VOICING-DEPENDENT VOWEL DURATION IN STANDARD ARABIC
AND ITS ACQUISITION BY ADULT AMERICAN STUDENTS

DISSERTATION

Presented in Partial Fulfillment of the Requirement for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

Lutfi Hussein, B.A., M.A.

* * * * *

The Ohio State University
1994

Dissertation Committee:
Brian Joseph
Mary Beckman
Frederic Cadora
Charles Hancock

Approved by
Advisor, Interdisciplinary Program
Copyright by
Lutfi Hussein
1994
ACKNOWLEDGMENTS

Many family members, professors, friends, and students have been of great help throughout my graduate career and in writing this dissertation. My wife Tarja and my brother Abdullah have been constantly supportive, caring, and infinitely patient. Without their support and care graduate school would have been much more difficult. Professor Frederic Cadora has been a source of encouragement from the first day I attended OSU. He often went out of his way to provide the support I needed to pursue my graduate studies. More importantly, he always opened his office, home, and heart for me and my many concerns. For all of this I am very grateful. My dissertation advisors Professors Mary Beckman and Brian Joseph made the writing of this dissertation a valuable learning experience. Mary guided me at different stages in designing the experiments and writing this thesis. Her advice and suggestions have been invaluable to this study. Brian patiently read the many drafts I wrote and provided invaluable comments and suggestions on each. Both advisors have been a source of inspiration through their hard work, integrity, and pleasant personalities. They have both contributed to my growth at the professional and personal levels. My deep gratitude goes to both of them. Professor Charles Hancock read the final draft and provided several constructive suggestions, especially on the pedagogical portion of this thesis. He was also willing to accommodate my needs and always managed to fit me in his busy schedule. For that I express my sincere thanks. I am also indebted
to my colleague professor Meihua Zhai of Ohio University who spent many hours designing and writing the statistical programs needed for this study. Finally, to the many students, friends, and colleagues who participated in the experiment, I acknowledge their assistance with deep appreciation.
VITA

1980 ..........................................................  B. A. Bethlehem University,
                                          West Bank.
1982-1984 ................................................  Graduate Teaching Associate,
                                          Ohio University, Athens, Ohio.
1984 ..........................................................  M.A., Department of Linguistics,
                                          Ohio University, Athens, Ohio.
1984-1992 ................................................  Graduate Teaching Associate,
                                          Ohio State University, Columbus,
                                          Ohio.
1986 ..........................................................  M.A., Department of Linguistics,
                                          Ohio State University, Columbus,
                                          Ohio.
1992- present .............................................  Instructor of Linguistics and Arabic
                                          Ohio University, Athens, Ohio.

PUBLICATIONS


**FIELDS OF STUDY**

Major Field: Interdisciplinary
  *Phonetics: Mary Beckman and Ilse Lehiste*
  *Historical Linguistics: Brian Joseph, Frederic Cadora*
  *Pedagogy: Charles Hancock*
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................. ii

VITA ......................................................................................... iv

TABLE OF CONTENTS ............................................................... vi

LIST OF TABLES .......................................................................... ix

LIST OF FIGURES ........................................................................ xii

CHAPTER PAGE

I.  INTRODUCTION ................................................................. 1

   1.0 Voicing-dependent vowel duration variation ............ 1

   1.1 The acquisition of the phonetic system of a second
       language in adulthood and foreign language accent .... 4

   1.2 Accounts for the causes of foreign accent ............... 8

   1.3 Experience and the acquisition of L2 phonological
       system ................................................................. 10

   1.4 Organization of the dissertation .............................. 11

II. VOICING-DEPENDENT VOWEL DURATION:
    LITERATURE REVIEW ......................................................... 13

   2.0 Introduction ................................................................. 13

   2.1 Production studies ..................................................... 14

   2.2 Perception studies ..................................................... 25

   2.3 Conditioning factors that affect vowel duration
       and confound the role of voicing on vowel
       duration ........................................................................ 32

   2.3.1 Types of vowel duration ......................................... 35

       2.3.1.1 Phonemic vowel duration ............................ 35

       2.3.1.2 Intrinsic vowel duration ............................ 35

       2.3.1.3 Environmentally-conditioned vowel duration .. 37

       2.3.1.3.1 Effect of adjacent consonants ............... 37

       2.3.1.3.2 Effect of stress .................................... 45
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.1.3.3 Effect of tempo on vowel duration</td>
<td>53</td>
</tr>
<tr>
<td>2.3.1.3.4 Effect of word length and syntactic boundaries on vowel duration</td>
<td>54</td>
</tr>
<tr>
<td>2.3.1.3.5 Effect of intonation on vowel duration</td>
<td>63</td>
</tr>
<tr>
<td>2.4 Voicing effect in Arabic</td>
<td>64</td>
</tr>
<tr>
<td>2.5 The voicing effect: universal or learned? Where does Arabic fit in this debate and does this study attempt to achieve?</td>
<td>71</td>
</tr>
<tr>
<td>III. THE EXPERIMENT</td>
<td>81</td>
</tr>
<tr>
<td>3.1 Subjects</td>
<td>81</td>
</tr>
<tr>
<td>3.1.1 Monolingual Arabic speakers</td>
<td>81</td>
</tr>
<tr>
<td>3.1.2 Bilingual American learners of Arabic</td>
<td>83</td>
</tr>
<tr>
<td>3.1.3 American English monolinguals</td>
<td>85</td>
</tr>
<tr>
<td>3.2 Test materials</td>
<td>86</td>
</tr>
<tr>
<td>3.3 Segmentation and measurement</td>
<td>93</td>
</tr>
<tr>
<td>IV. RESULTS, INTERPRETATION, AND DISCUSSION</td>
<td>106</td>
</tr>
<tr>
<td>4.1 Voicing-dependent vowel duration in Standard Arabic</td>
<td>106</td>
</tr>
<tr>
<td>4.1.1 Voicing-dependent vowel duration before stops</td>
<td>106</td>
</tr>
<tr>
<td>4.1.2 Voicing-dependent vowel duration before fricatives</td>
<td>111</td>
</tr>
<tr>
<td>4.1.3 Discussion</td>
<td>116</td>
</tr>
<tr>
<td>4.2 The acquisition of the voicing effect in Arabic by adult American students</td>
<td>122</td>
</tr>
<tr>
<td>4.2.1 The acquisition of the voicing effect on vowel duration before stops</td>
<td>123</td>
</tr>
<tr>
<td>4.2.2 The acquisition of the voicing effect in the pre-dental fricatives /s/ and /z/</td>
<td>132</td>
</tr>
<tr>
<td>4.2.3 The acquisition of the voicing effect in the pre-velar fricatives /x/ and /ɣ/</td>
<td>140</td>
</tr>
<tr>
<td>4.2.4 Discussion</td>
<td>148</td>
</tr>
<tr>
<td>4.3 The effect of learning L2 on L1 (Reverse phonological interference)</td>
<td>157</td>
</tr>
<tr>
<td>4.3.1 Method</td>
<td>160</td>
</tr>
<tr>
<td>4.3.1.1 Subjects</td>
<td>160</td>
</tr>
<tr>
<td>4.3.1.2 Test materials and procedure</td>
<td>160</td>
</tr>
<tr>
<td>4.3.1.3 Recording, measurement and segmentation</td>
<td>161</td>
</tr>
<tr>
<td>4.3.2 Results</td>
<td>161</td>
</tr>
<tr>
<td>4.3.3 Discussion</td>
<td>167</td>
</tr>
</tbody>
</table>
4.3.4 Conclusions ................................................................. 171
4.4 Characteristics of Arabic produced by the bilingual
groups .............................................................................. 171

V. CONCLUSIONS, DISCUSSIONS, AND IMPLICATIONS .......... 180

5.1 Summary of conclusions and discussion ....................... 180
5.2 Pedagogical implications of the study ......................... 185
5.3 Final remarks and recommendations ............................. 191

BIBLIOGRAPHY ..................................................................... 194
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Average vowel duration before /t/ and /d/ as reported by Heffner in 1937</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Average vowel duration before voiced voiceless consonants as reported by Peterson and Lehiste (1961)</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>Average vowel length in ms in all four languages as reported in Chen (1970)</td>
<td>22</td>
</tr>
<tr>
<td>2.4</td>
<td>Average vowel duration as a function of place of articulation of the following consonant as reported by Peterson and Lehiste (1960)</td>
<td>39</td>
</tr>
<tr>
<td>2.5</td>
<td>Average vowel duration as a function of place of articulation reported by Crystal and House (1988)</td>
<td>41</td>
</tr>
<tr>
<td>2.6</td>
<td>Average vowel duration as a function of manner of articulation for the following consonant reported by House and Fairbanks (1953)</td>
<td>43</td>
</tr>
<tr>
<td>2.7</td>
<td>Average ratios of vowel duration as a function of manner of articulation of the following consonant as reported by Laeufer (1992)</td>
<td>44</td>
</tr>
<tr>
<td>2.8</td>
<td>Table 2.8. Reduction in vowel length as more syllables are added to the word (reported in Lehiste 1970)</td>
<td>56</td>
</tr>
<tr>
<td>2.9</td>
<td>Average vowel durations in mono- and bisyllabic words as reported by Klatt (1973)</td>
<td>56</td>
</tr>
<tr>
<td>2.10</td>
<td>Average vowel durations in different positions in the sentence before voiced and voiceless consonants as reported by Laeufer (1992)</td>
<td>62</td>
</tr>
</tbody>
</table>
2.11 Average vowel length in Standard Arabic before voiced voiceless consonants and their ratios as reported by Port et al. (1980 at three different tempos .... 66

3.1 Target words and carrier sentences used in the study for Arabic and English ......................................................... 87

4.1 Average vowel durations in pre-dental and pre-fricative consonants and the ratios for each subject .......................................................... 113

4.2 Average mean duration in ms for short and long /æ/ in focused and unfocused pre-dental fricative environments ........................................ 119

4.3 Ratio of vowel duration in pre-voiceless to pre-voiced for all groups in the pre-stop focused context ...................... 127

4.4 Overall length of the vowel across the two vowel lengths for the focused and the unfocused contexts given in ms ...................................................... 128

4.5 Means, standard deviations, and ratios for long and short vowels in pre-stop environment for all groups in focused (A) and unfocused contexts (B) ................. 131

4.6 Means, standard deviations, and ratios for long and short vowels in pre-dental fricative environment for all groups in focused (A) and unfocused contexts (B) ........ 138

4.7 Overall means in ms for long and short /æ/ in focused and unfocused contexts for all the groups participating in the study ........................................... 139

4.8 Average focus-dependent vowel differences in ms for all the groups participating in the study ............................. 145

4.9 Means, standard deviations, and ratios for long and short vowels in pre-velar fricative environment for all groups in focused (A) and unfocused contexts (B) ............ 148

4.10 List of target words and carrier sentences used in the study .............................................................. 147
4.11 List of target words and carrier sentences used in the study ................................................................. 161
# TABLE OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Tracings of /l/ as produced by four native speakers of English, French, German, and Spanish</td>
<td>6</td>
</tr>
<tr>
<td>2.1</td>
<td>Effect of stress on intensity (a), frequency (b) and duration (c)</td>
<td>49</td>
</tr>
<tr>
<td>3.1</td>
<td>Subjects distributed according to years of experience and ratings by judges</td>
<td>84</td>
</tr>
<tr>
<td>3.2</td>
<td>A spectrogram showing the cursor at both end of the vowel</td>
<td>92</td>
</tr>
<tr>
<td>3.3</td>
<td>Spectrographic representation of the sentence &quot;say bad again&quot; produced a native speaker of American English</td>
<td>95</td>
</tr>
<tr>
<td>3.4</td>
<td>Position of the cursor at the end of the vowel when followed by /s, z, x, γ, h, and ʔ</td>
<td>96</td>
</tr>
<tr>
<td>3.5</td>
<td>Samples of /fæɪ/ produced by native speakers of Arabic</td>
<td>98</td>
</tr>
<tr>
<td>3.6</td>
<td>A set of programs produced by the bilinguals and shows no distinction between /γ/ and /x/</td>
<td>100</td>
</tr>
<tr>
<td>3.7</td>
<td>A spectrogram produced by a bilingual and showing laryngeal break in the middle of the vowel and the change in vowel quality</td>
<td>101</td>
</tr>
<tr>
<td>3.8</td>
<td>Hyper long /æ/ produced by two bilinguals</td>
<td>103</td>
</tr>
<tr>
<td>3.9</td>
<td>Aspiration follows /b/ in the word /bæːd/</td>
<td>104</td>
</tr>
<tr>
<td>4.1</td>
<td>Effect of voicing on vowel duration before stops as produced by monolingual Arabs for (a) short /æ/ in</td>
<td></td>
</tr>
</tbody>
</table>
focused context, (b) short /æ/ in unfocused context, 
(c) long /æ/ in focused context, and (d) long /æ/ in unfocused context ........................................... 108

4.2 Effect of voicing on vowel duration before fricatives as produced by monolingual Arabs for (a) short /æ/ in focused context, (b) short /æ/ in unfocused context, (c) long /æ/ in focused context, and (d) long /æ/ in unfocused context ........................................... 112

4.3 Effect of voicing on vowel duration before stops for long and short /æ/ in focused context as produced by (a) monolingual Arabs, (b) the beginning group (c) the intermediate group, and (d) the advanced group ....... 124

4.4 Effect of voicing on vowel duration before stops for long and short /æ/ in unfocused context as produced by (a) monolingual Arabs, (b) the beginning group (c) the intermediate group, and (d) the advanced group ....... 125

4.5 Effect of voicing on vowel duration before dental fricatives for long and short /æ/ in focused context as produced by (a) monolingual Arabs, (b) the beginning group (c) the intermediate group, and (d) the advanced group ....... 133

4.6 Effect of voicing on vowel duration before dental fricatives for long and short /æ/ in unfocused context as produced by (a) monolingual Arabs, (b) the beginning group (c) the intermediate group, and (d) the advanced group ....... 134

4.7 Effect of voicing on vowel duration before velar fricatives for long and short /æ/ in focused context as produced by (a) monolingual Arabs, (b) the beginning group (c) the intermediate group, and (d) the advanced group ....... 141

4.8 Effect of voicing on vowel duration before stops for long and short /æ/ in unfocused context as produced by (a) monolingual Arabs, (b) the beginning group (c) the intermediate group, and (d) the advanced group ....... 142

4.9 Role of reverse phonetic interference on the voicing-dependent vowel duration in English before stops as produced by (a) a group of monolingual Americans,
(b) the beginning bilingual group, (c) the intermediate bilingual group, and (d) the advanced bilingual group .......... 162

4.10 Role of reverse phonetic interference on the voicing-dependent vowel duration in English before fricatives as produced by (a) a group of monolingual Americans, (b) the beginning bilingual group, (c) the intermediate bilingual group, and (d) the advanced bilingual group .............................................................................. 166

4.11 Correlation between subjects’ ratios and their ratings, (a) for environment before stops and (a) before fricatives .... 169

4.12 Extra long /æ:/ in the word /bæːd/ (over 400 ms) .................. 173

4.13 Extra long /æ:/ before /γ/ .......................................................... 173

4.14 Extra long /æ:/ before /ŋ/ .......................................................... 174

4.15 Devoiced and long /γ/ ............................................................... 174

4.16 One of the variants of /γ/ ............................................................ 175

4.17 /γ/ is substituted by /g/ by the subject rated 35 ....................... 175

4.18 /sæːγ/ as produced by the subject rated 18 ............................... 176
Chapter I
Introduction

This dissertation addresses two main issues: the voicing-dependent vowel duration variation in Standard Arabic and how much of this variation is acquired by adult American students learning Arabic. In addition, it assesses any possible effect learning Arabic may have on the voicing effect in English among the American students of Arabic (i.e., any possible reverse interference, as it has begun to be known lately). Among the factors that affect the voicing effect and the acquisition process that are evaluated in this study are the student's level of advancement in Arabic, the role of focused versus unfocused environments, intrinsic vowel length, place of articulation, and manner of articulation of the following consonant. It is expected that a degree of phonetic interference will be carried over from the native language (English) into the target language (Arabic), particularly at the early stages of learning Arabic, and perhaps a smaller degree of reverse interference (i.e., from Arabic into English) might be detected, especially among the advanced students. In this chapter I briefly introduce the two major issues and pave the way for a detailed literature review in Chapter II.

1.0 Voicing-dependent vowel duration variation

One of the well-known phonetic features that has long been recognized and frequently studied is the vowel duration variation in pre-
obstruent environments. Research has shown that the magnitude of the variation as a result of voicing of the following consonant varies from one language to another. In English, a language characterized as having greater vowel duration variation than most languages, the ratio of vowel duration in pre-voiceless consonant environment to that of pre-voiced has been reported to be 76% (Heffner 1937, calculated from the reported results) 68.8% (House and Fairbanks 1953), 66.3% (Peterson and Lehiste 1960, calculated from their results), 61% (Chen 1970), 53.41% (Mack 1982), 62.5% in utterance-final position, 69.15% in mid utterance focused position, and 79.36% in mid utterance unfocused position (Laeufer 1992), etc. By contrast, corresponding ratios were found to be 87% in Spanish (Zimmerman and Sapon, 1958), 71% in German (Kohler and Kunzel 1978), 74.05% in French (Mack 1982), 87% also in French, 82% in Russian, 78% in Korean (Chen 1970), 97% in Saudi Arabic (Flege 1979), etc.

Arabic is one of the least studied languages with regard to this phenomenon. To date, there have at least been main four studies which addressed this phonetic feature. The first study was Flege's dissertation (1979), which two years later appeared in a short article (Flege and Port 1981), the second study is that of Port, Al-Ani, and Maeda (1980), the third is that of Mitelb (1984), and the fourth is that of Flege (1984). The first and the third studies investigated the duration of /æ:/ in monosyllabic stressed words in pre-stop consonant environment for Saudi Arabic and in pre-fricative consonant environment for Jordanian Arabic, respectively. Both studies concluded that voicing of the following consonant does not have a significant effect on the duration of /æ:/ . Flege found the ratio between pre-voiced and
pre-voiceless environments to be 97%, and Mitleb found only a 5-ms
difference between the two environments. Port, Al-Ani, and Maeda, on the
other hand, investigated the effect of voicing on short and long /æ/ in
trisyllabic Standard Arabic words produced at slow, neutral, and fast rates and
found the ratios at the neutral rate, the relevant rate for our purposes, to be
82% for short /æ/ and 92% for its long counterpart.

Since early 1980s it has been widely accepted among most phoneticians
and Arabists that Arabic does not show voicing-dependent vowel duration
(Flege and Port 1981, Flege 1984, Mitleb 1983, 1984,.)¹ This wide acceptance has
(1) placed Arabic at the opposite extreme of English, along with Czech and
Polish (Keating 1979) as the group of languages that do not show the voicing
effect, and (2) shattered the "universal" hypothesis which claims that vowel
duration variation (the strict version of the hypothesis) or some degree of the
variation (the subsequent modified version) is conditioned by inherent
physiological constraints and should be found in all languages.

It should be noted that the results reported for two of the three studies
that investigated Arabic, Polish, and Czech were based on one reading of the
data and in one context only.

In this study I first re-examine this widely accepted view using extensive
data from Standard Arabic (SA). To avoid the drawbacks of previous studies I
have (1) expanded the data to include all contrastive pairs of stops and
fricatives found in SA, (2) incorporated long and short /æ/, and (3) expanded
the context to include focused and non-focused sentential contexts. The data

¹For some reason, findings of Port, Al-Ani, and Maeda (1980) have been overlooked and
rarely mentioned as a proof for the existence of voicing effect in Arabic, though all the
studies conducted on Arabic came from Indiana University and in most cases with the
participation of Port.
were produced by six native speakers of Arabic whose knowledge of English is minimal. The results of this part of the study thus either confirm or refute early established results, but in any case establish a norm to which non-native bilingual productions (the second half of the study) can be compared. In the second portion of the study I examine the effect of voicing on the same data as produced by American students of Arabic. The purpose of this portion is to examine the degree to which adult American learners of Arabic would acquire the Arabic voicing-dependent vowel duration as a result of experience. Three groups of American students and professors, classified according to the level of competence in Arabic, have participated in the study. Results from their recordings assess the degree of acquisition. Third, I compare the production data of American bilinguals of English with those of a group of monolingual Americans. The purpose of this portion of the study is to examine the possibility of reverse interference from L2 (i.e., Arabic) into L1. It should be noted that reverse interference is a newly studied phenomenon and that there have been only a few recent studies that provided evidence to support the existence of reverse interference (Flege 1987, Jun and Cowie 1991, Peng 1993).

1.1 The acquisition of the phonetic system of a second language in adulthood and foreign accent

It is quite common, and perhaps natural, for second language (L2) adult learners to speak the target language with a distinct accent. This is clear in the production of /faɪf/ for "five" by many Arabs, /sɪŋk/ for "think" by many French speakers, /kʊd/ for "good" by many Finns, /hænt/ for "hand" by
many Germans, /lait/ for "right" by many Koreans, etc. It is indeed atypical to have L2 adult learners pronounce the target language in a way indistinguishable from its native speakers. It has been suggested that L2 learners seem never "capable of ridding themselves entirely of foreign accent" (Scovel 1969:245).2

One source of the foreign accent has been viewed as the differences between the native language of the learner (L1) and the target language (L2). The differences can be segmental (phonemic), subsegmental (phonetic), or suprasegmental, (Flege 1979, 81, Port and Flege 1981). At the segmental level, languages differ with regard to their phonemic inventories. Some segments that exist in a language are not found in another. For example, the Arabic pharyngeals cannot be found in English, German, French, or any modern Indo-European language. Similarly, French front high rounded vowels are not found in English, Arabic, Spanish, etc. At the subsegmental level, the differences are more subtle and harder to learn (Flege 1987). Many segments represented cross-linguistically by the same phonemic symbol are not identical in all phonetic aspects. These segments have been classified by phoneticians as "similar" (Flege, 1987). Consider Figure 1.1 below for illustration. The figure shows tracings of /l/ as produced by a native speaker of each of the four language represented in the figure in the context of the word film. The tracings are taken from X-ray pictures (Delattre 1964). As can be seen in the figure, the tongue tip is most anterior in French and Spanish

---

2It should not go unmentioned that we occasionally encounter or hear of a second language learner who acquired a L2 with no trace of accent. So far I have not come across any systematic studies that were designed and executed to examine the accents of these people nor have I ever come across any of these people. Therefore, such mention of free-accent adult learners of L2 remains anecdotal and unsubstantiated experimentally.
(dental), less so in German and least anterior in English (alveolar). English shows concave tongue shape in relation to the roof of the mouth, German shows a flat tongue, and Spanish and French show convex, yet not identical, shapes. Other examples show that /i/ has higher formant frequencies in Danish than in English (Malécot 1974), voiced implosive bilabial in Hausa differs from the voiced implosive bilabial in Kalabari in degree of voicing (Ladefoged 1980), French /u/ is more rounded than English /u/, English /æ/ has second formant frequency much higher than its counterpart in Standard Arabic, 1660 and 1250 Hz, respectively (Ladefoged 1982; Al-Ani 1970), etc.

![Tracings of /l/ as produced by four native speakers of English, French, German, and Spanish.](image)

It should be emphasized that phonetic (non-distinctive) differences are numerous and may exist in every cross-linguistic similarly-looking/sounding
phoneme. These differences are usually overlooked by adult L2 learners who tend to perceive similar sounds in terms of phonemes found in their L1 with less attention to the phonetic details (Flege 1981, 87; Major 1987).

Examples at a subsegmental level include the role of a feature as distinctive in one language and non-distinctive in another. For example, vocalic nasality is distinctive in French yet non-distinctive in many other languages. Aspiration of voiceless stops, to take another example, is non-distinctive in English but distinctive in some Korean dialects (Malécot 1974).

At the suprasegmental level, languages differ with regard to stress placement rules, prosodic structure, vowel quantity, etc. For example, word stress in Arabic is associated with the long vowel found in a word. Thus, Arab students learning English tend to associate incorrectly the stressed syllable in English with the one having the longest syllable nucleus. In the production of words like apologize and analyze they tend to place the primary stress on the final syllable instead of the second syllable for apologize and the first syllable for analyze. The final syllable in these verbs attracts the stress because it has a diphthong, the longest vowel in the word. Such suprasegmental differences across languages undoubtedly contribute to the perception of an accent.

Cross-linguistic phonetic differences can also result from combining segments into various sequences within the same language. For example, voiced stop consonants tend to devoice word finally in German, Polish, and Russian, but not in English or Arabic. The domain within which a phonetic segment or feature affects the quality of neighboring segments varies from one language or dialect to another. Port, Al-Ani, and Maeda (1980) report that while consonants affect the duration of the preceding vowels in English, they
affect the duration of the preceding and following vowels in Japanese. Younes (1991) and Davis (1991) reported that the domain of pharyngeal spreading in Arabic differs from one dialect to another. In one dialect the spread is blocked by a syllable boundary and in another it is blocked by some consonants or a word boundary. Also there seems to be an interaction between the direction of spreading (to the right versus to the left) and the blocking factors.

1.2 Accounts for the causes of foreign accent

One cannot help but wonder as to why most adult L₂ speakers seem to retain a foreign accent regardless of their experience in that language. A few theories/hypotheses have been offered to account for this "mystery". It was first suggested that L₂ learners cannot acquire the native pronunciation of a language after puberty (Penfield and Roberts 1966; Lenneberg 1967; Scovel 1969). According to this hypothesis, the acquisition of native pronunciation by L₂ learners is blocked by the neurological maturation and the establishment of cerebral lateralization for language functions. Thus, most adults who start learning a second language in adulthood would naturally speak that language with a distinct accent. Flege (1981) took issue with this hypothesis and offered one of his own called “Phonological Translation Hypothesis”. According to this hypothesis, adult learners approach L₂ with fully developed phonological system of L₁. They tend to translate the sounds of L₂ in terms of sounds found in L₁. Obviously, the two systems tend to have several sounds that are not identical in the two languages. Such
translation often leads to departure from the norms of the native language, thus creating a distorted acquisition of L2 sounds.

The major criticism of this hypothesis is its incapability of accounting for the fact that most children, who have already developed a full phonological system of L1, still do acquire L2 phonological system in a manner indistinguishable from that of the native speakers.

An alternative view is that of Natural Phonology. According to Donegan (1978), adult speakers of L2 apply the phonological processes that have been developed and modified to describe the native language. The phonological system of any native language, according to this approach, represents the residue of the processes that have gone unsuppressed or altered to fit that language. Applying these processes to L2 would certainly lead to the production of L2 with features from L1. Students of L2 will not be able to apply processes that are characteristic of L2 because these processes have been suppressed over the years and they are no more accessible to the learner.

It should be acknowledged that the first two views are more popular among second language educators than the last one. In fact, several studies have concluded that Flege's translation hypothesis applies to adult learners of L2. Flege himself (1987) found that learners of L2 tend to acquire "new" L2 sounds more authentically than sounds that have "similar" correspondents in L1. Similar results were reached by Cowie and Jun (1991) and Peng (1993). This behavior testifies to the fact that L2 learners perceive some sounds across the two languages as identical, thus overlooking the phonetic details that differentiate the two.
1.3 Experience and the acquisition of L2 phonological system

Knowing that most L2 learners tend to maintain a foreign accent, the question remains as to whether adult L2 learners would improve their accent (i.e., become more of a native-like) as a result of experience or adhere to phonetic features of L1 as they transfer them to L2? There have been two main lines of research to answer this question. The first line is perception-based. Native speakers would be asked to listen to samples of their language produced by adult non-natives and rate these productions according to their degree of accentedness. Experience, according to this approach, has been shown to have helped adult non-natives improve their accent (Asher and Garcia 1969; Snow and Hoefnagel-Hohle 1977, 78). It seems though that improvement has limits (Oyama 1976). The other line of research is production-based. Here researchers studied the acoustic properties of segments and the articulatory gestures for L2 as produced by L2 learners and compared their results with those of native speakers. Literature on this topic shows numerous studies conducted in this area using all kinds of instrumental methods that range from spectrographic display of the speech signal to modeling of articulatory gestures. Results have not always been consistent. Acquiring native-like sounds seems to depend on the feature under the study and the individual differences among the learners. For example, it has been shown that most L2 learners tend to recognize the overall vowel length of the target language and gradually approximate it as they advance in learning L2. By contrast, only some learners tend to succeed in acquiring the voicing-dependent vowel duration of L2 (Mack 1982, Laeufer 1992).
1.4 Organization of the dissertation

The remainder of this dissertation is organized as follows: Chapter II reviews the pertinent literature and the factors that contribute to vowel duration. In this chapter I provide the background relevant for the study described in the rest of this thesis. First, I review the literature available on voicing-dependent vowel duration as it exists in various languages and attempt to determine whether the voicing effect is universal (i.e., inherently conditioned) or a language-specific phenomenon. Second, I review the literature pertinent to L2 acquisition, and in particular as it applies to the phenomenon investigated in this study. Third, I provide all the linguistic and paralinguistic factors that help determine vowel length and contribute to the voicing effect on vowel duration. Finally, I present the hypothesis of the study reported in this dissertation.

Chapter III provides a detailed description of the experiment. It presents the data used in the study, describes the method used in collecting the data, the subjects, the recording procedure, various measurements, and the statistical analyses used in the study. It also elaborates on the difficulties associated with the segmentation process, especially those difficulties associated with identifying segment boundaries for the Arabic data as produced by L2 learners. I conclude the chapter by listing the segmentation rules adopted in the measurement in order to avoid producing inconsistent results.

Chapter IV presents the results, interpretation, and the discussion. The chapter has been divided into four parts. Part one presents the results of voicing effect on vowel duration in Arabic. It shows the voicing effect on
long and short /æ/ in pre-stop and pre-fricative consonants as they occur in focused and unfocused contexts. Part two describes the acquisition of voicing-dependent vowel duration of Arabic by American students and the possible interference effect of L₁ on L₂ with regard to the voicing effect. It compares the results of the voicing effect for the native speakers of Arabic with those of three groups of non-native American learners of Arabic classified as beginning, intermediate, and advanced to assess for possible interference effect. It attempts to determine whether advancement in learning Arabic would cause/help American students of Arabic acquire the Arabic-like voicing-dependent vowel duration or not. Part three shows the results of possible opposite interference; that is, interference from L₂ on L₁. Results in this section show whether American learners of Arabic would produce English data with voicing effect closer to Arabic as a result of studying Arabic. Part four describes some of the atypical acoustical features of the Arabic data as produced by the non-native learners of Arabic. Obviously, these features must be limited to those relevant to vowel duration and the effect of voicing.

