).

The LVC has been described as a constraint on lax vowels in open syllables. This is clear when we consider word-final syllables. But lax vowels are also found in non-final syllables, such as in the words *civil, lemon, cover* or *sugar*. Onset maximization principles would predict that the intervocalic consonant would syllabify with the following vowel. However, that syllabification would leave the lax vowel in an open syllable. An ambisyllabicity analysis of medial consonants best satisfies both requirements. This is the view advanced by a number of approaches (e.g. Pulgram, 1970; Giegerich, 1992; Hammond, 1997, or the theory embodied in Webster's dictionary).
These theories relate ambisyllabicity to the need to close syllables that include lax vowels, that is, a medial consonant is syllabified with the preceding lax vowel so that the lax vowel is found in a closed syllable. However, ambisyllabicity is not needed when the vowel is tense. Similarly, under a moraic account, the LVC can be defined as a bimoraic requirement fulfilled not by lengthening the vowel but by attracting a following consonant (Hammond, 1997).

Giegerich (1992) also relates the allophonic variation found with voiceless stops with the ambisyllabicity of medial consonants after lax vowels, claiming that medial consonants following a stressed lax vowel pattern as both onset and coda stops. For instance, the /t/ in *petrol* can be glottalized (like a coda), and be followed by a devoiced approximant (like an onset /t/), but the /p/ in *apron* is followed by a devoiced approximant and is not glottalized, thus patterning like an onset only. Trammell (1993) relates ambisyllabicity to a number of allophonic variations involving English consonants and vowels (including stop aspiration, vowel nasalization, vowel length, consonant length, flapping) and advocates the teaching of ambisyllabicity as a helpful tool in the acquisition of L2 English phonology. Other studies show the relation between ambisyllabic or coda syllabification of medial consonants and the phonetic realization of the consonant, e.g. Boucher (1988) on duration differences between ambisyllabic and non-ambisyllabic consonants, Davis and Van Summers (1989) on aspiration of intervocalic stops or Maddieson (1985) on allophones of the lateral in English. However, these studies do not link the ambisyllabicity analysis to the lax-tense vowel contrast.

There is thus no consensus in the account of ambisyllabicity and the ways it is phonetically manifested. A detailed examination of the precise nature of potentially
ambisyllabic consonants and the actual phonetics of it lies beyond the scope of this study. The next section discusses psycholinguistic experimental research that investigates the relationship between the Lax Vowel Constraint and syllabification of intervocalic consonants. The results support a close link between the LVC and coda syllabification of medial consonants.

2.1.4.1. Psycholinguistic evidence for CVC-VC syllabification

One source of evidence for the validity and productivity of the LVC comes from experiments involving syllable manipulation tasks such as syllable repetition, syllable reverse order, and break insertion tasks. A number of studies show a preference to syllabify a medial consonant with a preceding lax vowel.

Westby (1984) tested syllabification strategies by means of a syllable repetition task and a syllable reverse order task. Her results were inconclusive in that there was no clear difference in syllabification between words containing a lax vowel and words containing a tense vowel. The lack of a clear effect of the LVC may be related to the procedure and objective of the task. Westby was interested in examining how subjects dealt with lax vowels in open syllables and thus responses involving coda syllabification of the medial consonants were discouraged.\(^3\) In addition, the experiments were run on a small number of subjects (four or eight). Thus, it is not possible to determine the effect of the LVC on syllabification from her findings.

Other psycholinguistic experiments show in fact that short/lax vowels tend to attract single medial consonants in syllable division tasks involving English. Fallows

---

\(^3\) For instance, in the syllable reverse order task the experimenter provided the first part of the response including the medial consonant (e.g., river \(\rightarrow\) ver + ...). In addition, subjects were instructed to avoid copying the medial consonant (e.g., river \(\rightarrow\) ver + riv).
(1981) ran a study on thirteen native American-English speaking children belonging to two groups, 4-5 year-olds and 9-10 year-olds. The children performed syllable division tasks involving the repetition of the first and the last syllable and the insertion of a pause in the middle of the word. She found that subjects syllabified medial consonants with preceding stressed lax vowels 85% of the time. The study concluded that the knowledge of phonotactic constraints, including the need to close off lax vowels in stressed syllables, is acquired around the age of five. This is consistent with the results from singular-plural picture labelling tasks in Wallner (1986), who tested children aged between four and eleven and found that children acquired the LVC by the age of six.

Treiman and Danis (1988) also looked at medial consonants in English. They performed a series of experiments to assess the status of single consonants in intervocalic position using both an oral string-inversion task and a written slash-insertion task. Like Fallows, they found the type of the preceding vowel and the presence or absence of stress to be relevant factors at play. A short (i.e., lax) vowel tended to attract consonants to its syllable, as did a stressed vowel. Treiman and Zukowski (1990) used written and oral syllabification tasks and analyzed spelling errors to study syllabification of medial consonant clusters in English. Their research shows that while onset maximization principles (e.g., the Obligatory Onset Principle (OOP)) were in general confirmed, other factors were also found to play a significant role, confirming the findings in Treiman and Danis (1988) and in Treiman (1986).

---

4. In the syllable inversion tasks subjects had to invert the order of the syllables of a word presented orally (e.g. melon --> onmel? lonme? lonmel?). There was no explicit reference to syllable structure in the instructions, but examples were given in which syllable and morpheme boundaries coincided. Responses were oral. The second task involved choosing the preferred breakpoint between two possible candidates (e.g. me/lon vs. mel/on). For each word subjects were asked to circle the better syllabification of the two choices provided according to how they would syllabify it in print.
Derwing (1992) extends Treiman and Danis’ study to other languages in order to test the potential universal relevance of the OOP. The languages selected for this study were: English, colloquial Cairo Arabic, Blackfoot (an Algonquian language spoken in southern Alberta), standard Seoul Korean and Swiss German. Derwing uses a new task, inspired by Treiman and Danis's tasks, but more suited to the kinds of languages under investigation: the pause-break task. In this task subjects were asked to choose which of two or three alternative divisions of a word sounded the most natural, e.g. English *melon*:

(a). me (pause) lon, (b). mel (pause) on, (c). mel (pause) lon. His study supports the Obligatory Onset Principle in most languages but finds some complications with English. The English data show an effect of the preceding vowel along the lines of what Treiman and Danis observed. English subjects show a preference for the medial consonants in onset position of the second syllable, thus supporting the OOP, except for words with a lax vowel preceding the medial consonant. In those cases, there is a tendency for a consonant to be linked with a lax stressed vowel as its coda, varying directly as a function of its reputed relative position on the sonority hierarchy. The greater the sonority of the consonant, the stronger the tendency to syllabify with the preceding vowel. This finding is corroborated in Derwing and Nearey’s (1991) study on vowel-consonant cohesion (or the ‘vowel stickiness’ phenomenon).

The findings of these psycholinguistic studies reveal the position of stress, the quality of the preceding vowel, the type of medial consonant and the spelling of the consonant in question as factors in placing syllable boundaries. These findings can shed some light on the conflicting views about syllabification in English and the placement of
syllable boundaries. The effect of the type of vowel supports the validity of the LVC as an active phonotactic constraint.

Finally, the results of the studies reviewed in this section indicate that the LVC is active not just as a constraint on word final syllables but on non-final syllables as well. Recall that word-internal vowel sequences already indicate a syllable-final rather than a word-final scope of the LVC (e.g., tense vowels in prevocalic open syllables such as in *pious* and *theatre*). Some tasks such as the reverse order task (e.g., *melon* → *lon-me* / *lon-mel* / *on-mel*) may be said to simply move the lax vowel to word final position and thus test the validity of the constraint in word-final position. However, the tasks that involve repetition of the first syllable (Fallows, 1981) and syllable-division choice (e.g., choosing between *me-lon*, *mel-on* and *mel-lon*) such as by Fallows’ (1981) pause insertion task, Treiman and Danis’ (1988) written slash insertion task and Derwing’s (1992) oral pause break task, tap more clearly into the existence of the constraint word internally.

This thesis will provide more empirical evidence for the validity of the Lax Vowel constraint, understood as a syllabically based phonotactic constraint, by means of two sets of experiments involving singular-plural picture labelling tasks and syllable division tasks. These will be conducted with native English speakers to shed more light on the productivity of the LVC and its relation to syllabification of medial consonants. In addition, the acquisition of the constraint by L2 English learners will be evaluated by testing Catalan learners of English. The phonological characteristics of the L1, Catalan, are described in the following section.
2.2. The L1: Catalan

2.2.1. The Catalan vowel system

This study is concerned with the dialect of Catalan spoken in the central and eastern part of Catalonia, the region around Catalonia’s capital city, Barcelona. Central Catalan has an eight vowel inventory with four degrees of height and no temporal contrast. The Catalan vowels are presented in Figure 2.1, which illustrates the seven full vowels plus [ə], a reduced vowel in unstressed position in most dialects including Central Catalan. The vowel reduction process neutralizes the mid front and low vowels (/e/, /ɛ/, /a/) to the central vowel [ə] and the high and mid back vowels (/ɔ/, /o/, /u/) to [u]. Thus, the seven vowel inventory is reduced to a three vowel inventory in unstressed position (/i/, /u/, /ə/). In addition to the monophthongal vowels, Catalan has a number of diphthongs involving combinations of the eight vowels and the semivowels /j/ and /w/ as rising sequences (e.g., /ja/, /wa/) or falling sequences (e.g., /aj/, /aw/) (Recasens, 1993).

The seven vowels in stressed position thus are not organized in a system of contrasts like the lax-tense contrast described for English. Vowels differ in height and backness. In other words, vowel contrasts are based on vowel quality alone; there are no
temporal contrasts. In terms of duration, low vowels tend to be longer than high vowels so that /e/ is longer than /e/ which is longer than /i/. Vowel formants for the Catalan vowels relevant in this study are given in Table 2.4; mean durations are given in Table 2.5. The values represent the means for six male speakers’ production of these vowels in /s+vowel+k/ syllables and three male speakers’ production in symmetrical CVC syllables. They were speakers of the central Catalan variety from the region of Tarragona. Their values for the high and mid front vowels are taken as representative of the Central Eastern Catalan dialect that constitutes the L1 in this study. Any difference between this and other Eastern Catalan varieties (e.g. the region of Barcelona) would only affect back vowels.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>/i/</td>
<td>276 (29)</td>
<td>2156 (156)</td>
</tr>
<tr>
<td>/e/</td>
<td>397 (49)</td>
<td>1982 (199)</td>
</tr>
<tr>
<td>/e/</td>
<td>544 (58)</td>
<td>1811 (257)</td>
</tr>
</tbody>
</table>

Table 2.4. Formant values of Catalan high and mid front vowels (Notes: a. Vowels in sVk sequences, 6 male speakers; b. Vowels in symmetrical CVC sequences, 3 male speakers. Standard deviations are given in brackets).

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/e/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>78.6</td>
<td>86.2</td>
<td>102.9</td>
</tr>
<tr>
<td>S.D.</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 2.5. Mean durations in milliseconds for Catalan high and mid front vowels (Recasens, 1984)
As the vowel formant structure and the mean duration differences illustrate, vowel quality is a more salient feature in the distinction of Catalan vowels than vowel quantity. There is no parallel to the duration contrast between English vowel pairs such as English /i/ and /ɪ/. In fact, the higher vowel in the English pair (/i/) is longer than the lower one (/ɪ/), opposite to what we find in Catalan.

2.2.2. Catalan syllable structure

Catalan has a total of twelve possible syllable shapes (Recasens, 1993). These range from a single vowel (V, e.g., a ‘to’) to a maximum of two consonants in the onset and three in the coda (CCVCCC, e.g., flascs ‘flasks’). There is no restriction on coda consonants except for the approximant variants of the voiced stops, which are always stops in coda position. Consonant clusters tend to involve sequences of obstruent plus liquid in the onset (preu ‘price’, blat ‘wheat’, fred ‘cold’) and liquid plus nasal or obstruent, /s/ plus stop or any non-sibilant consonant plus /s/ in the coda (e.g., arc ‘arch’, salm ‘psalm’, fosc ‘dark’, dits ‘fingers’). With respect to syllabification of medial consonants, single intervocalic consonants always syllabify with a following vowel both word internally and across a word boundary within a phrase. Consonants in a cluster syllabify with the following vowel so long as the sequence is a possible onset cluster.

Thus, a variety of possible syllable structures is possible in Catalan and most consonant sounds are allowed in coda position. The crucial fact about Catalan syllabification and, as discussed below, the characteristic that distinguishes it from English syllabification is that intervocalic singleton consonants are always syllabified in
the onset together with the following vowel. This is true regardless of the preceding vowel. Furthermore, there is no syllabic restriction for any kind of vowel; all vowels can appear in open syllables and in closed syllables. Some examples of the distribution of Catalan vowels are given in (4) and of syllabification of medial consonants in (5). As the examples illustrate, syllabification of medial singleton consonants is straightforward and all vowels can be found in open and in closed syllables.

(4)  

<table>
<thead>
<tr>
<th>Example</th>
<th>Open syllables</th>
<th>Gloss</th>
<th>Example</th>
<th>Closed syllables</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>vi</td>
<td>[bi]</td>
<td>pine</td>
<td>nit</td>
<td>[nit]</td>
<td>night</td>
</tr>
<tr>
<td>vida</td>
<td>['biòa]</td>
<td>life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bè</td>
<td>[be]</td>
<td>well</td>
<td>nét</td>
<td>[net]</td>
<td>grandson</td>
</tr>
<tr>
<td>néta</td>
<td>['neta]</td>
<td>granddaughter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>be</td>
<td>[be]</td>
<td>lamb</td>
<td>net</td>
<td>[net]</td>
<td>neat, clean</td>
</tr>
<tr>
<td>neta</td>
<td>['neta]</td>
<td>clean (fem.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>va</td>
<td>[ba]</td>
<td>S/he goes</td>
<td>mag</td>
<td>[mak]</td>
<td>magician</td>
</tr>
<tr>
<td>casa</td>
<td>['kazɔ]</td>
<td>house</td>
<td>poc</td>
<td>[pɔk]</td>
<td>little</td>
</tr>
<tr>
<td>bo</td>
<td>[bɔ]</td>
<td>good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bona</td>
<td>['bɔna]</td>
<td>good (fem.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>por</td>
<td>[po]</td>
<td>fear</td>
<td>sóc</td>
<td>[sɔk]</td>
<td>I am</td>
</tr>
<tr>
<td>boca</td>
<td>['bɔkɔ]</td>
<td>mouth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nu</td>
<td>[nu]</td>
<td>naked</td>
<td>puig</td>
<td>[putʃ]</td>
<td>hill, mountain</td>
</tr>
<tr>
<td>pluja</td>
<td>['pluʒɔ]</td>
<td>rain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mare</td>
<td>['marɔ]</td>
<td>mother</td>
<td>préssec</td>
<td>['presɔk]</td>
<td>peach</td>
</tr>
</tbody>
</table>
2.3. Acoustic comparison of Catalan and English vowels

Having described the characteristics of the L1 and L2, this chapter finishes with an acoustic comparison of the English vowels /i/, /ɪ/, /ɛ/, /ɛ/ and the closest Catalan vowels (/i/, /ɛ/ and /ɛ/) on the F1-F2 dimension. Phonetic similarity will be assessed by means of a perceptual task in the next chapter. However, an acoustic comparison yields preliminary results. The F1 and F2 values are plotted in Figure 2.2, which is based on the frequencies presented in Tables 2.1 and 2.4 above. The values for the English vowels are represented by the symbols i, I, E, ey (for /i/, /ɪ/, /ɛ/ and /ɛ/ respectively) and the Catalan vowels are indicated with the letter c (e.g., ci, ce, cE (for /i/, /ɛ/ and /ɛ/).
Figure 2.2. Distribution of English and Catalan vowels in the F1-F2 vowel space according to the values presented in Tables 2.1 and 2.4.

The comparison indicates a strong parallelism in the formant structure of the high and mid vowels in both languages. Although a larger sample of Catalan data would be necessary to carry out a more reliable comparison, it is clear that the Catalan vowels /i/, /e/, and /e/ fall within the acoustic vowel space of the English vowels /i/, /u/ and /e/, respectively. Catalan /i/ and English /i/ cluster together, as do Catalan /e/ and English /u/. The overlap is even more evident in the case of Catalan and English /e/. The steady portion of the English vowel /e/ seems to be closer to the Catalan higher mid vowel /e/. This, together with the gliding nature of this English vowel, indicates that the vowel /e/
may more closely approximate the Catalan diphthong consisting of the vowel /e/ and the high front glide, i.e., the diphthong /ei/.

Thus, the comparison based solely on the acoustic description of the vowels indicates that all four L2 vowels may be fairly straightforwardly assimilated to existing L1 vowel categories. The perceptual reality of these acoustic similarities will be examined in the next chapter.

2.4. Conclusion

This chapter has presented the relevant characteristics of the L1 and the L2 in this study, i.e., Catalan and English, illustrating some differences with respect to the vowel properties and syllable structure. English has a lax-tense vowel contrast that is manifested in a difference in vowel quality and duration. In addition, tense and lax vowels also differ in distribution given that the lax vowels are restricted to closed syllables by the Lax Vowel Constraint. Although the LVC is not inviolable, converging evidence from the treatment of borrowings, English accent in L2 French, syllabification strategies and psycholinguistic experiments support its status as an active constraint in English.

Catalan differs from English in that there is no temporal contrast in the Catalan vowel inventory and no phonotactic restriction on any vowel type. Finally, whereas Catalan has clear syllable boundaries with onset syllabification of medial consonants, English has syllabification of post-tonic intervocalic consonants with a preceding lax vowel. Catalan syllabification always obeys maximal onset.

A preliminary contrastive analysis based on the acoustic values of the L1 and L2 vowels indicates that the two languages have comparable inventories with respect to the
high and mid front vowels. On the basis of this comparison the Catalan learners of
English are expected to draw from their native vowel inventory in their categorization of
the English vowels. The specific differences between English and Catalan make the
learning of L2 English by Catalan speakers an ideal testing ground to evaluate hypotheses
about the role of features that are not distinctive in the L1 as well as non-segmental
constraints. The effect that the differences in segmental and non-segmental structure
between the two languages have on the acquisition process will be examined in the
following chapters.
Chapter Three

Perceived similarity between Catalan and English vowels

3.0. Introduction

An acoustic comparison of the English and the Catalan high and mid front vowels illustrated that all four L2 vowels have an acoustically comparable match in the L1 (see Chapter 2). This chapter further examines the phonetic similarity using a perceptual assimilation task, following current trends in second language acquisition research. In this task, native Catalan speakers identify English vowels in terms of the Catalan vowels. To conduct a deeper investigation of the relation between the two vowels sets, English native speakers’ perception of Catalan vowels is also evaluated. The results of these tasks are discussed in light of the predictions outlined in Chapter 1.

3.1. Measuring phonetic similarity

There have been attempts to describe L1/L2 phonetic relationships using different units of analysis and levels of abstraction (Strange, 1999). For instance, phonetic similarity has been defined on the basis of phonetic symbols, that is, a similar sound is a sound that is represented by the same IPA symbol in two languages. But this approach is rather unreliable given crosslinguistic differences in transcription systems and uses (Rochet, 1995; Flege, 1992, 1991).

More commonly, similarity has been measured on the basis of acoustic comparisons in terms of first and second formant measurements plotted in a two-dimensional space (Flege, Munro and Fox, 1994; Flege, 1997b; Bohn and Flege, 1997,
among others). Thus, a new vowel phoneme can be understood as a vowel that occupies a portion of the acoustic space that is unoccupied by the realizations of any L1 vowel. Nevertheless, defining phonetic similarity in terms of acoustic measures (F1 and F2 values) alone is complicated by the variability produced by a number of factors including stress, speaking rate, formality and vocal tract size (Flege, Bohn and Jang, 1997). Further, speech sounds involve a variety of acoustic cues in addition to F1 and F2 values. Researchers would first need to assess the crucial cues for crosslinguistic equivalence comparisons and to find a precise, consistent, and reliable way of measuring them.

As discussed in Chapter 1, researchers have increasingly been employing perceptual assimilation tasks (PATs) to determine the degree of phonetic similarity. In these tasks, listeners with no L2 experience are presented with L2 speech stimuli, and asked to indicate to which L1 phonetic category (if any) each L2 token is most similar, and rate its "goodness" as an exemplar of that category. PATs allow a classification of the L2 sounds as identical, similar or new with respect to the L1 sound system (Strange, 1999; Flege, Bohn and Jang, 1997; Ingram and Park, 1997). Recall that the proposed classification is established as follows:

a) Identical (strongly assimilated): an L2 vowel that is consistently categorized as a good instance of one L1 vowel (that is, consistently assimilated to one L1 vowel) and obtains good ratings,

b) Similar (weakly assimilated): an L2 vowel that is consistently assimilated to one L1 vowel but with poor goodness ratings, or

c) New (non assimilated): an L2 sound that is identified as instances of multiple L1 sounds with poor goodness ratings.
Research has in fact evidenced a disparity between the results of perceptual similarity and acoustic similarity. Stevens et al. (1996) tested the correspondence between cross-language perceptual similarity and cross-language acoustic similarity. Acoustic measurements indicated that North German front rounded vowels are intermediate between American English front and back vowels. American English listeners judged German front vowels as "poor" matches to any English vowel category, as predicted given the acoustic difference, but listeners also categorized the L2 vowels as more similar to American English back vowels than to front vowels, which cannot be accounted for by their intermediate acoustic position between American English front and back vowels. Flege (1991) found that acoustic comparisons between Spanish and English vowels did not correspond to the perceptual comparisons. According to their acoustic data, English /i/ was expected to be identified with Spanish /i/, and English /u/ most likely as Spanish /e/. While the results for /i/ were as predicted (94%), Spanish monolinguals identified English /u/ more often with Spanish /i/ (68%) than with Spanish /e/ (19%), and English /e/ was assimilated to Spanish /e/ 81% of the time. Similarly, Bohn and Flege (1992) found that F1 and F2 analyses of L2 vowel production which fell within the natives speakers’ values were not necessarily found to reach corresponding high rates of intelligibility by native listeners (as in the case of English /æ/ produced by German speakers). Finally, Bohn, Strange and Trent (1999) conclude from a review of different studies that acoustic comparisons do not always predict assimilation patterns and advocate that direct assessments of assimilation patterns are better tools to predict perceptual difficulty.
The next sections assess the perceptual similarity between Catalan and English vowels. The results of the perceptual assimilation task will be contrasted with the results of the acoustic comparison in the previous chapter, allowing an evaluation of the two methods of measuring phonetic similarity.

3.2. Perceptual assimilation of English vowels to Catalan vowels

3.2.1. Stimulus preparation

English vowel stimuli were elicited from two male speakers of Canadian English. Each talker read a list containing English target vowels in monosyllabic words of the form /h/ + vowel (e.g., hee (/hi/)), in the case of the tense vowels and /h/ + vowel + /b/ in the case of the lax vowels (e.g., hib (/hɪb/), heb (hɛb/)). The lax vowels had to be elicited in a CVC sequence given the restriction on lax vowels in open stressed syllables. The consonant environment was chosen in order to minimize C to V and V to C tongue coarticulation (Strange et al., 1998), facilitating the extraction of the vowel in order to prepare the vowel stimuli for the experiment. Five tokens of each word were elicited.

Data were recorded using a SONY digital-audio tape recorder. The words were subsequently digitized at a 10kHz sampling rate with 16-bit resolution and normalized for peak intensity. The vowel portions were then edited out from each test word so as to minimize the effect of consonant properties that might create an impression of a foreign accent. As pointed out in earlier studies, results from accent rating tasks may be affected by the mispronunciation of neighbouring sounds (Munro, 1993; Flege, Munro and MacKay, 1995). The vowel portion was extracted leaving intact cues to vowel identity insofar as possible. Signal editing was carried out visually and aurally with CSL Kay
Elemetrics software by examining both the amplitude of the waveform and the formant structure on a spectrogram. The vowel portion included the first to the last positive peak in the periodic portion of the signal as indicated by an increase/decrease in overall amplitude and waveform complexity. Initial and final splices were made at zero crossings. Any vowel that was shorter than 140ms was modified by lengthening it to that duration, as was necessary with English lax vowels.

The best tokens from each talker were selected for the experiment on the basis of auditory judgment and spectrographic analysis. The same procedure was followed to prepare the Catalan vowels used for control purposes. In this case, stimuli were elicited from two male speakers of central Catalan and vowels were produced in isolation.

3.2.2. Procedure

The subjects were presented with randomized tokens of the four English vowels under study and had to choose from four options representing the Catalan high and mid front vowels and the Catalan diphthong /ei/ in conventional Catalan orthography, namely, ‘i’, ‘é’, ‘è’ and ‘ei’ (representing /i/, /e/, /e/ and /ei/, respectively). Subjects chose the alternative that best corresponded to the vowel they heard. After selecting an option, the same stimulus was heard again and subjects selected a goodness rating from a seven point scale according to how closely that sound approximated the Catalan vowel they had just selected. A score of ‘1’ corresponded to a poor exemplar of the chosen response vowel, and ‘7’ corresponded to a good exemplar. Subjects also heard and rated actual Catalan vowels, which were mixed in the task with the English vowels, for control purposes.
They were told the task consisted of identifying and providing goodness ratings for Catalan vowels and were not told that some vowels were not in fact Catalan.

There were two different talkers per vowel and each vowel appeared three times such that there were six tokens of each vowel. The vowel stimuli were presented in a randomized order. Each trial consisted of two presentations of each stimulus with an inter-stimulus interval of two seconds. After each trial subjects prompted the program for the following trial by clicking on the instructions on the screen. The task was preceded by a training period to familiarize the subjects with the procedure and adjust the listening level. Subjects took on average about twenty minutes to perform the PAT, which they did using an IBM laptop computer and good quality headphones (Sennheiser Headmax HD 400).

A preliminary test was conducted using a group of native Catalan speakers who lived in Ontario, Canada (Experiment 1). The test was also administered to a group of native Catalan speakers in Catalonia so as to test at least one group that had had little or no exposure to English (Experiment 2). It has been proposed that perceived similarity is best tested on monolingual speakers of the L1 because experience with the L2 may increase the listeners’ sensitivity to discern differences between the L1 and L2 sounds (Flege, 1991; Ingram and Park, 1997; Flege, 1995; Fox, Flege and Munro, 1995). Following this assumption, it is predicted that the Catalan speakers with experience in English will differentiate L1 and L2 sounds more readily (and thus perceive them as less similar) than the Catalans with no experience.
3.3. Experiment 1. Catalan native speakers in Ontario

3.3.1. Subjects

The subjects were Catalan native speakers living in the region of Toronto, Canada. The total number of subjects was twelve, all of them late learners of English who had spent a number of years in Canada, ranging from 7 to 39 with an average of 24 years. The group’s mean age was 53 (ranging from 31 to 67) and the average age of arrival in Canada was 29 (20 to 45). The subjects’ self-reported hearing ability was assessed during the training session by checking the correct perception of the first few trials, which involved real Catalan vowels. The volume was adjusted to the listener’s preferred level.

3.3.2. Results

The results of the PAT are given in Table 3.1, which provides the percentages of responses for each English target vowel. Table 3.2 shows the means collapsed across Talker for each vowel target. The results corresponding to the expected results on the grounds of acoustic similarity in F1 and F2 values are shown in boldface. Recall the values of the Catalan orthographic symbols: i: /i/, ei: /ei/, é: /e/, è: /e/.

<table>
<thead>
<tr>
<th>English target vowels</th>
<th>/i/</th>
<th>/e/</th>
<th>/ɪ/</th>
<th>/ɛ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses T1 GR T2 GR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% i</td>
<td>100</td>
<td>5.3</td>
<td>97</td>
<td>6.2</td>
</tr>
<tr>
<td>% ei</td>
<td>3</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% é</td>
<td>6</td>
<td>4.5</td>
<td>8</td>
<td>5.0</td>
</tr>
<tr>
<td>% è</td>
<td>8</td>
<td>6.3</td>
<td>6</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Table 3.1. PAT results with the Catalan group in Canada. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings (GR) (T indicates talker).
Table 3.2. PAT results with the Catalan group in Canada. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings (GR) across talkers.

<table>
<thead>
<tr>
<th>Responses</th>
<th>/i/</th>
<th>/e\textsuperscript{1}/</th>
<th>/u/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>% i</td>
<td>99</td>
<td>5.8</td>
<td>18</td>
<td>3.6</td>
</tr>
<tr>
<td>% ei</td>
<td>1</td>
<td>2.5</td>
<td>86</td>
<td>5.0</td>
</tr>
<tr>
<td>% é</td>
<td>7</td>
<td>4.8</td>
<td>58</td>
<td>3.5</td>
</tr>
<tr>
<td>% è</td>
<td>7</td>
<td>5.1</td>
<td>21</td>
<td>3.8</td>
</tr>
</tbody>
</table>

The English vowel /i/ was identified as Catalan /i/ on virtually all occasions (99%) and obtained a mean goodness rating of 5.5 out of 7, obtaining the same results as the control Catalan vowel /i/ which was correctly identified 94% of the time and given a mean rating of 5.5. These high assimilation scores of English /i/ to Catalan /i/ were followed closely by identification scores of English /e\textsuperscript{1}/ as the Catalan diphthong /ei/ (86%, mean rating of 5.0). It is noteworthy that this English vowel was consistently assimilated to a diphthong (/ei/) rather than to a monophthong (/e/), indicating that similarity is based not only on the steady state of the vowel but also on the formant movement. English /i/ yielded the greatest variation in responses, being most often heard as Catalan /e/ (58%), then as /e/ (21%) and /i/ (18%), and obtaining lower ratings (between 3.5 and 3.8). English /e/ was identified as Catalan /e/ 78% of the time (mean rating of 4.1) and 22% of the time as /e/ (mean rating of 3.9). Figure 3.2 below illustrates these results.
The proportion of trials on which the L2 target vowel was identified as the spectrally closest L1 vowel (e.g., English /i/-Catalan /i/, English /e/-Catalan /e/, English /e/-Catalan /e/, English /e/-Catalan /ei/) was calculated. These data were submitted to a 2 x 4 analysis of variance (ANOVA) with Talker and Vowel as within group factors. The main effect of Vowel was significant (F(3,33) = 6.71, p < .01). The main effect of Talker only reached marginal significance (F(1,11) = 4.12, p < .07) and the interaction was not significant. A Tukey HSD post-hoc test (collapsing across talkers, given that Talker did not yield a main effect nor an interaction) yielded significant differences between /i/ and /i/, and between /e/ and /e/.

That is, the assimilation scores for /i/ to the acoustically closest Catalan vowel (/ei/) were significantly lower than those for /i/ (p < .001), and for /e/ (p < .05).

The results indicate that English /i/ is strongly assimilated to Catalan /i/, whereas English /e/ appears to be the least similar with the greatest variation in responses and the
lowest goodness ratings. English /ε/ was most often assimilated to Catalan /e/ but with lower assimilation scores and goodness ratings than either /i/ and /e/. Further, the English vowels /i/ and /e/ consistently obtained higher goodness ratings than /ɪ/, stressing the greater perceived dissimilarity of vowel /ɪ/. Although the acoustic values predicted a fairly equal similarity between each pair of English-Catalan vowels, i.e., /i/-/ɪ/, /ɪ/-/e/, and /ε/-/ε/, the preliminary PAT results suggest a difference in perceived similarity. In order to more finely assess crosslinguistic phonetic distance and the potential influence of exposure to the L2, the same perceptual assimilation experiment was administered to a group of Catalan native speakers in Catalonia who had little knowledge of English.

