Experimental study of emphasis and voicing in the plosives of Yemeni Spoken Arabic with some implications for foreign language teaching and learning.

BY

ABDULRAHMAN MOHAMED AL-NUZAILI

Thesis submitted in accordance with the requirements for the degree of Ph.D.

THE UNIVERSITY OF LEEDS
DEPARTMENT OF LINGUISTICS AND PHONETICS
December, 1993.

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Contents

Page

PART ONE: PHONOLOGICAL BACKGROUND AND SURVEY OF LITERATURE

Chapter 1: Phonetic and Phonological Background of MSA and YSA

Chapter 2: Literature Survey for Plosives:

2.2 Survey of perceptual studies on plosives with reference to emphasis and voicing. ✓

2.3 Survey of articulatory and aerodynamic studies of voicing and emphasis in the plosives of Yemeni Spoken Arabic:

PART TWO: EXPERIMENTAL INVESTIGATION.

Chapter 3: Acoustic characteristics of emphasis and voicing in YSA.

3.2 Acoustic study of the formant transitions for the emphatic and non-emphatic plosives in YSA.

3.3 Other temporal characteristics of plosives in YSA besides VOT.

Chapter 4: Perceptual characteristics of voicing and emphasis in YSA.

4.2 Perceptual study of emphasis in Yemeni Spoken Arabic.

Chapter 5: Articulatory and aerodynamic characteristics of emphasis and voicing in YSA.

5.3 Results for experiment 2: Aerodynamic patterns for the voicing categories in YSA.

Chapter 6: Summary and Conclusions.

Appendices:

Abstract

This is an experimental study of two major distinctive features: emphasis and voicing in the plosives of Yemeni Spoken Arabic. It investigates some of their acoustic, perceptual and arodynamic correlates and aims, at least in part, to find the language-specific aspects in as far as these phonetic phonemena are concerned. It falls into two related parts.

Part One consists of two chapters. **Chapter one** gives a general background of YSA with special reference to the phonemic significance of emphasis and voicing in the plosives and their interaction with various contextual factors and positions. Phonological definitions of these features are given. Various theoretical approaches are also dealt with. The syllable structure and the stress patterns in both Modern Standard Arabic and Yemeni Spoken Arabic are presented. **Chapter two** reviews critically some of the hypotheses and interpretations of voicing mechanisms and the factors affecting their realizations in various languages. Some of the relevant aspects reviewed are voice onset time in various languages, formant transitions, closure durations, temporal relationship between consonants and vowels, categorical perception and the phoneme boundary, aerodynamic factors and their role in the production of plosives. The two features are also reviewed in relation to vocalic context, place of articulation, stress, gemination and phonetic position.

Part Two consists of four chapters representing the main body of this study. **Chapter three** is an investigation of the acoustic characteristics of the voiced/voiceless and emphatic/nonemphatic categories in words embedded in a contextual frame sentence. **Chapter four** is a perceptual investigation of the above contrasts by means of synthetically generated speech using the Klatt Synthesizer. It examines the role played by VOT, the relative onset time between the release and the onset of voicing, in the accurate identification of the voicing cognates. Another experiment attempts to evaluate the role of the second formant particularly its onset frequency and steady state portion in the emphatic/nonemphatic distinction. The relationships between perception and production are described and the theory of 'categorical perception' in relation to our data is also discussed.

Chapter five investigates aerodynamic patterns and aerodynamically derived estimates of articulation for the emphatic/nonemphatic and the voiced/voiceless consonants in two experiments. Since there are several variables involved in this investigation, the results in both experiments are subjected to analyses of variance to obtain the effects of the independent variables on the dependent ones.

In **chapter six** the findings of the previous three chapters are summarized. Some implications for foreign language teaching and learning are also discussed. The study ends with a section on the limitations and suggestions for future research.

Acknowledgements

I am deeply indebted to my research supervisor Dr. C. Scully for her constant guidance, friendliness and encouragement during the whole period of this research. Her suggestions and comments on the various parts of this thesis are invaluable and I have benefited from them greatly.

