BEST COPY

AVAILABLE

Variable print quality
THE ACOUSTIC ANALYSIS OF ARABIC SPEECH

by

K.S.K. SHAHEEN

Thesis submitted to the
UNIVERSITY OF WALES
in support of the application for the degree of
PHILOSOPHIAE DOCTOR

Department of Linguistics
University College of North Wales
Bangor
1979
2.3.4. Duration 46
2.3.5. Transitions 48
2.3.5.1. Direction of transition 60
2.3.5.2. Extent of transition 64
2.4. Conclusion 70
2.4.1. General discussion 70
2.4.2. Coarticulation and word-position effects 72
2.4.2.1. The effect of word position on fricative spectra 72
2.4.2.2. The effect of the adjacent vowel on fricative spectra 73
2.4.2.3. The effect of the adjacent fricative and word position on the vowel spectrum 74
2.4.2.4. Factors affecting fricative temporal properties 75

Chapter 3: The stops 77
3.1. Articulatory mechanism of stop production 79
3.2. The shutting phase: The role of transitions 80
3.3. The opening phase: The role of bursts 82
3.4. The voiced/voiceless distinction 86
3.4.1. The presence vs. the absence of voicing during the closure phase 87
3.4.2. Relative duration between burst and onset of vowel resonance 88
3.4.3. Relative duration of vowel resonance before the closure phase 90
TABLE OF CONTENTS

DECLARATION i
ACKNOWLEDGEMENTS ii
SUMMARY iii

Chapter 1: Introduction 1
1.1. The perception of speech 2
1.2. The analysis of speech 5

Chapter 2: The fricatives 12
2.1. The articulatory mechanism 14
2.2. The perception of the fricatives 15
2.2.1. The search for the acoustic cues 15
2.2.2. The cues in the friction part 16
2.2.3. The cues in the transition part 20
2.2.4. Intensity of fricative consonants 25
2.2.4.1. The role of intensity in fricative identification 28
2.3. The acoustic analysis of the Arabic fricatives 29
2.3.1. Frequency positions of the energy-density maxima 29
2.3.2. The effect of position on the spectra of the Arabic fricatives 35
2.3.3. The effect of the adjacent vowel on fricative spectra 42
3.4.4. The relative intensity of the stop burst 90
  A. Tenseness and laxness 90
  B. Duration and intensity of stop explosion 95
3.4.5. Relative duration of the closure phase 101
3.4.6. Extent of F1-transition 102
3.5. The acoustic analysis of the Arabic stops 106
  3.5.1. The effect of position on the spectra of the Arabic stops 106
    3.5.1.1. The voiced stops 107
    3.5.1.2. The voiceless stops 107
    3.5.1.3. The glottal stop 108
    3.5.1.4. The velar/palatal pair /k,g/ 109
3.5.2. The effect of the adjacent vowel on the spectra of the Arabic stops 111
3.5.3. Duration 111
3.5.4. Transitions 117
  3.5.4.1. Direction of transition 127
  3.5.4.2. Extent of transition 130
3.6. Conclusion 134
  3.6.1. Coarticulation and word position 134
    3.6.1.1. The effect of word position on stop spectra 135
    3.6.1.2. The effect of the adjacent vowel on stop spectra 136
    3.6.1.3. The effect of the adjacent stop and word position on the vowel spectra 136
3.6.1.4. Factors affecting stop temporal properties 137

Chapter 4: The trill 139
4.1. Articulatory mechanism 140
4.2. The acoustic theory of trill production 141
4.3. Acoustic analysis 142
4.4. Perception 144
4.5. Acoustic analysis of the Arabic r-sound /ɾ/ 145
4.5.1. The non-pharyngealized [ɾ]. 147
4.5.1.1. The effect of position on the spectrum of [ɾ]. 147
4.5.1.2. The effect of the adjacent vowel on [ɾ]. 148
4.5.1.3. Duration 148
4.5.1.4. Transitions 150
4.5.2. The pharyngealized [Ӆ]. 150
4.5.2.1. The effect of position on the spectrum of [Ӆ]. 150
4.5.2.2. The effect of the adjacent vowel on [Ӆ]. 153
4.5.2.3. Duration 153
4.5.2.4. Transitions 155
4.6. Conclusion 155

Chapter 5: The lateral 161
5.1. The articulatory mechanism 163
5.2. Acoustic theory of lateral production 164
5.2.1. A consonant or a vowel 164
5.2.2. Cavity affiliations 165
5.2.3. Formant continuity 167
5.3. Perception 169
5.4. The acoustic analysis of the Arabic lateral 171
5.4.1. The non-pharyngealized lateral 171
5.4.1.1. The effect of position on the spectrum of [l]. 172
5.4.1.2. The effect of the adjacent vowel on [l] 172
5.4.1.3. Transitions 172
5.4.1.4. Duration 174
5.4.2. The pharyngealized lateral [l]. 176
5.4.2.1. The effect of position on the spectrum of [l]. 176
5.4.2.2. Duration 177
5.4.2.3. Transitions 177
5.5. Conclusion 177
5.5.1. Spectral differences between the pharyngealized and the non-pharyngealized lateral 179
5.5.2. The effect of word position on the spectrum of the lateral 179
5.5.3. The effect of the adjacent vowel on lateral spectra 180
5.5.4. The effect of the lateral and word position on vowel spectra 180
Chapter 6: The nasals

6.1. Articulatory mechanism

6.2. Acoustic mechanism of production of nasals
   6.2.1. The contribution of the oral side chamber
   6.2.2. Formant continuity

6.3. Perception

6.4. The acoustic analysis of the Arabic nasals
   6.4.1. The non-pharyngealized bilabial nasal
       6.4.1.1. The effect of position on the spectrum of [m].
       6.4.1.2. The effect of the adjacent vowel on [m]
       6.4.1.3. Duration
       6.4.1.4. Transitions
   6.4.2. The pharyngealized bilabial nasal [m]
       6.4.2.1. The effect of position on the spectrum of [m]
   6.4.2.2. Duration
       6.4.2.3. Transitions
   6.4.3. The non-pharyngealized dental nasal
       6.4.3.1. The effect of position on the spectrum of the nasal [n].
6.4.3.3. Duration
6.4.3.4. Transitions
6.4.4. The pharyngealized post-dental nasal n
6.4.4.1. The effect of position on the spectrum of [n].
6.4.4.2. Duration
6.4.4.3. Transitions
6.5. Conclusion

Chapter 7: The semivowels
7.1. The articulatory mechanism
7.2. Acoustic analysis
7.3. Perception
7.4. The acoustic analysis of the Arabic semivowels
7.4.1. The effect of position on the spectrum of the semivowel
7.4.2. The effect of the adjacent vowel on the spectra of the semivowel
7.4.3. Duration
7.4.4. Transitions
7.5. Conclusion

Chapter 8: Conclusion
8.1. General discussion
8.2. General spectral and temporal properties of Arabic sounds
8.2.1. The effect of word position on consonant spectra 248
8.2.2. The effect of the adjacent vowel on the consonant spectra 250
8.2.3. The effect of the consonant on the adjacent vowel spectra 252
8.2.4. Factors affecting duration 253
8.3. Suggestions for further work 254

BIBLIOGRAPHY 256
I am most grateful to my supervisor, Mr. R.A.W. Bladon, for his advice, encouragement and criticisms at every stage in the preparation of this study.
This study is an investigation of the spectral and temporal characteristics of the sounds of the Egyptian dialect of Arabic, mainly as revealed by acoustic spectrography.

The consonant sounds of Arabic speech are divided into groups according to their manner of articulation and are studied in separate chapters accordingly.

Each chapter is divided into two main parts. In the first part we look into the articulatory mechanism for the production of the consonant under study, as well as at its acoustic theory of production. We also report the results of the acoustic analyses of sounds produced with the same mechanism in other languages, as well as the results of synthetic speech experiments.

In the second part of each chapter we report the results of the acoustic analysis of the Arabic consonant under investigation. In this part we study the temporal and spectral characteristics of the particular consonant in different word positions. We also study its effect on the direction and extent of the second-formant transitions of the adjacent vowels, as well as the effect of the adjacent vowel on the frequency positions of the components of the consonant spectrum.

This part is then followed by a concluding section in which we provide an overall view of the specifics of the acoustics of that consonant and discuss the relevance of our findings on its acoustics to the general theories of speech production and perception.
Chapter 1

Introduction

1.1. The perception of speech.

1.2. The analysis of speech.
Introduction

1.1. The perception of speech.

Phonetics could be defined as the science that examines the production, transmission, and perception of speech sounds and the relations between them.

It is true that until the beginning of the second half of this century phonetic research largely concentrated on the articulatory aspect of speech, while research into the physical nature of speech sounds and the relation between speech wave characteristics and our perception of them lagged behind. This, however, was the result of the lack of adequate experimental means.

With the rapid development of electronic devices of measurement phoneticians found themselves able for the first time to undertake with precision the study of the physical nature of speech sounds. However, the mass of information poured out of these devices did not in itself provide solutions as to the problem of the relationship between speech wave characteristics and our perception of speech sounds. That is to say, the fact that speech sounds are perfectly amenable to acoustic measurement with respect to the dimensions of frequency, intensity, and duration, does not mean that simply by making acoustic measurement we can discover how acoustic processing is carried out in speech communication.
There are two major reasons for this. The first is that any cue in speech will depend upon relative and not on absolute values. As Fant (1968) points out "the particular F-pattern (F1, F2, F3, et al.) of phonemically 'one and the same' vowel varies considerably with speaker type, sex, and age." Thus if we take a sample of the 'same' vowel uttered by a man, a woman, and a child, and make acoustic measurements on each of these, we shall not obtain absolute measurements that are identical or even particularly close to each other. There will be marked differences in overall intensity, spectral distribution of energy, fundamental frequency and duration. The fact that any listener will recognize the 'same' vowel in all three cases is due to his reliance upon acoustic cues based on relative values, relations within each utterance, and most important of all, relations between this particular utterance and others which might come from the same speaker.

The second reason is that there is a plurality of acoustic cues in any one given phonemic unit. Thus spectral analysis will reveal that in the case of English /p/ there will always be a brief interruption in the flow of acoustic energy from the mouth which corresponds to the phase of complete closure. Second, a short silence whose duration is affected by its immediate phonetic context. Third, a brief burst of noise somewhere in the region of 1000-2000 Hz corresponding to the explosion of air. And, fourth, if a vowel immediately follows or precedes the /p/ the spectral distribution of energy during the vowel will show certain features, including a rather rapid change in the
frequency of the second formant at the point nearest to the /p/, and
the moment at which periodicity begins will bear a certain time
relation to the burst of noise.

These observations do not justify us in saying that we have
here the acoustic cues which a listener employs when he perceives
an incoming sound as /p/. At the very most we have an indication
of possible cues and we are left with the task of demonstrating
that a listener employs any or all of them in given conditions of
communication.

The need for such demonstration has given rise to the develop-
ment of a number of experimental methods. One which has been widely
used for many years is that of subtracting information from the
acoustic signal by filtering, or by adding masking noise and measuring
the effect of this on the perception of various speech sounds.

Subtracting information from the speech signal tells us more
about the process of speech perception than straightforward analysis,
but it is still not enough to isolate the acoustic cues because what
is left in the signal is still a complex which is likely to contain
more than one cue of a given recognition. That is, none of the fil-
tering or masking conditions could serve to isolate a cue dimension.
In fact, the only way in which this can seemingly be effectively done is by the use of speech synthesis, that is, by creating artificially speech-like sounds in which only one cue for a given recognition is represented and is systematically varied.

1.2. The analysis of speech.

In order to attempt a synthesis of speech sounds we must have first carried out an analysis of them. That is, analysis provides us with information on the various components of speech sounds. It is out of this mass of components that we would then extract the elements thought to be essential, the parameters, and carry out the synthesis.

One can distinguish two approaches to the analysis of speech sounds. The first, which may be called 'segmental' is a 'phoneme-dominated' approach which looks at speech as consisting of a sequence of segments, each one of which representing a phoneme, which are selected from a finite inventory. The second, which may be called 'parametric', rejects the segmental approach to the analysis of speech on the grounds that it does not reflect the fact that speech is a composite of variables which are always present and always changing in value.
A speech wave is a continuum with no breaks or natural points for segmentation. Phoneticians' attempts to segment this continuum either started from within by attempting to segment the acoustic matter by reference only to the observable features in the continuum and with no reference to anything outside it, or else, they started from without by seeking to find in the acoustic matter the correlates of the articulatory features responsible for the production of speech sounds or the acoustic features used for differentiating between phonemes.

The attempt to segment the acoustic material from recurrent internal features faces the difficulty that the location of these transitions and segments in the sound stretch is not always obvious or indisputable. A segmentation according to these principles is not, consequently, unambiguous. A segmentation of a piece of connected speech may provide a larger number than phonemes of the message and it may also show a smaller number of segments than phonemes.

A stop sound, for instance, may have an occlusion which is partly voiced and the explosion phase of the release may be followed by a short fricative segment and an aspiration segment. The presence of the particular stop sound generally affects the source of the F-pattern in the segments preceding and following the occlusion. Thus a single phonetic segment of the speech wave may be influenced by
several successive phonemes of the message and conversely each single phoneme is generally signalled by cues spread out over several successive segments.

There are two different methods to extract from the acoustic medium those features that are relevant to the differentiation between phonemes. Both of them, however, are based on the assumption that the phonetic form of phonemically opposed speech sounds is different. The first method seeks to identify the acoustic correlates of the relevant features on the articulatory level. The second method, by contrast, attempts to arrive directly at the acoustic features used for establishing phonemic contrasts.

An investigation into the acoustic correlates of the relevant articulatory features of the phonemes of a language is presented by Fant (1968). Fant writes that the purpose of his study is "to extend the old concepts of manner and place of articulation to a categorization of speech production events and to summarize rules for translating from production category to speech wave characteristics."

According to this theory, connected speech may be divided into a succession of segments displaying a temporal contrast of 'manner' features. To these features are added 'place' features which vary continuously within segment boundaries. This method, then, assumes
that the effects of the relevant articulatory features are encoded in the total sound wave output and that an investigator can decipher the code. Such an assumption, however, faces the difficulty of accounting for the fact that there is not a one-to-one correspondence between articulatory events and their acoustic result.

Fant, in fact, does refer to this problem. He writes that in an experiment by Liberman $g$ appeared to have not one locus but two: one at about 3000 Hz when it occurred before front vowels and a second at around 1200 Hz before back vowels. He went on to say that "The conclusion of Liberman was that there existed a true discontinuity on the acoustic level not paralleled on the perceptual and the articulatory level." He points out, however, that "In the light of present knowledge, however, the outcome of Liberman's $g$ experiment is just what could be expected but the view of a discontinuity does no longer hold."

Another problem that faces this method is that of attributing individual component parts of an acoustic spectrum to specific phonemes. A single phonetic segment of the speech wave may be influenced by several successive phonemes of the message and, conversely, each single phoneme is generally signalled by cues spread out over several successive segments. A stop sound, for instance, may have an occlusion which is partly voiced and the explosion phase of the release may be followed by a short fricative segment and an aspiration segment.
Fant, however, points out that the absence of one-to-one simple relations between phonemes and speech wave units is the result of the absence of a one-to-one correspondence between phonemes and articulatory events. This refers to the fact that the phonetic realization of a sound segment is dependent on the phonetic context of preceding and following sounds. In an experiment by Öhman (1966), for instance, it was shown that the F-pattern of an intervocalic stop varies according to the phonetic nature of the preceding and following vowels.

Fant, however, does not believe that the apparent phonetic variability on the acoustic and articulatory levels invalidates the concept of invariable phonemic commands on the neural level of speech production. In fact, he believes that the factors responsible for this phonetic variability (such as coarticulation and vowel reduction) "eventually can be fully described and contained in models and formulas that enable a prediction of the physical speech event given the linguistic message transcript." (1968).

The last method of speech analysis that we shall deal with is that which attempts to segment speech sounds with reference to describing features of their acoustic spectrum through determining these acoustic features which are relevant in establishing phonemic oppositions. To the proponents of this method, a detailed account of
which has been proposed by Jakobson, Fant and Halle (1952), a speech sound is a bundle of simultaneous acoustic characteristics called 'distinctive features' which arise from the nature of the input source or sources used or from the effects of the resonatory transmission of power from these sources.

The method, obviously, avoids the drawbacks of any attempt to correlate relevant articulatory features with components of the acoustic spectrum. It implies, however, that the function of articulation is to produce sounds with different acoustic structures in phonemic oppositions, and that the listener's brain may identify different speech sounds from their different acoustic structures without any necessary attempt to derive articulatory correlates from the whole or part of those structures.

As we will have observed, analysis can be done in a variety of ways and on different grounds. The choice of one method rather than another depends ultimately upon the purpose which the investigator has in view. It can be seen, however, that an approach that starts out by a search for basic and independent parameters of speech production to determine their acoustic correlates as well as their possible spectral significance and distinctive function on the message level, or likewise an approach that attempts to determine what
acoustic features are relevant in distinguishing phonemes, is one that aims at descriptive completeness in the first place and at economy of description only in the second place. On the other hand, an approach that attempts to describe the different possible phonetic realizations of the phonemes of a language is one that is more economical, and, I would say, more suited to the elementary stages of investigation.
Chapter 2

The fricatives

2.1. The articulatory mechanism.

2.2. Perception of the fricatives.

2.2.1. The search for the acoustic cues.

2.2.2. The cues in the friction part.

2.2.3. The cues in the transition part.

2.2.4. Intensity of fricative consonants.

2.2.4.1. The role of intensity in fricative identification.

2.3. The acoustic analysis of the Arabic fricatives.

2.3.1. Frequency positions of the energy-density maxima.

2.3.2. The effect of position on the spectra of the Arabic fricatives.

2.3.3. The effect of the adjacent vowel on fricative spectra.

2.3.4. Duration.

2.3.5. Transitions.

2.3.5.1. Direction of transition.

2.3.5.2. Extent of transition.

2.4. Conclusion.

2.4.1. General discussion.
2.4.2. Coarticulation and word-position effects.
2.4.2.1. The effect of word position on fricative spectra.
2.4.2.2. The effect of the adjacent vowel on fricative spectra.
2.4.2.3. The effect of the adjacent fricative and word position on the vowel spectrum.
2.4.2.4. Factors affecting fricative temporal properties.
2.1. The articulatory mechanism.

Fricative sounds are produced by turbulent air flow caused by a constriction in the vocal tract at some point in or above the larynx. The constriction may vary as to position in the tract, sectional area of constriction, length of constriction, and shape of orifice. Further, the air stream may vary as to pressure or rate of flow.

A fricative sound is thus a function of an amalgam of variables. A variation of one or a combination of these variables can be expected to cause a variation in the physical nature of the resulting sound in any of three ways: by altering the spectrum of the noise source; by altering the filter function of the vocal tract as a whole; or by altering the intensity of the acoustic energy produced.

The spectrum of the sound produced varies jointly with the sectional area of constriction, length of constriction and shape of orifice. The overall acoustic intensity is proportional to rate of air flow for fricatives produced with the same vocal tract configuration.
2.2. Perception of the fricatives.

2.2.1. The search for the acoustic cues.

The fricative consonants in CV syllables appear on spectrograms to be made up of two successive segments - a period of noise (the friction), followed by a segment with well-marked formant structure (the transition). Cues for identifying the fricative phonemes might well reside in either or both of these two portions.

In an attempt to evaluate the relative importance of cues in the friction and transition portions, K.S. Harris (1954) split the friction segment away from the transition segment and then recombined friction and transition portions from different syllables, on the assumption that the more important part of the sound would determine which phoneme a listener would hear.

Harris made recordings of the English fricatives /f, ð, s/ and /ʃ/ as well as their voiced counterparts /v, ð, z/ and /ʒ/ in CV syllables which were recombined in the manner described above and then presented them to listeners for judgement. The results of the test demonstrated that while the friction part provided the necessary and sufficient cues for the identification of the alveolar and post-alveolar fricatives, the cues necessary for the identification of the dental and labio-dental fricatives resided in both the friction and transition parts.
2.2.2. The cues in the friction part.

One of the early studies to investigate the nature of the acoustic cues that reside in the friction part of a fricative was that reported by Hughes and Halle (1956). Hughes and Halle studied the spectra of the English labio-dental, alveolar and post-alveolar fricatives in isolation. They did not, however, extend their investigation to the dental fricatives in the belief that "a solution to this problem will come only after the mechanism involved in their production is more fully understood."

Hughes and Halle's analysis of the spectra of these fricative consonants revealed two points. First that there are consistent differences among the three classes of fricatives (labio-dental, alveolar, and post-alveolar). And, second, that these differences are maintained only within the speech of a single speaker. In the speech of several speakers, however, there exists such a degree of overlapping between the fricatives to the extent that the /s/ of one speaker can be recognized as the /ʃ/ of another.

Hughes and Halle's analysis of fricative spectra regarded differences between fricative categories as a function of the frequency position of the most prominent energy-density maxima. Furthermore, by viewing these maxima as resonances of the most effective portion of the vocal tract (that is, the portion in front of the source), they could expect an inverse relationship to hold between the length of the effective portion of the vocal tract and the frequency position of the energy-density maxima.
The fricatives do, indeed, pattern in such a way that the most prominent peaks in the spectra of the labio-dental fricatives are higher than those of the alveolar fricatives, which are, in turn, higher than those for the post-alveolar fricatives. Such an approximation, however, does not account, as Hughes and Halle point out, for the low-frequency peaks in the spectra of the labio-dental fricatives.

In order to check whether the isolated steady-state portions of different fricatives could be identified correctly by listeners, Hughes and Halle recorded the gated steady-state portions of the fricatives /f, s, and s/ in random order and presented them to a group of listeners who were asked which of the three phones had been spoken. Analysis of listeners' responses showed that the percentage of correct responses reached 71.

However, the results of these perceptual tests, as far as Harris is concerned, are questionable. She argues (1958) that "The result is not surprising for /s/ and /s/, since we had concluded that friction provided the necessary and sufficient cues for their identification. Furthermore, one would expect that /f/ friction would be discriminable in the set of alternatives presented, since in our experiment, /f/ was not confused with any friction except /θ/, and /θ/ was not a possible response in Hughes' and Halle's experiment."

The fricatives were again the subject of investigation in a later study by Peter Strevens (1960). This was also a study that restricted its search for the acoustic cues to the friction
part of fricatives in isolation. The analysis of the data in Streves' experiment supported the observation made earlier by Hughes and Halle that differences between the various fricative categories tend to be more systematic within the speech of a single speaker than within that of several speakers.

Strevens' strategy in separating fricative categories is, however, different from that of Hughes and Halle in one important respect. While phonemic distinctions of fricative consonants are regarded in Hughes and Halle's analysis as simply the function of the frequency location of the energy-density maxima, in Streves' analysis the separation of fricative categories is effected primarily on the basis of what he calls "upper and lower limits of energy" while relegating the role of the frequency position of energy peaks in distinguishing fricative classes to a less effectual position.

It is true that Streves (1960) writes that the sounds investigated "are shown to be capable of description in terms of the frequencies of the lower and upper limits of energy present, the presence or absence of formant-like concentrations of energy, and the overall relative intensity of the sounds." Yet the terms in which he distinguishes between the three major groups (\{\textit{p}, \textit{θ}, \textit{f}\}), \{\textit{s}, \textit{j}, \textit{zh}\}, and \{\textit{v}, \textit{x}, \textit{h}\}) into which he divides the fricatives of his experiment as well as between members of the same group are (with the exception of the group \{\textit{φ}, \textit{f}, \textit{θ}\}) primarily based on the notion of upper and lower limits of energy alone.
Strevens's strategy in describing the fricative consonants in terms of the upper and lower limits of energy raises an important question which relates to the general function of acoustic analysis.

One could argue that the justification for the choice of one method of analysis, rather than another, does not depend on the degree of success of that method in separating the different categories of speech sounds but on its success in increasing our understanding of how these speech sounds are produced.

The analysis of fricative consonants in terms of the frequency positions of their energy-density maxima has the backing of a highly developed acoustic theory of speech production. It is based on the assumption that "Acoustically, the common denominator of all sounds produced from a resonator system of a prescribed configuration is the particular set of formant frequencies of the vocal tract, i.e., the F-pattern." (Fant, 1960).

More specifically, it agrees with Fant's suggestion (1960) that since "the typical fricative is a noise sound, the spectral energy of which is largely contained in formants from cavities in front of the articulatory narrowing," the essentials of fricative production may be specified in terms of source and filter characteristics just as for vowels."
The specification of fricative quality in terms of the frequency positions of their energy-density maxima does not only develop our understanding of the mechanism of fricative production. By virtue of its being part of a general theory of speech production it also helps develop our understanding of speech production in general.

Stevens's strategy, on the other hand, might indeed have been successful in separating fricative categories. It fails, however, to illuminate our understanding of how they are produced.

2.2.3. The cues in the transition part.

One of the first studies that set out to investigate the role of the transition part in fricative identification was that reported by Heinz and Stevens (1961). Heinz and Stevens studied the mechanism of fricative sounds production with the aid of an idealized mechanical model of the human vocal organ. On the basis of the model studies, the vocal tract is approximated by a simple electric circuit whose transfer functions are characterized by poles and zeros. By proper selection of the centre frequencies and bandwidths of these poles and zeros, Heinz and Stevens were able to generate spectra of fricatives that matched reasonably well with the spectra of the fricatives of natural speech.

Heinz and Stevens started by generating the spectra of the fricative sounds \( [s, q, s, f] \) and \( [\theta] \) in isolation. The resonant
frequency of the circuit was varied through several values from 2000 Hz to 8000 Hz. For each resonant frequency, the bandwidth was given three different values that encompassed the range of bandwidths observed in the spectra of the fricatives of natural speech. The stimuli were tape recorded in random order and presented to a group of listeners who were asked to identify each by making one of the responses [s, s, q, f] and [θ].

The results of the test revealed three points. First, that the responses were relatively independent of bandwidth over a 2:1 range of bandwidth. Second, that a consistent shift of responses from [f] to [s] to [f, θ] is obtained as the resonant frequency is increased. Third, that distinctions between [θ] and [f] could NOT be made for these isolated stimuli.

The inability of listeners to distinguish between [f] and [θ] on the basis of differences in energy distribution among the frequencies suggested that the identification of a fricative consonant might be dependent not only on this factor but also on (1) the formant transitions of adjacent vowels and (2) on the intensity of the fricative relative to that of the vowel.

Since the role of these cues cannot be studied in isolated fricatives, it was necessary to generate stimuli consisting of
syllables. These were generated with varied overall level of friction intensity and varied F2-transition. They were then tape recorded in random order and presented to a group of listeners who were asked to identify the initial consonant in each syllable as one of \( [f, \theta, s, o] \).

The results of the test revealed that once again there is a consistent shift in the pattern of responses from \([s]\) to \([s]\) to \([f, \theta]\) as a function of the increase in resonant frequency. More importantly, the distinction between \([f]\) and \([\theta]\) could this time be effected by the use of different F2-transitions. Thus when an F2-transition starting at a frequency of 900 Hz is used the sound is perceived as \([f]\), while if an F2-transition starting at a frequency of 1700 Hz or 2400 Hz is used the sound is perceived as \([\theta]\). The high locus associated with \([\theta]\) could be accounted for, as Heinz and Stevens point out, by the fact that "in the production of this sound the tongue moves from a position in the vocal tract at a higher point vertically than that used for \([f]\) when the tongue normally lies flat in the mouth."