3.4. Experiment 2. Catalan native speakers in Catalonia

Recent second language speech models claim that the acquisition of a second language phonology may affect the L1 phonology to a certain extent (e.g., Flege, 1995). Some models purport that success in establishing second language categories carries with it a shift in the L1’s phonetic categories, especially when the L1 and L2 vowels are relatively similar. Thus, a second PAT experiment was run with native Catalan speakers in their native land to investigate the possibility that the knowledge of English influenced the results of Experiment 1.

3.4.1. Subjects

Subjects were 20 native speakers of Catalan with little or no knowledge of English, or at least with little exposure to native English pronunciation. The subjects were
native speakers of the Central Eastern Catalan variety and although they were bilingual in Catalan and Spanish, they were all Catalan-dominant. This was assessed by means of a questionnaire about their language background and a brief interview with the experimenter. The age of the subjects ranged from 19 to 47, with a mean age of 28. This second set of subjects performed the same PAT as in Experiment 1, and thus the methodology and procedure were exactly as described in the previous section.

### 3.4.2. Results

The results do not differ much from those obtained in Experiment 1. Percentages of responses and goodness ratings by vowel and talker are given in Table 3.3. Table 3.4 shows the results collapsed across Talker. Figure 3.2 illustrates these results.

<table>
<thead>
<tr>
<th>Responses</th>
<th>/i/</th>
<th>/e/</th>
<th>/i/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>% i</td>
<td>100</td>
<td>5.7</td>
<td>98</td>
<td>6.6</td>
</tr>
<tr>
<td>% ei</td>
<td>2</td>
<td>1.0</td>
<td>82</td>
<td>5.0</td>
</tr>
<tr>
<td>% é</td>
<td>16</td>
<td>4.6</td>
<td>10</td>
<td>3.3</td>
</tr>
<tr>
<td>% è</td>
<td>2</td>
<td>1.5</td>
<td>3</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3. PAT results with the Catalan group in Catalonia. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings (GR) (T indicates talker).
Table 3.4. PAT results with the Catalan group in Catalonia. Mean percentages of classification of English vowels in terms of Catalan vowels and mean goodness ratings (GR) across talkers.

<table>
<thead>
<tr>
<th>Responses</th>
<th>/i/</th>
<th>/e\textsuperscript{i}/</th>
<th>/u/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>% i</td>
<td>99</td>
<td>6.2</td>
<td>13</td>
<td>2.7</td>
</tr>
<tr>
<td>% ei</td>
<td>1</td>
<td>1.0</td>
<td>84</td>
<td>4.6</td>
</tr>
<tr>
<td>% é</td>
<td>13</td>
<td>4.0</td>
<td>67</td>
<td>3.5</td>
</tr>
<tr>
<td>% è</td>
<td>3</td>
<td>2.8</td>
<td>21</td>
<td>3.0</td>
</tr>
</tbody>
</table>

The English vowel /i/ obtained the highest assimilation scores to the acoustically closest Catalan vowel (/i/) and the highest goodness ratings (99% and 6.2, respectively). This was followed by English /e/ and then /e\textsuperscript{i}/, which was again consistently identified with the Catalan diphthong (/ei/). Vowel /u/ was the least readily assimilated to an L1
vowel and obtained the lowest goodness ratings. As in Experiment 1, these results reveal a discrepancy between perceptual and acoustic comparisons because the latter would predict the same degree of assimilation for all four L2 vowels on the basis of the existence of an acoustically close L1 counterpart for each L2 vowel.

An analysis of variance on the percentage of assimilations to the acoustically closest vowel was performed with Vowel and Talker as within-subject factors. The main effect of Vowel was highly significant ($F(3,57) = 8.52, p < .001$). Neither the main effect of Talker nor the Vowel x Talker interaction was significant. With respect to differences between vowels, a Tukey HSD post-hoc test showed that the only significant differences between vowels involved vowel /u/, which obtained significantly lower assimilation scores to an L1 vowel than /i/ ($p < .001$), as was the case with the Catalan speakers in Toronto, and /e/ ($p < .01$).

These results from Catalan speakers with little or no knowledge of English are very comparable to those obtained in Experiment 1 involving the group of Catalan speakers in Toronto. Indeed, identification scores and goodness ratings were very close for /i/ and /e/ (99% and 84-86%, respectively). The lowest assimilation scores were observed for vowel /u/ in both cases, though scores were higher in Experiment 2 than in Experiment 1 (67% vs. 58%). The goodness rating provided by each group for /u/ was equivalent (3.5) and was the lowest rating in the experiments. A statistical analysis of the results of both groups taking Group as a between factor indicated that there was no main effect of Group and no interaction involving Group. Thus, the two groups did not differ significantly. The main effect of Vowel was significant, as expected ($F(3,90) = 14.17, p <$
The main effect of Talker was marginally significant, as was the Vowel x Talker interaction (p < .07 in both cases). The interaction is due to the fact the Talker was marginally significant for the Toronto group but not at all for the Barcelona group.

The results from the two experiments show that the two Catalan groups pattern very similarly, indicating that the effect of being bilingual speakers of Catalan and English in the case of the Catalans in Toronto did not affect their assimilation patterns greatly, at least in the case of the English vowels /i/ and /e/. Assimilation scores for /i/ and /e/ were numerically lower for the Catalan speakers of English than for the Catalans in Catalonia, apparently supporting earlier findings that learning English heightens learners’ awareness of phonetic distance between English and L1 vowels (Flege (1991) on L1 Spanish speakers, Ingram and Park (1997) on L1 Korean speakers). However, these differences did not reach significance in the current study. The Speech Learning Model’s prediction that L2 vowel acquisition may affect learners’ L1 vowel categories finds some support in that the Catalans in Catalonia were more successful than the Catalans in Toronto in their identification of the Catalan vowels (88-100% vs. 72-97% correct identification, respectively). Given the similarity in results in the two experiments, the rest of the discussion will consider mainly the results from Experiment 2 for the sake of simplicity and of conformity to the suggested preference for monolingual speaker data in perceptual assimilation tasks.

Therefore, according to the results of the perceptual assimilation task, English /i/ is perceived as identical to Catalan /i/ given the high assimilation scores (99%) and comparatively very high goodness rating (6.2). The English vowel /u/ appears to pattern as a weakly assimilated, similar vowel given that it is assimilated to the acoustically
The closest Catalan vowel (/e/) less often than the other English vowels and obtains the poorest ratings. Vowel /e/ obtains high identification scores (93%) but lower goodness ratings than /i/ (4.2) and English /ey/ obtains somewhat lower assimilation scores than /e/ (84%) yet higher ratings (4.6). The classification of these two vowels is not as straightforward since there is no clear criterion to determine what percentage of assimilation or what value of the goodness rating is the minimum to be considered an identical or highly assimilated sound. One alternative is to compare these results to the results obtained for the Catalan vowels, that is, the case of actual Catalan vowel identification. Catalan vowels were correctly identified between 88 and 100% of the time and their goodness ratings ranged from 4.8 to 6.3. Thus, scores for /ey/ do not fall within the identification range for native vowels and in the case of /e/ assimilation scores are as high as identification scores for native vowels but the goodness rating is lower.

To further investigate the relationship between the English and Catalan vowels, a third experiment was conducted involving the reverse of the perceptual assimilation task, that is, testing the perceptual assimilation of Catalan vowels to English vowels. Given that L1 for L2 vowel substitution is expected to go unnoticed by native L2 speakers in the case of identical vowels, a perceptual assimilation test of Catalan vowels to English vowels can predict whether vowels /e/ and /ei/ really pattern as identical to Catalan /e/ and /ei/. Thus, the perceptual assimilation task was administered to a group of native English speakers. Since the task includes both Catalan and English vowels, English native speakers’ identification and goodness ratings of the English vowels can also be obtained and thus compared to the Catalan perceptual assimilation scores.
3.5. Experiment 3. Perceptual assimilation of Catalan vowels to English vowels

The aim of this experiment was twofold. The first was to make sure that the English vowels were easily recognized by native English speakers and to obtain English vowel identification scores by native speakers to compare to the English-to-Catalan assimilation scores. The second goal was to test the perceptual similarity of L1 and L2 vowels from the perspective of the L2 vowels system to see if any of the Catalan vowels would be accepted as good exemplars of an English vowel, thus supporting the “identical” classification.

3.5.1. Subjects and procedure

A group of 20 native English speakers of Southern Ontario English participated in the experiment, mostly undergraduate students at the University of Toronto. The mean age was 21. The subjects were tested on the same stimuli using the same task as in Experiments 1 and 2 but in this case they identified the vowel stimuli by means of the keywords ‘beat’, ‘bit’, ‘bait’ and ‘bet’.

3.5.2. Results

With respect to the correct identification of the English vowels, scores ranged from 91 to 100%, similar to the correct vowel identification scores of the Catalan vowels by the Catalan native speakers (88-100%). The goodness ratings ranged from 4 to 5.9, slightly lower than the goodness ratings given to Catalan vowels by Catalan speakers (4.8-6.3). The identification scores were 99%, 97%, 100% and 91% for /i/, /i/, /e/ and /e/
respectively, with corresponding goodness ratings of 5.8, 4.6, 5.9, and 4.0. The results are shown in Table 3.5, which also includes the assimilation scores of Catalan vowels to English vowels. These results show that English vowels were readily identified by the native English speakers but varied in the goodness ratings they obtained. Lax vowels, especially /e/, obtained lower ratings. The goodness rating obtained for the English vowel /e/ in Experiment 2 (4.2) is in fact very close to the rating obtained with the native English speaker group. Thus, in terms of goodness ratings, English vowel /e/ is as good an exemplar of Catalan vowel /e/ as it is an exemplar of English vowel /e/, underscoring the high assimilation of English /e/ to Catalan /e/. With respect to English /e/, however, English native speakers’ identification scores and goodness rating for this vowel are higher than the Catalans’ assimilation and goodness rating (100% and 5.9 vs. 84% and 4.6, respectively). The perceptual similarity results (judging Catalan vowels in terms of English vowels) yielded very high assimilation scores for the Catalan vowels /i/, /e/ and /ei/, and lower scores for /e/, as shown in Table 3.5.

<table>
<thead>
<tr>
<th></th>
<th>English vowels</th>
<th>Catalan vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
<td>/e/</td>
</tr>
<tr>
<td>responses</td>
<td>%</td>
<td>GR</td>
</tr>
<tr>
<td>‘beat’</td>
<td>99</td>
<td>5.8</td>
</tr>
<tr>
<td>‘bait’</td>
<td>100</td>
<td>5.9</td>
</tr>
<tr>
<td>‘bit’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘bet’</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3.5. Identification and goodness ratings of English vowels and perceptual assimilation of Catalan vowels to English vowels by native English speakers (GR means goodness rating).
The high assimilation scores and goodness ratings indicate that the Catalan vowels /i/ and /e/ as well as the diphthong /ei/ would most likely go unnoticed when produced by Catalan speakers in place of English /i/, /e/ and /eɪ/ respectively in their L2 English speech. These results favour the interpretation of those three English vowels as identical to Catalan vowels. The Catalan high mid vowel /e/ obtained the lowest and most varied assimilation scores, emphasizing the weaker perceptual similarity between Catalan /e/ and English /ɪ/, as already seen in Experiments 1 and 2. Taken together, the results from the three experiments indicate that the English vowels /i/, /eɪ/ and /e/ are strongly assimilated to Catalan vowel categories whereas the English vowel /ɪ/ is a weakly assimilated vowel.

3.6. Discussion and conclusions

The first finding from this set of experiments concerns the discrepancy between acoustic and perceptual similarity. Although the acoustic data shows a very close correspondence between Catalan and English vowels, the perceptual assimilation task has shown that the English vowels assimilate to the Catalan vowels to different degrees. This discrepancy is consistent with the findings in previous studies (Stevens et al., 1996; Flege, 1991; Bohn, Strange and Trent, 1999), as discussed above, and illustrates that direct assessments of assimilation patterns are necessary for predicting perceptual difficulty. Further research is necessary to determine the precise relationship between acoustic information and vowel category formation. The current acoustic analysis has been limited
to F1 and F2 frequencies. A more thorough examination including formant movement patterns, modal pitch and voice quality features would be necessary to fully evaluate the relationship between acoustic and perceptual similarity.

With respect to the effect of experience, the results of the perceptual assimilation tasks do not detect any statistically reliable differences in assimilation patterns between the Catalans with no knowledge of English and the Catalan learners of English. This finding runs counter to earlier findings that L2 learning heightens the ability to discern crosslinguistic differences (Flege, 1991; Ingram and Park, 1997; Fox, Flege and Munro, 1995). However, it may be necessary to evaluate a larger number of vowels to better assess the effect of experience. The small number of vowels that were tested and the fact that three of them were strongly assimilated may account for the absence of significant difference between the two groups.

Although the current results are based only on vowels in isolation and a more detailed analysis of phonetic similarity between L1 and L2 sounds would need to consider contextual and allophonic variation, some initial conclusions can be drawn. Of the four target English vowels, three are strongly assimilated to Catalan vowels (/i/, /ɛ/ and /eɪ/) in identification scores and relative goodness ratings and thus can be considered identical or nearly identical L2 vowels. I will refer to these vowels as strongly assimilated. This classification is most clear with English vowel /i/, consistently categorized in terms of Catalan /i/ and with high goodness ratings. English /ɛ/ is also strongly assimilated despite lower goodness ratings on the grounds that its ratings were the same as those generated by native English speakers. Vowel /eɪ/ seems at first to be less strongly assimilated. However, the fact that the Catalan diphthong /ei/ is consistently
heard as English /ɛ/ by English listeners gives support to the classification of the English /ɛ/ as a strongly assimilated, nearly identical vowel as well. The finding that English /ɛ/ is more frequently assimilated to Catalan /ei/ than to the corresponding monophthong /ɛ/ indicates that vowel movement is a crucial feature in the categorization of this vowel by the Catalan learners. The remaining vowel (/i/) was not as consistently assimilated to an L1 vowel and obtained the lowest goodness ratings. I will refer to this vowel as weakly assimilated. The substitution of English /ɛ/ by the closest Catalan vowel (/e/) would not go unnoticed in this case given the perceptual dissimilarity evidenced in the L1-to-L2 perceptual assimilation (Experiment 3).

According to Flege’s (1995) Speech Learning Model, the greater the perceived phonetic dissimilarity between the L2 and L1 sounds, the greater the likelihood that L2 learners perceive that difference and consequently that a new vowel category is established. Thus, the model predicts that Catalan learners are more likely to establish a new category for the weakly assimilated /i/ than for the other three vowels. The Perceptual Assimilation Model (Best and Strange, 1992; Best, 1995) makes the same predictions as the Speech Learning Model in this case: given sufficient experience with the L2, a more authentic category will be created for the similar vowel than for the nearly identical vowels. Finally, a contrastive analysis approach would predict the opposite, that is, that /i/ would be more difficult to acquire because it is the most dissimilar.

The current proposal agrees with the prediction of the Speech Learning Model and the Perceptual Assimilation Model that learners are more likely to establish a new category for /i/ than for the other three vowels. This is because the strongly assimilated
vowels will be categorized in terms of their L1 counterparts. Vowel quality is the main feature learners will rely on in categorizing these vowels because it is the contrastive trait in the L1. Success in perception and production of these vowels will depend on whether the L1 category is based on the same vowel features and feature weights as the target vowel. The results from Experiment 3 suggest that production of the English vowels will involve L1 for L2 sound substitution, which will likely go unnoticed by English listeners. With respect to /i/, success in the acquisition of this vowel will entirely depend on the ability to create a new vowel category given the absence of an identical L1 vowel. It is proposed that in this case learners may develop sensitivity to non-L1 based properties, such as duration, given the unavailability of strong assimilation based on the L1 contrastive feature of vowel quality. Finally, it is also hypothesized here that category creation will be governed by the need to maintain and establish contrasts between vowels in the system so that even strongly assimilated vowels may be affected by the creation of a neighbouring new category.

The results from the perceptual assimilation experiments and their predictions will be tested in the following chapters by means of vowel perception and production experiments.
Chapter Four. Perception of L2 vowels

4.0. Introduction

The experiments on phonetic similarity in Chapter 3 have shown that the English vowels /i/, /e/ and /ɛ/ are strongly assimilated to the Catalan vowels /i/, /e/ and the diphthong /ei/, respectively, whereas the English vowel /ɛ/ is weakly assimilated to the Catalan /e/. This assimilation is based on perceived similarity between L1 and L2 vowels in terms of vowel quality. Given the perceived similarity between L2 and L1 sounds, the following predictions can be made. No new vowel categories will be established for the strongly assimilated vowels, which will be perceived in terms of the L1 vowel categories. Success in perception of the strongly assimilated L2 vowels will depend on whether the L1 counterparts are based on the same vowel features and feature weights as the target vowel. In the case of the weakly assimilated vowel, successful acquisition will depend upon the learners’ ability to create a new target-like phonetic category. In addition, Catalan speakers will make more use of vowel features other than vowel quality in their categorization of the weakly assimilated vowel (/ɛ/). Finally, the need to create a new category for the weakly assimilated vowel may affect neighbouring vowel categories, regardless of their degree of assimilation.

The ability to identify L2 vowels is tested on two groups of Catalan speakers, a group of Catalan immigrants in Canada and a group of undergraduate university students specializing in English Philology at a Catalan university. This allows the examination of the effect of experience and L2 input on the acquisition of the target language. Both groups consisted of adult learners of English but differed in two important ways. Firstly,
the amount of exposure to the L2 and the number of years of learning the L2 was greater
with the Catalans in Canada, predicting a clearer effect of experience on the
categorization of L2 vowels (according to both Flege’s (1995) Speech Learning Model
and Best’s (1995) Perceptual Assimilation Model). The group of university students,
however, had substantial formal instruction in the L2. Therefore, the effect of experience
vs. instruction will be contrasted. Secondly, the Catalans in Toronto were exposed mostly
to Canadian English whereas the undergraduate students had L2 input from different
English varieties. As discussed below, the effect of the L2 input will be crucial when
evaluating whether learners’ reliance on duration in distinguishing the tense-lax vowel
contrast is determined by the L2 variety that constitutes the main input.

This chapter examines the perceptual ability of Catalan learners of English by
means of a series of vowel identification tasks. Three experiments will evaluate
perception of natural stimuli in three different environments. The first experiment tests
vowels in isolation (Experiment 4), the second examines vowel perception in a
monosyllabic environment (Experiment 5) and the third one tests perception in a
bisyllabic environment (Experiment 6). Finally, given the different acoustic cues
involved in the lax and tense vowel contrast, the effect of the relative weighting of
spectral and temporal cues will be examined in a vowel identification experiment
involving synthetic stimuli manipulated in terms of vowel quality and duration
(Experiment 7).
4.1. Vowel identification involving natural stimuli

Perception of the target vowels was tested by means of vowel identification tasks in which stimuli were identified as one of four target vowels. The tasks differed with respect to the environment in which the vowels were presented. In the first task (Experiment 4) vowels were presented in isolation. This task was intended to test the perception of the vowels with no intervening segmental material. Testing vowels in isolation also made possible the comparison to Experiment 7, which involves synthetic vowels presented in isolation (section 4.2). Experiment 5 examines perception of the target vowels in CVC environment so as to evaluate perception in a more natural environment, particularly in the case of the lax vowels for which the environment without a following consonant is not permissible. In addition, studies of vowel perception show that vowels are better perceived in CVC contexts than in isolation (Gottfried and Strange, 1980; Strange and Gottfried, 1980). Lastly, Experiment 6 tests perception in a bisyllabic environment. This third environment was chosen to determine perceptual ability in longer words in which overall vowel duration is shorter than in monosyllables and vowel duration differences are thus smaller. In addition, testing different environments allows an evaluation of the relationship between segmental and non-segmental structure. Recall from Chapter 1 that the monosyllabic environment is hypothesized to be a better environment for maintaining the tense-lax contrast given the clear tautosyllabicity of the vowel-consonant sequence, whereas interference from L1 syllabification in the CVCVC environment may leave the first vowel in an open syllable, triggering mispronunciation of the lax vowels as tense vowels. This will be tested by assessing the L2 learners’ knowledge of the LVC and of English syllabification (Chapter 5) and by testing the
production of the vowels in the different word environments (Chapter 6). In order to fully assess the validity of this claim, the perception experiments also include the environment variable.

Finally, the use of natural stimuli from different talkers will allow the examination of the relative importance of vowel properties. As will be discussed below, the talkers’ vowel productions varied in length and diphthongization for vowel /e/. This variation was exploited to assess the relative weighting of phonetic cues on vowel identification by native English speakers and Catalan speakers.

The same subjects participated in all three experiments (Experiments 4, 5, and 6). The stimulus elicitation and preparation, and the task procedures were also the same for the three experiments. The subjects and the methodology are thus discussed next.

4.1.1. Subjects

Three groups participated in the experiments: a group of 18 Catalan native speakers residing in Toronto (TC group), a group of 30 Catalan native speakers in Barcelona, Spain (BC group), and a control group of 13 native English speakers from Southern Ontario tested in Toronto (EN group). The TC subjects were Catalan native speakers living in and around Toronto, Canada, all of them late learners of English who had spent a number of years in Canada, ranging from 7 to 39 with an average of 25 years in Canada. The group’s mean age was 55 (ranging from 34 to 72) and the average age of arrival in Canada was 31 (ranging from 20 to 45). The subjects in Catalonia (BC group) were undergraduate students in English Philology at the Universitat Autònoma de Barcelona, in Barcelona, Spain. They were in their third or fourth university year with a
mean age of 22. They were speakers of L2 English, English being the language of instruction and study in most of their courses. In addition, many had spent between a few weeks and a year in an English speaking country. Both BC and TC were bilingual in Catalan and Spanish, but they were Catalan-dominant bilinguals as assessed by a questionnaire about their language background and a brief interview with the experimenter. Finally, the English native speaker group was mostly made up of undergraduate and graduate students at the University of Toronto with a mean age of 36.

4.1.2. Stimuli

English vowel stimuli were elicited from four male speakers of Canadian English. The use of multiple stimuli is not uncommon in vowel perception experiments and it is used to encourage listeners to process vowels in a general mode (Flege, Munro and Fox, 1994). Each talker read a list containing English target vowels in monosyllabic words of the shape /h/ + vowel (with tense vowels only, e.g., ‘hee’ (/hi/) and ‘hay’ (/he/) ), /h/ + vowel + /b/ (e.g., ‘heeb’ (/hib/), ‘hib’ (/hɪb/), ‘habe’ (/heb/), ‘heb’ (/hɛb/)) and in CVCVC words with the target vowel in a stressed h_b first syllable followed by an unstressed syllable with a reduced vowel (e.g., ‘heebut’ (/hibʌt/), ‘hibbut’ (/hɪbʌt/), ‘haibut’ (/heɪbʌt/), ‘hebbut’ (/hɛbʌt/)). The h_b consonant environment was chosen in order to minimize C to V and V to C tongue coarticulation (Strange et al., 1998), and to aid in the extraction of the vowel in order to prepare the vowel stimuli for the first experiment. Five tokens of each word were elicited. Stimuli were recorded using a SONY digital-audio tape recorder. The words were subsequently redigitized at a 10kHz sampling rate with 16 bit resolution and normalized for peak intensity. The stimuli for the
experiment involving vowels in isolation (Experiment 4) were the same as the English vowel stimuli used in the perceptual assimilation tasks described in Chapter 3 (Experiments 1, 2 and 3), plus a set of tokens from an additional third speaker. Thus, the vowel portion was extracted leaving intact cues to vowel identity insofar as possible. Signal editing was carried out visually and aurally with CSL Kay Elemetrics software by examining both the amplitude of a waveform and the formant structure on a spectrogram. The vowel portion included the first to the last positive peak in the periodic portion of the signal as indicated by an increase/decrease in overall amplitude and waveform complexity. Initial and final splices were made at zero crossings.

In addition to facilitating vowel extraction for Experiment 4, the choice of words of the shape h_b and h_but for Experiments 5 and 6 permitted the use of a set of non-sense words including all critical vowels. This was desirable so as to rule out any lexical interference in the vowel identification task resulting from a positive effect of word frequency. With nonsense words, higher scores could not be attributed to lexical status, or at least, such an effect was minimized. The best tokens from each talker were selected for the experiment, on the basis of both auditory judgment and spectrogram analysis.

Importantly, this analysis revealed some notable talker differences, mostly in vowel length and formant transitions. The most important talker differences involved the pronunciation of the vowel /eɪ/. This is clear when comparing the initial (steady state) and final (offglide) F1 and F2 values for the different talkers, as shown in Tables 4.1 to 4.3. The offglide trajectory is notably shorter in the case of Talker 2 (T2) and Talker 4 (T4). Although all talkers were native Canadian English speakers living in Toronto, the difference in diphthongization may have arisen from dialectal differences given that T2
and T4 were originally from western and northwestern Ontario, respectively. On the other hand, T4’s much shorter vowel durations may also account for the absence of greater formant movement. Given that the relative importance of acoustic properties in the categorization of L2 vowels is central to the current study, the effect of these differences was incorporated by comparing talkers with different vowel properties in each experiment.

<table>
<thead>
<tr>
<th>Talker</th>
<th>F1 Initial</th>
<th>F2 Initial</th>
<th>F1 Final</th>
<th>F2 Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talker 1</td>
<td>488</td>
<td>2018</td>
<td>360</td>
<td>2230</td>
</tr>
<tr>
<td>Talker 2</td>
<td>488</td>
<td>1968</td>
<td>440</td>
<td>2100</td>
</tr>
<tr>
<td>Talker 3</td>
<td>475</td>
<td>2050</td>
<td>300</td>
<td>2300</td>
</tr>
</tbody>
</table>

Table 4.1. Formant values for vowel /e/ in Experiment 4 (isolated vowels).

<table>
<thead>
<tr>
<th>Talker</th>
<th>F1 Initial</th>
<th>F2 Initial</th>
<th>F1 Final</th>
<th>F2 Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talker 1</td>
<td>488</td>
<td>2100</td>
<td>340</td>
<td>2350</td>
</tr>
<tr>
<td>Talker 2</td>
<td>488</td>
<td>2000</td>
<td>422</td>
<td>2100</td>
</tr>
</tbody>
</table>

Table 4.2. Formant values for vowel /e/ in Experiment 5 (CVC words).

<table>
<thead>
<tr>
<th>Talker</th>
<th>F1 Initial</th>
<th>F2 Initial</th>
<th>F1 Final</th>
<th>F2 Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talker 1</td>
<td>480</td>
<td>2150</td>
<td>360</td>
<td>2275</td>
</tr>
<tr>
<td>Talker 4</td>
<td>530</td>
<td>2200</td>
<td>515</td>
<td>2160</td>
</tr>
</tbody>
</table>

Table 4.3. Formant values for vowel /e/ in Experiment 6 (CVCVC words).
On account of the strong assimilation of English /e/ to the Catalan diphthong /ei/ evidenced in the perceptual assimilation experiments, we can predict that Catalan learners may be less successful at identifying more monophthongal /e/ tokens, that is, tokens that are less like the L1 diphthong /ei/ to which it is strongly assimilated. This prediction follows from the hypothesis that strongly assimilated L2 vowels are categorized in terms of L1 features and that they will be accurately perceived and produced when the L1 and L2 vowels are based on the same cues.

With respect to duration, T2 and T4’s vowels were shorter than the other talkers’, particularly in the case of /i/. In CVCVC words, vowels produced by T4 were considerably shorter than those produced by T1. Duration values for all talkers and vowels are given in Table 4.4, grouped by experiment. Although the effect of duration will be more specifically tested in the experiment involving synthetic stimuli (Experiment 7), these differences with natural stimuli allowed further testing of the effect of different cues on L2 perception.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Talker</th>
<th>/i/</th>
<th>/i/</th>
<th>/e/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolated Vs. (Exp. 4)</td>
<td>Talker 1</td>
<td>255</td>
<td>140</td>
<td>340</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>Talker 2</td>
<td>218</td>
<td>140</td>
<td>350</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>Talker 3</td>
<td>280</td>
<td>140</td>
<td>323</td>
<td>142</td>
</tr>
<tr>
<td>CVC words (Exp. 5)</td>
<td>Talker 1</td>
<td>276</td>
<td>164</td>
<td>276</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>Talker 2</td>
<td>235</td>
<td>116</td>
<td>234</td>
<td>132</td>
</tr>
<tr>
<td>CVCVC words (Exp. 6)</td>
<td>Talker 1</td>
<td>172</td>
<td>93</td>
<td>166</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>Talker 4</td>
<td>85</td>
<td>48</td>
<td>98</td>
<td>75</td>
</tr>
</tbody>
</table>

Table 4.4. Vowel durations in ms by talker for each experiment.
4.1.3. Procedure

The three experiments involved vowel identification tasks. The task methodology was the same for the three experiments, the difference being the kind of stimuli presented (e.g., vowels in isolation, in CVC words and in CVCVC words respectively for Experiments 4, 5 and 6). Subjects had to identify the aural stimuli in terms of one of the four English vowels presented as possible responses (four-alternative forced-choice tasks). Response alternatives were four pairs of English words representing each of the four target English vowels under study (/i/, /u/, /e/, /e/): <see/beat>, <sit/bid>, <say/bait> and <set/bed>. Subjects chose the alternative that contained the vowel they heard (Experiment 4) or contained the same vowel as the (stressed) vowel in the stimulus word. A trial consisted of two presentations of each stimulus with an inter-stimulus interval of two seconds. After each trial subjects prompted the program for the following trial by clicking on the instructions on the screen. All tasks were preceded by a practice period to familiarize the subjects with the procedure and adjust the listening level. Subjects took on average about twenty minutes to complete the three experiments, which they did using an IBM laptop computer and good quality headphones (Sennheiser Headmax HD 400).

4.1.4. Experiment 4. Vowels in isolation

The first identification task involved vowels in isolation. There were four repetitions of each of the four vowels produced by three different talkers (that is, 4 repetitions x 4 vowels x 3 talkers = 48 tokens). The percentages of responses for each target vowel and for each talker obtained for the native English speakers, the Catalan
group tested in Toronto and the Catalan group tested in Barcelona are given in Tables 4.5, 4.6 and 4.7, respectively, which also include the percentages across talkers. Correct identification percentages are highlighted in bold (T1, T2, T3 stand for Talker 1, 2 and 3).