Chapter V presents the conclusions of the study, the limitations on the study, offers ideas for further research, and finally it discusses the pedagogical implications that this study may have on teaching pronunciation to L₂ students.
Chapter II
Voicing-Dependent Vowel Duration: Literature Review

2.0 Introduction

The effect of voicing on vowel duration has been of interest to phoneticians since it was first experimentally proven to exist in late thirties (Heffner 1937, Rositzke 1939). Since then many production (i.e., acoustic and articulatory) and perceptual studies have been conducted on various languages in an attempt to understand the scope and the nature of the effect (references for these studies and their contribution are given in the next three sections). These studies investigated the effect of voicing on vowel duration in (1) monosyllabic versus bisyllabic versus polysyllabic words; (2) in connected speech versus word listings; (3) in stressed versus unstressed environments; (4) in various positions of the utterance/sentences; (5) in various accentual patterns; and (6) at various rates of speech. Some studies have gone further to investigate vowel duration variation (as a function of voicing) in whispered speech, in neutral environments (i.e., in pre-devoiced and pre-flapped environment), and in children's speech in an attempt to argue for or against the universality of the voicing effect. In the first section of this chapter I survey some of these studies and the development of interest in vowel duration variation as a function of voicing. The presentation starts with some of the production studies, given in chronological order. The contribution and drawbacks of each study will be highlighted. Next, a similar
description of some of the perceptual studies will be presented, and finally I
discuss the role of voicing on vowel duration in pre-neutralized
environments. In the second section, I will explore the role of the linguistic
and extralinguistic factors that condition vowel duration and have been
either overlooked or marginalized in the early studies on voicing-dependent
vowel duration. In the third section, I discuss the voicing-dependent studies
that have been conducted on Arabic and highlight the differences and
similarities between Arabic and English with regard to this point. Fourth, I
address the issue of whether vowel duration variation is universal or learned
and conclude the section with the questions to be addressed in this
dissertation.

2.1 Production studies

Voicing-dependent vowel duration variation was first investigated
instrumentally in American English in mid 1930s\(^3\). Heffner (1937)
investigated the role of quantity versus that of quality in distinguishing
vowels in American English. In that study he found out that vowels were
substantially longer before voiced consonants than before their voiceless
counterparts. His results are summarized in Table 2.1 below.

\(^3\) It should be noted that voicing-dependent vowel duration was observed and documented
long before 1930s. Among the first investigators were Sweet (1877) and Meyer (1903) for
English, both cited in Javkin, 1976 and Heffner 1937, Meyer (1904) for German, Meyer and
Gombocz (1909) for Hungarian, cited in Elert 1964, Gregoire (1911) for French, Metz (1914)
for Italian, Navarro Tomas (1916) for Spanish, and Einarsson (1927) for Icelandic, also cited
in Elert. In reviewing the literature relevant to vowel duration in this thesis I limit myself
to those studies that were conducted since mid 1930s and up.
Table 2.1. Average vowel duration before /t/ and /d/ as reported by Heffner in 1937.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Vowel duration in ms before /t/</th>
<th>Vowel duration in ms before /d/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/I/</td>
<td>150 ms</td>
<td>200 ms</td>
</tr>
<tr>
<td>/U/</td>
<td>150 ms</td>
<td>200 ms</td>
</tr>
<tr>
<td>/i/</td>
<td>210 ms</td>
<td>280 ms</td>
</tr>
<tr>
<td>/u/</td>
<td>230 ms</td>
<td>290 ms</td>
</tr>
</tbody>
</table>

Interestingly enough, Heffner found most of the lax vowels (viewed at that time by many phoneticians as short) to be longer before /d/ than the tense/long counterparts before /t/, though they may not seem to be so from the overall results as given in the table above.

...there is an almost complete overlapping of the values of the so-called short vowels before /d/ with those of the so-called long members of the pairs before /t/. For instance, of 106 cases of /I/ before /d/, 96 were as long as or longer than the minimum value of /i/ before /t/, and of 101 cases of /i/ before /t/, 99 were as short as or shorter than the maximum value of /I/ before /d/.” (Heffner 1937: 133)

Based on these results, Heffner concluded that quality, not quantity, is the distinctive factor in American English in vowel distinctions.

Rositzke (1939) investigated vowel duration in what he called General American English using five speakers and 1740 tokens of monosyllabic words. He too found that vowels were considerably longer before voiced consonants than before their voiceless counterparts.

"In all idioms heretofore investigated experimentally, a vowel before a (lenis) media is invariably longer than before a (fortis) tenuis.” (Rositzke 1939:104)
Similar to many more recent studies, he found that vowel duration variation is greater for tense vowels than for lax ones. More specifically, tense/long vowels (i, u, e, o, ə, ɒ, ɔ) were found to be 60% longer before voiced consonants than before voiceless ones, and short/lax vowels (I, U, ə, ɛ) were found to be 39% longer before voiced consonants than before voiceless ones.

The first attempt to directly investigate the effect of voicing on vowel duration was that of House and Fairbanks in 1953. In that study House and Fairbanks constructed a list of 72 bisyllabic nonsense words that involved 12 consonants in combination with six vowels. Bisyllabic words were used in the study. The first syllable for all the words consisted of the unstressed /hə/. The second syllable was the focus of the study and consisted of various stressed combinations of the consonants and the vowels used in the study. The consonants involved were six stops (p, b, t, d, k, and g), four fricatives (f, v, s, and z), and two nasals (m and n), and the vowels were /i/, /e/, /æ/, /ə/, /o/, /u/. Ten American subjects participated in the study. The goal of the study was not limited to investigating the effect of voicing on vowel duration, it also examined the effect of place and manner of articulation. The study confirmed the findings of Heffner and Rositzke in that vowels were significantly shorter before voiceless consonants than before their voiced cognates. In addition, it showed that vowel duration is subject to manner and place of articulation of the following consonant. Vowels were found to be longer before fricatives than before stops (203 ms before stops versus 239 ms before fricatives). Similarly, vowels were the shortest before velars (198 ms), a
little longer before labials (205 ms), and the longest before labiodental and postdental positions, 234 and 232 ms, respectively.

Zimmerman and Sapon (1958) were among the first scholars to examine the voicing effect in a language other than English to "determine" whether the effect is physiologically conditioned (conditioned henceforth) or "the relationship (between voicing and vowel duration) is primarily a matter of linguistic structure", that is, the effect can be interpreted on phonological grounds and may differ from one language to another. They examined the effect of voicing on vowel duration in Spanish as well as in English. They concluded that both languages exhibit the voicing effect, but to different degrees. Thus, while the mean difference between vowels in pre-voiceless versus pre-voiced consonants was 83.2 ms for English it was only 18.2 ms for Spanish. This finding led the authors to the conclusion that the voicing effect, or the degree of the effect, would most likely be a language-specific phenomenon. That is, the effect is not conditioned by physiological factors and therefore may not be found in all languages.

...while there may be a physiologically induced lengthening of a tonic vowel preceding a sonant, the size of the effect is determined by linguistic structure, i.e., the nature of the phonemic contrasts employed by a given language, and that the amount of durational variation is conditioned by the extent of sonant-surd contrasts in the language." (Zimmerman and Sapon 1958: 153)

In a fairly extensive study, Peterson and Lehiste (1960) re-examined vowel duration in American English using real English words. They recorded and measured over 1600 words, some of which were minimal pairs, as they occurred in a stressed position in an identical frame sentence. Their results
for the minimal pairs can be seen in the following table (taken from their article).

Table 2.2. Average vowel duration before voiced voiceless consonants as reported by Peterson and Lehiste (1961)

<table>
<thead>
<tr>
<th>The following consonant</th>
<th>Duration of short vowels</th>
<th>Duration of long vowels(^4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>138 ms</td>
<td>188 ms</td>
</tr>
<tr>
<td>b</td>
<td>203</td>
<td>307</td>
</tr>
<tr>
<td>t</td>
<td>147</td>
<td>210</td>
</tr>
<tr>
<td>d</td>
<td>206</td>
<td>318</td>
</tr>
<tr>
<td>k</td>
<td>145</td>
<td>200</td>
</tr>
<tr>
<td>g</td>
<td>243</td>
<td>314</td>
</tr>
<tr>
<td>f</td>
<td>192</td>
<td>261</td>
</tr>
<tr>
<td>v</td>
<td>321</td>
<td>374</td>
</tr>
<tr>
<td>θ</td>
<td>208</td>
<td>265</td>
</tr>
<tr>
<td>ř</td>
<td>260</td>
<td>381</td>
</tr>
<tr>
<td>s</td>
<td>199</td>
<td>269</td>
</tr>
<tr>
<td>z</td>
<td>262</td>
<td>390</td>
</tr>
<tr>
<td>č</td>
<td>145</td>
<td>198</td>
</tr>
<tr>
<td>́</td>
<td>191</td>
<td>300</td>
</tr>
</tbody>
</table>

Mean duration\(^a\) 204 ms 284 ms

\(^a\) (Excluded from the original table are nasals, liquids, š and Ž)

As can be seen in the table, the average mean difference between vowels in pre-voiced and pre-voiceless consonants is 80 ms, a finding that is not different from that of Zimmerman and Sapon for English. In terms of ratio, this difference translates into .72. That is, the average mean duration of vowels before voiceless consonants constitutes .72 the average mean duration of vowels before voiced consonants. This ratio is slightly different

\(^4\) The terms long and short are those of Peterson and Lehiste. They used them to refer to long and short syllable nucleus they discovered in their study.
from that reported by Peterson and Lehiste, namely, 2 : 3 ratio or 197 ms before voiceless and 297 ms before voiced consonants, due to the exclusion of some consonants that are irrelevant to the comparison of vowel duration in pre-voiced versus pre-voiceless consonants. The excluded consonants are nasals, liquids, and alveolar-palatal fricatives. The general finding, however, still shows a significant voicing effect. Also, it can be seen from the table that vowels are shorter before stops than before fricatives. This tendency seems to be common in English and many other languages, as discussed below. Peterson and Lehiste examined also the effect of the initial consonant on the following vowel and found it to be negligible.

Based on these findings and those of Zimmerman and Sapon, Peterson and Lehiste concluded that the effect of voicing on vowel duration is conditioned, but the extent of the effect is language-specific. In other words, all languages should show such an effect, but the degree of its influence differs from a language to another.

Delattre (1962) evaluated the effect of voicing on vowel duration as being either physiologically/articulatory conditioned or learned. In doing so, he looked at five works of those who studied the effect of voicing before him (Rositzke 1939, Heffner 1937, 1940, 1941, 1942, and 1943; Zimmerman and Sapon 1958, Peterson and Lehiste 1960, and House 1961). He criticized the results and conclusions of Zimmerman and Sapon as being basically inaccurate and misleading. First, Spanish, he claimed, did not have true voicing contrast in word final and in intervocalic positions. Thus, in comparing /p/ with /b/ in Spanish, a person is actually comparing /p/ with /β/, where /b/ has been "fricativated", to use Delattre's word. Of course, that
was exactly what Zimmerman and Sapon did (compared voiceless stops with fricativated sounds). Second, Zimmerman and Sapon compared vowels in monosyllabic words in English with vowels in bisyllabic words in Spanish. Obviously such a comparison is ill-founded since the lengthening/shortening effect of a vowel in a bisyllabic word is considerably less than in a monosyllabic word (this fact was proven later in the works of Sharf 1962, Klatt 1973, 1975, Harris and Umeda 1974). An accurate comparison should be either limited to monosyllabic or bisyllabic words in the two languages. Third, in calculating the average temporal difference between vowels in pre-voiceless and pre-voiced consonants, Zimmerman and Sapon included the vowels in pre-nasal and pre-liquid environments. A more accurate way would be to compare the two languages in one environment only. For example, if the authors limited their comparison to voiceless versus voiced (fricativated) stops, the average difference would have risen to 32 ms\textsuperscript{5}. It could be argued, based on this finding, that the average reduction in vowel duration in bisyllabic words in English would not be that much different from Spanish, and thus the effect of voicing on vowel reduction would be found "equally" significant in the two languages. Fourth, investigating the effect of the place of articulation of the following consonant on vowel duration would discover agreement between the two languages (English and Spanish) with regard to the direction of the effect, the further back in the mouth we move, the greater the variation becomes. This finding agrees with that of Peterson and Lehiste (1960).

\textsuperscript{5} The average 32 ms was calculated by averaging vowel durations before /p/ (93 ms), /t/ (104 ms), and /k/ (108 ms), the voiceless stops and averaging vowel duration before /β/ (130 ms), /ð/ (136 ms), and /γ/ (137 ms) all voiced/fricativated stops. The difference between the two average amounts to 32 ms.
Based on these arguments, Delattre concluded that Zimmerman and Sapon's conclusion that the effect of voicing can be learned should be rejected. Instead, he proposed that the effect should be conditioned and could be interpreted through the "Force of Articulation" hypothesis of Belasco (1953) (description of this hypothesis is given in section 2.6 below in this chapter).

Interested in the same idea of whether voicing effect is learned or conditioned, Chen (1970) examined the effect in four languages: English, French, Russian, and Korean. He contended that if the effect of voicing on vowel duration were to be found in all four languages, then the effect would be physiologically conditioned. Conversely, if it were not to be found in any of these languages, the opposite would be true, that is, the effect of voicing would be language-specific or learned. Chen's results indicate that voicing effect was common in all four languages but with varying degrees. Consider the following table, taken from Chen, for the actual average measurements and ratios.

Based on these findings and other findings reported for German (0.90), Spanish (0.86), and Norwegian (0.82), Chen concluded that the effect of voicing is conditioned, but the extent of it is learned.

---

6 This ratio was arrived at by F. A. Meyer (1903) and reported in Fintoft 1961.

7 This ratio was apparently calculated from Zimmerman and Sapon study on English and Spanish in 1958.

8 Fintoft reported this ratio in his study on duration of some vowel in Norwegian in 1961.
In view of the above observation (results as presented in the table above), we may tentatively conclude that (a) it is presumably a language-universal phenomenon that vowel duration varies as a function of the following consonant, and (b) the extent, however, to which an adjacent voiced or voiceless consonant affects its preceding vowel durationwise is determined by the language-specific phonological structure." (Chen 1970: 139)

Table 2.3 Average vowel length in ms in all four languages as reported in Chen (1970).

<table>
<thead>
<tr>
<th>Language</th>
<th>Before voiceless consonants</th>
<th>Before voiced consonants</th>
<th>Mean differences</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>146 ms</td>
<td>238</td>
<td>92</td>
<td>0.61</td>
</tr>
<tr>
<td>French</td>
<td>354 ms</td>
<td>407</td>
<td>53</td>
<td>0.87</td>
</tr>
<tr>
<td>Russian</td>
<td>131 ms</td>
<td>160</td>
<td>29</td>
<td>0.82</td>
</tr>
<tr>
<td>Korean</td>
<td>91  ms</td>
<td>119</td>
<td>28</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Unfortunately, Chen's study was not free of flaws. In many cases he compared (1) vowel duration in monosyllabic words in one language with that of bisyllabic words in another language, and (2) vowel duration in pre-consonantal clusters in one language with those of pre-single-consonant context in another. In one case he compared the duration of the vowel /e/ in a Russian word that had six consonants (i. e., /vnešnyj/) with the length of /e/ in three-consonant words in other languages. In addition, Chen failed to allow for the fact that Russian tends to devoice voiced stops word finally. Therefore, vowel duration in the Russian words /god/, /lug/, /gieb/, /glub’,/ and /l’od/ should not in principle be different from their voiceless
counterparts unless the voicing effect occurs at the underlying (phonological) level or as a reflection of spelling.

As a result of these flaws, it is hard to accept his results as accurate and valid. Further research becomes necessary to either verify or refute Chen's results.

In two controversial studies, Keating and Flege, both conducting their studies separately on three different languages in 1979, found the voicing effect to be virtually non-existent in Czech, Polish, and Saudi Arabic. Keating investigated the duration of /a/ in the Polish minimal pair rata and rada and found it to be 167.4 ms before /t/ and 169.5 ms before /d/. She also examined the effect of voicing on duration of /a/ and /a:/ before /t/ versus before /d/ in bisyllabic Czech words. She found duration to be 193.7 ms before /t/ and 204.2 ms before /d/. Statistical tests did not show significant difference between the two contexts in the two experiments. Obviously, these findings provide strong support for the claim that the effect of voicing cannot be conditioned. Both findings, however, could be criticized for (1) being taken from a single reading recorded by 24 subjects for Polish and 3 subjects for Czech, and (2) the fact that the measurements were taken from a word list reading. Both cases obviously have their limitations\(^9\) in acoustic studies.

---

\(^9\) Keating's studies do not follow the standard methodology of recording several readings by the same person for the same item(s), they lack the contextual variations in which the target words occur, they are limited to one pair, etc. Therefore, they appear to be more of pilot rather than full-fledged studies.
Similarly, Flege (1979) examined the effect of voicing on vowel duration in Saudi Arabic and found a minimal (non-significant) voicing effect\textsuperscript{10}. In this study Flege investigated the duration of long /æː/ in monosyllabic words in pre-stop environment. All words were read in a focused position within an identical frame sentence. The actual ratio between the duration of /æː/ in pre-voiceless stop to that of pre-stop was found to be .97.

The findings of these studies provided a strong claim that voicing effect is indisputably learned. This claim was questioned by Laeuefer (1992) who re-reexamined English and French. In her study, which was initially triggered by Mack (1982), Laeuefer argued that vowel duration variation as a function of voicing should be conditioned. Laeuefer realized that earlier phoneticians who studied differential vowel duration in French tended to either miss or overlook some language-specific variables that always masked the effect of voicing on vowel duration. These variables were (1) the prosodic structure of French (as it compares to that of English), (2) the release of the final voiceless consonant in French, (3) resyllabification, and (4) intrinsic vowel length. Laeuefer concluded that when these factors are controlled voicing effect in French would not be significantly different from that of English. Based on the findings of this study, it can be proposed that similar language-specific features might be found in Arabic, Polish, and Czech that mask or reduce the effect of voicing on vowel duration. Laeuefer concluded that all languages should show similar voicing effect on the temporal manifestation of the preceding vowel. These effects can be masked by language-specific features.

\textsuperscript{10} It should be noted that the two citations given here refer to the same study. The first citation refers to Flege's dissertation and the second one is an article that summarizes the findings of the dissertation.
that, if pinpointed, would resolve the debate over whether the phenomenon is learned or conditioned.

It is clear, then, that interest in the voicing effect on vowel duration is not new. Indeed, production literature on the topic is by no means limited to the studies cited here. Many other studies, mention of which at this point would be superfluous, have been conducted on various languages. It suffices to say that these studies have covered, besides the languages already cited here, Armenian, Assamese, Bengali, Danish, Dutch, German, Hindi, Hungarian, Icelandic, Italian, Japanese, Marathi, Norwegian, and Swedish and that all these languages show various levels of vowel duration variation that is dependent on voicing (Maddieson and Gandour 1976, Maddieson 1977).

It should be emphasized that despite all these studies, the nature and the scope of the effect of voicing on vowel duration is not fully understood yet. As can be detected from this introduction, each of the studies described above was built on the one(s) before and made its own contribution. Still, though, the problem of the effect being learned or conditioned is not resolved. Further studies are still needed, so that the study described in this dissertation becomes necessary.

Investigation of voicing-dependent vowel duration has not been limited to production. Many phoneticians have extended the investigation to perception. In particular, they have addressed the issue of whether duration of the vowel would be essential for identifying the voicing characteristics of the following consonant, thus highlighting the interaction and the dependency relationship between voicing and vowel length. The following
section provides a brief review, evaluation, and critique of some of the key studies done on perception.

2.2 Perception studies

Scholars dealing with speech synthesis and perception recognized as early as the mid 1950s that in identifying the voicing of a consonant they could not always rely on vocal cord vibration as the sole determinant for identification. Rather they realized that other factors, and particularly the duration of the preceding vowel, tend in many cases to outweigh glottal pulsing in determining whether the consonant would be perceived as voiced or voiceless. It was found that a voiced consonant in a vowel-consonant sequence would be perceived as voiceless if the preceding vowel had been shortened beyond a certain length. Inversely, a voiceless consonant would be perceived as voiced if the preceding vowel had been lengthened to a certain level. This dependency relationship between vowel duration and perception of voicing has been viewed as a reflection of speech production mechanisms by which voicing increases the duration of the preceding vowel (Raphael 1975, Javkin 1976).

Among the pioneers to realize and investigate this phenomenon was Denes in 1955. Using the synthesized pair (the) use /juːs/ and (to) use /juːz/, he looked primarily at the interactions between vowel and consonant lengths as cues for identifying the voicing of the consonant. Four vowel lengths (i.e., 50 ms, 100 ms, 150 ms, and 200 ms) combined with five voiceless consonant lengths (i.e., 50 ms, 100 ms, 150 ms, 200 ms, and 250 ms) were used in the experiment. The total number of combinations was twenty. Thirty-three
subjects listened to all combinations and identified the final fricative as /s/ or /z/. The general finding, and the one relevant to the study described here, was that more subjects perceived the consonant as voiced as the duration of the vowel was increased. For example, none of the subjects perceived /s/ of 150 ms length as /z/ when preceded by a 50 ms-long vowel. When vowel duration was increased to 200 ms and the consonant length was maintained at 150 ms, the number of subjects that perceived the consonant as /z/ jumped to 70%.

It should not be forgotten that this experiment was performed on an English-speaking population using English data. Therefore, this outcome may reflect the role voicing has on vowel duration in English only and not necessarily in the rest of the world's languages.

Raphael (1971) extended Denes work to include stops and clusters in word final position. He also used synthesized monosyllabic minimal pairs that varied with regard to the length of the steady state of the preceding vowel. Vowel duration ranged between 150 ms and 350 ms. Consonants appeared to have been kept at the same length for each minimal pair. All synthesized words were recorded and played back to twenty-five subjects for identification. The identification was done in a forced-choice format. Raphael found that "all the final consonants and clusters were perceived as voiceless when preceded by vowels of short duration and as voiced when preceded by vowels of long duration" (Raphael 1971: 1298). He also found vowel duration to be not only a sufficient cue for the perception of voicing of the following consonant but a necessary cue. Vowel duration was found to be more important as a perception cue than voicing of the final consonant or cluster.
Hogan and Rozsypal (1980), following Denes and Raphael's line of work, used a natural speech sample instead of a synthesized one. The data used in this study consisted of 24 Canadian English monosyllabic word pairs of CVC or CVCC structure. Each member of each pair was distinguished from the other member in the voicing of the final consonant or cluster. The primary goal of the study was that of Denes and Raphael, namely, to investigate the effect of vowel duration on the perception of the voicing of the following consonant or cluster. In addition, the researchers also looked at the contribution of three other factors, namely, the duration of the voice bar, the duration of silent closure of the following obstruent, and the duration of the burst and/or the friction of the following consonant, to the distinction of voicing. Of concern here is the contribution of the vowel duration only.

The data were first recorded, apparently in a word list format, by a female Canadian, and waveform spectra were made for all the recordings. Average measurements for each pair showed significant vowel variation in the two contexts, with pre-voiced vowels always longer than their pre-voiceless cognates. Four vowels consisting of two short and two long ones and six consonantal pairs that differed in voicing only were used in the study. Two tokens of each word were used for perception. Only words ending in a voiced obstruent or cluster were used.

Using a digital gating technique, the researcher examined the effect of vowel duration on the perception of voicing at five levels for each token. The levels ranged from the original length of the syllable nucleus, being the longest, to the shortest gated vowel duration. The levels were chosen independently for each word. The average shortest gated duration for
intrinsically short vowels was 64 ms, and for long vowels was 83 ms. In terms of percentage, these numbers stood for 30% of the original vowel duration and 45% of the average vowel duration in pre-voiceless context. Fourteen subjects listened to the readings (240 in total) and circled their choices in a two-alternative forced-choice format. For example, when listening to a form of the word "bud", subjects would either circle the choices "bud" or "but".

Interestingly enough, the results of this experiment did not totally support those of Denes and Raphael. Vowel duration did not appear to be the only determining factor in the perception of voicing. The correlation between vowel duration and the percentage of voicing identification was only .46 as opposed to .66 for voice bar duration, .66 for silent closure duration, and .69 for burst/friction duration. It was suggested that natural speech could be different from synthesized speech and therefore all the above factors combined tend to cue the voicing of the following consonant. It was also found that vowel duration was most effective in the perception of voicing in pre-fricative contexts. This effectiveness was attributed to the fact that vowel duration variation was the greatest before fricatives.

In conclusion, it could be argued that though vowel duration did not play the determinant role in this study, it still played an important role. The findings still support the generally well-accepted fact that voicing affects vowel duration.

Perceptual investigation was not restricted to the role of vowel duration in identifying voicing of the following consonants only. In some cases, researchers assessed vowel length as perceived in the context of pre-voiced
versus in pre-voiceless environments. Javkin (1976), following Lisker (1974), and in an attempt to provide a reliable account for the effect of voicing, rejected all the physiological/production accounts given up to that moment and offered a perceptual account instead. He proposed that the continuation of voicing (in the production of the vowel) into the following voiced consonant may make the vowel appear to the listener longer than when voicing is abruptly stopped in the context of a voiceless consonant. The perception of a longer vowel in pre-voiced consonant tends to eventually carry on to production. Thus, people end up producing greater vowel durations before voiced obstruents than before voiceless ones.

To substantiate this account, Javkin reviewed a number of production-based proposals and highlighted the weaknesses of each. Then he described three perceptual experiments that he conducted in this study; two of which (the first and the third) were deemed relevant to the study described in this dissertation. He used in the first experiment the synthesized words "his" /hlz/, and "hiss" /hIs/, and used actual production of these words by the subjects who participated in his work in the third experiment.

The duration of the initial /h/ and the final sibilants were kept constant while four vowel durations were given for each word in the first and second experiments. These lengths were 95 ms, 125 ms, 155 ms, and 185 ms. Transitions from the vowel to the following consonant were included in these lengths. Subjects heard a stimulus (i.e., /hIs/, /hlz/, or /hæːz/) followed by a tone given at 250 Hz. The tone length could be adjusted by the listeners to match that of the vowel in the stimulus and then it could be recorded by pushing a button on a computer. Results of the first experiment
showed average durations for all the tones to be 232 ms before /s/ and 258 before /z/, clearly showing that speakers of English perceive vowels as longer before voiced consonants than before their voiced cognates.

In the third experiment subjects produced and recorded the stimuli /hls/ and /hIz/ and then adjusted the length of the tone that followed to match the length of their self-produced vowel. Results of this experiment showed the average duration of tones that matched /I/ before /s/ to be 255 ms and those that matched /I/ before /z/ to be 340 ms, a difference of statistical significance at $P = 0.01$.

Regardless of the soundness of Javkin's proposal, both experiments provided strong evidence showing that vowels are perceived as longer before voiced obstruents than before voiceless ones.

In conclusion, perceptual evidence, unlike production evidence, lends strong and indisputable support to the lengthening effect of voicing on vowel duration. Greater length has been shown to be essential for the identification of voicing characteristic of the following consonant. In synthesized speech, it has proven to be more important than voicing of the following consonant.

The last two sections have shown that induced-voicing vowel duration has attracted the attention of many phoneticians over the last five decades. There is obviously a general agreement among most of these phoneticians that voicing of the following obstruent tends to lengthen the preceding vowel. The extent of that lengthening effect remains a point of dispute. Production-based studies have shown a minimal effect for Arabic, Czech, and Polish and a significant effect, with varying magnitudes though, in the rest of languages examined. Perceptual studies, on the other hand, agreed
unanimously that (1) vowels are longer before voiced consonants than before their corresponding voiceless ones, and (2) vowel length is necessary for identifying the voicing characteristics of the following consonant. Studies done on neutralized environments also showed a significant voicing effect for English, German, and Polish, but not for Catalan (Sharf 1962, 1964, Naeser 1970, Sheldon 1973, Fox and Terbeek 1977, Port et al. 1981, Krause 1982, Slowiaczek and Szymanska 1982, O'Dell and Port 1983, Charles-Luce 1985, Slowiaczek and Dinnsen 1985, Laeuffer 1992). The variation in the results for the various languages and within the same language, and the lack of agreement on the magnitude of the voicing effect, especially in the production studies, has been attributed to lack of control over many confounding factors that tend to decrease or increase the voicing effect. Understanding these factors and ultimately eliminating them is essential for the reliability of the results. In the following section I discuss some of these factors and elaborate on their role in affecting vowel duration and the voicing effect. The section starts with a general background that highlights some of the problems associated with some of the early studies, and differentiates vowels according to their phonemic and phonetic length. Next, I briefly discuss some of the confounding factors that have been reported as contributing significantly to vowel duration.

2.3 Conditioning factors that affect vowel duration and confound the role of voicing on vowel duration

A survey of the studies done on voicing-dependent duration in English reveals considerable variability in the results that range from as
low as a ratio of .53 (Mack 1982) to the absence of the effect in certain contexts (Klatt 1975; Crystal and House 1982, 1988). In other words, voicing caused a difference in Mack's study by which the average length of the vowels before voiceless consonants constituted only 53% their average length before voiced consonants. In Crystal and House's study, on the other hand, voicing did not show any significant effect. Other studies have shown ratios that lie between these two extremes. Thus Zimmerman and Sapon (1958) reported a ratio of .63, Peterson and Lehiste (1960) reported a ratio of .663, Chen (1970) reported a ratio of .61, Mitlleb (1984) reported a ratio of .80, Crystal and House (1982) reported no difference in connected speech, Laeuffer (1992) reported a ratio of .82 in pre-fricative environment for mid-utterance unfocused position, etc. This variability in the results raises some questions, three of which are directly relevant to the study of this thesis: What causes the variability? How does this variability bear on cross-linguistic studies? How does this variability bear on the universality claim of the voicing effect? The answer to the first question lies in the fact that there are confounding factors that tend to either increase or decrease vowel duration in general and affect the magnitude of the voicing effect in particular. These factors include, but are not limited to, the characteristics of the adjacent (mainly following) consonants, inherent vowel duration, mode of speech (reading data in a list versus in connected speech), word length, position of test word in the sentence or utterance, stress placement, rate of speech, the prosodic structure of the utterance, syllable structure, and the psychological state of the speaker (Klatt, 1971, 1973, 1975; Harris

In response to the second question, these factors should be controlled in any cross linguistic study. As a matter of fact, many of them have been overlooked in previous studies. For example, equating the duration of vowels in monosyllabic English words with that of bisyllabic Spanish words represents this practice (Zimmerman and Sapon, 1958). Other examples include equating differential vowel duration in pre-contrastive environments in English with that of pre-neutral environments in Russian (Chen 1970), equating vowel duration in pre-stop English environments with that of pre-fricative Spanish environments (Zimmerman and Sapon, 1958), ignoring the differences in intrinsic vowel length between English and French, the prosodic structure of the two languages, and the syllable structure of the words used in the study (Mack 1982), etc. Examples of this kind are numerous in the literature. Lack of control over these factors has often yielded conflicting and incompatible results. These results, which did not always account for all the different interactions among the various uncontrolled factors that influence vowel duration, have naturally prompted frequent, and usually more comprehensive, follow-up studies. The study reported in this thesis is one of these follow-up ones. In this section I elaborate on the role each of these factors plays in either enhancing or obscuring vowel duration. First, though, some brief background on the types of vowel duration is needed.
2.3.1 Types of vowel duration

A distinction between three types of vowel durations is essential: vowels can be (1) quantitatively contrastive (long versus short phonemically), (2) inherently long or short, and (3) environmentally-conditioned (i.e., their duration can be decreased or increased based in the surrounding environment, (phonetic distinction)).

2.3.1.1 Phonemic vowel duration

Vowel length operates contrastively in some languages but not in others. For example, Arabic, Finnish, Korean, Japanese, and Estonian\(^1\) distinguish phonemically between long and short vowels, but English and Spanish, among many other languages, do not. This distinction is extremely important in cross-linguistic studies and cannot be overlooked, especially when the comparison is held between a language with phonemic distinction and another that lacks this distinction. It tends to place certain constraints on the coarticulatory effects in speech production. For example, long vowels followed by a voiceless consonant may not shorten beyond a certain limit in order not to overlap with their corresponding short cognates (Klatt 1975; Laeufer 1992).