The observed shift of the pattern of responses from \([s]\) to \([s]\) to \([f, \theta]\) as a function of the increase in resonant frequency agrees with the analysis of Hughes and Halle who found that for any single speaker the spectra of \(/s/\) have peaks at consistently lower frequencies than those of \(/s/\), which in turn have peaks at lower frequencies than those of \(/f/\).
The results of Heinz and Stevens' investigation of the effect of F2-transition of adjacent vowels on fricative identification are also in good agreement with the results of fricative identification tests reported by Harris. Both studies lead to the conclusion that the important cues for the alveolar and post-alveolar fricatives are given by the noise, but that differentiation between the dental and labio-dental fricatives is accomplished primarily on the basis of cues contained in the transition part in the syllable.

There are good reasons, however, to doubt the validity of such a conclusion. A small experiment was run to test listeners' ability to discriminate between spoken isolated versions of the Arabic fricatives, including /θ/, in the absence of any acoustic cues other than those in the steady-state friction part, and without the aid of visual cues either.

In spite of the fact that /θ/ is of very limited occurrence in Cairene Arabic—being found mainly in the speech of the educated class and then only in classicisms—and is almost always replaced by /s/, the results of the test showed, first, that all the Arabic fricatives are discriminable with the aid of the cues in the friction part alone, and, second, that there are no differences in the rates of correct responses for the different fricatives.
The results of our perceptual tests do not, thus, support Harris' argument that /f/ was discriminable in Hughes and Halle's experiment only because /f/ was not a possible response in the set of alternatives presented to listeners. In fact, Harris' experiment only shows that in the presence of contradicting cues listeners rely on cues resident in the friction part for the identification of /s/ and /θ/, and on cues resident in the friction and transition parts for the identification of /f/ and /θ/.

The fact that the listeners in our experiment managed successfully to identify the spoken versions of Arabic /f/ and /θ/ on the basis of cues contained in the friction part alone, while those in Heinz and Stevens' experiment relied primarily on cues in the transition part of the synthesized versions of these two fricatives does not seem to be the product of differences in the quality of English and Arabic fricatives. More likely, it suggests that in synthesizing these fricatives the critical difference between the two sounds was damaged.

The results of our perceptual tests thus support Lobanov's observation (1971) that "It is now generally accepted that the most important information for the perception of fricatives in a CV context is the central frequency Fc of the fricative noise and the frequency F1 of the second formant at the beginning of the transition from the fricative to the vowel."
2.2.4. Intensity of fricative consonants.

Intensity levels for speech are decided by the degree of loudness at which speech is to be produced. The intensity level for shouting, for instance, is higher than that for quiet speech. A speech sound may thus occur with various degrees of intensity.

The existence of intensity differences in vocal output is a well-known phenomenon and was not of interest in this study. A far more interesting phenomenon is the fact that there are consistent differences in acoustic energy between speech sounds of different categories (House and Fairbanks, 1953; Strevens, 1960; and Heinz and Stevens, 1961).

In the case of the fricatives in particular Strevens found that the overall relative intensities of the fricatives could be ranked in a hierarchical order. "It is of importance," as Strevens points out, "to discover whether this is because of inherent differences of acoustic energy or because of other factors such as those arising in the initiation and modification of the pulmonic air-stream."

It was with such an aim in view that an experiment to investigate the relationship between the rate of air flow of a fricative and its overall relative intensity was undertaken. As a first step, the overall intensities of the Arabic voiceless fricatives were determined.
These all occurred in initial position and were spoken with the same degree of physical effort. The results of the analysis suggested that the relative intensities of these sounds might be arranged as in Table 2.1:

Table 2.1

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>/a/</td>
</tr>
<tr>
<td>2</td>
<td>/s/</td>
</tr>
<tr>
<td>3</td>
<td>/ʃ/</td>
</tr>
<tr>
<td>4</td>
<td>/z/</td>
</tr>
<tr>
<td>5</td>
<td>/h/</td>
</tr>
<tr>
<td>6</td>
<td>/r/</td>
</tr>
<tr>
<td>7</td>
<td>/θ/</td>
</tr>
<tr>
<td>8</td>
<td>/n/</td>
</tr>
</tbody>
</table>

The second step was to study the rates of air flow of these fricatives. For this purpose, the eight Arabic voiceless fricatives in initial position were spoken with about the same degree of physical effort and their rates of air flow in front of the mouth recorded by means of the electrokymograph. The values for rates of air flow were then compared for different fricatives and an order of relative strength established starting from highest to lowest as Table 2.2 shows:
A comparison of the order of relative strength of rate of air flow with the order of overall relative intensity immediately reveals that variations in acoustic intensity among fricatives are not directly proportional to variations in rate of air flow. Indeed, in view of the fact that the arrangement of the fricatives in one table is almost the reverse of their arrangement in the other, a statement that acoustic intensity among fricatives is inversely related to rate of air flow would be the more representative of the facts. However, since it is unreasonable for an increase in rate of air flow to cause a decrease in acoustic intensity, or vice versa, one has to concede that the arrangement is coincidental.

Strevens' statement, then, that "variations in the air flow of speech have a major effect upon intensity but only a negligible
effect upon the spectrum of the sound produced" would seem to apply only to those cases where the same fricative is produced with the same vocal tract configuration but with different rates of air flow.

In the case of a given vocal tract configuration, as Flanagan (1965) points out, "the overall intensity of an unvoiced sound is directly proportional to the amplitude of the unvoiced source" and hence to rate of air flow. In the case of different vocal tract configurations, however, the contribution of the source, and hence rate of air flow, to the overall intensity would seem to be supplementary rather than primary. Indeed, Fant (1962) has indicated that "formant intensities as well as the intensities at any part of the spectrum will be predictable from the frequency locations of the resonances and anti-resonances supplemented by the additional information on the intensity and spectral composition of the source."

2.2.4.1. The role of intensity in fricative identification.

In section 2.2.3 brief mention was made of the role of overall relative intensity in fricative identification. It was pointed out that listeners' inability to distinguish between /f/ and /θ/ on the basis of energy distribution among the frequencies alone suggested that fricative identification may be dependent not only on this factor but also on the intensity of the fricative relative to the vowel.
Heinz and Stevens investigated the role of intensity in fricative identification by generating stimuli consisting of noise spectra followed by the synthetic vowel /a/. The overall level of the fricative was given three values: -5, -15, and -25 db. The stimuli were tape recorded in random order and presented to a group of listeners for identification.

As expected, analysis of listeners' responses revealed that the identification of /f/ and /θ/ greatly improved when the intensity of the fricative was low relative to that of the vowel. In particular, no /f/ responses could be obtained when the overall intensity of the fricative relative to the vowel was -5 db.

2.3. The acoustic analysis of Arabic fricatives.

This section reports the results of the investigation into the spectral and temporal properties of the Arabic fricative sounds /f/, /s/, /z/, /θ/, /ð/, /s/, /z/, /θ/, /ð/ and /h/. The plan of study involved the analysis of each fricative in both initial and final positions. The effect of the Arabic long vowels on the fricative, as well as the effect of the fricative itself on the adjacent vowel, are studied.

2.3.1. Frequency positions of the energy-density maxima.

For the analysis of the frequency positions of the most prominent
energy-density maxima in the spectra of Arabic fricatives, it was decided to produce "sections" which portray the amplitude vs frequency at a certain point in time in the steady-state friction part of a fricative. The decision to use "sections" rather than normal spectrograms in this part of the work resulted from the fact that it is often very difficult to determine with any accuracy the frequency positions of such energy-density maxima from normal spectrograms.

Eight sections were thus made for each of the Arabic fricatives in initial position only, followed by the vowel /iː/. The first step in the analysis was to note down, for each fricative, the frequency positions of all the energy-density maxima in its spectrum. Maxima occurring at similar frequency positions were then grouped together and the average frequency position of the group calculated. Only those maxima which occurred within a range of ± 100 Hz of the average frequency position for the group were included. Furthermore, if the number of occurrences of the items of a group was less than 50% of the cases studied, the group was excluded from the analysis. The amplitudes of the maxima of the group were then calculated and their sum divided by the number of the members of the group. The frequency positions of the most prominent maximum in the spectrum of a fricative was then calculated and plotted in Figures 2.1 and 2.2.

Figure 2.1 presents the frequency positions for the most prominent energy-density maxima in the spectra of the Arabic voiceless fricatives.
Figure 2.1. Frequency positions for the most prominent energy-density maxima in the spectra of the Arabic voiceless fricatives.
with the exception of /f/. Analysis of the Arabic voiceless fricatives has shown /f/ to have its most prominent energy-density maximum at 2325 Hz. Such a frequency position, however, was thought to be too low for the main resonance of a tube that is only 2 cm long at the utmost. Furthermore, previous analyses of fricative spectra (Hughes and Halle, 1956) have shown that the frequency position of the first major energy-density maximum in f-spectra is often located above 10 KHz; well beyond the range of our spectrograph. It was in consideration of these facts that it was decided to exclude /f/ from Figure 2.1.

Figure 2.2, on the other hand, presents the frequency positions for the most prominent energy-density maxima in the spectra of the Arabic voiced fricatives. The purpose of presenting the voiceless fricatives separately from the voiced ones is twofold. First, such a procedure facilitates comparison between the results of the acoustic analysis of the Arabic fricatives and those of previous studies, (Strvens, 1960; Hughes and Halle, 1956), which concentrated on the voiceless fricatives. Second, it was obvious that a separation of the two classes of fricative would bring out more clearly the relationships that hold between their members.

The results of the acoustic analysis of the Arabic fricatives are in good agreement with those reported by Hughes and Halle (1956). Hughes and Halle's data show that the frequency position for the
Figure 2.2. Frequency positions for the most prominent energy-density maxima in the spectra of the Arabic voiced fricatives.
first major resonance for /$/$ to be around 2200 Hz, while that for /s/ is in the range 3500 Hz to 6400 Hz. These values are reasonably close to the frequency positions of the first major resonances for Arabic /$/$ at 2665, and Arabic /s/ at 5975 Hz.

More importantly, the results of the acoustic analysis of Arabic fricative spectra support Hughes and Halle's, as well as Heinz and Stevens' observation that the frequency positions for the first most prominent energy-density maxima in fricative spectra are inversely related to the length of the cavity in front of the fricative obstruction in the vocal tract.

Indeed, both Figures 2.1 and 2.2 show that the general pattern for the distribution of the most prominent energy-density maxima among the frequencies is one where the farther back in the vocal tract a fricative is produced the lower the frequency position of its major energy-density maximum.

It is true that the frequency positions for the first major energy-density maxima in the spectra of /s/ and /h/ do not fit into the general pattern. One would have expected the relationship between /s/ and /g/ (Fig. 2.1) to be similar to that between /z/ and /z/ (Fig. 2.2), where the frequency position for the first major energy-density maximum of the alveolar fricative is higher than that of the
post-alveolar fricative. One would have also expected the frequency position for the first major energy-density maximum in the spectrum of /h/ to be lower than that for /h/. These, however, are rather the exception than the rule.

In general, the results of the analysis of Arabic fricatives support the observation made by Hughes and Halle (1956) and Heinz and Stevens (1961) on the presence of an inverse relationship between the length of the cavity in front of the fricative obstruction in the vocal tract and the frequency position of the most prominent energy-density maximum in fricative spectra. The present study shows, furthermore, that this relationship holds between the voiced as well as the voiceless fricatives, and extends beyond the front fricatives to the back fricatives as well.

2.3.2. The effect of position on the spectra of the Arabic fricatives.

Both in initial and final positions the Arabic fricative consonants appear on the spectrograms as random noise which is slightly more intense and concentrated in those regions close to the formants of the adjacent vowel.

The effect of position on the spectra of the Arabic fricatives is most evident in the case of the voiced fricatives where the
presence of voicing is greatly determined by the position of the 
fricative in the word. In initial position, the Arabic voiced 
fricatives are characterized by the presence of a low-band component 
which appears along the baseline at or below 300 Hz. In final 
position, by contrast, this component is sometimes absent either partly 
or wholly.

The position of the fricative in the word, furthermore, affects 
the frequency position of its vowel-like formants. Figures 2.3 to 
2.5 show this effect for the voiceless fricatives, while Figure 2.6 
shows it for the voiced ones.

Figures 2.3 to 2.5 present the frequency positions for the 
first vowel-like formant (Fig. 2.3), the second vowel-like formant 
(Fig. 2.4), and the third vowel-like formant (Fig. 2.5), as a function 
of their occurrence in initial position (solid lines) or final 
position (broken lines).

The voiceless pharyngealized post-alveolar fricative /ʃ/ has 
not, however, been included. This is due to the fact that it was not 
possible to detect any formant-like concentrations of energy in its 
steady-state friction part when it occurred in initial position. 
In fact, even in final position it was possible to detect only two 
vowel-like formants; one at 570 Hz and the other at 1350 Hz.
Figure 2.3. Frequency positions for the first vowel-like formants in the spectra of the Arabic voiceless fricatives in initial position (the solid line) and final position (the broken line).
Figure 2.4. Frequency positions for the second vowel-like formant in the spectra of the Arabic voiceless fricatives in initial position (solid lines) and in final position (broken lines).
Figure 2.5. Frequency positions for the third vowel-like formant in the spectra of the Arabic voiceless fricatives in initial position (solid lines) and in final position (broken lines).
Figure 2.3 shows that the frequency position of the first vowel-like formant is higher in initial position than in final position. In fact, this is true of all the Arabic fricatives with the exception of the back fricative /h/.

Figure 2.4, on the other hand, shows that the frequency position of the second vowel-like formant is lower in initial position than in final position. Once again, this is true of all the Arabic voiceless fricatives with the exception of the back fricatives /x/ and /h/.

Figure 2.5 shows that, similar to the vowel-like F2, the frequency position of the vowel-like F3 is lower in initial position than in final position. Similarly again, this is true of all the voiceless fricatives with the exception of the back fricatives /h/ and /h/.

As Figure 2.6 shows, the pattern of variation displayed by the first three vowel-like formants in the friction part of the Arabic voiced fricatives is similar to that displayed by their voiceless counterparts. Thus, the frequency positions of the second and third vowel-like formants for both /s/ and /z/ are lower in initial position than in final position. Also, the frequency positions of the first and third vowel-like formants for both /x/ and /ʃ/ are lower in initial position than in final position. No comparison, however, is possible in the case of /ʒ/ and /ʒ/.
Figure 2.6. Frequency positions for the first three vowel-like formants in the spectra of the Arabic voiced fricatives in initial position (the solid lines) and final position (the broken lines).
In general, the pattern of variation exhibited by the first three vowel-like formants of the Arabic fricatives (both voiced and voiceless) is one where the frequency position of F1 is higher in initial position than in final position, whereas the frequency positions of F2 and F3 are lower in initial position than in final position.

2.3.3. The effect of the adjacent vowel on fricative spectra.

Figures 2.7 to 2.8 present, for each of the Arabic fricatives, the frequency positions of the first three vowel-like formants as a function of the adjacent vowel. The lines connect points representing the average values for 20 occurrences in both initial and final positions.

In general, the frequency position of a particular vowel-like formant in the spectrum of a fricative varies directly with the frequency position of the adjacent vowel formant. The degree of variation exhibited depends, however, on the number of that formant and the role of the tongue in the production of that particular fricative.

Firstly, the degree of variation exhibited by the second vowel-like formant in fricative spectra is much greater than that displayed by either the first or third vowel-like formants. This reflects the
Figure 2.7. Average frequency positions for the first three vowel-like formants in the spectra of /i:/, /e:/, /a:/, /o:/, and /u:/ as a function of the adjacent vowel.
Figure 2.8. Average frequency positions for the first three vowel-like formants in the spectra of /a/, /e/, /i/, /o/, /u/ and /s/ as a function of the adjacent vowel.
the fact that the range of frequencies occupied by the second formants of adjacent vowels is much greater than that occupied by either the first or third formants.

Secondly, and more importantly, the results of the analysis of the effect of the adjacent vowel on the spectra of Arabic fricatives suggest that while the frequency position of a vowel-like formant in the spectrum of a fricative generally varies directly with that of the adjacent vowel formant, the degree of such variation depends on the role of the tongue in the production of that fricative. More specifically, a vowel-like formant in the spectrum of a fricative shows the maximum degree of variation with the adjacent vowel formant when the tongue does not play an active role in the production of that fricative /f, x, k, h/ and /h/, and the minimum degree of variation when it does /s, z, ñ, s, z/ and /s/.

This, indeed, is not surprising. The tendency of those fricatives where the tongue plays an active role in their production to show the most resistance to coarticulation derives from the fact that in their production the tongue has to make a particular gesture towards a particular point in the vocal tract and stay there for the duration of the fricative. As such, the tongue is neither free to assume a during their production, an anticipatory position for the next sound, nor to remain at the position assumed for the previous sound.
In the production of the fricatives /f, x, v, h/ and /h/, on the other hand, there are no such constraints on the movement of the tongue with the result that it is either wholly free /f, h/ and /h/, or partly so /x and v/. Furthermore, the absence of any significant differences in the degree of coarticulation of these two kinds of fricatives suggests that the tongue, or the main part of it, has much the same degree of freedom of movement in both cases.

2.3.4. Duration.

Figure 2.9 presents the duration values for the Arabic fricatives in initial position (solid lines) and in final position (broken lines). In each case the top values represent maximum durations, the bottom values represent minimum durations, while the dashes represent average durations. The figure makes clear that the duration of an Arabic fricative varies with word position, voicing, and point of articulation.

First, for all the Arabic fricatives and in common with most other Arabic speech sounds studied in this work, the duration of a fricative is considerably greater in final position than in initial position. In fact, apart from /h/, the minimum duration value for a fricative in final position is always greater than the maximum value for the same fricative in initial position.
Figure 2.9. Maximum and minimum duration values for the Arabic fricatives in initial position (solid lines) and final position (broken lines) in milliseconds.
Second, the voicing of a fricative seems to have a considerable effect on its duration. As the figure makes clear, the Arabic voiceless fricatives both in initial and final positions have consistently greater duration than their voiced counterparts.

Third, the duration of a fricative seems to vary with its point of articulation. The direction of variation depends, however, on the presence or absence of voicing. Thus, while in the case of the voiceless fricatives the farther back in the vocal tract a fricative is produced the greater its duration, in the case of the voiced fricatives the opposite is true.

The tendency of the voiceless fricatives to have greater duration as their point of articulation moves farther back in the vocal tract could be related to the fact that the agility of the tongue decreases in the same direction. The reasons why this should not happen in the case of the voiced fricatives as well are, however, unclear to me.

2.3.5. Transitions.

Figures 2.10 to 2.19 present, for each of the Arabic fricatives, the average steady-state frequency positions as well as the onset and
Figure 2.10. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /f/ (left-hand side) and before word-final /f/ (right-hand side)
Figure 2.11. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /s/ (left-hand side) and before word-final /s/ (right-hand side).
Figure 2.12. The average onset, offset, steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /z/ (left-hand side) and before word-final /z/ (right-hand side).
Figure 2.13. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /s/ (left-hand side), and before word-final /s/ (right-hand side).
Figure 2.14. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /w/ (left-hand side), and before word-final /w/ (right-hand side).
Figure 2.15. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /S/ (left-hand side) and before word-final /S/ (right-hand side).
Figure 2.16. The average onset, offset, and steady-state frequency positions for the second formant of the Arabic long vowels after word-initial /x/ (left-hand side), and before word-final /x/ (right-hand side).
Figure 2.17. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /b/ (left-hand side), and before word-final /b/ (right-hand side).
Figure 2.18. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /h/ (left-hand side), and before word-final /h/ (right-hand side).
Figure 2.19. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /h/ (left-hand side), and before word-final /h/ (right-hand side).
offset frequency positions of the Arabic long vowels when they occur after word-initial fricatives and before word-final fricatives. Steady-state frequency positions are represented by squares while onset and offset frequency positions are represented by circles.

As the figures show, the direction of F2-transition varies with the quality of the vowel, the point of articulation of the fricative, and the position of the fricative in the word.

For example, while the Arabic fricative /f/ has a negative transitional influence on the second formant of /i:/, it has a positive transitional influence on that of /u:/, This, of course, is due to the fact that the /f/-locus is lower than the steady-state frequency position of the second formant of /i:/ and higher than that of /u:/. The F2-transition of /a:/ is negative after /f/, neutral after /z/, and positive after /s/. The variation is due in this case to the difference between the steady-state frequency position of the second formant of /a:/ and the frequency positions of the loci of /f/, /s/, and /s/.

The direction of F2-transition also varies with the position of the adjacent fricative in the word. For example, when /s/ occurs
in word-initial position it has a negative transitional influence on the second formant of /a:/, whereas, by contrast, it occurs in word-final position it has a neutral transitional influence.

The extent of transition similarly varies with the quality of the vowel, the frequency position of the fricative locus, and the position of the fricative in the word.

In general, the extent of transition of the second formant of an Arabic vowel is a direct function of the degree of closeness of that vowel. The extent of transition of the vowel second formant is greater in the vicinity of those fricatives which have a locus with a stable frequency position than in the vicinity of those which have one with a shifting frequency position. The effect of word-position on the extent of F2-transition is determined, however, by the role of the tongue in the production of the adjacent fricative.

2.3.5.1. Direction of transition.

Table 2.3 presents in a condensed form the transition directions for the second formants of the Arabic long vowels when they are preceded by word-initial fricatives (the top row for each vowel), and when they are followed by word-final fricatives (the bottom row for each vowel). Plus signs (+) are used for positive transition, minus signs (-) for negative transitions, and blank spaces for neutral
transitions. The vowels are arranged on the left-hand side starting with the close front vowel moving anti-clockwise to the close back vowel. The fricatives, on the other hand, are arranged on the bottom row according to their point of articulation, starting with the labiodental fricative and moving backwards in the vocal tract to the glottal fricative.

Table 2.3

<table>
<thead>
<tr>
<th></th>
<th>f</th>
<th>s</th>
<th>z</th>
<th>x</th>
<th>ʔ</th>
<th>n</th>
<th>ɬ</th>
</tr>
</thead>
<tbody>
<tr>
<td>i:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>e:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>a:</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>o:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>u:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

61
In general, Table 2.3 presents a picture of what one would expect with negative transitions dominating at the top, positive transitions at the bottom, and neutral transitions in the middle. This, of course, could only be due to the fact that the frequency positions for the second formants of both /i:/ and /e:/ are higher than those for the loci of all the Arabic fricatives, while those for /o:/ and /u:/ are, by contrast, lower. /a:/, on the other hand, occupies a middle position in this set.

Looking at Table 2.3 column by column one finds that /f/ seems to have a negative transitional influence on the front vowels, and a positive transitional influence on the back vowels. The homorganic pair /s,z/ on the other hand, exert a negative transitional influence on /i:/ and /e:/ only, and a positive transitional influence on the back vowels. Their influence on /a:/ can be negative or neutral.

The pharyngealized pair /ʃ,z/ behave in a nearly similar fashion to that of /s,z/. They exert a negative transitional influence on /i:/ and /e:/, and a basically positive transitional influence on /o:/ and /u:/.

The palatal fricative /ʃ/ is different from all other Arabic fricatives in that it exerts a positive transitional influence on the second formant of /a:/ in both word-initial and word-final
positions. Similar to other fricatives, however, it exerts a negative transitional influence on /i:/ and /e:/ and a positive transitional influence on /o:/ and /u:/.

As expected, the F2-transition directions for the back fricatives are mainly negative. This derives from the fact that as the point of articulation for a fricative moves farther back in the vocal tract, the frequency position for the associated locus would decrease. Consequently, the Arabic fricatives would increasingly show a tendency to exert a negative transitional influence on the second formants of the adjacent vowels.

Apart from word-final /h/ when preceded by /o:/, the Arabic back fricatives all exert a negative transitional influence on all the Arabic front vowels. Their effect on the back vowels is, however, mainly positive. In fact, apart from /t/ which seems to exert a negative transitional influence on all vowels regardless of whether or not they are back or front, the back fricatives show the expected positive transitional influence on the position of the second formants of the adjacent back vowels.

Finally, the Arabic fricatives seem to have struck a fine balance between their effect on the direction of F2-transition in initial and
final positions. As Table 2.3 shows, the Arabic fricatives normally exert either a positive or a negative transitional influence, and the limited number of neutral transitions that they exert is shared equally between the initial and final positions.

Basing our evaluation of the directionality of coarticulatory effects of the Arabic fricatives solely on Table 2.3 one has to conclude that the coarticulatory effects of the Arabic fricatives do not seem to be significantly influenced by their position in the syllable; they are as powerful in a vowel-consonant context as in a consonant-vowel one.

2.3.5.1. Extent of transition.

In investigating the effect of the Arabic fricative consonants on the extent of transition of the second formants of the Arabic long vowels, the average onset, or offset, frequency position for the second formant of each vowel was subtracted from its steady-state frequency position following, or preceding, a particular fricative.

To arrive at a measure of the extent of transition of the second formant for a particular vowel following, or preceding, all the Arabic fricatives, these values were squared, added up, and their sum divided by the number of the Arabic fricatives. Finally, the square root of this quotient was calculated and plotted in Figures 2.20 and 2.21.
The squaring of the difference between the onset or offset, frequency position and the steady-state frequency position was the result of the fact that this difference was positive in cases of negative transition and negative in cases of positive transition. If these values were then added up, without first squaring them, the negative values would have cancelled the positive ones, and the resulting average would not have been a true representative of the extent of transition.

Figure 2.20 presents the extent of F2-transition as a function of the vowel. As expected, the figure shows that the closer a vowel is the greater the extent of transition of its second formant. Thus, the extent of transition for the second formant of the close front vowel /i:/ is greater than that for the second formant of the half-close vowel /e:/.

Similarly, the extent of transition for the second formant of the close back vowel /u:/ is greater than that for the half-close vowel /o:/.

The extent of transition for the second formant of the open vowel /a:/ is, predictably enough, the smallest.

Figure 2.21 presents, on the other hand, the extent of transition for the second formants of the Arabic long vowels as a function of the adjacent fricative. As Figure 2.21 shows, the effect of a fricative is greatly determined by whether it belongs to one of two groups of fricative. The first group comprises those fricatives where the
Figure 2.20. The extent of F2-transition of the Arabic long vowels after word-initial fricatives (solid lines) and before word-final fricatives (broken lines).
Figure 2.21. The extent of F2-transition of the Arabic long vowels as a function of the adjacent fricative when it occurs in word-initial position (solid lines), and when it occurs in word final position (broken lines).
tongue plays an active part in their production. The second group, by contrast, comprises those where it does not.

The two groups differ remarkably in their effect on the extent of F2-transition. As Figure 2.21 shows, the extent of F2-transition for vowels adjacent to the first group is twice as much as that for vowels adjacent to the second group.

The two groups differ again in the way their word position affects the extent of F2-transition of the adjacent vowel. Thus, the extent of F2-transition for a vowel adjacent to a fricative from the second group is greater when that fricative occurs in word-final position than when it occurs in word-initial position. In the case of the fricatives of the first group, by contrast, the effect of word position on the extent of F2-transition is negligible.

The observed greater extent of F2-transition for vowels adjacent to those fricatives in which the tongue is the active articulator is related to a phenomenon already encountered in Section 2.3.3. There we found that it is precisely in the case of these fricatives that the frequency positions of the vowel-like formants in the steady-state friction part show the most resistance to coarticulation with the varying adjacent vowel. The frequency positions of the vowel-like
formants of the other group, by contrast, always shifted in sympathy with the frequency positions of the formants of the adjacent vowel.

This suggests that the greater influence of the first group on the extent of F2-transition is due to the stable locus position of the fricative. In the case of the second group, by contrast, the smaller extent of F2-transition would be the result of a situation where the fricative shifting locus position is always nearer to the frequency position of the second formant of the adjacent vowel than would have been otherwise.

The reasons why this second group of fricatives should exert more influence on the extent of F2-transitions when they occur in word-final position than in word-initial position are, however, unclear to me.