<table>
<thead>
<tr>
<th>Responses</th>
<th>English target vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
</tr>
<tr>
<td></td>
<td>T1</td>
</tr>
<tr>
<td>% beat/sea</td>
<td>100</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>8</td>
</tr>
<tr>
<td>% bait/say</td>
<td></td>
</tr>
<tr>
<td>% bed/set</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.5. English native speakers. Identification scores for English target vowels in isolation and means across talkers

<table>
<thead>
<tr>
<th>Responses</th>
<th>English target vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
</tr>
<tr>
<td></td>
<td>T1</td>
</tr>
<tr>
<td>% beat/sea</td>
<td>78</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>21</td>
</tr>
<tr>
<td>% bait/say</td>
<td>4</td>
</tr>
<tr>
<td>% bed/set</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.6. Catalans in Toronto (TC group). Identification scores for English target vowels in isolation and means across talkers
Responses | English target vowels
---|---

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/ɪ/</th>
<th>/eɪ/</th>
<th>/e/</th>
<th>Across 3 Talkers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td>/i/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɪ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/eɪ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Across 3 Talkers</td>
<td>100</td>
<td>84</td>
<td>93</td>
<td>9</td>
<td>92</td>
</tr>
</tbody>
</table>

% beat/sea

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>/i/</th>
<th>/ɪ/</th>
<th>/eɪ/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>100</td>
<td>84</td>
<td>93</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɪ/</td>
<td>16</td>
<td>7</td>
<td>90</td>
<td>91</td>
<td>88</td>
<td>2</td>
<td>9</td>
<td>2</td>
<td>13</td>
<td>11</td>
<td>2</td>
<td>8</td>
<td>90</td>
</tr>
<tr>
<td>/eɪ/</td>
<td>1</td>
<td>1</td>
<td>98</td>
<td>73</td>
<td>98</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>18</td>
<td></td>
<td>87</td>
<td>89</td>
<td>96</td>
<td>6</td>
<td>6</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7. Catalans in Barcelona (BC group). Identification scores for English target vowels in isolation and means across talkers

The control group of native English speakers obtained very high vowel identification scores, indicating that for English speakers there was little or no difficulty in distinguishing the vowels and recognizing the vowel stimuli as the intended vowel targets. Non-native speakers obtained lower correct identification scores than the native speaker group. The two non-native groups pattern somewhat differently in that the TC group had lower scores. Nevertheless, correct vowel identification was fairly high (91% for BC and 80% for TC). Overall, the BC group obtained high identification scores with the exception of the lower scores for the tense vowels produced by Talker 2 (T2), especially in the case of /eɪ/. The TC group showed a clearer difference between vowels, obtaining better identification scores for /eɪ/ and /e/ than for /i/ and /ɪ/, with a similar drop in correct identifications for the tense vowels produced by T2.

An analysis of variance on the percentage of correct identifications with Language Group as the between factor and Talker and Vowel as within factors revealed that all main effects were significant (Language Group: F(2,58) = 17.22, p < .001; Vowel: F(3,174) = 10.11, p < .001; Talker: F(2,116) = 19.21, p < .001), as were all
interactions. A Tukey HSD post-hoc test showed that all three groups differed significantly (TC vs. EN, p < .001; BC vs. EN, p < .05; TC vs. BC, p < .001). The post-hoc tests also showed that there were no significant differences between vowels either for EN or BC, but scores for vowels /i/ and /u/ were significantly lower than those for /e/ and /e/ in the case of TC (p < .001), hence the Vowel x Language Group interaction (F(6,174) = 10.6, p < .001. In addition, the TC group’s scores for /i/ and /u/ were significantly lower than those obtained by EN and BC (p < .001 in all cases). Post-hoc tests showed that Talker had an effect for the two L2 groups but no effect for the native speaker group, which explains the Talker x Language Group interaction, F(4,116) = 3.76, p < .01. The vowels produced by T2 obtained significantly lower identification scores than the vowels produced by the other two talkers (p < .01 for the BC group and p < .001 for the TC group). The effect of Talker is clear in the case of /i/ and /e/, but not /u/ and /e/, which explains the Vowel x Talker interaction (F(6,348) = 6.68, p < .001). Per vowel, talker differences reached significance in the case of /i/ for the TC group and /e/ for the BC group (thus the Vowel x Talker x Language Group interaction, F(12,348) = 3.65, p < .001.

It is interesting to note that there was no effect of Talker for native speakers. The differences between talkers in the pronunciation of the tense vowels, especially /e/, was picked up by the two non-native groups to the point that it affected the correct identification of vowels, but did not affect the native speakers. As shown above, T2’s production of /e/ had a clearly shorter offglide trajectory than the other talkers’. Thus, Catalan learners were less successful at identifying the highly assimilated vowel /e/ when
it was less like the L1 diphthong /ei/ to which it is strongly assimilated. This shows that offglide transitions constitute a more important cue for Catalan speakers than Canadian English speakers. In addition, it supports the prediction that accurate perception requires the L2 category to be based on the same vowel properties and weights as the target vowel. With respect to T2’s /i/, there was no clear difference in F1-F2 structure but T2’s /i/ was somewhat shorter than the other /i/ tokens (218 ms vs. 255-280 ms). This may account for the lower scores because, as will be shown below, duration plays an important role in the identification of English /i/ by Catalan speakers.

As expected, native speakers scored higher than non-native speakers. Although not as good as the native speakers, as a group, the 30 English Philology undergraduate students (BC) outperformed the 18 Catalans living in Ontario (TC). The superiority of BC’s results, which was also found in subsequent experiments, will be discussed below. For the BC group, identification scores were similar for all four vowels, with some improvement for the tense vowels if scores for T2’s vowels are disregarded. On the other hand, TC shows different results for different vowels. Identification scores are better for /eɪ/ and /ɛ/ than for /i/ and /u/, the latter vowel obtaining the lowest identification scores. These vowel differences go in the direction predicted by the perceptual assimilation results in the case of the highly assimilated /eɪ/ and /ɛ/ and the weakly assimilated /u/, but the lower identification scores for /i/ are unexpected for a strongly assimilated vowel. Before discussing results more globally, let us turn to the results for the remaining vowel identification tasks.
4.1.5. Experiment 5. Vowels in monosyllables

The participants and procedure for this vowel identification task were the same as in Experiment 4. There were five repetitions of each of the four vowels produced by two different talkers (that is, 4 vowels x 2 talkers x 5 repetitions = 40 tokens). The talkers were T1 and T2 from the previous experiment. The difference was that in this case vowels were presented in a CVC context, namely between /h/ and /b/. The purpose of this task was to examine the perception of the target vowels in a closed syllable environment and test whether it would be a facilitating context in the perception of lax vowels. This environment permitted testing the perception of these vowels in a word-like sequence, a more natural environment than vowels in isolation, while still avoiding the use of real words where frequency effects could have an influence in rate of identification.

The native speaker group had no difficulty in identifying the vowels and stimuli were correctly identified 100% of the time. The results for the two non-native groups are presented in Tables 4.8 and 4.9.

<table>
<thead>
<tr>
<th>Responses</th>
<th>English target vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
</tr>
<tr>
<td>% beat/sea</td>
<td>80</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>18</td>
</tr>
<tr>
<td>% bait/say</td>
<td>1</td>
</tr>
<tr>
<td>% bed/set</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.8. Catalans in Toronto (TC Group). Identification scores for English target vowels in the CVC context and means across talkers
Table 4.9. Catalans in Barcelona (BC Group). Identification scores for English target vowels in the CVC context and means across talkers

<table>
<thead>
<tr>
<th>Responses</th>
<th>/i/</th>
<th>/ɪ/</th>
<th>/ɛ/</th>
<th>/ɛ/</th>
<th>Across Talkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>% beat/sea</td>
<td>100</td>
<td>11</td>
<td>13.5</td>
<td>1</td>
<td>94  6.5  21  0.5</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>13</td>
<td>88</td>
<td>5.5</td>
<td>2</td>
<td>6   93  17  2</td>
</tr>
<tr>
<td>% bait/say</td>
<td></td>
<td>81</td>
<td>40</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>% bed/set</td>
<td>1</td>
<td>3</td>
<td>97</td>
<td>98</td>
<td>0.5 1 97.5</td>
</tr>
</tbody>
</table>

The analysis of variance on the percentage of correct responses with Language Group as the between factor and Talker and Vowel as within factors showed that all main effects and interactions were significant (Language Group: F(2,58) = 14.77, p < .001; Vowel: F(3,174) = 10.38, p < .001; Talker: F(1,58) = 17.19, p < .001; Language Group x Vowel: F(6,174) = 5.00, p < .001; Language Group x Talker: F(1,58) = 28.00, p < .001; Vowel x Talker: F(3,174) = 6.49, p < .001; Language Group x Vowel x Talker: F(6,174) = 9.73, p < .001). A Tukey HSD post hoc test indicated that both non-native speaker groups differed significantly from the native speaker group (p < .001) but did not differ from one another. Thus, while the BC group outperformed the TC group in the previous experiment, both non-native speaker groups pattern together in this experiment. Nevertheless, the scores for the BC group are still numerically higher than those for the TC group with the exception of the results for vowel /ɛ/, as discussed below. The main effect of Talker is clear in the lower responses for T2 in the case of tense vowels. The interactions involving Talker are explained by the absence of a Talker effect for EN, and the fact that Talker was only significant for /ɛ/ in the BC and TC groups, as indicated by
a follow-up Tukey HSD post-hoc test. The main effect and interactions involving Vowel
are explained by the higher identification scores obtained by /e/1, followed by /i/ and /u/,
and the lower scores for /e/1. The BC group obtained significantly lower identification
scores for /e/1 than for the other three vowels (p < .001 on a Tukey HD post hoc test). In
the case of the TC group, /e/ obtained significantly higher scores than /i/ (p < .05).
Comparing groups, TC scored significantly lower than EN in the case of vowel /i/
(p < .001), and /u/ and /e/1 (p < .05), while BC scored lower than EN only with respect to
vowel /e/1 (p < .001). Finally, the two L2 groups differed in their scores for /i/ and /e/1.
The BC group’s scores for /i/ were higher (p < .001), as in Experiment 4, but their scores
for /e/1 were lower (p < .001).

The lower scores obtained for tense vowels produced by Talker 2 parallel the
results in Experiment 4, especially with respect to vowel /e/1, the case in which Talker
differences reached significance. As discussed above, given the categorization of English
/e/1 as the Catalan diphthong /ei/, more monophthongal /e/1 tokens are less readily
identified by the Catalan speakers. The effect of Talker is greater in the current
experiment than in the previous experiment. The greater difficulty in identifying /e/1 in
the CVC environment than in isolation is related to the fact that the former was clearly
longer (351ms vs. 234ms), having been extracted from an open syllable (h+vowel) frame.
In addition, the lower identification rates in the CVC syllable may be connected with the
difficulty in distinguishing the glide of the diphthong from the glide-like transition to /b/.
Importantly, the results for /e/ illustrate another important fact, namely the influence of the L2 input. The fact that TC outperforms BC with respect to perception of /e/, while obtaining worse results for all other vowels, can be related to a difference in L2 input. Recall that TC consisted of Catalan subjects living in Ontario. Through exposure to this dialect, the TC group may have become more attuned to the characteristics of this vowel in the talkers’ dialect.

To sum up, the results for vowels /i/, /u/ and /e/ parallel the results obtained in the previous experiment. BC obtains good identification scores for all three vowels whereas TC is most successful with /e/ obtaining lower scores for the other two vowels. The difference between the two experiments lies in the lower scores for /e/ explained by the low identification rates obtained for /e/ produced by Talker 2.

Finally, with respect to a possible facilitating effect of the CVC context in the perception of lax vowels, context does not affect vowel /e/, whose identification rates were already high when perceived in isolation in Experiment 4. In the case of vowel /u/, the closed-syllable environment is not relevant for the BC group, whose scores were high in both experiments (90% in isolation and 93% in CVC) but is seemingly beneficial for the TC group, whose scores in Experiment 4 had been lower (64% vs. 82%). The issue of the environment is discussed in more detail after discussing Experiment 6.

4.1.6. Experiment 6. Vowels in bisyllabic environments

This experiment extended the investigation of L2 vowel perception to the bisyllabic environments. The duration of a stressed vowel in a bisyllable is shorter than in
a monosyllable. This task thus also allows us to examine if perception of lax and tense vowels by L2 learners, who have been found to rely often on vowel duration differences (see Chapter 1), is affected by absolute or relative duration. This was further investigated by including stimuli from two talkers (T1 and T4) who differed not only in the characteristics of /e/, but also in speech rate and thus in vowel duration, as illustrated in Table 4.4 above. In addition, as discussed above, testing different environments is important for evaluating the hypothesized relationship between the transferability of L1 syllabification patterns and the ability to produce the lax vowels in closed syllables.

The participants and procedure were the same as in Experiments 4 and 5. There were five repetitions of each of the four vowels produced by two different talkers (that is, 4 vowels x 2 talkers x 5 repetitions = 40 tokens). The target vowels were presented in CVCVC non-sense words (e.g., heebut, hibbut, haibut, hebbut). The results are presented in Tables 4.10, 4.11 and 4.12.

<table>
<thead>
<tr>
<th>Responses</th>
<th>English target vowels</th>
<th>Across Talkers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>T1</td>
<td>T4</td>
<td>T1</td>
</tr>
<tr>
<td>% beat/sea</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>% bait/say</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% bed/set</td>
<td>4.5</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.10. English native speakers. Identification scores for English target vowels in the CVCVC context and means across talkers
Table 4.11. Catalans in Toronto (TC Group). Identification scores for English target vowels in the CVCVC context and means across talkers

<table>
<thead>
<tr>
<th>Responses</th>
<th>English target vowels</th>
<th>Across Talkers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
<td>/u/</td>
</tr>
<tr>
<td>T1  T4</td>
<td>T1  T4</td>
<td>T1  T4</td>
</tr>
<tr>
<td>% beat/sea</td>
<td>83  53</td>
<td>27  4.5</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>13  46</td>
<td>60  69</td>
</tr>
<tr>
<td>% bait/say</td>
<td>4   6.5 4.5</td>
<td>93  59</td>
</tr>
<tr>
<td>% bed/set</td>
<td>1   6.5 22</td>
<td>1   12</td>
</tr>
</tbody>
</table>

Table 4.12. Catalans in Barcelona (BC Group). Identification scores for English target vowels in the CVCVC context and means across talkers

<table>
<thead>
<tr>
<th>Responses</th>
<th>English target vowels</th>
<th>Across Talkers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/i/</td>
<td>/u/</td>
</tr>
<tr>
<td>T1  T4</td>
<td>T1  T4</td>
<td>T1  T4</td>
</tr>
<tr>
<td>% beat/sea</td>
<td>100 56</td>
<td>26  1</td>
</tr>
<tr>
<td>% bid/sit</td>
<td>44  74 97</td>
<td>43  2 1</td>
</tr>
<tr>
<td>% bait/say</td>
<td>100 39</td>
<td>1</td>
</tr>
<tr>
<td>% bed/set</td>
<td>2   11</td>
<td>97  99</td>
</tr>
</tbody>
</table>

The ANOVA on the percentage of correct identifications in this case also showed that all main effects and interactions were significant (Language Group: $F(2,58) = 23.53$, $p < .001$; Vowel: $F(3,174) = 6.49$, $p < .001$; Talker: $F(1,58) = 28$, $p < .001$; Language Group x Vowel: $F(6,174) = 5$, $p < .001$; Language Group x Talker: $F(2,58) = 12.87$, $p < .05$; Vowel x Talker: $F(3,174) = 26.12$, $p < .001$; and the three way interaction: $F(6,174) = 9.73$, $p < .001$). EN obtained significantly higher identification scores than the two L2 groups, and BC’s scores were also significantly higher than TC’s (as shown in
a Tukey HSD post hoc test: EN vs. BC: p < .001; EN vs. TC: p < .001; BC vs. TC: p < .01). Thus, the BC group performed better than the TC group with vowels in isolation and in CVCVC words (Experiments 4 and 6) but both groups performed similarly with CVC stimuli (Experiment 5). Tukey HSD post-hoc tests also showed that the EN group’s identification scores were significantly higher than those of the TC group for all four vowels, and for vowels /i/ and /e/ in the case of the BC group. The BC group also scored significantly higher than the TC group in the case of the lax vowels /ɪ/ and /e/. Vowel had an effect for the BC group but not for the other two groups: scores for /e/ were significantly lower than scores for the other three vowels, and scores for /i/ were also significantly lower than scores for /e/. With respect to the main effect of Talker, Talker 4 yielded lower identification scores than Talker 1 and, as in the previous two experiments, the difference was significant for the two Catalan groups. The Tukey post hoc test showed that Talker had an effect for vowels /i/ and /e/ in the TC group and for vowels /i/, /ɪ/ and /e/ in the BC group.

It is clear from the recurrent talker effect in the three vowel identification experiments that the high offglide is a crucial cue in the identification of vowel /e/. More monophthongal /e/ tokens are often misheard as /i/ or /ɪ/. Notice also that, as with T2’s productions, the TC group outperforms the BC group in the identification of T4’s /e/ (59% vs. 39%, respectively). As discussed above, this can be explained by the familiarity with Canadian English varieties on the part of TC. On the other hand, English native speakers’ perception of this vowel does not rely on the offglide as much as on the
vowel’s steady state formant structure. This illustrates a difference in acoustic cue reliance between native speakers and L2 learners and supports the view that accurate cue weighting is a prerequisite for target-like categorization. On the other hand, T4’s much shorter vowel durations (see Table 4.4 above) may also account for the absence of greater formant movement. The disparity in vowel duration between the two talkers’ productions provides another example of the importance of acoustic cue reliance. Talker differences in vowel duration do not affect the EN group’s responses. Catalan subjects, on the other hand, seem to prefer longer tense vowels, thus readily identifying T1’s /i/ and /e/ tokens but failing to identify T4’s; T4’s /i/ was in fact shorter than T1’s /u/. Thus, learners seem to be listening to the vowels’ absolute duration rather than their relative duration in comparison to other vowels. This kind of Talker effect in L2 perception studies is not totally unexpected. Ingram and Park (1997) found that individual talker differences in absolute vowel duration influenced the patterns of assimilation of English vowels to Korean and Japanese vowels. The relative weighting of acoustic cues will be examined further in the last perception experiment in this chapter (section 4.2).

In brief, the CVCVC environment yields overall lower vowel identification scores for both Catalan groups. The effect of the environment is analyzed in the following section, which compares the results for the three experiments. As in previous experiments, the BC group outperforms the TC group, especially in the case of lax vowels. The weakly assimilated vowel /u/ appears as the most problematic for the TC group, but is comparatively well identified by the BC group, in accordance with results from the previous experiments. On the whole, of the three highly assimilated vowels, /e/ is the most successfully identified by non-native speakers regardless of talker differences
in duration. Differences in duration and diphthongization affect Catalan subjects’ perception of the other two highly assimilated vowels, namely /i/ and /e/, pointing to the need to examine relative cue weighting in L2.

### 4.1.7. Conclusions

This section provides a summary of the main conclusions from the three perception experiments. Table 4.13 presents the correct identification means for all three experiments.

<table>
<thead>
<tr>
<th></th>
<th>English speakers</th>
<th>Catalans in Toronto</th>
<th>Catalans in Barcelona</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp. 4</td>
<td>Exp. 5</td>
<td>Exp. 6</td>
</tr>
<tr>
<td>#V#</td>
<td>hVb</td>
<td>hVbut</td>
<td>Mean</td>
</tr>
<tr>
<td>/i/</td>
<td>97</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>/u/</td>
<td>99</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>/e/</td>
<td>100</td>
<td>100</td>
<td>95</td>
</tr>
<tr>
<td>/e/</td>
<td>99</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4.13. Correct identification by vowel and group for each experiment.

With respect to group differences, the first result is that native English speakers obtained extremely high percentages of vowel identification, often reaching 100%. These high identification scores are in accordance with studies involving vowel identification by native speakers such as Hillenbrand et al. (1995), whose identification scores for /i/, /u/, /e/ and /e/ were 99.6%, 98.8%, 98.3% and 95.1%, respectively. Of the two L2 groups, the BC group was considerably more successful in the vowel identification tasks than the
TC group (this difference reached significance for vowels in isolation and in CVCVC environments (Experiments 4 and 6)). This finding is surprising given that the TC group is made up of individuals who live in an English speaking country and who reported frequent use of English (which is expected given the small size of the Catalan community in Ontario). This indicates a greater effect of formal instruction over experience with the L2 on correct vowel identification. A possible explanation may lie in a facilitating effect of formal linguistic knowledge on the part of the BC group. Although BC subjects had presumably had less continuous exposure to native English input than the TC group, they were undergraduate students in English Philology with a fairly advanced level of English and had taken courses in English Phonetics and Phonology. Thus, it is plausible that their metalinguistic knowledge of the English vowel system may have enhanced their perceptual ability in these tasks even though they may or may not have internalized that knowledge to establish native-like vowel categories.

The perceptual assimilation task (Chapter 3) indicated that English /i/, /e/ and /e'/ were strongly assimilated to Catalan vowels. The remaining vowel /u/ was not as strongly assimilated to an L1 vowel. The prediction was that Catalan learners would not establish new phonetic categories for the highly assimilated vowels and they would perceive and produce them in terms of the Catalan vowel categories. In the case of the weakly assimilated vowel, successful acquisition would depend on the learners’ ability to create a new target-like category.

The results for /e/ and /e'/ support the hypothesis that strongly assimilated vowels will be categorized and perceived in terms of the L1 vowels and that accuracy of perception depends on target-like use of phonetic cues. The vowel /e/ was the most
consistently and successfully identified by both L2 groups, reinforcing the identity results obtained in the perceptual assimilation tasks in Chapter 3. With respect to /eɪ/, Catalan learners of English prefer clear diphthong-like formant transitions, in accordance with its categorization as the Catalan diphthong /ei/. Native English speakers do not rely on the offglide as much, as shown by the high identification scores regardless of talker differences. Thus, /eɪ/ represents the case of a strongly assimilated vowel for which the L1 category is not based on exactly the same cue weights as the assimilated L2 vowel. This results in less accurate perception, as predicted. An interesting effect of L2 input emerges with respect to /eɪ/. Although scores for the TC group were overall lower than those for the BC group, the TC group was more successful in the case of more monophthongal /eɪ/ tokens. This appears to reflect a positive effect of exposure to this particular English variety with a certain variation in the diphthongal quality of this vowel.

The effect of talker differences in duration on the identification of /i/ was not expected given that the strongly assimilated /i/ was predicted to be categorized in terms of the same features as in the L1, that is, vowel quality. This effect of duration in the categorization of /i/ will be discussed in more detail in the discussion of Experiment 7 (section 4.2). The results for the weakly assimilated vowel /i/ do not differ from the results for strongly assimilated vowels. This does not support the predictions of the Speech Learning Model (Flege, 1995) that the degree of perceived dissimilarity determines the likelihood of new category formation and that experience has an effect on the categorization of the more dissimilar but not of the more similar vowels (the latter is also predicted by the Perceptual Assimilation Model (Best and Strange, 1992)). The
models would predict better results for /i/ than for the other three vowels. In fact, the results show that the TC group, the group with more experience with the L2, obtained comparatively lower identification scores for the weakly assimilated vowel.

Finally, comparing across experiments, the bisyllabic environment emerges as the most problematic for the Catalan subjects, with lower overall results. This can be attributed to the shorter duration of stressed vowels in the bisyllabic stimuli. Per vowel, /e/ is best perceived in the vowel in isolation environment, which is explained by the longer duration and greater formant movement of this vowel in open syllables. Vowel /i/ displays a positive effect of the monosyllabic environment, but this is only found with the TC group, for whom the identification scores for /i/ were overall lower. The current results do not support the observation in some vowel perception studies that vowels are better perceived in CVC sequences than in isolation (Gottfried and Strange, 1980; Strange and Gottfried, 1980). This may be related to the finding that the beneficial effect of context is greater with low vowels than with high vowels (Rankerd, Verbrugge and Shankweiler, 1984). In addition, some experiments have shown that vowels in isolation may be readily identified by listeners, and even gated vowel centres may be well identified, as shown in Assmann, Nearey and Hogan (1982). Moreover, in the current experiment, stimuli for the vowels in isolation environment were edited out leaving intact cues to vowel identity insofar as possible so that vowel identity was as detectable as in CVC words.

The results so far are neutral regarding the prediction that the lax-tense vowel distinction is more readily maintained in CVC than in CVCVC environment due to the interaction of the Lax Vowel Constraint and L1 syllabification of medial consonants as
onsets. The identification scores for lax vowels do tend to drop in the bisyllabic environment. However, as we have seen, a number of factors seem to influence perception of the target vowels, including vowel stimuli differences in vowel duration. Thus, the drop in correct identification scores in the bisyllabic environment could be due to the shorter duration of the vowel stimuli and the concomitant increase in discrimination difficulty. The effect of syllabification and the LVC on the production of the lax vowel /ɒ/ will be investigated in greater detail by examining Catalan learners’ production of the target vowels in Chapter 5. But first we will explore further the effect of different cues in vowel perception.

4.2. Experiment 7. Vowel identification involving a synthetic vowel continuum

Previously we have seen that Catalan learners of English and native English speakers differ in their reliance on acoustic-phonetic cues for vowel perception, such as vowel duration. Catalan learners tend to be more successful at identifying longer tense vowels and shorter lax vowels. The effect of the relative weighting of spectral and temporal cues is the object of study of this last perception experiment.

Recall from Chapter 2 that English tense and lax vowels differ in both spectral and temporal cues and that native English speakers tend to rely mostly on spectral cues. In addition, as discussed in Chapter 1, learners of English have been found to note temporal differences more readily than spectral differences, even if duration is not part of their L1 (Bohn, 1995; Bohn and Flege, 1992). This is in opposition to the view in McAllister, Flege and Piske (1999) and Flege (1995) that an L2 contrastive category will be difficult to acquire if it is based on a phonetic feature not exploited in the L1. This
study follows Bohn’s proposal that learners will be more sensitive to temporal differences when the vowel contrast in the L2 is distinguished by both spectral and temporal properties and the spectral contrast has no counterpart opposition in the L1. Thus, the prediction is that Catalans may exploit durational differences in the distinction of the English /i/-/ɪ/ contrast because the two vowels are not equally matched to L1 categories (as determined by the degree of perceived similarity).

Importantly, an issue that needs to be addressed is the effect of different L2 input on the salience of the different cues. Studies of native English speakers are based on American English varieties and it has been claimed that British English and American English may differ precisely in the role of quantity and quality in their vowel systems (Lindsey, 1990). Lindsey (1990) provides evidence from the treatment of modern loans to support the claim that whereas standard British English retains an underlying long/short distinction, American dialects implement a distinction that is based mainly on vowel quality. Accordingly, if the implementation of the lax-tense distinction varies from one dialect to another, the source of the L2 input of the second language learners becomes a crucial issue.

Escudero (2001) tested the effect of exposure to different varieties of English on the acquisition of the English lax-tense vowel contrast. Based on the fact that in the Scottish dialect of English /i/ and /ɪ/ differ almost exclusively in spectral cues, Escudero tested the relative weighting of temporal and spectral cues in native speakers of Scottish English, Southern England English and Spanish learners of English. Results indicated that Spanish learners of English varied in their reliance on durational and spectral cues and this variation was explained as a function of predominance of exposure to one dialect.
or the other. Nevertheless, though her findings for native speakers yielded a greater reliance on spectral cues for the Scottish English speakers than the Southern English speakers, the latter were still found to rely as much or more frequently on spectral cues than duration cues. This was unexpected since Southern English speakers were expected to rely mostly on duration. The result was explained in terms of indirect influence from Scottish English because the Southern English speakers were residing in Scotland.

The effect of the target language input is examined here by testing two different L2 learner groups, the Catalans in Canada, exposed to a variety expected to rely mostly on spectral differences, as found in previous research (Wang and Munro, 1999), and the Catalans in Catalonia, exposed to different varieties of English, including British English.

Perceptual ability was tested using a synthetic /i/-/ɪ/-/ɛ/ vowel continuum varying in temporal and spectral steps. Vowel /ɛ/ was not included in this test for two reasons. First, this vowel is strongly assimilated to an L1 diphthong rather than a monophthong so that the main cue to its identification, a change in quality, is probably different from the main cue for the other L2 vowels. The experiments using natural stimuli have shown in fact that Catalan speakers identify this vowel mostly on the basis of its gliding nature. In addition, the lax mid front vowel counterpart, vowel /ɛ/, obtained the highest overall identification scores. This shows that the /ɛ/-/ɛ/ vowel pair is distinguished mostly in terms of vowel quality differences, as expected from strongly assimilated vowels. The relative importance of acoustic cues is more crucial with the high vowels given that /i/ is a weakly assimilated vowel. In addition, vowel duration differences affected identification of /i/ in the previous experiments. Secondly, the acoustic characteristics of
/e/ complicate its inclusion in the continuum; there is a considerable overlap between /e/’s steady state and vowel /u/. Furthermore, /e/’s formant transitions are greater and in the opposite direction (F1 drops and F2 rises) from its neighbouring vowels /u/ and /e/, which have a more centralized off-glide (F1 rises and F2 drops slightly). Thus, relative weighting of acoustic cues was tested on the /i/-/u/ pair with the inclusion of /e/ for which vowel quality alone is expected to be the main cue.

4.2.1. Subjects

The same subjects who participated in the perception experiments involving natural stimuli participated in this experiment. In addition, two more Catalan speakers living in Toronto and seven more native Southern Ontario English speakers, undergraduate students at the University of Toronto, participated in the experiment. Thus, the total number of subjects was 30 Catalan learners of English tested in Barcelona (the BC group), 20 Catalan residents in Ontario tested in Toronto (the TC group), and 20 native English speakers (EN group).

4.2.2. Stimuli

The two-dimensional /i/-/u/-/e/ English continuum consisted of 11 vowel quality steps and 4 temporal steps, giving a total of 44 different vowels. With respect to the 11 quality steps, vowels 1, 6 and 11 corresponded to the prototypical spectral values for English /i/, /u/ and /e/, respectively, based on Peterson and Barney’s (1952) values for male voices, given in Table 4.14. Intermediate vowels were calculated in linear steps. In
order to render the stimuli as authentic as possible, vowels included offglide transitions, which were slightly centralized in the case of /ɪ/ and /ɛ/, and slightly rising for /i/, following acoustic descriptions (Klatt, 1980). The four different durations were 100 ms, 150 ms, 200 ms and 250 ms. The synthetic continuum was created following Klatt’s (1980) parameters for vowels in isolation, and using Computerized Speech Lab (CSL/ASL) software.

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>270</td>
<td>2290</td>
<td>3010</td>
</tr>
<tr>
<td>/ɪ/</td>
<td>390</td>
<td>1990</td>
<td>2550</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>530</td>
<td>1840</td>
<td>2480</td>
</tr>
</tbody>
</table>

Table 4.14. Formant values for English high and mid front vowels.

### 4.2.3. Procedure

The task was a three-alternative forced-choice task in which each response alternative consisted of English words written in English orthography and representing one of the three target English vowels (/i/, /ɪ/, /ɛ/), namely ‘beat’, ‘bit’, and ‘bet’. Stimuli were drawn from the /i/-/ɪ/-/ɛ/ spectral and temporal continuum described above. The task was performed using an IBM laptop computer and good quality headphones (Sennheiser Headmax HD 400). A trial consisted of two presentations of each stimulus with an inter-stimulus interval of two seconds. After hearing a given stimulus, subjects selected a response from the alternatives presented on the screen. Subjects prompted the program for the following trial by clicking on the instructions on the screen. Each stimulus appeared five times (five trials). The task was preceded by a practice period
including 12 trials to familiarize the subjects with the procedure and adjust the listening level. In addition, prior to the practice session, there was a training session of 12 trials.