I would like to thank all members of staff and fellow research students in the Department of Linguistics and Phonetics who in one way or another contributed to my training and education. Thanks are also due to the departmental secretaries for their kindness and assistance with all matters related to adminstrative arrangements.

I also acknowledge with gratitude the kind help and advice of Eric Brearley (Chief technician) during the experimental work and the preparation of the experimental set-up in the Laboratory. Thanks are also due to all members of staff in the Speech Lab in the Department of Psychology for allowing me to use their facilities during the preparation of the perceptual tests and for providing me with a study room there during that period of my reseach.

My thanks are also extended to Professor P. J. Roach, of the Department of Psychology, Dr. Nawal Ghali, of the Speech Lab. and Mr. Paul Nicolson (Chief Statistician in the University of Leeds Computing Service) for helping me with various aspects of the statistical analyses particularly in the creation of data files on the SPSS package and the treatment of the (ANOVA) results presented in chapter 5.

Special gratitude and appreciation are also due to my wife and children for their patience, understanding and support all through the years of this research.

Last but not least, my sincere thanks are extended to the British Council for covering the tuition fees and all other financial aspects related to this scholarship. I am particularly grateful to Pam Simpson, my programme officer in the British Council, for helping me in every way she could. Without their help this research would not have been materialized.

ABDULRAHMAN M. AL-NUZAILI

List of Tables and Figures

Abbreviations and definition of terms

Acoustic cues: the necessary acoustic patterns of speech used by the listeners for deciding between sounds, words or sentences.

A : area of constriction.

AH: aspiration noise.

ANOVA: analysis of variance.

Av.: average.

B.C: beginning of contoid.

B.V: beginning of vocoid.

C: consonant.

CC: a long consonant or a geminate consonant.

c. circa. : approximately.

C.A.: Classical Arabic.

CF: citation forms.

cf.: compare.

Ch.: channel.

CL: acoustic closure.

cm/s: centimeter per second.

Cong.: congress.

Conf.: conference.

Cont.: contoid.

CS: connected speech.

dB: decibel.

E: emphasis.

(eds.): editors.

(ed.): editor.

et. al.: and the others.

f. : female.

F: frequency.

F0: the fundamental frequency of voicing.

F1: first formant.

F2: second formant.

Fric: frication noise.

F-test: the name given to a statistical test after the great mathematician Fisher who invented it. It is particularly useful for the analysis of variance (ANOVA). It differs from the t-test in that it allows the comparisons of several group means simultaneously. More generally the F-test can be obtained by calculating the ratio of the two sources of variability as follows: (between group variance over within group variance).

$$F_{obs} = S^2 \text{ between}/S^2 \text{ within.}$$

H.P: High pass filter.

Hz: Hertz.

JND: Just noticeable difference.

J.: Journal.

ibid. In the same reference.

I.L.: Intensity level.

imp. : imperative.

inter.: international.

I.P.A: International Phonetic Association.

k : an empirical constant.

kHz: kilo hertz.

loc. cit. in the page mentioned.

L.P: low pass filter.

L/s.: Liter/second.

Lx: Laryngograph signal.

M.I.T: Massachusetts Institute of Technology.

MSA: Modern Standard Arabic.

msec: millisecond.

Mingo: Mingograph.

M: mean.

N.B.: narrowband spectrogram.

(n. d.): no date specified.

op. cit.: in the reference mentioned.

p.: page

P: probability.

Pd : pressure duration.

PERILUS: Phonetic Experimental Research, Institute of Linguistics, University of Stockholm.

pl.: plural.

Po. : peak oral air pressure (cm H^2O).

Δp : pressure drop across the tongue constriction.

Proc: Proceedings.

PR: progress report.

R: release transient.

Sgm: spectrogram.

S. dev.: standard deviation.

SPL: speech laboratory report.

Sp: sound pressure signal.

SR : status report on speech research. Haskins Laboratories.

QPR-RLE: Quarterly Progress Report of the Research Laboratory of Electronics Massachusetts Institute of Technology (M. I.T.).