2.3.1.2 Intrinsic vowel duration.

Intrinsic duration refers to the length provided by the quality of the vowel itself. Place and degree of oral constriction of the vowel are the primary factors that affect intrinsic vowel duration. There is ample evidence

---

\(^1\) Lehiste (1970) reports that Estonian has actually three vowel duration distinctions that she labels as short, long, and hyper long.
from various languages which shows that high vowels are shorter than low ones when all conditioning factors are kept equal. Studies done on English (House and Fairbanks, 1953; Peterson and Lehiste, 1960; House, 1961, Sharf 1964; Klatt 1975; Crystal and House 1982, 1988; Luce and Charles-Luce 1985, Fourakis, 1991), on Danish (Fischer-Jørgensen, 1964), on French (O'Shanghnessy 1981), and on Swedish (Elert, 1964) support this view. Results from Luce and Charles-Luce (1985) show the average mean duration for English /a/ and /i/ to be 179 and 149 ms, respectively. Results from Crystal and House (1988) found the average duration for /i/ and /u/ in connected speech to be 108 ms and that of /æ/ and /a/ to be 132 ms. Similarly, Elert (1964) reported mean duration for the high Swedish vowels /i/, /y/, and /o/ across 89 occurrences to be 140 ms, while the mean duration for the low vowels /å/, /œ/, and /a/ to be 164 ms. O'Shanghnessy (1981) reports that vocalic duration in French tends to vary inversely with vowel height. Similar results are found in Peterson and Lehiste's study for English (1960)\textsuperscript{12}.

It has been suggested that the differences in intrinsic vowel length are physiologically conditioned. Low vowels require greater articulatory movement than high ones, and naturally more time is needed for the movements to be executed (Fischer-Jørgensen, 1964). Lindblom (1967) attributes the inherent length of the low vowels to lowering of the jaw, a

---

\textsuperscript{12} It might be of relevance to note at this point that these differences most probably exceed the \textit{difference limens} which constitutes the minimal duration difference that can be picked up by the human auditory system. This degree of difference obviously highlights the importance of the vocalic quality in determining the intrinsic duration of the vowel (For more on difference limens, see Lehiste 1970: 11-16).
characteristic absent in the production of high vowels. Lehiste (1970) states the differences "constitute a phonetic universal".

In addition, some languages like English distinguish between a set of long vowels (referred to commonly as tense) and a set of short (lax) vowels. Acoustic studies have consistently shown that /ɪ, ɛ, "wedge", ʊ/ are shorter than the rest of the vowels (with the exception of schwa) in English (Peterson and Lehiste 1960; Sharf 1964; Klatt 1975). However, to the average native speaker of American English it is quality, not quantity, that distinguishes vowels from each other.

2.3.1.3 Environmentally-conditioned vowel duration

As shown in the first three sections of this chapter, vowel duration in continuous speech has shown to be affected by the surrounding linguistic environment. Many factors such as the characteristics of the surrounding consonants, the syllable structure, tempo, stress placement, position within utterance, utterance-length, be it a word, a phrase, a sentence, or a longer discourse, etc., tend to affect vowel duration. These factors also tend to affect voicing-dependent vowel duration. In this section a list of these factors and a discussion of their role on vowel duration and the voicing effect on vowel duration is presented.

2.3.1.3.1 Effect of adjacent consonants

Production and perception studies conducted on various languages have shown that POA and MOA of the following consonant tend to affect vowel and voicing-dependent vowel durations (Fairbanks and House 1953; Maack

The first scholar to investigate directly the effect of POA of the following consonant on vowel duration was Maack (1953). He conducted his study on German and concluded that the effect is a result of interaction between the POA of the consonant and place of constriction for the vowel. Front vowels are generally shorter before dentals than before labials and velars. On the other hand, back vowels are shorter before velars than before labials and dentals. This conclusion motivated Fischer-Jørgensen (1964) to investigate the same phenomenon in Danish. She concluded that vowel duration depended on the distance and time consumed by the articulators (mostly the tongue) to move from the vowel position to the position of the following consonant, and on the mobility of the muscles involved in the production of the VC sequence. The greater the movement is and the more time it takes, the longer the preceding vowel becomes. Her conclusion was based on the fact that vowels followed by /b/ were shorter than vowels followed by /g/, which, in turn, were shorter than vowels followed by /d/, especially when /d/ is preceded by a back vowel. In the production of /Vb/, it was argued that two independent articulators (the tongue and lips) are involved. Therefore, the articulatory movement from the vowel position to the following /b/ takes practically no time at all, and consequently the vowels are shortest in this context. In the production of /Vd/, as opposed to /Vg/, where V stands for a front vowel, the apical part of the tongue is much more mobile than the back portion, and, as expected, the vowels are shorter. However, in cases
where the vowel was /u/, its duration in /ug/ was shorter than that of /u/ in /ud/. It is believed that the relative proximity with regard to POA for both /u/ and /g/ compensates for the sluggish movement of the back portion of the tongue. It should not go unnoticed that these findings do not agree totally with Maack's and assume that vowels are always the shortest before labials.

Peterson and Lehiste reached the following conclusions in their investigation of English (1960).

Table 2.4. Average vowel duration as a function of place of articulation of the following consonant as reported by Peterson and Lehiste (1960)

<table>
<thead>
<tr>
<th>Consonant</th>
<th>Short Vowels</th>
<th>Long Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>138 msec</td>
<td>188 msec</td>
</tr>
<tr>
<td>t</td>
<td>147</td>
<td>210</td>
</tr>
<tr>
<td>k</td>
<td>145</td>
<td>200</td>
</tr>
<tr>
<td>b</td>
<td>203</td>
<td>307</td>
</tr>
<tr>
<td>d</td>
<td>206</td>
<td>318</td>
</tr>
<tr>
<td>g</td>
<td>243</td>
<td>314</td>
</tr>
</tbody>
</table>

It is obvious from these findings that vowels are shorter before labials than before dentals or velars. Thus, they lend some support to those of Fischer-Jørgensen with regard to this point. In one case, however, the difference is minimal, and perhaps statistically insignificant (203 before /b/ for short vowels as opposed to 206 before /d/). However, it is not possible based these findings alone, to refute or support the rest of Fischer-Jørgensen's conclusions which associate the degree of consonantal influence with the
range of the articulatory movement. The table shows means taken for the front and back vowels combined; and, as a result, the interaction between the mobile front portion of the tongue with front vowels on one hand, and the less mobile dorsum portion with back vowels, on the other, cannot be factored out.

Zimmerman and Sapon (1958) reported the following vowel duration findings for Spanish: 93 ms before /p/, 104 ms before /t/, and 108 ms before /k/. They concluded that there was a positive correlation between vowel duration increase and moving POA further back in the mouth. This conclusion seems to be very general and overlooks the interactions manifested in the other studies (i.e., the mobility and sluggishness of various portions of the tongue).

In a more recent study, Crystal and House (1988) reached similar conclusions to those of Maack’s. They found that back vowels were longer than their corresponding front ones before labial and dental consonants and shorter before velars. Their findings are given in Table 2.5 below for illustration.

In contrast to these views, some phoneticians reached the conclusion that POA has no effect on the duration of the vowel.

"The place of consonantal production is shown to have a negligible influence on the duration of the vowel.” (House 1961: 372)

In sum, it is apparent that there is no unanimity among phoneticians as to whether POA plays a major role or not. However, two trends of influence have attracted attention: (1) the portions of the articulators involved in
speech production, and (2) the closeness of POA of the consonant to the place of constriction of the following vowel. Thus, both trends should be taken into account when investigating conditioned vowel duration.

Table 2.5. Average vowel duration as a function of place of articulation reported by Crystal and House (1988)

<table>
<thead>
<tr>
<th>POA</th>
<th>Average duration of back vowels</th>
<th>Average duration of front vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consonants in word-final position</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labials [p,b,m]</td>
<td>128 ms</td>
<td>116 ms</td>
</tr>
<tr>
<td>Dentals [t,d,n]</td>
<td>125 ms</td>
<td>84 ms</td>
</tr>
<tr>
<td>Velars [k,n,n]</td>
<td>67 ms</td>
<td>92 ms</td>
</tr>
<tr>
<td>Consonants in word-final nonprepausal position</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labials</td>
<td>122 ms</td>
<td>100 ms</td>
</tr>
<tr>
<td>Dentals</td>
<td>105</td>
<td>71</td>
</tr>
<tr>
<td>Velars</td>
<td>65</td>
<td>87</td>
</tr>
<tr>
<td>Consonants in word-final pre-pausal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labial</td>
<td>158</td>
<td>169</td>
</tr>
<tr>
<td>Dental</td>
<td>187</td>
<td>145</td>
</tr>
<tr>
<td>Velars</td>
<td>...</td>
<td>110</td>
</tr>
</tbody>
</table>

Attempts aimed at establishing the influence of the preceding consonants on vowel duration concluded that the influence is often negligible and its study is problematic, especially when the preceding consonant is a stop (Fischer-Jørgensen 1964; Lehiste 1970; Umeda 1975; O'Shanghnessy 1981).\(^{13}\)

\(^{13}\) A few exceptions to this commonly accepted effect have been reported in the literature. For example, Port et al. (1980) report that preceding consonants tend to affect vowel
Phoneticians often disagreed on whether the release of the preceding plosive is part of the vowel or not. Some considered vowel duration to include all the release portion (explosion and friction), some included the friction portion only, and others excluded both the explosion and friction portions.

The second contextual factor that influences vowel duration is the manner of articulation (MOA) of the following consonant. There is almost unanimous agreement across all the languages reviewed in this section that vowels are longer before fricatives than before stops. In particular, vowels fall in the following order (from the shortest to the longest) as a function of the manner of articulation of the following consonant: shortest before voiceless stops, longer before voiceless fricative, a little longer before voiced stops, and the longest before voiced fricatives\textsuperscript{14}. Results from House and Fairbanks (1953) showed the following vowel durations for English.

\textsuperscript{14} Discussion in this section is limited to stops and fricatives since they are the only manners of articulation included in this study.
Table 2.6. Average vowel duration as a function of manner of articulation for the following consonant reported by House and Fairbanks (1953)

<table>
<thead>
<tr>
<th>Duration (ms)</th>
<th>Consonant</th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td>/t/</td>
</tr>
<tr>
<td>197</td>
<td>/s/</td>
</tr>
<tr>
<td>258</td>
<td>/d/</td>
</tr>
<tr>
<td>291</td>
<td>/z/</td>
</tr>
</tbody>
</table>

Similar results were reported by Peterson and Lehiste (1960) where vowel duration averaged 147 ms before /t/, 199 ms before /s/, 206 ms before /d/, and 262 ms before /z/. House's figures (1961) showed that vowel duration seems to average 160 ms before /t/, 200 ms before /s/, 260 ms before /d/, and 325 ms before /z/. The same order of influence has been observed in several more recent studies (Umeda 1975, Klatt 1975, Crystal and House 1988). Further support for English comes from whispered speech. Sharf (1964) states that "vowels preceding fricatives are longer on the average than those preceding stops in normal speech by 41 ms and in whispered speech by 40 ms". (p. 91)

Elert (1964) found that vowels in Swedish were 13 ms longer before /d/ than before /t/, and 14 ms longer before /s/ than before /t/. Lehiste (1970), citing the results of Navorro Tomás (1916), reported that vowels in Spanish are 1.13 times longer before voiceless fricatives than before voiceless stops. These vowels are also 1.27 times longer before voiced fricatives than before voiceless stops. Laeufer (1992), in reviewing the effect of voicing on vowel

---

15 These figures represent approximations calculated from the graphs given in House's study.
duration in the various contexts, states: "Voiced fricatives are commonly described by French linguists as having the greatest lengthening influence on preceding vowels and the voicing effect is stronger in the context of fricatives than stops." (p. 412)

Laeufer's statement summarizes the importance of controlling for MOA in conducting cross-linguistic comparisons; the magnitude of the voicing effect on vowel duration tends to be greater before fricatives than before stops. And as expected, the results of her study came out in support of this view. Table 2.7 below summarizes these results.

Table 2.7 Average ratios of vowel duration as a function of manner of articulation of the following consonant as reported by Laeufer (1992)

<table>
<thead>
<tr>
<th>Language</th>
<th>Ratio before stops</th>
<th>Ratio before fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>.74</td>
<td>.65</td>
</tr>
<tr>
<td>English</td>
<td>.64</td>
<td>.61</td>
</tr>
</tbody>
</table>

b. Medial position focused (target word occurs in middle of utterance and carries the main accent in the sentence)

<table>
<thead>
<tr>
<th>Language</th>
<th>Ratio before stops</th>
<th>Ratio before fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>.78</td>
<td>.70</td>
</tr>
<tr>
<td>English</td>
<td>.72</td>
<td>.68</td>
</tr>
</tbody>
</table>

c. Medial unfocused position

<table>
<thead>
<tr>
<th>Language</th>
<th>Ratio before stops</th>
<th>Ratio before fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>.91</td>
<td>.86</td>
</tr>
<tr>
<td>English</td>
<td>.82</td>
<td>.76</td>
</tr>
</tbody>
</table>
It is clear from these results that the ratios of vowel duration before voiceless fricatives to those of pre-voiced fricatives are smaller (i.e., showing greater variation) than their corresponding pre-stop ratios. Similar results can be found in Belasco (1953); House and Fairbanks (1953), Peterson and Lehiste (1960); Raphael (1972), Klatt (1975); and Summers (1987).

2.3.1.3.2 Effects of stress

The presence or absence of stress is one the major factors that tends to influence vowel duration. Evidence from various languages shows that stressed syllables are generally longer than their corresponding unstressed ones. This statement should not be taken to imply that duration increases as a result of stressing, or decrease as a result of lack of stress, is identical cross-linguistically. Phonologists and phoneticians agree that there are two types of vowel reduction that result from lack of stress, phonological and phonetic. Phonological reduction, a process by which some unstressed vowels lose their identity and become a schwa in some contexts, is determined by the phonological structure of the language. In English, for example, the vowel /ɛ/ in the prefix tele-, pronounced as [tɛle], maintains its identity when stressed in telegraphy and becomes schwa when destressed in telegraphic. This form of reduction is a common characteristic of languages recognized as stress-timed. Lindblom (1963) states that "vowel reduction is said to be a characteristic feature of languages with heavy stress, such as, for instance, English and Swedish. In these languages, stress is known to be positively correlated with duration" (p. 1773). The same view is shared by Tiffany (1959), Lieberman (1960),
Oller (1971), Gay (1978), Miller (1981), Ladefoged (1982), and Fourakis (1986, 1991). Phonetic reduction, on the other hand, applies to unstressed vowels whose phonemic identity has been maintained. This form of reduction seems to apply to all languages, though with variable degrees of reduction magnitude. In English, for example, stressed vowels have been reported to be twice as long as their unstressed counterparts (Parmenter and Trivino 1935; Oller 1971). Klatt (1975) shows that the average vowel length for all unstressed vowels in English constitutes 53% the average length of all the vowels when stressed. Other phoneticians report that the effect is not as great. Fourakis (1991) found the average ratio of reduced vowels to that of the corresponding stressed ones to be .63 and .71 for vowels produced at normal and fast rate, respectively. In Standard Arabic, duration has been shown to be a major cue for stress, more so than the f0 (Al-Ani 1991). It is the phonetic reduction that we are interested in this study.

Another important distinction is necessary when talking about stress and duration. The domain of stress can be a lexical item or a sentence. Sentence stresses can be emphatic, intended to "highlight" certain word in that sentence. Klatt (1975) reports that vowels are generally longer in words with prominence than in other positions. Prominence words are those that convey new information. Vowels in these words are generally longer the first time they are mentioned. More repetition of these words tends to strip these words of prominence as these words will not be viewed as new any more (Coker, Umeda, and Brownman 1973). In this section I briefly review some of the studies that have been done on lexical and sentence stress, examine the role of stress on duration, and conclude
with a summary on the role of stress on voicing-dependent vowel
duration variation.

There have been numerous attempts to investigate the role of
duration in identifying stress, mostly in comparison to the roles of
intensity, formant structure, $f_0$, as well as the rate of speech. Though most
studies tend to show that all these factors contribute positively to the
stressed syllable, some of these factors, as emerges later in this section,
have greater weight than others in the various languages. For example, $f_0$
tends to have a greater weight in identifying a stressed syllable in English
than duration (Morton and Jassem 1965). In other languages, such as
Polish and Arabic, duration has a greater weight than in English (Morton,
Jassem, and Steffen-Batog 1968; Al-Ani 1991). It should be noted that a
debate over which cue plays a greater role is not immediately relevant to
the study described here. The important point to highlight is the changes
in duration as a function of stress.

Fry, in a series of classic studies conducted on English (1955, 58, and
65), found that duration is a major cue in the perception of stress. In his
first study (1955) he used pairs of English words that contrast in stress only.
Examples of these pairs are: permit versus permit, digest versus digést,
object versus objet, etc. Using a synthesizer, he provided five duration
ratios for the first and the second syllables that were approximately
adapted to actual human production. The ratios used for syllables 1 and 2
in the word permit were 0.5, 0.75, 1.00, 1.50, and 2.00 and for the word
contract were 0.2, 0.4, 0.6, 0.8, and 1.00.
Results showed that listeners increasingly identified the word as a noun as the ratio increased. This is a clear evidence that English speakers tend to associate vowel length with stress. In this experiment, Fry also showed that duration overrides intensity as a perceptual cue of stress. Similar results from his second experiment (1958) confirmed that duration is a major cue despite the fact that fundamental frequency tends to outweigh duration. In comparing duration with the formant structure of the vowel in his third experiment (1965), Fry found that duration overrides formant structure as a cue for stress.

Al-Ani (1991a and 1991b) concluded that stressed syllables produced by speakers of Standard Arabic are longer than unstressed ones. He performed two production experiments, one at the lexical level and the other at the phrase level. In both experiments he found that intensity and duration increase in stressed syllables while frequency remains unchanged. Figures 2.1a, b, and c below, taken from his second study, illustrate his conclusions. Each of these figures shows the production of the phrase /kitaːb kæbiːr/ "a big book" as produced by three speakers of Arabic. SHA is Iraqi, SAT is Jordanian, and AS is Saudi. Figure 2.1a shows the change in intensity as a result of stress, Figure 2.1b shows frequency changes, and Figure 2.1c shows the change in duration. The stress in this phrase falls on two syllables, namely, taab and biir. It is clear from the figures that intensity and duration have been considerably increased in these two syllables as a result of being stressed. By comparison, the frequency (Figure 2.1b) remains practically unchanged; all three subjects show close number of cycles for all syllables.
Figures 2. Effect of stress on intensity (a), frequency (b) and duration (c).
In two experiments by Morton and Jassem on English speakers (1965), and by Morton, Jassem, and Steffen-Batog (1968) on Polish speakers, duration was found to be a major cue for identification of stress too, more so for Polish than for English. The researchers used the following nonsense words /sisi/, /sɔsɔ/, and /sasa/, where the same vowel is used in both syllables in each word in order to control for any intrinsic vowel intensity in the two experiments. Three parameters, intensity, duration, and f0, were varied systematically. Both English and Polish speakers showed that f0 has greater effect than duration or intensity in identifying the stressed syllable. Variation in duration had the second greatest effect, and intensity was shown to have the least effect of the three.

Using a tape-slicing techniques, Westin, Buddenhagen, and Obrecht (1966) conducted a study on Swedish to assess the contribution of duration, f0, and intensity to the prominence of the syllable. The data consisted of words that have two meanings differentiated by stress placement. The first and second syllables were varied according to high versus low intensity, the difference being 6 dBs, frequency of either 122 or 144 Hz, and duration of either 195 or 300 msec. Results indicated that f0 has the greatest contribution. Both duration and intensity are still effective, but to a lesser degree.

More recent production studies on vowel reduction in unstressed syllables show that vowels are significantly shorter in unstressed syllables than in stressed ones within a sentence or in longer connected discourse (Klatt 1975; Fourakis 1986, 1991). Klatt (1975) measured the length of all the vowels of English in a passage that consisted of 13 sentences, 236
words, or 1200 segments. He found that the average vowel duration for all vowels in stressed environments to be 132 ms as opposed to 70 ms for unstressed environments. Thus, stressed vowels are almost twice their length in unstressed environments. This magnitude of difference is rather surprising since it applies to connected speech where sounds are far shorter than their idealized length\textsuperscript{16}. A comparison between vowel length as reported by Peterson and Lehiste (1960) and that of Klatt shows that vowel duration in connected speech is close to half that of vowel duration when produced in words in isolation (245 ms for vowels produced in CVC utterances versus 132 ms for vowels in connected speech). According to Klatt, vowels in connected speech should be more resistant to reduction than vowels produced in isolation due to the compressibility effect. That is, vowels are more compressed (shortened) in connected speech than when they are produced in isolation or in a frame sentence, and therefore they are unlikely to compress significantly further. Klatt also found that syllables with secondary stress to be only 5% shorter than the ones with primary stress.

By comparison, Fourakis (1986) investigated the length of Greek vowels in stressed and unstressed environments and found that unstressed syllables are 25% shorter than stressed ones. In a second experiment (1991), he investigated the effect of stress versus that of tempo on vowel duration in American English. The data consisted of nine monophthong vowels used in the contexts /h-d/ and /b-d/, produced at two rates of speech by eight native speakers each of Midwestern American

\textsuperscript{16} Miller (1981) identifies the idealized length (target) of the vowel as its length when produced in isolation (p. 42).
English in a focused and nonfocused environments. The results show that vowel durations in the nonfocused environment as produced at slow rate for the contexts /b-d/ and /h-d/ were 35% and 37%, respectively, shorter than those of the focused environment. On the other hand, vowels produced at the fast rate were 29% shorter in nonfocused environment than those in focused environment for both contexts /b-d/ and /h-d/. In fact, changes in duration have shown to have "slightly greater effect on vowel duration than changes in tempo". (p. 1825).

In conclusion, all the studies indicate that vowels are longer in stressed syllables than in unstressed ones. On the whole, fundamental frequency seems to provides the strongest evidence for the presence of stress in most languages, duration provides the second strongest evidence, and intensity provides the weakest evidence. In comparing the effects of stress and tempo, lack of stress seems to have greater shortening effect than tempo. Fourakis (1991) attributes this greater effect to the fact that the effect of tempo is global while that of stress is local. That is, tempo tends to shorten all the segments in the utterance as opposed to one syllable in the case of stress. The extent of reduction in one syllable is greater than the average reduction to all vowels in a sentence.

Stressed syllables have not proven to be only longer than their unstressed counterparts, but they tend to show greater voicing effect than the unstressed syllables. Laeufer (1992) reported that the voicing effect on the duration of the preceding vowels has dropped by 15% for French and by 10% for English in the unfocused context. To elaborate, the voicing effect for stressed syllables in mid utterance position was reported as
74.55% and 69.15 for English and French, respectively. These ratios jumped to 89.99% for French and 79.36% for English for the same syllables when the focus shifted to another syllable in the utterance.

2.3.1.3.3 Effects of tempo on vowel duration

Like the effect of stress, vowel duration can be increased or decreased based on the rate of speech. It is commonly believed that the faster we speak the shorter the speech segments become, though the shortening effect is not equally proportional across all the segments. Vowels tend to compress proportionally greater than consonants (Gay 1978, Port et al. 1980). Several studies have been carried out to examine the effect of fast speech production on vowel duration empirically. Lindblom (1963) examined the phenomenon of vowel reduction on eight vowels in Swedish at slow and fast rates. He concluded that vowels tend to centralize/neutralize as a result of fast speech. Gay (1978) compared vowel duration of American English as produced at two rates (slow and fast) in CVC utterances. He found that the average vowel length produced at fast rate constitutes 75% of its duration when produced at a normal rate. By comparison, consonant reduction in fast speech averaged only 10%. Similar results were found for Standard Arabic by Port et al. (1980). The overall ratio of stressed vowel duration produced at a fast rate to that of at a normal rate was .75. Short vowels showed a reduction of 12% only (ratio = .88) and long vowels 31% (ratio = .69). On the other hand, the consonant reduction averaged only 12%. Thus, according to these two studies, tempo seems to have the same effect on Standard Arabic and English.
Rate of speech seems to have the same effect in Greek. Fourakis (1986) reported that vowels in Greek shorten by 25% in fast speech. His research on English (1991), however, shows greater effect of tempo than what has been shown. He has reported an average vowel reduction of 29% as subjects change from slow to fast rate in the two contexts used in the study (i.e., /b-d/ and /h-d/).

Effect of fast speech rate in long texts was reported as less effective than in carrier sentences. Crystal and House (1982, 1988b) reported an overall vowel reduction of 16% and 10% in both studies, respectively. This should not be surprising since the magnitude of compressibility is limited in connected speech (Klatt 1975).

Research has shown that the faster the rate is the smaller the effect of voicing becomes (Harris and Umeda 1974; Port 1981). Port (1981) reports the difference in the duration of /i/ in pre-/b/ and pre-/p/ to have dropped from 16.5 ms to 5 ms as a result of fast rate. The average duration of /i/ was recorded as 121.5 ms before /b/ and 105 ms before /p/ at normal rate. At fast rate, the duration of /i/ has become 83 ms before /b/ and 78 ms before /p/.

Similar results were found for /I/ which was recorded as 80 ms before /b/ and 68.5 ms before /p/ at a normal rate and 61 ms before both /b/ and /p/ at a fast rate. The difference has disappeared completely for /I/ at the fast rate.

2.3.1.3.4 The effect of word length and syntactic boundaries on vowel duration

In addition to the adjacent consonants, stress, and tempo, a number of other contextual factors, perhaps not as immediate as the above-discussed
ones, tend to significantly affect vowel duration. These factors are word length, word boundaries, syntactic boundaries.

Vowels shorten as the number of syllables in a word increases (Sharf 1962; Lehiste 1972; Klatt 1973; Umeda and Harris 1974). Sharf (1962) investigated vowel length in mono- and bisyllable words using words like nap versus napping and nab versus nabbing. He concluded that the average vowel duration for all tense vowels in English was 222 ms in monosyllabic words and 139 ms in bisyllabic words. Similarly, Lehiste (1972) observed that vowels in base words such as stick, sleep, shade, and speed are longer than the same vowels in the derivatives sticky, sleepy, shady, and speedy, respectively, which, in turn, are longer than vowels in stickily, sleepily, shadily, and speedily, respectively. Lehiste concluded that the more syllables are added to the base the shorter all the segments in the base become, especially the vowels. Similar results were arrived at in an earlier study conducted by Tarnoczy (1965). Tarnoczy studied the impact of adding several suffixes, one at a time, to a base word in Hungarian on the duration of the base vowels. Results of his study are given in Table 2.8 below (reported in Lehiste 1970).

Klatt (1973) studied the interaction of word length and voicing characteristic of the following consonant on vowel length. His relevant conclusions are given in Table 2.9 below.

Results in the table confirm that (1) vowels in bisyllabic words are much shorter than they are in monosyllabic words, and (2) the magnitude of reduction correlates with the length of the vowel. Thus, the difference between vowels in pre-voiced environment as a function of word length is greater than that in pre-voiceless environment due to the relatively longer
vowels in pre-voiced environment (.66 versus .78). Also, the difference between vowel durations in monosyllabic words as a function of voicing is greater than that found in their bisyllabic counterparts due to the relatively longer vowels in these environments (.67 versus .79). Klatt attributes this differential variation to the compressibility effect (a brief description of this effect is given at the end of this section).

Table 2.8. Reduction in vowel length as more syllables are added to the word (reported in Lehiste 1970)

<table>
<thead>
<tr>
<th>The base word and the derivatives</th>
<th>Length of base vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>/tats/</td>
<td>210 msec</td>
</tr>
<tr>
<td>/tats+og/</td>
<td>180 msec</td>
</tr>
<tr>
<td>/tats+og+at/</td>
<td>140 msec</td>
</tr>
<tr>
<td>/tats+og+at+o:k/</td>
<td>120 msec</td>
</tr>
<tr>
<td>/tats+og+at+o:k+nak/</td>
<td>110 msec</td>
</tr>
</tbody>
</table>

Table 2.9. Average vowel durations in mono- and bisyllabic words as reported by Klatt (1973).

<table>
<thead>
<tr>
<th>Voicing</th>
<th>Monosyllabic words</th>
<th>Bisyllabic words</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>-voice</td>
<td>132 ms</td>
<td>103 ms</td>
<td>.78</td>
</tr>
<tr>
<td>+voice</td>
<td>198 ms</td>
<td>131 ms</td>
<td>.66</td>
</tr>
<tr>
<td>Ratio</td>
<td>.67</td>
<td>.79</td>
<td></td>
</tr>
</tbody>
</table>
Umeda and Harris (1974) compared the effect of word length on vowel duration in a carrier sentence and in connected speech. Their target word in the carrier sentence consisted of a monosyllable word to which they added progressively another three syllables; and for connected speech they had recorded a thirty-minute-long text. Vowel durations in words consisting of one, two, and three syllables were measured. They found vowel duration in monosyllabic words for one speaker, for example, to be 226 ms. The same vowel was reduced to 151 ms when a second syllable was added, then to 130 ms when a third syllable was added, and to 119 ms in a four-syllable word. By contrast, duration differences in connected speech were found much smaller at any word length (133 in monosyllables, 127 ms in bisyllables, and 127 ms in trisyllables). The authors concluded that it is not the number of syllables in a word that determines vowel duration as much as it is the prosodic structure of the carrier sentence. They do not, however, provide any explanation as to how the prosodic structure is different in the two modes of speech and how it affects the temporal aspect of the vowel.

Results from Umeda and Harris, Tarnoczy, and Lehide, are in general agreement with Klatt's model of vowel compressibility concerning the addition of more syllables to the base monosyllabic word. Klatt (1973) found that the degree of vowel compressibility is the greatest after the addition of the a second syllable to the monosyllabic word (30%). This degree of compressibility decreases gradually until vowel length reaches a minimum of 45% its inherent length then it stops. Inherent length, according to Klatt, refers to the length of the vowel in monosyllabic words before a voiced consonant as these words are produced in isolation.
Another major factor that affects vowel duration and the magnitude of voicing-dependent differential vowel duration is the position of word in the utterance. For example, vowels tend to get longer as a result of being at the end of a phrase or utterance than anywhere else within a phrase or an utterance. Numerous studies have examined this point (Fonagy and Magdics 1960; Geitenby 1965; Lehisute 1972; Lindblom and Rapp 1973; Klatt 1975; Umeda 1975; Huggins 1975; Cooper et al. 1981; Crystal and House 1982, 1988; Luce and Charles-Luce 1985; Rakard et al. 1977, 1987; Laeuffer 1992). Obviously, discussing everyone of these studies separately will yield unnecessary redundancy. Therefore, I limit the discussion to the most representative.

Huggins (1975) investigated the shortening effect of unstressed syllables on the duration of a stressed syllable in the same metrical foot in English. Using data that fit the metrical foot with regard to alternations of stressed and unstressed syllables, he found that stressed syllables maintain their length if the following unstressed syllable(s) is across a major syntactic unit such as a NP or VP, and get significantly shortened if the unstressed syllable(s) is within the syntactic phrase. He concluded that phrase boundaries have lengthening effect on the last syllable in the phrase. Similar results were found by Cooper et al (1977) who examined the same phenomenon in sentences like (1) and (2) below.