4.2.4. Results

The results are illustrated in Figures 4.1 to 4.9, which provide the percentages of ‘beat’ (Figures 4.1-4.3), ‘bit’ (Figures 4.4-4.6) and ‘bet’ (Figures 4.7-4.9) responses for each vowel stimulus. The eleven vowel quality steps from /i/ to /u/ to /e/ are represented on the x-axis while the four durations are represented by the lines in the graph. The actual percentages are given in Tables 4.15 to 4.23. As shown in Figures 4.1, 4.4 and 4.7, the EN group displayed clear crossovers from /i/ to /u/ and from /u/ to /e/, indicating a consistent pattern of vowel categorization. Thus, for example, Figure 4.1 shows the percentages of ‘beat’ responses for each vowel in the continuum obtained by EN. We can see that the number of ‘beat’ responses for Vowel 1 (prototypical vowel /i/) was very high (99-100%) for all four lengths (100ms, 150ms, 200ms and 250ms). Scores for Vowel 2 were also very high for all lengths (96-100%) and they drop quickly from Vowel 3 to Vowel 5 as the quality of the vowel becomes similar to prototypical /u/ (Vowel 6). Thus Vowels 1 and 2 are good tokens of the vowel in ‘beat’ (/i/) and as we move away from prototypical /u/ (Vowels 3 to 5), percentages of ‘beat’ identification drop dramatically. More importantly, length is not a factor in the identification of /i/ given that vowels that are spectrally close to prototypical vowel /i/ (Vowel 1 and 2) are identified as the vowel in ‘beat’ regardless of their length. Both the 100ms Vowel 1 and the 250ms Vowel 1 obtain 99% ‘beat’ responses. This allows us to conclude that the EN
group base their response choices on the quality of the vowel. The same pattern of responses is found with the middle part of the continuum, corresponding to vowel /ɪ/, and the lower end of the continuum, that is, vowel /ɛ/. The only difference was that the longest /ɪ/ tokens had lower percentages of correct identifications (i.e., ‘bit’ responses) indicating some effect of length, as discussed below.

The pattern found with the L2 groups is different, however. Although their identification scores for /ɛ/ closely resemble those of native speakers (Figures 4.8 and 4.9), L2 learners’ results display a noticeable effect of vowel duration for /ɪ/ and /ɪ/, whose percentages of correct identifications are not consistent across temporal differences. This is evident if we examine the percentages of ‘beat’ responses for the high end of the continuum (Figures 4.2 and 4.3). For both L2 groups, the number of ‘beat’ responses increases as a function of Length. That is, longer Vowel 1 and Vowel 2 tokens yield much higher ‘beat’ identifications than shorter tokens. Thus, with equal spectral characteristics, the longer the vowel, the more likely it is to be identified as ‘beat’. This effect of length is also present in the identification pattern of vowel /ɪ/ (Vowel 6 in the continuum, Figures 4.5 and 4.6). In this case, shorter tokens of this vowel obtain higher ‘bit’ responses than longer tokens. Thus, with equal spectral characteristics, the shorter the vowel, the more likely it is to be identified as the vowel in ‘bit’. The responses for vowel /ɛ/, at the low end of the continuum, show that in this case length was not a factor and in fact L2 learners’ responses were very similar to the native speakers’ (Figures 4.7 to 4.9).
Figure 4.1. Percentages of ‘beat’ responses for native English speakers.

<table>
<thead>
<tr>
<th>% beat</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>99</td>
<td>96</td>
<td>71</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>150 ms</td>
<td>99</td>
<td>97</td>
<td>86</td>
<td>22</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200 ms</td>
<td>100</td>
<td>99</td>
<td>94</td>
<td>34</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>250 ms</td>
<td>99</td>
<td>100</td>
<td>97</td>
<td>51</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.15. Percentages of ‘beat’ responses for native English speakers.

Figure 4.2. Percentages of ‘beat’ responses for Catalan subjects in Barcelona.
### Table 4.16. Percentages of ‘beat’ responses for Catalan subjects in Barcelona.

<table>
<thead>
<tr>
<th>% beat</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>55</td>
<td>50</td>
<td>38</td>
<td>21</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>150 ms</td>
<td>78</td>
<td>59</td>
<td>53</td>
<td>34</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200 ms</td>
<td>97</td>
<td>94</td>
<td>86</td>
<td>54</td>
<td>30</td>
<td>16</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>250 ms</td>
<td>99</td>
<td>97</td>
<td>89</td>
<td>73</td>
<td>43</td>
<td>23</td>
<td>19</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 4.3. Percentages of ‘beat’ responses for Catalan subjects in Toronto.

### Table 4.17. Percentages of ‘beat’ responses for Catalan subjects in Toronto.

<table>
<thead>
<tr>
<th>% beat</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>53</td>
<td>44</td>
<td>34</td>
<td>21</td>
<td>11</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>150 ms</td>
<td>70</td>
<td>58</td>
<td>52</td>
<td>32</td>
<td>9</td>
<td>14</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>200 ms</td>
<td>76</td>
<td>78</td>
<td>72</td>
<td>54</td>
<td>27</td>
<td>25</td>
<td>11</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>250 ms</td>
<td>88</td>
<td>88</td>
<td>82</td>
<td>74</td>
<td>40</td>
<td>17</td>
<td>20</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4.17. Percentages of ‘beat’ responses for Catalan subjects in Toronto.
Figure 4.4. Percentage of ‘bit’ responses for native English speakers.

<table>
<thead>
<tr>
<th>Time (ms)</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>1</td>
<td>4</td>
<td>29</td>
<td>85</td>
<td>97</td>
<td>100</td>
<td>95</td>
<td>70</td>
<td>12</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>150 ms</td>
<td>1</td>
<td>3</td>
<td>14</td>
<td>76</td>
<td>97</td>
<td>97</td>
<td>84</td>
<td>41</td>
<td>15</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>200 ms</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>63</td>
<td>86</td>
<td>88</td>
<td>63</td>
<td>24</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>250 ms</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>46</td>
<td>84</td>
<td>74</td>
<td>50</td>
<td>16</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.18. Percentage of ‘bit’ responses for native English speakers.

Figure 4.5. Percentage of ‘bit’ responses for Catalan subjects in Barcelona.
Table 4.19. Percentages of ‘bit’ responses for Catalan subjects in Barcelona.

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>45</td>
<td>50</td>
<td>62</td>
<td>79</td>
<td>90</td>
<td>90</td>
<td>88</td>
<td>85</td>
<td>43</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>150 ms</td>
<td>22</td>
<td>41</td>
<td>47</td>
<td>63</td>
<td>83</td>
<td>87</td>
<td>77</td>
<td>65</td>
<td>41</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>200 ms</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>42</td>
<td>49</td>
<td>54</td>
<td>51</td>
<td>47</td>
<td>13</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>250 ms</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>19</td>
<td>38</td>
<td>52</td>
<td>45</td>
<td>25</td>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 4.6. Percentage of ‘bit’ responses for Catalan subjects in Toronto.

Table 4.20. Percentages of ‘bit’ responses for Catalan subjects in Toronto.

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>46</td>
<td>56</td>
<td>66</td>
<td>77</td>
<td>81</td>
<td>82</td>
<td>75</td>
<td>46</td>
<td>21</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>150 ms</td>
<td>29</td>
<td>41</td>
<td>48</td>
<td>63</td>
<td>77</td>
<td>79</td>
<td>65</td>
<td>35</td>
<td>23</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>200 ms</td>
<td>23</td>
<td>22</td>
<td>27</td>
<td>41</td>
<td>50</td>
<td>53</td>
<td>42</td>
<td>23</td>
<td>6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>250 ms</td>
<td>11</td>
<td>11</td>
<td>17</td>
<td>18</td>
<td>36</td>
<td>49</td>
<td>25</td>
<td>12</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 4.7. Percentages of ‘bet’ responses for English native speakers.

<table>
<thead>
<tr>
<th>% bet</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
<th>V11</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ms</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>30</td>
<td>88</td>
<td>98</td>
<td>100</td>
</tr>
<tr>
<td>150 ms</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>16</td>
<td>59</td>
<td>85</td>
<td>98</td>
<td>99</td>
</tr>
<tr>
<td>200 ms</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>37</td>
<td>76</td>
<td>98</td>
<td>99</td>
<td>98</td>
</tr>
<tr>
<td>250 ms</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>26</td>
<td>50</td>
<td>84</td>
<td>97</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4.21. Percentages of ‘bet’ responses for English native speakers

Figure 4.8. Percentages of ‘bet’ responses for Catalan subjects in Barcelona.
A statistical analysis was conducted on the percentage of ‘bet’ responses obtained for the high end of the continuum (Vowel 1, prototypical /i/), the percentage of
‘bit’ responses for the prototypical vowel /ʌ/ in the middle the continuum (Vowel 6) and of ‘bet’ responses for the low end of the continuum (Vowel 11, prototypical vowel /ɛ/), in other words, on the percentage correct responses for each prototypical vowel. A three way analysis of variance was performed with Language Group as the between groups factor (EN group, TC group and BC group), and Length (four durations) and Vowel (/i/, /u/ and /ɛ/) as the within groups factors. All main effects and interactions proved significant. Language Group was significant, as is evident in the higher identification scores obtained by the EN group (F(2,67) = 14.52, p < .001). A Tukey HSD post-hoc test showed that the two Catalan groups did not differ significantly, but the EN group differed significantly from both Catalan groups (p < .001). The overall higher identification scores for /ɛ/ across groups explains the significance of Vowel (F(4,134) = 44.74, p < .001), and the effect of Length on the number of responses for Vowels 1 and 6 accounts for its statistical significance (F(3,201) = 4.97, p < .01). The interactions with Language are clear in that neither Vowel nor Length appeared to have an effect for the EN group, whereas they did for the TC and BC group (Language x Vowel: F(2,134) = 5.01, p < .001; Language x Length: F(6,201) = 3.1, p < .01). The two way interaction Length x Vowel (F(6,402) = 34.45, p < .001) is explained by the absence of the Length effect with vowel /ɛ/. Finally, the three way interaction (F(12,402) = 3.95, p < .001) is due to the fact that Length had an effect for only two of the three vowels and for only two of the three groups.

The post hoc analyses also confirmed that the two Catalan groups differ from the English group in the correct identification scores for vowel /i/ and /u/ (p < .001) but not
for /ε/. As is clear from Figures 4.8 and 4.9, Catalan subjects’ results for /ε/ closely resemble those for English native speakers. The BC group’s scores for /ε/ were significantly higher than those for /i/ (p<.01) and /u/ (p<.001). Scores for /i/ were significantly higher than those for /u/ (p<.05). In the case of the TC group, scores for /ε/ were significantly higher than those for the other two vowels (p<.001 in both cases).

Length affected vowels /i/ and /u/, but not /ε/. The percentages of correct responses for the prototypical vowels in the four different durations are repeated in Table 4.24 below. Catalan learners patterned like English listeners in their reliance on spectral cues in the identification of /ε/, as illustrated by their consistently high percentages of ‘bet’ responses for Vowel 11 (prototypical /ε/) across all four lengths. A Tukey HSD post hoc test indicated that Length was significant for the two non-native groups with respect to vowels /i/ and /u/ and it also reached significance for EN with respect to vowel /u/. In the case of /i/, the effect of Length is clear, as discussed above and shown in Table 4.24: 250 ms Vowel 1 (prototypical vowel /i/) was identified as the vowel in ‘beat’ 88%/99% of the time (for TC and BC groups, respectively) and as the vowel in ‘bit’ only 11%/1% of the time. On the other hand, 100 ms-/i/ yielded only 53%/55% ‘beat’ responses and reached 46%/45% /u/ responses. Length is thus a clear factor for Catalan speakers in the identification of /i/, but not for native speakers, whose scores do not vary as a function of length (99-100% correct across lengths).
<table>
<thead>
<tr>
<th>Correct Identification</th>
<th>Group</th>
<th>100ms</th>
<th>150ms</th>
<th>200ms</th>
<th>250ms</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prototypical /i/ (Vowel 1)</td>
<td>BC</td>
<td>55</td>
<td>78</td>
<td>97</td>
<td>99</td>
<td>83</td>
</tr>
<tr>
<td>% ‘beat’ responses</td>
<td>TC</td>
<td>53</td>
<td>70</td>
<td>76</td>
<td>88</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>99</td>
<td>99</td>
<td>100</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>Prototypical /a/ (Vowel 6)</td>
<td>BC</td>
<td>90</td>
<td>87</td>
<td>54</td>
<td>52</td>
<td>71</td>
</tr>
<tr>
<td>% ‘bit’ responses</td>
<td>TC</td>
<td>82</td>
<td>79</td>
<td>53</td>
<td>49</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>100</td>
<td>97</td>
<td>88</td>
<td>74</td>
<td>90</td>
</tr>
<tr>
<td>Prototypical /e/ (Vowel 11)</td>
<td>BC</td>
<td>94</td>
<td>95</td>
<td>97</td>
<td>99</td>
<td>96</td>
</tr>
<tr>
<td>% ‘bet’ responses</td>
<td>TC</td>
<td>95</td>
<td>93</td>
<td>98</td>
<td>96</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>100</td>
<td>99</td>
<td>98</td>
<td>100</td>
<td>99</td>
</tr>
</tbody>
</table>

Table 4.24. Correct identifications of each prototypical vowel in each length condition.

As for vowel /i/ and the number of ‘bit’ responses, it is noteworthy that length had an effect for native English speakers too. This is evident in the fact that whereas identification rates for the shortest /i/ (100 ms Vowel 6) tokens were close to 100%, identification dropped to 74% in the case of 250 ms-long /i/ (the remaining 26% involved ‘bet’ responses). Native speakers’ identification scores of 250ms /i/ differed significantly from their identification scores at 100ms (p < .01) and marginally from 150ms /i/ (p < .07). This means that very long /i/ tokens sound less natural to native speakers than average or shorter /i/ tokens. Nevertheless, Catalan speakers still pattern quite differently from native English speakers since Catalans’ identification rates are much lower and the effect of length is greater, with only 49% and 52% ‘bit’ responses for 250 ms-long /i/ for the TC and BC groups respectively as opposed to 82%/90% with 100 ms /i/. Longer /i/
tokens were identified as the vowel in ‘beat’ around 17-29% of the time and as the vowel in ‘bet’ another 17-34% of the time. The Language x Vowel x Length interaction is due to the fact that Length was significant for all three groups with respect to vowel /i/, but only for the Catalan groups with respect to /i/ and never reached significance in the case of /e/.

4.2.5. Conclusions

The results of Experiment 7 show that native English speakers, at least native speakers of Southern Ontario English, rely mostly on spectral cues in their identification of the English vowels /i/, /u/ and /e/. This finding is in accordance with studies on acoustic cue reliance involving North American English speakers. Nevertheless, a certain effect of Length for the native speaker group is visible in the drop in identification scores for the longest /u/ tokens.

These results resemble the results from Experiment 4 (vowels in isolation) in that the BC group obtained higher rates of identification than the TC group with respect to /i/ and also in that /e/ was the least problematic vowel. In fact, if we compare the two tasks we can see that identification scores for long synthetic /i/ were close to the scores obtained in Experiment 4 (88% vs. 85% respectively for the TC group and 99% vs. 96% for the BC group). Likewise, scores for /u/ follow a similar pattern: 75-82% for short /u/ vs. 67% in the current task for the TC group, and 87-90% vs. 89% respectively for the
BC group. It seems then that vowel /u/, and vowel /i/ for the TC group, are more difficult to identify than /e/ both with synthetic stimuli and with natural stimuli.

The two L2 groups differ from the native English speakers in their reliance on durational cues to distinguish between English /i/ and /u/. Catalan subjects do in fact make use of spectral cues, as illustrated by the decrease in ‘beat’ responses and the increase in ‘bit’ responses as we progress from Vowel 1 to Vowel 6, but they still differ from native English speakers in their strong reliance on vowel duration. Thus, this finding supports the prediction that Catalan speakers will exploit temporal differences in the distinction of English /i/ and /u/. This is in accordance with the view in Bohn (1995) and Flege and Bohn (1992) that durational differences can be salient for L2 learners even if not part of the L1. In contrast, it argues against McAllister, Flege and Piske (1999) and Flege (1995) who do not predict the availability of a non-L1 feature to adult learners.

In the case of /u/, the effect of duration is in accordance with the prediction for a weakly assimilated vowel. Non-L1 features were expected to play a greater role in the categorization of weakly assimilated vowels than strongly assimilated vowels. Strongly assimilated vowels were predicted to be categorized in terms of L1 contrastive properties, mainly vowel quality in this study. This is the case of the vowel /e/, whose identification was not affected by duration. As in the previous experiments, the fact that two strongly assimilated vowels, /i/ and /e/, obtain different results, and that a strongly assimilated vowel (/i/) and a weakly assimilated vowel (/u/) obtain similar results does not support the predictions of the Speech Learning Model (Flege, 1995) or the Perceptual Assimilation Model (Best and Strange, 1992) that accurate categorization varies as a function of
perceived dissimilarity. In addition, the facilitating role of experience predicted by these two models is not supported by the current data either. More exposure to the native language in the case of the TC group, made up of Catalan learners residing in an English speaking environment, does not translate into better identification or a more native-like weighting of cues. Finally, the results for /i/ show that Catalan learners do not simply rely on quality but also on duration. This effect is consistent with the finding in the previous experiments that talker differences in duration affect the perception of /i/. The fact that the weakly assimilated vowel /u/ is categorized in terms of its temporal opposition to /i/ may affect the categorization of the latter and may override its strong assimilation to an L1 vowel, in turn exploiting non-L1 features to define the vowel category. This was predicted by the hypothesis that even those L2 categories that are strongly assimilated may be affected by the category formation of neighbouring vowels. This issue is discussed in more detail after evaluating the production data in the next chapter.

The use of temporal cues has also been found with Mandarin learners of English, both in perception and in production (Wang, 1997; Wang and Munro, 1999). In this case, researchers found a different pattern for different L2 vowel pairs. Wang examined the production of vowel pairs by Mandarin speakers and found that they greatly exaggerated the durational differences in the production of the high front but not the high back lax-tense vowel pair. The difference found in the use of duration for front and back vowels in Wang and Munro (1999) was explained as a result of the higher functional load of the front pair and the result of ESL instruction. This raises the issue of whether sensitivity to duration is primarily due to crosslinguistic perception patterns or if it is primarily the result of pedagogical approaches to the teaching of the contrast. The current study shows
that the setting of acquisition does not seem to have an effect on the acquisition of these vowels and the relative weighting of the acoustic cues. The TC group, despite greater exposure and experience with the L2, did not obtain better results than the BC group. Further, the effect of duration is independent of the L2 input given that both the TC group, whose L2 input is Canadian English, and the BC group, with a mixed input and more presence of British English, pattern exactly the same way with respect to reliance on acoustic cues. Thus, as exemplified by the TC group, categorization of the /i/-/u/ contrast in terms of a predominantly duration distinction takes place in a naturalistic setting, even when duration is not the predominant cue in the native speech.

Segmental representations are typically investigated using measures of both perception and production. Thus, the results will be compared with the results from production experiments in the following chapter.
Chapter Five. Production of L2 vowels

5.0. Introduction

In Chapter 4, the acquisition of the L2 vowels was examined by evaluating the learners’ ability to distinguish and differentiate the target L2 vowels perceptually. This chapter looks at acquisition from the perspective of their ability to produce the L2 vowels. Let us review the main hypotheses examined in this study which are relevant to the production of L2 vowels. First, strongly assimilated L2 vowels will be produced as their L1 counterparts and no new categories will be created for them. The accuracy of production will depend on whether the target vowels are based on the same cues and cue weighting as the L1 vowels they assimilate to. Non-L1 features are predicted to be more relevant in the categorization of weakly assimilated vowels. An example of this is the prediction that learners will exploit temporal cues in distinguishing a lax-tense vowel pair for which no spectral contrast is available in the L1. Finally, new category formation is governed by the need to establish and maintain vowel oppositions. As a result, the creation of a new vowel category may affect neighbouring vowels, regardless of their degree of L1-L2 assimilation.

The results from the perception experiments have provided some initial support for these hypotheses. First, the vowels /ɛ/ and /ɛʲ/ illustrate the case of strongly assimilated vowels that differ in the accuracy of their phonetic representation in terms of an L1 vowel. Whereas the L1 /ɛ/ provides an accurate representation for the L2 vowel, the categorization of /ɛʲ/ in terms of the Catalan diphthong /ei/ evidences a greater importance of formant movement for Catalan speakers than for English speakers. This
difference in cue weighting results in a failure to perceive /e/ in a native-like manner. In addition, unlike native speakers, Catalans tend to rely on temporal cues in the identification of the /i/-/i/ vowel pair. Sensitivity to non-L1 features is expected for the weakly assimilated vowel /i/. The fact that /i/ is categorized in terms of its temporal opposition to /i/ explains the role of durational cues in the identification of the latter. These findings will be tested against the production data in order to further ascertain the validity of the hypotheses.

The second important issue addressed in this chapter concerns the hypothesized relationship between segmental and non-segmental structure. As discussed in Chapter 1, Cebrian (1996, 1999) found that Catalan learners of English were more successful at producing the /i/-/i/ contrast in monosyllables than in bisyllables with a single intervocalic consonant. This was explained by the fact that the Catalan speakers had some knowledge of the Lax Vowel Constraint but, due to the strong influence from L1 syllabification, they treated the first syllable of such bisyllables as an open syllable. As a consequence, the lax vowel was mispronounced as its tense counterpart (/i/). This chapter examines the possible effect of environment on production by eliciting the target vowels in monosyllabic and bisyllabic environments.

Three experiments are reported in this chapter. The first (Experiment 8) elicited Catalan speakers’ production of the target vowels in the three syllabic environments. The utterances were analyzed acoustically to determine the accuracy of both spectral and temporal characteristics. Production ability was then assessed perceptually by having native speakers identify the vowels (Experiment 9) and a second set of native speakers provided goodness ratings (Experiment 10).
5.1. Experiment 8. Production of L2 vowels

5.1.1. Subjects

Subjects were the same 30 native speakers of Catalan (learners of English as a second language) who had participated in the perception experiments (the BC group), thus making the perception and production data as comparable as possible. In addition, four native speakers of Southern Ontario English, two male and two female, also participated in the experiment for control purposes.

5.1.2. Procedure

The production of tense and lax high and mid English vowels was tested in monosyllabic and bisyllabic environments to examine a possible effect of L1-based syllabification in the production of lax vowels. Vowels were elicited in three environments, two CVC words (one word and one nonword), /h_d/ and /h_b/ (e.g., ‘heed’ (/hid/), ‘heeb’ (/hib/)) and a CVCVC nonword with the target vowel in the first syllable followed by an unstressed syllable with a reduced vowel (e.g., ‘heepus’ (/hipɔs/).

Bisyllables were all of the same ‘h_pus’ shape. The method of elicitation was a mixture of repetition and vowel insertion. Subjects first repeated a h_d word and then produced two new sets of words by inserting the vowel in the h_d word into h_b and h_pus environments. For example, given the test word ‘heed’, the subjects produced three repetitions of ‘heed’, ‘heeb’ and ‘heepus’. One motivation for this methodology was that it allowed us to elicit the target vowels in the same environments as they appeared in the

---

1 Thirty native Catalan speakers participated in the experiment but due to technical problems the data for one of the speakers was lost. Thus, production results will be based on the data elicited from remaining 29 native Catalan speakers.
perception tests (e.g., between /h/ and /b/), thus facilitating the comparison between the two experiment sets. Secondly, this repetition and insertion technique was chosen so as to minimize direct orthographic interference. In addition, the use of non-words helped avoid contamination from effects of word frequency. Finally, as in the case of the stimulus preparation for the perception tasks, this consonant environment was chosen in order to minimize C to V and V to C tongue coarticulation (Strange et al 1998), permitting a clearer analysis and extraction of the vowel portion for the intelligibility assessment of L2 production.

The test words had been elicited from a male native North American English speaker and were recorded using a SONY digital-audio tape recorder. The words were subsequently digitized at a 10kHz sampling rate and normalized for peak intensity. The prompt words were presented aurally via good quality headphones (Sennheiser Headmax HD 400), played from an IBM laptop computer. Subjects were also provided with a card containing the list of h_d test words if they wished to check that they had understood the word they had heard. In addition, the subjects were also given a card containing the structure of the nonsense words, i.e. h__b and h__pus so as to facilitate the vowel insertion task. The experiment was preceded by a practice session with tokens including different vowels so as to familiarize the subjects with the experimental procedure. In addition, subjects were told that they could repeat their responses as many times as they wanted until the nonsense words sounded as natural or wordlike as possible. The list of

---

2 The /b/ was replaced with a /p/ in the two syllable word in the current experiment because of the difficulty on the part of Catalan native speakers to produce a plosive /b/ in intervocalic position due to highly transferable L1 spirantization process (cf. Cebrian 2000). So as to avoid a fricative consonant in this context, which would have complicated the analysis of the vowel portion and its subsequent extraction for the vowel identification task, a /p/ was chosen as the intervocalic consonant.
test words included ‘hide’, ‘head’, ‘hid’, ‘hayed’, ‘heed’, and ‘had’. Recordings took place in a soundproof booth in the phonetics laboratory at the Philology Department of the Universitat Autònoma de Barcelona.

Subjects’ responses were digitized at a 10kHz sampling rate, low-pass filtered at 4.8 kHz with 16-bit resolution, and saved as audio files for both acoustic analysis and to provide the stimuli for the intelligibility and goodness rating tasks. The digitized files were examined for signal quality. Each subject’s best two tokens per environment and vowel were selected for analysis, both on the basis of auditory judgment and acoustic signal examination. This yielded six different productions of each vowel per subject. The use of multiple stimuli in vowel perception experiments is a common strategy to ensure that the results can be generalized beyond the particular characteristics of a specific token (Flege, Munro and Fox, 1994). Given that L2 production in the current study is assessed by native speaker identifications of L2 vowels, the use of multiple tokens seemed appropriate.

5. 2. Acoustic analysis

An acoustic analysis of the production data was conducted by examining the formant frequencies and duration of each vowel. In addition to allowing a description and evaluation of the production data, this analysis also aimed to explore the relative weighting of temporal and spectral cues. Recall that the Catalan learners relied more on duration than on vowel quality in distinguishing between /i/ and /u/, as illustrated by the results in the vowel identification task involving a synthetic continuum. An acoustic analysis of the production data will reveal if this also holds for L2 production.
5.2.1. Vowel quality

F1 and F2 frequencies were estimated by placing a 25ms hamming window at the temporal midpoint of the vowel, except for vowel /e/ in which case frequency values were estimated closer to the onset of the vowel during the steady state. The mean F1 and F2 frequencies for the 25 female speakers and the four male speakers are given in Tables 5.1 and 5.2, respectively. The tables also include the means for the native speakers in the current study, as well as the reported values given in Chapter 2 from Deterding (1997), Peterson and Barney (1952) and Hillenbrand et al. (1995) for comparison purposes. The former provides the values for British English speakers and the latter two describe American English. The data from all three studies were obtained from ‘h_d’ words and thus the values are listed on the row for that environment. There is a certain discrepancy between the values in the two American English studies, as discussed in Chapter 2, but in both studies the vowels were readily identified in listening tasks.

As mentioned above, there was an uneven distribution of male and female speakers in the current study given that only four out of thirty subjects were male. Gender is not a consideration in this study, and the group of L2 learners has been considered as a whole without within group differences. However, when it comes to formant frequency analysis, obvious differences in vocal tract dimensions require separate analyses for male and female speakers. Given the small number of male speakers in the current study and given that there is no clear consensus on a good speaker normalization procedure (Deterding, 1990; Adank, van Hout and Smits, 2001), I present the data for female and male speakers separately and focus on the female data for statistical analysis. F1 and F2
values are then presented in Tables 5.1 and 5.2 for female and male speakers, respectively.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Condition</th>
<th>L2 speakers</th>
<th>English sp.</th>
<th>PB 1952</th>
<th>Hill. et al. 1995</th>
<th>Deterding, 1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>h_d</td>
<td>343 2725</td>
<td>318 2819</td>
<td>310 2790</td>
<td>437 2761</td>
<td>319 2723</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(47) (99)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>354 2676</td>
<td>307 2743</td>
<td></td>
<td>(57) (138)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>359 2643</td>
<td>338 2673</td>
<td></td>
<td>(58) (90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/u/</td>
<td>h_d</td>
<td>474 2413</td>
<td>463 2487</td>
<td>430 2480</td>
<td>483 2365</td>
<td>432 2296</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(68) (169)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>462 2461</td>
<td>434 2439</td>
<td></td>
<td>(102) (184)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>438 2450</td>
<td>434 2325</td>
<td></td>
<td>(86) (178)</td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>h_d</td>
<td>537 2439</td>
<td>516 2618</td>
<td>536 2530</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(79) (142)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>535 2405</td>
<td>558 2534</td>
<td></td>
<td>(70) (209)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>534 2381</td>
<td>445 2635</td>
<td></td>
<td>(71) (224)</td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>h_d</td>
<td>677 2219</td>
<td>675 2351</td>
<td>610 2330</td>
<td>731 2058</td>
<td>645 2287</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(68) (225)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>703 2224</td>
<td>683 2290</td>
<td></td>
<td>(73) (152)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>705 2160</td>
<td>736 2104</td>
<td></td>
<td>(75) (123)</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1. F1 and F2 frequencies in Hertz for the female L2 and native speakers in this study, plus reported data. Values in parentheses indicate standard deviations.
Table 5.2. F1 and F2 frequencies in Hertz for the male L2 and native speakers in this study, plus reported data.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Condition</th>
<th>L2 speakers</th>
<th>English sp.</th>
<th>PB 1952</th>
<th>Hill. et al. 1995</th>
<th>Deterding, 1997</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F1</td>
<td>F2</td>
<td>F1</td>
<td>F2</td>
<td>F1</td>
</tr>
<tr>
<td>/i/</td>
<td>h_d</td>
<td>279</td>
<td>2353</td>
<td>300</td>
<td>2358</td>
<td>270</td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>283</td>
<td>2326</td>
<td>292</td>
<td>2338</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>281</td>
<td>2314</td>
<td>318</td>
<td>2343</td>
<td></td>
</tr>
<tr>
<td>/u/</td>
<td>h_d</td>
<td>439</td>
<td>1943</td>
<td>469</td>
<td>2032</td>
<td>390</td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>442</td>
<td>1902</td>
<td>473</td>
<td>2009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>433</td>
<td>1940</td>
<td>496</td>
<td>1904</td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>h_d</td>
<td>491</td>
<td>1925</td>
<td>509</td>
<td>2118</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>488</td>
<td>2008</td>
<td>504</td>
<td>2127</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>476</td>
<td>1959</td>
<td>487</td>
<td>2171</td>
<td></td>
</tr>
<tr>
<td>e/</td>
<td>h_d</td>
<td>576</td>
<td>1816</td>
<td>635</td>
<td>1924</td>
<td>530</td>
</tr>
<tr>
<td></td>
<td>h_b</td>
<td>570</td>
<td>1795</td>
<td>646</td>
<td>1808</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h_pus</td>
<td>585</td>
<td>1743</td>
<td>693</td>
<td>1651</td>
<td></td>
</tr>
</tbody>
</table>

Statistical analyses were conducted separately for the F1 and F2 values of the female L2 group with Vowel and Environment as factors. The two-way ANOVA of F1 values yielded a significant effect of Vowel (F(3,72) = 167.4, p < .001) and a significant Vowel x Environment interaction (F(6,144) = 4.6, p < .001), but no main effect of Environment. A Tukey HSD post-hoc test showed that all vowel comparisons were significant at the p < .001 level. This means that the Catalan subjects were successful at producing the height differences among vowels, because F1 is a spectral indication of vowel height. Environment as a whole had no effect, but it interacted significantly with Vowel. A Tukey HSD test showed that the source of this interaction was the fact that
environment significantly affected vowel /u/ but not the other vowels. In the case of the high lax vowel, /u/’s F1 values in h_pus were significantly lower than in h_d (p < .01). Though not significantly, h_pus F1 values for /u/ were also lower than in h_b. Lower F1 in this case means a vowel quality closer to vowel /i/. Thus, the target vowel /u/ was produced as a significantly higher vowel in the h_pus condition than in the other two conditions. There is then some evidence that Catalans tend to produce /u/ with a more /i/-like quality in a bisyllable than in a monosyllable. This is the hypothesized ‘tensing’ of /u/ due to the failure to syllabify the post-vocalic consonant with the lax vowel and thus leaving the vowel in an open syllable. Nevertheless, the Catalans’ lower F1 values in this environment condition fall within the range of the native speakers’ values as illustrated in Table 5.1 but in that case the F2-F1 distance is smaller.