The observed disagreement between our findings on the direction of F2-transition and those on the extent of F2-transition raises the interesting question of which to take as a measure of coarticulatory influence. This, however, should be the subject of a separate study.
2.4. Conclusion.

This section is naturally divided into two main parts. The first evaluates the relevance of some of the findings on the acoustics of the Arabic fricatives to the general theories of fricative production and perception. The second, on the other hand, attempts to provide an overall view of the specifics of the Arabic fricatives dealt with separately in the previous sections.

2.4.1. General discussion.

The acoustic characteristics of the Arabic fricatives - in particular, the fact that acoustic energy in their spectra is organized in well-defined bands, and the general pattern of distribution of these bands of energy among the frequencies for the different fricatives - support Fant's suggestion (1960) that "the typical fricative is a noise sound, the spectral energy of which is largely contained in formants from cavities in front of the articulatory narrowing." (See Sections 2.2.2. and 2.3.1)

The results of the acoustic analysis of the Arabic fricatives also support the observation made by Hughes and Halle (1956) and Heinz and Stevens (1961) on the presence of an inverse relationship between the length of the cavity in front of the fricative obstruction
in the vocal tract and the frequency position of the most prominent energy-density maximum in fricative spectra. Furthermore, they also show that this relationship holds between the voiced as well as the voiceless fricatives and extends beyond the front fricatives to the back fricatives as well. (See Sections 2.2.2 and 2.3.1)

The results of our perceptual tests on fricative identification are, however, in disagreement with those reported by Harris (1954) and Heinz and Stevens (1961). Both studies have concluded that while the important cues for /s/ and /ʃ/ are given by the friction part, those for /f/ and /θ/ are contained primarily in the transition part of the syllable. Our perceptual tests have, by contrast, shown that both /f/ and /θ/ could be discriminated with the aid of cues in the friction part alone. (See Sections 2.2.1, 2.2.2, and 2.2.3)

The success of the listeners in our experiment to correctly identify the various fricative stimuli may be attributed to the fact that these stimuli were spoken rather than synthesized. This is, no doubt, an indication of the frailty of the cues in the friction part which seem to be difficult to reproduce faithfully on the one hand, or to be easily damaged in the process of cutting and recombination on the other. This frailty seems to support the common belief that the cues in the transition part are the overdominant ones. More
importantly, at least to an acoustic phonetician, the discrepancy between the results of natural speech perceptual tests and those of synthetic speech is a warning against too readily accepting synthetic speech as a true copy of natural speech.

2.4.2. Coarticulation and word-position effects.

The effects of coarticulation and word position determine both the spectral and temporal characteristics of the Arabic fricatives as well as the spectral properties of the adjacent vowel.

In this section we shall first deal with the variation of the spectral characteristics of the fricatives as a function of word position and the adjacent vowel. We shall then discuss how the vowel spectrum is itself affected by the adjacent fricative and word position. Lastly, we shall examine the factors that determine the temporal properties of the Arabic fricatives.

2.4.2.1. The effect of word position on fricative spectra.

The acoustic analysis of the Arabic fricatives has revealed that their spectral characteristics are significantly determined by their position in the word. The spectrum of a fricative varies as a function of this factor either by the total or partial loss of one of its components, or by a change in the locations of their frequency positions.
Firstly, the spectra of the Arabic voiced fricatives in initial position are characterized by the presence of a low-band component which appears along the baseline at or below 300 Hz. In final position, by contrast, this component is often absent either partly or wholly.

Secondly, the frequency positions of the vowel-like formants of the Arabic fricatives are influenced by word position in such a way that, for all fricatives, the frequency position of F1 is higher in initial position than in final position, whereas the frequency positions of F2 and F3 are lower in initial position than in final position. (See Section 2.3.2)

2.4.2.2 The effect of the adjacent vowel on fricative spectra.

The quality of the adjacent vowel significantly affects the frequency positions of the vowel-like formants of the Arabic fricatives. In general, the frequency position of a particular vowel-like formant in the spectrum of a fricative varies directly with the frequency position of the adjacent vowel formant. The degree of variation depends, however, on the number of that formant and the role of the tongue in the production of the fricative.

Firstly, the degree of variation exhibited by the second vowel-like formant in fricative spectra is much greater than that displayed by either the first or third vowel-like formants. This reflects the
fact that the range of frequencies occupied by the second formants of the adjacent vowels is much greater than that occupied by either the first or third formants.

Secondly, our investigation into the acoustic characteristics of the Arabic fricatives has revealed that the degree of coarticulation between the fricative and the adjacent vowel is greatly determined by the role of the tongue in the production of the fricative. Where the tongue plays an active role in the production of a fricative, the spectral characteristics of the fricative show the least degree of variation with the adjacent vowel. Where it does not, the opposite is true. (See Section 2.3.3)

2.4.2.3 The effect of the adjacent fricative and word position on the vowel spectrum.

The spectral characteristics of the vowel are similarly affected by the quality of the adjacent fricative, at least as far as the direction and extent of F2-transition are concerned.

First, the direction of F2-transition is mainly the product of the interaction between the frequency position of the highest energy-density maximum in the spectrum of the fricative and the frequency position of the steady-state second formant of the adjacent vowel.
Where the former occupies a higher frequency position than the latter the direction of F2-transition is positive, and, conversely, where it occupies a lower frequency position the direction of F2-transition is negative.

The extent of F2-transition is, on the other hand, determined by the quality of the vowel and the role of the tongue in the production of the fricative. First, the closer a vowel is the greater the extent of its F2-transition. And, second, the extent of F2-transition is much greater in the vicinity of those fricatives where the tongue plays an active role in their production than in the vicinity of those where it does not.

Both the direction and extent of F2-transition do not seem to be significantly influenced by the position of the fricative in the word. In general, the coarticulatory effects of the Arabic fricatives seem to be as powerful in a vowel-consonant context as in a consonant-vowel one.

2.4.2.4 Factors affecting fricative temporal properties.

The duration of an Arabic fricative varies with word position, voicing, and the point of articulation of that fricative. First, for all fricatives, the duration of a voiceless fricative is consistently greater in final position than in initial position. Second, the
Arabic voiceless fricatives both in initial and final positions consistently have greater durations than their voiced counterparts. And, third, the farther back in the vocal tract a fricative is produced the greater its duration. This last phenomenon, however, does not apply to the voiced fricatives where the opposite is true.
Chapter 3

The stops

3.1. Articulatory mechanism of stop production.

3.2. The shutting phase: The role of transitions.

3.3. The opening phase: The role of bursts.

3.4. The voiced/voiceless distinction.

3.4.1. The presence vs. the absence of voicing during the closure phase.

3.4.2. Relative duration between burst and onset of vowel resonance.

3.4.3. Relative duration of vowel resonance before the closure phase.

3.4.4. The relative intensity of the stop burst.
   A. Tenseness and laxness.
   B. Duration and intensity of stop explosion.

3.4.5. Relative duration of the closure phase.

3.4.6. Extent of F1-transition.

3.5. The acoustic analysis of the Arabic stops.

3.5.1. The effect of position on the spectra of the Arabic stops.
3.5.1.1. The voiced stops.
3.5.1.2. The voiceless stops
3.5.1.3. The glottal stop.
3.5.1.4. The velar/palatal pair /k, g/.
3.5.2. The effect of the adjacent vowel on the spectra of the Arabic stops
3.5.3. Duration.
3.5.4. Transitions.
3.5.4.1. Direction of Transition.
3.5.4.2. Extent of transition.
3.6. Conclusion.
3.6.1. Coarticulation and word-position.
3.6.1.1. The effect of word position on stop spectra.
3.6.1.2. The effect of the adjacent vowel on stop spectra.
3.6.1.3. The effect of the adjacent stop and word position on the vowel spectrum.
3.6.1.4. Factors affecting stop temporal properties.
THE STOPS

3.1 Articulatory mechanism of stop production.

There are three phases in the articulation of a stop: (1) the shutting phase, (2) the closure phase, and (3) the opening phase. In the shutting phase, the articulators move to the particular point of articulation of the stop. This is followed by the closure phase where the flow of air is completely interrupted (since there is a velar closure as well as the articulatory stricture of complete closure) and where pressure is built up. In the final phase, i.e., the opening phase, the barrier to the air stream is removed rapidly thus releasing the air blocked behind the point of complete closure usually in the form of an explosion. During the closure phase the vocal cords may, or may not, vibrate: if they do, we have a voiced stop; if they do not, we have a voiceless stop.

The movement of the articulators causes changes in the vocal cavity size and shape thus affecting the short-time energy spectrum. These changes appear on the spectrograms in the form of formant transitions. The closure phase, on the other hand, appears on the spectrograms as a gap denoting the absence of energy in all bands. In the case of voiced stops, however, a low band component appears on the spectrograms along the baseline below 300 Hz. The sudden release of the articulators usually produces a short burst of noise which appears on the spectrograms as a noise component.
3.2. The shutting phase.

The role of transitions.

The development of the concept of "transition" illustrates well the difficulties encountered by researchers in the early days of spectroscopy. At that time, speech was often portrayed as a succession of discrete units for the production of which the organs of speech assume certain postures for the consonants different from those for the vowels. The rapid movements of the articulators from one posture to the other were thought of as no more than the necessary 'transitions' between the postures that serve to identify the successive phonemes. This meant that the corresponding rapid shifts in the acoustic output are mere nulls which dilute or even confuse the acoustic message.

Experiments with synthetic speech suggested, however, that these rapid changes might be heard as important distinguishing characteristics of the sound stream, and might, indeed, serve as a principal cue for the perception of speech sounds.

The importance of formant transitions for the perception of speech sounds was established after it was found out that the distinction among the voiceless stops /p, t, or k/ and the voiced stops /b, d, or g/ and the nasal resonants /m, n, or q/ is a function of the direction and extent of F2-transition.

At the same time there were attempts to accommodate the findings of these experiments within the classificatory systems of articulatory phonetics. It was thus suggested, among other things, that F2-transition
correlates with articulatory place of production. It was not long, in fact, before what had been tentatively suggested in 1952 (Cooper, Delattre, Liberman, Borst and Gertsman) was taken as an established fact in later years.

The equation of F2-transition with articulatory place of production assumes that there is a simple one-to-one correspondence between articulation and sound. Liberman (1957), in fact, points out that "to the extent that there is a one-to-one correspondence between articulation and the acoustic result", the correlation of F2-transition with articulatory place of production is "neither more nor less than we should expect."

Unfortunately, however, there is not entirely a one-to-one correspondence between articulation and the acoustic result. A specific articulation may well have a fairly constant effect on the spectrum of the phone being produced, but this effect need not be unique to that articulation. Thus, rounding of the lips may always produce a fall in the frequency of F2, but such a fall may also result from other, entirely different articulations, the pharyngealization of a phone, for example.

In the production of the pharyngealized consonants of Arabic, the root of the tongue moves towards the pharynx, thus enlarging the volume of the vocal cavity. As acoustic theory tells us, there are two factors that affect the effective resonance of a cavity - its volume and its opening: the larger the cavity and the smaller the opening, the lower the note of resonance; and vice versa. The large volume of the vocal cavity thus results in a fall of the frequency of F2, even in the absence of lip rounding.
The phenomenon of pharyngealization in Arabic is, in fact, a good example for illustrating the fallacy of correlating F2-transition too readily with articulatory place of production. In the case of the non-pharyngealized consonant /d/, F2-transitions of adjacent vowels point to a locus frequency of about 1800 Hz. In the case of the pharyngealized consonant /d/, by contrast, F2-transitions of adjacent vowels point to a much lower locus frequency at about 1200 Hz. Yet, in spite of this extensive displacement in the direction of F2-transition, the fact still remains that both the pharyngealized and the non-pharyngealized pair have approximately the same primary place of production.

To sum up, the concept of transition is a useful tool in the description and presentation of acoustic data, and there are certainly cases where the direction of any one of these transitions matches up with an articulatory place of production. The attempt, however, to generalize and correlate F2-transition throughout with articulatory place of production can only harm our understanding of how language operates.

3.3. The opening phase.

The role of bursts.

The role of bursts as a cue for the identification of different stops has been investigated by many researchers. In an experiment with natural speech, Halle, Hughes and Radley (1957) selected words ending in stops that were produced with a burst and without vocal cord vibration. They then gated out and recorded the first
20 ms of each stop burst. The gated stop bursts were presented to a group of listeners with instructions to judge them as /n/, /t/ or /k/. The percentage of correct responses by the listeners was so high that the experimenters concluded that the bursts in isolation are identifiable as particular stops.

The investigators also concluded that the cues that make possible the identification of the bursts as different stops must reside in the spectrum. Further examination of bursts spectra did, in fact, reveal that they have different spectral properties which correlate with their different places of articulation. The differences can be stated as follows:

The postdental stops have either a flat spectrum or one in which the higher frequencies (above 4000 Hz) predominate, aside from an energy concentration in the region of 500 Hz.

The palatal allophones of /k/ and /g/ have energy concentration in the region between 2000 Hz and 4000 Hz, while the velar allophones have energy concentration at much lower frequencies.

The labial stops have a primary concentration of energy in the low frequencies (500-1500 Hz).

The findings of the above experiment illustrate the general expectation that there is a one-to-one correspondence between our perception of speech sounds and the spectral properties of these sounds. This, however, does not seem the whole case.
In an experiment with synthetic speech, Liberman, Delattre and Cooper (1952) investigated the effect on perception of minimal changes in the frequency position of a brief burst of noise presented in initial position in the syllable. Bursts at twelve different frequency positions were used, each one paired with seven synthetic vowel sounds to form the test syllables. These 84 syllables were presented in random order to a group of subjects who were asked to identify the initial stop sound as /p/ /t/ or /k/. Bursts at 3000 Hz were identified as /t/ before every vowel. In this range the conclusion can be drawn that the frequency position of the burst determines how it will be heard, regardless of the context. Below 3000 Hz, however, whether the burst was heard as /p/ /t/ or /k/ depended on the combination of burst plus vowel, not on the burst alone. The bursts which yielded /k/ responses varied from a frequency of 2880 Hz before /i/ to as low a frequency as 720 Hz before /u/. The 720 Hz burst was heard as /p/ before /a/, as /k/ before /u/. The 1440 Hz burst before /i/ and /u/ was heard as /p/, and before /a/ as /k/. The 2880 Hz burst before /i/ was heard as /k/ before /a/ as /t/.

These results indicate clearly that in the frequency range below 3000 Hz, identification of a stop consonant depends not only on the burst, but on the burst in relation to the vowel which follows it. The consonant-vowel combination is of primary importance in the perception of these synthetic stops: placing a burst before two different vowels changes the way in which the burst is perceived.

The assumption that speech perception depends upon relative and not on absolute values has been strengthened by the results of
experiments with natural speech as well. Schatz (1954) recorded the three syllables /ski/, /ska/ and /sku/. The /sk/ segment was then cut away from the vowel in the recorded syllables, and spliced back before each of the three sequences /id/, /ar/ and /ul/. The three /sk/ segments, each combined with the three vowels, yielded a total of nine combinations. Five samples were made up of each of these combinations and the 45 syllables were presented in random order to 20 subjects, who were asked to identify the initial cluster as /sk/, /sp/ or /st/.

In the case of those syllables in which the /k/ burst was combined with the vowel which it originally preceded, the burst was perceived as /k/ in every case. In other combinations, however, /k/ was not the preferred judgement. The burst from /ski/ was perceived as /t/ when combined with /a/ and as /p/ when combined with /u/. The burst from /ska/ was perceived as /p/ when combined with both /i/ and /u/. The burst from /sku/ was perceived as /p/ before /a/ and as both /p/ and /k/ before /i/. "These responses indicate," as Schatz (1954) points out, "that the perception of a spoken /k/, as /k/ or as some other stop, depends very much on the vowel following and not on the /k/ alone."

The results of experiments both with natural and synthetic speech thus indicated that the context of an initial voiceless stop is an important factor in its perception. Furthermore, the fact that "an acoustically identical stimulus gives the impression of different consonants before different vowels", as Brosnahan and Malinberg point out (1970), has the important implication that "there is a lower limit to segmentation in the acoustic aspect, and this limit is not at the level of the discrete speech sound or phone."
3.4. The voiced/voiceless distinction.

The discussion presented so far deals only with those cues that serve to distinguish among the different classes of stop consonants - bilabial, dental, or velar - rather than with those cues that serve to distinguish within the classes. The frequency position of the burst and the direction and extent of F2-transition serve as cues to distinguish /k/ for instance, from /d/ and not from /g/.

The task of looking in the acoustic layer for the cues that might possibly be used by listeners for distinguishing /d, g/ from /t, k/ has been made relatively easy thanks to the pioneering work of various researchers in the field of acoustic phonetics. An extensive study of the literature on the subject has revealed that the search for the cues separating /d, g/ from /t, k/ may concentrate on:

1. The presence vs the absence of voicing during the closure phase (Potter, Kopp and Kopp; 1947);
2. The relative duration between the burst and the onset of vowel resonance (Potter, Kopp, and Kopp; 1947);
3. The relative duration of vowel resonance before the closure phase (P. Denes; 1955);
4. The relative intensity of the burst (Halle, Hughes and Radley; 1957);
5. Rate and extent of formant transitions (Leigh Lisker; 1957);
6. The relative duration of the closure phase (Leigh Lisker; 1957);
7. The extent of F1-transition (Cooper, Delattre, Liberman and Gertsman; 1952).
3.4.1. The presence vs. the absence of voicing during the closure phase.

Spectrograms made of Arabic stops in word-initial position show that it is quite easy to separate /d דלת/ from /t典雅k/ on the basis of a simple voiced/voiceless opposition. In this position, /d דלת/ are completely voiced, while /t典雅k/ are completely voiceless.

In other word positions, however, a simple voiced/voiceless opposition fails to separate both categories of stop consonants. Thus, although it is often the case that in word-final position /d דלת/ are found to be completely or partly voiced while /t典雅k/ are always completely voiceless, there are cases where /d דלת/ occur as completely voiceless as /t典雅k/.

In intervocalic position, furthermore, the voiced/voiceless opposition again fails to separate /d דלת/ from /t典雅k/, though for a different reason this time. It is true that /t典雅k/ do sometimes appear on spectrograms with virtually no glottal activity during the closure phase, while /d דלת/ are always marked by the presence of a low-band component during this phase. Yet, it is too often the case that voicing continues uninterrupted from the preceding vowel into the closure phase of a following stop, regardless of its nature, for the opposition to operate.

In short, acoustic energy referable to a simple voiced/voiceless opposition can be used quite successfully to separate both categories of Arabic stop consonants in word-initial position. In other positions, however, it fails.
3.4.2. Relative duration between burst and onset of vowel resonance.

An initial attempt was made to separate both categories of stops by using the relative duration between the burst and the onset of vowel resonance as a measure in interpreting the spectrograms. However, it was soon realized that the use of this measure is not possible for two reasons. First, it is often impossible to detect a stop burst as distinct from neighbouring vowel resonance in the case of the voiced stops /d ɡ/ or from aspiration in the case of /t ʈ k/. Second, there sometimes occur more than one burst both in the case of /d ɡ/ as well as in the case of /t ʈ k/. The choice of one stop burst rather than another would certainly be arbitrary and is impossible to justify when the stop burst or click occurs in the middle of the closure phase. The procedure had, therefore, to be abandoned and another procedure which takes into consideration only the duration of the aspiration phase was used.

As has already been observed, in word-initial position it is often impossible to detect a stop burst of any sort, and vowel resonance (or aspiration) starts immediately after the closure phase. In this position the vertical striations characteristic of vowel resonance appear to start immediately after the closure phase for /d ɡ/. The closure phase of /t ʈ k/, on the other hand, is always separated from vowel resonance by a period of aspiration. The duration of this aspiration period varies, as expected, from one class of stop consonant to the other. Thus it is about 30 ms on average for the dental and post-dental stops, and about 45 ms on average for the velar stops.
The presence vs. the absence of aspiration separates both categories of Arabic stop consonants in intervocalic position as well. Unexpectedly, however, in this position it is the dental stop /t/ that has a longer period of aspiration than that of /k/. Duration values for /t/ fall in the range 45-75 ms, with an average value of 60 ms, while values for /k/ vary from 30 ms to 60 ms, with an average value of 45 ms.

These data point out that for Arabic at least aspiration is a more powerful cue than voicing in separating /d d g/ from /t t k/ at least in word-initial and intervocalic positions. They also suggest that the differences between both categories of stop consonants can be stated simply in terms of an aspirated/unaspirated opposition.

In final position, however, the situation is not so simple. Thus, in spite of the fact that /t t k/ are often distinguished from /d d g/ by longer periods of aspiration and also by the almost total absence of energy in the lower frequencies, while /d d g/ are characterized by shorter periods of aspiration and by the presence of a low-band component along the baseline, both categories may occur unexploded and unaspirated.

Thus neither the adoption of an aspirated/unaspirated contrast nor the use of a more aspirated/less aspirated procedure will succeed in adequately separating /d d g/ from /t t k/. This, however, should not mean that the categories are not distinct from each other. It merely means that the use of a single measure fails to separate them.
3.4.3. **Relative duration of vowel resonance before the closure phase.**

Accepted phonetic descriptions of English, along with experiments with natural speech (P. Denes; 1955) indicate that the relative durations of vowel and final consonant can be used as a cue for hearing the final sound as voiced or voiceless. This suggested that it might also be the case that this feature operated in Arabic as well to separate the voiced stops from their voiceless counterparts.

An experiment was thus made to investigate the presence or absence of this cue. 20 words ending in stop + vowel + stop were used. They consisted of five groups. Each group contained one of the five Egyptian Arabic long vowels, and four words ending in /t, d, k, or g/.

The investigation revealed that the duration of vowels does not vary with the voicedness or voicelessness of the final stop consonant. The relative duration of vowel resonance before the closure phase is not thus an operative cue in the voiced/voiceless distinction in Egyptian Arabic.

3.4.4. **The relative intensity of the stop burst.**

A. **Tenseness and laxness**

In diagnosing the phonetic basis for listeners' ability to distinguish between categories of stop consonants, linguists and phoneticians have often invoked the dimension of "force of articulation" to call one category "fortis" or "tense" and the other "lenis" or "lax".
By contrast with other phonetic dimensions, such as voicing and friction, which are characterized by unambiguous physical correlates, the correlates of the tense/lax distinction have always been a subject of controversy. Hudgins and Stetson (1935), for instance, point out that "the fundamental differentiation of surd and sonant occlusives is the result of vertical movements of the larynx-hyoid unit. The larynx begins to move downward at the closing of the sonant occlusive; it may even start slightly in advance of the closure. It continues to move downward throughout occlusion ... During the surd occlusion, on the other hand, the larynx-hyoid unit does not move downward." Hardcastle (1973), on the other hand, has found that during the closure phase of the tense unaspirated stops of Korean /P T K/ "the larynx is sharply lowered."

Statements about the role of aspiration in implementing the opposition of tense and lax stops have also been conflicting. For Heffner (1964), for instance, tenseness of a stop is determined by the degree of its aspiration. Jakobson, Fant and Halle (1952), however, state that in Lezgian and Ossete the "feature of aspiration marks the lax stops in contradistinction to the tense."

In fact, even the status of the tense/lax distinction itself has always been somewhat of a controversial issue among linguists and phoneticians. Sweet (1906), Jones (1918), Stetson (1951), Jakobson, Fant and Halle (1952), Ladefoged (1964), Perkell (1969), Malecot (1970), Han and Weitzman (1970), Hardcastle (1973) have all claimed that tensity is an important, independent differentiating feature. Lisker and Abramson (1964) have, on the other hand, questioned the independent
status of the tensity feature and have argued that, for stop articulation at least, tensity is a redundant feature because "no one of the physical measures, whether physiological or acoustic, that have been proposed as correlates of the fortis/lenis dimension has been shown NOT to be significantly connected with voicing or aspiration."

The search for the acoustic correlates of the fortis/lenis dimension is, in fact, often hampered due to the fact that investigators have not been particularly explicit in their treatment of force of articulation. Heffner (1964), as mentioned previously, seems to link wrongly, tensity with aspiration. For him, the degree of energy with which a stop is uttered seems to be proportional to the duration of its aspiration. And, as the following quotation makes clear, the acoustic correlates of tensity during the release phase are to be looked for in the aspiration rather than in the explosion of the stop:

"The precise nature of the movements made to release a stop consonant depends also (a) on the positions of the organs during the hold, and (b) the positions to which they next move. But there is, in the case of the release, the further factor of the dynamic nature of these movements. If they are impulsive, or sudden, the rush of the air out of (pressure stops), or into (suction stops) the stopped cavity may be vigorous and puff-like. In the case of the pressure stops this puff may readily become the distinctive mark of the stop. Stops which have it are called aspirated, stops which lack it are called unaspirated stops. Evidently the degree of energy displayed by these puffs depends upon the degree of compression or rarefaction achieved during the occlusion. On the basis of such differences in energy we describe some stops as strong or fortis, others as weak or lenis. When necessary, strongly aspirated stops may be written by adding h to the symbol, thus ph, th, while weakly aspirate stops may be written p', t', k' ."
Jakobson, Fant and Halle (1952), on the other hand, state clearly that the search for the acoustic cues distinguishing between fortis and lenis stops should concentrate on the explosion and not on the aspiration of the stop since "In consonants, tenseness is manifested primarily by the length of their sounding period, and in stops, in addition, by the greater strength of the explosion." Likewise, Lisker and Abramson (1964) indicate that listeners' judgment of articulatory force "depends directly on the audible features of closure duration and the loudness of the stop explosion."

Pike (1943) also agrees that "Fortis movement of an initiator tends to make relatively loud sounds, and brings acoustic judgments to bear on the fortis nature of sounds." Pike, however, makes the important observation that although fortis articulation results in relatively loud sounds, relative loudness is not necessarily the result of fortis articulation. "Thus," he writes, "the hissing of English s is loud, but not fortis, in comparison with z; the lingual strictures seem to be the same (hence neither fortis nor lenis), but the noise of the hissing of the s is louder because it has a stronger air stream at the lingual stricture (the partial glottal closure for voicing reduces the strength of the air stream for z)."

Pike's remark that it is not justified to speak in terms of tenseness and laxness when differences in the spectra of a homorganic pair of consonants are due to the presence or absence of vocal fold vibrations contradicts the statements of other writers (Halecot, 1966; Westell, 1969) who observe that lenis consonants are lenis precisely because of the reduction in the rate of air flow from below the glottis, which is the result of the less frequent puffs of air.
It is true that Alecot (1966) indicates that actual voicing is a dispensable differentiating feature. Yet his statement that "glottal adduction for the voiced consonants and abduction for their voiceless cognates constitute the fundamental class difference" seems to justify Lisker and Abramson's observation that "the ensemble of acoustic features that are used as evidence for a dimension of articulatory form may be plausibly grouped together without any need for positing an independent dimension," and may instead be associated with "differences in the position and activity of the glottis during the various phases of stop production". (1964).

It was precisely this conclusion which led Lisker and Abramson to combine voicing and aspiration into one parameter which they called voice-onset time (VOT), which refers to the temporal relation between the release of the stop occlusion and the onset of glottal pulsing. Lisker and Abramson (1971) went on to claim that VOT is "the single most effective measure whereby homorganic stop categories in languages generally may be distinguished physically and perceptually."

Hardcastle (1973), however, has pointed out that recent studies on the production and perception of stops have thrown some doubt as to the primacy of VOT as a category differentiating feature, and have given additional evidence to the presence of tensity as another important feature in the production and perception of stops.