(1) The police kept Clinton till three o'clock.

(2) The police kept Clint until three o'clock.

The vowel in Clint in sentence (2) was found to be significantly longer than that in Clint in Sentence (1). They concluded that the shortening effect of the

17 A metrical foot in English is defined as a stretch of speech that consists of one stressed syllable followed by zero, one, or two unstressed syllables.
following unstressed syllable is blocked in sentence (2) because of the phrase boundary. Cooper et al. examined the effect of phrase boundary on vowel length across a variety of syntactic phrases boundaries including NP, VP, PP, etc. The conclusion they arrived at was consistent in that vowels are longer at phrase boundaries than their counterparts within a phrase.

Rakerd et al. (1987) viewed Cooper et al.'s study as having confounded the effect of word boundaries with that of the phrase boundaries. They pointed out that Clint in sentence (1) is part of a word that consists of two syllables (i.e., Clinton) whereas Clint in sentence (2) is a monosyllable word whose increased vocalic length may stem from word boundary rather than a phrase boundary, as Cooper et al. claimed. Therefore, they replicated Cooper's experiment using monosyllabic words within a phrase and at the end of phrase boundaries to assess whether vowel duration increase in words like Clint stems from a word or phrase boundary. The following is a sample of their data.

A. Data across major syntactic boundary
   1. His first date (a)roused some anxiety (for obvious reasons).
   2. That young duke (dis)armed his subjects (against the advice of his counselors).

B. Data within major syntactic boundaries
   3. The strong peach (de)light was unpleasant (but better than nothing).
   4. John must bike (a)round the block (because it is too far to walk).

In both groups (i.e., A and B) the italicized words are ones that have the target vowel. Group A shows the target word at the end of a NP and group B shows
the vowel within the NP and VP, respectively. This experiment, however, has its own shortcomings\textsuperscript{18}. The results were similar to those of Cooper et al.'s. Vowels were found to be longer in phrase final position than within the phrase. The actual difference between vowels in the two positions was found to be 45 ms. Rakerd et al., however, found that the interaction between the foot structure and the syntactic boundary to be statistically insignificant. Therefore, they concluded that this difference (i.e., 45 ms) must have resulted from word boundary, not from phrase boundary. Of course, such a conclusion is unexpected if we take into account the widely accepted assumption that phrase boundaries contribute to vowel length. Lehiste (1972), however, concluded that the effects of morpheme boundaries and syntactic boundaries cannot be separated from each other.

In comparing vowel length at the end of a sentence (in pre-pausal position) with that of the median\textsuperscript{19}, Klatt (1975) found that 8 vowel occurrences out of a total of 13 exceeded the median by the length of 1.4 times. The number of occurrences was increased by 1 when he reduced vowel length to 1.2 times that of the median. In exact numbers, the difference was translated into an average of 40 ms or 30\% longer than vowel duration average in mid-utterance position. In addition, he examined vowel length at the end of a NP, VP, before a conjunction, between N and PP, and before an embedded clause. He concluded that all these boundaries cause vowels to be longer than in mid-category position. For example, 11 occurrences out of 13 were reported to be 1.2 times longer than the median at the end of a NP and

\textsuperscript{18} The phonetic properties of the target vowel, for example, are not maintained in all the sentences used in the study.

\textsuperscript{19} Median in this experiment stands for the middle value for the lengths of a vowel spread in a text of connected speech.
VP (medians were calculated for all occurrences of a vowel in a passage consisting of 13 sentences). Similarly, 9 occurrences out of 11 that occurred before a conjunction were at least 1.2 times longer than the median.

Crystal and House (1988) found the average vowel length for all vowels in non-final-utterance position in connected speech to be 89 ms and 99 ms in pre-voiced and pre-voiceless environments, respectively. By contrast, this average jumped to 179 ms and 161 phrase finally in pre-voiced and pre-voiceless obstruents, respectively. Equally important, the effect of voicing on vowel duration was observed in pre-pausal position only. Other positions in this study on connected speech did not show any voicing effect.

Luce and Charles-Luce (1985) examined the impact of the position of the test word on vowel duration in what they called "controlled connected speech." Two contexts where the target words were once in a pre-pausal position and the other in mid-utterance position were used in this study. They found that vowels, on the average, were 69 ms longer in clause-final position than in mid-clause position. Similarly, the average voicing-dependent vowel duration variation was shown to have correlated positively with vowel length (i.e., 42 ms in mid-utterance position versus 68 ms in pre-pausal position).

Lauefer (1992) investigated the same phenomenon in English and French, but in mid- and final-utterance focused positions and in mid-utterance unfocused position. The following are the results of her study for the focused context in mid- and final-utterance position.20

---

20 Measurements for the medial unfocused position have been excluded due to lack of matching measurements for the phrase-final position.
Table 2.10 Average vowel durations in different positions in the sentence before voiced and voiceless consonants as reported by Laeufer (1992)

<table>
<thead>
<tr>
<th></th>
<th>Final position</th>
<th>Medial Position</th>
<th>Difference in ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+voice</td>
<td>306.65</td>
<td>213.15</td>
<td>93.5</td>
</tr>
<tr>
<td>-voice</td>
<td>195.23</td>
<td>153.52</td>
<td>41.71</td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+voice</td>
<td>371.03</td>
<td>232.15</td>
<td>138.12</td>
</tr>
<tr>
<td>-voice</td>
<td>227.13</td>
<td>157.69</td>
<td>69.44</td>
</tr>
<tr>
<td><strong>French</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+voice</td>
<td>207.22</td>
<td>171.84</td>
<td>35.38</td>
</tr>
<tr>
<td>-voice</td>
<td>154.23</td>
<td>133.8</td>
<td>20.43</td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+voice</td>
<td>222.63</td>
<td>159.75</td>
<td>62.88</td>
</tr>
<tr>
<td>-voice</td>
<td>145.75</td>
<td>112.24</td>
<td>33.51</td>
</tr>
</tbody>
</table>

It is clear from this table that vowels in both languages are longer at the end of the phrase than in the middle of it when all other factors are kept equal. Two observations about these results cannot be overlooked. First, the difference is consistently greater in English than in French. This difference can be accounted for by the fact that vowels in English are on the average intrinsically longer than vowels in French. Second, the differences are consistently greater before fricatives than before stops confirming that MOA plays a major role in determining vowel duration and the effect of voicing on vowel duration as well.
In conclusion, it has been demonstrated that word length, and phrase boundaries contribute significantly to vowel duration and its variability as a function of voicing. Therefore, any attempt to investigate voicing-dependent vowel duration cannot overlook any of these factors. Phrase boundary has been deemed to have comparable effect on vowel length to that of reading words in isolation; that is, words occurring in phrase-final position have the same length as if they were read in isolation (Klatt, 1976: 1211).

2.3.1.3.5 The effect of intonation on vowel duration

The position of the test word in the prosodic structure has shown to either intensify or reduce vowel duration and voicing-dependent differential vowel duration. Syllables occurring in focused positions (holding the prominent part of the utterance) tend to have longer vowels than those occurring in unfocused positions. Laeufer (1992) compared the vowel duration in focused and unfocused positions in both English and French. Two focused positions were used in the study, one at the end of the intonational phrase in both languages, which happened to be also the end of the carrier sentence used in the study, and the other was in phrase medial focused positions. The unfocused position was in mid-phrase position for both languages too. Results showed the mean vowel duration of English in pre-voiced and pre-voiceless obstruents in final focused context to be 213.83 ms and 150.6 ms, respectively. For French, mean vowel durations in the same context were 186.89 ms and 143.24 ms. In medial focused context the average means for English were 221.71 ms and 153.31 in pre-voiced and pre-voiceless contexts, respectively, and for French they
were 166.66 ms and 124.25 ms. By contrast, mean vowel durations in
medial unfocused positions were significantly shorter. For English, the
means were 164.21 ms and 130.32 ms before voiced and voiceless
consonants, respectively, and for French the were 112.32 and 97.06,
respectively.

Close examination of these means reveals that the effect of voicing on
vowel duration diminishes significantly in the unfocused position in both
languages. For example, the pre-voiceless/voiced ratio has been increased
for English from 62.5% in the final focused position to 79% in the
unfocused medial position. In French, the effect has been reduced to 10%
only, an outcome that has been shown to be statistically insignificant.

2.4 Voicing effect in Arabic

The voicing effect on vowel duration has not been extensively
(satisfactorily) investigated for Arabic. There have been three production-
based and one perception-based studies that attempted to investigate this
phenomenon (Flege 1979, 84, Port, Al-Ani, and Maeda 1980, Port and Flege
1981, Mittleb 1983, 84). In two of these studies (Port et al.'s and Mittleb's 1984)
the examination of the voicing effect was a by-product of experiment designed
to investigate other phonetic issues. I review these studies and their
contributions in this section to pave the way for the importance of the study
described in this dissertation.

The first phonetician to tackle the issue of voicing effect in Arabic was
Flege (1979). In that study (his Ph. D. dissertation) he compared, among a few
other things, vowel duration in English and Saudi Arabic in pre-consonantal
environments. His study was limited to monosyllabic words read in a stressed environment in an identical frame sentence. He also limited the study to the long low vowel /æ:/ in pre-stop environments. The outcome of that study showed a difference of 6 ms only between /æ:/ in /gæ:t/ and /æ:/ in /gæ:d/ and 7 ms between /æ:/ in /sæ:k/ and /æ:/ in /sæ:g/. In both cases the vowel before the voiced stop was longer than before the voiceless counterpart (177 ms before /t/ versus 183 ms before /d/ and 167 ms before /k/ versus 173 before /g/). The difference did not reach the level of significance at \( p > 0.05 \) in either case. In terms of the durational ratio, the length of /æ:/ before the voiceless stop was 97% its length before voiced stops. These findings led Flege to conclude that Saudi Arabic does not exhibit a significant voicing effect on the preceding vowel. Obviously, this conclusion, should it prove to be true, would create serious problems for all those who believe(d) that the voicing effect is physiologically conditioned and thus should be found in all languages. It shows that some languages do not exhibit the voicing effect.

The next study on the voicing effect in Arabic was done in 1980 by Port et al. They used Standard Arabic trisyllabic words positioned in stressed environment. The study included the short and long low vowel /æ/ in pre-dental stop environment only. The stimuli were produced at three different rates of speech (i.e., slow, neutral, and fast) by two speakers of Egyptian Arabic, two speakers of Iraqi Arabic, and one Kuwaiti. Results of this study are given in the following table.
Table 2.11 Average vowel length in Standard Arabic before voiced voiceless consonants and their ratios as reported by Port et al. (1980) at three different tempos

<table>
<thead>
<tr>
<th>Rate of Speech</th>
<th>Vowel</th>
<th>Following consonant</th>
<th>Vowel length</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow</td>
<td>/æ/</td>
<td>/t/</td>
<td>72 ms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/d/</td>
<td>104 ms</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>/æ/</td>
<td>/t/</td>
<td>244 ms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/d/</td>
<td>287 ms</td>
<td>85%</td>
</tr>
<tr>
<td>Neutral</td>
<td>/æ/</td>
<td>/t/</td>
<td>60 ms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/d/</td>
<td>73 ms</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>/æ/</td>
<td>/t/</td>
<td>155 ms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/d/</td>
<td>168 ms</td>
<td>92%</td>
</tr>
<tr>
<td>Fast tempo</td>
<td>/æ/</td>
<td>/t/</td>
<td>53 ms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/d/</td>
<td>70 ms</td>
<td>76%</td>
</tr>
<tr>
<td></td>
<td>/æ/</td>
<td>/t/</td>
<td>106 ms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/d/</td>
<td>117 ms</td>
<td>91%</td>
</tr>
</tbody>
</table>

As can be seen in the table, results of this study are not in agreement with those of Flege: voicing seems to affect vowel duration at all three rates and for the two vowel lengths. The results are more in agreement with those of Chen (1970) for French, Russian, and Korean. Taking the results at the neutral rate, the rate at which most studies on vowel duration are conducted, we find the voicing effect to be significant at P> = 0.01. Indeed, statistical analysis for all the subjects except for one of the pairs, as reported by the authors, was significant at P> = 0.01.
Three important observations about the results in the table above should be noted. First, the table shows the ratios of vowel duration variation to be consistently greater for short /æ/ than for its long counterpart. This finding is in disagreement with the generally-accepted assumption that long vowels tend to exhibit greater variation as a function of voicing than do short vowels (Peterson and Lehiste 1960, Delattre 1962, Mack 1982, Laeufer 1992). Second, the results show greater ratios for the vowels when produced at the fast rate than at the neutral rate. This finding also disagrees with the fact that faster rate of speech causes smaller variation than neutral or slow rates (Gay 1978, etc.). Third, the ratios found in this study were for vowels occurring in trisyllabic words. This finding creates serious problems for Flege who found the voicing effect not to exist in monosyllabic words. It would be rather odd to believe that voicing affects vowel duration in trisyllabic words but not in monosyllabic ones (see Section 2.4.1.3.4 on word length and its effect on vowel duration and the voicing effect for explanation). Indeed, it would have been more acceptable had the result been the other way around with the voicing effect greater in monosyllabic words than in trisyllabic ones.

Two explanations can be given to account for this unexpected magnitude of the voicing effect in trisyllabic words. Either Flege's or Port et al.'s study is faulty, so that only one of them represents the truth. Second, it could be assumed that Standard Arabic (SA) shows the voicing effect but Saudi Arabic does not. The second explanation is an unlikely scenario because in both studies the data were collected through reading. The only variety of Arabic that has a written representation is SA. Thus, it can be assumed that the variety under investigation must be SA in both studies. Additionally, the
subjects who participated in the Port et al.'s study represent three different spoken dialects. It would be unlikely that all three dialects would show the voicing effect but Saudi Arabic would not, though admittedly, in principle there is no reason why all three dialects could not be different from SA. Still, I tend to accept Flege's study as the more representing of the role of voicing in Arabic due to the details and size of the study described in his dissertation.

Mitleb, in a series of studies conducted in 1981, 1983 and 1984, the most relevant of which is that of 1983, extended the investigation of the voicing effect to pre-fricative contexts. He examined the duration of /æ:/ before /s/ versus before /z/ in the two monosyllabic words /kæ:s/ "a glass, pitcher" and /kæ:z/ "gas". These words were inserted in a carrier sentence and read by eight Jordanian speakers of Arabic. By the fact that the data were elicited through reading, one would assume that the variety Mitleb investigated was Standard Arabic. It should be noted that the word /kæ:s/ is limited in its use to the spoken variety, having the form /kæ?s/ in Standard Arabic. The findings of this study were in total agreement with those of Flege. The average duration variation as a function of voicing did not exceed 5 ms, as can be calculated from the figure provided in the study, and did not show a significant difference, as reported by Mitleb. Mitleb concluded that the effect of voicing differs from one language to another and the claim that it is universal should be reconsidered.

Though Mitleb's findings support those of Flege's they cannot be decisive in determining whether voicing has a significant effect in Arabic or not. The study was limited to one pair only, produced in one context, and showing a mixture of Standard and Colloquial varieties. Additionally, Mitleb seems to
have forgotten that it was he who reported, in Port and Mitleb (1983), a
difference of 11 ms between /a/ in "sop" versus "sob" (78 versus 89 ms) and a
difference of 13 ms between /l/ in "sit" versus "sid" (66 versus 79 ms) when
produced by Jordanian teachers of English who had never been to an English
speaking country. Such differences cannot be attributed to the acquisition of
authentic English due to lack of exposure by these teachers to authentically
spoken English. By contrast, Jordanian teachers of English who received
some of their education in the United States and who participated in the
same study showed an average difference of 4 and 8 ms only for /a/ and /l/,
respectively. Therefore, it is obvious that the conflict over the role of voicing
remains unresolved.

Flege (1984) carried the investigation of the voicing effect in Arabic to
perception. Based on the findings of his previous studies (1979 and Port and
himself 1981), he hypothesized that speakers of Arabic may not be able to
perceive the voicing contrast of consonants as a result of varying the duration
of the preceding vowel. To prove this hypothesis empirically, he conducted a
perceptual experiment in which he varied the duration of the vowel /i/ and
the fricative /s/ in the synthesized pair peace and peas in a way similar to
that of Denes (1955). Five vowel durations that ranged from 150 ms to 350 ms
and five consonant durations that ranged from 100 ms to 300 ms were used in
this study. Three groups of subjects participated in the study: ten
monolingual speakers of American English, five Saudi learners of English
(described as inexperienced in English), and a group of six advanced Arab
learners of English. In varying the vowel duration and maintaining the
consonant length at 200 ms, Flege found that Saudi beginning learners of
English showed a crossover identification point for /z/ almost identical to that of the monolingual speakers of American English. Thus, perception evidence shows that vowel duration operates as a perceptual cue in an identical manner for both "monolingual" Arabic speakers and monolingual English speakers. Such a conclusion conflicts with the production findings reached by Flege and Mitleb, as described earlier. In the meantime, this finding adds to the confusion as to whether voicing has a significant effect on vowel duration in Arabic or not.

In summary, findings from the studies presented thus far leave the reader with a chaotic picture: Flege's and Mitleb's production experiments argue for the absence of a voicing effect and those of Port et al. and Flege's perception experiments argue for a significant voicing effect. Which of the two views would be the correct one? Obviously, to answer this question, further, and naturally more extensive, studies are needed. The study described in this dissertation was in part a response to this unresolved situation. It was hoped that by extending the study to cover two contexts (focused and unfocused), two vowel lengths (long and short), and all contrastive (in voicing) stops and fricatives, in addition to expanding the number of tokens to 4800 I would be able to come up with a more reliable answer (a detailed description of the experimental design is given in Chapter 3 below).

It should be admitted though that regardless of the conflicting views, it can be deduced from these studies that Arabic shows a tendency, small though it may be, for voicing to affect vowel duration: none of the studies described above shows a case where vowels are longer before voiceless consonants than before voiced ones. The real argument is therefore over
whether the voicing effect can be strong enough to place Arabic at levels comparable to those of French (as described by Chen or Mack), Spanish, Russian, etc. or weak enough to place Arabic at the extreme opposite end of English. Again, it is hoped that this study would contribute to a better description for this phenomenon.

2.5 The voicing effect: universal or learned? Where does Arabic fit in this debate and what does this study attempt to achieve?

As has been shown in the first part of this chapter, the different voicing-dependent vowel duration variations found in many languages have created an ever-lasting debate among phoneticians as to whether the vocalic variation is physically conditioned, and thus universal, or language-specific. In this section I present some of the various accounts given by phoneticians to support either view (i.e., universal versus learned phenomenon).

Based on results from the early studies, the overwhelming view developed to be that the voicing effect or at least some of the effect must be universal. As a result, a host of hypotheses/explanations, none of which has thus far been completely convincing, have been offered to account for this phenomenon. Some of these hypotheses were production-based, some were auditory-based, and some were based on language-specific phonological factors that supposedly masked the variation. Brief description of the most prominent of these hypotheses is provided in the following paragraphs.

Delattre (1962) argued strongly for the "Force of Articulation" hypothesis, originally proposed by Belasco (1953), as the cause for the variation. According to his hypothesis, some consonants and consonant classes require
more articulatory force to be produced than others. Belasco classified consonant classes according to the degree of articulatory force needed for their production into eight categories, with voiceless consonants requiring the greatest amount of force and /r/ requiring the least amount of force. Meanwhile, anticipation is a common process in speech production. Therefore, in the production of "bat" or "bad", for example, speakers tend to anticipate the consonants /t/ and /d/ in advance (perhaps while they are still in the process of producing /æ:/ or even the /b/) and thus adjust their production gestures accordingly. In anticipation of a /t/, as opposed to a /d/, speakers end up "trimming" the end of the vowel in order to provide sufficient force needed for production. Such trimming would be significantly less in the case of /d/ than in the case of /t/ due to less needed force for production.

Voiceless consonants require greater force to be produced than their voiced counterparts. "The anticipation of a consonant requiring a 'strong' force of articulation will tend to shorten the preceding vowel since more of the total energy needed to produce the syllable is concentrated in the consonant" (Belasco 1953: 1016)

This hypothesis has been criticized as suffering from three major drawbacks. First, it lacks an agreed-upon measure for what constitutes physiological energy or force of articulation (Lisker 1974, Kluender et al. 1988). Second, this hypothesis lacks a definite domain in which it operates. Thus, while it might function properly in a VC sequence, it does not account for the lack of a voicing effect on the following vowel in a CV sequence. It would be logical to assume that in a CV sequence, where C can be any consonant, some consonants would require greater force than others to be produced. These
consonants, according to this hypothesis, would naturally affect the duration of the following vowel due to greater force they would draw upon in their production. Thus far, such effect has not proven to be of significant value. Indeed, most studies testify to the lack of it (see Lehiste 1970 and the literature cited there for further discussion). Third, this hypothesis implies that all languages should behave in the same manner and that the voicing effect should be constant cross-linguistically. Most empirical evidence provided by the tens of studies cited in the literature argues otherwise.

Halle and Stevens (1967) and later Chomsky and Halle (1968) attributed vowel lengthening in a pre-voiced environment to the time consumed as a result of laryngeal adjustment aimed at maintaining glottal pulsing during oral constriction or closure. Like the Force or Articulation hypothesis, this view has been criticized for its inability to account for the lack of voicing effect on the following vowel. By analogy, if there would be a laryngeal adjustment from the vowel to the following voiced consonant, there should be similar or counter adjustment from the preceding voiced consonant to the following vowel that would consume some time and therefore lead to a lengthened following vowel. Again empirical evidence does not support a lengthening effect in post voiced consonant position. Moreover, Lisker (1974) pointed out that fiber-optic and EMG studies do not show any laryngeal adjustment from a vowel to the following voiced consonant. To the contrary, these studies showed laryngeal adjustment from the vowel to the following voiceless consonant. Also, Sharf (1964) reports longer vowels before underlyingly voiced consonants than before voiceless ones in whispered speech where there would be no glottal pulsing. Similar reports were reported for
neutralized environment (Sharf 1962; Fox and Terbeek 1977; Port et al, 1981). Lastly, this view would not be able to account for the lack of voicing effect in some languages as reported by Keating (1979), Flege (1979), and Mitleb (1983).

Chen (1970) compared vowel duration variation in four languages and found that variation is not constant cross-linguistically. He concluded that the effect of voicing on the preceding vowel is universal but the magnitude of the effect is language-specific. In accounting for the universality of the voicing effect, he evaluated six hypotheses and ended up rejecting five and favoring the "Rate of Closure Transition" hypothesis as a potential explanation. According to this hypothesis, supraglottal pressure builds up faster during the closure of a voiceless consonant due to an open glottis than during the closure of a voiced consonant where the glottis is semi-closed. Articulators move accordingly faster when followed by voiceless than when followed by voiced consonants to withstand such a greater oral pressure. The fast movement tends to shorten the vowel in pre-voiceless consonants more than in pre-voiced ones.

According to Javkin (1976), "this argument fails on several grounds." First, it predicts that high front vowels in pre-dental stop environments, for example, should not cause as much shortening effect as front or back low vowels in the same position due to shorter articulators' movement. This has not proven to be true. Both Peterson and Lehiste (1960) and Sharf (1962) have shown the voicing effect of the dental consonants to be approximately the same on both /i/ and /a/. Second, Javkin states that the difference in the rate of closure between voiced and voiceless consonants is not sufficiently great to account for the vowel variation. Third, some studies have shown that the
combined rate of movement for the lips and jaw in the production of labials after /i/ to be faster for voiced labials than for voiceless labials (Sussman, MacNeilage, and Hanson 1973). These arguments make it necessary that a new and more reliable account for the voicing effect be sought.

Lisker (1974) rejected all these accounts and provided one of his own. He argued that vowels tend to become shorter before voiceless consonants in order to prevent these consonants from becoming pre-aspirated. Vocal cords tend to separate towards the final portion of the vowel. Consonant closure for the voiceless consonants takes place immediately after the separation of the vocal cords to prevent pre-aspiration, an act which is not permitted in English. Lisker claims that the coordination between the vocal cords and the vocal tract leads to shortened vowels.

Like other accounts, Lisker's explanation does not account for all different aspects of vowel duration variations. It is limited to some cases of English only. For example, it cannot account for the different context-dependent variations. Klatt (1975) found the variation to be greater before a syntactic boundary than in a phrase-medial position. Laeufer (1992) found greater vowel duration variations in focused and phrase final contexts than in unfocused and phrase-medial position. Since, according to Javkin, "the injunction against pre-aspiration in English does not vary according to these environments, it cannot be responsible for variable vowel duration differences" (P. 81). More importantly, since Lisker appears to have developed his explanation for English only, we are left with no explanation for vowel duration variation in other languages.
Scully (1974, cited in Javkin 1976) proposed an aerodynamic model that would account for the variation in all languages. According to this model, vowel amplitude decreases faster before a voiceless consonant than before a voiced one in VCV sequence. Javkin pointed out two major problems with this model. First, the model predicts that vowels would shorten more in post consonant positions than in pre-consonant ones. Second, the model predicts that the duration of the vowel should be constant regardless of the voicing of the preceding consonant. This prediction does not account for the aspiration period in the case of voiceless stops. Obviously, if aspiration is to be counted as part of the vowel the vowel becomes longer than if aspiration were to be dropped.

As can be seen from these studies, attempts to support the universal hypothesis on physiological grounds have generally been unsuccessful. This failure pushed some scholars to seek perception-based accounts instead of production-based ones. Javkin (1976) proposed, in support of the universal view, that the main cause of the variation is auditorily-based, and should not be attributed to physiological and articulatory constraints. As indicated earlier, he believes that voicing during the closure or constriction of the consonant is perceived as extension to the preceding vowel. Thus, vowels are heard longer than they actually are. Javkin supported his claim by a series of experiments in which subjects created tones that matched in their length the vowel /I/ in the synthesized tokens /hIs/ and /hIz/. This perception is usually carried over to production. This proposal is consistent with Ohala's proposals (1974, 85) that sound change starts with errors made by the listener who continuously attempts to copy the sounds he/she hears.
This proposal has been criticized as being inadequate (Kluender et al. 1988). It cannot account for why listener/speaker would exaggerate in perceiving vowels beyond the actual duration of the vocal vibration (Kluender et al. 1988).

"By itself, Javkin's proposal is not an adequate explanation for the vowel length effect. If glottal pulsing during closure makes the preceding vowel seem longer, then, according to the argument, listeners/speakers will tend to copy this illusory lengthening by producing a longer vowel. But notice that in copying the perceived utterance, listeners/speakers will no doubt produce a closure pulsing (and hence illusory vowel lengthening), and therefore any significant actual vowel lengthening should lead to a perceptibly longer vowel than the one being copied. Although listeners/speakers may be assumed to make copying errors of various kinds, it is unclear why detectable errors in apparent length would not be corrected. Thus, what remains unexplained by Javkin's account is why listeners/speakers would tend to exaggerate perceived vowel lengthening beyond what is afforded by glottal pulsing alone" (Kluender et al. 1988: 156)

After this criticism, Kluender et al. (1988) offered another perceptual explanation. They believe that speech communities tend to vary the duration of the vowel to enhance the perception of the closure duration. This model is based on the durational contrast in psychology. Accordingly, long vowels tend to make listeners perceive the adjacent (following) short consonant even shorter and a short vowel tend to make listeners perceive the following long consonant even longer. This interrelatedness in relations becomes clear if we realize that closures for voiced consonants are generally shorter than those of their voiceless cognates. Thus, the differential vowel duration is an enhancing cue that operated optimally in the context of VC sequences. The authors add that it is not the number of the auditory cues as much as it is the proper interaction of these cues with the context they are in that makes them
operate optimally. Languages may differ in the weight they put on these cues to signal phonological contrasts just as they may differ in their phonemic inventories. Based on this explanation, it can be argued that some languages like English apparently place more weight on vowel duration variation as a cue for the perception of the following consonant than languages like Arabic where such a cue is absent.

Upon closer examination of Kluender et al.'s explanation, one would find that this model has not been set to support the universality view as much as to show how the variation happens. Indeed, by implication, it assumes that the voicing effect is not identical cross-linguistically and thus must be learned. It further implies that learners of a foreign language have to learn which cues are used in L2 to account for the variation whether these cues be the same in their native language or not.

In contrast to the above view, it should be mentioned that the notion of vowel duration being learned (i.e. not universal) was accepted by many phoneticians as early as late fifties and early sixties (Zimmerman and Sapon 1958; Peterson and Lehiste 1960). This acceptance became more and more popular as subsequent studies revealed vocalic variations different from (mostly smaller than) those found in English. In fact, the prevailing view since early seventies has been that at least some of the variation is learned (Chen 1970, Flege 1979, Port, Al-Ani, and Maeda 1980, Port and Flege 1981, Mack 1982, Borden and Harris 1984). This view was questioned by Laeufer (1992) who compared vowel duration in French and English. Laeufer realized that earlier phoneticians who studied differential vowel duration in French tended to either miss or overlook some variables that always masked the
effect of voicing on vowel duration. Some of these variables are (1) the prosodic structure of French (as it compares to that of English), (2) the release of the final voiceless consonant in French, (3) resyllabification, and (4) intrinsic vowel length. Laeuf er concludes that when these factors are controlled the voicing effect in French would not be considerably different from that of English.

Though Laeuf er's hypothesis is interesting and might be promising, the supporting evidence for it is practically limited to French. She encourages further studies that would seek language-specific features that may mask the voicing effect. Based on this conclusion and the available literature on the topic, it can be argued that Laeuf er's conclusion is still tentative and inconclusive in terms of its applicability to the rest of languages. It remains then that languages tend to differ with regard to the effect of voicing on the duration of the preceding vowel. Taking in this study Arabic and English, two languages which lie at the opposite ends of two extremes, one can investigate whether native speakers of one language learning the other in adulthood ever acquire vowel duration variation as it exists in L2. Observation indicates that adult learners of L2 tend to improve over time yet paradoxically most of them tend to maintain an accent in L2 regardless of the time they spend learning that language\textsuperscript{21}. This study has been designed to (1)

\textsuperscript{21} A number of studies have addressed similar relevant issues that deal with different strategies people at various ages use in categorizing sounds, words, pictures, etc. (Bruner 1964; Bruner, Oliver, and Greenfield 1966; Nelson 1974, Kuhl 1980; Hillenbrand 1983, 84, cited in Flege's 1987). It is commonly believed that children rely on sensory information when encountering a new sound while adults (after the age of puberty) rely more on cognitive processes. Such reliance on cognitive processes leads them automatically to classify sounds according to categories they already know rather than according to their acoustic characteristics.
re-examine whether Arabic manifests significant vowel duration variation ascribable to the voicing of the following consonant; (2) investigate whether duration of exposure and advancement by adult American learners of Arabic would help them acquire vowel duration variation as it is in Arabic; (3) investigate whether learning L₂ would affect vowel duration variation in L₁; and (4) present some of the salient characteristics of Arabic as produced by L₂ learners.