The analysis of the F2 values also yielded a main effect of Vowel (F(3,72) = 77.6, p < .001), indicating that vowels are also produced with distinctly different F2 frequencies and thus that as a whole L2 learners do distinguish the four English vowels, at least in terms of steady state formant values. The only exception was the vowel pair /u/ and /e/, whose F2 values did not differ significantly according to a post hoc test. Nevertheless, as discussed below, these two vowels are clearly distinguished by other more salient cues such as the longer vowel duration and the high off-glide that characterizes /e/. The main effect of Environment reached significance (F(2,48) = 4.48, p < .05) and the Environment x Vowel interaction yielded marginal significance (F(6,144) = 1.9, p < .08). This pattern is due to the significantly lower F2’s in the
bisyllabic condition for vowels /i/ and /e/. This is probably the result of the faster speech rate in the two-syllable word.

The results indicate that the (female) L2 speakers produce vowels with spectral characteristics that fall within the range of native English values obtained both in the current experiment and in other studies. Thus, if we only consider the mean values for the L2 learners in this study, we could infer that the L2 vowels were very target-like. However, there was considerable variability in the L2 data (as illustrated by the standard deviations). In order to examine the extent of this variability, the F1 and F2 frequencies of the vowels produced by all (female) subjects per environment were plotted. The results for the h_d condition, the h_b condition and the h_pus condition are shown in Figures 5.1, 5.2 and 5.3 respectively. Each ellipse represents the acoustic space of the vowels /i/, /u/ and /e/. The symbol in the centre of the plot represents the mean value across subjects. Also plotted are the means for the native speakers in this study, represented by the vowel symbols preceded by the letter “n” (for native). The values for vowel /e/ have been excluded because the relevant cues are not static F1 and F2 values but more salient cues such as gliding and duration (as illustrated in the perception experiments; duration is discussed in more detail in section 5.2.2). In terms of steady state F1 and F2 values there was a considerable overlap between /e/ and /u/ as is evident in Tables 5.1 and 5.2 above.
Figure 5.1. F1 and F2 values for vowels /i/, /u/ and /e/ in the h_d condition. Ellipses represent the L2 data, symbols indicate the mean values. An ‘n’ preceding a symbol indicates the native speakers’ mean value for that vowel.

Figure 5.2. F1 and F2 values for vowels /i/, /u/ and /e/ in the h_b condition. Ellipses represent the L2 data, symbols indicate the mean values. An ‘n’ preceding a symbol indicates the native speakers’ mean value for that vowel.
Figure 5.3. F1 and F2 values for vowels /i/, /u/ and /e/ in the h_pus condition. Ellipses represent the L2 data, symbols indicate the mean values. An ‘n’ preceding a symbol indicates the native speakers’ mean value for that vowel.

According to the vowel plots, vowel /u/ displays the greatest variability, as determined by the size of the acoustic space it occupies. The graphs also illustrate a difference across environment conditions. This is especially clear in the relation between the high tense and the high lax vowel. The real word condition (h_d) displays the greatest differentiation between vowel spaces, in terms of less overlap and more equidistant mean values. There is more overlap between /i/ and /u/ in the h_b condition, with a large portion of the /i/ vowel space intersecting with the lax vowel area. The better results for h_d compared to h_b may be related to the effect of real vs. nonsense words or to a sensory memory effect given that the former was produced immediately after the aural prompt whereas the latter required the process of inserting the vowel into a new environment.
Finally, the overlap is almost absolute in the h_pus condition: the acoustic space occupied by /i/ is almost completely within the range of the variation observed for /u/. This means that a large number of /i/ productions were not distinct from /i/ productions, at least in terms of their F1 and F2 values. In addition, the means for /i/ and /u/ are closer than in the other conditions. It is noteworthy that the greatest confusion corresponds to the bisyllabic condition. This provides support for the hypothesized negative effect of this environment on the implementation of the lax-tense contrast for high front vowels. It remains to be seen if vowels could nonetheless be differentiated perceptually on account of other acoustic cues such as vowel length.

### 5.2.2. Vowel duration

The other acoustic parameter measured was vowel length. The duration of each vowel was calculated from a dual spectrogram and waveform display from the first to the last positive peak in the periodic portion of the signal, as indicated by an increase/decrease in overall amplitude and waveform complexity. Tables 5.3 and 5.4 provide the mean duration for each vowel by environment for the L2 learners and native speakers in the study. Duration values for male and female speakers were combined.

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/u/</th>
<th>/i/-/u/ ratio</th>
<th>/e/</th>
<th>/e/-/e/ ratio</th>
<th>means</th>
</tr>
</thead>
<tbody>
<tr>
<td>hVd</td>
<td>270 (62)</td>
<td>178 (24)</td>
<td>1.53</td>
<td>289 (42)</td>
<td>212 (26)</td>
<td>1.37</td>
</tr>
<tr>
<td>hVb</td>
<td>243 (64)</td>
<td>153 (33)</td>
<td>1.62</td>
<td>257 (45)</td>
<td>183 (36)</td>
<td>1.44</td>
</tr>
<tr>
<td>hVpus</td>
<td>136 (41)</td>
<td>99 (21)</td>
<td>1.39</td>
<td>157 (23)</td>
<td>109 (21)</td>
<td>1.47</td>
</tr>
<tr>
<td>Means</td>
<td>216</td>
<td>143</td>
<td>1.51</td>
<td>234</td>
<td>168</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Table 5.3. Mean vowel duration and tense/lax vowel duration ratios in the L2 data (values are given in ms).
A statistical analysis was conducted on the L2 data. The two-way ANOVA (Vowel x Environment) yielded a significant main effect of Vowel ($F(3,84) = 102.4, p < .001$) and Environment ($F(2,56) = 268, p < .001$) and significant interaction ($F(6,168) = 10.44, p < .001$). A Tukey HSD post-hoc test showed that all vowel comparisons were significant. Vowel /e/ was the longest, followed by /i/ and then by the lax vowel /ɛ/ and finally /u/. Although the main effect of Vowel is qualified by the interaction, it is relevant to point out that the pattern of vowel differences found in the L2 data clearly replicates the pattern found in native speech (e.g., /ɛ/ > /i/ > /ɛ/ > /u/), which is consistent with previous studies that show that low vowels are longer than high vowels and tense are longer than lax (e.g., Umeda, 1975; Crystal and House, 1988; Hillenbrand et al. 1995; Giegerich, 1992; Lindsey, 1990, among others). The overall duration of the vowels produced by the native speakers was shorter than the that of the L2 group. However, it is difficult to assess the significance of that difference due to the small sample of native speakers. Further, this difference may simply reflect a faster speech rate on the part of native speakers even in the production of isolated words.

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/ɪ/</th>
<th>/i/-/ɪ/ ratio</th>
<th>/e/</th>
<th>/ɛ/</th>
<th>/ɛ/-/ɛ/ ratio</th>
<th>means</th>
</tr>
</thead>
<tbody>
<tr>
<td>hVd</td>
<td>279</td>
<td>179</td>
<td>1.64</td>
<td>289</td>
<td>201</td>
<td>1.44</td>
<td>237</td>
</tr>
<tr>
<td>hVb</td>
<td>257</td>
<td>144</td>
<td>1.79</td>
<td>292</td>
<td>169</td>
<td>1.77</td>
<td>215</td>
</tr>
<tr>
<td>hVpus</td>
<td>82</td>
<td>61</td>
<td>1.35</td>
<td>103</td>
<td>73</td>
<td>1.43</td>
<td>80</td>
</tr>
<tr>
<td>Means</td>
<td>206</td>
<td>128</td>
<td>1.59</td>
<td>228</td>
<td>148</td>
<td>1.54</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4. Mean vowel duration and tense/lax vowel duration ratios for the native English speakers (values are given in ms).
The duration differences between vowels are further illustrated in Figures 5.4 and 5.5 below. The graphs also illustrate the main effect of Environment. Turning now to that effect, a Tukey HSD post hoc test showed that mean vowel durations in all three environment conditions differed significantly. Duration was greatly reduced in the bysillabic words for all vowels, a clear result of the shorter duration of the stressed vowel in the two-syllable words. Durations in the h_d condition were also longer than those in the h_b position for all vowels. This pattern was repeated in the native speaker data except for vowel /e\textsuperscript{i}/ in the two monosyllabic environments. The difference between the two monosyllabic conditions could be due to lexical status (real word vs. non-word) or the effect of the place of articulation of the final consonants. The Tukey HSD post hoc test was used to explore the significant Vowel x Environment interaction. The test showed that for all four vowels the comparisons between environments were significant (e.g., h_d vs. h_b, h_d vs. h_pus and h_b vs. h_pus for each vowel). Vowel comparisons were also significant, except for /i/ vs. /e\textsuperscript{i}/, in which case duration differences were not significant in the h_b condition and only marginally significant in the h_d and h_pus condition, hence the interaction.
Figure 5.4. Mean vowel durations per vowel and environment conditions for the Catalan learners of English (‘iy’ = /i/, ‘I’ = /ɪ/, ‘ey’ = /eɪ/, ‘E’ = /ɛ/).

Figure 5.5. Mean vowel durations per vowel and environment conditions for the native English speakers (‘iy’ = /i/, ‘I’ = /ɪ/, ‘ey’ = /eɪ/, ‘E’ = /ɛ/).

A useful index of length contrasts is obtained by calculating a ratio of tense to lax vowel duration. The tense-lax vowel duration ratios were 1.51 for /i/-/ɪ/ and 1.43 for /eɪ/-/ɛ/. The ratios for the native speakers were 1.59 for /i/-/ɪ/ and 1.54 for /eɪ/-/ɛ/. Table 5.5
below provides the duration ratios for native speakers obtained in a number of studies, including the results from the current study. Recall from Chapter 2 that according to some studies duration differences are greater in British English than in American English for some vowel pairs. The current study’s ratios for native speakers are greater than those reported in Munro (1993), 1.20 in both cases (averaged from b_t and b_d environments produced by 23 talkers), and Hillenbrand et al. (1.28 for /i/-/ɪ/ and 1.34 for /e/-/ɛ/ in h_d words), but very close to the ratios in Flege et al. (1994), namely 1.6 for /i/-/ɪ/ and 1.43 for /e/- /ɛ/ (from /CVto/ words spoken by five native English speakers) and Giegerich (1992), between 1.57 and 1.9 for /i/-/ɪ/. Duration differences between tense and lax vowels vary across studies and may be greater with British English speakers. Nevertheless, the comparison shows that the ratios obtained for the L2 learners in this study fall within the range of the ratios found with native speakers.

<table>
<thead>
<tr>
<th>/i/-/ɪ/ ratio</th>
<th>/e/-/ɛ/ ratio</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.30 – 1.35</td>
<td></td>
<td>Lindsey (1990), American English</td>
</tr>
<tr>
<td>1.50 – 1.85</td>
<td></td>
<td>Lindsey (1990), British English</td>
</tr>
<tr>
<td>1.57-1.94</td>
<td>1.43</td>
<td>Giegerich (1992), British English</td>
</tr>
<tr>
<td></td>
<td>1.60</td>
<td>Flege, Munro &amp; Fox (1994), American English</td>
</tr>
<tr>
<td>1.20</td>
<td>1.20</td>
<td>Munro (1993), American English</td>
</tr>
<tr>
<td>1.28</td>
<td>1.34</td>
<td>Hillenbrand et al. (1995), American English</td>
</tr>
<tr>
<td>1.59</td>
<td>1.54</td>
<td>Canadian English (current study)</td>
</tr>
<tr>
<td>1.51</td>
<td>1.43</td>
<td>Catalan learners of English (current study)</td>
</tr>
</tbody>
</table>

Table 5.5. Duration ratio /i/-/ɪ/ as reported in different studies.
The duration difference in the Catalans’ L2 English is remarkable given that temporal contrasts are not present in the Catalan vowel inventory. In fact, the lower vowel /ɛ/ is on average 24ms longer than the high vowel /i/ (Recasens, 1984). Thus, L2 learners are implementing an L2 based duration contrast. This finding replicates in production the results from the perception experiments where Catalan learners of English made predominant use of vowel length in discriminating between the tense vowel /i/ and the lax vowel /ɪ/, as illustrated in the perception task involving synthetic stimuli of varying lengths.

An analysis of variance was conducted on the duration ratios for the two pairs (/i/-/ɪ/ and /ɛ/-/ɛ/) with Vowel Pair and Environment as factors. The mean ratios were 1.53, 1.62 and 1.39 for /i/-/ɪ/ in the h_d, h_b and h_pus conditions respectively and 1.37, 1.44 and 1.47 for /ɛ/-/ɛ/, respectively. The analysis yielded no main effects but a significant interaction (F(2, 56) = 4.98, p < .05). The post-hoc test in this case revealed that the only significant difference was between the environments h_b and h_pus for the high vowels. The /i/-/ɪ/ ratio was indeed smaller in the bisyllabic condition than in the monosyllabic conditions. This again lends support to the greater difficulty in maintaining the lax-tense contrast in bisyllabic words. However, though based on only four speakers, the /i/-/ɪ/ ratio was also smaller in the bisyllabic environment for the native speaker data.

To sum up, evidence that length is a distinctive vowel differentiation cue is clearly found in the L2 production data. This is true to some degree even in the case of the bisyllabic environment where it was hypothesized that vowel /ɪ/ would be produced like
vowel /i/ due to L1 syllabification (and subsequent ineffectiveness of the Lax Vowel Constraint). However, some suggestive evidence for the hypothesized effect comes from the significantly smaller /i/-/ɪ/ ratio in the bisyllabic condition, showing that in that environment duration differences between the tense and the lax vowel are smaller than in the monosyllabic environment. A closer look at the results for individual subjects analyzing both the spectral and temporal cues is necessary to determine if subjects varied in that respect.

5.3. Experiment 9. Intelligibility task

As mentioned earlier, previous research has shown it is important to complement acoustic analyses with perceptual measures. For this reason, the production data from Experiment 8 were evaluated by native speakers of English. Intelligibility tests of L2 data are in fact a common way of judging vowel production in L2 research (Munro, 1993; Munro, Flege and MacKay, 1996; Flege, 1997b; Flege, Bohn and Jang, 1997, among others).

5.3.1. Stimulus preparation

Stimuli consisted of the vowel portions edited out from each test word produced by speakers in the previous experiment so as to minimize the effect of consonant properties that might create an impression of a foreign accent. As pointed out in earlier studies, results from accent rating tasks may be affected by the mispronunciation of neighbouring sounds (Munro, 1993; Flege, Munro and MacKay, 1995). The vowel portion was extracted leaving intact cues to vowel identity insofar as possible. Signal
editing was carried out visually and aurally with Praat software v. 4.0.5 by examining both the amplitude of a waveform and the formant structure on a spectrogram. The vowel portion was defined as the portion from the first to the last positive peak in the periodic portion of the signal as indicated by an increase/decrease in overall amplitude and waveform complexity. The selected vowel portion was windowed out smoothing the onset and offset with a ramping function and making the initial and final splices at zero crossings.

5.3.2. Listeners

Eight native speakers of Southern Ontario English participated in the vowel identification task at three different times, as explained below. This group of listeners consisted mostly of undergraduate students at the University of Toronto.

5.3.3. Procedure

The experiment was spread over three different sessions, each session corresponding to each of the three elicited environments: h_d, h_b and h_pus. The procedure and length of the three sessions was the same. The task was a forced choice vowel identification task. Listeners were asked to identify the vowels they heard as the vowel in one of the six word alternatives that appeared on the screen, namely, ‘had’ (/æ/), ‘heed’ (/i/), ‘hid’ (/u/), ‘hayed’ (/e/) ‘head’ (/e/) and ‘hub’ (/ʌ/). Thus they were to choose the English word that contained the vowel which was most similar to the one they heard. Each stimulus appeared twice in the session. Each session consisted of eight blocks of 62
vowels. The order of the stimuli was randomized but maintained an equal distribution of female, male, native and non-native talkers across blocks.

The tasks were performed using an IBM laptop computer and good quality headphones (Sennheiser Headmax HD 400). A trial consisted of two presentations of each stimulus with an inter-stimulus interval of two seconds. After hearing a given stimulus, subjects selected a response from the alternatives presented on the screen by clicking on it. After selecting a response, subjects prompted the program for the following trial by clicking on the instructions on the screen. The task was preceded by a training period to familiarize the subjects with the procedure and to adjust the listening level. Each session lasted approximately forty-five to fifty minutes.

5.3.4. Results

Table 5.6 presents the percentages of correct identifications for the L2 learner group in each word environment. Across environments, English vowel /e/ is the one that obtained the highest scores, with an average of 99% correct identification. This vowel was followed by /æ/ (88%), with similar rates in all three environment conditions. When not identified as intended, misidentified /e/’s were heard as the vowel in ‘had’ or ‘hid’.

Vowel /i/ obtained overall lower identification scores (77%) and some variation with respect to environment. When not identified, it was most commonly heard as the vowel in ‘hid’. Finally, vowel /u/ was the least successfully produced according to the lower identification rates (70%), which were lowest in the h_pus condition, and was most often heard as the vowel in ‘head’ or ‘heed’, depending on the environment, as will be discussed below. The percentages of correct identifications are illustrated in Figure 5.6.
Table 5.6. Percentages of correct identification of the target vowels by the L2 learners, listed by word type. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/u/</th>
<th>/e/</th>
<th>/ɛ/</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>h_d</td>
<td>81.7 (25)</td>
<td>72.3 (25)</td>
<td>99.4 (2)</td>
<td>90.3 (10)</td>
<td>85.9</td>
</tr>
<tr>
<td>h_b</td>
<td>72.8 (33)</td>
<td>71.8 (33)</td>
<td>98.7 (4)</td>
<td>86.3 (9)</td>
<td>82.4</td>
</tr>
<tr>
<td>h_pus</td>
<td>77.0 (31)</td>
<td>64.9 (28)</td>
<td>98.8 (2)</td>
<td>88.5 (9)</td>
<td>82.3</td>
</tr>
<tr>
<td>Means</td>
<td>77.2</td>
<td>69.7</td>
<td>99.0</td>
<td>88.4</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.6. Correct identification of the target vowels by the L2 learners in each environment (‘iy’ = /i/, ‘I’ = /u/, ‘ey’ = /e/, ‘E’ = /ɛ/).

The identification scores of the control native speaker group were higher, reaching 96% for /i/, 94% for /e/, 92% for /ɛ/ and 87% for /u/. The comparatively lower scores for /u/ may be due to the fact that one of the four speakers yielded a rate of only 74% for /u/ (the other subjects obtained more than 90%). If we look at the identification of native English vowels by native English listeners in the perception experiments (Chapter 4) the percentages of correct identifications are higher, reaching 97-100%. Those tasks differed,
however, in that only native vowels were presented and that the task was much shorter. Additional data on native English speakers’ identification of native English production comes from the perceptual assimilation task administered to English speaking listeners for control purposes (Chapter 3). Recall that in that case the test included both Catalan vowels and English vowels. The percentages of identification of English vowels by English listeners in that case were between 91 and 100% (99% for /i/, 100% for /e/, 97% for /u/ and 91% for /e/). In any case, although the control group of English speaking talkers in the current experiment was very small, results from three different tasks using the same methodology illustrate that L2 learners’ production of English vowel /e/ and /e/ obtained identification scores within the range of correct identification of native vowels, whereas their productions of English /i/ and /u/ obtain clearly lower correct identification scores than native English productions.

A statistical analysis was conducted on the number of correct identifications per vowel (four target vowels) and source environment (h_d, h_b and hpus). The two-way analysis of variance (Vowel x Environment) revealed a main effect of Vowel (F(3,84) = 14.32, p < .001) but only a marginal effect of Environment (F(2,56) = 2.73, p < .08) and no interaction. A Tukey HSD post-hoc test revealed that the scores for /e/ were significantly higher than for /i/ and /u/ (p < .001) and the scores for /e/ were also higher than for /u/. These results are in accordance with the lower identification rates for vowel

---

3 Due to the small number of native English speakers in the control group, this ANOVA was run on the group of L2 learners only. Nevertheless, an ANOVA including all subjects and Language Group as a between factor yielded a marginal effect of Group (p < .08), a main effect of Vowel (p < .01) and no effect of Environment and no interaction. Further, a separate ANOVA on the four native speakers yielded no significant effects.
/i/ by Catalan learners of English in the perception tasks and the more successful identifications of /e/ and /e'/ (if we ignore the effect of the talkers with a more monophthongal production of /e'/). The results for /i/ are more surprising given the high degree of identification in the perceptual assimilation tasks (96-100%). Nevertheless, the perception experiments did indicate a lower percentage of correct identifications of this vowel, often as a function of vowel duration, in particular the perception task involving synthetic stimuli of varying durations (Experiment 7). A more detailed discussion of the results comparing perception and production will be presented after discussing the results from the goodness ratings task in Experiment 10.

The identification rate results show no noticeable difference between the monosyllabic and bisyllabic word environments, contrary to the hypothesized more accurate production of /i/ in monosyllabic than in bisyllabic environments. However, some evidence of that tendency may come from misidentifications. In the two monosyllabic environments, /i/ is misheard as /e/ more often than as /i/ (18% vs. 5% in h_d and 17% vs. 10% in h_b), whereas in the bisyllabic environment /i/ was heard as the tense vowel more often (18% as /i/ vs. 13% as /e/). It is likely that a closer analysis of individual subjects’ performance will determine if subjects vary with respect to the effect of environment. This will be discussed in the Chapter 7.

5.4. Experiment 10. Goodness Ratings Task

It is clear that not all utterances that are identified with the same target vowel in a listening test are in fact phonetically equivalent (Hillenbrand et al. 1995). Some tokens
may be poor renderings of a given vowel but still identifiable, even if by elimination, while other tokens may sound very target-like. Thus, L2 speech may be accented yet highly intelligible (e.g., Munro and Derwing, 1995; Munro, Flege and MacKay, 1996). In order to distinguish between intelligibility and native-like production, the L2 productions were also assessed for accuracy by means of a goodness rating task. In this task, as explained below, English speaking listeners rated the vowels in terms of how good an exemplar of the intended vowel they were. Goodness ratings are also frequent in L2 research (Munro, 1993; Flege, Bohn and Jang, 1997; Southwood and Flege, 1999). The goodness rating task was very similar in procedure to the previous task. The difference was that in this case listeners were not asked to identify the vowel, but they were to rate labelled vowels for accuracy of production.

5.4.1. Listeners

A new set of eight native speakers of Southern Ontario English participated in the goodness rating task, which was spread over three different sessions. In this case, the group of listeners consisted mostly of graduate students in Linguistics and other members of the department of Linguistics at the University of Toronto. Thus, the raters were obviously familiar with the phonetic symbols representing the vowels.

5.4.2. Procedure

The stimuli were the same as in Experiment 9. The task was divided into three sessions, one per environment condition (h_d, hD and h_pus). Listeners were asked to provide a goodness rating for the vowels they heard using a seven point scale: ‘1’
represented a poor exemplar of the target vowel, ‘7’ represented a good English-sounding vowel. Each session consisted of eight blocks and each block grouped 62 tokens of the same target vowel. There were two blocks per vowel. The order of the stimuli was randomized but with an equal distribution of female, male, native and non-native talkers across blocks. Each stimulus appeared twice in the session.

Again, the task was performed using an IBM laptop computer and good quality headphones (Sennheiser Headmax HD 400). For each block subjects had a card with the phonetic symbol of the vowel to be rated in that block and some examples of words containing that target vowel. A trial consisted of two presentations of each stimulus with an inter-stimulus interval of two seconds. After hearing a given stimulus, subjects selected a response from the alternatives presented on the screen (that is, ‘1, 2, 3, 4, 5, 6, or 7’) by clicking on it. After selecting a rating, subjects prompted the program for the following trial by clicking on the instructions on the screen. The task was preceded by a training period to familiarize the subjects with the procedure and adjust the listening level. Each session lasted approximately forty to forty-five minutes.

5.4.3. Results

The results of the goodness ratings task for each vowel by condition are given in Table 5.7 and Figure 5.7. The four native speakers obtained higher and more consistent ratings across vowels: 5.6, 5.5, 5.6 and 5.7 for /i/, /u/, /e/ and /e/, respectively.
Table 5.7. Mean goodness ratings per vowel and environment for Catalan subjects. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/u/</th>
<th>/e(^i)/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>h_d</td>
<td>4.7 (1.1)</td>
<td>3.8 (1.0)</td>
<td>5.2 (0.5)</td>
<td>4.9 (0.8)</td>
</tr>
<tr>
<td>h_b</td>
<td>4.5 (1.3)</td>
<td>3.9 (1.3)</td>
<td>5.5 (0.5)</td>
<td>5.2 (0.7)</td>
</tr>
<tr>
<td>h_pus</td>
<td>4.5 (1.0)</td>
<td>4.2 (1.2)</td>
<td>5.6 (0.4)</td>
<td>5.2 (0.5)</td>
</tr>
<tr>
<td>Means</td>
<td>4.6</td>
<td>4.0</td>
<td>5.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Figure 5.7. Mean goodness ratings per vowel and environment for Catalan subjects (‘iy’ = /i/, ‘I’ = /u/, ‘ey’ = /e\(^i\)/, ‘E’ = /e/).

A statistical analysis was conducted on the L2 data with Vowel and Environment condition as within factors. The analysis revealed a main effect of Vowel (F(3,84) = 21.49, p < .001) and also Environment (F(2,56) = 5.63, p < .01), but no interaction. A Tukey post-hoc test indicated that all vowels differed significantly in their goodness ratings, except for /e\(^i\)/ and /e/. Thus, /e\(^i\)/ and /e/ were the most accurately or nativelike produced, followed by /i/, and finally /u/. The effect of Environment is less evident in the data. As illustrated in Figure 5.7, the h_d environment seems to elicit lower ratings for
vowels /u/, /e/ and /e/, whereas this is reversed for vowel /i/. The Tukey HSD post hoc test yielded as significant only one difference, namely /u/, whose scores for h_pus were significantly greater than in h_d.

Overall, results from the goodness ratings corroborate the intelligibility test results. One exception is the difference in goodness ratings between the bisyllabic condition and the monosyllabic condition for vowel /u/. The higher goodness ratings in the bisyllabic environment contrast with the lower percentage of correct identification. A possible explanation could be that listeners may have been at times confused in their vowel evaluations when intended vowels sounded like good exemplars of another vowel. For instance, some /u/ vowels sounded like good /i/’s, since they were identified as /i/ 18% of the time in the identification task, and listeners may have given those /i/-sounding /u/ tokens a high goodness rating in a moment of confusion. Although listeners had a card reminding them of the target vowel they were rating in each block, one listener at least reported becoming momentarily confused about the vowel he was supposed to be rating in the middle of an /u/ block. If this is the cause for the discrepancy between intelligibility and goodness ratings for /u/ in the bisyllable condition, it illustrates a methodological problem that could be solved by more obvious indications of the vowel to be rated in each block and by a different organization of blocked vowels per session.

On the whole, the goodness ratings results clearly reflect and emphasize the vowel differences observed in the intelligibility test given that the vowels that obtained the highest correct identification rates in the intelligibility test also obtained the highest goodness ratings. For instance, /e/ was the vowel that obtained the highest identification
scores (99%) and the highest mean goodness ratings (5.4). It is followed by vowel /e/ which was correctly identified 88% of the time with fairly high accuracy (5.1). Vowel /i/ follows with lower identifications (77%) and slightly less accuracy (4.6). Finally /u/ is the vowel that poses more of a difficulty for the L2 group, being the least frequently identified (70%), and reaching the lowest ratings of accuracy (4.0).

A statistical analysis was conducted on the possible correlations between intelligibility scores and goodness ratings for each vowel and environment. The results for Experiment 9 and Experiment 10 were highly correlated for vowels /i/ and /u/ for each environment and for the means across environments. This means that the subjects that obtained the highest intelligibility scores for /i/ and /u/ also obtained the best goodness ratings for those two vowels, and those with the lowest intelligibility scores yielded the lowest ratings. Values are given in Table 5.8, which summarizes the results of both experiments providing ranges and standard deviations. In the case of /e/ and /e/, the only significant correlations involved the h_d environment, though only marginally significant in the case of /e/. The fact that the correlation between goodness ratings and the intelligibility test was not always significant in the case of /e/ and /e/ was due to small variance given the high intelligibility and accuracy scores across subjects. When taken as a whole collapsing across vowels, the correlation between intelligibility and goodness ratings is positive and significant (r(29) = .45, p < .05).
### Table 5.8. Listening tests results and correlations.

<table>
<thead>
<tr>
<th></th>
<th>Intelligibility test</th>
<th>Goodness rating</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Range</td>
</tr>
<tr>
<td>/i/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h_d</td>
<td>82</td>
<td>25</td>
<td>94</td>
</tr>
<tr>
<td>h_b</td>
<td>73</td>
<td>33</td>
<td>97</td>
</tr>
<tr>
<td>h_pus</td>
<td>77</td>
<td>28</td>
<td>100</td>
</tr>
<tr>
<td>Mean</td>
<td>77</td>
<td>24</td>
<td>95</td>
</tr>
<tr>
<td>/i/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h_d</td>
<td>72</td>
<td>25</td>
<td>94</td>
</tr>
<tr>
<td>h_b</td>
<td>72</td>
<td>33</td>
<td>100</td>
</tr>
<tr>
<td>h_pus</td>
<td>65</td>
<td>31</td>
<td>100</td>
</tr>
<tr>
<td>Mean</td>
<td>70</td>
<td>26</td>
<td>82</td>
</tr>
<tr>
<td>/e'/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h_d</td>
<td>99</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>h_b</td>
<td>99</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>h_pus</td>
<td>99</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>99</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>/e/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h_d</td>
<td>90</td>
<td>10</td>
<td>44</td>
</tr>
<tr>
<td>h_b</td>
<td>86</td>
<td>9</td>
<td>34</td>
</tr>
<tr>
<td>h_pus</td>
<td>88</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td>Mean</td>
<td>88</td>
<td>6</td>
<td>23</td>
</tr>
</tbody>
</table>

Finally, the results for the three environments in the goodness rating task were highly correlated for all four vowels. In the case of the intelligibility tests, however, significant correlations existed for the three environments in the case of vowel /i/ and for the h_d and h_pus conditions for vowel /i/. Results for the different environments were not correlated in the case of vowels /e'/ and /e/.
5.5. Relationship between listening test results and acoustic measurements

The effect of the different acoustic properties was compared to the results from the listening tests for vowels /i/ and /u/, which had the lowest overall scores and more variability among subjects. In the case of /i/, there were clear correlations between listening test results and temporal and spectral cues. Intelligibility scores correlated significantly with vowel duration ($r(29) = .49, p < .01$), F1 values ($r(25) = -.73, p < .001$) and the F2-F1 difference ($r(25) = .58, p < .01$). The same acoustic measurements correlated significantly with the goodness ratings scores: vowel duration ($r(29) = .44, p < .05$), F1 values ($r(25) = -.81, p < .001$) and F2-F1 difference ($r(25) = .59, p < .01$) (collapsing across contexts). These correlations were in fact significant for each of the three environments (h_d, h_b and h_pus). Thus, longer duration, lower F1 and greater F2-F1 difference were correlated with better identification and judgments for /i/.