STL - QPSR: Speech Transmission Laboratory, Quarterly Progress and Status Report.

t-test: a statistical test used to show the significance level of the difference between the means of two samples.

Uo : peak oral airflow (cm^3/sec).

Uc : articulatory closure duration.

Ur : articulatory release duration.

VOC.: vocoid.

Vol.: volume.

VOT: voice onset time.

vs: versus.

VV.: long vowel.

W.B: wideband spectrogram.

YSA: Yemeni Spoken Arabic.

< less than.

> more than.

≤ equal or less than.

≥ equal or more than.

[ ]: phonetic unit.

/ /: phonemic unit.

' ': orthographic unit.

* : statistically significant.

** : Highly significant.

(`) : primary stress.

(′) : secondary stress.

Note : cm H2 O is used to mean cm H O.

Journals:

Cleft Palate Journal.

Folia Phonetica.

Janua

Journal of the Acoustical Society of America (JASA).

Journal of the American Oriental Society.

Journal of Experimental Psychology (J. Exp. Psych.).

Journal of Phonetics (J. Phon.).

Journal of Semitic Studies.

Journal of Speech and Hearing Research (JSHR).

Journal of Speech and Hearing Disorders.

Language.

Language and Speech.

Miscellanea Phonetica.

Moslem World.

Perception and Psychophysics (Per. and Psychphys.)

Phonetica.

Psychological Review (Psycho. Rev.).

Psychological Monograph.

Studia Linguistica.

Tarbiz.

TESL = Teaching English as a Second Language.

Word.

Reading Transcription

The phonetic symbols employed in this study are listed below. They are used to facilitate the reading of the transcribed words. The emphatic consonants are distinguished from their non-emphatic counterparts in that they are represented by capital letters. IPA symbols are used in almost all other situations.

Consonants:

/b/ voiced bilabial plosive

/t/ voiceless denti-alveolar nonemphatic plosive

/T/ voiceless denti-alveolar emphatic plosive

/d/ voiced denti-alveolar nonemphatic plosive

/D/ voiced denti-alveolar emphatic plosive

/k/ voiceless velar nonemphatic plosive.

/g/ voiced velar plosive.

/q/ voiceless uvular emphatic plosive

/ʔ/ glottal stop.

/f/ voiceless labio-dental fricative.

/θ/ voiceless interdental fricative

/ð/ voiced interdental fricative

/δ/ voiced interdental emphatic fricative

/s/ voiceless denti-alveolar nonemphatic fricative.

/S/ voiceless denti-alveolar emphatic fricative.

/z/ voiced denti-alveolar fricative.

/ʃ/ voiceless palato-alveolar fricative.

/dʒ/ voiced palato-alveolar affricative.

/x/ voiceless uvular fricative.

/ɤ/ voiced uvular fricative.

/ħ / voiceless pharyngeal fricative.

/ʕ/ voiced pharyngeal fricative.

/h/ glottal fricative.

/m/ voiced bilabial nasal.

/n/ voiced denti-alveolar nasal.

/l/ voiced alveolar-palatal lateral.

/L/ voiced alveolar emphatic lateral.

/w/ voiced labial-velar semi-vowel.

/j/ voiced palatal semi-vowel.

/r/ voiced alveolar flap.

Vowels:

/ii/ a long close front unrounded vowel.

/i/ a short close to half-close front unrounded vowel.

/ee/ a long half-close to half-open front unrounded vowel.

/aa/ a long open front unrounded vowel.

/a/ a short half-open to open front unrounded vowel.

/ oo/ a long half-close to half-open back rounded vowel.

/uu/ a long close back rounded vowel.

/u/ a short close back rounded vowel.