In order to avoid some of the problems associated with some of the previous studies and provide more reliable results, I have been careful to (1) avoid any of the potential confounding factors discussed above; (2) include a relatively large number of tokens (i.e., 4800 tokens); (3) extend the context to focused and unfocused environments; (4) include long and short vowels; and (5) extend the experiment to include all stops and fricatives found in Arabic. The following chapter gives a detailed description of the experimental design, subjects, data, and data collection procedures.
Chapter III
The Experiment

This chapter presents the data used in the experiment and the contexts in which they were used, describes the methodology used in acquiring the data, the subjects that participated in the experiment, the segmentation processes and the difficulties associated with these processes.

3.1 Subjects

To assess the acquisition of Arabic-like voicing-dependent vowel duration, or lack of it, and to assess whether $L_2$ features can be carried over into $L_1$ production, it was necessary to have three groups participate in the experiment: essentially monolingual Arabs, essentially monolingual Americans, and incipient bilingual American learners of Arabic.

3.1.1 Monolingual Arabic speakers

Six native speakers of Levantine Arabic participated in the experiment. All six speakers were Palestinians who had just arrived in the United States. The longest period of time anyone of them spent in the US before he/she was recorded was three weeks. Their knowledge of English was limited to what they had learned at middle and high school when they were in either Jordan or the West Bank. Such knowledge can be described at best as familiarity with sporadic words and structures.
More importantly, their English teachers were either Jordanians or Palestinians who learned English in adulthood in a foreign language environment: either in the West Bank or in Jordan. Thus, the exposure of these subjects to native American English is basically nil.

It should be noted though that passive exposure to American English through TV can be common in these areas. However, personal experience confirms that interest in the development of the story and the constant availability of subtitles tend to divert people's attention from the language itself, especially at the beginning levels. Based on these observations, the subjects were viewed as essential monolinguals. They will be referred to in the rest of this study as native speakers of Arabic.

Four of these subjects were males and two were females with ages that ranged from twenty-two to thirty-four. Their mastery of Standard Arabic (SA), the variety under investigation in this study, should not have been different from any educated Arab. That is, they did not speak it natively in daily interaction due to the diglossic situation they lived in, yet they were capable of producing it natively if the need arose like any other educated Arab\(^\text{22}\) (Ferguson 1959, Mitchell 1986).

Results of their production data serve two purposes. First, they provide evidence that either support previous findings that argue for the absence of voicing-dependent vowel duration in Arabic or refute it.

---

\(^{22}\)This statement should not be taken to imply that uneducated Arabs are incapable of producing Standard Arabic. In some cases, depending on the interests of the individual, uneducated people may have better command of SA than the educated. It is the common understanding, however, that educated people receive several years of instruction in SA, they are assumed to have relatively good command that is superior than uneducated ones. Also, educated people tend to have constant exposure through reading and writing, a privilege not available to uneducated people.
Second, they will constitute the norms to which the production of the bilingual non-native groups are compared.

3.1.2 Bilingual American learners of Arabic

Eighteen American students and professors studying or teaching at the Ohio State University, University of Pennsylvania, or Princeton University participated in the study. All of them came from families of European ancestry and their exposure to Arabic prior to starting college was nil. Their exposure to Arabic ranged from a minimum of one and a half years to a maximum of twenty-two years. They were classified into three groups consisting each of six subjects and described as beginners, intermediate, and advanced. The bilinguals were classified into these levels based on a rating done by ten native speakers of Arabic who had no experience teaching Arabic to non-natives and were unfamiliar with accented Arabic. These speakers listened to a passage read by each subject and rated them according to their degree of accentedness on a scale that ranged from zero to ten, with zero standing for heavily accented or unintelligible and ten standing for native competency of Arabic. The ratings, which would range from zero to one hundred when they were added up, were added up for each speaker. Then, speakers were classified into three groups: learners rated from 0 - 34.99 were considered beginners, learners rated between 35 - 64.99 were classified as intermediate, and those rated above 65 were classified as advanced. Figure 3.1 below shows the distribution of subjects according to their years of experience and their ratings by the ten judges. It can be seen from the figure that there is an
overall correlation between the scores obtained from judges and the years spent on learning Arabic ($r = .72$). This correlation disappears in the advanced group ($r = .11$). This disappearance can be attributed to the fact that advanced learners differ from the beginning learners with regard to their learning strategies. Beginning students are usually exposed to the

![Figure 3.1. Subjects distributed according to years of experience and rating by judges](image)

same instructional materials by virtue of following structurally sequenced syllabi and being under immediate supervision of the instructors. Advanced learners, on the other hand, are more self-reliant; they are freer
to choose the materials they want to learn, focus on the language area (i.e., speaking, writing, pronunciation, etc.) they like most, adopt the learning strategy that fits them, etc. Their level of mastery of a foreign language reflects their individualistic styles and interests rather than collective exposure to the same instructional materials as is often the case for beginning students.

3.1.3 American English Monolinguals

Six American monolinguals also participated in the experiment. All six subjects were natives of the Midwest who at least finished high school. Their ages ranged from nineteen to forty-six. Some of them, according to self-report, studied a European language in high school, but none of them had any exposure to Arabic. Results from their production data serve as norms to which production data of the bilingual groups are compared. Such comparisons were designed to assess the influence of L2 on L1 with regard to voicing-dependent vowel duration, if there would be indeed any influence. Any interference from L2 to L1 will most likely show up in the production of the advanced bilingual group. In the meantime, results from this group will be compared with those of native Arabs to assess the magnitude of difference in voicing-dependent vowel duration between the two languages. The motive for the elicitation and incorporation of data from English monolingual speakers was the variability in results reported in the literature. Previous studies have shown ratios as low as 53.41% (Mack 1982) and as high as 79.36% for mid utterance unfocused position (Laeufer 1990). Therefore, it was necessary to have the same
English data produced by the monolingual and the bilingual groups under the same conditions in order to establish norms that fit the study better than any other findings reported in the literature.

3.2. Test materials

Eight monosyllabic Standard Arabic minimal pairs were used as stimuli. All of them were real Arabic words taken from a popular contemporary dictionary of SA called Al-Munjid. It should be noted that some of these pairs were deemed somewhat rare and unfamiliar to native speakers. The pairs differed from each other with regard to the voicing of the post vowel (final) consonant (voiced versus voiceless), manner of articulation of the post vowel consonant (stops versus fricatives) and vowel length (long versus short). Two pairs ended in stops and six ended in fricatives. The fricatives represented three places of articulation: dental (i.e., /s/ and /z/), velar (/x/ and /γ/), and pharyngeal (/h/ and /γ/). In addition, two English monosyllabic minimal pairs ending in stops similar to those used in the Arabic data were also used. Table 3.1 below present all the data used in the study.

It should be noted that the pair /bæs, bæs/, used in the SA data, may not be found in SA but is found in all vernacular varieties of Arabic that the writer is familiar with. It was, however, matched acoustically\(^{23}\) with the pair /dæs, dæs/, found in SA, for verification of similar voicing effect.

\(^{23}\)The pair /dæs/ and /dæs/ was recorded five times in the two contexts given above by two native speakers. Measurements were taken from spectrograms. Results did not show any significant difference from those found for /bæs/ and /bæs/. 
Table 3.1 Target words and carrier sentences used in the study for Arabic and English.

A. Arabic Data:
1. List of Target words
   - Stops
     /bæd/ "to keep legs apart" /bæt/ "to decide"
     /bæd/ "to exterminate" /bæt/ "to stay overnight"
   - Dental fricatives
     /bæz/ "to embezzle" /bæs/ "enough, only"
     /bæz/ "a type of bird" /bæs/ "to kiss"
   - Velar fricatives
     /saɛx/ "to insert ... into" /sæɣ/ "to move a (wedge) in place"
     /saɛx/ "to melt" /sæɣ/ "pure, exact"
   - Pharyngeal fricatives
     /fæh/ "voice/noise of snake" /fæʃ/ 
     /fæh/ "to spread (the word)" /fæʃ/

2. Carrier sentences
   (a) katabat ... ṣala llawḥ (focused environment)
   (b) lam taktub ... ṣala llawḥ (unfocused environment)

A. English Data:
1. List of target words/minimal pairs
   - Stops
     /bæt/ "bat" /bæd/ "bad"
   - Fricatives
     /bæs/ "bass" /bæz/ "baas, Plural of baa"

2. The carrier sentences
   (a) Say ... again (focused environment)
   (2) Don't say ... again (unfocused environment)
Results for both pairs did not show any noticeable difference. It has also been established that preceding consonants should not have any significant effect on the duration of the following vowel (See Lehiste 1970 and literature cited there)\textsuperscript{24}.

All words were inserted into two carrier sentences that represent two contexts, focused and unfocused environments. Focused in this experiment refers to the portion of the utterance that has the highest degree of prominence in that utterance, and it usually conveys the new information. It is acoustically characterized as having the highest F\textsubscript{0} value in the utterance. Carrier sentence (a) represents the focused context and carrier sentence (b) represents the unfocused context (See Table 3.1 for the two environments).

Seventy-two Arabic words, fifty-six of which were distracters, and twenty English words, sixteen of which were distracters, were inserted in the carrier sentences and typed on 3X5” cards in the respective language. Cards with Arabic data were shuffled randomly and passed to subjects for recording. Each card was recorded a minimum of five times with the following constraint on ordering: no token could be repeated unless each of the other 71 tokens had been repeated at least once for any given cycle. The same procedure was applied to the English data. The total number of tokens from the two languages included in the study were 4800 tokens distributed according to the following:

\textsuperscript{24}There appears to be only one study in the literature which shows some effect of the preceding consonants on the following vowel. That study was done by Port et al. (1980) and the language that shows the effect is Japanese.
(1) 16 Arabic words X 2 contexts X 18 bilingual Americans X 5 times
    (readings) = 2880 tokens

(2) 4 English words X 2 contexts X 18 bilingual Americans X 5 times
    (readings) = 720 tokens

(3) 16 Arabic words X 2 contexts X 6 Arabic monolinguals X 5 times
    (readings) = 960 tokens

(4) 4 English words X 2 contexts X 6 English monolinguals X 5 times
    (readings) = 240 tokens

Recording occurred in an anechoic chamber at the phonetics laboratory at the Ohio State University under the direct supervision of the experimenter. Prior to the recording each subject listened to the author modeling (recording) a sample of the data. Each modeling session was prompted by the appropriate question(s) which would produce answers in either focused or unfocused context depending on the prompt. For example, the question \textit{mæ:ðæ kætæbaet sæleæ l-læẉh} “What did she write on the board?” would produce an answer with the target word in a focused context (i.e., kætæbaet \textit{bæd sæleæ l-læw} with \textit{bæd} being in focused situation). On the other hand, a question of the form \textit{ṣu ?illi mæ simliθu:ð} “What is it that she did not do?”, would yield an response with a focus on the negation portion rather than on the target word. The typical response would be \textit{læm tæktUb bæd sæleæ l-læẉh}. (I have used a question in the colloquial variety of Arabic because it was deemed as the most suitable form of question to elicit the proper response. Any question from the standard variety would not have yielded as much of an authentic response). Subjects were then instructed individually in their native
language to read each utterance at a normal rate and with as much natural intonation as possible, in a way similar to that modeled by the researcher. Purpose of the experiment was not explained to any of the subjects prior to the recording. All recordings were done on professional-quality cassette tape recorders.

Wide-band spectrograms were made on DSP Sona-graph Model 5500 for all 4800 tokens. Measurements were taken for the vowels in all the tokens by placing cursors at both ends of the second formants for the vowel. The machine provides automatically the duration of the segment between the cursors in milliseconds (ms) (See Figures 3.2a and 3.2b). Other techniques were also used in identifying the boundaries and taking the measurement. These techniques are all explained below under the section Segmentation.

It should be noted that the only vowels used in the study were the short and long /æ/. Both /i/ and /u/ and their short counterparts, the only other monophthongal vowels found in SA, were excluded from the study due to (1) absence of corresponding words in both languages (English and SA) that have either /i/ or /u/ as is the case with words such as bat, bad, bass, and baas, (2) both English and long Arabic /æ/ seem to exhibit near-equivalence in duration (Flege and Port 1981). Adherence to the use of real SA words has been an important part of the study. No nonsense words could be used in a study of this type for fear that subjects may use their native language background instead of the second language background in the production of the data. For example, if we were to use a

---

25 Both /æ/ in Arabic and /æ/ in English were reported to have 181 ms and 177 ms, respectively (Flege 1979).
word such as *food*, which exists in English but is absent in Arabic, we would not be able to tell whether the bilinguals would use their English or Arabic backgrounds in the production though the script (Arabic versus English) in which the word was written may dictate that subject use the language of that script.

All measurements were taken, means were calculated, and several comparisons were made. Graphs were made to show the differences graphically and many statistical tests were run to check for significance. The following comparisons were made;

A. Comparisons exclusive to Arabic monolinguals

1. Voicing effect on short /æ/ before stops in focused and unfocused environment.

2. Voicing effect on long /æ/ before stops in focused and unfocused environment.

3. Voicing effect on short /æ/ before fricatives in focused and unfocused environment.

4. Voicing effect on long /æ/ before fricatives in focused and unfocused environment.

5. Voicing effect on long versus on short /æ/ across all consonant classes and contexts.

6. Voicing effect in focused versus unfocused environments.

7. Voicing effect before fricatives versus before stops across the two contexts.

8. Effect of place of articulation of vowel duration as shown by the fricatives.
Figures 3.2. Two spectrograms showing the cursors at both ends and a panel showing the duration in ms
B. Comparisons between the monolingual Arabic speakers and the bilingual groups using SA data (interference from L1 to L2)

1. Interference of voicing effect on the duration of long and short /æ/ in pre-stop focused and context.
2. Interference of voicing effect on the duration of long and short /æ/ in pre-stop unfocused context.
3. Interference of voicing effect on the duration of long and short /æ/ in pre-dental fricative focused context.
5. Interference of voicing effect on the duration of long and short /æ/ in pre-velar fricative focused context.

C. Comparisons between the bilingual groups and the monolingual English speakers using English data (interference from L2 on L1)

1. Voicing effect on /æ/ in pre-stop focused and unfocused context.
2. Voicing effect on /æ/ in pre-fricative focused and unfocused context.

3.3 Segmentation and measurement

Segmentation is used in this study to refer the process of identifying a segment by marking its boundaries with two cursors as it occurs in connected speech (See figures 3.2a, b earlier in this chapter). Measurement refers to the process of finding the length of the segment in millisecond
In this study the segment under investigation will be the vowel /æ/ and its long counterpart.

Identifying the boundaries of a segment in speech analysis is not always a straightforward process, especially so when dealing with data from L2 speakers. Overlap between adjacent segments due to coarticulatory effects is a common process in connected speech. Yet identifying these boundaries (i.e., boundaries of a segment) to the most accurate degree possible is extremely important in duration studies. After all, it is the goal of the study to find out the actual duration of the segment(s) under investigation. Thus, it was deemed necessary that the following background be provided in order for the reader to understand the difficulties associated with segmentation, and in particular the difficulties associated with the data described in this study. In the meantime, the current study appears to have provided segmentation problems of its own that should be of interest to anyone concerned about segmentation of L2 production.

First, it should be recognized that cues signaling segment boundaries are unambiguous in many cases. In fact, they can be easily identified. Consider the segmentation of /æ/ in the sentence "Say bad again" as produced by a native speaker in Figure 3.3 above. Boundaries of /æ/ are marked with two cursors at the beginning and at the end.

On the other hand, not every spectrogram included in the study shows clear-cut boundaries as those shown in Figure 3.3. The degree of ease or difficulty with which a boundary of the target vowel can be identified, especially the boundary signaling the end of the vowel, seems to correlate
Figure 3.3 Spectrographic representation of the sentence "Say bad again" produced by a native speaker of American English

with the consonantal class of the following sound (stop versus fricative), place of articulation, and whether the speaker is a native speaker of the language or non-native. This study has shown that end boundaries of a vowel followed by a stop are usually the easiest to identify. By contrast, ease of identification of boundaries "separating" a vowel from the following fricative is determined by the place of articulation for that fricative as well as what appears to be the acoustic characteristics of the
Figures 3.4 The position of the cursor at the end of the vowel when followed by /s/ (a), /z/ (b), /ʃ/ (c), /ʒ/ (d), /h/ (e), and /ŋ/ (f).
Figure 3.5 Samples of /fæʃ/ produced by native speakers of Arabic.
vowel under investigation. The study has shown that boundaries separating /æ/ from the following dental /s/ or /z/ are relatively easy to identify; boundaries separating /æ/ from the following velar fricatives /x/ and /γ/ are occasionally problematic; and boundaries separating /æ/ from the following pharyngeal /h/ and /ŋ/ are often impossible to identify, especially when /æ/ is followed by /ŋ/ (See Figures 3.4a, b, c, d, e, and f below). The difficulty of identifying consistent boundaries in the context of /ŋ/ forced the writer to drop the contrastive pair /fæŋ, fæŋ/ from the study. It was just not possible to have reliable measurements.

By contrast, identifying the boundaries for the Arabic data produced by the bilingual groups was far harder than those produced by native speakers26. The bilinguals showed characteristics that are atypical of the data under investigation. These characteristics did add to the difficulty of identifying the boundaries. Among the characteristics are the following:

(1) Lack of distinction between /x/ and /γ/ in many instances. The common practice was to produce /γ/ as /x/27. In a few cases /γ/ was produced as a velar stop (See Figure 3.6 below). In measuring the duration of /æ/ in either case, the measurement was taken as it was supposedly intended by the speaker. That is, if the card reads /γ/ but the production

---

26It should be noted that segmentation for data as produced by the bilingual groups can be a study of its own. Here I simply list some of the observed difficulties and provide a brief description of each. Such description, in my estimate, is rudimentary and by no means exhaustive.

27The interesting point concerning this substitution is that it was always unidirectional; that is, it was always /γ/ that became /x/, never the other way around. This note should not obscure the fact that /γ/ was often used to substitute for /γ/, as indicated in the text above.
showed /x/, I took it as /γ/. Lack of distinction in these cases was viewed as a performance error rather than an underlying voicing difference.

Figure 3.6. Substitution of /γ/ by /g/.

(2) No distinction was made between the focused and unfocused contexts in some cases, especially for the beginning group.

(3) Insertion of irregular pulsing (i.e., successive glottal stops) in the middle of the vowel creating a rather long vowel with breaks. Two subjects, both belonged to the beginning group, produced their long vowels with breaks in the middle. Sometimes, the break would show up as a single glottal stop and often as successive glottal stops. These irregularities were considered part of the vowel (See Figure 3. 7 below for
a sample with a break in the middle of /æ/ in the word /sæ:x/ produced by a bilingual beginner).

Figure 3.7 A spectrogram produced by a bilingual and showing a laryngeal break in the middle of the vowel and a change in vowel quality

(4) Producing extra long /æ:/ in many instances. Some subjects, apparently in attempt to highlight the quantity difference between the short and long /æ/, produced a hyper, to use Lehiste's term, long /æ:/, produced a hyper, to use Lehiste's term, long /æ/. In many cases the duration of this vowel exceeded 400 ms, and in a few cases it exceeded 500 ms. This process was individualistic; that is, individual subjects, regardless of their level, systematically lengthened the vowel. Therefore, it was possible to find students in the advanced group behaving with regard to this feature in a manner similar to students in the middle or beginning groups. Consider Figure 3.8a and 3.8b for illustration. Figure
3.8a was produced by a female student from the advanced group and
Figure 3.8b was produced by a female student from the beginning group.
Notice that in Figure 3.8b the duration exceeds half a second and the
vowel shows a change in quality with attenuated formant representation
half way through. It looks to me that this change represents a rather long
transition from a pure vowel to a z-colored transition that leads to /z/.
(5) Some students produced a pause after every word, thus producing data
that should in principle resemble data produced in isolation. No
intonation pattern can be found in these data. Reading the sentences does
not resemble connected speech.

To overcome the problem of inconsistency in segmentation and to
maximize the validity of the results, the author established some
segmentation rules and tried to be as consistent as possible in following
these rules. In cases where wideband spectrograms did not show clear
boundary cues, the author employed other spectral displays such as wave
form display, narrow-band display, amplitude display and occasionally
identified the boundaries by listening to the target segment several times
as the cursor has been moved back and forth accordingly.

The segmentation rules were (1) the inclusion of the transitions as
part of the vowel, (2) the exclusion of the aspiration period of the previous
consonant, (3) the inclusion of gaps found in the middle of the vowel for
some speakers as well as the gaps found occasionally between the vowel
and the following fricative, and (4) reliance on voicing as it shows up on
the voicing bar and the striations of higher formants as markers of
boundaries.
Figures 3.8. Hyper long /æ/ produced by two bilinguals.
It is worth mentioning at this point that aspiration is usually associated with voiceless stops in English. The stimuli used in this study do not show any voiceless stops in pre-vocalic position, and consequently the second rule given above should, in principle, be considered irrelevant. However, the data in this study showed that an aspiration period (i.e., 10 - 20 ms) can follow voiced stops as well (See Figure 3.9 below for illustration).

Figure 3.9 Aspiration follows /b/ in the word /bæ:d/ in a way similar to, yet shorter, than aspiration after voiceless consonants. The utterance was produced by an American bilingual subject.
Finally, it should be admitted that regardless of how consistent a person tries to be in identifying the boundaries, a level of subjectivity cannot be avoided in some cases. This study should not be viewed as an exception. In some cases the author had to rely on his experience and background in drawing the boundaries. The subjectivity level of drawing the boundaries, however, should in principle have no impact on the reliability of the results because errors tend to happen in the two opposite directions and thus "cancel" each other out. There is no reason why errors are always committed in the same direction.
Chapter IV

Results, Interpretation, and Discussion

Chapter 4 presents the results, interpretations, and discussions for the three aspects of the study described in this dissertation in addition to showing some of the features of L2 speech. Thus, the chapter is divided into four parts. Part one shows the effect of voicing on vowel duration in Standard Arabic before stop and before fricative environments. Part two assesses the acquisition of voicing-dependent vowel duration in Arabic by American students and shows any potential interference from L1 into L2. Part three assesses potential reverse phonological interference (i.e., interference from L2 (Arabic) into L1 (English)). Finally, part four presents some of the characteristics of American learners' production of Arabic.

4.1 Voicing-dependent vowel duration in Standard Arabic

4.1.1 Voicing-dependent vowel duration before stops

Figures 4.1a to 4.1d display the effect of voicing on vowel duration before stops for short and long /æ/ (Figures 4.1a and 4.1b versus Figures 4.1c and 4.1d) and in focused and unfocused environments (Figures 4.1a and 4.1c versus Figures 4.1b and 4.1d). The vertical axis of each figure shows the average vowel duration in milliseconds (ms henceforth) for each subject and the horizontal axis shows the subjects numbered from 1 to 6. The scale for vowel duration on the vertical axis ranges from 0-350 ms. The two bars
assigned to each subject show vowel duration before the voiceless stop /t/ (slash-lined bars) and before the voiced stop /d/ (the dark bars).

The voicing effect would be manifested with the dark bars being on the average significantly longer than the slash-lined bars.

Figure 4.1a shows the effect of voicing on the short /æ/ before /t/ and /d/ in the focused environment. As can be seen from the figure, the effect varies slightly from one subject to another but it is consistently very small for subjects 1 through 5 and disappears for subject 6. The average duration for all the subjects amounted to 95.56 ms before /t/ and 100.96 ms before /d/.

Though this average difference between the pre-voiced and pre-voiceless environments is small, one-tail t-test run on the overall average ratio for the six subjects shows a significant difference at the p <0.05 level (t(5) = 3.00, p <0.05).28

Two observations about the results in this figure should not go unnoticed. First, the data show a consistent voicing effect that cannot be random particularly since it is significant. The majority of the subjects produce the vowel /æ/ slightly longer before voiced consonant than before its voiceless counterpart. Second, subjects 5 and 6 show longer overall vowel duration than the rest of the subjects. Both of these subjects are the only two females in the group. Perhaps gender plays a role in determining the style of speech production among Arab females; they might be adhering to a more careful style than Arab males. Both observations are maintained in the rest of the figures and the rest of the study.

---

28 The t-test was applied to the average mean of the ratios for all six subjects and compared to 1.0 (no effect) to test for significance. The same test has been repeated on all the figures given in this section of the dissertation.
Figure 4.1. Effect of voicing on vowel duration before stops as produced by monolingual Arabs for (a) short /æ/ in focused context, (b) short /æ/ in unfocused context, (c) long /æ/ in focused context, and (d) long /æ/ in unfocused context.
Figure 4.1b shows the effect of voicing on the short vowel /æ/ before /t/ and /d/ in the unfocused environment. The figure shows similar results to those shown in Figure 4.1a. All subjects with the exception of 6 show relatively small voicing effect. The overall ratio of vowel length before /t/ to that of before /d/ was .95. This effect, however, did reach significance at the p < 0.05 level (t(5) = 2.33, p < 0.05).

The focused context has contributed to the overall length of the vowel but not to the magnitude of the effect of voicing. The average vowel duration for /æ/ in the focused context across the pre-voiced and pre-voiceless environment amounted to 98.26 ms. By contrast, the same average has been 90.2 ms in the unfocused context. This context-dependent difference in vowel duration is much smaller than what has been reported in the literature for other languages (Klatt 1975; Fourakis 1991; Laeufer 1992). The magnitude of the voicing effect, on the other hand, did not show any noticeable difference that can be attributed to context. Both ratios of vowel duration in pre-voiceless to that of pre-voiced amounted to .95 for the focused and unfocused contexts. This result is not in agreement with what has been reported in the literature. Laeufer 1992, for example, shows that the effect of voicing has been reduced from a medial focused context to medial nonfocused context by 10% for English and by 15% for French. The lack of such a difference in Arabic can be attributed to (1) the fact that the voicing effect in Arabic has been shown to be in general very small, thus allowing for minimal or no variation, and (2) the small effect of the focused context on vowel duration in general.

The effect of voicing on long /æ/ is shown in Figures 4.1c and 4.1d. Figure 4.1c shows the effect in the focused context and Figure 4.1d in the unfocused
context. In both figures the effect of voicing is consistently small; in fact smaller than the effect on short /æ/. The overall ratio in the focused and unfocused contexts has been .97 and .98, respectively. Statistical tests again did surprisingly show significant difference at \( P < 0.05 \) \( (t(5) = 2.238, p < 0.05; (t(5) = 3.03, p < 0.05, \text{respectively}) \). The average vowel length, on the other hand, turned out to be 14% shorter in the unfocused context than in the focused context. This difference is very close to that found in the pre-velar environments which will be presented in the next section.

It should be noticed that subject 5 has shown greater vowel length before voiceless stop than before the voiced counterpart in both contexts. She is though the only subject in the two figures to show reverse effect. Her performance has been considered exceptional.

The findings in these four figures are in agreement with those of Flege (1979) who found the average ratio of vowel length before voiceless to that of pre-voiced stop to be .97 for Arabic produced by Saudis. In fact, the average length of the long /æ/ came out almost identical in the two studies (177 ms in Flege’s study and 177.7 ms averaged across the two contexts in this study).

In conclusion, the sample used in this study confirms that Standard Arabic has a small yet significant voicing effect on vowel duration before stops. This conclusion still supports the hypothesis suggested earlier that Arabic contrasts sharply with English with regard to this feature (i.e., the effect of voicing) and thus makes it a good example for the study of phonological interference and L2 acquisition.
4.1.2 Voicing effect on vowel duration in pre-fricative environment

Figures 4.2a to 4.2d display the effect of voicing on vowel duration in pre-fricative environments. The figures differ from the previous four in that each subject is represented by four bars instead of two. The first two bars (the filled and crossed bars) show vowel duration in the pre-dental fricatives /s/ and /z/ and the next two bars (the slash-lined and dark bars) show vowel duration in the pre-velar fricatives /x/ and /γ/.

Figure 4.2a shows the effect of voicing on the short /æ/ in the focused context. The figure shows a consistent voicing effect for all the subjects before dentals and velars, with the effect being greater for the latter. The effect seems to vary according to the place of articulation of the fricative. The average ratio for the dentals across all the subjects has been .90 and for the velars .79. T-tests showed the difference for the pre-dentals to be significant at p <0.05 (t(5) = 3.136, p <0.05) and the difference for the pre-velars is significant at p <0.01 (t(5) = 3.82, p <0.01).

The findings for the pre-dentals vary considerably from a subject to another. Some subjects show a minimal voicing effect that is not significant while others show much greater effects. Table 4.1 below details the performance of each subject as well as the ratios for each word pair.

The table shows subjects 1, 2, 4, and 5 with high ratios (small differences) that are close to those found in before stop environment. It is clear that subject 3 has deviated from the group. Thus, it is fair to conclude that the majority of this group shows a relatively small voicing effect before dentals. This conclusion confirms that of Mitleb (1983, 1984) who found that vowels
Figure 4.2. Effect of voicing on vowel duration before fricatives as produced by monolingual Arabs for (a) short /æ/ in focused context, (b) short /æ/ in unfocused context, (c) long /æ/ in focused context, and (d) long /æ/ in unfocused context.
Table 4.1. Average vowel durations in pre-dental and pre-fricative consonants and the ratios for each subject

<table>
<thead>
<tr>
<th>Subject</th>
<th>/baes/</th>
<th>/beiz/</th>
<th>Ratio</th>
<th>/sæx/</th>
<th>/sæy/</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>82 ms</td>
<td>87</td>
<td>.94</td>
<td>87.2</td>
<td>107.2</td>
<td>.81</td>
</tr>
<tr>
<td>2</td>
<td>104.2</td>
<td>108.2</td>
<td>.96</td>
<td>88</td>
<td>127</td>
<td>.69</td>
</tr>
<tr>
<td>3</td>
<td>82.2</td>
<td>112.2</td>
<td>.73</td>
<td>84.4</td>
<td>107.8</td>
<td>.78</td>
</tr>
<tr>
<td>4</td>
<td>94.2</td>
<td>100.2</td>
<td>.94</td>
<td>82.2</td>
<td>95.6</td>
<td>.86</td>
</tr>
<tr>
<td>5</td>
<td>151</td>
<td>156</td>
<td>.97</td>
<td>127.2</td>
<td>131.6</td>
<td>.97</td>
</tr>
<tr>
<td>6</td>
<td>133.2</td>
<td>150.8</td>
<td>.88</td>
<td>126.2</td>
<td>185.6</td>
<td>.68</td>
</tr>
</tbody>
</table>

Overall Ratios: .90 .79

in Jordanian Arabic, a dialect usually characterized as very similar to one spoken by the subjects who participated in this study, to be only 5 ms on the average longer before /z/ than before /s/. However, it differs from Mitleb in showing that this small difference is statistically significant.

By comparison, the majority of the subjects show a greater consistent voicing effect in the pre-velar environment that is not different from French (Chen 1970, Mack 1982, Laeuffer 1992), Russian, or Korean (Chen 1970). In other words, the voicing effect in pre-velars in Arabic does not seem to be that different for the rest of languages that show voicing effect. In fact, if we drop subject 5, the voicing effect in this context in Arabic would be close to some ratios reported for English.

It is worth noting that this is the first study that attempted to investigate the voicing effect before velars in Arabic. It certainly raises many interesting questions: Does place of articulation play that much of a role to cause greater
voicing effect before velars than before dentals? If this were true, what articulatory or auditory mechanism/process determines this effect in velars but not in dentals? Could the difference be attributed to the phonological structure of Arabic (i.e., dentals might have some features that are absent in the velars or vise versa)?