In the case of vowel /u/, better results in the listening tests were correlated with the opposite characteristics, that is, with shorter vowel duration, higher F1, lower F2 and smaller F2-F1 difference. However, in this case the degree of correlation was lower, approximating significance only in the comparisons involving goodness ratings and F1 values ($r(25) = .38, p < .06$), F2 values ($r(25) = -.35, p < .09$) and F2-F1 difference ($r(25) = -.40, p < .05$), collapsing across context conditions. The correlation with duration was marginally significant for the goodness ratings in the h_b environment ($p < .06$). Correlations involving F1 and F2 values reached marginal significance in the h_d environment ($r(25)= .38, p < .07$ and $r(25) = -.39, p < .06$, respectively) and bisyllabic (h_pus) condition ($r(25) = .39, p < .06$ and $r(25) = -.38, p < .06$, respectively). The more

---

4 The correlations involving formant values were run on the female subjects’ data only, thus sample size in those cases was 25.
moderate correlations involving vowel /i/ are due to the fact that lower scores on listening
tasks are related to varying deviations from the target acoustic values (as also shown by
the error responses). Some mispronounced /i/ tokens had the characteristics of /i/ (with
lower F1 and higher F2 than /i/) and other poorly rated /i/ tokens were closer acoustically
to /e/ (higher F1 and lower F2 than /i/). Thus, low intelligibility and goodness ratings
scores for /i/ were the consequence of both too high and too low F1 and F2 values.

5.6. Discussion and conclusions

5.6.1. Production of strongly and weakly assimilated vowels

The results of the production experiments show that /eɪ/ and /æ/ are produced more
accurately than /i/ and /u/. In the case of /eɪ/ and /æ/, the production of the Catalan vowel
in place of the English target vowel goes unnoticed by native English listeners. This was
predicted by the high L1 to L2 assimilation rates obtained in Experiment 3 (Chapter 3),
100% (goodness rating of 5.7) and 98% (goodness rating of 5.2) respectively for /eɪ/ and
/æ/. Recall that L1 categories are predicted to be used for highly assimilated vowels.
However, the vowel /eɪ/ illustrates a case in which the assimilation of the L2 vowel to an
L1 category yields better results in production than in perception. The perception results
showed that Catalans relied mostly on vowel formant movement in the identification of
this vowel. This example illustrates the importance of cue weighting in category
formation and supports the claim that accurate representation of an L2 vowel, even in the case of strongly assimilated vowels, ultimately depends on a target-like use of features and feature weights.

The weakly assimilated vowel /u/ obtains lower identification and goodness ratings, showing that an accurate representation for this vowel has not yet been established. As found with perception, duration, a non-L1 feature, is important in the categorization of this weakly assimilated vowel, as predicted. The most notable result involves vowel /i/, whose poorer results in production are incongruent with the very high assimilation rates in the perceptual assimilation task (99% assimilation to of Catalan /i/ to English /i/ and 100% assimilation of English /i/ to Catalan /i/, with goodness of fit of 6.2 and 5.6, respectively). One interesting effect in this regard is a reliable negative correlation in the success of producing the vowels /i/ and /u/ (though this was only significant for the goodness rating results for the real word environment (r(29) = -.39, p < .05)). In other words, the best /u/ producers are the worst /i/ producers. It is possible that the attempt to categorize the weakly assimilated vowel /u/ in terms of a non-L1 feature, namely duration, and the establishment of that category in terms of a vowel opposition with /i/, entails the establishment of a category for /i/ distinct from the L1 vowel /i/. The case of /i/ clearly illustrates that vowels are acquired as part of a system, supporting the claim that the formation of a new category vowel may affect the categorization of another related vowel, even of a strongly assimilated vowel.

It is not suggested, however, that Catalan learners of English will be unable to ultimately produce the vowel /i/. They may initially assimilate English /i/ to Catalan /i/ in view of the perceived similarity. As learners become aware of the temporal difference
between /i/ and /u/, the contrast begins to be implemented as a duration contrast. Concurrently, the gradual realization of the spectral characteristics of /u/ and the creation of a more target-like internal representation for this vowel may at first carry with it the deterioration of its tense (long) counterpart in the temporal contrast, i.e., /i/. Cases of deterioration of very similar sounds have been reported before. Major (1987) found that Portuguese subjects’ production of English /æ/ deteriorated as production of /æ/ improved. Similarly, Flege (1992) observed that Dutch speakers’ production of English /u/ became worse as they gained proficiency. If the contrast is eventually reanalyzed as both a temporal and spectral contrast, more accurate representation and production of /u/ and /i/ may be achieved. The relevance of duration in the implementation of the /i/-/u/ contrast is discussed in the next section.

Finally, these results do not support any of the models that relate the likelihood of L2 vowel category formation to the degree of phonetic similarity between L1 and L2 sounds (e.g., the Speech Learning Model (Flege, 1995) or the Perceptual Assimilation Model (Best, 1995; Best and Strange, 1992)). This is clear in the fact that a strongly assimilated vowel and a weakly assimilated vowel (e.g., /i/ and /u/) obtain the same results and that strongly assimilated vowels (e.g., /i/ vs. /æ/ and /e/) obtain different results. They do not support a contrastive analysis either, which would predict better production of /i/ than /u/. Failure to find a different pattern for vowels with different degrees of similarity to L1 vowels is reported in other studies. For example, Flege and Bohn (1992) found that identification rates of Spanish speakers’ production of English /i/ and /u/, considered on the basis of spectral measurements to be similar and new,
respectively, was 57% for /i/ and 61% for /u/ for experienced late learners and 69% /i/ and 51% /u/ for inexperienced late learners, whereas the similar vowel /e/ obtained 91-99%.

Clearly, this failure in prediction underscores the importance of factors that interact with perceived similarity that such as adequate cue weighting and categorization of neighbouring vowels.

5.6.2. Acoustic cues to the lax-tense contrast

The duration of the tense and lax vowels in the Catalan subjects’ production falls within the range and the ratios found with native English speakers. Perhaps the most consistent finding of the production experiments has to do with the apparent acquisition of the length contrast by Catalan learners of English. The L2 learners in this study as a whole seem to have acquired vowel duration differences which are not found in the native language. The current study allows a comparison of the use of duration in perception and production. As we saw in the vowel identification task involving synthetic vowels of varying durations, Catalan learners of English make use of duration cues to distinguish between /i/ and /u/ just as they implement length distinctions in their productions of the two vowels. These results provide support for the hypothesis that adult learners may implement a non-L1 duration distinction to establish a contrast that has no spectral counterpart in the L1 (Bohn, 1995). This is consistent with results from studies involving L2 learners of a variety of backgrounds (Minnick Fox and Maeda, 1999; Bohn and Flege, 1990; Flege and Bohn, 1992; Wang and Munro, 1999; Flege, Bohn and Jang, 1997; Cebrian, 1999). On the other hand, this argues against McAllister, Flege and Piske (1999) and Flege’s (1995) Speech Learning Model, who posit that adult L2 learners are
unlikely to perceive and exploit L2 features not used to signal phonological contrast in the L1.

In addition to duration, intelligibility and goodness of Catalans’ production of /i/ and /u/ was found to be related to formant structure, namely low F1 values for /u/. This is consistent with earlier studies. Flege, Bohn and Jang (1997) found that listeners were influenced by both spectral and temporal cues. Munro and Nearey (1991) found clear relations between accentedness scores and acoustic properties including F1 values, formant movement and vowel duration. Munro (1993) found that the frequency of F1 was the most important predictor of vowel accentedness at least in the discrimination of front vowels, which may result from the fact that F1 reflects vowel height.

5.6.3. Relation between perception and production of L2 vowels

An important issue that emerges in L2 speech research, also of importance to first language acquisition research and phonological theory, is the relation between perception and production. As seen above, perceptually-based approaches to L2 speech limitations put an emphasis on this relation and try to determine the extent of the interaction between these two aspects of speech learning.

Results examining the perceptual origin of articulator errors in L1 acquisition yield contradictory results. Although children's misarticulation of sounds cannot usually be traced to deficits in auditory acuteness (Powers, 1957), it is clear that perception precedes production since infants discriminate L1 sounds before they can produce them distinctly (Flege, 1995). The task of L1 acquisition is then one of bringing perception and production into a perfect alignment. For example, Fox (1981) and Bell-Bertri et al. (1979)
report a clear link between perception and production of vowels for their adult monolingual subjects.

In the early stages of L2 learning perception seems to precede production. L2 learners may differentiate a new vowel contrast perceptually without differentiating this contrast in production. Researchers disagree, however, on the extent to which production and perception are correlated in experienced L2 speech. Strange (1995) observed that production errors can often be predicted from perceptual errors in the case of inexperienced learners, but perceptual difficulty may persist after production has been mastered. Thus, as claimed by Bever (1981), perception and production may be uncorrelated with more experienced learners, and the ability to align segmental production and perception may become unavailable once the L1 sound system has been established. The Speech Learning Model (Flege, 1995) proposes a closer link between perception and production in L2 acquisition. This model predicts that accurate perception precedes accurate production of L2 sounds since the production of a sound eventually corresponds to the properties represented in its phonetic category representation and L2 phonetic categories are formed on the basis of L1 perceptual categories. Further, although L2 production and perception may not be brought into the perfect alignment that characterizes L1 acquisition, modest correlations between production and perception will exist for highly experienced speakers of an L2 (Flege, 1999).

Recall from Chapter 4 that the BC group obtained overall good results in the perception experiments: vowel /e/ was the most consistently identified, identification scores for vowels /e/ and /i/ varied as a function of talker differences in diphthongization and length of the tokens, and vowel /i/ was comparatively well identified. On the whole,
the subjects who obtained the best results in the perception experiments yielded also the best results in the production experiment. Significant correlations between the results for the two sets of experiments were found for three of the four vowels, that is, vowel /i/, /u/ and /e/. Collapsing across vowels, the correlation yielded marginal significance when comparing the overall perception results with the production results from the intelligibility test ($r(29) = .36, p < .06$). Taken separately, the results for /i/ obtained a significant correlation between intelligibility scores and perception results collapsing across environments in both cases ($r(29) = .43, p < .05$), while the correlation between perception and goodness ratings was marginally significant ($r(29) = .33, p < .08$). The intelligibility results for the real word environment (h_d) were highly correlated with the perception for vowel /i/ in isolation ($r(29) = .73, p < .001$) and with the means for perception ($r(29) = .41, p < .05$). The means for intelligibility scores and perception were also correlated for /e/ ($r(29) = .37, p < .05$), as well as the results for the environment h_b in both experiments ($r(29) = .61, p < .001$). Finally, with respect to vowel /u/, the mean perception results are also correlated with the production intelligibility score for this vowel in the real word condition ($r(29) = .44, p < .05$) and marginally with the goodness ratings ($r(29) = .36, p < .06$). In addition, the intelligibility scores for the real word condition also correlated with the perception results for the bisyllabic context, the context that obtained the wider range of variation and overall lower correct identification scores in the perception experiments, though the significance was only marginal ($r(29) = .34, p < .08$).
Correlations between perception and production are thus found for three of the four vowels. The absence of correlations for vowel /e/ can be related to the very small variation in results, as illustrated in Table 5.8 above. These results indicate that the best perceivers in target vowel identification tasks tend to be the most successful producers as judged by intelligibility and goodness rating tasks. The moderate correlations in the current study are in agreement with similar findings from studies on vowel and consonant perception and production (Flege and Schmidt, 1995; Schmidt and Flege, 1995; Flege, 1993; Rochet, 1995; Flege, Bohn and Jang, 1997; Rochet and Putnam Rochet, 1999; Flege, MacKay and Meador, 1999; and McAllister, Flege and Piske, 1999), where the correlations average about $r = 0.50$ (Flege, 1999). Flege (1999) points out some possible explanations for the absence of higher correlations, including methodological factors. For example, Johnson et al. (1993, in Flege 1999) found that native speakers chose as best instances of a vowel category vowels that tended to be more peripheral in the vowel space than their own production. Nevertheless, the current results indicate that the emerging L2 system in the Catalan subjects is at a point where it reflects at least some evidence of alignment between perception and production.

5.6.4. ‘Tensing’ of /h/ in bisyllables

With respect to the hypothesized effect of syllabic environment on the production of the lax-tense distinction, an examination of the production data of the L2 learners as a group yields some tentative evidence. Recall that the prediction is that as a result of Catalan syllabification of CVCVC words, the stressed vowel in the first syllable in the bisyllables will be left in an open syllable. If Catalans have acquired the phonotactic
constraint that restricts lax vowels from occurring in open syllables, their production of lax vowels in bisyllables will be affected. This is expected to affect the lax vowel /u/ rather than /e/ because the former is a weakly assimilated vowel whereas the latter is strongly assimilated to an L1 vowel with no syllabic restriction.

Some evidence for this prediction comes from the acoustic measurements which show that /u/’s F1 values were significantly lower in the bisyllabic condition than in monosyllabic condition. A lower F1 implies a more /i/-like vowel. This is in accordance with the greater number of /i/ for /u/ errors in bisyllabic environments as opposed to /e/ for /u/ errors in monosyllabic environments. There was also greater overlap in the F1-F2 vowel space than in monosyllabic environments. Recall from the vowel plots in figures 5.2, 5.3 and 5.4 above that the greatest overlap between vowel spaces for /u/ and /i/ corresponded to the hpus tokens where the plots of the values for the whole group showed that /u/ occupied the entire vowel space for /i/. Further, the number of correct identifications for /u/ in the intelligibility test was numerically smaller in the bisyllabic environment than in the monosyllabic environments. Finally, the /i/-/u/ duration ratio was also smaller in the bisyllabic environment than in monosyllabic environments. However, an important assumption in this prediction is that L2 learners have acquired knowledge of the English Lax Vowel Constraint. To this point, knowledge of the constraint (and its interaction with syllabification) has been assessed only indirectly in its effect on the acquisition of segmental categories. A more direct test is needed to investigate the
subjects’ knowledge of the English Lax Vowel Constraint and of English syllabification strategies. This will be examined in Chapter 6.
Chapter Six. Acquisition of non-segmental structure.

6.0. Introduction

Acquiring an L2 phonological system is not limited to the construction of an inventory of sound contrasts. One of the goals of this thesis is to evaluate the interaction between the acquisition of nonsegmental structure and the acquisition of segmental categories. Two types of nonsegmental information are relevant to the lax-tense vowel contrast in English: the phonotactic constraint on lax vowels in stressed open syllables (the Lax Vowel Constraint) and the ambisyllabicity of medial intervocalic consonants. These are examined in the following experiments. Two types of task are used: syllable manipulation tasks and singular-plural picture matching tasks. The aim of these experiments is thus to examine the subjects’ performance with respect to the placing of syllable boundaries in English, and to evaluate knowledge of the Lax Vowel Constraint as assessed by word and syllable well-formedness judgements. In both types of tasks, the objective is to examine whether subjects avoid responses in which lax vowels are left in stressed open syllables.

6.1. Syllable manipulation tasks

The syllable manipulation tasks were intended to test the syllabification of medial singleton consonants. In addition, they tested the Lax Vowel Constraint (LVC) by examining whether subjects would allow lax vowels in open syllables. The tasks were a syllable repetition task in which subjects had to repeat the first syllable of a two-syllable word (Experiment 11) and a reverse order task in which subjects had to reverse the order
of the syllables in a two-syllable word (Experiment 12). Test words differed in the type of vowel they contained in the first syllable. Thus, there were words with a tense vowel in the first syllable and words with a lax vowel in the first syllable. In both tasks, the crucial test words were the latter type. If subjects respect the LVC, they should manipulate the test word in such a way that the post-vocalic consonant would not be separated from the lax vowel. Further, the syllable division tasks test whether L2 subjects’ syllabification of English words follows mainly L1 patterns (e.g., CV-CVC) or whether non-L1 patterns emerge.

6.1.1. Subjects

The subjects were a group of 25 native English speakers from Southern Ontario and a group of 30 speakers of Central Catalan from the region of Barcelona, who spoke English as a second language. The Barcelona group was made up of undergraduate students who were majoring in English. The same group of subjects performed the perception and production tests reported in Chapters 4 and 5 (i.e., the BC group). The same subjects participated in all of the following four experiments.

6.1.2. Stimulus preparation

Test words for both tasks were recorded from a male native North American English speaker who read the words. The words were taped using a digital audio tape recorder and transferred to computer files. Examples for each task were also taped and digitized. Stimuli were presented via good quality headphones. Responses were taped for subsequent analysis.
6.1.3. Experiment 11. Repetition task

6.1.3.1. Procedure

In this task subjects heard a word and had to respond by first repeating the first part of the word and then the whole word (the term ‘syllable’ was not used in the instructions). The example words and their responses are given in (1).

(1) Examples:

napkin → nap  napkin
elbow → el  elbow
velvet → vel  velvet

The examples contained words with two closed syllables so that the syllable division was uncontroversial, falling between the two consonants. Subjects were told to follow the example and to provide responses that would sound as natural as possible in English. For instance, taking the first example in (1), it was explained to the subjects that the response “nap-napkin” was a more natural response than other alternatives such as “napk-napkin” because “napk” is an improbable combination of sounds in English. Subjects then practised with seven practice words of the same kind as the ones they heard, that is, with uncontroversial syllable divisions. This was done so as to allow subjects to familiarize themselves with the methodology and procedure.

The actual test words consisted of twenty words which included seven words with a lax vowel preceding one intervocalic consonant (e.g., lemon), seven words with a tense
vowel preceding one intervocalic consonant (e.g., beaver) and six filler words. Test words can be found in Appendix B.¹

6.1.3.2. Results

The percentage of CVC-VC syllabifications was taken as the scoring criterion. CVC-VC syllabification means that the response of a CVCVC word would be CVC – CVCVC (e.g., “lemon → lem-lemon” or “beaver → beav-beaver”). In the crucial case of words containing lax vowels, this is the response that would illustrate the effect of the LVC. Table 6.1 thus presents the percentages of CVC-VC syllabification, that is, the number of times the first syllable was considered to include the medial consonant. The results clearly indicate that the two groups differ in their syllabification preferences: the English group yielded a higher number of closed-syllable syllabifications than the Catalan group. In addition, the effect of vowel type was more evident among native speakers.

<table>
<thead>
<tr>
<th>Subject group</th>
<th>Tense vowel word</th>
<th>Lax vowel word</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>38 (29)</td>
<td>82 (31)</td>
</tr>
<tr>
<td>Catalan</td>
<td>19 (29)</td>
<td>23 (35)</td>
</tr>
</tbody>
</table>

Table 6.1. Percentages of CVC-VC syllabifications in the Repetition task. (S.D.’s in parentheses)

An analysis of variance was conducted taking the number of CVC-VC breaks as the dependent measure, Vowel Type as a within factor (tense vs. lax vowels), and

¹ Some of the filler words did in fact contain lax or tense vowels in similar contexts but were not counted as test words because they were either spelled with a double consonant (e.g., pillow) or involved morphological boundaries (e.g., loser). It would be interesting to test those cases too but the small number of such word types in the current experiments does not permit a thorough analysis.
Language Group as a between factor. The ANOVA illustrated that Language Group was significant (F(1,53) = 24.83, p < .001 in the subject analysis and F(1,12) = 141.71, p < .001 in the item analysis) and so was Vowel Type (F(1,53) = 55.33, p < .001 in the subject analysis and F(1,12)=21.6, p<.001 in the item analysis). The two-way interaction was also significant (F(1,53)=38.92, p<.001 in the subject analysis and F(1,12)=34.6, p<.001 in the item analysis). Tukey HSD post hoc tests indicated that the EN group scored significantly higher than the Catalan group for both word types (i.e., tense vowel words and lax vowel words) and that there was a significant difference between lax vowel and tense vowel words in the case of the EN group but not in the case of the Catalans. Native English speakers disfavour responses that would leave lax vowels in open syllables. Thus, vowel type constrained the preferred syllabification pattern for native speakers but not for non-native speakers.

6.1.4. Experiment 12. Reverse order task

6.1.4.1. Procedure

In this task, subjects were asked to break the test word into two parts and reverse the order of the two parts, following a set of examples (as in the previous task, the term syllable was not used in the instructions). The examples are given in (2). Subjects listened to three examples before starting the practice session.

(2) Examples:

napkin → kin nap
elbow → bow el
velvet → vet vel
The examples were the same as in the previous task, consisting of two closed syllables so that the syllable division was uncontroversial, falling between the two consonants. Subjects were told to follow the example and to provide responses that would sound as natural as possible in English. Taking the first example in (2), it was explained to the subjects that the response “kin-nap” was a more natural response than other alternatives such as “in-napk” or “pkin-na”. Subjects then practised with the same seven practice words used in the previous experiment. The actual test words were different from those in the Repetition task but they had the same characteristics. Specifically, there were seven words with a lax vowel preceding one intervocalic consonant (e.g., melon), seven words with a tense vowel preceding one intervocalic consonant (e.g., demon), and six filler words. All test words can be found in Appendix B.

6.1.4.2. Results

Table 6.2 presents the results by group and word type showing the percentage of responses that involved a CVC syllabification of the first syllable, e.g. “melon \(\rightarrow\) (l)on-mel” or “demon \(\rightarrow\) (m)on-dem”. When reversing the order of the word, some subjects repeated the medial consonant as in “melon \(\rightarrow\) lon-mel”. This could be taken as support for an ambisyllabicity analysis of medial intervocalic consonants. One purpose of this task was to evaluate the strength of the requirement that lax vowels be found in closed syllables. Therefore, the important consideration when computing the responses was whether the medial consonant was syllabified with a preceding lax vowel, and thus both “lon-mel” (ambisyllabicity) and “on-mel” (no ambisyllabicity) responses were counted as CVC syllabification of the first syllable. An inspection of the responses provides support
for ambisyllabicity analyses of English syllabification (Pulgram, 1970; Giegerich, 1992; Hammond, 1997, among others, see Chapter 2). The number of ambisyllabicity responses for the EN group was 48% of the total (66% of the closed syllable syllabifications) in the case of lax vowels, and 15% of the total (40% of the closed syllable responses) in the case of the tense vowels. Considering the possible effect of the subjects’ knowledge that these words are spelled with a single medial letter, the percentage of ambisyllabicity responses is remarkable, especially in the case of lax vowels. Importantly, very few ambisyllabicity responses were obtained in the case of the L2 groups, amounting to only 1% of the total (ca. 10% of the closed syllable responses).

<table>
<thead>
<tr>
<th>Subject group</th>
<th>Tense vowel word</th>
<th>Lax vowel word</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>37 (25)</td>
<td>73 (32)</td>
</tr>
<tr>
<td>Catalans</td>
<td>8 (18)</td>
<td>7 (20)</td>
</tr>
</tbody>
</table>

Table 6.2. Percentages of CVC-VC syllabifications in the Reverse Order task. (S.D.’s in parentheses)

An analysis of variance was conducted using percentage of CVC-(C)VC breaks as the dependent measure, Vowel type as the within factor (tense vs. lax vowels), and Language Group as a between factor. Language Group was significant (F(1,53) = 63.26, p < .001 in the subject analysis and F(1,12) = 241.24, p < .001 in the item analysis) and so was Vowel Type (F(1,53) = 40.24, p < .001 in the subject analysis and F(1,12) = 13.7, p < .01 in the item analysis), and the two-way interaction was also significant (F(1,53) = 44.73, p < .001 in the subject analysis and F(1,12) = 37.36, p < .001) in the item analysis). Tukey HSD post hoc tests indicated the same pattern of significant differences as in Experiment 11. The English group scored significantly higher than the Catalan group for
both word types (i.e., tense vowel words and lax vowel words) and there were significantly more CVC syllabifications with lax vowel words compared to tense vowel words in the case of the English group but not in the case of the Catalans.

The results from Experiment 12 thus parallel those of Experiment 11. First, Catalan learners of English are using mostly L1 syllabification strategies with a clear preference for onset syllabification of the medial consonant. As discussed in Chapter 1, this was predicted in view of earlier evidence that L1 prosodic structure is highly transferable to the L2 (Tarone, 1987; Chela-Flores, 1996; Trammell, 1993; Sato, 1984; Broselow, 1987). Vowel type was a factor for native speakers but not for non-native speakers. Thus, L2 speakers make no distinction between lax and tense vowels when it comes to syllabifying on these tasks. Native English speakers, however, do show an effect of lax-vowel stickiness and disfavour responses that leave lax vowels in open syllables. It may be the case, however, that the Catalan learners of English have some knowledge of the Lax Vowel Constraint but the difficulty in overcoming highly transferable L1 syllabification patterns may impede the use of amabisyllabicity (or coda syllabification) as a strategy to provide a closed syllable environment for lax vowels. Knowledge of the LVC will be tested with different and possibly more sensitive measures in the following experiments.

6.2. Picture labelling tasks

After performing the syllable manipulation tasks, subjects were tested on their knowledge and productivity of the LVC by means of picture labelling tasks. These tasks involved English-sounding nonsense words and pictures of novel animal-like or human-
like creatures. The tasks involved matching the words with the pictures. The pictures depicted either one or two identical creatures, so that the subject’s job was to determine whether the test word could function as a singular or a plural noun. Details of each task will be explained in each section.

The purpose of these tasks was to test the validity and productivity of the LVC with a method that is free from external influences such as L1 syllabification interference. In addition, the tasks in Experiments 11 and 12 used real words and may consequently be more easily affected by word frequency effects and interference from the L1 in words that were similar to words in the L2 (e.g., ‘clinic’, ‘melon’, ‘lemon’, etc).

6.2.1. Stimuli preparation

As in Experiments 11 and 12, test words for the picture labelling tasks were elicited from a male native North American English speaker who read the list of words selected for the experiment. The experimenter first checked that the reader pronounced each word as intended. Words were taped using a digital audio tape recorder and transferred to computer files. Files were checked for pitch and volume.

6.2.2. Experiment 13. Picture matching task

6.2.2.1. Procedure

In this task, subjects were presented with two images and heard a pair of words associated with these images. They had to determine which image each word referred to.
One of the images contained a novel creature. The second image depicted two creatures that were identical to the first one. Thus, subjects had to associate each member of the pair of words with a singular or a plural meaning. They provided their response by inserting the words, without modifying them, into the following carrier sentences:

(3) a. These are two __________
    b. The one on the left is called __________

The crucial word pairs contained a monosyllable ending in a lax vowel plus /z/ (henceforth, a lax vowel word) and another word which ended either in a tense vowel plus /z/ (a tense vowel word) or in a consonant plus an alveolar fricative of the same voicing quality. Most trials contained a lax vowel word and a tense vowel word. The order of presentation of the two words in each trial varied throughout the experiment. This was done to avoid a possible effect of order or memory that would make subjects produce the words simply in the order they heard them. Stimuli were never presented in written form. The test included a total of ten word pairs. Test words and pictures can be found in Appendices C and D.

The response that would obey the LVC would be to insert in sentence (3a) a word that sounds more like a plural form, e.g., a tense vowel word or a word ending in a consonant plus ‘s’, and to insert the lax vowel word into sentence (3b). For example, with the forms /spiz/ and /gréz/, although both words would be acceptable in sentence (3b), choosing /gréz/ for (3a) would be less acceptable because it would be an irregular plural form. This is because reversing the regular plural rule would leave us with a violation of
the LVC (e.g., two /grez/ → one /gre/). Thus, the LVC-motivated alternative would be to insert /spiz/ in (a), given that it is easy to imagine it as the plural form of one /spi/, and to place /grez/ in sentence (b), as follows:

(4) a. These are two /spiz/.
   b. The one on the left is called /grez/.

Subjects were told that they would hear two nonsensical but possible English words, i.e., two new English words, and that they would have to insert them in the carrier sentences to produce plausible English sentences. Every time they heard a new pair of words, they saw a new picture on the computer screen so that the carrier sentences would refer to that new picture.

The procedure was the following. After the pair of images was presented, subjects heard the pair of words twice and then repeated each word to make sure their perception of the word was accurate. Then they heard the words again in the following order: Word 1, Word 2, Word 2, Word 1. This was done to help subjects remember the words and to minimize the effect of the order in which the words appeared in each trial. Subjects then produced the two sentences. Test words were played again if subjects had difficulty remembering them or producing them accurately. Each pair of new words was associated with a new pair of images.

Before starting the actual experiment, subjects practised first with real words and subsequently nonsense words. For example, they would hear “bobby” and “dogs” and would respond “These are two dogs. The one on the left is called Bobby.” The practice
included examples such as the pair “pinkies” and “pig” to reinforce that test words could not be changed so that a response like “These are two pigs. The one on the left is called Pinkie.” would not be acceptable because test words could not be modified. Thus the expected alternative would be “These are two pinkies. The one on the left is called Pig.”

The audio stimuli were presented via good quality headphones. Responses were taped for subsequent analysis. Prior to each experiment, subjects performed a few practice trials with real words and nonsense words to become familiar with the experimental procedure.

6.2.2.2. Results

Table 6.3 below shows the percentage of responses that obeyed the LVC, that is, how often a pair consisting of a tense vowel word and a lax vowel word, such as /spiz/-/grɛz/, were inserted in the sentences as illustrated in (4) above. Responses are listed separately according to the order the words appeared in, that is, whether the tense vowel word or lax vowel word was heard first (e.g., /spiz/-/grɛz/ vs. /grɛz/-/spiz/). This is indicated separately because it was necessary to analyze whether the order in which the test words appeared was a determining factor in the choice of response for each sentence. If subjects were simply responding in the same order as they heard the words they would obtain close to 0% LVC-obeying responses when the lax vowel word was presented first and close to 100% in the opposite case.
Table 6.3. Percentages of LVC-motivated responses (e.g., lax vowel word as singular word and Tense vowel word as plural word) in the picture matching task by both groups. (S.D.’s in parentheses)

<table>
<thead>
<tr>
<th>Subject group</th>
<th>tense vowel word 1st</th>
<th>lax vowel word 1st</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>84</td>
<td>89</td>
<td>87 (17)</td>
</tr>
<tr>
<td>Catalans</td>
<td>74</td>
<td>55</td>
<td>63 (21)</td>
</tr>
</tbody>
</table>

The results indicate that the LVC has a stronger effect on native English speakers than on L2 speakers of English. The order of the test words does not seem to affect the responses of the native speaker group. However, Order seems to affect the non-native speaker group to a certain extent, as can be seen from the drop in LVC-obeying responses when the lax vowel word was heard first. An analysis of variance was conducted using the percentage of responses that obeyed the LVC as the dependent measure, Language Group (English vs. Catalans) as a between subjects factor, and Order as a within subjects factor. The analysis showed that the main effect of Language Group was significant (F(1,53) = 17.87, p < .001 in the subject analysis and F(1,5) = 61.93, p < .001 in the item analysis). Order did not reach significance. However, there was a significant two-way interaction (F(1,53) = 5.9, p < .05 for the subject analysis and F(1,5) = 32.81, p < .01 in the item analysis). The interaction is due to the fact that Order constrained responses for the L2 group but not for the EN group, as confirmed in a Tukey HSD post-hoc test. The post-hoc test also showed that Catalans scored significantly lower than native English speakers in both order conditions.

Thus, for the EN group, responses involving a lax vowel word as a plural noun (e.g., two /grɛz/) were not favoured, amounting to only 13% of the responses, regardless of the order of the words in the pair. The L2 group’s responses obeyed the LVC less
consistently, allowing for 37% of responses of lax-vowel words to be chosen as plurals, and this percentage was higher (45%) when the lax vowel word was heard first. A t-test indicated that the Catalans’ responses were significantly different from chance (i.e., 50%) when the tense vowel word was heard first (t(29) = 4.92, p < .01) but not when the lax vowel word was heard first (t(29) = 0.95, p = .18). Collapsing across Order conditions, the Catalans’ results differed significantly from chance (t(29) = 3.39, p< .01). This indicates that although the LVC is not as strong in the case of L2 learners as for native speakers, as a group the Catalan subjects show some knowledge of the LVC in this task, while they did not in the syllable manipulation tasks. Further, both groups has high percentages of LVC obeying responses when the pair involved a word ending in a consonant plus ‘s’ and a lax vowel word (e.g. /flımps/ vs. /laz/). In this case, and regardless of the order of the test words, both groups chose the word ending in a consonant plus ‘s’ (/flımps/) as the plural word and the lax vowel word (/laz/) as the singular word in more than 90% of the cases (94% the EN group and 93% the Catalan group).