Introduction and Aims

Modern Standard Arabic (MSA) has been and still is the first language of all the Arab speaking countries for the last 15 centuries. Like many other languages it was under constant pressure for change apart from being a source of influence on other languages during all that period. Deviations from the standard, as a result, actually started to emerge in different parts of the Arabic Speaking world. For example, a comparison between the Yemeni Spoken Arabic variety and those of other Arab countries (e.g. Algerian, Egyptian, or Iraqi) would show the extent of such variations at all levels: phonetic, syntactic and semantic. Several factors contributed to that deviation from the MSA. Some of them were political ; others were religious or economic. Politically, the effect might be attributed to, at least, two factors: (1) the foreign invasions such as the three Turkish invasions of North Yemen and the British occupation of South Yemen and (2) differences in the ideological systems that separated one Arab country from another. Religiously, mutual linguistic influences occurred between Arab-Muzlims and non-Arab-Muzlims. Consequently, each Arab country started to develop its own Arabic spoken variety. Even inside each country various accents started to emerge. Thus, in Yemen one could often hear division such as Sana'ani accent, Adani accent, Hodeidah accent and HaDramauti accent in the north, south, west and east of Yemen respectively. Linguistically, however, only two varieties were said to be found in Yemen (Bravmann Meïr, 1942). From research obtained during his stay in the Yemen , Bravmann distinguished an east Yemeni variety spoken in Sana'a and the variety of the coastal areas spoken in Hodeidah. This classification might be extended to include the /gultu/ variety of Sana'a and most of the north east region where the MSA consonant /q/ is replaced by /g/ in this variety. The /qultu/ variety of Aden and most of the south-west region where the MSA consonant /q/ is kept unchanged in this variety. These are only minor variations and both Yemeni varieties share many common features which collectively characterize Yemeni Spoken Arabic. Some of these common features of YSA will be the topic of this study.

In particular, the study will aim at investigating the acoustical, perceptual, articulatory and aerodynamic characteristics of plosive consonants in YSA. Their examination will be related to some important phonological features namely that of emphasis and voicing. Other phonetic and phonological variables will be touched upon. It is hoped that the study of such features will provide further evidence for the distinctive feature theory. The theory of distinctive features is not only useful for the description of structural linguistics and the development of a system free of redundancy, but also it is important in determining the phonemic characteristics and information which are relevant for speech communication and speech technology. Of all the phonological oppositions in all the languages of the world, the voicing distinction between cognate plosive consonants is probably the most frequently used. The complexity of the mapping of a large number of acoustic parameters onto a variable of voicing categories makes the opposition an essential test case for the theory. Since this theory awaits to be tested in many languages, this study will be a step towards that goal. The findings, it is hoped, will provide a basis by which cross-language comparisons can be made particularly between English and YSA. The marrying of physical events and underlying linguistic units challenges rigid notions of universal phonetics. It urges for the role of language specific learning of phonetic realization rules and hence for the need to see parameters as constituting just as essential a part of the native speaker's systematic knowledge as does knowing the correct bounding structures in syntax or how to construct syllables in phonology. From an educational point of view, these rules may be incorporated as an essential part of pronunciation teaching. They may facilitate the teaching and learning of various cross-linguistic phenomena.

Other reasons for undertaking a research on this topic can be summarized as follows. First, the plosive consonants as a class of sounds perform multiple grammatical functions in YSA. Thus, the consonant /t/ (1) indicates gender when occurring at the end of nouns and in non-pausal forms such as /ʃadʒaratʊn/ 'a tree' or /ʃadʒarah/ where the /t/ is realized as /h/ in pausal forms; (2) it also indicates the second person singular, that acts as the subject, particularly when joined with certain classes of verbs as in /laʕibt/ 'you played'; (3) it is also used to distinguish the past from the present tenses. It indicates the present when it occurs at the beginning of verbs such as /taʕmal/ 'she works' or 'you work', and the past tense when it occurs word-finally such as /ʕamalt/ 'I worked'.

Secondly, the particle /ba/ in YSA is used to indicate the future as in this example /ba nizuurhum/ 'we will visit them'. Conversely, the particle /qa/ is used to indicate the past in YSA words as in /qa xallaS/ ' he had finished doing something'.

