ABSTRACT

The primary objective of this study was to investigate the degree of influence of Arabic duration on the production of English when English is learned as a second language. The specific questions raised were:

i. What is the average Arabic syllable duration in "emphatic" and "non-emphatic" contexts. Is length variation a function of the adjacent consonant?

ii. Is duration of English syllables as produced by Jordanian students rule-governed i.e. having the same acoustic features as those of English?

INTRODUCTION

Although much is known about the duration of English vowel sounds (Cf. Elsendoorn, 1984), very little is known about the duration of Arabic vowel sounds or about English vowel sounds as produced by Arab learners of English. It is generally established that vowel duration is longer when the English vowels precede voiced consonants than when followed by voiceless consonants. For American English, it is known that a voiced consonant has a considerable lengthening effect on a preceding vowel (Cf. Peterson and Lehiste 1960, House 1961).

Vowel duration can also be influenced by factors other than post-vocalic consonants, e.g. the syllabic structure of the word, the position of the word in the sentence and many other important factors such as "emphasis" stress and intonation. Klatt (1973) demonstrated that the addition of a syllable to a word in which the final consonant is voiced, or the replacement of the final voiced consonant by its voiceless counterpart leads to an equivalent shortening of the vowel duration. The shortening effect becomes greater when both changes were combined. In contrast to English, no analytical data have been provided to show the quantitative differences between Arabic long and short vowels, or how the complex phonetic features of "emphasis" - "non-emphasis" interact as factors determining vowel length and syllable-quantities or the extent to which vowel length varies according to the presence or absence of preceding or following "emphatic" consonants. Unless a systematic description of Arabic vowel duration is provided, one cannot possibly determine whether an Arab learner possesses an internal criterion or implicit knowledge which governs vowel duration in their production of English vowel sounds.

The Arabic system of stressed vowels is generally agreed to consist of ten phonologically distinctive elements of relatively pure quality. These ten monophthongs are traditionally divided into two classes: one comprising five relatively long vowels [i:, a:, e:, o:, u:] and one consisting of five relatively short vowels: [i, a, e, o, u].

It is, however, recognised that the phonological contrast is never dependent on a difference in duration alone. Thus the difference in duration is always accompanied by a certain difference in quality. It is always equally important to take into consideration the phonetic features of "emphasis" whose minimum domain is the sequence CV-, in order to take a fuller account of Arabic vowel and syllable-duration.

PROCEDURE

Four Jordanian students (1st year university students: two females) were recorded on a tape-recorder in the studio of the language laboratory of the University of Jordan. Forty minimal pairs of contrastive Arabic and English forms were pronounced in real sentence frames:

Arabic carrier sentence: la(a) ..... wa lal(a)
English carrier sentence: It's not ... but ...
The word list was read twice by all speakers. Analysis was made from the second reading. All speakers were of the same age group and in the same class.

The collected samples were analysed on a wide band sound spectrogram (Voiceprint 700 spectrograph)

RESULTS

Inspection of the spectrograms obtained and the measurement of syllabic duration for the test words reveal durational difference between "non-emphatic" and "emphatic" syllables. Durational differences arise from the fact that "emphasis" has an overall shortening effect on the pattern is apparent in all Arabic syllable-types investigated: CVV, CVCC and CVV. Within each syllable-type, however, there is considerable variation in length according to the type of consonantal articulation being uttered and the mutual dependency obtaining between successive consonantal articulations within the syllable structure. Thus, syllables of the

* University of Jordan, Department of English, Amman, Jordan.
phonological structure CV.C in which the final C

1 position is realised as nasal or voiced pharyngeal

2 or approximant are longer in duration than when C

2 position is filled by a plosive consonant.

### Non-emphatic

<table>
<thead>
<tr>
<th></th>
<th>Emphatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>saam</td>
<td>427 ms</td>
</tr>
<tr>
<td>damm</td>
<td>315</td>
</tr>
<tr>
<td><em>baa</em></td>
<td>330</td>
</tr>
<tr>
<td>tobb</td>
<td>270</td>
</tr>
<tr>
<td>last</td>
<td>155</td>
</tr>
<tr>
<td>raad</td>
<td>287.5</td>
</tr>
</tbody>
</table>

### CV-CV

<table>
<thead>
<tr>
<th></th>
<th>Emphatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>faa-di</td>
<td>250-95 ms</td>
</tr>
<tr>
<td>saa-di</td>
<td>230-97.5</td>
</tr>
<tr>
<td>saa-ri</td>
<td>220-92.5</td>
</tr>
<tr>
<td>faa-ri</td>
<td>225-112.5</td>
</tr>
<tr>
<td>taa-ri</td>
<td>210-102.5</td>
</tr>
<tr>
<td>laa-ri</td>
<td>210-97.5</td>
</tr>
</tbody>
</table>

(The figures above represent average syllable
duration for diphthongical structure CV-CV in
which the first consonantal articulation is
whether "emphatic" or "non-emphatic" each figure
represents the average of four tokens).