It is important to recognize some limitations in using this task with non-native speakers. The main problem is that we cannot be certain of the subjects’ ability to distinguish accurately the vowels in the stimuli correctly. Given that the quality of the vowel is what determines the choice of word in the task, being unable to distinguish vowels may make the task ineffective as a measure of phonotactic knowledge. For example, given the word pair /driz/-/griz/, it is clear for a native speaker that /dri/ is a better singular form than /grı/ because the latter has a vowel that cannot be found in a stressed open syllable. But if the L2 listener was unable to distinguish between the two
vowel sounds, choosing either /grɪ/ or /dri/ as the singular form cannot assess knowledge of the lax vowel constraint. Indeed, some speakers’ responses revealed this confusion by repeating the two words with the same vowel, e.g., /driz/ and /griz/. The difficulty lies in determining whether this is a perception problem or a production problem, that is, a difficulty in perceiving or producing the vowel contrast. A close look at their performance on the perception tests, however, indicates that this group of L2 learners (the BC group) performed overall very well with natural stimuli in the vowel identification tasks, with identification scores ranging from 80 to 100% (and from 90-100% if only monosyllables are considered, as observed in Experiments 4, 5, and 6). Thus, it may be safe to take their responses as valid assuming they are able to distinguish the vowels in perception, despite production inaccuracies. Notwithstanding, in order to be more certain of the presence or absence of an LVC effect, another analysis of the data was performed taking into account only those responses that indicated the correct identification of the vowel on the part of the L2 listener evident in a faithful repetition of the stimuli. These data were compared to the previous results. A statistical analysis comparing the two groups of responses revealed that they did not differ significantly.

To summarize, despite some possible shortcomings of this task, the results reveal a strong effect of the LVC among native speakers and a weaker effect with non-natives. A closer examination of individual subject data is necessary to determine if subjects differ with respect to knowledge and productivity of the constraint. But first let us consider the results of the second picture labelling experiment.
6.2.3. Experiment 14. Singular derivation task

6.2.3.1. Procedure

In this experiment, subjects provided a singular form for a given plural form. Subjects heard a single test word and saw two adjacent pictures on the computer screen. One picture contained two identical individuals or creatures and the written sentence “Here are two ....”. The other picture contained just one of those individuals or creatures and the sentence “And here is one ....”. Subjects had to insert the test word into the first sentence frame (“here are two X”) without modifying the word and then read the second sentence inserting the same word but with the possibility of modifying it this time. Words were of the same type as those in Experiment 13, that is, one-syllable words ending in a tense vowel plus /z/ (tense vowel words, e.g., /te/z/), words ending in a lax vowel plus /z/ (lax vowel words e.g., /krɛz/) and words ending in consonant plus ‘s’ (e.g., /frɛlz/). In the case of tense vowel words and words ending in a consonant plus ‘s’, the most likely response would be “Here are two /teɛz/. And here is one /teɬ/” and “Here are two /frɛlz/. And here is one /frɛl/”, applying the reverse of the plural formation rule based on the interpretation of the final alveolar fricative as a plural marker. This is possible because after removing the plural marker we are left with a well-formed English word. The crucial test items were of course those involving lax vowel words. In this case, interpreting the final /z/ as a plural marker and reversing the plural rule leaves an ill-formed English word, as in “Here are two /krɛz/. And here is one /kɛɾe/”. The purpose of the task in this case was to examine the way lax vowel words would be dealt with.
The procedure was as follows. Subjects heard the test words via good quality headphones and saw the pictures on the computer screen which included the carrier sentences. They heard the test word twice and repeated it to check for perception accuracy. Then they heard it once more and afterwards read the two sentences. The test word was played again if the subject was unable to recall it. The task procedure was illustrated with a few examples, followed by a practice session, including real and nonsense word examples. The real word examples included straightforward plural formations such as “cats” or “bees”, but also irregular plurals such as “fish” so as to illustrate that responses for the second sentence did not necessarily simply imply removing a final fricative. Responses were spoken into a microphone and recorded for later analysis.

6.2.3.2. Results

Table 6.4 shows the percentage of times that the final alveolar fricative (/s/ or /z/) was not removed from the test word in deriving the singular form, thus leaving the word unmodified in the singular. Notice that applying the reverse of the plural formation rule, that is, removing the final /s/ or /z/, would result in a possible English word in the case of words ending in a tense vowel plus /z/ (tense vowel words, e.g., /te'z/ → /te/ ) or words ending in a consonant plus /s/ or /z/ (e.g. /frɛlz/ → /frɛl/), but it would give an impossible English singular form in the case of test words ending in a lax vowel plus /z/ (lax vowel words, e.g., /krez/ → /kre/ ). Responses that did not involve removing the final consonant always involved leaving the word unmodified. Thus, knowledge of the LVC would
predict a predominance of unmodified responses with lax vowel words and fewer with
tense vowel words and words ending in a consonant plus ‘s’.

<table>
<thead>
<tr>
<th>Subject group</th>
<th>tense vowel word</th>
<th>lax vowel word</th>
<th>C + ‘s’ word</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>10 (16)</td>
<td>97 (12)</td>
<td>2</td>
</tr>
<tr>
<td>Catalans</td>
<td>21 (27)</td>
<td>60 (27)</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6.4. Percentage of final unmodified responses in the singular derivation task by both groups. (S.D.’s in parentheses)

These results show that in the case of words ending in a consonant plus ‘s’ both subject groups readily reversed the plural formation rule and obtained logical singular forms since unmodified responses amounted to only 2%. As in Experiment 13, words ending in a consonant plus ‘s’ were dealt with equally and straightforwardly by both groups. Although the number of items testing this condition in both experiments was in fact very small (two in each experiment), a consistent pattern of responses across subjects and groups was expected given that the reverse of the plural formation rules did not violate any phonotactic constraint, not only in the L2 (English) but also for the most part in the L1 (Catalan).

Responses of the two groups differed with the other types of test words. Tense vowel words were most often analyzed as plural forms, especially by the native speakers, though a number of unmodified responses were also obtained. Notice though that the unmodified responses do not in fact violate any particular constraint since they could be analogous to irregular plurals such as “fish” or “series”. It is interesting to note, though, that L2 learners opted for this kind of response more often than native speakers. This is probably due to the effect of the training session, which laid out the possibility of leaving
the test words unchanged. The crucial examples, then, are those involving words containing lax vowels (lax vowel words). In this case, the native speakers clearly disfavoured regular plural-form interpretations implying “s”-dropping, which reached only 3% (97% unmodified responses). Non-native speakers, however, allowed a much higher number of such interpretations (40%) but still left lax vowel words unmodified 60% of the time.

A 2x2 ANOVA was conducted on the percentage of unmodified responses. Language Group was the between subjects factor (English vs. Catalans) and vowel type the within factor (lax vowel word vs. tense vowel word). Both main effects and the two-way interaction reached significance (Language Group: F(1,53) = 6.69, p < .05 in the subject analysis and F(1,10) = 177.07, p < .001 in the item analysis; Vowel Type: F(1,53) = 473.45, p < .001 in the subject analysis and F(1,10) = 24.13, p < .001 in the item analysis). The source of the Language Group x Vowel Type interaction was that the EN group scored lower with respect to the tense vowel words but much higher with respect to the lax vowel words than the L2 group (F(1,53) = 68.92, p < .001 in the subject analysis and F(1,10) = 7.76, p < .05 in the item analysis). A Tukey HSD post hoc test in both subject and item analyses showed that vowel type had an effect for each group, so that the unmodified responses for lax vowel words were significantly higher than those for tense vowel words for both language groups. In addition, the EN group’s scores for lax vowel words were significantly higher than the scores obtained by the L2 group. The difference between the two groups with respect to the scores for tense vowel words was significant in the subject analysis but not in the item analysis.
The results from Experiment 14 show that the LVC is seldom violated by the native speakers, who hardly ever gave responses that would involve a stressed lax vowel in an open syllable. The L2 group violated the LVC significantly more often than the native speakers, but still left lax vowel words unmodified 60% of the time. The outcome that unmodified responses were significantly less frequent with tense vowel words than with lax vowel words indicates that though weaker than in native speakers, the LVC is respected by L2 learners. Catalan subjects followed a relatively consistent strategy of “s”-dropping with tense vowel words but were more uncertain about their responses involving lax vowel words. L2 learners may have a sense that ending a word in a stressed lax vowel does not appear very natural in English, or not as natural as ending a word in a stressed tense vowel. It is possible too that the instructions and practice prior to the test may have highlighted the opportunity for unmodified responses and that then L2 learners may have resorted to such responses in cases of unclear singular form derivations. Finally, a closer look at the individual subject data is necessary in order to examine if the LVC is moderately productive across all subjects or if it is strong with some and nonexistent with others. This will be done in Chapter 7.

6.3. General discussion and conclusions

The results obtained for the English native speaker group and the Catalan group are repeated and summarized in Table 6.5. The percentages refer to the responses for each task that could be judged to be motivated by the LVC, i.e., a closed-syllable strategy in the syllable manipulation tasks and a singular meaning interpretation in the picture labelling tasks. Table 6.5 also includes the results for tense vowel words and the
difference in scores between the two word types (except for Experiment 3, which involved a 2 words – 2 pictures matching technique). The difference between the two word types is of interest given that it shows whether the LVC-motivated strategy was indeed motivated by the LVC, that is, only applied with lax vowel words, or if it was a more general way of responding regardless of word type.

<table>
<thead>
<tr>
<th>Subject group</th>
<th>% CVC-</th>
<th>% SG interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp. 11</td>
<td>Exp. 12</td>
</tr>
<tr>
<td></td>
<td>Repetition Task</td>
<td>Reverse Order Task</td>
</tr>
<tr>
<td>TV-word</td>
<td>LV-word</td>
<td>Diff.</td>
</tr>
<tr>
<td>EN</td>
<td>38</td>
<td>82</td>
</tr>
<tr>
<td>CAT</td>
<td>19</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 6.5. Percentages of LVC-obeying responses per group and task (S.D.’s in parentheses). TV-word and LV-word refer to words containing a tense or a lax vowel respectively.

Task differences. With respect to the relation between the different tasks, in the case of the native English speakers the statistical analysis showed no correlation between the two syllable division tasks (repetition vs. reverse order task), but it yielded a significant correlation between the two picture labelling tasks when comparing the results for Experiment 13 and the word type difference scores in Experiment 14, $r(25) = .54$, $p < .01$. Correlations were not significant for the L2 group. Only the comparison between the syllable repetition task (Experiment 11) and the matching tasks (Experiment 13) reached even marginal significance ($r(29) = .32$, $p < .09$). These subjects appear to have
performed differently in different tasks. Although both types of tasks examine the 
violability of the LVC by forcing judgments on syllable structure well-formedness, they 
tap into different abilities. The repetition and reverse order tasks test the subjects’ 
syllabification preferences by forcing the subjects to divide the test words into two 
syllables. The two syllable division tasks also revealed different degrees of difficulty, the 
reverse order task involving a greater processing load. On the other hand, the picture 
labelling tasks examine syllable well-formedness by forcing a choice between potential 
singular and plural forms (Experiment 13) or asking subjects to derive a singular form 
from a plural form (Experiment 14). In addition, the stimuli in the tasks were also quite 
different, with syllable manipulation of real words in one case and picture labelling tasks 
invoking nonsense words in the other.

The LVC in English. The results for native English speakers are consistent with previous 
research that shows a preference for coda syllabification of medial consonants when the 
preceding vowel is lax. As discussed in Chapter 2, a number of studies on syllabification 
strategies in English have found that coda syllabification of intervocalic consonants is 
more common with preceding lax vowels than with tense vowels (Treiman and Danis, 
syllabification with English speaking children. Cebrian (1999) found varying percentages 
of coda syllabifications for different native speakers, but coda syllabifications were 
always more common with preceding lax vowels (31-73%) than with preceding tense 
vowels (5-32%). The somewhat lower scores of coda syllabification in Cebrian (1999) 
may be due to the fact that the subjects were not monolingual English speakers but also
spoke Spanish, and in some cases Catalan, fluently. Their knowledge of Spanish and Catalan may have influenced their responses. Finally, native speakers in the current study had more LVC-motivated responses than in Westby (1984) in the syllable manipulation tasks. Westby reported only 34% of closed syllable responses with lax vowels in a syllable response task. The difference between the studies may reside in the small number of subjects in Westby’s study (8 vs. 25 in the current study). Further, the test methodology also differed given that in Westby’s study responses involving coda syllabifications of intervocalic consonants were discouraged in an attempt to expose lax vowels in open syllables. For instance, in the syllable reverse order task the experimenter provided the first part of the response including the medial consonant (e.g., river Æ verv + ...). In addition, subjects were instructed to avoid copying the medial consonant (e.g., river Æ ver + riv).

With respect to picture labelling tasks, only one task showed an LVC effect in Westby’s study, namely the forced double matching technique involving two words and two pictures, which corresponds to Experiment 13 in the current study. Results for that task were very similar to the results obtained here; indeed, native speakers gave 87% LVC-obeying responses. Similar results were obtained by Wallner (1986), who reported scores from 25 native English speakers performing similar picture labelling tasks. In this case, the number of LVC-motivated responses was high in all tasks (82-88%).

The productivity of the LVC for English native speakers, or at least its strong influence on the preference for well-formed syllable structures, is supported by the current study and is in accordance with most previous works. This supports the view of the Lax Vowel Constraint as an active constraint for native English speakers, which
determines their knowledge of syllable and word well-formedness. In addition, the effect of the LVC on the syllabification of medial consonants supports models of English syllable structure that propose coda syllabification or ambisyllabicity of a medial consonant following a lax vowel (Pulgram, 1970; Giegerich, 1992; Hammond, 1997, among others).

The LVC and L2 English. The experiments also investigated whether L2 learners of English possessed knowledge of the LVC and if it had any influence in the way they structured English words. The results show that Catalan learners of English have some knowledge of the phonotactic restriction on lax vowels but still have a preference for onset syllabification of medial consonants.

The results of the syllable manipulation tasks show that the Catalan subjects follow L1 syllabification strategies regardless of vowel type. In a preliminary study, Cebrian (1999) found that Catalan learners of English produced coda syllabifications 0-35% of the time with lax vowels and 14-0% with tense vowels. Although based on a much smaller number of subjects, these results for Catalan learners of English (and native English speakers) are generally consistent with the current study. These results are also consistent with L2 acquisition studies showing that L2 learners often show a strong influence of the native language’s syllabic structure (Tarone, 1982; Rochet, 1999; Chela-Flores, 1996; Trammell, 1993; Sato, 1984; Broselow, 1987).

The picture labelling tasks, thus, are crucial in determining the status of the LVC in L2 English. The results in fact indicate an effect of the LVC in that words with lax vowels are treated differently from words with tense vowels. Although it is not as strong
as with native speakers, Catalans do show a tendency to avoid lax vowels in open syllables. This is clearly not an L1 feature, and it provides evidence that the LVC is actually learned and is productive, at least with respect to word structure preference. Taken together, the results for both native speakers and learners of English show that the Lax Vowel Constraint is part of the phonological system of both L1 and L2 English. The LVC is fully established in L1 English but less strongly manifested with L2 learners, resembling the way in which native speakers perceive and produce vowels more accurately than L2 learners.

One of the objectives of this thesis is to evaluate the interaction between segmental and non-segmental structure in the shaping of an L2 phonology. One way of examining this interaction was to investigate the effect of L1 syllabification and knowledge of L2 phonotactics on the ‘tensing’ of the lax vowel /i/. As discussed in Chapter 1, Cebrian (1996, 1999) found that Catalan learners of English were more successful at producing the /i/-/ɪ/ contrast in monosyllables than in bisyllables with a single intervocalic consonant. This was explained by the fact the Catalan speakers had some knowledge of the Lax Vowel Constraint but, due to the strong influence from L1 syllabification, they treated the first syllable of such bisyllables as an open syllable. As a consequence, the lax vowel was mispronounced as its tense counterpart (/i/). The assumptions underlying this explanation are supported by the findings in this chapter. Specifically, the results of the experiments directly show that Catalans demonstrate knowledge of the LVC but fail to syllabify in an English-like fashion. These outcomes also support the previous interpretation of the findings from the perception and
production experiments concerning the mispronunciation of the lax vowel /a/ as tense in bisyllabic environments. It is important to recall, however, that some of the effects reported in those experiments were relatively modest. To this point we have not explored whether subject variability accounts for the absence of stronger effects. A closer examination of the individual data is required to evaluate whether the productivity of the LVC is moderate across all subjects or if it is strong with some and absent with others. Potential subject differences should also be related to perception and production results to fully assess the relationship between segmental and non-segmental structure. The following chapter will include an investigation of the interaction between segmental and non-segmental acquisition by evaluating the subject data across experiments.
Chapter Seven. Global analysis and conclusions

7.0. Introduction

At the beginning of this investigation two main research questions were presented, namely:
- Question 1. How do second language learners implement a contrast that is largely based on a distinction that is not found in the L1?
- Question 2. How does the acquisition of non-segmental information interact with the acquisition of segmental categories?

To answer these questions, the current chapter provides a comprehensive evaluation of the experimental findings and investigates relationships across the set of results.

7.1. Acquisition of segmental structure

This thesis has examined the acquisition of the lax-tense contrast in English high and mid vowels by native speakers of Catalan by means of perceptual assimilation tasks and tests of L2 perception and production. The results from these experiments are summarized in Table 7.1, which shows the mean percentages and goodness ratings (where applicable) for each vowel by experiment, indicating the relevant subject groups in each case. These results are discussed by reviewing the main findings in each experiment set and summarizing the main conclusions in light of the hypotheses and predictions outlined in the first chapter.
Phonetic similarity. The perceptual assimilation task administered to Catalan speakers with little or no knowledge of English indicated that English /i/ is strongly assimilated to Catalan /i/, with very high assimilation scores and goodness ratings. The assimilation scores and goodness ratings for the mid vowels were lower but they still indicated a high degree of perceived similarity between English /e/ and Catalan /e/ as well as English /e′/ and the Catalan diphthong /ei/. The latter outcome, along with the finding that English /e′/ is more frequently assimilated to Catalan /ei/ than to the corresponding monophthong /e/, indicates that vowel movement is the main feature used in the categorization of this English sound. Importantly, the vowel /i/ appears to be the most weakly assimilated.

**Table 7.1. Summary of results for each vowel by experiment (% = percentage correct identifications or percentage perceptual assimilation; GR = goodness ratings; TC = Catalan speakers of L2 English in Toronto; BC = Catalan speakers of L2 English in Barcelona).**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group</th>
<th>/i/</th>
<th>/i/</th>
<th>/e′/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Perceptual assimilation of English vowels to the acoustically closest Catalan vowel (/i, e, e, ei/) (Experiments 1 and 2).</td>
<td>(a) Catalans no L2 Engl.</td>
<td>%</td>
<td>99</td>
<td>67</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GR</td>
<td>6.2</td>
<td>3.5</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>(b) L2 Engl. speakers</td>
<td>%</td>
<td>99</td>
<td>99</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GR</td>
<td>5.8</td>
<td>3.5</td>
<td>5.0</td>
</tr>
<tr>
<td>2. Perceptual assimilation of the acoustically closest L1 vowels (/i, e, e, ei/) to English vowels (Experiment 3).</td>
<td>English native sp.</td>
<td>%</td>
<td>100</td>
<td>64</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GR</td>
<td>5.6</td>
<td>4.3</td>
<td>5.7</td>
</tr>
<tr>
<td>3. Perception means (natural stimuli) (Experiments 4, 5, 6).</td>
<td>(a) BC group</td>
<td>%</td>
<td>88</td>
<td>89</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>(b) TC group</td>
<td>%</td>
<td>70</td>
<td>70</td>
<td>83</td>
</tr>
<tr>
<td>4. Perception means (synthetic stimuli: range of identification scores for the prototypical vowels including all durations) (Experiment 7).</td>
<td>(a) BC group</td>
<td>%</td>
<td>55-99</td>
<td>52-90</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(b) TC group</td>
<td>%</td>
<td>53-88</td>
<td>49-82</td>
<td>-</td>
</tr>
<tr>
<td>5. Production means ( intelligibility test and goodness ratings) (Experiments 8, 9).</td>
<td>BC group</td>
<td>%</td>
<td>77</td>
<td>70</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>4.6</td>
<td>4.0</td>
<td>5.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>
Although it is most often categorized in terms of the Catalan vowel /e/, this categorization is inconsistent and the corresponding goodness ratings are lower than for the other vowels.

In order to fully assess the identity relationship between strongly assimilated sounds, the perceptual assimilation of L1 sounds to L2 sounds was also evaluated (Experiment 3). In fact, the perceptual assimilation of Catalan vowels to English vowels ((2) in Table 7.1) lends support to the classification of vowels /i/, /e/ and /ɛ/ as nearly identical given that the substitution of the Catalan vowel for the English vowel goes unnoticed by English listeners.

Based on these outcomes and the hypotheses presented in Chapter 1, it was predicted that new vowel categories would not be created for the highly assimilated vowels (/i/, /e′/, /ɛ/) and that the use of the L1 vowel categories would allow comparative success in perception and production so long as they are based on the same use and weighting of vowel properties as in the target language. With respect to /u/, success in the acquisition of this vowel would depend on the ability to create a new vowel category given the absence of an identical L1 vowel. In this case, then, there is the potential to develop sensitivity to non-L1 based properties.

Interestingly, the assimilation results do not conform to expectations based on acoustic comparisons, which would predict the same level of similarity between L1 and L2 sounds. This is in accordance with the discrepancies between acoustic comparisons (based on F1-F2 values) and perceptually-based measures of similarity found in previous studies. Additional research is necessary to determine the acoustic correlates of perceptual distance.
Further, the perceptual assimilation tasks do not detect any differences in assimilation patterns between the Catalans with no knowledge of English and the Catalan learners of English ((1a) and (1b), respectively, in Table 7.1). Although the assimilation rates for /i/ and /e/ are numerically lower for the latter group, the difference was not significant. This finding runs counter to earlier findings that L2 learning heightens the ability to discern crosslinguistic differences. However, it may be necessary to evaluate a larger number of vowels to better assess the effect of experience on discrimination ability. The fact that only four vowels were tested here, three of which were very strongly assimilated to L1 categories, may account for the absence of significant difference between the two groups.

Perception. In the perception experiments, vowel /e/ was the most consistently identified. Identification scores for vowels /e/ and /i/ varied as a function of talker differences in diphthongisation and length of the tokens, and vowel /u/ was comparatively well identified by the group of Catalan speakers in Barcelona but obtained lower identifications by the group of Catalan speakers in Toronto. The results of the vowel identification task involving a synthetic continuum with varying spectral and temporal properties showed that native Canadian English speakers relied predominantly on spectral cues in discriminating vowels /i/, /u/ and /e/. Catalan learners of English patterned like native English speakers with respect to /e/. In the case of /i/ and /u/, however, the results indicated heavy reliance on durational cues.
The results of the perception experiments show that strongly assimilated (nearly identical) vowels are not necessarily easier to perceive than weakly assimilated vowels. These results do not support any of the models that relate the likelihood of L2 vowel category formation relationship to the degree of phonetic similarity between L1 and L2 sounds (e.g., the Speech Learning Model (Flege, 1995) or the Perceptual Assimilation Model (Best, 1995; Best and Strange, 1992)). This is clear in the fact that a strongly assimilated vowel and a weakly assimilated vowel obtain the same results (e.g., /i/ and /u/) and that two strongly assimilated vowels obtain different results (e.g., /e/ and /e/u/).

It appears then that accurate predictions must take cues and cue weighting into account. Non-L1 features were expected to play a greater role in the categorization of weakly assimilated vowels than strongly assimilated vowels. This was found for /u/, for which the predominant cue to identity is duration, a non-L1 feature. Strongly assimilated vowels were predicted to be categorized in terms of L1 distinctive features, mainly vowel quality in this study. This is the case for the vowel /e/, whose identification was not affected by duration in Experiment 7. The results for /e/ illustrate a case in which the main cues to the identity of a strongly assimilated vowel are not target-like. Specifically, diphthongization is the main cue for the Catalan learners of English, but not for the native English speakers. Consequently, the Catalan learners are more successful in identifying vowel tokens produced with larger formant movement than more monophthongal productions. This is not a factor for native English speakers, who are equally successful at identifying all /e/ tokens, indicating that they rely more on other cues such as the steady state portion of the vowel. The examples of /e/ and /e/ support the prediction that
target-like categorization with strongly assimilated vowels will be found if the L1 and L2 sounds are based on the same cues and cue weighting. Finally, the results for /i/ show that Catalan learners do not simply rely on quality but also on duration, which was not predicted because English /i/ and Catalan /i/ are already good matches on the basis of spectral qualities.

An explanation for the lower degree of success found with the perception of /i/ can be found in the prediction that adult learners will note temporal differences between L2 vowels despite the absence of contrastive length in the L1 if the L2 vowels differ both in duration and spectral properties and if there is no spectral match for both L2 vowels in the L1, in agreement with Bohn (1995). This is what we observe with the /i/-/ɪ/ contrast. Therefore, the fact that the weakly assimilated vowel /ɪ/ is categorized in terms of its temporal opposition to /i/ may affect the categorization of the latter and may override its strong assimilation to an L1 vowel. This too was predicted by the hypothesis that even those L2 categories that are a good match to L1 categories may be affected by the category formation of neighbouring vowels. The fact that Catalan learners are sensitive to duration differences is consistent with Bohn (1995) and Flege and Bohn (1992) and runs counter to the prediction by McAllister, Flege and Piske (1999) and Flege’s Speech Learning Model (1995) that features not used to signal phonological contrast in the L1 will be difficult for late learners to perceive.

Finally, a comparison of the results for the two groups of learners reveals some important facts. The first is the relationship between L2 input and salience of durational differences. The Catalans in Canada (the TC group) differed from the Catalans in Barcelona (the BC group) in two important ways: the amount of exposure and experience
with the L2 and the kind of L2 input they received. Thus, the TC group had been learning English for a longer time (a mean of 25 years in Canada) and was exposed to a variety of English that makes more use of spectral cues than durational cues in distinguishing /i/ from /ɪ/. The BC group was a group of foreign language learners in their native country and the amount of time spent in an English speaking country did not exceed one year. In addition, the BC group was exposed to a variety of English dialects, including British English (which is said to make more use of temporal differences). Nevertheless, Experiment 7 showed that both groups rely equally on duration for the distinction between /i/ and /ɪ/. This indicates that neither amount of exposure nor the kind of L2 input has an effect on the sensitivity to durational differences. This argues against Escudero’s (2001) claim that L2 English learners may vary in their reliance on spectral or durational cues as a function of the English variety they are exposed to and strongly emphasizes the salience of duration cues proposed in this study, as evidenced by the fact that even the TC group, exclusively exposed to a variety of English that relies mainly on spectral cues, were more sensitive to temporal cues.

Note, too, that while the effect of pedagogical approaches to the teaching of the /i/-/ɪ/ contrast as a length-based contrast would account for the reliance on duration in the case of the Catalan learners of English in Barcelona (the BC group), with a lot of formal instruction, this is less plausible as an account of the results from the group of Catalan immigrants in Canada. Thus, the predominant reliance on acoustic cues can be related to a greater salience of durational cues.

Secondly, the Catalans in Canada appear to be better perceivers of the English vowel /eɪ/ than the Catalan learners of English in Barcelona. The former were more
successful at identifying more monophthongal tokens than the latter. This indicates that increased exposure to the L2 facilitates a more English-like categorization of the L2 vowel in terms of the importance of cues. Nevertheless, the TC group obtained overall worse results than the BC group with respect to the other vowels, especially /i/ and /u/. An explanation is offered based on the role of instruction. The TC group was composed of university students with substantial formal instruction in English, including courses on English phonetics and phonology. Their metalinguistic knowledge of the English vowel system may have helped their performance in the identification of the English vowels.

Production. Production of the four target English vowels was examined by means of acoustic measurements of L2 production and listening tests with native English speakers as evaluators. The results indicated that vowels /e/ and /ei/ are the best produced. Being strongly assimilated to L1 vowels, the use of the Catalan vowel categories (/e/ and /ei/) is not detected by the English listeners, as predicted by the perceptual assimilation of L2 sounds to L1 sounds (Experiment 3). Vowel /i/ receives low intelligibility and goodness scores, which indicates that on the whole subjects have not been very successful at forming a target-like category for this weakly assimilated vowel. The creation of this new category appears to draw from non-L1 features such as duration, again as predicted for a weakly assimilated vowel. The acoustic analysis yielded significant differences in vowel duration, showing that Catalan learners make use of this cue both in perception and production of the /i/-/u/ distinction. As with the results from perception, the hypothesis that adult learners may exploit a non-L1 feature to establish a contrast that has no spectral
counterpart in the L1 is supported. Further, the acoustic analyses also show that variations in F1, F2-F1 distance and duration account for the degree of accentedness and intelligibility of the L2 production, in accordance with previous research.

The results for both perception and production of /i/ and /u/ support the idea that the categorization of the weakly assimilated vowel /u/ in terms of a temporal opposition to /i/ may affect the categorization of the latter despite its strong assimilation to an L1 vowel. This suggests that the acquisition of one vowel can directly affect the acquisition of another, and, as discussed in the previous section, supports the hypothesis that the formation of new vowel categories may even affect strongly assimilated vowels. Thus the lower scores for /i/ result from the fact that as a consequence of the need to establish a category for /u/ and the fact that /i/ is temporally contrastive with /i/, a reanalysis of /i/ in terms of a non-L1 feature may be triggered. Consequently, /i/ no longer patterns as a strongly assimilated vowel.

Subsequent analyses indicated that success in the production of these two vowels is in fact negatively correlated. As discussed in Chapter 5, it may be the case that the gradual realization of the spectral characteristics of /u/ and the creation of a more target-like representation for this vowel may carry with it the deterioration of its tense (long) counterpart in the temporal contrast, i.e., /i/. If the contrast is eventually reanalyzed as both a temporal and spectral contrast, more accurate representation and production of /i/ and /i/ are possible. The assumption, then, is that in order to achieve an accurate representation of /i/ learners go through a stage in which /i/ deteriorates.
With respect to the relationship between perception and production, significant correlations across the results of perception and production tasks were found for vowels /i/, /u/ and /e/, and on the whole, perception results are better than production results. This supports the view that accurate perception precedes accurate production in L2 (Flege’s (1995) Speech Learning Model; Rochet, 1995). However, there is an exception to this pattern: perception results are worse than production results in the case of the vowel /e/. This has been explained by the fact that the strong assimilation of this vowel to the closest L1 counterpart is based on a feature that is not the main feature in the target language, namely the formant movement that characterizes the offglide. Native English speakers’ production of this vowel has been found to vary in diphthongal quality resulting in the fact that the more diphthongal the production, the better it is identified by the Catalan learners. On the other hand, Catalans’ clearly diphthongal production is readily identified by the native English speakers and also obtains high goodness ratings, comparable to ratings obtained by native English speakers. This indicates that inaccurate categorization of this vowel in terms of greater relevance of formant movement than in native speech affects the ability to identify it but allows for accurate production. This fact is in accordance with Strange (1995) who observed that perceptual difficulties may persist even after segmental production has been mastered. In addition, it agrees with the Speech Learning Model’s proposal that production of a sound will correspond to the properties represented in the perceptual representation of the sound. Importantly, this is another example of how accurate or target-like categorization requires not only the same vowel properties as in the target language but also the same relative weighting of these properties.
7.2. Acquisition of non-segmental structure

The acquisition of non-segmental structure was examined by evaluating the extent to which the learners displayed knowledge of L2 phonotactics and syllabification strategies. Knowledge of non-segmental structure was tested on the same group of Catalan undergraduate students that participated in the perception and production experiments and on a group of native English speakers. The results for both English speakers and Catalan speakers are repeated in Table 7.2.