One way to confirm the validity of the results in the pre-velar environment is to replicate the experiment using dental and velar stops. Unfortunately, Standard Arabic does not have the voiced velar stop /g/. Therefore, the absence of this sound in this variety of Arabic prevents us from extending the experiment to verify the findings reported in Figure 4.2a and the rest of the figures.

Figure 4.2b shows the corresponding effect of voicing in the unfocused context. The effect of voicing in pre-dentals decreases in this context to a bare significance at the p <0.05 level (t(5) = 2.36, p <0.05). The overall ratio of vowels before /s/ to that before /z/ is .92. The weak voicing effect can be attributed to the great variability among the subjects. Four of the subjects show a relatively small voicing effect. Only subjects 2 and 3 show considerable voicing effect. The performance of these two subjects must have swayed the overall ratio.

By contrast, the effect before velars is significant in this context at p <0.01 (t(5) = 5.453, p <0.01). The vowel /æ/ is consistently longer before /γ/ than before /x/ for all the subjects. This finding confirms the results shown in the previous figure. The overall ratio in pre-velars for this figures is .78, almost identical to that found in previous figure (.79).
The lack of focus in this context has caused the vowel to drop in length by 13%. That is, the average vowel length in the focused context is 13% longer than in the unfocused context. This drop is smaller than what has been reported for other languages such as English and Greek (Fourakis 1986, 1991).

Figures 4.2c and 4.2d display the voicing effect on the long /æ:/ in focused and unfocused environments, respectively. Figure 4.2c shows the effect in pre-dental and velar fricatives in the focused context. The average length across all the subjects before /s/ and /z/ amounts to 161.2 and 209.5 ms, respectively. These averages show an average difference of 48.3 ms. This relatively large difference has shown significant difference at $p < 0.05$ only ($t(5) = 3.61$, $p < 0.05$), apparently due a high degree of variability across the subjects. A closer look at the figure and the averages for each subject reveals that subjects 3 through 6 show a consistent voicing effect. By contrast, subjects 1 and 2 show a smaller effect.

Averages for the pre-velars, on the other hand, show less variability between the two environments (voiced and voiceless environments) for all the subjects than those of pre-dentals, and thus statistical testing has confirmed a significant difference at $p < 0.01$ ($t(5) = 3.87$, $p < 0.01$). This statistical outcome is as strong as those shown in the rest of the figures for the pre-velar environment. In terms of overall ratio, the average length of /æ:/ in pre-/x/ constituted .89 of its average before /γ/.

Figure 4.2d shows the voicing effect on /æ:/ before dentals and velars in the unfocused environment. In both environments (pre-dentals and pre-velars) the effect is significant at the $p < 0.05$ and 0.01 levels, respectively ($t(5) = 2.91$, $p < 0.05$; $t(5) = 4.45$, $p < 0.01$). The overall vowel duration across all the
environments is shorter than in the focused environment by 20 ms. Vowels averaged 182.05 ms in the focused environment and 162.50 ms in the unfocused environment. The effect of voicing on vowel duration has also decreased considerably in the unfocused environment (48.3 ms in the focused pre-dentals versus 25 ms in the corresponding unfocused pre-dentals, and 21 ms in focused pre-velars versus 17 ms in the corresponding unfocused pre-velars). Both differences fall within the expected patterns. This pattern has been reported for English, Greek, and French (Laeufer 1992; Fourakis 1991).

4.1.3 Discussion

It is clear from Figures 4.1a - 4.2d that a pattern of a voicing effect has been established for the Arabic data used in this experiment. This pattern can be characterized as minimal yet significant in pre-dental stop and in pre-dental fricative environments, and consistently greater and naturally significant in pre-velar fricatives. The question to be raised now is about the causes that determine and shape this pattern. On the face of it, it seems that either some or all of the following three factors determine this pattern: manner of articulation (MOA), place of articulation (POA), and/or the phonological structure of Arabic.

Taking MOA first, it does not seem that this factor can determine and contribute to this pattern. MOA cannot differentiate between the effect of stops versus fricatives in the pre-dental environment since the effect is consistently small in both environments. POA seems responsible for some of the effect, but the major factor, I believe, is the phonological structure of Standard Arabic. Arabic has dental consonants, not alveolars (Cadora 1979).
Dentals are usually produced by the tip of the tongue and not the blade like alveolars. They tend to have smaller tongue-teeth contact area than alveolars and thus shorter duration. Consequently, the difference in duration between a dental voiced stop, for example, and a voiceless one is smaller than the difference between the corresponding voiced and voiceless alveolars. In a VC sequence in a language like English the shortening effect of the voiceless consonant on the preceding vowel would be compensated for by the length of that voiceless consonant. Conversely, when the following consonant is voiced, the longer preceding vowel would be followed by a shorter consonant (Port and Dalby 1982). This balancing act between the vowel and the following consonant tends to maintain the duration of the sequence VC at relatively the same length regardless of the voicing characteristics of the following consonant (Belasco 1953, Delattre 1961, Chen 1970, Port et al. 1984). Carrying this analysis to Arabic, we find that the following dental consonant, whether it is voiced or voiceless, affect minimally the length of the preceding vowel due to lack of variation in the duration of the consonants. Therefore, it is very likely that the effect of voicing on the preceding vowel is very small because the relative length of the VC sequence would be maintained with just the minimal voicing effect shown. Moreover, Standard Arabic has been classified as a syllable-timed language where syllables are expected to maintain similar length (Kharma and Hajjaj 1989; Dauer 1983). This classification adds to the fact that the smaller variation in length between voiced and voiceless consonants would naturally lead to smaller variation in the duration of the preceding vowel to maintain the same syllable length throughout.
This explanation does not apply to the velars /x/ and /g/. Both of these consonants require raising the back portion of the tongue towards the velum. The duration of the tongue movement and area of the contact is much greater than the that of the dentals. Also, the duration of the consonant would vary according to voicing; voiceless consonants would be longer than voiced ones due to the greater force needed in their production. The variation in the length of consonant would be compensated for by a reverse variation in the length of the preceding vowel to maintain similar syllable length. Thus, vowels preceding voiceless consonants would be expected to be shorter than when they are followed by their corresponding voiced ones.

Another factor that might limit the effect of voicing on vowel duration is the contrastive vowel length found in Arabic. Contrastive length might limit the reduction magnitude of vowels, especially the long ones, beyond a certain level to prevent duration overlap with their corresponding short ones. The findings of this study support this view strongly. Table 4.2 below shows the average mean durations for long and short vowels in focused and unfocused pre-velar fricative environments.

The table shows greater duration variation for the short /æ/ than for the long /æ/ in both focused and unfocused contexts. This variation can be seen in the ratios of the long /æ/; they are 10 points on average smaller than those of the short /æ/. This can be explained by the fact that the reduction of the short vowels would not lead to overlap with any other (shorter) vowel, unlike long vowels. Also, the number of vowels found in Standard Arabic allows for greater variation without causing overlap with each other.
Standard Arabic has three short vowels only thus leaving plenty of room for these vowels to vary without losing their identities.

Table 4.2. Average mean durations in ms for short and long /æ/ in focused and unfocused pre-dental fricative environments.

<table>
<thead>
<tr>
<th></th>
<th>/ææ/</th>
<th>/æː/</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused context</td>
<td>99.2 ms</td>
<td>125.7</td>
<td>.79</td>
</tr>
<tr>
<td>Unfocused context</td>
<td>84.5</td>
<td>108.43</td>
<td>.78</td>
</tr>
<tr>
<td>Ratio</td>
<td>.85</td>
<td>.86</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>/æːæ/</th>
<th>/æːː/</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused context</td>
<td>168.2</td>
<td>189.3</td>
<td>.89</td>
</tr>
<tr>
<td>Unfocused context</td>
<td>143.1</td>
<td>161</td>
<td>.89</td>
</tr>
<tr>
<td>Ratio</td>
<td>.85</td>
<td>.85</td>
<td></td>
</tr>
</tbody>
</table>

This analysis may seem to run counter to Klatt’s degree of compressibility suggested in his 1975 work. According to Klatt, long vowels tend to shorten more than short ones because of their intrinsic length. Support for this hypothesis comes from works like Rositzke (1939), House (1961), and Crystal and House (1982). Rositzke found that long/tense vowels were 60% longer before voiced than before voiceless consonants and short vowels vary by only 39% in the same environments. Similarly, House found that lax vowels were 43% longer before voiced than before voiceless consonants and long vowels were 48% longer in the same environments. Crystal and House found significant voicing effect on vowel duration in pre-stop environment when
the preceding vowel was long/tense only. In all these studies, and many others like them, long/tense vowels show greater variation than short ones.

This view would be misleading if it were applied to Arabic. In other words, it would be misleading to expect vowels in Arabic to behave in the same way. Vowels in English and Arabic have different features. While length is contrastive in Arabic, it is not in English. Quality, not quantity, is the primary distinctive feature in English. That being the case, vowel compression can be greater in English as a function of voicing\(^{29}\) than in Arabic due to the secondary role of quantity in English. In Arabic, on the other hand, quantity is the primary feature and overlap in duration can cause a loss of this feature.

Table 4.2 also shows that the focused context affects the overall length of the vowel but not the magnitude of the voicing effect. The long and short /æ/\(^s\) are approximately 15\% each longer in the focused than in the unfocused context. Ratios of pre-voiced to pre-voiceless vowel duration, on the other hand, have maintained their values unchanged. The effect of focused context

---

\(^{29}\) There have been historically two ways to interpret the voicing-dependent vowel duration variation. The first view considers the effect of voicing as that of lengthening; that is, voicing of the following consonant tends to increase the duration of the preceding vowel (Chomsky and Halle 1968, Javkin 1976). The second view is the opposite: the lack of voicing of the following consonant tends to shorten the preceding vowel (Belasco 1953, Delattre 1961, Chen 1970). The conflict over which view is correct has not been resolved yet. It is not clear whether vowel duration variation in the pre-obstruent environment is a process of lengthening or shortening.

Though the work in this dissertation does not permit one to decide for either view over the other and indeed the whole debate falls outside the scope of the dissertation, it has thus far been implicitly assumed that the effect of voicing is that of lengthening. In citing Klatt’s degree of compressibility at this point to account for the degree of the voicing effect on the short versus that on the long vowels in Arabic a shift may seem to have been made towards the shortening view. However, since neither view is supported over the other in this dissertation the citation of Klatt’s “degree of compressibility” does not constitute a shift so much a continuation of the search for the most logical interpretation for the results.
on vowel length and its absence on the magnitude of the voicing effect has been the prevailing pattern throughout the whole data. The overall percentage of vowel reduction across all the environments and subjects combined has been 11%. This level of reduction is much smaller than what has been reported for English and French (Laeufer 1992, Fourakis 1991), and Greek (Fourakis 1991). The low effect can be attributed to the position of the target word in the carrier sentence. Both the focused and unfocused contexts fall at the end of a major syntactic boundary (at the end of a verb phrase) causing often times a short pause to follow. Syllables at the end of a phrase are always longer than when they are in a mid utterance position (Klatt 1973, 75, 76; Umeda 1975; Cooper and Danley 1981; Luce and Charles-Luce 1985; Crystal and House 1982, 88). Since all of the test words used in this experiment are monosyllabic and fall at the end of a phrase the effect of focus could have been limited by their position in the phrase.

In conclusion, it could be argued that Arabic does not differ from languages like French or Spanish with regard to the voicing effect. That is, the magnitude of the voicing effect ranges between 10 - 20% of the overall length of the vowel. This effect is decreased in the pre-dental environment due to the phonological structure of Arabic as explained above. Most existing studies done on Arabic reported no effect because they were limited to the pre-dental environments (Port and Flege 1981, Mitleb 1983, 1984). In the meantime, the effect is not as great as it is in English whose voicing effect ranges from 20 to 50% depending on the context and mode of recording the data. This outcome, though it differs from other studies done on Arabic, still confirms a considerable difference between the two languages. The acquisition of this
difference is the focus of the next two sections. In the next immediate section I discuss whether adult American learners of Arabic would rid themselves of the great voicing effect found in English when they produce Arabic or transfer this difference from English into Arabic. The following two questions will be addressed: will advancement in learning Arabic help American students acquire the relatively small voicing effect found in Arabic? Will the effect be subject to enhancing and weakening factors such as focused/unfocused environments, intrinsic vowel length, or MOA? The discussion will be limited to the pairs of words viewed as similar in the two languages.

4.2 The acquisition of the voicing effect in Arabic by adult American students

Figures 4.3a - 4.8d display the results for this section. The figures are grouped into three sets of eight each according to the place and manner articulation of the following consonant and will be presented in this section as such. The first set which consists of Figures 4.3a - 4.4d shows the acquisition of voicing-dependent vowel duration before stops. The second set consists of Figures 4.5a - 4.6d and shows the acquisition of voicing-dependent vowel duration in the pre-dental fricatives /s/ and /z/. And the third set consists of Figures 4.7a - 4.8d and shows the acquisition of voicing-dependent vowel duration in the pre-velar fricatives /x/ and /γ/. The figures are set in a way identical to those discussed in the previous section. The following description has been repeated for convenience. The vertical axis shows the duration of vowels given at a fixed range from 0 - 500/550 ms, and the horizontal axis shows the subjects (native speakers are represented by numbers 1 - 6 and non-
natives are represented by their ratings). The four bars assigned to each subject stand for the duration for short /æ/ before the voiceless consonants (filled bars), the duration for short /æ/ before the voiced consonant (crossed bars), the duration for long /æ/ before the voiceless consonant (slash-lined bars), and the duration for long /æ/ before the voiced consonant (dark bars).

The acquisition of the Arabic voicing effect would be manifested in the difference in duration between the even and the odd bars in each grouping. The even bars display vowel duration before voiced consonants and the odd bars before their voiceless counterparts. The smaller the average difference between these two sets of bars, the closer non-natives are to the natives in decreasing the difference, and thus showing greater magnitude of acquisition. On the other hand, the greater the difference between the two sets of bars the greater the degree of interference from the native language (English) into Arabic.

4.2.1 The acquisition of the voicing effect on vowel duration before stops

Figures 4.3a - 4.4d display the effect of voicing on vowel duration for the native and non-native groups before stops. The first four figures (i.e., 4.3a - 4.3d) show the effect in the focused environment and the last four figures (i.e., 4.4a - 4.4d) show the effect in the unfocused environment. Starting with the focused context, Figure 4.3a displays the voicing effect for the monolingual Arab group. The voicing effect of this group serves as the norm to which the performance of the rest of the groups will be compared.
Figure 4.3. Effect of voicing on vowel duration before stops for long and short /æ/ in focused context as produced by (a) monolingual Arabs, (b) the beginning bilingual group, (c) the intermediate group, and (d) the advanced group.
Figure 4.4. Effect of voicing on vowel duration before stops for long and short /æ/ in unfocused context as produced by (a) monolingual Arabs, (b) the beginning bilingual group, (c) the intermediate group, and (d) the advanced group.
As discussed in the previous section, the voicing effect has turned out to be very small for the long and short /æ/ in this context (i.e., focused context). The average vowel length across all the subjects for short /æ/ has amounted to 96\textsuperscript{30} ms before /t/ as opposed to 101 ms before /d/. Similarly, the average vowel length across all the subjects for the long /æ/ amounted to 176 ms before /t/ and 180 ms before /d/. In both cases, however, t-tests did show significant difference at the .05 level.

Figures 4.3b, 4.3c, and 4.3d show vowel duration variation for the non-native groups in the focused context. Figure 4.3b shows the performance of the beginning group, Figure 4.3c show the performance for the intermediate group, and Figure 4.3d shows the performance for the advanced group. It is obvious that there is no linear positive correlation between the level of the learners as rated by the judges and the voicing-dependent difference they show (r = .20 for the short /æ/ and .16 for the long /æ/). However, the behavior of some groups seems, on the average, to show possible acquisition of Arabic voicing-dependent vowel duration. Table 4.3 below shows the ratios of vowel duration in pre-voiceless to those of pre-voiced consonant for all three groups in the focused context.

The table shows the advanced group to have raised the ratio to .84 for both short and long vowels, thus reducing the effect of voicing and becoming the closest to the monolingual group. The beginning and intermediate groups, by contrast, have maintained considerably lower ratios, closer to those reported for English (Peterson and Lehiste 1960; Mitleb 1984; Luce and Charles-Luce 1985; Laeufer 1992). In fact, the ratios between the latter two

\textsuperscript{30}The fractions have been rounded up to the closest number/point. This practice has been maintained throughout the rest of this dissertation.
groups do not show any noticeable difference. It could be argued then that
level of advancement in Arabic has helped the advanced group reduce the
difference, at least in this context. The effect of voicing for the advanced group
has also been reduced to a greater degree for the short /æ/ before stops in
unfocused position; it has reached a ratio of .87.

Table 4.3. Ratio of vowel duration in pre-voiceless to pre-voiced for
all groups in the pre-stop focused context (calculated by dividing the
overall means in pre-voiceless context over the overall means
for the pre-voiced context for each group).

<table>
<thead>
<tr>
<th>Monolinguals</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>long</td>
<td>short</td>
<td>long</td>
</tr>
<tr>
<td>.95</td>
<td>.97</td>
<td>.76</td>
<td>.78</td>
</tr>
<tr>
<td>.95</td>
<td>.97</td>
<td>.76</td>
<td>.74</td>
</tr>
<tr>
<td>.84</td>
<td>.84</td>
<td>.84</td>
<td>.84</td>
</tr>
</tbody>
</table>

Inter-subject variability cannot be ignored in these graphs. Some subjects
show very little variability, especially with short vowels, while others do not
seem to differ from English. For example, the subjects rated 19 and 34 from
the beginning group, 49 and 55 from the intermediate group, and 67 and 68
form the advanced group show either very little or no difference in vowel
duration for short /æ/ ascribable to the voicing characteristic of the following
consonant. By contrast, the rest of the subjects show considerable voicing
effect.

In addition to showing the least voicing effect, the advanced group shows
the closest overall vowel length to that of the monolingual group. The
beginning group shows the greatest overall vowel length of 215 ms averaged
across the two vowel lengths. This average length is decreased for the intermediate group to 175 ms, and is decreased further for the advanced group to 157 ms. By comparison, the monolingual group averages 138 ms. This outcome is also observed across the unfocused context as well. Table 4.4 below presents the overall length of the vowel for all the groups across the two vowel lengths in the focused and unfocused contexts.

Table 4.4. Overall length of the vowels across the two vowel lengths for the focused and the unfocused contexts given in ms

<table>
<thead>
<tr>
<th></th>
<th>Monolinguals</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused</td>
<td>138 ms</td>
<td>215 ms</td>
<td>175 ms</td>
<td>157 ms</td>
</tr>
<tr>
<td>Unfocused</td>
<td>123 ms</td>
<td>193 ms</td>
<td>158 ms</td>
<td>153 ms</td>
</tr>
</tbody>
</table>

It can be concluded from this table that the bilinguals tend to gradually perceive and acquire the overall vowel length in Arabic before they might acquire the voicing effect. The acquisition process correlates with the duration of exposure to Arabic\(^{31}\). This conclusion lends more support to that of Mack's

\(^{31}\) It could be argued that the gradual approximation of the overall vowel length has resulted from developing a better reading proficiency in Arabic by the advanced group. In other words, the approximation of this phonetic feature may not be a case of acquisition as much as it a case of improvement in the reading of a foreign script. Though such an argument may be true and cannot certainly be ignored in this study, we cannot be sure from the data and the results we have that it is the sole cause or just a contributing cause for the approximation. The role of acquisition could also be the sole cause for this approximation. It is obvious at this point that preferring one interpretation over the other is not possible. Therefore, though the gradual approximation could be confounded with improvement in the reading ability, I will continue attributing it in the rest of the study to the acquisition leaving for further studies the investigation into which interpretation is definitely the correct one.
(1982) who found that French-English bilinguals did not acquire the voicing effect of English but produced overall vowel lengths of English that were closer to English than to French.

The subjects rated 67 and 77 in the advanced group have produced longer vowels than the rest of the subjects and might have skewed the overall average. Without these two subjects the overall average of the advanced group would not be different from the monolingual average. It is perhaps worth noting that both of these subjects are professors of Arabic who had already spent several years teaching it prior to their participation in this experiment and they might have learned to pay more attention to the clarity of the message than to its authenticity.

Figures 4.4a -4.4d display the effect of voicing on the duration of long and short /æ/ before stops in the unfocused environment. Figure 4.4a shows the effect of voicing for the monolingual group, Figure 4.4b shows the effect for the beginning group, Figure 4.4c shows the effect for the intermediate group, and Figure 4.4d shows the effect for the advanced group. The figures display a pattern of performance for all four groups that is similar to that found in the focused context. The only difference seems to be in the overall length of the vowel between the two contexts. As shown in Table 4.4 above, vowels are longer in the focused context than in the unfocused context for all the groups. Statistical testing shows the variable "focused context" to be barely significant at P <0.05 (F(1,159) =3.94, P <0.049). The voicing effect, on the other hand, has shown a systematic focus-dependent decrease. In fact, all the groups except for the beginning group show greater voicing effect on the long /æ/ in the unfocused context than in the focused one (See Table 4.5 below). This
outcome is in conflict with that of Laeuf er's (1992) who found that unfocused context tends to systematically reduce the effect of voicing in both French and English.

Long /æ/ has been consistent in showing a smaller voicing effect than short /æ/. This outcome, though might not be as clear in the environment before stops, is extremely consistent in the pre-velar and pre-dental fricatives as we will see a little later. It is also in conflict with the expected performance. That is, long/tense vowels have been showing greater voicing effect than short/lax ones (Rositzke 1939; Peterson and Lehiste 1960; House 1961; Klatt 1975; Crystal and House 1982). The reason for this reverse outcome in this experiment can be ascribed to the temporal phonemic distinction found in Arabic. Most American students as well as native speakers of Arabic tend to put more emphasis on long /æ/ in production to distinguish it from short /æ/. By so doing, they end up confounding the voicing effect, at least to some degree, with the required length for the vowel to be quantitatively distinctive. Similar results were found earlier by Port et al. (1980) for Arabic. Port et al. found that ratio of vowel length for short /æ/ in pre-voiceless of pre-voiced stands at .82. By comparison, the same ratio for long /æ/ stands at .92. Level of advancement does not contribute to improvement with regard to this point. All groups exhibit similar behavior.

Table 4.5 shows a positive correlation between vowel length and the variability in the scores. There are greater standard deviations for long /æ/ than for its shorter counterpart. Variability is also greater for non-native groups than for the monolinguals, an indication of lack of stable acquisition of the vowel length.
Table 4.5. Means, standard deviations, and ratios for long and short vowels in pre-stop environment for all groups in focused (A) and unfocused contexts (B).

<table>
<thead>
<tr>
<th></th>
<th>A. Focused Context</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monoling. Group</td>
<td>Beginning Group</td>
<td>Intermed. Group</td>
<td>Advanced Group</td>
</tr>
<tr>
<td><strong>Short vowels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>95 (25)</td>
<td>132 (34)</td>
<td>105 (29)</td>
<td>98 (41)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>101 (23)</td>
<td>173 (49)</td>
<td>137 (43)</td>
<td>117 (40)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.95</td>
<td>.76</td>
<td>.76</td>
<td>.84</td>
</tr>
<tr>
<td><strong>Long Vowels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>176 (48)</td>
<td>244 (69)</td>
<td>195 (52)</td>
<td>188 (46)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>180 (45)</td>
<td>311 (85)</td>
<td>262 (76)</td>
<td>224 (52)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.98</td>
<td>.78</td>
<td>.74</td>
<td>.84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B. Unfocused Context</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monoling. Group</td>
<td>Beginning Group</td>
<td>Intermed. Group</td>
<td>Advanced Group</td>
</tr>
<tr>
<td><strong>Short vowels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>88 (19)</td>
<td>119 (35)</td>
<td>104 (24)</td>
<td>95 (37)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>92 (15)</td>
<td>158 (48)</td>
<td>129 (38)</td>
<td>110 (36)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.95</td>
<td>.75</td>
<td>.81</td>
<td>.87</td>
</tr>
<tr>
<td><strong>Long Vowels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>152 (43)</td>
<td>224 (64)</td>
<td>170 (44)</td>
<td>173 (55)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>158 (43)</td>
<td>272 (70)</td>
<td>231 (73)</td>
<td>234 (86)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.96</td>
<td>.82</td>
<td>.73</td>
<td>.74</td>
</tr>
</tbody>
</table>

The graphs also show that the bilinguals tend to recognize the temporal phonemic distinction found in Arabic at the very early stages in commencing to learn Arabic. All groups show a distinction between the short and long /æ/
that is greater yet very close to that of the monolinguals: the beginning group produced an average of short /æ/ that constituted .55 of the length of its long cognate, the intermediate group had an average ratio of .56, and the advanced group had an average ratio of .52. By comparison, the monolingual group had a ratio of .57. Similar results were found in the pre-dental and pre-velar fricatives. Such findings indicate that the quantity distinction found in Arabic is among the phonetic features to be most easily learned by American students of Arabic.

4.2.2 The acquisition of the voicing effect in the pre-dental fricatives /s/ and /z/

Figures 4.5a - 4.6d display the effect of voicing on the duration of long and short /æ/ in the pre-dental fricative environment for all the groups participating in the experiment. The figures are set up in a way identical to figures 4.3a - 4.4d. Figures 4.5a - 4.5d show the voicing effect in the focused context and Figures 4.6a - 4.6d for the unfocused context.

Similar to the results found for vowel duration before stops, the subjects do not show a linear correlation between the voicing effect and the subject ratings as given by the judges (r = .09 for the short /æ/ and .001 for the long /æ/ in the focused context and r = .15 and .11 for the short and long /æ/, respectively, in the unfocused context). As groups, however, the advanced group has shown the closest overall vowel duration averages to those of the monolingual group. The beginning group shows the overall average for the short /æ/ across the voiced and voiceless consonant and across focused and
Figure 4.5. Effect of voicing on vowel duration before dental fricatives for long and short /æ/ in focused context as produced by (a) monolingual Arabs, (b) the beginning bilingual group, (c) the intermediate group, and (d) the advanced group.
Figure 4.6. Effect of voicing on vowel duration before dental fricatives for long and short /æ/ in unfocused context as produced by (a) monolingual Arabs, (b) the beginning bilingual group, (c) the intermediate group, and (d) the advanced group.
unfocused contexts to be 185 ms followed by the intermediate group with an average of 152 ms, and last the advanced group with an average of 128 ms. By comparison, the monolingual group shows an overall average of 108 ms. The same pattern is maintained for the long /æ/. The beginning group has an overall average of 298 ms followed by the intermediate group with an average of 246 ms, and the shortest belongs to the advanced group with an overall average of 238 ms. The monolingual group has an average of 184 ms. This outcome along with that found before stops provide strong support for the hypothesis that advancement in Arabic helps American students acquire the Arabic overall vowel duration. This form of acquisition fits within the "approximative system" described by Flege (1981). According to this system, learners of a foreign language tend to approximate the authenticity of some L2 sounds as they advance in learning the language. Therefore, the more advanced a student becomes the more authentic his/her production of some L2 sounds or phonetic features becomes.

More support for the acquisition of the overall length of the vowel comes from the degree of inter-subject variability found within some of the groups. The beginning group shows the highest degree of variability for long /æ/ in both focused (S.D. = 99) and unfocused (S.D. = 113) contexts. This degree of variability, which can be noticed clearly in Figures 4.5b and 4.6b where the subjects rated 18, 22, and 33 differ greatly from the rest of the subjects, provides a strong indication that some members of this group have not acquired a stable and consistent duration distinction for long and short /æ/. They are still providing extra long /æ/ s in an attempt to distinguish it from its short cognate.
In the meantime, it should not go unnoticed that the advanced group has a rather large degree of variability for long /æ/ as well. However, the variability for the advanced group can be accounted for by the deviation of the two subjects rated 67 and 77. As mentioned earlier, both of these subjects are professors of Arabic who adhered to a rather deliberate style in their recording, thus producing a distinct long /æ/. Their behavior is consistent in all contexts (in pre-voiced, pre-voiceless, and in both focused and unfocused contexts). When these two subjects are eliminated, the inter-subject variability level drops to 53, thus becoming very close to that of the monolingual group.

The monolingual group shows a consistent voicing effect in all environments and contexts for both long and short vowels. The ratios range from .86 for the long /æ/ in the unfocused context to .90 for the short /æ/ in the same context. This voicing effect is very close to that of French (.87) as reported by Chen 1970 or Spanish as reported by Zimmerman and Sapon 1958. Zimmerman and Sapon reported an average difference of 18 ms between the pre-voiced and pre-voiceless context. In this experiment the average difference across all the contexts and environments has been 19 ms. The difference is still far smaller than that found in English and the two languages differ considerably with regard to this point.

The findings of the voicing effect do not show consistent and gradual acquisition process. Results for long /æ/ in focused and unfocused contexts show a progressive, yet weak, parallel between the group level and the magnitude of the voicing effect. Students in the advanced group show the highest average ratio, hence the closest to the monolingual group. The same
trend is not maintained for the short /æ/. In fact, the advanced group shows the lowest ratios in the focused and unfocused contexts, .69 and .66, respectively. In the meantime, the average differences between the beginning and the intermediate groups do not exceed two points. Such findings show a split between acquisition and the lack of it. Results of the long /æ/ support a level of approximation similar to that of Nemser (1971), Flege (1981) and Mitleb (1984), regardless of how weak it is. Results of the short /æ/ do not show any signs of acquisition. On the contrary, they show adherence, especially as manifested by the advanced group, to the voicing effect as it is found in English. This outcome is not different from Port and Flege (1981) who found out that the Saudi Arabic speakers had not acquired the voicing effect of English. Nor is it different from that of Mack's (1982) and Lauerfer (to appear) who found that French-English bilinguals produced English data with voicing effect characteristic of French.

The effect of focused context has been limited in this section to the overall length of the vowel and not to the voicing effect. Thus, the monolingual group shows an average of 112.5 ms in the focused context for short /æ/ as opposed to 104 ms in the unfocused context. Other groups show similar, yet smaller difference, 192 versus 197 ms for the beginning group, 154.5 versus 150 ms for the intermediate group, and 131 versus 127 for the advanced group. Similar results but with greater differences have been found for the long /æ/. The following table shows these results.
Table 4.6 Average vowel duration for all the groups for long and short /æ/ in focused and unfocused contexts and the average differences between the two contexts.

<table>
<thead>
<tr>
<th></th>
<th>Monolingual</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused</td>
<td>195.5 ms</td>
<td>310.5 ms</td>
<td>258.5 ms</td>
<td>259.5 ms</td>
</tr>
<tr>
<td>Unfocused</td>
<td>173 ms</td>
<td>285 ms</td>
<td>234 ms</td>
<td>236.5 ms</td>
</tr>
<tr>
<td>Difference</td>
<td>22.5 ms</td>
<td>25.5 ms</td>
<td>24.5 ms</td>
<td>23 ms</td>
</tr>
</tbody>
</table>

This outcome conflicts in part with that of Laeufener (1992) who found that both the overall vowel length and the voicing effect were greater in the focused than in the unfocused context for both French and English as produced by monolinguals of the two languages. In the meantime, in another study (to appear) on the acquisition of some L2 phonological features by L2 learners she found that voicing-dependent variation is one of the last features to be learned, when it is learned.