Hardcastle points out that VOT cannot be the major discriminating one between stop categories for two reasons. First, VOT values characteristic of different Korean stops have been reported to overlap by
Ki (1965) and by Han and Weitzman (1970). Second, in an experiment by Han and Weitzman (1970) where subjects were asked to identify stimuli which were made by cutting off portions of the aspirated part of a heavily aspirated tense stop, subjects did not perceive an unaspirated tense stop even when the VOT was reduced to less than 10 ms.

Hardcastle then states that "If VOT were the major discriminating cue between the three types of stops one would expect judgments to go in the direction of the unaspirated stops." He goes on to say that in view of these findings "It seems reasonable to speculate therefore that there is another differentiating feature, tense, present in the production of these stops that may also have some relevance as a perceptual cue."

Hardcastle's conclusion that tense has relevance as a perceptual cue since listeners would not confuse between heavily aspirated and unaspirated stops of Korean is impossible to justify. Han and Weitzman's experiments have shown that a reduction in VOT values turns the heavily aspirated 'tense' stops into slightly aspirated 'lax' stops. The fact that no matter how much further VOT values are reduced the heavily aspirated 'tense' stops would not turn into unaspirated 'tense' stops does not mean, however, that tense has any relevance as a perceptual cue. Both stops are, after all, tense.

B. Duration and intensity of stop explosion.

In our discussion of the tense/lax distinction in stop consonants we have seen how most writers regard the loudness of the stop explosion
as a mark of tenseness. We have also seen that there are many good reasons to suspect whether tenseness is a really independent category differentiating feature. The investigation into the role of loudness as a perceptual cue in distinguishing between different categories of stop consonants, which is the primary concern of this section, does not, therefore, assume that loudness underlies a tense/len distinction. It merely attempts to look into the role it might play in speakers' identification of stop consonants.

The perceptual dimension of loudness depends primarily upon intensity. However, it is known from Bekesy's experiments (1960) that the sensation of loudness is not independent of the duration of the sound perceived. With the increase of duration the loudness of the sound perceived rises.

We should note, furthermore, that the relation between loudness and intensity is further complicated by the fact that the sensitivity of the ear to differences in intensity differs greatly in the different frequency regions. The ear is most sensitive to frequencies between 1000 and 6000 Hz. Within this range we can perceive tones which have relatively little intensity. Tones of lower and higher frequencies, however, need to be of much greater intensity to be perceived. This relation between frequency and intensity does not concern us here since there are no frequency differences between a pair of homorganic stops.

In the examination of the duration of stop explosion, we were looking for confirmation of two basic assumptions:
1. First, that the duration of the explosion decreases in the order velar > dental > labial.

2. Second, that the duration of the explosion of voiceless stops is longer than that of their voiced counterparts.

First, differences in the duration of the explosion in relation to the place of articulation are expected for two reasons. First, we can speculate that "the bulk of the tongue cannot move as quickly as the tip or the lips" (Jorgensen, 1954). Second, these differences are again expected due to the fact that the volume of the supraglottal cavity behind the point of constriction for the velar stop is smaller than behind the constriction for the dental or bilabial stops, so the pressure there at the beginning of the release phase will be relatively greater. This means that, other things being equal, the time taken to achieve the pressure stabilisation in the supraglottal cavity would be longer in the case of the velar stop than the dental or bilabial ones.

Second, the expected difference between the duration of the explosion of voiceless stops and that of their voiced counterparts rests on the fact that the supraglottal pressure behind the point of constriction is greater in the case of the voiceless stops. This could be the result either of smaller cavity volume, or absence of resistance to air flow at the glottis. Once again, the time taken to achieve pressure stabilisation in the supraglottal cavity would be longer in the case of a voiceless stop than in the case of a voiced one.

Although the task of measuring the duration of the stop explosion was made relatively easy due to the fact that it was limited to the
examination of stops in initial position only, we have met various problems. First, it is not uncommon to find stops for which acoustic evidence of their explosion cannot be detected with any certainty. Second, it was not possible to measure the duration of the explosion consistently at one single frequency. This was the result of the fact that formants of following vowels in the case of the voiced stops, and concentrations of energy during the aspiration phase in the case of the voiceless stops, have various frequency regions. Taking into consideration the fact that it is not possible to detect any stop explosion in the vicinity of these regions of energy concentration, the measurement of the duration of the explosion has to be applied to a different frequency position in the case of different environments.

A preliminary analysis of stop explosions revealed that for all stop consonants the explosion lasts normally for the duration of a single pulse ('striation') at the onset of a periodic waveform, that is, approximately 5 ms, except in the case of the pharyngealized stops /\a/ and /\i/ where it usually lasts slightly less than 10 ms.

The exceedingly short duration characteristic of stop explosions did not allow us, however, to quantify more precisely the duration values characteristic of different categories of stop consonants. Neither was it possible to find the expected differences between the stop explosions of voiced and voiceless stops, or those between different places of articulation.
It was, therefore, decided to use a different technique of analysis which could enable us to measure more precisely the duration of the stop explosion. According to this technique stops were recorded at a speed of 15 in/s and played back at 7½ in/s.

Measurement of duration of stop explosions was undoubtedly made quite easy through this technique, since the reduced speed made the duration appear twice as long as the real duration. It focussed attention, however, on a problem of a fundamental sort. That is, what is the acoustic correlate of a stop burst?

We may decide that the acoustic correlate of a stop burst is that period characterized by high intensity spread among the frequencies. Alternatively, we may decide that a stop burst appears on spectrograms in the form of a single spike of high intensity.

Apart from the fact that the latter decision is an arbitrary one, it is also highly unjustifiable for there is no reason why we should limit the explosion to only one noise spike while spectrograms (especially of /k/) show regularly more than one. Furthermore, its adoption in the measurement of explosion duration is useless, since the duration of the noise spike is the same for all stops - about 4 ms.

It was in view of these facts that we decided to regard the period, rather than the spike, of high intensity spread among the frequencies as the acoustic correlate of stop burst. The adoption of this procedure did not, however, put an end to our problems. In fact, it gave rise to some of its own. For instance, in terms of such
an analysis it turns out that there are no duration differences between the explosions of the voiced stops and their voiceless counterparts, or between the pharyngealized and non-pharyngealized members of a homorganic pair of stop consonants. In fact, all stops turn out to have an explosion with an average duration of about 5 ms, with the exception of /k/ whose explosion has a duration of about 10 ms on average.

These results are, to say the least, unexpected. One advantage that came out of playing back the stop recordings at a lower speed was, however, that this brought out intensity differences between the explosions of voiced stops and those of their voiceless counterparts. Explosions of the voiceless stop consonants /t ɾ/ take the form of a column having a duration of about 5 ms. Explosions of the voiced stop consonants /d ɾ/, on the other hand, appear in the form of two spikes (each lasting about 1 ms) separated by a period of little or no energy that has a duration of about 3 ms. In the case of the velar stops, however, there is no difference between the explosion spectra of a voiced stop and a voiceless one. In each case, /k/ or /ɾ/ appears in the form of three spikes or clicks.

Spectrograms of Arabic stops do, indeed, consistently show that the explosion of a voiceless stop is more intense than that of a voiced one. This is undoubtedly due to the greater pressure build-up behind the point of constriction in the case of a voiceless stop. Exact measurement of the difference was not, however, possible.
To sum up, the investigation into the acoustic characteristics of stop explosions in Arabic has failed to reveal the expected temporal differences in stop explosions due to different modes of articulation. Differences in relative intensity, however, were observed; the voiceless stops being characterized by explosions of higher intensity than their voiced counterparts. Taking into account the fact that stop explosions do not normally last more than 10 ms, we may conclude that the role they play as perceptual cues in the voiced/voiceless distinction of Arabic stops is a minor one.

3.4.5. Relative duration of the closure phase.

Variations in the extent of closure duration—that is, the time interval between the termination of the vowel-formant transitions preceding the stop and the onset of the transition to the following vowel—have been shown (Leigh Lisker, 1957) to be relevant in the voiced/voiceless distinction within the stop consonants in English.

While it was recognized at the outset that this difference could not possibly turn out to be of major cue value in more than a restricted number of environments, it appeared none the less worth while to try to determine whether or not it has cue value in the case of Arabic stop consonants as well.

Inspection of spectrograms of Arabic stop consonants in post-stressed intervocalic position has revealed that closure durations for /d/ fall in the 70-75 ms range, with an average value of 70 ms, while values for /t/ vary from 55 to 105 ms, with an average value
of 85 ms. Closure durations for the velar stops, on the other hand, fall in the 45-90 ms range for /g/, with an average value of about 75 ms, while /k/ has values that vary from 70 to 85 ms, with an average value of about 80 ms.

These data thus indicate that the closure duration values of the voiceless stop consonants in Arabic are, on average, longer than those of their voiced counterparts. However, the observed overlap in the values of closure duration throws some doubt as to the primacy of this cue in the voiced/voiceless distinction in Arabic stop consonants. It is obvious, none the less, that the perceptual value of this cue can only be determined by testing listeners' responses to differences in closure durations of the sort reported.

3.4.6. Extent of F1-transition.

In the early days of speech analysis, the interpretation of consonant-vowel transitions was, as Cooper, Delattre, Liberman and Gertsman (1952) put it, "a major problem." It was not clear at the time whether these transitions were, as the name implies, no more than the necessary transitions between the sounds that serve to identify successive phonemes and, as such, mere nulls which dilute the acoustic message, or whether they serve as a principal acoustic cue for the perception of the consonant-vowel combination.

In an attempt to explore the function of these transitions, Cooper, Delattre, Liberman and Gertsman (1952) prepared two series of synthetic spectrograms of transition-plus-vowel, then converted these
spectrograms into sound and played the recordings to a group of listeners. In the first series F1 had a considerable negative transition, while in the second it had a slight negative transition. The agreement among the listeners was, in general, sufficient to allow the investigators to report that "the transitions of the first formant appear to contribute to voicing of the stop consonants."

Later experimentation with synthetic speech was also soon to suggest a correlation between F1 negative transition and voicing of stop consonants. The experiments of Delattre, Liberman and Cooper (1955), for instance, showed that a strong voiced stop could be produced only by starting F1 very low on the frequency scale.

The voiced/voiceless distinction within the stop consonants thus began to be attributed to variations in the extent of F1-transition. Harris, Hoffman, Liberman, Delattre, and Cooper (1958) and Howard Foffman (1968), for instance, suggested that the primary function of F1 negative transition is to mark the class of voiced stops.

Such an assertion needs, however, to be qualified in two main aspects. The first concerns the fact that negative F1-transition is not the acoustic correlate of voicing. The second, on the other hand, concerns the perceptual importance of negative F1-transition relative to the delay in its onset time.

First, Fischer-Jørgensen's analysis (1954) of the acoustic properties of Danish stops has revealed that while negative F1-
transition is characteristic of Danish /b, d and g/. A negative F1-transition cannot be the marker for the class of voiced stops for the simple reason that Danish /b, d and g/ are as voiceless as their Danish counterparts /p, t and k/.

Fischer-Jorgensen then accounted for the negative F1-transition by pointing out, first, that the height of F1 is closely related to the degree of opening of the vocal tract at the lips (that is, vertical displacement of the jaw) and, second, that the distance between the explosion of /b, d and g/ and the beginning of the vowel is short. These two factors together result in lowering the frequency of F1 since the maximum opening is not reached until after the beginning of the following vowel. In the case of /p, t and k/, by contrast, there is sufficient time during the aspiration phase for the opening movement to be completed before the vowel.

Jorgensen's explanation of the physiological correlate of F1-transition thus reveals that what was taken to be the acoustical correlate of voicing is, in fact, the acoustical correlate of an entirely different articulatory feature, that is, aspiration.

Second, there are reasons to believe that the distinction has more to do with the starting time of F1 than with the starting frequency of F1. In an experiment with synthetic speech, Liberman, Delattre, and Cooper (1958) prepared synthetic spectrograms of the voiced stop consonants /b, d and g/. These were altered, first by removing the voice bar and then through successive cutbacks of F1. They were then converted into sound and presented to listeners for judgment as /b, d or g/ or /p, t or k/.

104
Listeners' responses indicated clearly that the F1 cutback effectively converted the voiced stops into voiceless ones. The investigators thus concluded that "the presence or absence of a 'cutback' in the first formant is a sufficient cue, and very likely an important one, in distinguishing voiced and voiceless stops in initial position."

It was also obvious, as the experimenters pointed out, that cutting back the beginning of F1-transition means, in actual fact, not only raising F1 starting point but also delaying "the time at which it begins relative to the other two formants." There was thus an obvious need to investigate the problem further in order to determine the role of each of these two variables in the effects described. The authors, therefore, carried out a number of experiments where they found that, apart from certain combinations of stop and back vowel, both the starting frequency of F1 and its onset time are important for the distinction. As they put it, "a rising first formant is a cue for the class of voiced stops, and a time delay in the first formant, without the rising transition, is a cue for the voiceless stops."

Later experiments with synthetic speech, however, suggest the overdominance of the onset time of F1 as a perceptual cue in the voiced/voiceless distinction. Stevens and Klatt (1974) have found that listeners differ in their response to differences in the spectra of synthetic stop consonant - plus - vowel syllables. "Some listeners," they write, "assign more weight to the absolute VOT, whereas others will not accept a stimulus with an appreciable transition at the onset of voicing as a voiceless consonant regardless of the VOT."
It is, therefore, significant that of the five listeners tested it was only for one that the boundary between voiced and voiceless "responses occurred at a fixed time relative to completion of the formant transitions, while for other listeners the boundary tended to be at a nearly fixed time relative to plosive release."

3.5. The acoustic analysis of the Arabic stops.

This section reports the results of the investigation into the spectral and temporal characteristics of the Arabic stops. The plan of study involved the analysis of each stop in both initial and final positions. In each position the effect of the stop on the second formant transitions of the adjacent vowel is investigated.

3.5.1. The effect of position on the spectra of the Arabic stops.

In reporting the results of the study of the effect of position on the spectra of the Arabic stops, this section will deal with the voiced stops separately from the voiceless ones. Such a division on the basis of the voiced/voiceless contrast is, indeed, a reflection of the fact that the effect of position on the spectra of the Arabic stops is dependent on whether, or not, these are voiced.

Since the glottal stop does not fit into a framework based on this distinction, a special sub-section has been allocated to deal with the effect of position on its spectrum. A further sub-section
has also been allocated to the treatment of two idiosyncratic properties of the velar/palatal pair /k, ɡ/.

3.5.1.1. The voiced stops.

Both in initial and final positions, the Arabic voiced stops appear on the spectrograms as gaps followed by spikes of high intensity noise. The presence or absence of aspiration and/or voicing in the spectrum of a voiced stop is, however, determined by its position in the word. In initial position, thus, the spike is immediately followed by vowel resonances without an intervening period of noise. In final position, by contrast, the spike is followed by a period of noise of considerable duration.

The distribution of acoustic energy during this period of noise following the spike in initial position is such that although it appears randomly diffused, it is, in fact, more concentrated, and intense, in those regions close to the formants of the preceding vowel.

The Arabic voiced stops in initial position are further characterized by a low-band component which appears along the baseline at about 200 Hz for /d/ and /d̪/, and 250 Hz for /b/ and /g/. In final position, by contrast, this component is often absent either partly or wholly.

3.5.1.2. The voiceless stops.

Unlike the Arabic voiced stops, the Arabic voiceless stops
appear on the spectrograms, in both initial and final positions as blank gaps followed by spikes of high intensity noise.

Unlike them again, the spectra of the Arabic voiceless stops are always characterized by a period of noise following the spike in both initial and final positions.

Similar to the voiced stops, however, the distribution of acoustic energy during the period of noise following the spike is such that although it appears randomly diffused, it is, in fact, more concentrated, and intense, in those regions close to the formants of the preceding (and, in this case, also the following) vowel.

Thus, by contrast with the Arabic voiced stops, the position of the Arabic voiceless stop in the word does not affect its spectrum.

3.5.1.3. The glottal stop.

A separate treatment of the glottal stop is warranted by the fact that within a framework based on the voiced/voiceless contrast the glottal stop is the odd one out since it is neither voiced nor voiceless. It is further justified by the fact that the pattern of behaviour of the glottal stop neither conforms to that of the voiced stops nor to that of the voiceless ones.

Like all the Arabic voiceless stops, the glottal stop appears on the spectrograms as a blank gap followed by a spike of high
intensity noise. Unlike them, however, its spike in initial position is followed immediately by vowel resonances without an intervening period of noise. In final position, the spike is followed, similar to all the other Arabic stops, by a period of noise.

Thus, similar to the voiced stops, the presence or absence of aspiration in the spectrum of the glottal stop is determined by its position in the word.

3.5.1.4. The velar/palatal pair /k, g/.

The homorganic pair of /k/ and /g/ have been singled out for separate treatment in this sub-section for two minor, though somewhat interesting, properties of their own.

Firstly, the spectra of /g/ differ from those for any other Arabic voiced stop in that they appear on the spectrograms, in both initial and final positions, as gaps followed by spikes of high intensity noise and then a period of noise.

Unlike the spectra of any other Arabic voiced stop, the spectra of /g/ alone are thus characterized by the presence of a period of noise following the spike in initial position. Thus, by contrast with the other Arabic voiced stops, the position of /g/ in the word has no effect on the presence or absence of its aspiration.
Secondly, and more importantly, the distribution of energy concentrations during the burst phase for both the /k/ and /g/ is, unlike that for any other stop, different in initial position from what it is in final position.

Mention has already been made of the fact that during the period of noise following the spike, acoustic energy is distributed in such a way that although it appears randomly diffused among all the frequencies, it is more concentrated close to the regions of the adjacent vowel formants. More particularly, the major concentrations of energy during this phase are located close to the frequency position of the second formant of the adjacent vowel.

In the case of /k/ and /g/, however, this is only true when they occur in initial position. In final position, by contrast, this close relationship between the frequency position of the major energy concentration in the burst phase and that of the second formant no longer holds.

For word-initial /k/, for instance, acoustic energy is mainly concentrated at about 2950 Hz for /i:/, 2650 Hz for /e:/, 2250 Hz for /a:/, 1100 Hz for /o:/, and 700 Hz for /u:/. For word-final /k/, by contrast, acoustic energy is concentrated around 2400 Hz for both /i:/ and /e:/, 1900 Hz for /a:/, and 1500 Hz for both /o:/ and /u:/.
3.5.2. The effect of the adjacent vowel on the spectra of the Arabic stops.

Apart from the fact that the frequency positions of energy concentrations during the aspiration phase are largely determined by the frequency positions of the formants of the adjacent vowel, the nature of the adjacent vowel does not seem to play an important part in determining the spectrum of an immediately adjacent stop.

In the case of /k/ and /g/, however, the nature of the adjacent vowel does effect a major change in their spectra. Thus, while the /k/ and /g/ are realized as palatal stops with loci around 2200 Hz in the vicinity of the front vowels, they are realized as velar stops with loci around 1000 Hz in the vicinity of the back vowels.

3.5.3. Duration.

This section deals with the duration of the closure phase and the aspiration phase of the Arabic stops in both initial and final positions. Data on the duration of the closure phase of the Arabic voiceless stops are not, however, included. This is due to the fact that it is not feasible to study by spectrographic means the duration of the closure phase of a voiceless stop in absolute initial position since there is no way of determining where it starts.

Figure 3.1 presents, therefore, the duration values only for the Arabic voiced stops in initial position. The top values represent maximum durations, the bottom values represent minimum durations, and
the dashes represent average durations.

As the figure shows, the range of variation is widest in the case of the bilabial stop, narrowest in the case of the velar/velaral stop, with the dental and post-dental stops occupying a middle range. Fig. 3.1 thus suggests that the farther back in the vocal tract a stop is produced, the narrower the range of variation of the duration of its closure phase.

Figure 3.2 presents the duration values for the closure phase of the Arabic stops in final position. The figure shows that the duration of an Arabic stop varies as a function of both voicing and point of articulation.

Thus, the closure phase of a voiced stop is, on average, greater than that of a voiceless one. The difference is not exclusive except in the case of the alveolar pair /t, d/. In the other cases, by contrast, there is a considerable degree of overlapping.

For both the voiced and voiceless stops, however, the farther back in the vocal tract they are produced, the greater the duration of their closure phase. This tendency is, as the figure makes clear, much more pronounced in the case of the voiceless stops than in the case of the voiced ones.

Figure 3.3 compares the duration values for the Arabic voiced stops in final position (solid lines) with those in initial position (broken lines). As the figure clearly shows, the duration values...
Figure 3.1. Duration of the closure phase (in milliseconds) for the Arabic voiced stops in initial position.
Figure 3.2. Duration of the closure phase (in milliseconds) for the voiced stops (solid and broken lines), voiceless stops (solid lines) and the glottal stop (broken line) in final position.
Figure 3.3. Duration of the closure phase (in milliseconds) for the Arabic voiced stops in initial position (broken lines) and in final position (solid lines)
Figure 3.4. The duration of the aspiration phase for the Arabic stops in initial position (broken lines) and in final position (solid lines) in milliseconds.
for the closure phase of the Arabic voiced stops in final position are always greater than those in initial position.

Figure 3.4 shows the average duration values for the aspiration phase of the Arabic stops in initial position (broken lines) and in final position (solid lines). Apart from /d/, the Arabic voiced stops in initial position are not aspirated at all. The voiceless stops, on the other hand, are characterized by the presence of a period of aspiration of considerable duration. As the figure also makes clear, the aspiration phase in initial position is considerably shorter than that in final position.

3.5.4. Transitions.

Figures 3.5 to 3.12 present, for each of the Arabic stops, the average steady-state, as well as the onset and offset, frequency positions of the Arabic long vowels after word-initial stops, as well as before word-final stops. Steady-state frequency positions are represented by squares while onset and offset frequency positions are represented by circles.

As the figures show, the direction of F2-transition varies with the quality of the vowel, the point of articulation of the stop, and the position of the stop in the word.

For example, while the Arabic stop /d/ has a negative transitional influence on the second formant of /i:/, it has a positive transitional
Figure 3.5. The average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /b/ (left-hand side) and before word-final /b/.
Figure 3.6. The average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /t/ (left-hand side) and before word-final /t/.

Figure 3.6
Figure 3.7. The average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /t/ (left-hand side) and before word-final /t/.
Figure 3.8. The average onset and offset frequency positions for the second formants of the Arabic logg vowels after word-initial /d/ (left-hand side) and before word-final /d/.
Figure 3.9. The average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /d/ (left-hand side) and before word-final /d/.
Figure 3.10. The average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /k/ (left-hand side) and before word-final /k/.
Figure 3.11. The Average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /g/ (left-hand side) and before word-final /g/.
Figure 3.12. The average onset and offset frequency positions for the second formants of the Arabic long vowels after word-initial /ʔ/ (left-hand side) and before word-final /ʔ/.
influence on that of /u:/. This, of course, is due to the fact that the
/d/-locus is lower than the steady-state frequency position of /i:/ and higher than that of /u:/.

As one would also expect, the F2-transition of /a:/ is negative after /b/, neutral after /d/, and positive after /g/. The variation is due to the differences between the steady-state frequency positions of the second formant of /a:/ and the frequency positions of the loci of /b/, /d/, and /g/.

The direction of F2-transition also varies with the position of the adjacent stop in the word. For example, when /g/ occurs in word-initial position it has, as already mentioned, a positive transitional influence on the second formant of /a:/.

When, however, it occurs in word-final position it has, by contrast, a neutral transitional influence on the second formant of the same vowel. Such a phenomenon could be accounted for in terms of a theory that posits that the coarticulatory effects of the Arabic stops are more powerful in a consonant-vowel context than in a vowel-consonant one.

The extent of transition of the second formant is similarly determined by the quality of the vowel, the frequency position of the stop locus, and the position of the stop in the word.

In general, the extent of transition of the second formant of an Arabic vowel is a simple direct function of the degree of closeness...
of that vowel. It is also greater on average when the vowel occurs after a word-initial stop than when it occurs before a word-final one. The relationship between the extent of transition and the frequency position of the stop locus is, however, more complex.

3.5.4.1. Direction of transition.

Table 3.1 presents in a condensed form the transition directions for the second formants of the Arabic long vowels when they are preceded by word-initial stops (the top row for each vowel) and when they are followed by word-final stops (the bottom row for each vowel). Plus signs (+) are used for positive transitions, minus signs (-) for negative transitions, and blank spaces for neutral transitions. The vowels are arranged on the left-hand side starting with the close front vowel through to the close back vowel. The stops are arranged on the bottom row in order of occurrence according to the point of articulation, starting with the bilabial stop and moving backwards in the vocal tract to the glottal stop.

In general, Table 3.1 presents a picture of what one would expect with negative transitions dominating at the top, positive transitions at the bottom, and neutral transitions in the middle. This, of course, is due to the fact that the steady-state frequency positions of the second formants of both /i:/ and /e:/ are higher than the frequency positions for the loci of most Arabic stops, while those for /o:/ and /u:/ are, by contrast, lower. /a:/, on the other hand, occupies a middle position in this set.
Looking at Table 3.1 column by column, one finds that /b/ seems to have a negative transitional influence on the front vowels, and a positive transitional influence on the back vowels. The homorganic alveolar pair /t, d/, on the other hand, exert a negative transitional influence on /i:/ and /e:/ only. Their transitional influence on /a:/ is basically neutral. This, no doubt, is due to the proximity of the steady-state frequency position of the second formant of /a:/ and the frequency position of the /t, d/-locus.

Table 3.1. The direction of transition of the second formant as a function of word-initial and word-final stops.

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>t</th>
<th>d</th>
<th>t</th>
<th>d</th>
<th>k</th>
<th>g</th>
<th>?</th>
</tr>
</thead>
<tbody>
<tr>
<td>i:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>e:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>a:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>o:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>u:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Looking at Table 3.1 column by column, one finds that /b/ seems to have a negative transitional influence on the front vowels, and a positive transitional influence on the back vowels. The homorganic alveolar pair /t, d/, on the other hand, exert a negative transitional influence on /i:/ and /e:/ only. Their transitional influence on /a:/ is basically neutral. This, no doubt, is due to the proximity of the steady-state frequency position of the second formant of /a:/ and the frequency position of the /t, d/-locus.
Table 3.1 also shows that the pattern of transitional influence exhibited by the pharyngealized post-alveolar homorganic pair /t, d/ is very similar to that exhibited by the alveolar pair /t, d/, especially in word-initial position where they are, indeed, identical. In word-final position, however, the alveolar pair consistently exerts a positive transitional influence on the back vowels /o:/ and /u:/ while the post-alveolar pair exerts, again consistently, a neutral transitional influence.

The pattern of transitional influence exhibited by the palatal/velar homorganic pair /k, g/ is different from that for any other Arabic stop. It is only in the case of this pair that we find instances of neutral transitional influence on the second formant of /i:/ and of positive transitional influence on that of /e:/ In all other cases the F2-transition direction of these two vowels is negative. The different directions this time are obviously due to the high frequency position of the palatal pair /k, g/. The transitional influence of the velar /k, g/ pair is, on the other hand, mainly neutral although word-initial /g/ has a positive transitional influence on both /o:/ and /u:/.

The pattern of transitional influence exerted by the glottal stop presents problems in more than one way. Apart from the fact that one would not expect the glottal stop to have a transitional influence of any sort, it is surprising to find that such an influence as it may have is more powerful in word-final position and not, as is the case for the other Arabic stops, in word-initial position.
In fact, for all the Arabic stops, with the exception of the glottal stop, the direction of F2-transition tends to be neutral more when they occur in word-final position than when they occur in word-initial position. This dominance of neutral transitions in word-final position suggests that the coarticulatory effects of the Arabic stops are more powerful in a consonant-vowel context than in a vowel-consonant one.

3.5.4.2. Extent of transition.

In investigating the effect of the Arabic stop consonants on the extent of transition of the second formants of the Arabic long vowels, the average onset, or offset, frequency position for each vowel was subtracted from its steady-state frequency position following, or preceding, a particular stop.