<table>
<thead>
<tr>
<th>Subject group</th>
<th>% CVC-</th>
<th>% SG interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Repetition Task</td>
<td>Reverse Order Task</td>
</tr>
<tr>
<td>ENG TV-word</td>
<td>Exp. 11</td>
<td>Exp. 12</td>
</tr>
<tr>
<td></td>
<td>TV-word</td>
<td>LV-word</td>
</tr>
<tr>
<td>ENG</td>
<td>38</td>
<td>82</td>
</tr>
<tr>
<td>CAT TV-word</td>
<td>19</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 7.2. Percentages of LVC-obeying responses per group and task (SD’s in parentheses). TV-word and LV-word refer to words containing a tense vowel or a lax vowel respectively; “Diff.” columns show the difference in scores between TV-words and LV-words.

The psycholinguistic experiments provided additional evidence that the Lax Vowel Constraint (LVC) is productive in native English phonology. Native English speakers show a preference for closed syllables and coda syllabification of middle consonants with lax vowels in syllable manipulation tasks. More importantly, this thesis has also shown evidence that L2 learners learn and represent phonotactic regularities in
the course of acquiring L2 phonology. The presence of the LVC in the learners’ L2 is weaker than in native speakers, but some evidence is found in the picture labelling tasks (Experiments 13 and 14). Specifically, L2 learners obey the constraint 60-63% of the time in these tasks (significantly above chance level) and a different pattern emerges for lax and tense vowels. Thus, the LVC appears to constitute part of the phonological knowledge of L2 learners. This result is consistent with recent research showing that adult learners can acquire phonotactic constraints (Onishi, Chambers and Fisher, 2002; Dell, Reed, Adams and Meyer, 2000).

The results from the syllabification tasks (Experiments 11 and 12) indicate a preference for L1 syllabification strategies that require an onset position for middle consonants. The difficulty in overcoming L1 syllabification patterns which disallow coda syllabification or ambisyllabicity is consistent with earlier studies showing that non-segmental phonological structure from the L1, such as syllabification, may be carried over to the L2 and affect L2 segmental acquisition (Tarone, 1987; Chela-Flores, 1996; Trammell, 1993; Sato, 1984; Broselow, 1987, among others). Thus, rather than a difference in knowledge and use of the LVC, the results from the syllable division tasks show that the native speakers and the learners differ in their syllabification strategies.

To conclude, the prediction that non-segmental structure would also be learned as part of the construction of the L2 phonology was confirmed by the results. Learners show some knowledge of L2 phonotactics yet appear to be less successful at acquiring L2 syllabification, specifically the syllabification of intervocalic singleton consonants. Acquisition of L2 syllabification may require more extensive exposure and experience with the L2 than the acquisition of phonotactics. In addition, as discussed above, the
difficulty in acquiring ambisyllabicity may stem from the strength of the interference from L1 syllabification.

7.3. Interaction between the acquisition of segmental and non-segmental structure

Having examined the results on the acquisition of the target vowels in terms of both perception and production ability, and having explored the acquisition of L2 phonotactics and syllabification, one remaining question is whether there is a relationship between the knowledge of L2 non-segmental structure and the acquisition of L2 vowels. The hypothesis was that the two components of phonological knowledge interact such that the acquisition of non-segmental structure (or lack thereof) may have an effect on segmental acquisition.

A series of statistical analyses was conducted to compare the results for the different experiments. First, correlations among the results from the different experiments were analyzed to determine if the acquisition of segmental structure proceeds in tandem with the learning of non-segmental structure, as expected. Only a subset of the comparisons (to be described below) were significantly correlated. Nevertheless, they indicate the expected relationship between the acquisition of vowel contrasts and knowledge of English syllabification and the LVC. It is important to note that the ability to detect correlations with this data set is limited by the fact that the overall higher results obtained for vowels /ɛ/ and /ɛ'/ (especially in production) were associated with very little variation. Consequently, the analysis focuses mainly on relationships between the perception and production of vowels /i/ and /ʌ/ and the results of the syllable division and picture labelling tasks (Table A.1 in Appendix C provides the means, ranges and standard
deviations for the results of all perception and production experiments and Table A.2 provides the same information for the psycholinguistic experiments).

With respect to psycholinguistic and production tasks, the results from picture labelling tasks were correlated with the intelligibility scores for the lax vowel /u/ in the real word condition (i.e. h_d): \( r(29) = .41, p < .05 \) in the case of the singular derivation task alone and \( r(29) = .49, p < .01 \) in the case of the two picture labelling tasks combined. This means that L2 learners of English with a more native-like knowledge of the LVC, as indicated by performance in the picture labelling tasks, are better at producing the lax vowel. The goodness ratings for this vowel (also in the real word condition) reached marginal significance when correlated with results in the syllable division tasks (collapsing across the two tasks): \( r(29) = .33, p < .08 \). These ratings were also significantly correlated with the difference in scores between the two word types (lax vowel word and tense vowel word) in the repetition task \( (r(29) = .41, p < .05) \) and reverse order task \( (r(29) = .40, p < .05) \). There is thus some relationship between the tendency to syllabify in an English-like fashion and superior production of the lax vowel /u/. As expected from the results of the psycholinguistic experiments, the correlation is stronger with the results of the picture labelling task results than with the results of the syllable division tasks because the latter tasks appear to allow for the incursion of the L1 syllabification patterns that block the creation of closed syllables.

Significant correlations were also found between the results of the psycholinguistic experiments and the perception tasks. In this case, mean perception scores for /u/ (that is, correct identifications of /u/ in perception tasks involving natural stimuli) were marginally correlated with the means for syllable division tasks \( (r(29) = .
Finally, within contexts, correct perception of /u/ in the bisyllabic environment was correlated with native-like syllabification in the repetition task \((r(29) = .36, p < .05)\) and with the mean values for both syllabification tasks \((r(29) = .40, p < .05)\). Subjects who were more successful at identifying vowel /u/ were also the ones who performed the syllabification and picture labelling tasks in a more native-like manner, and also those who produced it better, at least in the real word condition. These results support the predicted relationship between the acquisition of segmental and non-segmental structure: the establishment of an accurate category for the most dissimilar vowel is accompanied by a clearer presence of the L2 phonotactic constraint and a greater tendency to syllabify in the L2 fashion (that is, greater success in overcoming interference from L1 syllabification patterns).

To more carefully establish the relationship between phonotactic knowledge and segmental acquisition, it is necessary to isolate the roles of perception, production, and phonotactic knowledge in these correlations. This was done by conducting a hierarchical multiple regression analysis using production accuracy of vowel /u/ (as measured in the intelligibility task) as the dependent measure. Production was used because this measure provides a more conservative estimate of acquired knowledge than perception (perception often precedes production abilities). Scores for the three perception tasks were entered in step 1 and scores for the two picture labelling tasks were entered in step 2. The analysis revealed that the picture labelling tasks collectively accounted for 21% of variance above and beyond what is accounted for by the perception results \((R^2 \text{ change} = .21, F(2,23) = 3.99, p < .05)\). This result suggests that knowledge of the LVC improves production accuracy independently of the accuracy of the learners’ perceptual representations for /u/
and that it is not simply the case that superior general performance accounts for the
described correlations. Note, however, that as the correlations discussed above suggest,
most of the variance accounted for by the picture labelling tasks was due to the singular
derivation picture labelling task.

With respect to the correlations for the results for vowel /i/, the analyses revealed
a negative correlation between production results and the two syllable manipulation task
results. This was true for production results from both the intelligibility test and the
goodness ratings test. Thus, the mean intelligibility scores across the three production
environments were negatively correlated with the results from the repetition task
\( r(29) = -.48, \ p < .01 \), and the reverse order task \( r(29) = -.39, \ p < .05 \). Negative
correlations also existed between the intelligibility scores for each production
environment and the results for each of the two syllabification tasks. The results of the
repetition task were also correlated with the results of the goodness rating task for the h_b
environment \( r(29) = -.40, \ p < .05 \) and marginally for the means for the three
environments \( r(29) = -.35, \ p < .07 \). Finally, results in the repetition task also correlated
negatively with the results of one of the perception experiments, namely perception of /i/
in CVC frame. Most generally, these outcomes indicate that the subjects with a more
native-like syllabification tend to have a less accurate category for /i/.

Recall that the production results for the vowels /i/ and /u/ were found to correlate
negatively. Although in fact only a subset of all possible comparisons between the two
vowels reached significance, the results from the current analysis reinforce the contrast
between successful /i/ producers and successful /u/ producers. On account of the greater
difficulty evidenced in the experiments, we can assume that more accurate production of
the weakly assimilated /i/ and, especially, more English-like syllabification correspond to a more advanced or native-like level of English. Thus, the production of /i/ in fact gets worse as learners gain proficiency. As discussed in Chapter 5, similar cases of deterioration of vowel production have been reported (Major, 1987; Flege, 1992). This does not mean, however, that learners may not eventually be able to establish a more native-like category for /i/, once the /i/-/ɪ/ opposition is interpreted as an opposition in both temporal and spectral properties. The ability to implement the /i/-/ɪ/ distinction in a native-like fashion, like the sensitivity to ambisyllabicity, may require a great amount of experience and exposure, particularly in the case of adult learners.

Another source of interaction between the acquisition of non-segmental and segmental structure was found in earlier studies (Cebrian, 1996, 1999), namely the “tensing” of the lax vowel /ɪ/ in bisyllabic words with one single intervocalic consonant. Recall that this was interpreted to be due to a joint effect of knowledge of the LVC, which disallows lax vowels in open syllables, and L1-based syllabification of a bisyllabic CVCVC frame as CV-CVC, which leaves vowels in open syllables. As stated earlier, the LVC has some status in the L2 phonology. In addition, the syllabification tasks have shown that, as a group, the Catalan learners of English syllabify the middle consonants in the onset position, following the constraints of the L1. Therefore, we may expect to find more evidence of this phenomenon, that is, better distinction between /i/ and /ɪ/ in CVC than in CVCVC environment.

The results from the perception experiments show a drop in identification scores for vowel /ɪ/ in the bisyllabic environment. However, the poorer results in this
environment may simply be the consequence of the shorter duration of the vowel in a longer word, which makes identification more difficult. Stronger evidence for a context effect would come from production. Here we find that intelligibility scores of L2 production in a bisyllabic environment (CVCVC) are lower than in a monosyllabic environment (CVC), but only in the case of /a/. This would mean that subjects are in fact more able to produce this vowel in CVC syllables. However, the drop in scores for /a/ in bisyllabic environments did not reach significance. Nevertheless, the error rates in those experiments provide some indication that a more tense vowel is produced in CVCVC environment in place of /a/: the English speaking listeners tended to hear mispronunciations of /a/ as /i/ (18%) more often than as /e/ (13%) whereas in the monosyllabic environments misidentifications tended to involve /e/ (17% in the h_d condition and 16% in the h_b environment) more often than /i/ (5% and 9%). This is in accordance with the acoustic measurements which showed that /a/’s F1 values were significantly lower (i.e., more /i/-like) in the bisyllabic condition than in monosyllabic condition. Accordingly, the greatest overlap between F1-F2 vowel spaces for /a/ and /i/ corresponded to the bisyllabic environment. Finally, the /i-/a/ duration ratio was also smaller in the bisyllabic environment than in monosyllabic environments.

Further evidence of the conflicting effect of L2 and L1 non-segmental structure on the pronunciation of /a/ in bisyllables would come from a correlation between acquisition of the lax vowel /a/ (as indexed by perception and production ability) and English like syllabification. Although the syllable manipulation tasks showed very little evidence of
English-like syllabification, production and perception of /u/ correlates with both knowledge of the LVC (results in picture labelling tasks) and more English-like syllabification, as discussed above. Thus, there is some support for the relationship between the acquisition of the L2 vowel contrast and knowledge of the LVC and English syllabification. Nevertheless, the correlations between production of /u/ and syllabification tasks involve the production of /u/ in monosyllabic environment. There is no correlation between production of /u/ in the bisyllabic environment and native-like syllabification. Thus, there is no direct statistical support for the relation between English like syllabification and better production in CVCVC environments. However, production results for the different environments are highly correlated, so that production results of /u/ in monosyllabic environment is correlated with results in CVCVC. A closer look at the individual results is required to clarify the nature of the relationships.

An inspection of the individual subject data was undertaken in order to examine further if the joint effect of knowledge of the LVC and failure to acquire ambisyllabic resulted in the mispronunciation of the lax vowel /u/ as its tense counterpart. Out of the 29 subjects, 15 had more or less consistently good production scores for vowel /u/. Out of the remaining 14 subjects, six had low production scores in all three environments but the other eight evidenced a clear drop in intelligibility scores from the monosyllabic conditions to the bisyllabic. An examination of the errors also indicates two distinct patterns. Five out of the six subjects with low scores in all environments tended to mispronounce the target vowel /u/ as /e/ according to the vowel identification responses of the English-speaking listeners (the remaining subject rendered it as /i/). In contrast, the
eight subjects for which there was a drop in intelligibility scores from the monosyllabic to the bisyllabic environment tended to mispronounce the target vowel /i/ as /i/ in the bisyllabic environment. These can be taken as indication that /i/ is produced better in CVC environment (at some stage in acquisition) and that L1 syllabification results in mispronunciation of the target lax vowel as tense in an open syllable and failure to maintain the lax-tense contrast.

T-test analyses were conducted on the two relevant subgroups, namely the group with good results in all three environments and the group with a drop in scores in the bisyllabic environment. The results showed that the two subgroups differed significantly in most production results both for intelligibility scores and goodness ratings. In addition, and more crucially, the two subgroups did not differ with respect to the picture labelling task results but they did in terms of the syllabification tasks results ($t = 17.16, p < .05$ for the repetition task and $t = 15.93, p < .05$ for the mean of both syllabification tasks). Thus, both groups have some knowledge of the LVC, and to the same extent. But only the group with consistent scores across all three environments shows a tendency towards CVC- syllabification. Thus, the subgroup with equally good scores across environments is the group that syllabifies in a more English-like way. The subgroup that has a drop in scores in bisyllabic environments performs significantly worse in the syllable division tasks. Thus, knowledge of the LVC together with English-like syllabification seems to correlate with better production in the bisyllabic condition. Critically, there was no evidence that certain subgroups produced a better contrast between /i/ and /i/ in CVCVC environments than in CVC environments.
To sum up, the global analysis of results comparing all experiments provides evidence for a positive, if moderate, correlation between the production and perception of the lax vowel /i/ and knowledge of the LVC, and, to a lesser extent, of English syllabification strategies. In addition, the subject analysis has shown that a drop in production scores in the bisyllabic environment can be related to less target-like syllabification patterns. This is explained by two competing non-segmental constraints: the L2 LVC, which disallows lax vowels in open syllables, and L1 syllabification, which restricts intervocalic consonants to onset position. The conflict between these two non-segmental tendencies results in segmental production errors because in the absence of am bisyllabic ity the lax vowel is rendered as a tense vowel. Importantly, this issue affects the weakly assimilated vowel /i/ but not the strongly assimilated vowel /e/, which obtains high scores in all environments. It thus seems plausible to conclude that the interaction of acoustic cues of the L2 with non-segmental factors related to a particular vowel contrast plays a greater role in the categorization of weakly assimilated vowels than strongly assimilated vowels.

7.4. Final conclusions and implications

This thesis makes several primary contributions to the field of second language acquisition. The first concerns the importance of examining vowel properties and cue weights when determining the phonetic similarity between L2 and L1 sounds. As shown, while assessments of perceived (dis)similarity may allow us to predict whether or not L2 learners will establish new categories, the accuracy of L2 categories will depend on the ability to establish those new categories based on the same vowel features and feature
weights as in the target language. We have seen that L2 vowels are initially assimilated to
L1 vowels on the basis of L1 contrastive features. This yields successful perception and
production in the case of strongly assimilated vowels that are identical in terms of vowel
properties and weights, as seems to be the case for English /ɛ/ and Catalan /e/. The vowel
/e/ illustrates the case of a very strongly assimilated sound for which the L1 vowel is not
based on exactly the same cue as the assimilated L2 vowel. This results in a less accurate
representation that is manifested in less success in perception. In addition it was also
shown that non-L1 cues may be more readily exploited in the categorization of weakly
assimilated vowels than of strongly assimilated vowels. Thus, tense-lax contrasts for
which there is no L1 spectral counterpart may be implemented at least initially in terms
of the non-L1 cue of duration.

Further, strong similarity, even in cases of near identical spectral qualities, does
not ensure invariant L2-L1 identity, as illustrated by English /i/ and Catalan /i/. This
illustrates that L2 vowels are acquired as part of a system where the acquisition of one
vowel affects the acquisition of a related vowel. As a result, the need to maintain vowel
contrasts may take precedence over initial vowel assimilation patterns such that even L2
categories that are a good match for L1 categories may be unstable. As discussed above,
it seems that the attempt to categorize /u/ by establishing a vowel duration contrast with /i/
entails the creation of a category for /i/ distinct from the L1 vowel /i/. This is because the
new opposition is based on a feature that is foreign to the L1 and thus both members of
the vowel contrast fail to be assimilated to L1 vowel categories (which are defined in
terms of vowel quality). If this is the case, then clearly vowels are not acquired in
isolation; the shaping of an L2 phonology involves establishing a set of vowel contrasts
which may overrule L1-L2 relations based on perceptual assimilation. An obvious methodological implication of this contrastive precedence effect is that even when investigating the creation of a specific L2 sound category, researchers ought to consider a range of potentially related sounds that may affect the course of acquisition, as well as carefully examine the vowel features and feature weights that determine the nature of the oppositions that characterize the L1 and L2 systems.

The next important contribution involves the acquisition of L2 phonotactics. A number of arguments have been provided supporting the claim that the English Lax Vowel Constraint is a distinct component of a speaker’s synchronic phonological knowledge. As discussed in Chapter 2, studies on infants, children and adults indicate that phonotactic constraints are learned concomitantly with the other components of the speaker’s grammar and are internally represented by the language speaker. This perspective is supported by the results from the experiments in Chapter 6, which demonstrate the strong productivity of the LVC with native English speakers. As for the possibility that phonotactics are acquired as part of the L2 phonology, previous evaluations have been limited and are indirect in nature. For example, Onishi, Chambers and Fisher (2002) tested the ability to extract a probabilistic phonotactic constraint while listening to a list of English-like words with a novel restriction on the distribution of consonants with respect to syllable position and adjacent vowels. They found that knowledge of the constraint was reflected in the latency of subjects' cued productions after the listening experience. While this is generally consistent with the idea that adults have the ability to learn these patterns from exposure, the current experiments go farther in evaluating whether knowledge of phonotactics is reflected in the course of real L2
language acquisition. The findings of the current research do indeed show that the English LVC is acquired, to some degree, by L2 learners. Interestingly, the evidence that Catalan learners were acquiring this constraint was more obvious than the evidence that they were acquiring an understanding of English syllabification patterns, possibly because of the strong interference from the L1. This outcome raises a range of interesting questions relating to the relative rates at which aspects of non-segmental phonological knowledge are acquired, and the reasons for these differences.

The final contribution identifies an important relationship between segmental and non-segmental structure. Statistical analyses show that knowledge of the LVC and acquisition of segmental structure, as evidenced by perception and production ability, are correlated. This is consistent with the idea that phonotactics is acquired in tandem with segmental structure. For example, experience with the sounds may trigger the creation of categories that at first may be rough approximations to native L2 categories. As learners continue to be exposed to the L2 sounds, they will begin to fine-tune these categories. One consequence of the fact that these sounds appear in words is that learners may start to detect and learn phonotactic constraints, which then become a source of evidence for creating categories. Some support has been found for the hypothesized interaction between the different phonological components in the acquisition of segmental structure. Catalan learners appear to have knowledge of the LVC before they are able to overcome L1 syllabification patterns, which results in a difference in production accuracy of the contrast between monosyllables and bisyllables. This implies that the acquisition of segmental structure is not independent of the knowledge of L2 prosodic structure and phonotactic constraints.
7.5. Directions for future research

This thesis has attempted to provide some answers to a range of questions on the acquisition of L2 contrasts and the relationship between different types of phonological knowledge. The attempt to answer these questions has brought to light new issues that deserve further attention. This section concludes this work by outlining the most relevant directions for future research.

With respect to non-segmental phonological knowledge, such as phonotactics and ambisyllabicity, there are many outstanding questions. For instance, how much evidence is required before learners can begin to extract these patterns, and do L2 learners achieve the same productivity as native speakers. This could be done, for example, by means of a study with a large number of speakers and a wide range of proficiency levels. In addition, the relationship between these and other components of the phonology can be further explored, for example by investigating the status of other phonotactic information such as the distribution of the velar nasal /ŋ/, the voiced palato-alveolar fricative /ʒ/ or /h/, or the restriction on tense vowels before velar nasals. Comparing the knowledge and productivity of these constraints may allow a better understanding of the role and the process of acquisition of phonotactics. Further, phonotactic constraints may be based on articulatory markedness or sonority sequencing preferences (e.g., the constraint on stop+stop onset sequences) or to other issues such as syllable weight requirements, as in the case of the LVC. More work needs to be done to determine if such factors have a bearing on their acquisition and role in the phonology.
With respect to second language acquisition research more generally, this thesis presents new proposals for the study of L2 phonology acquisition that need to be tested further. The use of two kinds of perceptual assimilation tasks (L2 to L1 and L1 to L2 assimilation) can be applied to other vowel studies and other language situations. Other proposals outlined in this study included the creation of L2 vowel categories on the basis of L2-L1 assimilation relationships, the effect of the L1 predominant features in the assimilation process, the salience of temporal differences when they accompany spectral differences, and the claim that vowels are acquired in a system where maintaining vowel contrasts takes precedence over individual category formation. These can be extended and explored further along different dimensions including degree of L2 proficiency (beginners and highly proficient), and more naturalistic data (e.g., testing perception and production in sentences rather than citation form), in addition to other languages with different inventory sizes. More specifically, the relationship between the Lax Vowel Constraint, ambisyllabiciry and segmental acquisition can be examined further by extending it to other cases such as the acquisition of the English high back vowels (/u/ and /u/) or to other languages with lax-tense vowel contrasts (e.g., Dutch or German).
Appendices

Appendix A

Descriptive statistics for the Catalan native speakers' results (the BC group).

<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>/1/</td>
<td>Perception</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>50</td>
<td>100</td>
<td>90</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>hVb</td>
<td>50</td>
<td>100</td>
<td>93</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>hVbut</td>
<td>60</td>
<td>100</td>
<td>86</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>43</td>
<td>100</td>
<td>89</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Intelligibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hVd</td>
<td>93.7</td>
<td>100</td>
<td>72</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>hVb</td>
<td>100</td>
<td>100</td>
<td>72</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>hVpus</td>
<td>100</td>
<td>100</td>
<td>65</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>82.3</td>
<td>99</td>
<td>70</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Goodness Rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hVd</td>
<td>4.4</td>
<td>5.6</td>
<td>3.8</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>hVb</td>
<td>4.7</td>
<td>5.7</td>
<td>3.9</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>hVpus</td>
<td>4.5</td>
<td>5.8</td>
<td>4.2</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>4.2</td>
<td>5.5</td>
<td>4.0</td>
<td>1.01</td>
</tr>
<tr>
<td>/i/</td>
<td>Perception</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>42</td>
<td>100</td>
<td>92</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>hVb</td>
<td>50</td>
<td>100</td>
<td>94</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>hVbut</td>
<td>50</td>
<td>100</td>
<td>78</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>33</td>
<td>100</td>
<td>88</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Intelligibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hVd</td>
<td>93.8</td>
<td>100</td>
<td>82</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>hVb</td>
<td>96.9</td>
<td>100</td>
<td>73</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>hVpus</td>
<td>100</td>
<td>100</td>
<td>77</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>95</td>
<td>100</td>
<td>77</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Goodness Rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hVd</td>
<td>4.3</td>
<td>6.3</td>
<td>4.7</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>hVb</td>
<td>5.2</td>
<td>6.8</td>
<td>4.5</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>hVpus</td>
<td>4</td>
<td>6.2</td>
<td>4.5</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>4</td>
<td>6.3</td>
<td>4.6</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table A.1a. Results of the perception and production experiments for Catalan native speakers (the BC group).
<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>/ɛ/</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>33</td>
<td>67</td>
<td>100</td>
<td><strong>91</strong></td>
<td>10</td>
</tr>
<tr>
<td>hVb</td>
<td>20</td>
<td>80</td>
<td>100</td>
<td><strong>98</strong></td>
<td>5</td>
</tr>
<tr>
<td>hVbut</td>
<td>20</td>
<td>80</td>
<td>100</td>
<td><strong>98</strong></td>
<td>5</td>
</tr>
<tr>
<td>Mean</td>
<td>16</td>
<td>84</td>
<td>100</td>
<td><strong>96</strong></td>
<td>5</td>
</tr>
<tr>
<td>Intelligibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hVd</td>
<td>44</td>
<td>56</td>
<td>100</td>
<td><strong>90</strong></td>
<td>10</td>
</tr>
<tr>
<td>hVb</td>
<td>35</td>
<td>66</td>
<td>100</td>
<td><strong>86</strong></td>
<td>9</td>
</tr>
<tr>
<td>hVpus</td>
<td>47</td>
<td>53</td>
<td>100</td>
<td><strong>88</strong></td>
<td>9</td>
</tr>
<tr>
<td>Mean</td>
<td>23</td>
<td>75</td>
<td>98</td>
<td><strong>88</strong></td>
<td>6</td>
</tr>
<tr>
<td>Goodness Rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hVd</td>
<td>2.9</td>
<td>3.2</td>
<td>6.1</td>
<td><strong>4.9</strong></td>
<td>0.8</td>
</tr>
<tr>
<td>hVb</td>
<td>3.2</td>
<td>3.4</td>
<td>6.6</td>
<td><strong>5.2</strong></td>
<td>0.7</td>
</tr>
<tr>
<td>hVpus</td>
<td>1.9</td>
<td>4.1</td>
<td>6</td>
<td><strong>5.2</strong></td>
<td>0.5</td>
</tr>
<tr>
<td>Mean</td>
<td>2.4</td>
<td>3.6</td>
<td>6</td>
<td><strong>5.1</strong></td>
<td>0.6</td>
</tr>
<tr>
<td><strong>/ɛʰ/</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td><strong>90</strong></td>
<td>15</td>
</tr>
<tr>
<td>hVb</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td><strong>61</strong></td>
<td>29</td>
</tr>
<tr>
<td>hVbut</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td><strong>70</strong></td>
<td>19</td>
</tr>
<tr>
<td>Mean</td>
<td>64</td>
<td>33</td>
<td>97</td>
<td><strong>74</strong></td>
<td>16</td>
</tr>
<tr>
<td>Intelligibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hVd</td>
<td>6</td>
<td>94</td>
<td>100</td>
<td><strong>99</strong></td>
<td>2</td>
</tr>
<tr>
<td>hVb</td>
<td>19</td>
<td>81</td>
<td>100</td>
<td><strong>99</strong></td>
<td>4</td>
</tr>
<tr>
<td>hVpus</td>
<td>6</td>
<td>94</td>
<td>100</td>
<td><strong>99</strong></td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td>6</td>
<td>94</td>
<td>100</td>
<td><strong>99</strong></td>
<td>1</td>
</tr>
<tr>
<td>Goodness Rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hVd</td>
<td>2</td>
<td>3.9</td>
<td>5.9</td>
<td><strong>5.2</strong></td>
<td>0.5</td>
</tr>
<tr>
<td>hVb</td>
<td>2.1</td>
<td>4.3</td>
<td>6.4</td>
<td><strong>5.5</strong></td>
<td>0.5</td>
</tr>
<tr>
<td>hVpus</td>
<td>1.3</td>
<td>5.0</td>
<td>6.3</td>
<td><strong>5.6</strong></td>
<td>0.4</td>
</tr>
<tr>
<td>Mean</td>
<td>1.6</td>
<td>4.4</td>
<td>6.0</td>
<td><strong>5.4</strong></td>
<td>0.4</td>
</tr>
</tbody>
</table>

Table A.1b. Results of the perception and production experiments for Catalan native speakers (the BC group).
Table A.2. Results for the psycholinguistic experiments for Catalan native speakers (the BC group).

Appendix B.

Syllable manipulation tasks: List of test words.

Repetition task

Examples:
napkin ➔ nap napkin
elbow ➔ el elbow
velvet ➔ vel velvet

Practice:
napkin
dentist
elbow
kidnap
velvet
splendid
wonder

Experiment items:
seldom ➔ beaver
depend ➔ lemon
Venus ➔ pillow
bacon ➔ ceiling
river ➔ meadow
loser ➔ heaven
balance ➔ divorce
Simon ➔ people
select ➔ even
clinic ➔ Philip
chicken ➔ basket
Reverse order task.

Examples:
napkin → kin nap
elbow → bow el
velvet → vet vel

Practice:
napkin
elbow
velvet
wonder
dentist
kidnap
splendid

Experiment items:
window
reveal
demon
label
liver
feeling
Alice
final
guitar
limit
widow

music
panic
fellow
pilot
leopard
seven
delay
deacon
evil
melon
Appendix C

Picture labelling tasks: List of test words.

Test words for Experiment 13. Picture matching task

dʒez  fets
tʌz  spiz
nez  plɔ'z
diz  θe'z
trans  flʌnz
griz  driz
plæz  ske'z
lʌz  flɪmps
sliz  bliz
kez  spuz

Picture labelling tasks: Pictures.

Test words for Experiment 14. Singular derivation task

felz  lars
ekaz  tʃjets
spiz  frelz
lʌz  te'z
krez  tʃæz
ʃe'z  pels
blez  druz
giz  glʌks
tez  fliz
kajz
Appendix D

**Picture labelling tasks: Pictures**
Pictures for Experiment 13. Picture matching task

1

![Picture 1](image1.png)

2

![Picture 2](image2.png)

3

![Picture 3](image3.png)
Pictures for Experiment 14. Singular derivation task.

Example
Here are two . . .

Here is one . . .

Example
Here are two . . .

Here is one . . .
Example.
Here are two ...

Here is one . . .

Example
Here are two ...

Here is one . . .

Image T2I1
Here are two ...

Here is one . . .
Image T2I2
Here are two …

Here is one . . .

Image T2I3
Here are two …

Here is one . . .

Image T2I4
Here are two

Here is one . . .
Here are two ...

Image T2I 5

Here is one . . .

Image T2I 6

Here are two ...

Image T2I 7

Here is one . . .
Here are two …

Here is one . . .

Image T2I9
Here are two ...

Here is one . . .

Image T2I 10
Here are two ...

Here is one . . .
Here are two ...

Here is one . . .

Here are two ...

Here is one . . .

Here are two ...

Here is one . . .
Here are two …

Here is one . . .

Here are two …

Here is one . . .

Here are two …

Here is one . . .
Bibliography