Investigation of the acquisition of the quantity distinction in this environment (i.e., pre-dental fricative) show results similar to those found before stops. All groups produced average ratios of short /æ/ over that of long /æ/ close to that of the monolinguals. In actual numbers, the beginning group had an average ratio of .62, the intermediate group had an average ratio of .61, and the advanced group had an average ratio of .51. The monolinguals, on the other hand, had an average ratio of .59.
Table 4.7. Means, standard deviations, and ratios for long and short vowels in pre-dental fricative environment for all groups in focused and unfocused contexts.

<table>
<thead>
<tr>
<th>A. Focused Context</th>
<th>Monoling. Group</th>
<th>Beginning Group</th>
<th>Intermed. Group</th>
<th>Advanced Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>106 (25)</td>
<td>160 (42)</td>
<td>131 (44)</td>
<td>105 (35)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>119 (28)</td>
<td>222 (59)</td>
<td>178 (45)</td>
<td>152 (44)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.89</td>
<td>.72</td>
<td>.74</td>
<td>.69</td>
</tr>
<tr>
<td>Long Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>182 (52)</td>
<td>271 (59)</td>
<td>224 (52)</td>
<td>235 (90)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>209 (48)</td>
<td>350 (99)</td>
<td>293 (75)</td>
<td>284 (90)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.87</td>
<td>.77</td>
<td>.76</td>
<td>.82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Unfocused Context</th>
<th>Monoling. Group</th>
<th>Beginning Group</th>
<th>Intermed. Group</th>
<th>Advanced Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>99 (19)</td>
<td>152 (42)</td>
<td>129 (34)</td>
<td>101 (34)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>109 (20)</td>
<td>206 (54)</td>
<td>171 (40)</td>
<td>153 (54)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.90</td>
<td>.73</td>
<td>.75</td>
<td>.66</td>
</tr>
<tr>
<td>Long Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>160 (47)</td>
<td>247 (58)</td>
<td>210 (51)</td>
<td>212 (85)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>186 (44)</td>
<td>323 (113)</td>
<td>258 (64)</td>
<td>261 (97)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.86</td>
<td>.76</td>
<td>.81</td>
<td>.81</td>
</tr>
</tbody>
</table>
4.2.3 The acquisition of the voicing effect in the environment before the fricatives /x/ and /γ/.

Figures 4.7a - 4.8d display the effect of voicing on vowel duration for the mono- and bilingual groups participating in the experiment. Figures 4.7a - 4.7d show the voicing effect in focused context and Figures 4.8a - 4.8d show the effect in the unfocused context.

As can be seen in Figures 4.7a and 4.8a, all even bars in each grouping for the monolinguals show greater duration than the odd ones; that is, they show a voicing effect on the duration of the preceding /æ/. This difference in duration shows that the voicing effect for the monolingual group in the pre-velar environment is consistently greater than those found in the pre-dental stop and fricative environments. The effect is, in fact, statistically significant at the .01 level for both short and long /æ/, though greater for short than for long /æ/ (See Section 4.1 for exact levels of power). The actual ratios are .79 and .78 for short /æ/ in the focused and unfocused contexts, respectively, and .89 for long /æ/ in both contexts. These results show again a consistent deviation form the expected pattern found for tense/long vowels in English where the effect has always been reported to be greater for tense/long than for lax/short correspondents (Rositzke 1939, Fairbanks and House 1953, Peterson and Lehiste 1960, Crystal 1982, Laeufer 1992). In Arabic, on the other hand, this is the second study in which the effect is consistently shown to be greater on short than on long /æ/. The other study was done in 1980 by Port et al. and reported the ratios of .82 for short /æ/ and .92 for long /æ/. No other study, to my knowledge, has compared the effect on long and short vowels in Arabic.
Figure 4.7. Effect of voicing on vowel duration before velar fricatives for long and short /æ/ in focused context as produced by (a) monolingual Arabs, (b) the beginning bilingual group, (c) the intermediate group, and (d) the advanced group.
Figure 4.8. Effect of voicing on vowel duration before velar fricatives for long and short /æ/ in unfocused context as produced by (a) monolingual Arabs, (b) the beginning bilingual group, (c) the intermediate group, and (d) the advanced group.
This "unexpected" effect could be attributed to the "hyperarticulation" of the long vowel by the native speakers to maintain the quantity distinction between it and its short cognate. In doing so, the voicing effect ends up confounded with the length distinction.

Surprisingly, the same trend of effect (i.e., greater for short /æ/ than for long /æ/) has been maintained by the non-native groups as well (See Table 4.9 for details). The ratios for the short /æ/ in the focused context do not exceed .70 whereas for long /æ/ they do not go below .81 for the different groups. In the unfocused context the ratios do not go above .68 for short /æ/ and do not go below .73 for the long /æ/. This outcome indicates that the non-natives have not transferred the greater effect found for tense/long English vowels, as opposed to the smaller effect for their lax/short cognates, into Arabic. They (non-natives) are most likely following the same strategy used by native speakers of Arabic of hyperarticulating the long /æ/ to distinguish it from its short correspondent.

The figures also show that not all monolinguals show the same effect in both contexts. Some subjects show greater effects than others. For example, subject number 6 shows a great effect that is not different from any non-native speaker. Sometimes the same subject shows the effect in one context but not in the other. Subject 4, for instance, shows hardly any effect in the focused context but a distinguishable one in the unfocused context. In general, however, all the subjects either show a distinguishable effect or a trend toward it. No subject shows opposite effect as it is occasionally the case in the pre-dental environments. This systematic outcome lends strong support for
the unquestionable existence of the voicing effect in the pre-velar environment.

The figures show that the non-native speakers have not acquired the voicing effect of Arabic regardless of their advancement level and the time they spent learning Arabic. All groups show greater voicing effect, in fact closer to English, than the native speakers. This outcome is consistent with the lack of acquisition in the other two environments included in the study (in the pre-dental stops and fricatives environments). Previous studies conducted on the same topic, and mostly on different languages, have shown conflicting results. Flege (1979), Mack (1982), and Suomi (1976) have reported lack of acquisition of the voicing-dependent vowel duration on the part of L2 learners. In all these studies L2 learners adhered to a voicing effect that is closer to the one found in their native languages than the one of the target language. On the other hand, Mitleb (1984) reports that advanced Jordanian learners of English have increased the difference from 5 ms as reported for Arabic to 20 ms for English data. This increase falls approximately half way between Arabic and English (the difference was reported at 50 ms in the same study for the monolingual American group). Mitleb interprets this difference as a gradual acquisition that falls within the approximative system suggested earlier by Nemser (1971) and Flege (1981). Laeufer (to appear) found that three out of five advanced French-English bilinguals had acquired English-like voicing-dependent vowel duration. One of the remaining two maintained a native-like difference and the last one showed a mid-point between the two languages. These findings from different studies attest to the lack of a
consistent pattern that bilinguals follow in learning the difference in the target languages.

Consistent with the findings of the previous two environments, the bilinguals tend to gradually acquire an overall vowel duration that is closer, but not identical, to that of the native speakers. Table 4.8 below provides the overall means for all the groups in focused and unfocused contexts for long and short /æ/.

Table 4.8. Overall means in ms for long and short /æ/ in focused and unfocused contexts for all the groups participating in the study.

A. Focused Context

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Monolingual</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/</td>
<td>111.5 ms</td>
<td>183 ms</td>
<td>133.5</td>
<td>123</td>
</tr>
<tr>
<td>/æ/</td>
<td>178.5</td>
<td>257</td>
<td>214</td>
<td>226.5</td>
</tr>
</tbody>
</table>

B. Unfocused Context

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Monolingual</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/</td>
<td>96.5</td>
<td>155.5</td>
<td>119</td>
<td>108</td>
</tr>
<tr>
<td>/æ/</td>
<td>152</td>
<td>223</td>
<td>179.5</td>
<td>201.5</td>
</tr>
</tbody>
</table>

With the exception of the performance of the advanced group for long /æ/, the bilinguals tend to progressively approximate the Arabic vowel length
as they advance in the language. The difference among all levels including the monolingual group has shown to be significant at P. \(<0.001\) (\(F(3, 159) = 15.21, P < 0.000\)). This outcome is consistent with those of the pre-dental stop and fricative environments. It is also in agreement with findings reported by Mack (1982), Elsendoorn (1984), and Laeufe (to appear).

The overall means of the advanced group for long /æ/ have been skewed by the performance of the subjects rated 67 and 77. When both subjects are excluded from the study the overall means of the advanced group drop to 188.4 for the focused context and 150.75 for the unfocused context. This drop, especially in the unfocused context, puts overall the means for the advanced group at a level that is not different from that of the monolinguals.

The pattern of the bilinguals acquiring the overall length of the vowel and the quantity distinction found in Arabic in all environments, and not the voicing-effect seems to indicate that some phonetic aspects of Arabic are easier than others for American students to learn. In particular, the overall vowel length and the temporal distinction seem to be among the easy ones and the voicing effect belongs to the difficult features. Similar results were found for French students of English (Laeufe, to appear). Laeufe found that vowel length of English is acquired early on by French students and therefore it falls within the easier set of features for native French. By contrast, voicing-dependent vowel duration belongs to the difficult set of features and thus requires a long time to be learned. It could be argued based on the results of these two studies as well as those of Elsendoorn (1978) for Dutch that the overall vowel length belongs to a "universal" set of easy-to-learn features and the voicing effect belongs to a set of difficult-to-learn features.
The role of focused context is evident on vowel length for the mono- and the bilingual groups alike. Vowel averages are significantly longer for all the groups in the focused context than in the unfocused one ($f (1, 159) = 10.70, P <0.009$). Table 4.9 below provides the average differences between the focused and unfocused contexts for each group. These average differences have been calculated by subtracting the overall means of the unfocused context from those of the focused context for each group.

Table 4.9. Average focus-dependent vowel differences in ms for all the groups participating in the experiment.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Mono.</th>
<th>Beginning</th>
<th>Intermed.</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/</td>
<td>15 ms</td>
<td>27.5 ms</td>
<td>14.5 ms</td>
<td>15 ms</td>
</tr>
<tr>
<td>/æ̃/</td>
<td>26.5 ms</td>
<td>34 ms</td>
<td>43.5 ms</td>
<td>25 ms</td>
</tr>
</tbody>
</table>

Results in this table support similar results reported earlier in the other two sections. They are also in agreement with those of Fourakis (1991) and Laeufer (1992). These scholars reported the vowel to be significantly longer in the focused context than in the unfocused one.
Table 4.10. Means, Standard Deviations, and Ratios for long and short /æ/ in pre-velar fricatives for all groups in focused and unfocused contexts.

<table>
<thead>
<tr>
<th>A. Focused Context</th>
<th>Monoling. Group</th>
<th>Beginning Group</th>
<th>Intermed. Group</th>
<th>Advanced Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>99 (21)</td>
<td>151 (24)</td>
<td>110 (53)</td>
<td>97 (44)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>124 (35)</td>
<td>215 (42)</td>
<td>157 (70)</td>
<td>149 (48)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.80</td>
<td>.70</td>
<td>.70</td>
<td>.65</td>
</tr>
<tr>
<td>Long Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>168 (47)</td>
<td>231 (50)</td>
<td>192 (49)</td>
<td>205 (58)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>189 (55)</td>
<td>283 (40)</td>
<td>236 (82)</td>
<td>248 (85)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.89</td>
<td>.82</td>
<td>.81</td>
<td>.82</td>
</tr>
</tbody>
</table>

Table 4.10 (continued)

<table>
<thead>
<tr>
<th>B. Unfocused Context</th>
<th>Monoling. Group</th>
<th>Beginning Group</th>
<th>Intermed. Group</th>
<th>Advanced Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>85 (15)</td>
<td>121 (31)</td>
<td>95 (40)</td>
<td>88 (36)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>108 (23)</td>
<td>190 (38)</td>
<td>143 (59)</td>
<td>128 (35)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.78</td>
<td>.63</td>
<td>.66</td>
<td>.68</td>
</tr>
<tr>
<td>Long Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Voice</td>
<td>143 (38)</td>
<td>189 (45)</td>
<td>153 (39)</td>
<td>177 (68)</td>
</tr>
<tr>
<td>+ Voice</td>
<td>161 (40)</td>
<td>257 (43)</td>
<td>206 (71)</td>
<td>226 (97)</td>
</tr>
<tr>
<td>Ratio</td>
<td>.89</td>
<td>.73</td>
<td>.74</td>
<td>.78</td>
</tr>
</tbody>
</table>

4.2.4 Discussion

Several issues have been addressed in this section. The most prominent among them is the acquisition of Arabic voicing-dependent vowel duration
by the non-native groups and the factors that either enhance or weaken this process. The factors contributing to the acquisition process, positively or negatively, include the role of a focused context, place and manner of articulation, and vowel length. Other issues that have shown to be of importance are the acquisition of the overall vowel length, the quantity contrast, and lack of uniformity among subjects in acquiring the differential vowel duration.

The acquisition of the voicing-dependent vowel length has been limited to the environment before stops for the advanced group only. The rest of the environments and contexts including the unfocused environment before stops for long /æ/ do not show any traceable pattern. In fact, I am more inclined to consider the performance of the advanced group in this environment (before stops) accidental rather than attribute it to acquisition. The lack of such acquisition should not come as a surprise, especially if we compare it with the results of the very few studies that investigated the same phenomenon cross-linguistically. Flege and Port (1981) arrived at similar results when they found that Saudi learners or Arabic did not acquire the English voicing-dependent patterns regardless of how many years they spent learning English and living in the United States. Similarly, Mack (1982) found that the French-English bilingual had not acquired the voicing-dependent pattern of English either. Laeufre (to appear) found that four out of the five bilinguals who participated in her study would identify the target sound with a corresponding variant in their native language (French) where the voicing effect is the greatest, thus the closest to English, and transfer this variant into English as a first step. Then, in the second stage, they would extend this
variant distributionally to the rest of the variants in English producing
differential vowel duration that is close but not identical, and in most cases
far from identical, to English. At the third stage, bilinguals would start
approximating the differential vowel duration found in English. According to
Laeufer, duration of exposure did help only one bilingual out of five that
participated in her study go beyond the distributional extension stage in
approximating the differential vowel duration of English though this stage
was reached in the acquisition of other "easier" phonetic aspects such as the
VOT of English by most subjects.

The question remains as to why bilinguals tend to fail in acquiring
differential vowel duration of the target language. One way to account for this
failure is to attribute it to what Flege (1987) calls "Equivalence Classification".
According to this model, bilinguals tend to identify the sound in the target
language, in this case /æ/, with a "similar" correspondent found in the native
language. By doing so, they overlook the phonetic details that differentiate
the two sounds and end up substituting the native sound for the target one.
This identification and the substitution tend to block L2 learners from
acquiring the target sound authentically. Duration of exposure and
advancement in learning L2 do not usually help the bilinguals acquire the
phonetic details when the two sounds in the two languages are similar.

Extending this model to the sound /æ/ used in this study, it would not be
surprising to suggest that American students had equated the Arabic /æ/ with
the American English /æ/ in all contexts. In doing so, it would be reasonable
to assume that authentic acquisition of the Arabic /æ/-variant has been
blocked by the variant found in English.
In addition to the similarity between the two /æ/ in English and Arabic, the two languages share a host of similarities with regard to the phonetic characteristics of the following consonants that might contribute to the blockage. For example, neither English nor Arabic allows for a strong aspiration of stops word finally the way French does. To clarify, in French the aspiration of stops word finally is so strong that it is often perceived as a schwa. This strong aspiration causes an underlying monosyllabic word to resyllabify creating a surface bisyllabic one (Lauefer 1992). The final stop, therefore, which initially constituted a coda for the monosyllabic word, would split from the first syllable and become an onset for the created second syllable. This process makes French saliently different from other languages like English and, according to the equivalence classification, easier for speakers of either language (English and French) to perceive and acquire since it is not shared in both languages. In addition, neither English nor Arabic devoices word finally in a way similar to German, Polish, or Russian. In sum, the similarity between English and Arabic in an /æC/ sequence, where C stands for a consonant, is not limited to the /æ/, but it extends to include some of the phonetic features of the following consonant. Thus, American learners of Arabic would equate not only the /æ/ in the two languages, but also the similarities in voicing of the following consonant. Such similarities between the two languages would lend more support to the equivalence classification and makes the acquisition of the voicing-dependent vowel duration in Arabic harder for American students than that of French, Russian, German, etc.
Another way to account for the lack of acquisition of differential vowel duration is to investigate it in comparison to the acquisition of other phonetic features under consideration. Following this approach, we find that some L₂ phonetic features are acquired early on in learning a foreign language and others are either acquired at an advanced stage or never acquired at all with all their phonetic details. For example, the acquisition of the quantity contrast between long and short /æ/ has been unanimous among all the groups in this study regardless of their level. All groups showed a systematic distinction between short and long /æ/ that is not different from that of the monolingual group in all contexts. Meanwhile, the study has shown a gradual approximation of the Arabic overall vowel length by the bilingual group as they advanced in Arabic. This gradual approximation has been uniform in all contexts and environments. It (the approximation) can be characterized as the longest for the beginning group, shorter for the intermediate group and shortest, and therefore closest to the monolingual group, for the advanced group\textsuperscript{32}. By contrast to these two features, there is no systematic acquisition of the voicing-dependent vowel duration.

The different levels of acquisition by the bilinguals of these three phonetic features of Arabic (i.e., quantity distinction, overall vowel length, and voicing-dependent vowel duration) indicate that some of these features are easier than others for the American students to acquire. Obviously, the phonemic distinction between long and short /æ/ seems to be the easiest of the three, the overall vowel length seems to fall next on the hierarchy, and

\textsuperscript{32} The overall averages for the advanced group will be shortest after the subjects rated 67 and 77 have been dropped. These two subjects tended to systematically produce rather long /æ/ s in all contexts that skewed the inter-subject averages.
the differential vowel duration resides on top of the scale as the most difficult.

Previous research done on the same topic has reached similar conclusions for other languages. Laeufer (to appear) has found that some phonetic features of English are harder than others for native French to acquire. Therefore, she established a model in which phonetic features of English fall into a scale that ranges from the easiest features to acquire to the most difficult. According to her model, the overall vowel duration of English is among the easy features for native French students to acquire and the differential vowel duration is among the most difficult. Though this model cannot be extended to other languages due to lack of a universal basis for the classification, similar models can be established for specific languages based on the findings of studies like the one reported in this dissertation, and eventually, when enough studies have been conducted on various languages a "universal" model might be reached. Following up on the same idea and based on the findings of the study described here, it can be posited that Arabic has a similar scale. According to this scale, the quantity contrast should be considered the easiest to acquire for American students, overall vowel length comes next, and the hardest is voicing-dependent vowel duration. If this scale were proven to be true, then it should have important pedagogical implications. Perhaps a corresponding scale can be followed in teaching the pronunciation of Arabic to American students. In other words, the syllabus may start with features that are easy to learn leaving those that require extensive experience till later stages. Additionally, there might be decisions as to what can and should be taught and what cannot and should not be taught.
Of interest and perhaps of direct relevance is the findings of other studies done on the same topic. They all (other studies) agree that the acquisition of voicing-dependent vowel duration is among the last features to be acquired, if it is ever acquired at all. These studies have covered languages as different as Arabic, English, French, and Dutch (Mack 1982, Elsendoorn 1980, Laeufer (to appear)). This common agreement may provide the basis for a claim that voicing-dependent vowel duration is hard to acquire across all languages, therefore it is intrinsically difficult to acquire.\footnote{It should be noted that the scale of difficulty is based primarily on intuitive appeal. For such a scale to become universal there should be an independent criteria to control and standardize the definition. These criteria can be arrived at in many ways such as investigating which phonetic features of $L_1$ are acquired early on (versus those acquired at a later stage) by children, investigating the acquisition of such features in a pidginized and creolized environment, etc.}

The role of focused context has been limited in this study to the length of the overall vowel and not the voicing effect. Both long and short $/æ/$ have been longer in the focused than in the unfocused context. This is understandable for all groups because focus induces, among several things, lengthening. Similar results were found by Summer (1987), Fourakis (1991), Laeufer (1992). One of the interesting consequences of this finding is the realization that non-natives tend to recognize the global structure of the utterances before they get into the minute phonetic details of the target language. In this case, they seem to have recognized the intonation structure and identified the words with the highest degree of prominence in these utterances before they could realize the small effect voicing has on the preceding vowel in Arabic. This practice should be taken as relative to that applied to most subjects because, as discussed later on, some subjects especially from the beginning group had a brief pause after each word.
The lack of a noticeable voicing effect that is due to context for both
groups can be attributed to two main factors, the small effect voicing has in
Arabic and the quantity distinction found in Arabic. Because the effect is
small it will allow for minimal variation due to context or any other factor.
The quantity distinction may block the effect, especially for long vowels, from
variation so that these vowels will not shorten enough to overlap with their
short cognates.

One of the unexpected outcomes of this study is the greater durational
effect final obstruent voicing has on the short /æ/ than on its long
correspondent, especially in the pre-velar fricative environment. This is
represented in the greater ratios for the long /æ/ than for its short
counterpart. Interestingly, the effect was systematic across the mono- and
bilingual groups alike. This outcome is in conflict with Klatt's compressibility
(1975) discussed early Chapter II. According to Klatt, longer vowels should
show greater effect because they can afford to due to their intrinsic length. In
other words, long vowels can be compressed to a greater degree than short
ones while still maintaining their identity. By contrast, short vowels cannot
compress as much because they may lose their identity and perhaps turn into
a schwa or another vowel. Numerous studies, mostly based on English, have
supported this view (Rositzke 1939, House 1961, Klatt 1975, Crystal and House
1982, just to mention a few). All these studies and many more have reported
smaller ratios, thus greater effect, for tense/long vowels than for short ones.
The results in this experiment as well as those of Port et al. (1980) have shown
that the opposite is true for Arabic. One way to account for this mismatch
between the two languages is to attribute it to the contrastive length found in
Arabic (and absent in English). Long vowels in Arabic cannot be compressed beyond a certain point to avoid overlapping with their short correspondents, and thus losing their identity. Short vowels, on other hand, will not overlap with any other vowel when shortened. Therefore, they may permit greater voicing effect. Native and non-native speakers alike seem to have internalized this difference and acted accordingly.

The study shows that there is no positive correlation between the global judgment of somebody's accent and his or her mastery of some of the phonetic details of L2. The support for this finding comes from the lack of correspondence between the subjects' ratings as given by the judges and the acquisition of the voicing-dependent vowel duration. Some of the bilinguals in the beginning group show smaller voicing effect than some in the intermediate or advanced group yet their ratings placed them at a beginning level. The same is true for some subjects in the advanced or intermediate groups. This outcome seems to represent a trend in cross-linguistic studies. It has been observed in the studies of Jun and Cowie (1991) and Peng (1993), Laeufer (to appear). This conclusion raises several questions: should there be a correspondence between perception and production of speech in all phonetic details? Do raters rely on global impression in judging the accentedness of L2 learners, or on specific deviations from the norm, or on both? Which features "count" more in these evaluations, if it is not a global impression that leads to perception? How does production of the phonetic details relate to the overall perception of the speakers of L2?

In conclusion, the study has shown that advancement in Arabic has not helped the bilinguals acquire the voicing-dependent vowel duration of
Arabic, but it has helped them approximate, to some degree, the Arabic overall vowel length. The effect of focused context has also been limited to the overall length of the vowel; it did not affect the voicing effect in any environment for any group. The study has also revealed that there is no correspondence between vowel length and the degree of the voicing effect, at least for Arabic. Unlike English, Arabic has shown greater voicing effect on short /æ/ and on its long cognate. Finally, it has been shown that global rating does not necessarily correspond to the acquisition of some phonetic details.

4.3 The effect of learning L2 on L1 (Reverse phonological interference)

Recent research on phonological and phonetic interference as well as personal observation have shown that interference is not only unidirectional (i.e., from L1 to L2). In many cases learning L2 can affect the pronunciation of some aspects of L1, thus making phonetic interference bidirectional (i.e., from L1 to L2 and vice versa) (Caramazza et al. 1973, Flege 1987, Flege and Efting 1987, Jun and Cowie 1991, Peng 1993).

To start with a personal example, I feel I have lost the authentic pronunciation of the Arabic trilled /r/ in certain contexts as a consequence of my futile attempts to produce the American /r/ authentically, and unless I consciously pay attention to its pronunciation, the chances that it will not come out as an authentic Arabic /r/ are very high. Similarly, reviewing some of the studies that have been conducted on phonological interference shows several

---

34 I realize that the personal example given here may stem from an extensive experience in learning and using English that is not comparable to the experience of most of the subjects that participated in this study. However, the purpose of giving this example is not to create a comparison. Rather it is meant to document that cases of reverse phonetic interference are possible, a subject that has not been recognized until very recently.
cases of reverse phonological interference. The following paragraphs describe some of the findings of the few studies that have addressed this issue and provide the empirical evidence to support my hypothesis.

Caramazza et al. (1973) found that French-English bilinguals have shifted the 50% cross-over point of VOT for French (their native language) to an intermediate point between English and French for some French stops, thus making French VOT value closer to English than it normally is. To clarify, English is well-known as having a rather long VOT (in the range of 50 - 80 ms) and French a considerably shorter VOT (in the range of 20 - 35 ms). These ranges were highlighted by the monolingual groups that participated in Caramazza et al.'s study and several other subsequent studies (Flege and Hellenbrand 1984; Flege 1987; Serniclaes et al. 1984). The most likely explanation for the shift of the bilingual group is the effect of L2 (i.e., English) on L1.

In the meantime, VOT operates as a strong cue for phonemic distinction between voiced and voiceless stops in English but not as strong in French. This conclusion is evident in the graphs provided in Caramazza et al.'s study. The monolingual English group shows a sharp monotonic slope that distinguishes the voiced from the voiceless sets while the French monolingual group shows a jagged slope that extends over a range of approximately 30 ms duration. The jagged line is an indication that VOT is not phonemic in French. Examining the bilingual group's production of the French sound /k/, for example, shows that their slope is as sharp as that of the English monolingual group. This clearly indicates that VOT has begun to occupy a stronger cue for phonemic distinction in the French production of this group.
Similar conclusions were arrived at by Flege (1987). He found that advanced native French speakers of English produced French /t/ with significantly longer VOT value (51 ms) than that of the monolingual French group (33 ms). Conversely, advanced native English speakers of French produced English /t/ with considerably shorter values (49 ms) than the monolingual English group (77 ms). In both cases we notice that VOT values of the groups’ native language are closer to those of the target language than those of the monolinguals’ values. The beginning group of native English speakers learning French, on the other hand, produced VOTs of equal values to both English and French that are not significantly different from those of the monolingual English group. Flege attributes shift of values for the advanced to the effect L2 had on L1.

In a different and more recent study Cowie and Jun (1991) found that experienced Korean speakers of English have altered the formant values of some Korean vowels. This alteration was viewed by the authors as an example of reverse phonological interference. Similarly, Peng (1993) found that some native speakers of Amoy, who were proficient in Mandarin, producing the sound /h/, native to Amoy and absent in Mandarin, with the acoustic characteristics of the Mandarin "similar" counterpart /x/. More specifically, proficient Amoy speakers of Mandarin produced the sound /h/ with high concentration of energy around F3 and lowered F2 and F4. These are acoustic characteristics of the Mandarin sound /x/. By contrast, less proficient Amoy speakers of Mandarin produced /h/ with high energy, represented in striations for higher formants that are intrinsically characteristic of /h/.
Following up on the idea of reverse phonological and phonetic interference, this part of this study was designed and implemented to assess whether advanced American learners of Arabic would produce voicing-dependent vowel duration of English with values closer to Arabic than those of monolingual English speakers.

4.3.1. Method
4.3.1.1 Subjects (See the section describing subjects in Chapter III)
4.3.1.2 Test materials and procedure

Two pairs of English monosyllabic minimal pairs were used as stimuli. The pairs differed from each other with regard to voicing characteristics of the final consonant and the manner of articulation of the final consonant (stops versus fricatives) (See Table 4.11 below for a list of the minimal pairs).

The words were recorded ten times each in two contexts, focused and unfocused environments. Means were calculated for all ten readings for each target word across the two contexts rather than calculated separately for each context.

The four target words, along with sixteen foils were inserted in the carrier sentences and typed on 3X5" cards. Cards were shuffled randomly and each card was read ten times by each subject. The total number of tokens included in the study was 960 (4 words X 10 readings X 24 subjects = 960).
Table 4.11 List of target words and carrier sentences used in the study

1. List of Target words
   - Stops
     \(/bæd/ \quad "bad" \quad /bæt/ \quad "bat"
   - Fricatives
     \(/bæz/ \quad "Pl.\ form\ of\ baa" \quad /bæs/ \quad "bass"

2. Carrier sentences
   (a) Say X again (focused environment, in a response to the question 'what am I to say again?')
   (b) Don't say X again (unfocused environment, in response to the question 'what am I not to do again?')

4.3.1.3 Recording, measurement and segmentation (Identical to those measures in Chapter III for the acquisition of voicing-dependent vowel duration by American bilinguals)

4.3.2. Results

A. Effect of voicing of vowel duration in the environment before stops

   Figures 4.9a - 4.9d below display the effect of voicing on vowel duration in environment before stops for all four groups. Figure 4.9a shows the effect of voicing on vowel duration for the monolingual group, Figure 4.9b for the beginning bilingual group, Figure 4.9c for the intermediate bilingual group, and Figure 4.9d for the advanced bilingual group. Subjects are displayed on the horizontal axis (monolinguals by numbers 1 through 6 and the bilinguals by their rating scores obtained from judges). Vowel duration is displayed on the vertical axis. Dark bars represent average vowel durations before voiced stops and bars with slanted lines represent average vowel duration before voiceless
Figure 4.9. Role of reverse phonetic interference on the voicing-dependent vowel duration in English before stops as produced by (a) a group of monolingual Americans, (b) the beginning bilingual group, (c) the intermediate
stops. Voicing effect would be represented with the dark bars being significantly higher than the ones with the slanted lines. Overall ratios for each group are given below each figure. The ratios stand for the outcome of dividing the overall means of vowel duration in pre-voiceless environment by the overall means of vowel duration in pre-voiced environment (both means are given below the figures also).

It should be acknowledged at the beginning that the ratio of .81 for the monolingual group is higher than what has been repeatedly reported in the literature. All of the studies I have examined, with the exception of Mitleb's (1984) and Laeuffer's conclusions (1992), show lower overall ratios. Mitleb reported a ratio of .80 and Laeuffer, incidentally, reported a ratio of .81 in the non-focused environment. All other environments investigated in her study show lower ratios, hence greater voicing effect. The outcome of this study should not be surprising, however, since the voicing effect in many studies has been reported to exhibit a rather wide range.

Figure 4.9b shows a ratio which is lower than that of the monolingual group, .78. This finding should not be unexpected since the students included in this group did not have extensive exposure to Arabic. Figure 4.9c shows a ratio for the intermediate group that is higher than any other group including the monolingual group. Figure 4.9d, on the other hand, shows an identical ratio to that of 4.9b, .78. This fluctuation in values of ratios shows a lack of consistency with regard to the effect of L2 on L1. In other words, the results do not show any indication that there is a gradual reverse effect on L1 as these
students advanced in learning Arabic. The strongest piece of evidence against
the interference comes the advanced group. If there were to be any reverse
interference, it would most likely be shown in the performance of this groups
since they had the most extensive exposure to Arabic. Correlation tests run on
all the subjects did not show any correlation between the performance of these
subjects and their ratings ($r = .10$).