To arrive at a measure of the extent of transition of the second formant for a particular vowel following, or preceding, all the Arabic stops, these values were squared, added up, and their sum divided by the number of the Arabic stops. Finally, the square root of this quotient was calculated and plotted in Figures 3.13 and 3.14.

As Figure 3.13 shows, the closer a vowel is the greater the extent of transition of its second formant. Thus, the extent of transition for the close front vowel /i:/ is greater than that for the half-close
Figure 3.13. Extent of F2-transition after word-initial stops (solid lines) and before word-final stops (broken lines).
front vowel /eː/. Similarly, the extent of transition for the second formant of the close back vowel /uː/ is greater than that for the half-close back vowel /oː/. The extent of transition for the second formant of the open vowel /æː/ is, predictably, the smallest.

Figure 3.13 also shows that although the extent of transition of the second formants for /iː/ and /eː/ is considerably greater when they precede word-final stops than when they follow word-initial ones, the extent of transition of the second formants of the Arabic long vowels is, on average, greater when they follow word-initial stops than when they precede word-final stops. This strengthens the suggestion that the coarticulatory effects of the Arabic stops are more powerful in a consonant-vowel context than in a vowel-consonant one.

Figure 3.14 presents the extent of transition as a function of the preceding or following stop. Apart from /b/, all the other stops seem to affect a greater displacement of the transitions of the second formant of adjacent vowels when they occur in word-initial position than when they occur in word-final position.

As Figure 3.14 also shows, the pharyngealized post-alveolar pair /tˤ, dˤ/ seem to affect the extent of transition most, followed by the non-pharyngealized alveolar pair /t, d/, the bilabial stop /b/, the glottal stop /ʔ/, and last the velar/palatal pair /k, มงคล/.

132
Figure 3.14. Extent of F2-transition as a function of word-initial stops (solid lines) and word-final stops (broken lines).
The observed small influence of the velar/palatal pair /k, g/ on the extent of transition of the second formants of the adjacent vowels is, no doubt, due to the fact that it has, unlike any other Arabic stop, two loci. This results in a situation where the locus of the stop is always nearer to the steady-state frequency position of the second formant than would otherwise have been. The reasons for the remarkable influence of the pharyngealized stops are not, however, clear to me.

3.6. Conclusion.

This section attempts to provide an overall view of the specifics of the Arabic stops treated separately in the previous sections.

3.6.1. Coarticulation and word position.

The effects of coarticulation and word position determine both the spectral and temporal properties of the Arabic stops as well as the spectral properties of the adjacent vowel.

In this section we shall first deal with the variation of the spectral characteristics of the stops as a function of word position and the adjacent vowel. We shall then discuss how the vowel spectrum is itself affected by the adjacent stop and word position. Lastly, we shall examine the factors that determine the temporal properties of the Arabic stops.
3.6.1.1. The effect of word position on stop spectra.

The acoustic analysis of the Arabic stops has revealed that their spectral characteristics are significantly determined by their position in the word. The effect of word position on the voiceless stops is, however, different from that on the voiced ones. While the position of the Arabic voiceless stops in the word does not affect their spectra, the presence or absence of aspiration and/or voicing in the spectrum of a voiced stop is determined by its position in the word.

The spectrum of a voiced stop in initial position is characterized by the presence of a low-band component which appears along the baseline at or below 250 Hz, and the absence of an intervening period of noise between the termination of the stop closure and the onset of vowel resonance. In final position, by contrast, this low-band component is often absent either partly or wholly, and the stop closure is followed by a period of noise of considerable duration.

The glottal stop shows, however, a pattern of behaviour that conforms neither to that of the voiced stops nor to that of the voiceless ones. Similar to the voiceless stops, its spectrum shows the absence of energy during the closure phase. Unlike them, however, the presence or absence of aspiration in its spectrum is determined by its position in the word.
3.6.1.2. The effect of the adjacent vowel on stop spectra.

Apart from the fact that the frequency positions of energy concentrations during the aspiration phase are largely determined by the frequency positions of the formants of the adjacent vowel, the quality of the adjacent vowel does not seem to play an important part in determining the spectrum of the immediately adjacent stop.

In the case of /k/ and /g/, however, the quality of the adjacent vowel does effect a major change in their spectra. Thus, while the /k/ and /g/ are realized as palatal stops with loci around 2200 Hz in the vicinity of the front vowels, they are realized as velar stops with loci around 1000 Hz in the vicinity of the back vowels.

3.6.1.3. The effect of the adjacent stop and word position on the vowel spectrum.

The spectral characteristics of the vowel are similarly affected by the nature of the adjacent stop, at least as far as the direction and extent of F2-transition are concerned.

Both the direction and extent of F2-transition are mainly the product of the interaction between the frequency position of the stop locus and that of the steady-state second formant of the adjacent
vowel. Where the former occupies a higher frequency position than the latter, the direction of F2-transition is positive and its extent is relatively greater. Conversely, where it occupies a lower frequency position the opposite is true.

Both the direction and extent of F2-transition are significantly influenced by the position of the stop in the word. First, for all the Arabic stops, with the exception of the glottal stop, the direction of F2-transition tends to be neutral more when they occur in word-final position than when they occur in word-initial position. And, second, apart from /b/, all the Arabic stops seem to effect a greater displacement of the transitions of the second formants of adjacent vowels when they occur in word-initial position than when they occur in word-final position.

Both these two facts suggest that the coarticulatory effects of the Arabic stops are more powerful in a consonant-vowel context than in a vowel-consonant one.

3.6.1.4. Factors affecting stop temporal properties.

The duration of an Arabic stop varies as a function of voicing, word position, and point of articulation. First, the position of the Arabic stop in the word affects the duration of both the closure
phase as well as the aspiration phase. In the case of the voiced stops, the duration of the closure phase in final position is always greater than in initial position. Similarly, in the case of the voiceless stops the duration of the aspiration phase in final position is considerably greater than in initial position.

The effect of point of articulation on stop duration manifests itself in two ways. First, the farther back in the vocal tract a stop is produced the narrower the range of variation in the duration of its closure phase. And, second, the farther back in the vocal tract a stop is produced the greater the duration of its closure phase. This last phenomenon is, however, more pronounced in the case of the voiceless stops than in the case of the voiced ones.
Chapter 4

The trill

4.1. Articulatory mechanism.
4.2. The acoustic theory of trill production.
4.3. Acoustic analysis.
4.4. Perception.
4.5. Acoustic analysis of the Arabic r-sound /r/
  4.5.1. The non-pharyngealized [r].
  4.5.1.1. The effect of position on the spectrum of [r].
  4.5.1.2. The effect of the adjacent vowel on [r].
  4.5.1.3. Duration.
  4.5.1.4. Transitions.
  4.5.2. The pharyngealized [ɹ].
  4.5.2.1. The effect of position on the spectrum of [ɹ].
  4.5.2.2. The effect of the adjacent vowel on [ɹ].
  4.5.2.3. Duration.
  4.5.2.4. Transitions.
4.6. Conclusion.
4.1. Articulatory mechanism.

A trill is a type of consonant segment produced by the interposition of a vibratile organ in the passage of the air stream. This vibratile organ is held loosely against another articulator so that the flow of air between them sets one or both in motion, alternately sucking them together and blowing them apart.

The speed with which a trill is made demands the participation of a particularly elastic organ. This in turn limits the number of vibratile organs to three: the uvula, the apex, and the lips. The uvula can roll against the back of the tongue, the apex can roll against the teeth, the alveolum, or the post-alveolar ridge, and the lips can roll against each other.

Brosnahan and Malmberg (1970) report that the number of vibrations per second in trills varies somewhat for different speakers but is normally in the region of 16-20 per second. They also state that the number of vibrations in any particular occurrence of a trill varies from language to language, but is normally from two to four, rising to six or even seven when deliberately stressed.

In practice, however, trills are often shortened to one vibration. Brosnahan and Malmberg (1970) have called such single vibrations flaps. Ladegoged (1975), on the other hand, distinguishes further
between a single vibration "in which one articulator strikes another in passing while on its way back to its position of rest," and another which "is caused by a single contraction of the muscles so that one articulator is thrown against another." In Ladefoged's terminology only the first is termed a flap, while the second is termed a tap. In the following pages the terms 'tap' and 'flap' will be used in the senses defined by Ladefoged.

4.2. The acoustic theory of trill production.

Fant (1960) states that the acoustic theory of the production of Russian /r/, which we may take to be a trill, is simple "provided turbulent noise generated at narrow passages can be disregarded." Voiced /r/, in this way, "can be treated by the same acoustic theory as vowels since there are no shunting chambers introducing anti-resonances."

In the production of the non-pharyngealized /r/ the tongue position determines a back cavity configuration of a wide unobstructed pharynx, and a gradual narrowing of the mouth cavity towards the region of articulatory constriction. Fant (1960) points out that "The relations of the first three formants of the /r/-spectrum to the vocal tract configuration are essentially the same as for the /l/.

"Thus, F2 is the half-wavelength resonance of the anterior cavities, and F3 is to a substantial degree dependent on the anterior mouth cavity."

141
In the production of the pharyngealized /ɹ/, on the other hand, the tip of the tongue makes a post-alveolar contact while the back of the tongue approaches the upper part of the rear pharynx wall, thus dividing the cavities behind the point of articulation. "As a result F2 will be more dependent on the anterior of the two mouth cavities. F3 remains affiliated to the mouth cavity in front of the tongue." (Fant, 1960).

4.3. Acoustic analysis.

To start with, studies dealing with the acoustical properties of r-sounds are limited in number. The majority of these (Potter, Kopp and Kopp, 1947; Joos, 1948; Tarnoczy, 1948; Fant, 1960; Lehiste, 1964; and Dalston, 1975) are, furthermore, mainly concerned with one particular variety of r-sounds; that of the frictionless continuant. The rest, on the other hand, deal with the acoustical properties of trills or rolls in general rather than with the acoustical properties of any one particular r-sound (Broshahah and Malmberg, 1970; O'Connor, 1973). Evidence of the acoustic form of taps [ɻ] and fricatives [ɻ] is, therefore, not available.

O'Connor (1973) indicates that apart from the fact that the production of trills (lingual, uvular, and, may be, otherwise. K.S.) is characterized by "several rapid interruptions of the air stream, i.e. complete closures, compressions and releases made much faster than for plosives," their spectra have "the same acoustic features" as
for plosives. "In addition," he adds, "a clear vowel-like formant structure is visible between the short 'silent' gaps."

Brosnahan and Malmberg (1970) concede the presence of formants in the spectra of trills (both lingual and uvular). They state that "maximum energy comes in the range about 1000-2500 c/s for [R], and somewhat lower, about 500-2500 c/s for [r], with weak but frequently distinguishable formants at frequencies about 300, 1100, and 2100 c/s in [R], and again somewhat lower in [r]." They point out, however, that the spectra of these trills are further characterized by "an uneven, though not random distribution of energy over a wide band of frequency."

Brosnahan and Malmberg's statement about the presence of unevenly distributed energy over a wide band of frequency might seem to contradict O'Connor's statement about the presence of a clear vowel-like formant structure in the spectra of trills. We should remember, however, that both O'Connor and Brosnahan and Malmberg have not specified what trill they are describing and in what word-position. The results of the acoustic analysis of Arabic /r/, for instance, show that the spectrum of /r/ varies widely depending upon its position in the word.

Acoustic analysis of the American frictionless continuant [j] suggests that the position of F3 is a major cue in the identification of this sound. Potter, Kopp and Kopp (1947) have remarked that the spectrum of [j] is easily recognizable by the fact that "Bars 2 and 3 are in very close proximity, a little below the centre of the pattern."
Indeed, Joos (1948) has gone so far as to suggest that "every American r-coloured sound ... is characterized by a third resonance band placed a little higher than formant 2."

The spectrum of [ʃ] is, furthermore, influenced by its position in the word. Potter, Kopp, and Kopp (1947) have noted that the formants of [ʃ] are higher in final position than in initial position. This is most obvious in the case of F1 which is usually on the baseline in initial position, and off the baseline in final position.

Potter, Kopp, and Kopp (1947) have also noted that due to the low frequency position of F1 of [ʃ] it will always have a negative transitional influence on F1 of adjacent vowels. The central position of F2 of [ʃ], by contrast, could result in either a negative or positive influence depending on the frequency position of F2 of the adjacent vowel. F3 of [ʃ], similar to F1, is always lower than F3 of the adjacent vowel and thus will always have a negative transitional influence.

4.4. Perception.

Initial attempts to produce two-formant synthetic spectra recognizable as [ʃ] vowel met with only partial success (O'Connor, Gertsman, Liberman, Delattre, and Cooper, 1957). Listeners were able to perceive [ʃ] only where the following vowel was a front unrounded one. That is, a stimulus was recognized as [ʃ] vowel only where F2-transition was a negative one.
The experimenters, therefore, concluded, first, that although F2-transition was a necessary cue for the perception of [t], it was not a sufficient cue, and, second, that recognition was a function of F2-transition direction rather than of the frequency position of F2-onset.

In order to produce a stimulus that could be recognized as [t] before all vowels, the experimenters found it necessary to synthesize three-formant spectra with a negative F3-transition.

The fact that O'Connor et al found it necessary to include a negative F3 in order to synthesize [t] before both front and back vowels is not, indeed, surprising. Potter, Kopp and Kopp (1947) have already indicated that [t] always has a negative transitional influence on the F3 of the adjacent vowel. Joos (1948) has stated that a low F3 is a major cue for r-colouring. More recent work on the subject has also confirmed that a negative F3 "appears to be a strong cue for /r/" (Ainsworth, 1968).

4.5. Acoustic analysis of the Arabic r-sound /r/.

The Arabic r-sound /r/ can be phonetically realized either as a frictionless continuant, a tap, or a fricative. Each one of these can, in turn, occur either as pharyngealized or non-pharyngealized. The phonetic realizations of the Arabic r-sound /r/ may thus be grouped into two main sets: a non-pharyngealized set, for which the symbol [r] will be used, and a pharyngealized set, for which the symbol
[r] will be used. The three members of the first set will have the symbols [J], [I], and [\/], while those of the second set will have [J], [I], and [\]. The following diagram illustrates this point.

This section reports the results of the investigation into the spectral and temporal characteristics of the Arabic r-sound. The plan of study involved the analysis of both the pharyngealized [r] and the non-pharyngealized [r] in three different word positions: initial, intervocalic, and final. In each position, the effect of the Arabic long vowels on the r-sound, as well as the effect of r-sound itself on the adjacent vowel, are investigated.
4.5.1. The non- pharyngealized \( r \).

4.5.1.1. The effect of position on the spectrum of \( r \).

The position of \( r \) in the word has a considerable effect on its spectrum. In initial position, acoustic energy is concentrated in well-defined formants only at the lower end of the spectrum. In 80 per cent of \( r \) occurrences \( F_3 \) was absent and acoustic energy above \( F_2 \) was unevenly distributed, though vaguely anticipating the formants of the following vowel.

The average steady-state position of \( F_1 \) was 305 Hz, while that of \( F_2 \) was 1310 Hz. \( F_3 \), in those cases in which it could be detected, had an average steady-state frequency position of about 2400 Hz.

The spectrum of \( r \) in this position is often characterized by the presence of a gap with an average duration of 20 ms immediately before the beginning of the vowel. This gap is, presumably, the acoustic correlate to the closure phase made at the end of the [f] articulation. In the absence of a gap in the spectrum, [r] appears on the spectrograms as a frictionless continuant [w].

In intervocalic position the spectrum of [r] is much the same as that of a stop. [r] appears on the spectrograms as a gap with no energy above the voice bar, save perhaps a shadow of the formants of the adjacent vowel. [f] has an average duration of about 25 ms.
In final position, the spectrum of \( [r] \) looks virtually the same as that of a voiceless fricative \( [\mathbf{\mathfrak{u}}] \). In roughly one half of the occurrences of \( [\mathbf{\mathfrak{u}}] \), acoustic energy is diffusely spread in the frequency range 2700-5000 Hz. In the other half of the occurrences, although acoustic energy is still diffusely spread among the frequencies, F1 and F2 could be detected in spite of their low intensity. The average steady-state frequency position of F1 was 250 Hz, while that of F2 was 1420 Hz. In common with \( [\mathbf{\mathfrak{u}}] \), the spectra of \( [\mathbf{\mathfrak{u}}] \) are also characterized by the presence of a silent gap immediately after the end of the vowel.

4.5.1.2. The effect of the adjacent vowel on \( [r] \).

Figure 4.1 presents the formant transitions of \( [r] \) as a function of the following vowel. The lines connect points representing average values measured for 60 occurrences of \( [r] \) in initial, intervocalic, and final positions. As may be seen from the figure, although the range of F1 variation is small, compared with that of F2, the values of the frequency positions of F1 show quite clearly that it varies directly with the first formant of the adjacent vowel. F2 and F3 also display a similar tendency to vary directly with the formants of the adjacent vowels. This time, however, the pattern of change is not as systematic as that displayed by F1.

4.5.1.3. Duration.

The position of \( [r] \) in the word has a considerable influence on its duration. \( [r] \) has an average duration of about 50 ms in initial position, 25 ms in intervocalic position, and 150 ms in final position.
Figure 4.1. Frequency positions for the first three formants in the spectra of r as a function of the adjacent vowel.
The effect of the nature of the adjacent vowel on the duration of \([r]\) is, by contrast, negligible.

4.5.1.4 Transitions.

Figure 4.2 presents the average steady-state as well as the onset and offset frequencies of the second formant of the five Arabic long vowels in the vicinity of \([r]\). In general, \([r]\) seems to have a negative transitional influence on the front vowels and a positive transitional influence on the back vowels, suggesting a locus value at about 1400 Hz.

As may be seen from figure 4.2, the extent of transition is greater when the \([r]\) occurs in initial position than when it occurs in final position. This suggests that the effect of \([r]\) on the second formant of an adjacent vowel is greater in a consonant-vowel context than in a vowel-consonant one.

4.5.2 The pharyngealized \([r]\).

4.5.2.1 The effect of position on the spectrum of \([r]\).

The position of \([r]\) in the word has a considerable effect on its spectrum. In initial position it is realized either as a frictionless continuant \([\mathcal{J}]\) or a tap \([r]\), while in intervocalic position it is always a tap. In final position, by contrast, it occurs as a fricative \([\mathcal{J}]\).
Figure 4.2. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial [r] (left-hand side), and before word-final [r].
In initial position, acoustic energy is concentrated in well-defined formants at the lower end of the spectrum. Above 1500 Hz, however, there is very little energy with the result that for the majority of the cases investigated only F1 and P2 are present.

In common with other pharyngealized sounds, the frequency position of F1 is relatively high, while that of F2 is relatively low. Thus, the average steady-state frequency position of F1 is 335 Hz, while that of F2 is 1175 Hz.

Similar to [r], the spectrum of [z] in this position is often characterized by the presence of a gap with an average duration of about 25 ms immediately before the beginning of the vowel. This gap is, presumably, the acoustic correlate of the closure phase made at the end of the [f] articulation.

In intervocalic position the spectrum of [z] is much the same as that of a stop. [z] appears on the spectrograms as a gap with little or no energy above the voice bar. In fact, the only difference between the spectrum of [z] in this position and the spectrum of a stop is the briefness of the gap in the case of [z]. This is consistent with a remark made by Brosnahan and Malmberg (1970) that the dividing line between a flap (in our terms, a tap) and a stop "can usually be drawn on the basis of the briefness or incompleteness of the flap articulation in comparison with a plosive articulation."

In final position, by contrast, the spectrum of [z] is not much different from, and very often identical with, that of a fricative.
Acoustic energy is concentrated in the frequency range 2700-2500 Hz with little energy above or below that range.

The voicing of [y] deserves special mention because it seems to depend to some extent on the nature of the preceding vowel. With the back vowels, devoicing starts immediately after the vowel. In the case of the front vowels, by contrast, devoicing starts about 70 ms after the vowel. Furthermore, during the voiced part, acoustic energy is concentrated in well-defined formants of low intensity. The average steady-state frequency position of F1 is 315 Hz while that of F2 is 1375 Hz. This suggests, as has been pointed to me by R.A.W. Bladon, that front vowels have a voiced diphthongal offglide onto [y].

4.5.2.2. The effect of the adjacent vowel on [u]

Figure 4.3 presents the formant transitions of [u] as a function of the following vowel. The lines connect points representing average values measured for 40 occurrences of [u] in initial and final positions. As may be seen from the figure, both F1 and F2 appear to vary directly with the first and second formant of the adjacent vowel.

4.5.2.3. Duration.

The position of [u] in the word has a considerable influence on its duration. [u] has an average duration of about 60 ms in initial position, 25 ms in intervocalic position, and 180 ms in final position. The effect of the nature of the adjacent vowel on the duration of [u] is, by contrast, negligible.
Figure 4.3. The average frequency positions for the first three formants in the spectra of [x] as a function of the adjacent vowel.
4.5.2.4. Transitions.

Figure 4.4 presents the average steady-state as well as the onset and offset frequency positions for the second formants of the five Arabic long vowels in the vicinity of \([\text{x}]\). In general, \([\text{x}]\) seems to have a negative transitional influence on the front vowels and a neutral transitional influence on the back vowels.

As may be seen from the figure, the extent of transition is greater when \([\text{x}]\) occurs in initial position than when it occurs in final position. This suggests, once again, that the effect of \([\text{x}]\) on the second formant of an adjacent vowel is greater in a consonant-vowel context than in a vowel-consonant one.

4.6. Conclusion.

The results of the acoustic analysis of the Arabic \(r\)-sound reported in the preceding sections indicate the following four basic findings.

Firstly, the position of /\(r\)/ in the word has a remarkable effect on its spectrum. The spectrum of /\(r\)/ in initial position is vowel-like with clear and well-defined formants. In intervocalic position, /\(r\)/ appears on the spectrograms as a brief stop with little acoustic energy above the voice bar. In final position, by contrast, /\(r\)/ has a spectrum of an undoubtedly fricative nature.
Figure 4.4. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial [x] (left-hand side), and before word-final [x].
Secondly, the adjacent vowel has a considerable effect on the steady-state part of /r/. In general, the formants of /r/ show a clear tendency to vary directly with the formants of the adjacent vowel.

Thirdly, the position of /r/ in the word also affects its duration. For both the pharyngealized and the non-pharyngealized phones, /r/ is shortest intervocalic position, longest in final position, and with an intermediate duration in initial position.

Fourthly, while both the pharyngealized and non-pharyngealized phones have a negative transitional influence on the front vowels, the pharyngealized [x] exerts a neutral transitional influence on the back vowels, while the non-pharyngealized [r] exerts a positive transitional influence.

The results of the acoustic analysis of the spectra of Arabic /r/ reveal, as expected, that they have much in common with the reported properties of r-sounds in general.

The spectra of Arabic /r/ bear out the observation made by Brosnahan and Malmberg (1970) that the trills are acoustically
characterized by "an uneven, though not random distribution of energy over a wide band of frequency." Both in initial and final positions, spectrograms of Arabic /r/ do, indeed, show this uneven distribution of energy at frequencies above 2500 Hz. In final position, however, the intensity of acoustic energy is far higher than it is in initial position.

Again, spectrograms of Arabic /r/ in intervocalic position illustrate very clearly Brosnahan and Malmberg's point (1970) that the distinction between a tap and a stop can sometimes be made "on the basis of the briefness or incompleteness of a flap articulation in comparison with a plosive articulation." The spectra of Arabic /r/ in intervocalic position are virtually the same as those of a cognate plosive, and it is only in terms of the briefness of closure duration that a distinction can, indeed, be made.

The results of the acoustic analysis of Arabic /r/ also support the claim made by Potter, Kopp and Kopp (1947) that the position of the American r-sound in the word affects the frequency positions of its formants. Potter, Kopp and Kopp report that the formants of the American r-sound are higher in final position than in initial position. Our findings, however, are more complex. F2 of Arabic /r/ is higher in final position than in initial position. In the case of F1, by contrast, the opposite is true.

158
The results of the analysis of Arabic /r/ differ from those reported for other r-sounds in two main points. The first relates to the strongly emphasized importance of the low position of the third formant of the steady-state part, while the second concerns the transitional influence of the first formant.

Potter, Kopp and Kopp (1947) report that the American r-sound always has a negative transitional influence on the third formant of the adjacent vowel. They indicate that "Bar 3 of the vowels never reaches as low a position as bar 3 of r, so that its curve from the initial position is always upwards." Joos (1948), on the other hand, states that "every American r-coloured sound ... is characterized by a third resonance band placed a little higher than F2."

In the case of Arabic /r/, however, the situation is different. In the first instance, Arabic /r/ exerts a negative transitional influence on the third formant of the adjacent vowel in only one half of the cases investigated, in the other half it exerts a neutral transitional influence. And, secondly, F3 is not placed a little higher than F2." In fact, it is widely separated from F2. In initial position, for instance, F2 at 1310 Hz, is closer to F1 at 305 Hz than it is to F3 at 2400 Hz.
The results of the acoustic analysis of Arabic /r/ are, again, in disagreement with the reported results as far as the transitional influence of F1 is concerned. Potter, Kopp and Kopp (1947) have indicated that due to the low frequency position of F1 of the American r-sound it will always have a negative transitional influence on the first formants of the adjacent vowels. The spectra of Arabic however, show quite clearly the first formant of /r/ exerting a positive transitional influence on the first formants of all the Arabic long vowels except /aː/. 
Chapter 5

The lateral

5.1. The articulatory mechanism.

5.2. Acoustic theory of lateral production.

5.2.1. A consonant or a vowel.

5.2.2. Cavity affiliations.

5.2.3. Formant continuity.

5.3. Perception.

5.4. The acoustic analysis of the Arabic lateral.

5.4.1. The non-pharyngealized lateral [1].

5.4.1.1. The effect of position on the spectrum of [1].

5.4.1.2. The effect of the adjacent vowel on [1].

5.4.1.3. Transitions.

5.4.1.4. Duration.

5.4.2. The Pharyngealized lateral [l].

5.4.2.1. The effect of position on the spectrum of [l].

5.4.2.2. Duration.

5.4.2.3. Transitions.

5.5. Conclusion.

5.5.1. Spectral differences between the pharyngealized and non-pharyngealized lateral.
5.5.2. The effect of word position on the spectrum of the lateral.

5.5.3. The effect of the adjacent vowel on lateral spectra.

5.5.4. The effect of the lateral and word position on vowel spectra.

5.5.5. Duration.
§.1. The articulatory mechanism.

"A lateral is a type of consonant segment produced by a stricture of complete closure in the centre of the vocal tract." (Abercrombie, 1967). This stricture of complete closure can be effected almost anywhere in the vocal tract from the lips to the uvula. It can also be accompanied by another, more open, stricture at some other point in the vocal tract.

The central blockage of the passage of the air stream means that the air cannot escape except laterally. This lateral passage of the air current can occur either on only one side or both sides of the obstruction. The resulting perceptual difference is, however, barely perceptible (Abercrombie, 1967; Malmberg, 1963).

The friction of the air current against the surfaces of the obstruction produces noise. As Brosnahan and Malmberg (1970) point out, "In the movement of air through any passage, provided a certain critical velocity is reached, the air particles close to the surface are set into turbulent eddying, whence result small variations of pressure which are transmitted through the surrounding air as sound waves." The amplitude of these pressure variations are directly proportional to the velocity of the air current. The velocity of the air current itself varies directly with the rate of air flow and inversely with the cross-sectional area of the constriction. This
means that the intensity of the noise characteristic of a lateral depends on the degree of approximation of the articulators and the rate of air flow.