### Non-emphatic

<table>
<thead>
<tr>
<th></th>
<th>Emphatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV-CV</td>
<td>CV-CV</td>
</tr>
<tr>
<td>9a-qa</td>
<td>105-95</td>
</tr>
<tr>
<td>9a-sa</td>
<td>100-95</td>
</tr>
<tr>
<td>9a-qa</td>
<td>140-80</td>
</tr>
<tr>
<td>1a-qa</td>
<td>150-250</td>
</tr>
</tbody>
</table>

As the figures indicate, features of consonantal
(e.g., nasal and pharyngeal approximant)
accompaniment in successive places have
considerable length; in both "emphatic" syllables. The
occurrence of a compatible plosive consonant at C
position has an inverse relationship as far as the
duration of the syllable is concerned. Differences
in duration between syllables closed by a voiced pharyngeal approximant or nasal are
obvious from the measurements obtained.

The exact duration of the Arabic syllable
types is not only determined by the type of
mutual dependency and co-occurrence restrictions
obtaining between various consonantal and vocalic
articulations but also by the total syllable
configuration. Thus, utterances in which the first
syllable is of structure CV- followed by a short
cvc syllable do not exhibit similar duration to
long syllables of the type CV-; long open
syllables (CV- are shorter in duration than
syllables of the type CV.C. Short syllables, on
the other hand, (CV- or Cvc) exhibit more or
less regular durational pattern ranging from a
minimum duration of 90 msec. to a maximum
duration of 135 msec.

On the basis of these observations, it
seems reasonable to conclude that syllable
duration is influenced appreciably by the follow-

i. in terms of quantity of the syllable, Arabic
vowels may be divided into two classes:
(a) Long Vowels (b) Short Vowels. Vowels in
general are shorter in duration when they
occur in "emphatic" contexts, i.e. preceded
or followed by an "emphatic" consonant.

ii. Arabic syllables may be divided in terms
of quantity into two types: Long Syllables
having as their nucleus Long Vowels (VW)
and Short Syllables whose nucleus is a short
vowel (v). "Non-emphatic" syllables require
longer duration to produce than their
corresponding short syllables. Within each
type of syllables, it is essential to recog-
nise the shortening effect caused by restric-
tions imposed on the occurrence of certain
consonants to the effective exclusion of
other consonants. Greater lengthening is
given, for example, to combinations allowing
nasals, pharyngeal approximants, trill and
laterals.

**DISCUSSION**

**ACOUSTIC AND ARTICULATORY FEATURES OF "EMPHATIC" AND "NON-EMPHATIC" CV-SEQUENCES**

The speech spectrograms obtained from
the experiment conducted provide a display of the
speech wave in terms of frequency—along the
vertical axis—and time—along the horizontal
axis—and intensity by the darkness of the markings. A
great deal of phonetic information from the
acoustic signal is a direct reflection of the
factors that contribute to the phonetic identi-
fies of speech sounds and their acoustic co-
relates. The constraints imposed by the phono-
logical rules on the allowable sequences help to
give better insights into how the speech signal
is processed. The proper identification of one
element in the sequence determines the other
element. Thus, for example, an Arabic "emphatic"
consonant is nearly always preceded or followed
by a more retracted "backer" vocalic articulation.
The denti-labial "emphatic" plosive T or D co-
cur with a preceding of following back open
evowel. The knowledge of such constraints on
allowable phoneme sequences enables the listener
to fill in phonetic details.

$\gamma$ has a high frequency peak. There is
a strong tendency for $\gamma$ to have greater disper-
sion that $\alpha$ and $\gamma$. The difference in centres
of gravity may be the only criterion found to
measure the difference between "emphatic" and
"non-emphatic" $\gamma$.

Measurement on spectrograms of locus of
formants of vowels following $\gamma$, s S, t T, d D,
l r, show great influence on formant transi-
tions after "emphatic" consonants. The spectral
characteristics of all F2 of vocalic elements
following "emphatics" indicate a lowering of the
formant. The locus is drastically lowered for
all the vowels. The fall of F2 in all vowels
following "emphatic" consonants is noticeable.
See Fig. 1 in which the $\gamma$ is followed by a
relatively long vowel with a reasonably high F1
and mid range F2. Note that the second formant
is considerably low due to the coarticulatory
effect of the "emphatic" 1 which tends to lower,
the second formant of the front vowel /a/.

It is noticeable that the FIRST FORMANT
frequency increases after "non-emphatic" phase
and the first formant decreases after the low
back phase. There is good correlation between
the coarticulatory effect of "emphatic" consonants
and the distance between the first two formants

---

EUROSPRach '89, Paris, France, September 1989

2337
which are far apart in front vowels and close together in back vowels.

Arabic long vowels are pronounced longer in the environment of a nasal, lateral or trill consonant, e.g. the long vowel i: is longer in [ti:in] than in [ti:n].

Formant frequencies of Arabic vowels are considerably affected by DOUBLE ARTICULATION (two places of articulation of equal importance) which is characteristic of all "emphatic" consonants. Double articulation has therefore an important effect not only on the duration of the vowel but also on its quality. Most formant frequencies are lowered when "emphatic" consonants are pronounced.