It should be noted, however, that inter-subject variability is very high
within each group. Some subjects show smaller voicing effect than others
regardless of their level. For example, the subject rated 61 in the intermediate
group shows no voicing effect at all. In fact, the average vowel duration before
voiceless stops is .8 of a point higher than that of the average vowel duration
before voiced stops. The absence of the voicing effect in this case can be
attributed to the fact that this subject flapped all the final /t/s and /d/s in the
context given in the study, hence both environments became neutralized
where the effect of voicing is usually greatly diminished (Sharf 1962, Fox and
contrast, the subjects rated 77 and 87 maintain English-like voicing effect
though both of them have been studying and teaching Arabic for 22 and 8 years,
respectively.

In sum, it can be concluded that no reverse phonetic interference has been
practiced in the performance of this group. This conclusion should not be
surprising in light of the fact that the acquisition of the voicing effect in a target
language has been proven to be a hard task to be achieved as argued in the
previous section of this chapter.
B. Effect of voicing on vowel duration in before fricatives

Figures 4.10a - 4.10d show the effect of voicing in pre-fricative environment for all four groups. The layout of these figures is identical to those of Figures 4.9a - 4.9d above.

Figure 4.10a shows a small voicing effect for the monolingual English group that has been unparalleled in any other study I am aware of. This small effect conflicts with the commonly accepted view that the voicing effect is greater in the environment before fricatives than before stop (House and Fairbanks, 1953; Peterson and Lehiste, 1960; House, 1961; Laeufer, 1992). Being as small as it is makes the effect of voicing appear marginal in the rest of groups in cases where there might be one.

Both beginning and intermediate bilingual groups show overall ratios that are smaller than that of the monolingual group. The advanced group, on the other hand, shows a ratio that is one point higher than that of the monolingual group. Such an increase is minute and obviously cannot be significant.

The results also show some subjects from the beginning and intermediate groups reducing the effect of voicing to a point that is consistent with that of the majority of the native speakers of Arabic in the pre-dental environment. For example, the subject rated 34 in the beginning group shows a ratio of .94.

The results also show some subjects from the beginning and intermediate groups reducing the effect of voicing to a point that is consistent with that of the majority of the native speakers of Arabic in the pre-dental environment. For example, the subject rated 34 in the beginning group shows a ratio of .94.
Figure 4.10. Role of reverse phonetic interference on the voicing-dependent vowel duration in English before fricatives as produced by (a) a group of monolingual Americans, (b) the beginning bilingual group, (c) the intermediate bilingual group, and (d) the advanced bilingual group.
Similarly, the subject rated 59 in the intermediate group shows an identical ratio. This performance, however, remains limited to two subjects out of eighteen. That is, the majority of the students, especially the advanced ones, do not show a voicing effect that can be characterized as close to that of the native speakers of Arabic. In fact, the correlation test that compared the performance of all the subjects with their ratings did not exceed .34 (i.e., $r = .34$). This poor correlation is a strong indication that advancement in $L_2$ did not result in gradual reverse effect.

In conclusion, the performance of the majority of the subjects show a lack of reverse effect that is not different from that absent in the environment before stops.

4.3.3. Discussion

The results reported in this study did not show any traceable trend of reverse phonetic interference. Figures 4.11a and b show the correlation between subjects ratios and their ratings. As can be seen in the figures, the correlation is extremely poor for both environments ($r = .10$ and .34 before stops and fricatives respectively). These results are not systematically consistent with the assumption that more experience in $L_2$ leads to greater reverse phonetic interference. The beginning group does not collectively show any voicing effect ascribable to Arabic in either environment (i.e., in before stops and before fricatives). To the contrary, it shows lower ratio than that of the monolingual group in both environments. The intermediate group, on the other hand, shows a higher ratio than that of the monolingual group in the environment before stops, but a lower ratio before fricatives. The advanced group, on the
other hand, shows the opposite relationship, a lower ratio in the environment before stops than that of the monolinguals and a slightly higher one before fricatives.

There seem to be several ways to account for the lack of significant reverse phonetic interference. First, the relatively high ratios observed for the monolingual group, especially in the pre-fricative environment, can be responsible for the absence of interference. These high ratios create a situation by which English would not look considerably different with regard to the effect of voicing from Arabic, unlike what has been frequently reported in the literature, and subsequently any form of interference would not appear to be significant. If the ratios of the bilingual groups were to be compared with the findings of other studies (e.g., Chen, 1970; Peterson and Lehiste, 1960; Mack 1982; Laeufer 1992), we would certainly report a significant interference effect. However, since our monolingual group was intended to represent the norm, our findings cannot be compared with any findings other than those of the monolingual group. The cause for the low voicing effect on vowel duration by the monolingual group, especially in pre-fricative environment, remains unexplainable.

Another reason for the lack of the interference could be the method and environment in which the subjects learned Arabic. Most of the subjects participating in the experiment learned Arabic in an English speaking environment. Thus, their exposure to Arabic was limited to the classroom setting. By comparison, those who showed reverse phonetic interference in Peng’s study, for example, happened to have lived for many years in the target
Figure 4.11. Correlation between subjects’ ratios and their ratings, (a) for inviolment before stops and (b) before fricatives.
language environment. Besides, the target language was their main medium of daily communication. Such extensive exposure could be more conducive to reverse phonetic interference.

A third factor that could have blocked the interference is the difficulty of the feature under consideration. Most of the research done on voicing-dependent vowel duration reported lack of acquisition of this feature by L2 students. That is, most L2 learners end up transferring the voicing effect characteristic of their native language to the target language. This trend was clear in the works of Elsendoorn (1980), Mack (1982), Laeufer (to appear), and in this study. It is reasonable to assume at this point that for such feature to have a reverse effect on the native language it has to be acquired first. Since it is rarely acquired in the first place, it would be quite natural for it not to have any reverse effect.

It should be noted that the voicing effect has shown to be in some cases subject to the manner of articulation within the same subject. For example, the subject rated 61 shows no voicing effect in pre-stop environment and considerable effect in pre-fricative environment. Flapping of the alveolars could be responsible for the lack of the effect. Conversely, the subject rated 59 shows much less variation in the pre-fricative than in the pre-stop environment. This finding conflicts with the generally accepted view that vowel duration variation is greater before fricatives than before stops cross linguistically.
4.3.4. Conclusions

The results of this study have shown no trace of reverse phonetic interference. It is possible that a greater degree of interference has been masked by the extremely high ratios shown for the monolingual group. More likely, the absence of interference could have been resulted from the lack of extensive and extended exposure to Arabic or the difficulty of the feature under consideration, or all the two factors combined. Inter- and intra-subject variability has been common in this study.

4.4 Characteristics of Arabic produced by the bilingual groups

In addition to showing the relevant information concerning the main issue of this dissertation (i.e., voicing-dependent vowel duration and the factors that affect it), the data have revealed several other characteristics that can be of help in understanding accentedness in a foreign language in general and the English-accented Arabic in particular. These characteristics range from a full substitution of a sound (phonemic interference) to a slight modification that can be limited, for example, to lengthening or shortening of that sound or the sequence in which that sound falls (phonetic interference). In this section a list of these characteristics are presented and described briefly. Samples of spectrograms are also presented as evidence for some of these characteristics.

It should be emphasized at the outset, though, that the list of characteristics provided in this section is not meant to be exhaustive nor is it applicable to every participant in the experiment. In some cases, however, a considerable number of participants within one level (beginning, intermediate, or
advanced) and some sometimes across levels reveal the same phenomenon. The purpose of this presentation is to highlight the main features of the sample of Arabic used in this experiment as produced by American students in the hope that it might be of help for either further investigation or for pedagogical purposes. It should also familiarize some readers with the techniques that American learners of Arabic use in the production of some Arabic, especially the velar and pharyngeal Arabic fricatives.

Among the most common techniques that the bilinguals used is the over lengthening of long /æ/. Eight participants out of total eighteen from the different groups showed above average length for long /æ/ in most contexts and for apparently different reasons. In words like /baː:t/ and /baːd/, the lengthening process cannot be more than an exaggerated attempt to distinguish it from its short cognate. All sounds found in these two words have "similar" counterparts in corresponding English words. Therefore, there is no reason, other than the lengthening distinction found in Arabic, not to equate these sounds with their English correspondents and produce /æ/ comparable in its length to that of English. Figure 4.12 is given below for illustration. It is a sample of the word /baːd/ as produced in the focused context by the subject rated 61. It shows that the duration of /æ/ exceeds 400 ms; a length that far exceeds the 177 ms length reported earlier in section 4.1.1 for the average monolinguals in the environment before stops. In some cases the subject went on lengthening until he/she ran out of breath and started laryngealizing the vowel (he started showing laryngeal striations at the end of the vowel).
Figure 4.22. Extra long /æ:/ in the word /bæd:/ (over 400 ms).

Figure 4.13 Extra long /æ:/ before /γ/. 
Figure 4.14 Extra long /æ:/ before /ʌ/.

Figures 4.15 Devoiced and long /γ/. 
Figure 4.16 One of the variants of /γ/

Figure 4.17 /ɡ/ substitutes for /γ/ in /sæɡ/ as produced by the subject rated 35
Figure 4.18 /sæːγ/ as produced by the subject rated 18.

Another context in which long /æ/ was lengthened above normal is when it was followed by a voiced velar or pharyngeal fricative. Figures 4.13 and 4.14 illustrate this lengthening. Figure 4.13 was taken from the subject rated 67 and Figure 4.14 was taken from the subject rated 61. In both cases /æː/ has been over lengthened beyond the average for the monolinguals. One of the common characteristics of the sequence /æːŋ/ (Figure 4.14) is that /ŋ/ would blend into /æː/ and the demarcation between the two would be completely lost.

In a few cases the whole sequence would consist of /æː/ that ends in a few laryngeal striations. Some students ended the sequence with a schwa. That is,
the sequence /æːs/ would be perceived as /æːsə/. I believe that the addition of /ə/ to the sequence represents a forceful attempt by the bilinguals to produce the pharyngeal fricative.

It should be noted that some bilinguals do not perceive the distinction between /æ/ and /ə/, especially when they are produced in connected speech. This should be understandable since the sounds are acoustically very similar. The schwa perceived at the end of the sequence could be an /æ/ that has been centralized due to its position in the word.

The production of /y/ in the sequence /æːy/, on the other hand, displays great variation within the data. It is often devoiced and longer than average. Figure 4.15 illustrates this process. This sample was produced by the subject rated 22. In other cases, /y/ takes the form of lower amplitude and weaker formants that extend into those of the preceding /æ/. In this case, no obvious transition interval exists between the two sounds (i.e., /æː/ and the following /y/). Lack of such transition interval is an indication of an unobstructed vocal tract that resembles the vocal tract shape used in the production of /h/. The only difference is the existence of a few "irregular" vocal cord pulses that appear on the spectrogram. Figure 4.16, produced by the subject rated 34, exemplifies this form of /y/.

Lengthening of sounds was sometimes extended to the word initial fricative. For example, the sound /s/ in words like /sæːγ/ was lengthened by some subjects. This lengthening was apparently a technique by the subjects to give themselves enough time to prepare for the production of the "new" sound found at the end of that word. For example, in producing the word /sæːx/, some subjects produced a rather long /s/ in preparation for the
production of /x/. In most cases the lengthened fricative was followed by a shortened /æ/ as if the process was reciprocal. In other words, when the fricative is lengthened the following vowel will be shortened and when the fricative is short or shows average length the vowel is lengthened. In either case the purpose, I think, is one of buying time to prepare for the production of the velar or pharyngeal. Another motive for the lengthening that might be more applicable to beginning students was unfamiliarity with the word used in the study. These students needed the time to be able to spell the word before them.

Another common technique used by the beginning students to help them produce the word correctly was slow reading; some beginning students paused briefly after each word. In so doing, these subjects could not follow the intonation pattern of the utterance and the focused/unfocused context difference was neutralized.

In addition to lengthening and shortening, substitution was a common technique among the bilinguials. The most frequent substitution examples used by the bilinguials were the use of /g/ for /ɹ/, /ʔ/ for /s/, /x/ for /γ/, and /h/ for /h/. Figure 4.18 shows the use of /g/ for /γ/ in the word /sæγ/ as produced by the subject rated 35.

One of the unexpected performances was a brief pause in the middle of the vowel by subject rated 18. This subject had a habit of starting with a central vowel (instead of the expected low front vowel) then pausing for a few milliseconds that were usually followed by a few irregular vocal cord pulses and then continuing with a pharyngealized front low vowel. This behavior was systematic in most long /æ/ s before stops and fricatives alike. The pause
and shift in vowel quality represent an unnecessary artifact that is hard to explain.

Several subjects centralized the short /æ/ in the words /bæd/ and /bæt/. They produced a schwa instead of the expected front low vowel, especially in the unfocused context. This centralization did not spread to words that end in a fricative. This behavior shows a clear case of interference from English where short unstressed vowels tend to centralize.
Chapter V

Conclusions, Discussion, and Implications

In this chapter I summarize the conclusions of the study, discuss whether they fit within the expected patterns, and present some of their implications on teaching pronunciation to L₂ students.

5.1 Summary of conclusions and discussion

The study conducted in this dissertation was designed to answer three questions: whether the voicing effect on the preceding vowel exists in Arabic just like most languages or not; whether adult American students learning Arabic acquire the voicing effect in Arabic as they progress in learning the language; and whether advancement in learning Arabic causes the advanced learners to alter the voicing effect in their native language (i.e., whether there will be a reverse effect on the native language as a result of advancing in L₂).

With regard to the first question, findings of this study have shown that the voicing effect in Arabic is similar to that found in most languages. These findings therefore support the universality of the voicing effect on the duration of the preceding vowel. The debate over whether the effect is physiologically conditioned (i.e., universal) or learned has been a long one. The effect was first documented instrumentally in 1937 by Heffner. The issue of it being universal was first argued for by Belasco (1953) who attributed it to the Force or Articulation hypothesis (See chapter II for a description of the hypothesis).
Zimmerman and Sapon (1958) investigated the effect in Spanish and reported that the effect was weak, thus concluding that the effect must be language-specific. Two years later Peterson and Lehiste (1960) concluded that the effect is conditioned but the degree of it could be learned. Delattre (1962) criticized Zimmerman and Sapon's study and argued strongly that the effect must be universal attributing it to the Force of Articulation hypothesis. Chen (1970) extended the investigation to French, Russian, Korean in addition to English and reached the same conclusion of Peterson and Lehiste. A series of studies were conducted in the seventies and eighties (Klatt 1973, 1975, &1976; Umeda 1974; Crystal and House 1982, 1988) on vowel duration in connected speech and found that the effect was non-existent or limited to specific environments such as phrase-final position. In essence, these studies provided evidence showing that the effect does not exist in most contexts of connected speech.

The strongest evidence against the universality of the effect was reported in 1979. Flege, who conducted his study on Saudi Arabic in that year, and Keating, who investigated the effect in Polish and Czech in the same year also, concluded that the voicing effect in these languages was insignificant statistically, therefore implying that the effect must be language-specific and thus learned. Finally, Laeufer (1992) re-examined the effect in French in various contexts and environments and concluded, unlike all previous studies conducted on French, that the effect in French could not be different in magnitude from that of English. According to her, previous studies seemed to have overlooked certain language-specific features that masked the effect and made French look as if it had smaller magnitude. In effect, she provided evidence supporting the universality of the voicing effect.
The current study has shown the effect not to be absent in Arabic as it had been reported by Port and Flege (1981) and Mitleb (1983). It was argued in Chapter IV, Section 4.1.3, based on the findings of this study as well as those of Port et al. (1980) and Mitleb (1984), and following Lleufer’s notion of seeking language-specific masking factors, that Arabic does not differ from most languages with regard to the voicing effect. This outcome, therefore, provides a strong indication that at least a portion of the voicing effect is conditioned. The fact that English remains the only language to show a greater effect than most languages can be attributed to a form of learning by which native speakers of English apply a low level phonetic realization rule that extends the effect beyond the conditioned level.

With regard to the second question, results from the acquisition experiment indicate a hierarchy of difficulty that American students of Arabic follow in acquiring the phonetic features/parameters of Arabic. According to this hierarchy, some features are easier than others and therefore are learned at the early stages of commencing to learn Arabic; and other features are more difficult and are either learned at later stages or never learned at all. Following this assumption, the quantity contrast found in Arabic between long and short \(/æ/\) seems to be among the easiest of features to be acquired. All groups regardless of their level have phonemic contrasts (see chapter for actual ratios) that are not that different from the monolingual performance. Vowel length \(/\) falls next on the scale of difficulty; it is harder than the quantity distinction to be acquired by the bilinguals. Results of the different groups showed gradual

---

34Studies conducted on various languages tend to report a difference in vowel duration that range from 10 to 25% between the pre-voiced and pre-voiceless contexts (See chapter two for reported differences for some of the language languages). Results from the pre-dental fricative and pre-velars in this study fit within this range.
acquisition of Arabic-like overall vowel length that parallel with the level of the
group. The more advanced the group is the closer its vowel length becomes to
that of the average for the monolinguals. By contrast, voicing-dependent vowel
duration variation seems to fall at the other end of the hierarchy. It is hard to
learn and only some subjects tend to recognize it and acquire it. Advancement
in the language does not guarantee gradual acquisition of this feature.

Though this hierarchy lacks the basis to be claimed universal, work on the
acquisition of the same feature in other languages has produced similar results
(Flege 1979, Mack 1982, Elsendoorn 1984, and Laeufer (to appear)). All these
studies agree with this study in that L2 learners tend to acquire the overall
vowel length before they acquire, if they ever do, the differential vowel
duration. This classification of phonetic features according the degree of
difficulty may provide good basis for pedagogical consequences, some of
which will be discussed in the next section of this chapter.

Another interesting finding of this study was the greater magnitude of
voicing effect on short /æ/ than on its long cognate. For the most part, both
natives and non-natives alike showed smaller ratios (i.e., greater effect) for
short /æ/ than for long /æ/. By comparison, studies conducted on English
tended to show greater effect on the tense/long vowels than on the short ones
(Peterson and Lehiste 1960, House 1961, Klatt 1975, Crystal and House 1982,
just to name a few). The difference between the two languages indicates that:
there might be some language-specific features that contribute to the voicing
effect. As explained in the previous chapter, the quantity distinction found in
Arabic is responsible for this "reversed" effect. To reiterate briefly for
convenience, natives and non-native, especially non-natives whose behavior
would supposedly be reversed in English, did not exercise greater effect on long /æ/ in order to avoid a possible neutralization; greater decrease in the duration of long /æ/ may result in it approaching the length of its short cognate.

One of the consequences of this interpretation is the assumption that some features carry more weight (i.e., are more basic) than others cross-linguistically. In this particular case, quantity distinction seems to be more basic than differential duration and therefore it blocks or, at east, diminishes the voicing effect that results in the variation. Evidence for this view comes from the performance of the bilingual groups who would be expected to show greater effect, parallel to that of the tense vowels in English, for long /æ/ than for its short cognate. Instead, these groups show reversed behavior for most part.

POA has demonstrated a significant role in determining the voicing effect in Standard Arabic. Vowels in the pre-velar environment are longer than those in the pre-dental environment (F(1, 352) = 10.99, p<0.001). It should be emphasized here that POA does not seem to have a uniform role cross-linguistically. Indeed, most languages do not follow the same pattern of Arabic, leaving phoneticians in a state of confusion (see Lehiste 1970 and the literature cited therein concerning this point).

The contribution of the other variables included in the study fall into the expected patterns. Focused context has affected the overall vowel length but not the voicing effect. Analysis of variance has shown that vowels are on the average longer in the focused than in the unfocused context for all groups (F(1, 351) = 8.28, p<0.01). Interaction between language and focus did not reach significant level indicating that focused context has the similar effect in the two
languages. The lack of effect on the voicing effect should be understandable in light of the fact that the magnitude of the voicing effect is smaller in Arabic than that of English and may not allow for significant variation.

MOA has yielded the expected results as well. Vowels were longer before the fricatives than before stops for all the groups (F(1, 351) = 14.44, P<0.001). Interaction between the native language and MOA did not show significant difference, an indication that both languages behave in the same manner. Similarly, the voicing effect was greater before fricatives than before stops.

Finally, the results have shown that there is great inter-learner variability in the acquisition of some L2 features. Some learners seem to recognize and acquire the features of L2 early on in their exposure to the language. Some take longer time and some may not acquire these features at all. Therefore, in investigating the acquisition of L2 features, it might be a good idea to do subject analysis when the study is limited to a few subjects or increase the number of subjects to a point where the generalizations cross subjects become representative and meaningful.

In answer to the third question, results from the experiment on the reverse phonological interference have not shown a systematic effect on L1. At best, it could be argued that there was a trend for the effect but this trend has not reached significant level.

5.2 Pedagogical implications of the study

One of the consequences of conducting this study is to find out whether it has any pedagogical implications that may be of help to L2 learners and
instructors. In order to do so, the findings must be tied to the current trends in foreign language education.

Most pedagogues believe nowadays that the goal of teaching a foreign language should be communication (Wilkins 1976, Savignon 1982, Omaggio 1986, just to name a few). Therefore, a balance between linguistic accuracy (i.e., phonological, morphological, and syntactic accuracy) and the cultural and pragmatic conditions that determine the appropriate use of language should be achieved to guarantee effective communication. To achieve this goal, the techniques and methods of teaching a foreign language have been geared to focus on "contextualized" and meaningful activities. Repetition of individual segments that are devoid of a communicative context have been proven to yield unsatisfactory results (MacCarthy 1978, Savignon 1982, Omaggio 1986). In the meantime, the demand for the use of authentic instructional materials has increased. Instead of using mechanical, and occasionally meaningless materials, pedagogues have started to demand the use of actual, often unabridged or unsimplified, authentic materials.

Carrying these principles to teaching pronunciation and particularly the sounds incorporated in this study, a foreign language educator would expect his/her students to achieve a satisfactory level of pronunciation that will not impede or sever communication. In essence, such an expectation can be interpreted as "clear" pronunciation, not necessarily native-like, particularly in light of the well-known fact that adult learners of a foreign language will most likely not be accent-free. Foreign language theorists and pedagogues aim therefore at a modest level of accentedness that does not impede comprehension. Obviously, the closer that level is to the native norms the
clearer and more comprehensible speech will be. With that in mind, learners and teachers of a foreign language should aim at the clearest level of pronunciation and yet be ready to tolerate a level of accented speech.

It is not my intention in this study to provide specific techniques and drills to promote the teaching of pronunciation for these methods and techniques can be found in the many books available on this topic (Allen and Valette 1972, Paulston and Bruder 1976, Rivers and Temperley 1978, MacCarthy 1978) but to lay out some principles that should be taken into account in designing or selecting a method or a technique to teach a phonetic feature.

First, I do not advocate teaching pronunciation in a systematic way similar to some syllabi used in teaching syntax or morphology. Indeed, I do not advocate any method that directly teaches grammar. Instead, I would rather focus on the mastery of the linguistics elements that serve a particular communicative function leaving the grammar to be learned indirectly. To facilitate the build-up of a coherent body of grammatical knowledge, however, I advocate a symmetrical coordination between the linguistic units to be used in the communicative functions included in a syllabus and the grammatical notions that tie with these units perhaps in a linear or notional fashion (see Hussein 1993 for explanation). Thus, teaching grammar becomes a by-product of teaching the language for communicative goals. According to this view, teaching pronunciation becomes more of a remedial endeavor. Its aim is to correct and improve upon those phonetic aspects that seem to either impede understanding or cause miscommunication. Once a feature or a segment has

---

35Grammar has been used in this context in its broad sense. This is, it is not limited to syntax as it is traditionally known; rather it incorporates the phonological, morphological, and syntactic knowledge as well.
been identified as causing difficulty in communication then a drill that is focused and contextualized can be either designed or selected to remedy this difficulty.

The second principle for teachers of a foreign language to realize is that students cannot be drilled on single segments in isolation. The effect of adjacent segments on each other makes it obligatory that segments be presented in various contexts. As has been described in the study and many others conducted on voicing-dependent vowel duration, the length of a vowel varies depending on the voicing characteristic of the following obstruent. Thus, teaching a segment in isolation assumes a fixed static vowel length that does not represent the actual use of that segment. Another example, perhaps a better well-known one which shows the effect on adjacent sounds on each other in Arabic, is the pharyngealization phenomenon. Arabic shows that the existence of a pharyngealized segment in a word may cause the whole word or at least some syllables of it to acquire the pharyngealization feature as a result of the spread of pharyngealization. In this particular case, teaching the pharyngealized segments in isolation would certainly overlook their interaction with and influence over the adjacent segments. To teach vowels and pharyngealized segments properly then is to have them presented in at least a syllable, a word, or better yet in a phrase that represents their various phonetic manifestations. The use of phrases, especially if these phrases express language functions, will be more acceptable by the proponents of the communicative approach.

The third principle to be realized is the level of accuracy expected of the learners. Research in phonetics has shown that some cross-linguistic segments
are "similar" acoustically to each other and others are considered "new" (Flege 1987). Flege's research as well as that of others (Peng 1993, for example) has shown that L2 students tend to perceive the "similar" sounds as identical in the two languages, thus overlooking the phonetic details that actually distinguish the two. On the other hand, new sounds are more likely to be acquired more authentically since students have no native correspondents to equate them with. Such a distinction, if it were true, implies that teachers should pay special heed to the similar segments and perhaps provide more creative techniques in presenting them. On the other hand, it can be argued, especially by the proponents of the communicative approach, that dwelling on the phonetic details is not necessary for communication and therefore should not be an issue, particularly so when the use of the similar sounds by L2 students does not create difficulty in communication. Moreover, it has been shown that duration of exposure to L2 does not always guarantee the acquisition of L2 segments authentically. Some L2 learners spent more than 20 years studying and living in the target language environment and still have not acquired some phonetic aspects of L2. The study described in this dissertation, among many others, testifies to this outcome. In such situations it might be wiser to save the time and energy put into such futile efforts, especially when we realize that classroom time is hardly sufficient for more constructive and achievable goals. In summary, then, the degree of phonetic accuracy expected of the learners is a subjective matter that can be determined by the ease of communication and occasionally the level of the learner.

The fourth principle to be taken into account in teaching the pronunciation of L2, particularly as it has been manifested in this study, is the inter-subject
variability. Some students seem to acquire the phonetic details of L2 at an early stage of their exposure to that language. Others, on the other hand, do not acquire these details till later stages or may not acquire them at all. Such differences among the students should be recognized by the instructors and pronunciation drills should be adapted to fit the various abilities when there is a need for them.

Finally, the study has shown that some phonetic features are easier than others for students to recognize and acquire. As mentioned before, most students seem to have recognized and acquired the quantity distinction found in Arabic but a few have acquired the voicing-dependent vowel duration variation. A similar classification was given by Laeuger (to appear) when she investigated the acquisition of voicing-dependent vowel duration in English by two groups of French bilinguals. Laeuger's classification included seven phonetic parameters, namely, overall vowel length, duration of syllable initial consonant, VOT of syllable initial stop, voicing-dependent vowel duration, duration of final consonant, flapping of intervocalic alveolar stop, and release of voiceless final stops. According to this classification, overall vowel length was the easiest, followed by VOT in the second place, followed by consonant release and flapping in the third place, etc. If such ranking is to be taken into account in teaching English to adult French students, then designing pronunciation drills that cover all these features and presenting them according to their level of difficulty would be the logical step to take. Equally importantly, these rankings would help teachers determine the expectations from each drill and at each level. It remains to be acknowledged that these rankings may vary
from a language to another. In other words, what might be a difficult feature for French speakers may not be that difficult for native Arabic.

These principles are not meant to be more than guidelines that are based primarily on the findings of the study described here and the relevant literature. A larger set of principles can be established to cover other areas and features not included in this study. For example, principles that guide speech perception and discrimination versus production might also be necessary. In short, according to these principles, teachers start first with identifying the problematic feature, decide how serious it is (whether it causes miscommunication), place it on the scale of difficulty, decide on the proper pronunciation drill that fits the level, then lounge on the application.

5.3 Final remarks and recommendations

It goes without saying that most studies tend to raise more questions than provide answers when they are finished and this study is not an exception. Thus, several questions remained unanswered. Most prominent among these is the nature of the voicing effect. What is it in voicing or in its absence that causes the fluctuation in the duration of the preceding vowel? So far there has been no completely trouble-free answer as to how/why voicing causes the preceding vowels to be longer than those in the pre-voiceless environment. Another question as to why the voicing effect is more pronounced in certain environments than others. For instance, the effect is greater in word list reading than in connected speech. A third question has to do with the difficulty of acquiring the effect by L2 students. Again, what makes one phonetic feature easier or harder to acquire than another? A fourth question is about the
contribution of the phonetic details to the global perception and evaluation of someone's accent. Do all phonetic details contribute equally to someone's accent or are some phonetic aspects more outstanding than others? What is it that native speakers rely on when they evaluate someone's accent? These are among the most pressing questions that require further investigation. I am sure there are many others that can be raised. However, it should be emphasized that this list of questions is not meant to be exhaustive, rather it is intended to exemplify some of the issues that remain unresolved.

Another interesting issue that is worth highlighting is the line of research that has been followed during the past few decades. Except for a few studies, most phoneticians have focused on the presence (versus the absence) of the effect on the preceding vowels in the various languages. It is apparent that the effect is universal though with varying degrees in the various languages. It appears that it would be futile to continue the same line exploring the effect in more languages. It is about time that researchers turn their attention to the nature of the effect and attempt to find out how it operates, particularly how it causes the duration of the vowel to increase (versus how its absence causes the vowel to decrease).

The study described in this dissertation has been planned and executed following the mainstream of trying to find out whether voicing has similar effect in Arabic like most languages or not. Though considerable amount of data has been included, I still consider the results to be tentative and inconclusive. Further follow-up work is strongly recommended and encouraged to verify or refute some of the findings of this study. I recommend though that any follow-up work should expand on context, number of vowels
included in the study, increase intonation contexts, vary number of syllables incorporated in the target word (i.e., monosyllabic versus polysyllabic), increase the length of the utterances used as frame sentences, and perhaps vary the method or collecting data and presenting them spectrally.

With its modest contribution, I hope that the study will benefit those who are interested in phonetics and those interested in teaching pronunciation to L2 students. It is hoped also that it will stimulate further research in these areas.
Bibliography


Fischer-Jørgensen, E. 1964. "Sound Duration and Place of Articulation". 
Zeitschrift fur Sprachwissenschaft und Kommunikationsforshung 17, 175-207.

Lengths." Language and Speech. 3. 179-192.

Speech Research, SR-2. 3.1-3.12.

Elsendoorn, B. 1980. "Vowel Duration as a Cue to a Dutch Foreign Accent 
in English CVC-Words". Progress Report from the Institute of Phonetics, 
University of Utrecht 5. 64-68.

A. S. Dil (ed.) "Language Acquisition and Communicative Choice." 

Language Learning 30,1. 117-133.

English." Language and Speech 24. 125-146.


Fourakis, Marios & Gregory Iverson. 1984. "On the 'Incomplete' Neutralization of German Final Obstruents". Phonetica 41. 140-149.


Fry, B. 1958. "Experiments in the perception of stress". Language and Speech 1, 126-152.