The intensity of the friction caused by the lateral passage can thus vary considerably. At one extreme, the friction can be so strong that the lateral is perceived as being distinctly a fricative. At the other extreme, the intensity of the friction can be so low as to be imperceptible.

The Arabic lateral sound has two main variants: a post-dental pharyngealized lateral and a dental non-pharyngealized lateral. The production of the non-pharyngealized variant only involves raising the tip of the tongue to make a contact with the alveolar ridge. In the production of the pharyngealized lateral, on the other hand, as the tip of the tongue is making a post-alveolar contact, the front of the tongue is depressed, and the back of the tongue approaches the pharynx. The pharyngealized variant thus has a secondary place of articulation dividing the cavities behind the point of articulation. The non-pharyngealized variant, on the other hand, determines a back cavity configuration of a wide unobstructed pharynx and a gradual narrowing of the mouth cavity towards the region of articulatory constriction.

5.2. Acoustic theory of lateral production.

5.2.1. A consonant or a vowel?

In the acoustic specification of laterals it has been usual practice to regard a voiced lateral pronounced with the minimum of
friction as being very nearly a vowel. (Fant, 1960; Malmberg, 1963; Joos 1948; Ladefoged, 1975; Brosnahan and Malmberg, 1970; Potter et al, 1947).

There are four main reasons for regarding a voiced lateral as a vowel. First, the mechanism responsible for the production of a voiced lateral is different from that of most other consonants in that the source is a voice source rather than a noise source. Second, a lateral is often perceived as a vowel (Martin Joos, 1948; Brosnahan and Malmberg, 1970). Third, experiments with synthetic speech have shown that "a perfectly good l can be produced without a zero" (Fant, 1960). And, fourth, the acoustic energy of the lateral is concentrated in well-defined formants with the result that it has a "pattern that is not essentially different from a vowel pattern." (Joos, 1948).

In spite of the strong similarity between a lateral and a vowel as illustrated by the above paragraph, a lateral, even one with absolutely no friction, has to be regarded as being basically a consonant. This is due to the fact that in the production of a lateral, unlike vowels, the air stream is completely obstructed at the centre of the vocal tract. Again, by contrast with the spectrum of a vowel, a "lateral is always marked either at its beginning or at its end and often at both extremes, by a relatively abrupt shift of resonance pattern - a sudden change of spectrum." (Joos, 1948). And, as Joos points out, "It is this abruptness that characterizes them acoustically as consonants."

5.2.2. Cavity affiliations.

Attention has quite often been drawn to the similarity between
the acoustic spectra (Potter, Kopp and Kopp, 1947) and articulatory positions of [l] and [i], on the one hand, and [l] and [u], on the other, the similarity between the vocal tract configurations responsible for the production of these sounds has also been referred to. Fant (1960) states that while in the case of the non-pharyngealized lateral "the F1- and F2-cavity dependencies are the same as for the vowel i", in the case of the pharyngealized lateral "The dependency of F1 on the lateral constriction, and of F2 on the pharyngeal constriction, is definite and similar to the conditions for the production of u".

Taking into account the fact that Fant (1960) has already pointed out that the first formant of both [i] and [u] is determined by the back cavity rather than by the front cavity, one would expect, on grounds of the suggested similarity between the conditions for the production of [l] and [i], and [l] and [u], that the first formant of both laterals would also be determined more by the back cavity than by the front cavity. It should be observed, however, that by contrast with the conditions for the production of [i] and [l] where the first formant is "almost completely determined by the back cavity volume" (Fant, 1960), in the case of the pharyngealized lateral the front cavity has a not unimportant role in determining the first formant, similar to the role it plays in determining the first formant of [u]. The dependency of F1 on the lateral passage, which is part of the anterior mouth cavity, has already been referred to.

The similarity between the conditions for the production of [l] and [l], on the one hand, and [u] and [l], on the other, extends to the
characteristic cavity affiliations of their second formants. Thus, in the case of [1], the second formant cavity dependency is the same as for [i], both representing "a half-wavelength standing wave of the combined mouth-pharynx system behind the point of articulatory closure." (Fant, 1960). The second formant of [i], on the other hand, is "approximately equally dependent on the cavities in front of and behind the pharyngeal constriction," in much the same way as there is "an equal dependency of F2 on the two cavities for u" (Fant, 1960).

The suggestion that the first formant of a lateral is primarily determined by the back cavity means that we would expect an inverse relationship to hold between back cavity volume and F1 frequency position. This is, indeed, borne out by the fact that the first formant of the pharyngealized lateral has a higher frequency position than that of the non-pharyngealized lateral. The back cavity volume for [l] is much smaller than that for [1].

Furthermore, the low position of the second formant of the pharyngealized lateral could be accounted for in terms of "the secondary articulation characteristics in the form of the divided cavity system behind the primary point of articulation" (Fant, 1960), and the narrow protruded lip opening.

5.2.3. Formant Continuity.

Fant (1960) has suggested that the spectral structure of a lateral is characterized by the presence of a zero at a frequency position very close to that of F3. "The anti-resonance," he states, "is evidently
due to the shunting effect of the mouth cavity behind the tongue blade." The presence of the anti-resonance very close to F3 would then have the expected effect of neutralizing F3. Fant is thus suggesting that the formant which appears on the spectrograms, and which is normally taken to be F3, is in fact F4.

Lehiste (1964) writes that "The presence of a zero in the region of the third formant might be detected by a discontinuity in the formant movements." That is, the presence of a zero might be inferred spectrographically in those cases where what appears to be the third formant of a lateral ends abruptly at the transition point and the third formant of the following vowel begins at a much higher or lower frequency position.

Lehiste, however, makes the point that "the typical articulations of speech sounds differ in various languages, and, furthermore, that within one language there may exist contextual variants of one phoneme that may exhibit considerable phonetic differences." In particular, she indicates that "It might be possible that some allophones of /l/ are characterized by the presence of a zero in the spectrum, while some other positional variants show a different spectral structure."

Thus, while Lehiste conceives the presence of zeros in the spectral structure of some laterals, she points out the possibility that they might be absent in the spectral structure of other laterals. Such a point of view is evidently in disagreement with Fant’s statement that the acoustic structure of laterals in general is characterized by the presence of a zero.
The question of the presence of absence of a zero in the acoustic spectrum of a lateral is representative of the wider issue of cavity affiliations. Lehiste (1954) voices the belief that "on the basis of information presently available, the calculation of resonances and the assignment of specific formants to particular cavity resonance is not feasible." This is not a lonely voice; as early as 1951 Delattre wrote, quoting Joos in the process, that "It is generally believed that "the shape of the filtering cavity is so very complex as to be mathematically unmanageable..."

5.3. Perception.

One of the early attempts to discover the acoustic cues essential to the recognition of /l/, as well as /w, j, and r/, was that made by O'Connor et al (1957). The experimenters started by examining spectrograms of the sounds under investigation in order to identify the acoustic cues that might possibly be responsible for their recognition. Synthetic spectra of these sounds in which only the first two formants were included, and where the starting frequency of the second formant was systematically varied, were presented to listeners for identification as /l/, /r/, /w/ or /j/.

Analysis of the results of the perceptual tests revealed that while the F2-transition alone was sufficient to cue the distinction between /w/, /j/, and /l, r/ , it failed to cue the distinction between /l/ and /r/.

The experimenters pointed out that the failure of the F2-transition to cue the distinction between /l/ and /r/ is not, in fact,
surprising. The starting frequencies of the second formant for /l/ are not very different from those for /r/. It is true that the starting point of the second formant transition should be somewhat higher for /l/ than for /r/. This, however, did not seem to provide a reliable differentia between the two phones.

We should note, however, that listeners' inability to distinguish between /l/ and /r/ on the basis of F2-transition alone does not mean, as O'Connor et al point out, that "the one-or-less common /r, l/ second formant will produce, with all vowels, a phoneme which is indifferently /r/ or /l/." An interesting result of the experiment was the discovery that it was not so much the frequency position of the F2-onset as it was the direction of the F2-transition. Whenever the F2 had a negative transition the sound was invariably recognized as /r/, and whenever it had a positive transition it was recognized as /l/.

Naturally enough, the frequency position of the third formant starting frequency would have to depend on the nature of the F2-transition. With a negative F2, where the stimulus is perceived as /r/, F3 would have to start at a low frequency very close to that of F2, if the stimulus generated was to be perceived as /r/. If on the other hand, the sound was to be perceived as /l/, the third formant onset "must be no lower than the third formant of the vowel." With a positive F2-transition, where the stimulus is perceived as /l/, F3 would have to start at a very low frequency position to produce /r/, and at a higher position to produce /l/.
5.4. The acoustic analysis of the Arabic lateral.

This section reports the results of the investigation into the spectral and temporal characteristics of the Arabic lateral. The plan of study involved the analysis of both the pharyngealized [ɓ] and the non-pharyngealized [l] in three different word positions: initial, intervocalic and final. In each position, the effect of the Arabic long vowels on the lateral, as well as the effect of the lateral itself on the adjacent vowel, are investigated.

5.4.1. The non-pharyngealized lateral [l].

5.4.1.1. The effect of position on the spectrum of [l].

Table 5.1 shows the average steady-state frequency positions of the first three formants of [l] in initial, intervocalic and final positions. It is obvious from the table that the position of [l] in the word has a considerable effect on its spectrum. This is most pronounced in the case of F1 which is highest in intervocalic position and lowest in final position. The degree of variation is not so marked in the case of F2 which remains the same in both initial and final positions and only shows a tendency to rise in intervocalic position. The frequency position of F3, likewise, remains unchanged in both the initial and intervocalic positions. It should be mentioned, however, that the frequency positions of F3 in initial and intervocalic positions have been averaged from a few occurrences only since F3 is absent in the majority of cases investigated in these two positions, and always absent in final position.
Table 5.1. Average steady-state frequency positions for the first three formants of [l] (in Hertz) in initial, intervocalic, and final positions.

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>315</td>
<td>1500</td>
<td>2300</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>420</td>
<td>1560</td>
<td>2300</td>
</tr>
<tr>
<td>Final</td>
<td>265</td>
<td>1500</td>
<td>--</td>
</tr>
</tbody>
</table>

5.4.1.2. The effect of the adjacent vowel on [l].

Figure 5.1 presents the formant positions of [l] as a function of the adjacent vowel. The lines connect points representing the average values measured for 80 occurrences of [l] in initial, intervocalic and final positions. As may be seen from the figure, the range of variation of F1 is slight. There is, however, a tendency to vary directly with the F1 of the adjacent vowel. Limited variation is again observed in the case of F2 of the lateral. This is, indeed, surprising, since one would expect F2 of [l] to show more extensive fluctuation similar to that observed in the case of other segments studied in this work. F2 of the lateral does not, furthermore, seem to vary directly with the adjacent vowel but seems to stay around 1500 Hz regardless of the nature of the vowel. The effect of the nature of the adjacent vowel on the lateral F3 could not, however, be studied since this is absent in most of the occurrences of [l].

5.4.1.3. Transitions.

Figure 5.2 presents the average steady-state as well as the onset and offset frequency positions of the second formant of the five Arabic
Figure 5.1. The average frequency positions for the first two formants in the spectra of [i] as a function of the adjacent vowel.
long vowels following initial and intervocalic [l] and preceding intervocalic and final [l]. In general, [l] seems to have a negative transitional influence on the close front vowels and a positive transitional influence on the back vowels. Its influence on the transition of [aː] however, depends on its position in the word. When [l] occurs prevocally, it has a negative transitional influence on [aː]. Post-vocalic [l], by contrast, has a slight positive transitional influence on [aː] when it occurs in intervocalic position, and a neutral transitional influence when it occurs in word-final position.

As may be seen from figure 5.2, the extent of transition is greater when the [l] occurs in initial position than when it occurs in final position. Indeed, the figure shows that even when the [l] is in intervocalic position the extent of the transition is greater when it precedes the vowel than when it follows it. This suggests that the effect of [l] on the second formant of an adjacent vowel is greater in a consonant-vowel context than in a vowel-consonant one.

5.4.1.4. Duration.

Table 5.2 shows the average durations of [l] in initial, intervocalic and final positions. The position of the lateral in the word has a considerable influence on its duration. The lateral has the longest duration in final position, the shortest in intervocalic position, with an intermediate value in initial position. The nature of the following vowel, by contrast, has little influence on the duration of the lateral. (Cf. Table 5.3)
Figure 5.2. The average onset, offset and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial and intervocalic [l] (left-hand side) and before intervocalic and final [l] (right-hand side).
Table 5.3 Average durations of $\text{l}$ (in milliseconds) as a function of the neighbouring vowel.

<table>
<thead>
<tr>
<th></th>
<th>[i:]</th>
<th>[e:]</th>
<th>[ɛ:]</th>
<th>[o:]</th>
<th>[u:]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>115</td>
<td>110</td>
<td>115</td>
<td>105</td>
<td>105</td>
</tr>
</tbody>
</table>

5.4.2. The pharyngealized lateral $\text{l}$.

5.4.2.1. The effect of position on the spectrum of $\text{l}$

Table 5.4 shows the average steady-state frequency positions of the first three formants of $\text{l}$ in initial, intervocalic, and final positions. $F_3$ is absent in all three positions. Similar to $\text{l}$, the frequency position of the first formant of $\text{l}$ in final position is lower than that in initial position which is, in turn, lower than that in intervocalic position. $F_2$, by contrast, is lower in final position than in intervocalic position which is, in turn, lower than that in initial position.

<table>
<thead>
<tr>
<th></th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>$F_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>400</td>
<td>1100</td>
<td>--</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>580</td>
<td>1035</td>
<td>--</td>
</tr>
<tr>
<td>Final</td>
<td>295</td>
<td>1000</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 5.4 Average steady-state frequency positions of the first three formants of $\text{l}$ (in Hertz) in initial, intervocalic and final positions.
5.4.2.2. Duration

Table 5.5 shows the average durations of \([l]\) in initial, intervocalic and final positions. Similar to \([l]\), the \([l]\) has the shortest duration in intervocalic position, the longest in final position, with an intermediate value in initial position.

<table>
<thead>
<tr>
<th>Initial</th>
<th>Intervocalic</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>70</td>
<td>160</td>
</tr>
</tbody>
</table>

Table 5.5 Average durations of \([l]\) (in milliseconds) in initial, intervocalic and final positions

5.4.2.3. Transitions.

Figure 5.3 presents the average steady-state as well as the onset and offset frequencies of the second formant of \([\alpha:]\) in the vicinity of \([l]\). As may be seen from the figure, \([l]\) has a neutral transitional influence on \([\alpha:]\). The position of \([l]\) in the word does not affect either the direction or the extent of the transition.

5.5. Conclusion.

The acoustic analysis of the Arabic lateral reported in the previous sections has shown that while the non-pharyngealized lateral has different spectral characteristics and exerts a different transitional influence from that of the pharyngealized lateral, both laterals show similar patterns of spectral and temporal variation with word position.
Figure 5.3. The average onset, offset, and steady-state frequency positions for the second formant of [æi] after word-initial and intervocalic [l] (left-hand side), and before intervocalic and final [l] (right-hand side).
In this section we shall deal first with the spectral differences between the pharyngealized and the non-pharyngealized laterals. We shall then examine the variation of the spectral characteristics of both as a function of word position and the adjacent vowel. We shall then discuss how the vowel spectrum is itself affected by the lateral and word position. Lastly, we shall examine the factors that determine the temporal properties of the Arabic lateral.

5.5.1. Spectral differences between the pharyngealized and the non-pharyngealized lateral.

The acoustic analysis of the Arabic lateral has revealed that there are significant differences between the spectral characteristics of the pharyngealized [l] and the non-pharyngealized [l]. First, the spectrum of the non-pharyngealized lateral shows the presence of its first three formants at the average frequency positions of about 330 Hz for F1, 1520 Hz for F2, and 2300 Hz for F3. The spectrum of the pharyngealized lateral, by contrast, is characterized by the absence of the third formant in all positions investigated. Furthermore, and in common with other pharyngealized sounds, the average frequency position of F1 at 425 is considerably higher than that for F1 of [l], while the average frequency position for F2 at about 1045 Hz is considerably lower than that for F2 of [l].

5.5.2. The effect of word position on the spectrum of the lateral.

By contrast with other speech segments studied in this work, the...
effect of word position on the frequency position for the first three formants of both [l] and [r] is more observed in the case of F1 than in the case of F2. F3, in those cases in which it could be detected, shows little or no variation at all.

Similar to other speech sounds studied in this work, however, the frequency positions of lateral formants are lowered in the spectra of final position laterals. For both [l] and [r], the frequency positions of their first two formants are higher in intervocalic position than in initial position, which are in turn higher than those in final position.

5.5.3. The effect of the adjacent vowel on lateral spectra.

By contrast with the other speech sounds studied in this work, with the exception of the lingual fricatives, the range of variation of the second formant is exceptionally limited. Furthermore, the second formant does not seem to vary directly with the adjacent vowel.

5.5.4. The effect of the lateral and word position on vowel spectra.

The transitional influence exerted by the pharyngealized lateral on the extent of F2-transition of the adjacent vowel is different from that exerted by the non-pharyngealized lateral. The extent of
transition of the second formant of the adjacent vowel is affected by the position of the non-pharyngealized lateral in the word in such a way that it is greater when the non-pharyngealized lateral precedes the vowel than when it follows. In the case of the pharyngealized lateral, by contrast, word position has no effect on the degree of influence exerted on the extent of F2-transition of the adjacent vowel.

5.5.5. Duration.

The duration characteristics of both [l] and [l] are very similar. [l] has an average duration of about 110 ms, while [l] has an average duration of about 100 ms. More importantly, both [l] and [l] show the same pattern of duration variation with word position. They have the longest duration in final position, the shortest in intervocalic position, with an intermediate value in initial position.
Chapter 6

The nasals

6.1. Articulatory mechanism.


6.2.1. The contribution of the oral side chamber.

6.2.2. Formant continuity.

6.3. Perception.

6.4. The acoustic analysis of Arabic nasals.

6.4.1. The non-pharyngealized bilabial nasal [m].

6.4.1.1. The effect of position on the spectrum of [m].

6.4.1.2. The effect of the adjacent vowel on [m].

6.4.1.3. Duration.

6.4.1.4. Transitions.

6.4.2. The pharyngealized bilabial nasal [n].

6.4.2.1. The effect of position on the spectrum of [n].

6.4.2.2. Duration.

6.4.2.3. Transition.

6.4.3. The non-pharyngealized dental nasal [n].

6.4.3.1. The effect of position on the spectrum of [n].

6.4.3.2. The effect of the adjacent vowel on the spectrum of [n].
6.4.3.3. Duration.

6.4.3.4. Transitions.

6.4.4. The pharyngealized post-dental nasal [n].

6.4.4.1. The effect of position on the spectrum of [n].

6.4.4.2. Duration.

6.4.4.3. Transitions.

6.5. Conclusion.
6.1 Articulatory mechanism.

A nasal consonant is a type of consonant segment during the production of which the air stream escapes only through the nose. During the production of a nasal consonant a stricture of complete closure is effected at some point in the oral tract thus preventing the escape of air through the mouth. The velum, however, is lowered permitting the air stream to escape freely through the nose.

6.2 Acoustic theory of production of nasals.

The acoustic system suitable for the production of a nasal consonant consists essentially of the pharynx and nasal cavities, which are maximally coupled, and of the oral cavity which acts as a closed side chamber.

Researchers agree that the formants of nasal consonants are formed chiefly by the pharynx-nose system, while the anti-formants are formed by the closed oral cavity, which functions as a side chamber coupled to the main path. (Cf. Fant, 1960; Hattori, 1958; Hecker, 1962; Fujimura, 1962; Su, Li and Fu, 1974)

Spectral properties of nasal consonants could be predicted by examining the dimensions and characteristics of the generating system. Studies based on such theoretical considerations (Fant, 1960; Fujimura, 1962; and Hecker, 1962) have, indeed, been able to predict
spectral properties of nasal consonants which are in general agreement with their observed spectra.

Such studies all point to the fact that it is the participation of the nasal branch in the generation of nasal spectra which determines the spectral characteristics that distinguish the nasals as a class from other speech sounds.

First, Fujimura (1962) has shown that the density of the formants of an acoustic system is a function of its length. According to him, the average spacing of the formants of an acoustic system such as the pharynx-nose system will be expected to be smaller (about 800 Hz) than the average spacing of the pharynx-mouth system (about 1000 Hz). The difference, he indicates, "is ascribed to the unequal lengths of the nasal and oral cavities."

Second, Fant (1960), Flanagan (1965), and Fujimura (1962) have pointed out that the proportionally greater surface area of cavity walls present in the nasal tract would result in greater damping factors of formants generated by the system. As Fant (1960) points out, "A greater damping of resonance in the nasal part than in the oral part of the vocal tract can be expected owing to the greater surface outline to area ratio of any cross-section except in the nasopharynx."

And, third, vocal tract model experiments (Fant, 1960) have shown that the transmission characteristics of an acoustic system simulating the pharynx-nose tract, with no oral side chamber, result in a formant
Acoustic analysis has, indeed, shown these features to be characteristic of the spectra of natural speech nasals in general (Potter, Kopp, and Kopp, 1947; Joos, 1948; Brosnahan and Malmberg, 1970). There is, however, widespread disagreement as to the number and frequency positions of the formants of nasals.

The reported frequency positions of the formants of nasal consonants differ widely from one study to the other. For instance, has been reported to have formants at 250, 1100, 1350, and 2000 Hz (Flanagan, 1965), at 250, 850, 1300, 1800, and 2250 Hz (Joos, 1948), and at 300, 1000, 2200, and 2500 Hz (Brosnahan and Malmberg, 1970).

Furthermore, in the frequency region extending from the bottom of the spectrum up to about 2400 Hz, a region where most vowels show only two formants, the formants of nasal consonants have been reported to number anything from one to five formants (Potter, Kopp, and Kopp, 1947; Fant, 1960; Hecker, 1962; Ladefoged, 1975; Brosnahan and Malmberg, 1970).

Differences as to the number and frequency positions of the formants of nasal consonants should not, however, be surprising for two reasons. First, researchers agree that it is generally difficult to extract any consistent features of the spectra of nasal consonants, especially when a comparison is made over several speakers (Fant, 1960; Fujimura, 1962; Dickson, 1962; and Nord, 1976). Secondly, Potter, Kopp and Kopp (1947) have noted that the number of formants of nasal consonants...
varies with the degree of stressing. Thus, while the stressed /m/ or /n/ "has a resonance area directly above the voice bar, another area which varies somewhat in position in the lower central part of the pattern, and a resonance area which resembles a double bar at about the same position as bar 3 of the vowel a", the unstressed /m/ or /n/ "may be portrayed either by the voice bar alone, or by the voice bar plus one or more weakly defined resonances."

6.2.1 The contribution of the oral side chamber.

The existence of the mouth cavity as a side chamber coupled to the pharynx-nose system is important in the determination of the transmission characteristics of that system, and variation of its volume will affect these characteristics, and hence the output spectra of the various nasal consonants.

Vocal tract model experiments (Fant, 1960) have shown that coupling the pharynx-nose system to the oral side chamber has the effect of increasing the total volume of the coupled system. As such the frequency positions of the formants of nasal consonants will be expected to vary inversely with the increase of volume of the oral cavity. This will thus account for the fact that the spectra of nasals show the lowest range of frequencies for [m], and the highest for [ŋ], with an intermediate range for [n].

Variation in the size and shape of the oral cavity are also the main factor in determining the frequency positions of the anti-formants of nasal consonants.
First, the location of the antiformants would be expected to be
different for different nasal consonants since the oral cavity takes
on a different shape and size for each nasal consonant. The frequency
positions of the antiformants of a velar nasal, for instance, are
higher than those of a dental nasal, which are in turn higher than
those of a bilabial nasal; the mouth cavity volume getting progressively
smaller as the obstruction moves towards the back of the mouth. As
Fuimura (1962) has shown, the antiformant of [m] is between 750 Hz
and 1250 Hz, that of [n] is between 1450 Hz and 2200 Hz, while that of
[ŋ] is above 3000 Hz. In general, the narrower the back opening of the
oral cavity, viz, the constriction made by the back of the tongue, and
also the larger the volume of the oral cavity, the lower the frequency
position of the antiformant.

Second, the frequency position of the antiformant of the same
nasal would be expected to be different depending on the nature of
the adjacent vowel. That is, the frequency of the antiformant of
a nasal is relatively high when the consonant is adjacent to a front
vowel and is lower when it is adjacent to a back vowel. When a nasal
precedes a front vowel, the anterior part of the oral cavity during
the production of the nasal is narrowed in anticipation of the vowel,
and acoustic theory would predict a consequent rise in the frequency
of the antiformant. In the case of a back vowel context, on the other
hand, the anticipatory tongue position results in a large mouth cavity
with a comparatively narrow neck, with a consequent decrease in the
frequency of the antiformant.
The configuration of the vocal tract also determines the bandwidth of the antiformant of nasal consonants. Fujimura (1962) accounts for the fact that the antiformant of /n/ has a very wide bandwidth as compared with the rather sharp antiformant of /m/, by pointing out that "In the case of /n/, the wedge-shaped termination of the oral cavity causes a gradual change in characteristic impedance and hence results in relatively large absorption of sound at the termination. The configuration of the oral cavity for /n/ on the other hand, has a rather abrupt termination, and the mouth cavity has a smaller ratio of surface area to volume. The acoustic losses in the oral cavity are, therefore, less appreciable and the bandwidth of the antiformant is smaller."

6.2.2 Formant continuity.

It will be apparent by now that the sound structure of the steady-state part of a nasal consonant - with its oral tract generated antiformants, nasal tract generated formants, and high damping factors - is basically different from that of a vowel. "This fact," point out Hattori, Yamamoto and Fujimura (1958) "causes sometimes an essentially discontinuous transition in the structure of the sound, or in other words, an abrupt change of timbre at the moment of release of the stop at the point of articulation, giving often an impression of plosion, even when no actual plosion can be heard."

This remark is reminiscent of a similar observation made earlier by Potter, Kopp, and Kopp (1947) who have noted that a rapid movement of the velum "results in sharp initiation and
termination of the vowel resonances and also in a brief spike between the nasal and the vowel."

Nakata (1959), however, has disputed the existence of a discontinuity in formant frequencies at the boundary between a nasal consonant and a vowel. He believes that if the coupling to the nasal tract at the velum is not too great, the second and higher resonance frequencies of the vocal tract will not be greatly affected, although their amplitudes might be influenced by the highly damped nasal tract resonances and the spectral zeros originating from the side chamber. "Thus," he points out "as the articulators shift from the consonant to the vowel configuration there may be relatively sharp changes in the amplitudes of the second and higher formants, but the discontinuities in formant frequencies are probably not too great."

There is no doubt that Nakata's rejection of the existence of a discontinuity in formant frequencies at the boundary between a nasal and a vowel, and the hypothesis on which his rejection is based, form a departure from the generally accepted theory of production of nasals. Most researchers in the field have, after all, assumed maximal coupling of the mouth system to the pharynx-nose system. (Cf. Fant, 1960 and Hecker, 1962)

It is precisely in such a situation that the results of the acoustic analysis of speech can be crucial in confirming, or refuting, one or the other of two contradicting hypotheses. Nakata has, after all, claimed that "Examination of the spectrograms of natural speech confirms" his hypothesis.
6.3 Perception.

In view of the fact, as we have already seen, that the "nasal consonants have different antiformants, which fall within different and distinctively defined regions" (Su, Li, and Fu, 1974), it is not unusual to expect that "the antiformant plays the major role in identification of the point of articulation during the closure period of nasal consonants" (Hattori, Yamamoto, and Fujimura, 1958). Portions of much shorter durations have, after all, been shown to have residual acoustic cues that carry place of production information — for instance, the explosion phase of a plosive (Halle, Hughes, and Radley, 1957).