Vowel Duration in English words of CV(V) Structure as Produced by Jordanian Students

To see how far Arabic rules of vowel duration are exploited when English vowels are produced in comparable forms, forty English minimal pairs were selected and recorded on tape by four undergraduate students of the University of Jordan. Particular attention was paid so that the structure of the English contrastive forms chosen for study would conform, to what may be regarded, as equivalent of the Arabic word structure. Particular attention was also given to the measurement of English front vowels and English back vowels in more or less similar phonetic environment to those of Arabic front and back vowels.

The tables below are data tables listing average vowel duration of English CV(V) words for the four subjects recorded. Each figure represents average duration in milliseconds as obtained from the spectrograms:

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>di</td>
<td>165 ms</td>
<td>da</td>
<td>210 ms</td>
</tr>
<tr>
<td>ti</td>
<td>217.5 ms</td>
<td>ta</td>
<td>189.5 ms</td>
</tr>
<tr>
<td>ra</td>
<td>105 ms</td>
<td>la</td>
<td>157 ms</td>
</tr>
<tr>
<td>ba</td>
<td>225 ms</td>
<td>da</td>
<td>190 ms</td>
</tr>
<tr>
<td>si</td>
<td>195 ms</td>
<td>ba</td>
<td>192 ms</td>
</tr>
<tr>
<td>te</td>
<td>170 ms</td>
<td>so</td>
<td>197.5 ms</td>
</tr>
<tr>
<td>si</td>
<td>197.5 ms</td>
<td>li</td>
<td>187.5 ms</td>
</tr>
<tr>
<td>ri</td>
<td>175 ms</td>
<td>la</td>
<td>212.5 ms</td>
</tr>
<tr>
<td>ha</td>
<td>185 ms</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be judged from the spectrograms, spectrographic evidence of the acoustic realisations of English consonants t, d, s, ð, l, r as "emphatic" consonants is quite obvious. The burst of energy following the stop gap is much stronger in amplitude and lower in frequency than a non-emphatic plosive.

When learning English as a second language negative transfer operates on corresponding English features so that English t, d, s, ð, l, r when followed by open or half-open back vowels may be realised as "emphatic" Arabic consonants with the typical canonic acoustic characteristics of Arabic "emphatic" consonants T, D, S, ð, L, R. Thus, depending on the immediate phonetic environment involving back vocalic elements, the English voiceless obstruents in the utterances bus-stop, sauce, roast are realised as /bæs- stɒb/ , /sɒ:st/ and /roʊst/. The acoustic realisations are seen to be entirely different from those produced by native speakers and may often lead to the distortion of some important acoustic cues.

In Jordanian English, the acoustic realisation of phonetic segments is highly context-sensitive and is largely influenced by Arabic context-sensitive rules. Thus the realisation of English /t/ as /T/ , /ð/ as /d/, /s/ as /s/ , /ʃ/ as /ʃ/, l as f and r as r is fairly systematic and can be captured by explicit rules:

i. /t/ becomes /T/ voiceless alveolar fricative becomes voiceless dental/uvular plosive when a back open or half-open vowel follows the segment, e.g. /tɒn/ is realised as /Tɒn/, /tʃɪm/ as /Tʃɪm/, and /bɑ:/ as /bɒ:/.

ii. /s/ becomes /ð/ voiceless alveolar fricative becomes "emphatic" dental/uvular fricative or in acoustic terms the centre of gravity becomes in the lower frequency ranges than in the higher frequency ranges, when the segment occurs in the environment of a back open or half-open vowel, e.g. sauce, roast, bus, fuss, sub, socks, loss, boss, Sussex, sausages, fast, are realised as /sɒʃ/, /ʃɒʃ/, bə s, fə s, ʃə b, sʊʃ, bʊʃ, suːs, fæ st, /ʃə bət/.

iii. /ʃ/ becomes /ʃ/ when it occurs in the environment of a back open or half-open vowel, e.g. mother, father, thus, are realised as /moʊðər, fæ ˈfɑːðər/.

The conclusion to be drawn from the measurements described in the previous sections is that the constraints imposed by the phonological system of Arabic will inevitably show itself in the Arab students' performance in English. This is particularly noticeable when English voiced and voiceless plosives are realised as Arabic dentis/lateralolar plosives: non-emphatic in the environment of non-emphatic sounds and are realised as emphatic in the environment of back open vowels. The durational characteristics of Arabic vowels sounds in the immediate phonetic environment involving "emphasis"/ "non-emphasis" is expected to be carried over to English sequences of comparable utterances. Furthermore as the figures in the previous sections indicate, Arabic vowels vary in duration not only in relation to "emphasis" within the domain of the syllable but they all vary in duration according to co-occurrence restrictions imposed by the syllabic structure of the word. Thus, syllables closed by voiced pharyngeal fricatives and nasals are longer than syllables closed by plosive consonants. Also, vowels occurring within the domain of CVC- syllable types are shorter than vowels occurring the domain of CVC syllable-types. (Cf. ta /b/ av. duration 215 ms vs. tan-li av. duration 187.5 ms.).
REFERENCES


Fry, D.B. "Duration and Intensity as Physical Correlates of Linguistic Stress". JASA, 27:765-766.