Experiments with natural speech (Malecot, 1956), and synthetic speech (House, 1957 and Nakata, 1959) have, indeed, demonstrated that the isolated steady-state parts of different nasals can be correctly identified. Malecot recorded gated steady-state portions of different nasals in random order and presented them to a group of listeners for identification as /m/, /n/, or /ŋ/. Analysis of listeners' responses showed that the percentage of correct responses was 96 for /m/, 56 for /n/, and 12 for /ŋ/. Much higher percentages, however, were reported by Nakata (1959) for his synthetic stimuli, with 90 for /m/, 80 for /n/, and 65 for /ŋ/. These results, together with those of another experiment by House (1957), seem to indicate that isolated synthetic nasals are more readily identified than the natural nasals.
Since the early days of spectrography, however, there were suggestions that it might be the transitions of adjoining vowels, rather than the steady-state parts, which are perceptually important as carriers of place of production information (Potter, Kopf, and Kopp, 1947 and Joos, 1948).

Such suggestions were later supported, in part, by the results of experiments with synthetic speech carried out at the Haskins Laboratories (Cooper, Delattre, Liberman, Borst, and Gertsman, 1952) and other research centres (Nakata, 1959). These experiments showed that the second formant transitions of adjacent vowels are perceptually important as acoustic cues conveying information on the point of articulation of nasal consonants.

Experimenters were thus able to show that information on the point of articulation of a nasal consonant is provided by both the steady-state and transition parts. The next step was, logically enough, to try to determine the relative importance of the acoustic cues resident in both parts. In other words, they wanted to know which part, if either, will be dominant and thus determine listeners' perception of the sound in those cases where the steady-state carried place cues different from those carried by the vowel transitions.

Such an experiment was carried out by Malécot (1956) who split the transition part away from the steady-state part and recombined transition and steady-state parts from different syllables, on the assumption that the more powerful part of the sound would determine
which phone a listener would hear. The results of the experiment showed that while the transitions of the adjoining vowels are over-dominant, the steady-state parts "are not completely neutral with respect to place identification; they contain a small amount of place information almost negligible in initial position but somewhat more important in terminal position."

In a later experiment by Nord (1976), however, the contribution of the steady-state part of natural speech nasals was found to be over-dominant in intervocalic position. Nord combined steady-state and transition parts from different stretches of vowel + nasal + vowel and presented them to listeners for judgment. Analysis of listeners' responses showed that they depended primarily on the steady-state part in their identification of the stimuli. The percentages of correct responses reached an impressive 80 for /n/ and 70 for /m/.

The results of Nord's experiment indicate quite clearly that the steady-state parts of nasal consonants carry strong place cues. They cannot, as such, support Nakata's argument (1959) that since in natural speech the spectra of the steady-state parts are often obscured by the presence of spectral zeros, high damping, or other extraneous factors, "The identification of nasal consonants in natural speech must, therefore, rely heavily on the second formant locus as observed in the transition of the adjacent vowel."

The results of Nord's experiment are highly significant in that they throw doubt on a nearly established tradition that has always
belittled the contribution of the steady-state part in the identification of the point of articulation in general. They are also surprising especially since previous acoustic analyses have pointed out the difficulty of differentiating /m/ spectra from /n/ spectra (Cf. Potter, Kopp, and Kopp, 1947 and Brosnahan and Malmberg, 1970). Indeed, the results of the acoustic analysis of the Arabic nasals reported in Section 6.4 show that although the frequency positions for the formants of /n/ are generally higher than those for /m/, they very often overlap.

6.4 The acoustic analysis of Arabic nasals.

This section reports the results of the investigation into the spectral and temporal properties of the Arabic bilabial and dental nasals [m] and [n], and their pharyngealized counterparts [g] and [ŋ]. The plan of study involved the analysis of each nasal in three different positions: initial, intervocalic and final. In each position the effect of the Arabic long vowels on the nasal, as well as the effect of the nasal itself on the adjacent vowel, are investigated.

6.4.1 The non-pharyngealized bilabial nasal [m].

6.4.1.1. The effect of position on the spectrum of [m].

Table 6.1 shows the average steady-state frequency positions for the first three formants of [m] in initial, intervocalic, and final positions. The frequency positions for F1 and F3 in final position are lower than those in intervocalic position, which are in turn lower than those in initial position. The frequency positions of F2, on the other hand, are higher in intervocalic and final positions than in initial position.
Table 6.1 Average steady-state frequency positions of the first three formants of [m] (in Hertz) in initial, intervocalic and final positions.

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>235</td>
<td>860</td>
<td>2015</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>215</td>
<td>1000</td>
<td>1725</td>
</tr>
<tr>
<td>Final</td>
<td>200</td>
<td>1000</td>
<td>1750</td>
</tr>
</tbody>
</table>

6.4.1.2 The effect of adjacent vowel on [m].

Figure 6.1 presents the formant positions of [m] as a function of the neighbouring vowel. The lines connect points representing average values measured for 80 occurrences of [m] in initial, intervocalic, and final positions. As may be seen from the figure, the first formant shows little variation and seems to vary independently of the adjacent vowel. The second formant, by contrast, shows a greater range of fluctuation and, apart from its exceptionally high frequency position in the vicinity of [æ], seems to vary directly with the second formant of the adjacent vowel. The third formant is invariably high and no systematic variation could be attributed to the influence of the adjacent vowel.

6.4.1.3 Duration.

Table 6.2 shows the average durations of [m] in initial, intervocalic, and final positions. The position of the nasal in the word
Figure 6.1. Average frequency positions for the first three formants in the spectra of [m] as a function of the adjacent vowel.
has a considerable influence on its duration. The nasal has the longest duration in final position, the shortest in initial position, with an intermediate value in intervocalic position. The nature of the adjacent vowel, by contrast, has little influence on the duration of the nasal. Furthermore, as may be seen from table 6.3, there is no discernible pattern in the variation of the duration of [m] which could be attributed to the influence of the adjacent vowel.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Intervocalic</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>75</td>
<td>100</td>
<td>190</td>
</tr>
</tbody>
</table>

Table 6.2 Average durations of [m] (in milliseconds) in initial, intervocalic, and final positions.

<table>
<thead>
<tr>
<th></th>
<th>iː</th>
<th>eː</th>
<th>ðː</th>
<th>oː</th>
<th>uː</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>115</td>
<td>120</td>
<td>115</td>
<td>130</td>
<td>115</td>
</tr>
</tbody>
</table>

Table 6.3 The average durations of [m] (in milliseconds) as a function of the neighbouring vowel.

6.6.1.4 Transitions.

Figure 6.2 presents the average steady-state as well as the onset and offset frequencies of the second formant of the five Arabic long vowels in the vicinity of [m]. In general, [m] seems to have a negative transitional influence on the front vowels and a neutral transitional
Figure 6.2. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial and intervocalic [m] (left-hand side) and before intervocalic and final [m] (right-hand side).
influence on the back vowels. This is, indeed, expected since [m], in
combination with other bilabial sounds, has a very low locus.

As may be seen from figure 6.2, the extent of transition is greater
in initial than in final position. Indeed, the figure shows that even
when [m] is in intervocalic position the extent of transition is
greater when the nasal precedes the vowel than when it follows it.
This suggests that the effect of [m] on the second formant of an
adjacent vowel is greater in a consonant-vowel context than in a
vowel-consonant context.

6.4.2 The pharyngealized bilabial nasal [m].

6.4.2.1 The effect of position on the spectrum of [m].

Table 6.4 shows the average steady-state frequency positions
of the first three formants of [m] in initial, intervocalic, and
final positions. F3 is absent in all three positions. In final
position, furthermore, F2 is also absent with the result that we are
left with only F1. Similar to [m], the frequency position of F1 in
final position is lower than that in intervocalic position, which is
in turn lower than that in initial position. In the case of F2,
however, the situation is different, with F2 in intervocalic position
lower than F2 in initial position.
Table 6.4  Average steady-state frequency positions of the first three formants of \( [a] \) (in Hertz) in initial, intervocalic and final positions.

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>195</td>
<td>900</td>
<td></td>
</tr>
<tr>
<td>Intervocalic</td>
<td>185</td>
<td>850</td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>180</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.4.2.2 Duration.

Table 6.5 shows the average durations of \( [m] \) in initial, intervocalic, and final positions. Similar to \( [m] \), the \( [m] \) has the shortest duration in initial position, the longest in final position, with an intermediate value in intervocalic position.

<table>
<thead>
<tr>
<th></th>
<th>Intervocalic</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>80</td>
<td>87</td>
</tr>
</tbody>
</table>

Table 6.5  Average durations of \( [m] \) (in milliseconds) in initial, intervocalic, and final positions.

6.4.2.3 Transitions.

Figure 6.3 presents the average steady-state as well as the onset and offset frequencies of the second formant of \( [a] \) in the vicinity of \( [m] \). \( [m] \) obviously has a negative transitional influence on \( [a] \). Furthermore, similar to \( [m] \), the extent of transition is
Figure 6.3. The average onset, offset, and steady-state frequency positions for the second formant for [ɔ:] after word-initial and intervocalic [m] (left-hand side), and before intervocalic and word-final [m] (right-hand side).
greater when the nasal precedes the vowel than when it follows it suggesting once again that the coarticulatory influence of [n] is greater in a CV context than in a VC one.

6.4.3 The non-phonorealized dental nasal [n].

6.4.3.1 The effect of position on the spectrum of the nasal [n].

Table 6.6 shows the average steady-state frequency positions of the first three formants of [n] in initial, intervocalic, and final positions. The frequency positions of F1 and F2 in intervocalic position are higher than those in initial position, which are in turn higher than those in final position. F3, by contrast, is higher in initial position than in intervocalic position and does not appear in final position.

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>245</td>
<td>870</td>
<td>1785</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>250</td>
<td>1010</td>
<td>1515</td>
</tr>
<tr>
<td>Final</td>
<td>220</td>
<td>800</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 6.6 Average steady-state frequency positions of the first three formants of [n] (in Hertz) in initial, intervocalic and final positions.

6.4.3.2 The effect of the adjacent vowel on the nasal [n].

Figure 6.4 presents the average steady-state frequency positions...
Figure 6.4. The average frequency positions for the first three formants in the spectra of \([n]\) as a function of the adjacent vowel.
of the first three formants of [n] as a function of the neighbouring vowel. The lines connect points representing average values measured for 80 occurrences of [n] in initial, intervocalic, and final positions. As may be seen from the figure, the first formant shows little variation and seems to vary independently of the adjacent vowel. The second formant, by contrast, shows greater fluctuation and, apart from its exceptionally high frequency position in the vicinity of [e:], it seems to vary directly with the second formant of the adjacent vowel. Again, with the exception of the uncommonly low value of F3 in the vicinity of [e:], F3 is invariably high and no systematic variation could be attributed to the influence of the adjacent vowel.

6.3.3.3 Duration

Table 6.7 shows the average durations of [n] in initial, intervocalic and final positions. The position of the nasal in the word has a considerable influence on its duration. [n] has the shortest duration in intervocalic position, a slightly longer duration in initial position, and the longest duration in final position. The nature of the adjacent vowel (See table 6.8), by contrast, has little influence on the duration of [n]. Furthermore, there is no discernible pattern in the variation of the duration of the nasal that could be attributed to the influence of the adjacent vowel.
Table 6.8 Average durations of \([n]\) (in milliseconds) as a function of the neighbouring vowel.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Inter-</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80</td>
<td>70</td>
<td>210</td>
</tr>
</tbody>
</table>

6.4.3.4 Transitions.

Figure 6.5 presents the average steady-state as well as the onset and offset frequencies of the second formant of the five Arabic long vowels in the vicinity of \([n]\). In general, \([n]\) seems to have a negative transitional influence on the front vowels, and a positive transitional influence on the back vowels. This is, in fact expected since \([n]\) has a higher locus than \([m]\).

As may be seen from Figure 6.5, the extent of transition is greater when the nasal occurs in initial position than when it occurs in final position. Indeed, the figure shows that even when \([n]\) is in intervocalic position the extent of the transition is greater when the nasal precedes the vowel than when it follows it. This suggests that the effect of \([n]\) on the second formant of an adjacent vowel is greater in a consonant-vowel context than in a vowel-consonant one.
Figure 6.5. The average onset, offset and steady-state frequency positions for the second formants of the Arabic long vowels after initial and intervocalic [n] (left-hand side) and before intervocalic and final [n] (right-hand side).
6.4.4 The pharyngealized post-dental nasal [n].

6.4.4.1 The effect of position on the spectrum of [n].

Table 6.9 shows the average steady-state frequency positions of the first three formants of [n] in initial, intervocalic, and final positions. The position of the nasal in the word has a considerable influence on its spectrum. The frequency position of F2 in initial position is higher than that in intervocalic position, which is in turn higher than that in final position. The frequency positions of F1 in initial and intervocalic positions are the same. They are, however, higher than F1 in final position. In the case of F3 no comparison can be made since it is present only in initial position.

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>250</td>
<td>1175</td>
<td>2000</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>250</td>
<td>1125</td>
<td>--</td>
</tr>
<tr>
<td>Final</td>
<td>200</td>
<td>1075</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 6.9 Average steady-state frequency positions of the first three formants of [n] (in Hertz) in initial, intervocalic and final positions.

6.4.4.2 Duration.

Table 6.10 shows the average durations of [n] in initial, intervocalic, and final positions. The duration of [n] is considerably
affected by its position in the word. It has the shortest duration in intervocalic position, the longest in final position, with an intermediate value in initial position.

<table>
<thead>
<tr>
<th>Initial</th>
<th>Intervocalic</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>60</td>
<td>225</td>
</tr>
</tbody>
</table>

Table 6.10 Average durations of \( n \) (in milliseconds) in initial, intervocalic, and final positions.

6.4.4.3 Transitions

Figure 6.6 presents the average steady-state as well as the onset and offset frequencies of the second formant of \( a:i \) in the vicinity of \( n \). The lines represent values measured for 16 occurrences of \( n \) in initial, intervocalic, and final positions. As may be seen from the figure, \( n \) has a slight positive influence on \( d:i \). The pattern of transitional influence displayed in figure 6.6 is the same as that displayed by other nasals, pharyngealized or otherwise, in which transitional influence is greater in a consonant-vowel context than in a vowel-consonant one.

6.5 Conclusion

The results of the acoustic analysis of Arabic nasals reveal, as expected, that their spectral properties have much in common with the predicted and observed spectra of nasal consonants in general.
Figure 6.6. The average onset, offset, and steady-state frequency positions for the second formant of [aɪ] after word-initial and intervocalic [a] (left-hand side) and before intervocalic and final [n] (right-hand side).
First, Arabic nasals always show a first formant at about 250 Hz. This agrees quite well with Fant's prediction (1960) that the transmission characteristics of an acoustic system simulating the pharynx-nose tract, with no oral side chamber, will result in a formant at about 300 Hz.

Second, in those cases where Arabic nasals display more than one formant the average spacing between the formants is, indeed, the same as that specified by Fujimura (1962), that is, about 800 Hz.

Third, in agreement with the findings of Potter, Kopp and Kopp (1947) the number of formants of Arabic nasals varies with the degree of stress. Analysis of overall intensity of Arabic nasals shows that they are considerably more intense in initial and intervocalic positions than in final position. This accounts for the fact that Arabic nasals usually display three formants in the initial and intervocalic positions and only one formant in the final position.

Fourth, the first two formants of [n] are, in general, higher than those of [n]. However, as Brosnahan and Halmberg have noted, the formant frequencies of these two sounds are very similar and very often overlap. In the case of the pharyngealized nasals, however, the first two formants of [n] are consistently higher than those for [ŋ], with the result that they can always be differentiated.

Fifth, spectra of Arabic nasals show an abrupt change of timbre at the moment of release of the nasal. This supports the assertion
made by Hattori, Yamamoto and Fujimura (1958) about the existence of an essentially discontinuous transition between the nasal and the vowel.

The acoustic analysis of Arabic nasals has, however, revealed two surprising results. The first relates to the fact that the frequency position for the second formant of [n] has been found to be higher than that for [n]. The raising of the second formant of the pharyngealized nasal is, undoubtedly, unexpected. One would expect [n], or any other pharyngealized sound for that matter, to have a lowered second formant and not a raised one.

The second, on the other hand, relates to the fact that the first formant of [m] has been found to be lower than that for [m]. Fant (1960) has suggested that the low F1 of nasal consonants represents the fundamental resonance of the pharynx cavity tuned by the nasal system as a resonator neck. In terms of this analysis one would expect — due to the constricted and hence smaller pharynx cavity — the pharyngealized bilabial nasal to have a higher F1 than that of the non-pharyngealized bilabial nasal. As the acoustic analysis of Arabic nasals shows, however, this is not the case.

Our investigation has also shown that the position of the nasal in the syllable affects the degree of coarticulation. For all nasals, the extent of the transition of the second formant of an adjacent vowel is greater when the nasal precedes the vowel than when it follows it. This suggests that the effect of the nasal on the second formant of an adjacent vowel is greater in a CV context than in a VC one.
In a more detailed vein, the results of the acoustic analysis of Arabic nasals in the preceding sections indicate that:

1. The position of the nasal in the word has an observed effect on its spectrum. For all nasals - pharyngealized or non-pharyngealized - the region of concentration of energy is lowered in frequency where the nasal occurs in intervocalic or final position.

2. The adjacent vowel has a considerable effect on the steady-state part of the nasal. This effect is marked only in the case of the second formant of the nasal part which varies directly with the frequency position of the second formant of the adjacent vowel. The first and third formants of the nasal part show, by contrast, little fluctuation.

3. The position of the nasal in the word has a considerable effect on its duration. For [n] and [m], the nasal is shortest in intervocalic position, longest in final position, and with an intermediate value in initial position. For [m] and [n], on the other hand, the nasal is shortest in initial position, longest in final position, and with an intermediate value in intervocalic position. The nature of the adjacent vowel, by contrast, has no systematic effect on the duration of the nasal.

4. [m] has a negative transitional influence on the front vowels and a neutral transitional influence on the back vowels. [n], by contrast, has a negative transitional influence on the front vowels and a positive transitional influence on the back vowels.
Chapter 7

The semivowels

7.1. The articulatory mechanism.

7.2. Acoustic analysis.

7.3. Perception.

7.4. The acoustic analysis of the Arabic semivowels.

7.4.1. The effect of position on the spectrum of the semivowel.

7.4.2. The effect of the adjacent vowel on the spectrum of the semivowel.

7.4.3. Duration.

7.4.4. Transitions.

7.5. Conclusion.
7.1. The articulatory mechanism.

Semivowels, like vowels, could be defined as "segments made with central passage of the air-stream and open approximation of the articulators, so that no noise of friction is produced." (Abercrombie, 1967). Potter, Kopf and Kopp (1947), however, would define them further as sounds "produced only while the articulators are in motion." According to Potter, Kopf and Kopp, then, semivowels would be essentially equivalent to formant transitions.

Apart from the fact that Potter, Kopf and Kopp advance no reasons for their decision to regard the semivowels as glides or formant transitions, such a viewpoint seems to be unjustifiable. The semivowels, unlike formant transitions, can be produced in isolation and can, as Potter, Kopf and Kopp themselves concede, be "continued from steady-state positions of the articulators." It also leads to unnecessary complications: vowels in the vicinity of semivowels (as thus defined) will no longer have formant transitions, since these are now the semivowels themselves; semivowels will have no transitional influence on adjacent sounds; a semivowel will also have no locus, for its initial and terminal frequencies will differ with every different adjacent speech sound. It is, therefore, necessary to reject the definition of the semivowels as provided by Potter, Kopf and Kopp and limit it to that provided by Abercrombie.
The mechanism involved in the production of the semivowels /j/ and /w/ is, one could safely assume, fully understood. The semivowel /j/ is produced with the lips spread or neutral, and the tongue assuming the position of a front half-close to close vowel. The semivowel /w/ on the other hand, is produced with the lips rounded and the tongue assuming the position for a back half-close to close vowel.

Our knowledge of the mechanism involved in the production of /i/ is, by contrast, far from complete. Until very recently, the Arabic /i/ - classical, modern, and in the various dialects - was always described as a pharyngeal fricative - the voiced counterpart of the Arabic voiceless pharyngeal fricative /h/. Specific references to Egyptian Arabic include Gairdner (1925) and Al-Ani (1978).

However, the results of the acoustic analysis reported in this work have shown that the /i/ is not a fricative in Egyptian Arabic. Spectra of /i/ in initial, intervocalic, and final positions show no trace of the random scattering of noise characteristic of fricative consonants. By contrast, these spectra are characterized by a clear formant structure very much like that of the Arabic vowels.

Thus, contrary to ALL accepted notions as to the nature of the Egyptian /i/ (Gairdner, 1925; Al-Ani, 1978), the results of the acoustic analysis reported in this work (Section 7.2) show that the Egyptian /i/ is, in fact, phonetically a vowel and not a fricative.
Physiological support for this finding comes from a cineradiographic study of the pharyngeal features in Lebanese Arabic. Delattre (1971) found that, contrary to what the traditional correlation between /ɣ/ and /h/ would have led us to expect, the pharyngeal stricture for /h/ is narrower than that for /ɣ/. This, he points out, "is to be expected since, in the absence of voicing, the friction noise must be loud enough to carry the load of perception alone."

Indeed, examination of the X-ray tracings of Delattre's experiment shows that, apart from the fact that "the front portion of the tongue dorsum for /ɣ/ is higher and more fronted than for /a/", the stricture between the tongue root and the pharyngeal wall is virtually the same for both these two sounds.

Further counter-evidence to the traditional view that Arabic /ɣ/, generally across dialects, is a voiced pharyngeal fricative has been provided for Iraqi Arabic by Al-Ani (1978) who reports that the results of his acoustic analysis show that the Iraqi /ɣ/ is realized either as a stop or a glide. While, as we have already mentioned, Egyptian /ɣ/ is realized as a glide, in none of the positions investigated did it occur once as a stop.

Egyptian /ɣ/ is, however, accompanied during its production by a sphincteric constriction of the pharynx musculature short of actual closure. This phenomenon has, indeed, been observed by Gairdner (1925) who noted that in the production of Egyptian /ɣ/ "the general tenseness of the pharynx is notably increased."
7.2. Acoustic analysis.

The similarity between the acoustic spectra of the semi-vowels /j/ and /w/ and the vowels /i/ and /u/ has been observed by Potter Kopp and Kopp (1947). Potter, Kopp and Kopp have found that the steady-state part of American English /j/ resembles the pattern of /i/, except that formants 2 and 3 are a little higher than for /i/. F1 of the steady-state part is, however, on the baseline as it is for /i/. The steady-state part of /w/, on the other hand, resembles the pattern of /u/, except that F2 occupies a slightly lower position than for /u/.

Potter, Kopp and Kopp have also observed that "Although the duration of the pattern of the sound made at the origin of a glide will vary with the rate of utterance, the positions of its resonance bars remain quite constant." In the case of /j/, however, they found that "The initial position of bar 2 of j shifts slightly with different vowel. It is slightly higher with the front vowels than with the back vowels."

Potter, Kopp and Kopp have also reported that a noise source is often added to the voice source used in producing /w/ and, to a lesser extent, /j/. This results in the appearance of "vertical striations either superimposed on the resonances or substituted for them," in the steady-state parts of the spectra of /w/ and /j/. In the case of /w/,"it is believed to be related to the tension and position of the lips in producing the steady-state part of the sound."
Potter, Kopp and Kopp have observed that /j/ always has a positive transitional influence on the second and third formants of the adjacent vowel. /w/, by contrast, always exerts a negative transitional influence on the second and third formants of the adjacent vowel. Since the first formant for both /j/ and /w/ starts on the baseline, it will have a negative transitional influence on the first formant of the adjacent vowel, unless this happens to be on the baseline too. In those cases of neutral transition, however, a shift in the width of the first formant of the adjacent vowel takes place.

The vocalic nature of Arabic /j/, or more precisely, of certain allophones of Arabic /j/, has been observed by Obrecht and Al-Ani. Obrecht (1968) has found that the spectra of Lebanese /j/ "can be described as four bands of resonance, at 600 cps, 1300 cps, 2700 cps, and 3300 cps - more or less corresponding to the subsequent formants, but with some blurring." Al-Ani (1978) has similarly found that the Iraqi /j/ in non-geminated intervocalic position is most often realized as "glide continuations of the preceding and following vowel formants."

The results of the acoustic analysis reported in this work have, as has already been indicated, similarly revealed the vocalic nature of Egyptian /j/. On the spectrograms, the Egyptian /j/ appears as a vocalic element with a clear formant structure having F1 at 650 Hz, F2 at 1270 Hz, and F3 at 2220 Hz.
In common with the spectra of both the Iraqi and Lebanese /ý/ - in those cases where they are realized as vocalic elements - the spectra of Egyptian /ý/ are characterized by wide spacing between their voice striations.

Auditory and kinesthetic impressions have suggested that the /ý/ might be produced, as has already been mentioned, by some restricted form of vocal fold vibration similar to that of 'creaky' voice. In order to investigate this point it was decided to run an experiment to examine the acoustic properties of /ý/ as compared to those of creaky voice. (See Spectrograms 1 to 5).

Sustained utterances of 'creaky' voice produced with a 'neutral' vocal tract configuration, and at various speeds, were thus recorded and analysed. The similarities between the spectra of creaky voice and those of /ý/ were so strong as to confirm the belief that the 'voice source' used in producing both is identical.

The results of the acoustic analysis of creaky voice were also useful in the interpretation of another property of /ý/ spectra. That is, spectra of /ý/ often show the presence of a number of low frequency high intensity spikes immediately preceding the /ý/ in initial position and following it in final position. The physiological
Spectrogram 1. A stretch of creaky voice produced at high speed

Spectrogram 2. A stretch of creaky voice produced at low speed
Spectrogram 3. /ai:d/

Spectrogram 4. /biːʃi/
These spikes had the quality of a voiceless stop, with the voiceless phase preceding the voiceless stop. In some cases, the voiceless phase of a voiceless stop, to precede the voiceless phase of a voiceless stop, could be verified by spectrographic analysis. The spectrograms of such a stop would show the presence of these low-frequency bursts, indicating that it is possible for a voiceless stop to precede a stop.

The sonorous equivalent of such a voiceless stop would be, of course, a voiceless final /l/. This would be characterized by a distinct and termination of the sound, and cessation of the voiceless quality before the stop.
explanation of these spikes presented a problem. That is, although these spikes had the familiar appearance of the explosion phase of a stop, they occurred immediately after the /r/ offset in final position that there was not the possibility of the presence of a closure phase of a stop. In initial position, on the other hand, while it was possible for a voiceless stop, or rather the closure phase of a voiceless stop, to precede these spikes - a possibility that cannot be verified by spectrographic means since both the shutting phase and the hold phase of a voiceless stop in absolute initial position do not show on the spectrograms - this was thought to be rather unlikely since the presence of such a stop would have to be inferred from the presence of these low frequency high intensity spikes whereas we already know that it is possible for them to occur after the offset of /r/ without a stop preceding them.

The acoustic analysis of the sustained utterance of 'creaky' voice provided a solution to this problem. Utterances of creaky voice produced at a relatively high speed had spectra similar to the familiar /r/ spectra. Those produced at a relatively slow speed were more similar to the onsets of initial /r/ and the offsets of final /r/. This suggested that the spikes observed at the initiation and termination of /r/ could be the result of gradual acceleration and deceleration of glottal vibration rather than the result of the explosion phase of some voiceless stop.
7.3. Perception.

O'Connor et al (1957) have reported that English /w/ and /j/ can be synthesized satisfactorily with only two formants. They comment that "this is no more than might be expected, bearing in mind the well-known possibility of two-formant vowel synthesis, and the articulatory relationship between /i/ and /j/ and between /u/ and /w/.

The synthesized spectra had a first formant onset frequency of about 240 Hz for both /w/ and /j/. The second-formant onset frequency on the other hand, was about 600 Hz for /w/, and about 2760 Hz for /j/.

The onset frequencies for the first two formants of /w/ and /j/ are, as one would expect, in the neighbourhood of those for /u/ and /i/. Further, it is obvious that it is the onset frequency of the second-formant transition in particular which determines listeners' recognition of /w/ or /j/. The onset of the first formant, by contrast, does not distinguish between /w/ and /j/ since its frequency is the same for both. Similarly, the "third formant contributes little to the perception of /w/ and /j/," (O'Connor et al, 1957) which is not unexpected in a system, like that of English or Arabic which does not have a rounded/unrounded contrast of the type /j, y/ as in French.

O'Connor et al have further reported that in their synthesis of /w/ and /j/ they found it necessary to synthesize at the beginning...
of each formant a relatively short steady state, which they called "a steady-state onset." They point out that steady-state onsets were found to be useful to avoid a potential confusion with clusters of stop + /w, j/. When the onset phase was omitted and the transition began immediately the resulting stimulus appeared to have an explosive beginning, a stop of some kind being clearly apprehended though not always identified with certainty. An onset duration of 30 ms was sufficient to eliminate the explosive beginning, while one of 40 ms and upwards gave the effect of a full vowel, /i/ or /u/.

Transition duration is another factor in the perception of /w/ and /j/. Experiments with two-formant /w/ and /j/ (O'Connor et al., 1937) have shown that durations of 50 ms and 100 ms were satisfactory though the latter gave a somewhat more realistic sound; below 50 ms the effect was of a vowel plus stop or flap. At 150 ms /j/ was still good, but an otherwise satisfactory /w/ pattern was heard as /wr/.

These results are in agreement with those reported previously by Liberman et al. (1956) who had also found that if the transition duration was below 40 ms for /w/ or below 50 ms for /j/ the resulting impression was that of a stop.
7.4. The acoustic analysis of the Arabic semivowels.

This section reports the results of the investigation into the spectral and temporal characteristics of the Arabic semivowels. The plan of study involved the analysis of each of the Arabic semivowels in three different word positions: initial, intervocalic and final. The effect of the Arabic long vowels, as well as the effect of the semivowels themselves on the adjacent vowel, are investigated.

7.4.1. The effect of position on the spectrum of the semivowel.

The position of the semivowel in the word has a considerable effect on the frequency positions of its formants. Table 7.1 gives the steady-state frequency positions of the first three formants of the Arabic semivowels /j/, /ɬ/ and /w/ in initial, intervocalic and final positions.

For /ɬ/, the frequency positions of the first and third formants are higher in intervocalic position than in initial or final positions. The second formant, by contrast, is higher in final position than in initial position, which is, in turn, higher than in intervocalic position.

For /w/, the frequency positions of the second and third formants are similarly highest in intervocalic position, lowest in final position,
<table>
<thead>
<tr>
<th>Pin</th>
<th>Int.</th>
<th>Pin</th>
<th>Int.</th>
<th>Pin</th>
<th>Int.</th>
<th>Pin</th>
<th>Int.</th>
<th>Pin</th>
<th>Int.</th>
</tr>
</thead>
<tbody>
<tr>
<td>275</td>
<td>250</td>
<td>275</td>
<td>250</td>
<td>275</td>
<td>250</td>
<td>275</td>
<td>250</td>
<td>275</td>
<td>250</td>
</tr>
<tr>
<td>725</td>
<td>700</td>
<td>725</td>
<td>700</td>
<td>725</td>
<td>700</td>
<td>725</td>
<td>700</td>
<td>725</td>
<td>700</td>
</tr>
<tr>
<td>225</td>
<td>200</td>
<td>225</td>
<td>200</td>
<td>225</td>
<td>200</td>
<td>225</td>
<td>200</td>
<td>225</td>
<td>200</td>
</tr>
</tbody>
</table>

*Table 7.1. The steady-state frequency positions of the first and second positions.*

Three formulas of the harmonic moment in initial, intermediate, and final positions.*
and with an intermediate value in initial position. Although the first formant similarly has the highest frequency position in intervocalic position, it is higher in final position than in initial position.

The same trend is again displayed by /w/ where the frequency positions of the first and second formants are higher in intervocalic position than in initial position, which are again higher than those in final position. No comparison could be made in the case of the third formant, however, since it is absent from the spectra of /w/ in both intervocalic and final positions.

In general, the frequency positions for the first three formants of the Arabic semivowels are higher in intervocalic position than in initial position, which are, in turn, higher than those in final position.

7.4.2. The effect of the adjacent vowel on the spectra of the semivowel.

Figures 7.1, 7.2 and 7.3 present the formant positions of the Arabic semivowels /j/, /q/ and /w/ as a function of the adjacent vowel. For all the semivowels, the figures show that F1 exhibits a tendency to vary directly with the frequency position of the first formant of the adjacent vowel. This tendency is much more pronounced in the case of F2, which also varies directly with the frequency position of
Figure 7.1. Frequency positions for the first three formants in the spectra of /w/ as a function of the following vowel.
Figure 7.2. The average frequency positions for the first three formants in the spectra of /i/ as a function of the following vowel.
Figure 7.3. The average frequency positions for the first three formants in the spectra of /j/ as a function of the following vowel.
the second formants of the adjacent vowel. Although F3 shows nearly the same trend of direct variation in the case of /j/, it is too often absent from the spectra of /q/ and /w/ for a trend to be detected.

7.4.3. Duration.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Intervocalic</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ʃ/</td>
<td>70</td>
<td>80</td>
<td>190</td>
</tr>
<tr>
<td>/q/</td>
<td>85</td>
<td>80</td>
<td>230</td>
</tr>
<tr>
<td>/w/</td>
<td>70</td>
<td>80</td>
<td>240</td>
</tr>
</tbody>
</table>

Table 7.2. Duration values for the Arabic semi-vowels in initial, intervocalic, and final positions (in milliseconds).

Table 7.2 presents the duration values for the Arabic semi-vowels /ʃ/, /q/, and /w/ in initial, intervocalic, and final positions. It is obvious that while the duration values of the Arabic semi-vowels vary little or not at all as a function of point of articulation, they vary considerably as a function of word position. In fact, as Table 7.2 shows, the Arabic semi-vowels have the same duration values in intervocalic position and very similar values in both initial and final positions. The effect of word position on their duration values is, by contrast, much more remarkable. In final position, thus, they are about three times as long as in initial or intervocalic positions.
In general, the Arabic semivowels are longest in final position, shortest in initial position, and with an intermediate value in intervocalic position.

7.4.4. Transitions.

Figures 7.4, 7.5, and 7.6 present, for each of the Arabic semivowels, the average steady-state frequency positions as well as the onset frequencies of the Arabic long vowels when they occur after the Arabic semivowels in word-initial and intervocalic positions. Figure 7.5 presents, in addition, the offset frequencies of the Arabic vowels before word-final /q/. These are not shown for /j/ or /w/ since they, unlike /q/, can occur finally in the syllable only after /a:/.

Steady-state frequency positions are represented by squares while onset and offset frequency positions are represented by circles.

In general, /j/ seems to have a slightly negative (and sometimes neutral) transitional influence on /i:/ and /e:/, and a positive transitional influence on the rest of the vowels. /q/, on the other hand, seems to have a negative transitional influence on the front vowels, and a positive transitional influence on the back vowels. In the case of /w/, by contrast, it seems to have a negative transitional influence on all following vowels, back as well as front ones. In intervocalic position, however, while it still has a negative transitional influence on the front vowels, it has a neutral transitional influence on /o:/ and a positive transitional influence on /u:/.
Figure 7.4. The average onset and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /w/ (left-hand side), and intervocalic /w/ (right-hand side).
Figure 7.5. The average onset, offset, and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial and intervocalic /i/ and before word-final /i/. 
Figure 7.6. The average onset and steady-state frequency positions for the second formants of the Arabic long vowels after word-initial /j/ (left-hand side) and intervocalic /j/ (right-hand side).
7.5. Conclusion.

The acoustic analysis of the Arabic semivowels reported in the previous sections has revealed the following findings.

First, spectra of Egyptian Arabic /q/ in initial, intervocalic, and final positions show no trace of the random scattering of noise characteristic of fricative consonants. By contrast, they show a clear formant structure very much like that of the Arabic vowels. These spectral properties do not support the traditional view that Egyptian /q/ is a pharyngeal fricative. Indeed, contrary to all accepted notions as to the nature of Egyptian /q/, the results of the acoustic analysis reported in this work have shown that it is, in fact, phonetically a vowel and not a fricative.

Second, spectra of /q/ differ from those of other Arabic vowels by the wide spacing between voice striations. This phenomenon has been observed by two previous studies on the phonetics of Arabic (Obrecht, 1968; Al-Ani, 1978). Neither has, however, attempted to offer a physiological reason for it. The present work is the first to show that this feature is not characteristic of /q/ spectra alone but also of spectra of creaky voice and to suggest, therefore, that the voice source used in the production of both /q/ and creaky voice is identical.
Third, the acoustic analysis of sustained utterances of creaky voice has revealed that both at the initiation and termination of this type of voice the size spacing between the striations tends to become still wider while the acoustic energy present in the striations tends to get more concentrated in the lower frequencies. While these initial and final striations may have the appearance of what we would normally call 'spikes' (that is, the acoustic correlate of the explosion phase of a stop), it was quite obvious that they are the result of a gradual build up of creaky voice at initiation and a gradual slow down at termination.

Interpreted in the light of this finding, the 'spikes' encountered at both the initiation of word-initial /j/ and the termination of word-final /ʕ/ were considered a characteristic of creaky voice and not an acoustic correlate of a stop explosion. The low-frequency high-intensity 'spikes' reported by Al-Ani (1978) to be present at the initiation and termination of word-initial and word-final Iraqi /ʕ/ would, from this point of view, be regarded as the result of a similar process rather than as the result of a voiceless stop explosion. And bearing in mind that Al-Ani has inferred rather than shown the presence of a closure phase of a voiceless stop at these two positions, and also that Iraqi /ʕ/ is the result of creaky voice, this suggestion seems to be the more likely.
In a more detailed vein, the results of the acoustic analysis of the Arabic semivowels reported in the preceding sections indicate:

1. The position of the semivowel in the word has an observed effect on its spectrum. For all the Arabic semivowels, the frequency positions of the first three formants are higher in intervocalic position than in initial position, which are in turn higher than those in final position.

2. The adjacent vowel has a considerable effect on the frequency positions of the formants of the semivowel. Both F1 and F2 vary directly with the frequency positions for the formants of the adjacent vowel. Variation of this kind is most observed in the case of F2 in particular. F3, on the other hand, is often absent.

3. The position of the semivowel in the word has a considerable effect on its duration. The Arabic semivowels are about three times as long in final position as in initial or intervocalic position. Although there is not much difference between their duration values in initial and intervocalic positions, the Arabic semivowels tend to be slightly longer in intervocalic position.

4. While /j/ exerts a negative transitional influence on /i:/ and /e:/ alone, and /j/ exerts a negative transitional influence on /i:/, /e:/, and /a:/, /w/ exerts a negative transitional influence on all following vowels.
Chapter 8

Conclusion

8.1. General discussion.

8.2. General spectral and temporal properties of the Arabic sounds.

8.2.1. The effect of word position on consonant spectra.

8.2.2. The effect of the adjacent vowel on consonant spectra.

8.2.3. The effect of the consonant on adjacent vowel spectra.

8.2.4. Factors affecting duration.

8.3. Suggestions for further work.
CONCLUSION

This chapter is divided into three main parts. The first evaluates the relevance of some of the findings on the acoustic of Arabic speech to general theories of speech production and perception. The second, on the other hand, attempts to provide an overall view of the specifics of the Arabic speech sounds dealt with separately in the previous chapters. The third, and last, contains a few suggestions for further study.

8.1. General discussion.

The results of the acoustic analysis of the Arabic speech sounds reported in the previous chapters reveal, as expected, that their spectral properties have much in common with the predicted and observed spectra of sounds produced with similar mechanisms in other languages.

First, the acoustic analysis of the Arabic fricatives - in particular, the fact that acoustic energy in their spectra is organized in well-defined bands, and the general pattern of distribution of these bands of energy among the frequencies for the different fricatives - support Fant's suggestion (1960) that "the typical fricative is a noise sound, the spectral energy of which is largely contained in formants from cavities in front of the articulatory narrowing."
The results of the acoustic analysis of the Arabic fricatives also support the observation made by Hughes and Halle (1956) and Heinz and Stevens (1960) on the presence of an inverse relationship between the length of the cavity in front of the fricative obstruction in the vocal tract and the frequency position of the most prominent energy-density maximum in fricative spectra. Furthermore, they also show that this relationship holds between the voiced as well as the voiceless fricatives and extends beyond the front fricatives to the back fricatives of Arabic as well.

Second, the spectra of Arabic /r/ bear out the observation made by Brosnahan and Malmberg (1970) that the trills are acoustically characterized by "an uneven, though not random distribution of energy over a wide band of frequency." Both in initial and final positions, spectrograms of Arabic /r/ do, indeed, show this uneven distribution of energy at frequencies above 2500 Hz.

Again, spectrograms of Arabic /r/ in intervocalic position illustrate very clearly Brosnahan and Malmberg's point (1970) that the distinction between a tap and a stop can sometimes be made "on the basis of the briefness or incompleteness of a flap articulation in comparison with a plosive articulation." The spectra of Arabic /r/ in intervocalic position are virtually the same as those of a cognate plosive, and it is mainly in terms of briefness of closure duration that a distinction can, indeed, be made.
Third, the acoustic analysis of the Arabic nasals has revealed that they always have a first formant at about 250 Hz. This agrees with Fant's prediction (1960) that the transmission characteristics of an acoustic system simulating the pharynx-nose tract, with no oral side-chamber, will result in a formant at about 300 Hz.

The acoustic analysis of the Arabic nasals has also revealed that, in those cases where the Arabic nasals show more than one formant, the average spacing between the formants is, indeed, the same as that specified by Fujimura (1962), that is, about 800 Hz.

Again, and in agreement with the findings of Potter, Kopp and Kopp (1947), the number of formants of the Arabic nasals has been found to vary with the degree of stress. Analysis of overall intensity of the Arabic nasals has shown that they are considerably more intense in initial and intervocalic positions than in final position. This accounts for the fact that the Arabic nasals usually display three formants in the initial and intervocalic positions and only one formant in final position.

The results of the acoustic analysis of the Arabic nasals also support the assertion made by Hattori, Yamamoto and Fujimura (1958) about the existence of an essentially discontinuous transition between the nasal and the vowel. Spectra of Arabic nasals show an abrupt change of timbre at the moment of release of the nasal.
The acoustic analysis of the Arabic nasals has further shown that as expected, the frequency positions for the first two formants of [n] are, in general, higher than those for [m]. However, as Brosnahan and Malmberg (1970) have noticed, the formant frequencies of these two sounds are very similar and very often overlap.

The acoustic analysis of Arabic speech has however turned out some unexpected results. First, the results of our perceptual tests on fricative identification are in disagreement with those reported by Harris (1956) and Heinz and Stevens (1961). Both studies have concluded that while the important cues for /s/ and /f/ are given by the friction part, those for /f/ and /θ/ are contained primarily in the transition part of the syllable. Our perceptual tests have, by contrast, shown that both /f/ and /θ/ could be discriminated with the aid of cues in the friction part alone.

The success of the listeners in our experiment to correctly identify the various fricative stimuli can probably be attributed to the fact that these stimuli were spoken rather than synthesized. This is, no doubt, an indication of the frailty of the cues in the friction part which seem to be difficult to reproduce faithfully on the one hand, or to be easily damaged in a process of cutting and recombination on the other. This frailty seems to support the common belief that the cues in the transition part are the overdominant
ones. More importantly, the discrepancy between the results of natural speech perceptual tests and those of synthetic speech is a warning against too readily accepting synthetic speech as a true copy of natural speech.

The acoustic analysis of the Arabic nasals has also revealed two surprising results. The first relates to the fact that the frequency position for the second formant of [n] has been found to be higher than that for [m]. The raising of the second formant of the pharyngealized nasal is, undoubtedly, unexpected. One would have expected n, or, for that matter, any pharyngealized sound, to have a lowered second formant and not a raised one.

The second, on the other hand, relates to the fact that the first formant of [m] has been found to be lower than that for [n]. Fant (1960) has suggested that the low F1 of nasal consonants represents the fundamental resonance of the pharynx cavity tuned by the nasal system as a resonator neck. In terms of this analysis one would expect — due to the constricted and hence smaller pharynx cavity — the pharyngealized bilabial nasal to have a higher F1 than that of the non-pharyngealized bilabial nasal. As the acoustic analysis of Arabic nasals shows, however, this is not the case.
The acoustic analysis of the Arabic semivowels has similarly revealed some interesting findings which have an immediate relevance to Arabic phonetics in addition to their intrinsic value to our understanding of speech sounds in general.

First, spectra of Egyptian Arabic /ɣ/ in initial, intervocalic, and final positions have been found to show no trace of the random scattering of noise characteristic of fricative consonants. They show, by contrast, a clear formant structure very much like that of the Arabic vowels. These spectral properties do not support the traditional view that Egyptian /ɣ/ is a pharyngeal fricative. Indeed, contrary to all currently accepted notions as to the nature of Egyptian /ɣ/, the results of the acoustic analysis reported in this work have shown that it is, in fact, phonetically a vowel and not a fricative.

Second, spectra of /ɣ/ differ from those of other Arabic vowels by the wide spacing between voice striations. This phenomenon has been observed by two previous studies of the phonetics of Arabic (Brecht, 1968; Al-Ani, 1978). Neither has, however, attempted to offer a physiological reason for it. The present work is believed to be the first to show that this feature is not characteristic of /ɣ/ alone but also of creaky voice and to suggest, therefore, that the voice source used in the production of both is probably identical.
Third, the acoustic analysis of sustained utterances of creaky voice has revealed that both at the initiation and termination of this type of voice the wide spacing between the striations tends to become still wider and the acoustic energy present in the striations to get more concentrated in the lower frequencies. While these initial and final striations may have the appearance of what we would normally call 'spikes' (that is, the acoustic correlate of the explosion phase of a stop), it was quite obvious that they are the result of a gradual build up of creaky voice at initiation and a gradual slow down at termination.

Interpreted in the light of this finding, the 'spikes' encountered at both the initiation of word-initial /ɣ/ and the termination of word-final /ɣ/ were considered a characteristic of creaky voice and not an acoustic correlate of a stop explosion. The low-frequency high-intensity 'spikes' reported by Al-Ani (1978) to be present at the initiation and termination of word-initial and word-final Iraqi /ɣ/ would, from this point of view, be regarded as the result of a similar process rather than as the result of a voiceless stop explosion. And bearing in mind that Al-Ani has inferred rather than shown the presence of a closure phase of a voiceless stop at these two positions, and also that Iraqi /ɣ/ is the result of creaky voice, this suggestion seems to be the more likely, and worthy of independent investigation.
8.2. General spectral and temporal properties of Arabic sounds.

The effects of coarticulation and word position have been shown to determine both the spectral and temporal characteristics of the Arabic consonant sounds as well as the spectral properties of the Arabic vowel sounds.

In this section we shall first deal with the variation of the spectral characteristics of the Arabic consonant sounds as a function of word position and the adjacent vowel. We shall then discuss how the vowel spectrum is itself affected by the adjacent consonant and word position. Lastly, we shall examine the factors that determine the temporal properties of the Arabic consonant sounds.

8.2.1. The effect of word position on consonant spectra.

The acoustic analysis of the Arabic consonant sounds has shown that their spectral characteristics are significantly determined by their position in the word.

The spectrum of an Arabic consonant sound varies as a function of word position in one of three ways. First, it may show a partial or total loss of one or more of its components. Second, it may show a change in the frequency positions of its components. And, third, it may show a radical change of its properties with the result that we have a totally different spectrum.
The first type of variation is illustrated by the Arabic /r/, the voiced fricatives, as well as the voiced stops. In initial position, these sounds are characterized by the presence of a low-band component which appears along the baseline at or below 300 Hz. In final position, by contrast, this component is often absent, either partly or totally.

The spectra of the Arabic voiced stops in final position are further characterized by the presence of a period of noise of considerable duration following the stop closure. In final position, by contrast, this period of noise is absent.

Similarly, the spectra of the Arabic nasals in initial and intervocalic positions characteristically display three formants. In final position, however, they show only one.

The second type of variation is shown by all the Arabic consonant sounds studied in this work. As a rule, the frequency positions of the first three formants of an Arabic consonant sound are higher in intervocalic position than in initial position, which are, in turn, higher than those in final position.

The third type of variation is illustrated by the Arabic /r/. In initial position, acoustic energy is organized in the form of clear and well-defined formants. In intervocalic position, on the other hand,
there is little or no acoustic energy above 250 Hz. In final position, by contrast, acoustic energy is present in the form of randomly scattered noise.

Both the first and third types of variation could be immediately related to such articulatory features as voicing, aspiration, stress, and manner of articulation. The physiological correlates of the second type of variation are not, however, clear to me. Variation in the frequency positions for the formants of a speech sound as a function of word position have been reported previously. Potter, Kopppand Kopp (1947) for example, have observed that the frequency positions for the first and second formants of the American /r/ in initial position are lower than those in final position. No physiological reasons for such a phenomenon have, however, been advanced.

8.2.2. The effect of the adjacent vowel on the consonant spectra.

The spectra of the Arabic consonant sounds have been found to show varying degrees of variation with the quality of the adjacent vowel. Such variation is most observed in the case of the second formant and, although to a much lesser extent, in the case of the first formant. The third formant, by contrast, either shows some unsystematic form of variation or stays unaltered.
For most of these sounds the frequency position of the second formant varies as a direct function of the frequency position for the second formant of the adjacent vowel. In every case, however, the second formant transitions all point to a common source, the consonant locus.

Only in the case of the stop consonants /k/ and /g/ does the quality of the adjacent vowel affect the spectrum of the consonant to such an extent that we have more than one locus. Thus, in the vicinity of the front vowels, the frequency position for the /k, g/ locus is about 2200 Hz. In the vicinity of the back vowels, by contrast, it is about 1000 Hz.

In the case of the Arabic lateral /l/ and the lingual fricatives /s, z, s, z, j/, however, the quality of the adjacent vowel does not seem to have a significant influence on their spectra. There is, on the contrary, a remarkable degree of resistance to coarticulation.

It was suggested (Chapter Two) that the absence, or near absence, of coarticulation between the lingual fricatives /s, z, s, z, j/ and the adjacent vowel is due to the constraints on the movement of the front of the tongue. The remarkable coarticulation observed in the case of the fricatives /f, x, y, h, h/, on the other hand, was attributed to the absence of such constraints.
Such a suggestion assumes, obviously, that the front of the tongue is more able to coarticulate with the adjacent sound than the back of the tongue. And, indeed, the results of the acoustic analysis of the Arabic stops (Chapter Three) provide further evidence in support of this suggestion. By contrast with the front stops /b, t, s, d, j/, the back stops /k, g/ are characterized by the greater durations of their aspiration and closure phases, as well as by the limited range of variation for the duration of their closure phases. All of which phenomena are attributed to the sluggishness of the back of the tongue. The absence, or near absence, of coarticulation between the lateral and the adjacent vowel would, similarly, be attributed to the participation of the front of the tongue in its production.

8.2.3. The effect of the consonant on the adjacent vowel spectra.

Our analysis of the spectral properties of the Arabic vowels has shown that both the direction and extent of their second formant transitions are significantly affected by the adjacent consonant.

Both the direction and extent of the vowel F2-transitions are also significantly affected by the position of the vowel relative to the consonant. In general, the direction of F2-transition tends to be neutral more when the vowel occurs before the consonant than after it. The extent of F2-transition, on the other hand, is greater when the vowel occurs after the consonant than before it.
Both these two phenomena indicate that the coarticulatory effects of the Arabic consonant sounds are more powerful in a consonant-vowel context than in a vowel-consonant one.

8.2.4. Factors affecting duration.

The duration of the Arabic consonant sounds has been found to vary as a function of voicing, word position, and point of articulation. First, for all the Arabic consonant sounds, the voiceless member of a homorganic pair has consistently greater duration values than its voiced counterpart.

Second, the duration of the Arabic consonant sounds varies as a function of word position. With the exception of /r/, the difference between the duration of an Arabic consonant sound in initial position and that in intervocalic position is slight. In final position, by contrast, the duration values for all the Arabic consonant sounds are consistently more than twice as long as in initial or intervocalic position.

And, third, the farther back in the vocal tract a voiceless stop or fricative is produced the greater its duration. This relationship, however, does not seem to exist in the case of the voiced stops and is reversed in the case of the voiced fricatives.
8.3. Suggestions for further work.

1. A study of fricative spectra in other languages with a sizable number of fricatives to verify our findings on the presence of an inverse relationship between the length of the cavity in front of the fricative obstruction and the frequency position of the most prominent energy-density maximum in fricative spectra.

2. A study of spoken stimuli of steady-state fricatives in other languages as compared with recorded and/or synthesized stimuli.

3. Harris (1956) has found that while /sl/ and /s/ could be identified on the basis of cues in the friction part alone, identification of /f/ and /θ/ relied primarily on the vowel. No physiological reason was, however, advanced.

The results of the acoustic analysis of the Arabic fricatives suggest that this might be due to the fact that, by contrast with the non-lingual fricatives - for example, /f/ - both /sl/ and /s/ coarticulate far less with the adjacent vowel and, as such, are independent of the vowel context. Such a hypothesis is well worth testing.

4. A physiological investigation of the precise role of both the pharynx and larynx in the production of Arabic /q/.
6. More work on the acoustics of Arabic /\gamma/ in other dialects of Arabic. In particular, whether the low-frequency high-intensity 'spikes' in the spectra of word-initial and word-final Iraqi /s/ denote the presence of voiceless stops or whether, as suggested in this work, they are a property of the initiation and termination of creaky voice.

7. An attempt should be made to explain the variation in frequency position of the formants of speech sounds as a function of word position.

8. A study of fricatives in other languages to test whether there are differences in the degree of coarticulation between fricatives produced with the front of the tongue and other fricatives, of the sort reported in this work.

9. A study of duration variation as a function of word position and whether voicing affects such a relation, and if so why.
Bibliography


257


Fant, G. (1956) "On the predictability of formant levels and spectrum envelopes from formant frequencies," in For Roman Jakobson. The Hague, Mouton


Fant, G. (1982) "Descriptive analysis of the acoustic aspects of speech," Logos. 5:3


Harris, K. S. (1958) "Cues for the discrimination of American English fricatives in spoken syllables," Language and Speech. 1:1


Lisker, L. (1957) "Closure duration and the intervocalic voiced-voiceless distinction in English," Language. 33:42

Lisker, L., and A. S. Abramson. (1964) "A cross-language study of voicing in initial stops," Word. 20:384

Lisker, L., and A. S. Abramson. (1971) "Distinctive features and laryngeal control," Haskins Laboratories Status Report on Speech Research. 27:133


Nestell, R. (1969) "Subglottal and intraoral air pressure during the intervocalic contrast of /t/ and /d/," Phonetics. 26:68

Nord, L. (1976) "Perceptual experiments with nasals," Speech Transmission Laboratory, Quarterly Progress and Status Report. 2


Schatz, C. D. (1954) "The role of context in the perception of stops," Language. 30:47


Tarnoczy, I. (1948) "Resonance data concerning nasals, laterals and trills," Word. 4:71