Thirdly, the plosives in MSA have been found to be the second most frequent group of sounds, particularly the denti-alveolars, after vowels (Aniis, 1971). From a phonological point of view, they are reported to carry more functional load than the others in terms of the different distinctive oppositions that a given phoneme can form with others (RaHiim, 1980). It is therefore our belief that a further investigation of this group of sounds on the lines described above will reveal more insights into their characteristics and how they may interact with others in the contexts determined by the selection of material for this study.

Linguistic origins of Modern Standard Arabic (MSA) and Yemeni Spoken Arabic (YSA).

Yemeni Spoken Arabic is historically and geographically considered to be a 'semitic' language. The term 'semitic' is a general designation of a group of languages that have in common several linguistic characteristics. The term was coined by the German scholar Schlöezer during the 19th century and is derived from the name Shem, son of Noah. Semitic languages are divided into three main branches (Rabin, 1971; FrayHa, 1981). Figure 1.1 below shows the relationship between these languages and the area in which each one of them was, or still is, spoken. It can be seen that YSA is one of those Arab dialects that came out of the Northern Arabic Language.

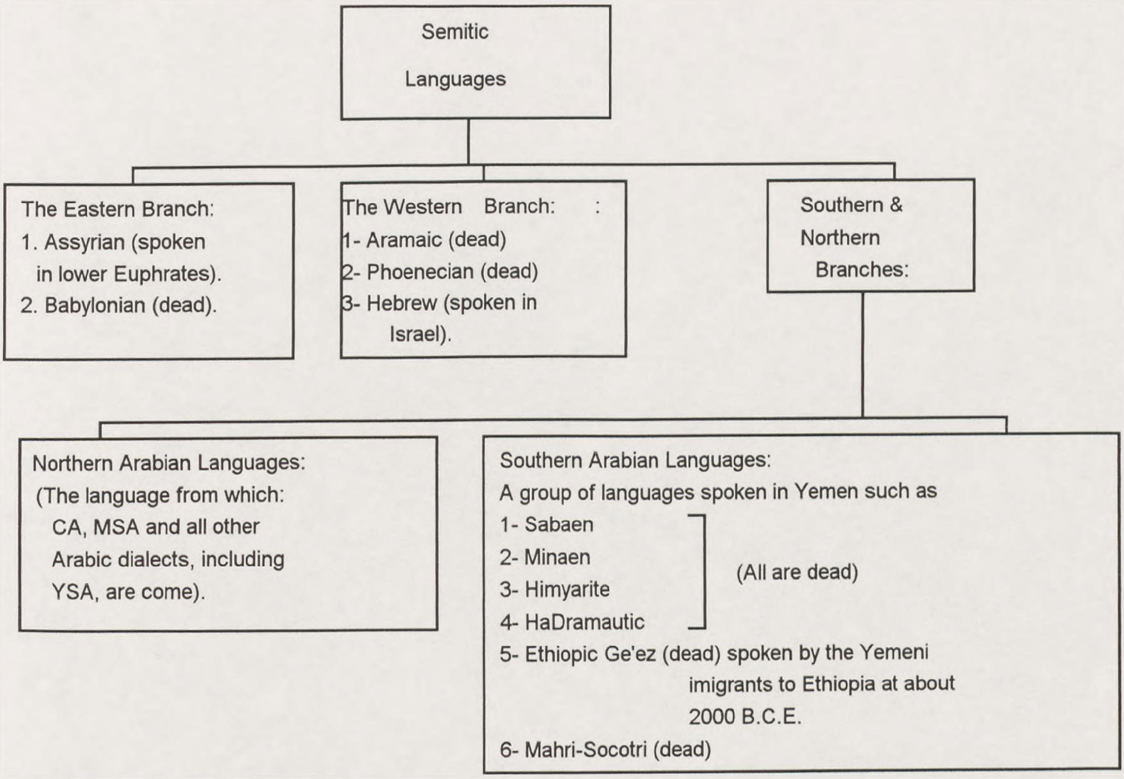


Figure 1.1 A tree-diagram of the Semitic Language Family.

PART ONE: PHONOLOGICAL BACKGROUND AND SURVEY OF LITERATURE

Chapter One: Phonetic and phonological background of MSA and YSA

1.1Voicing:

1.1.1 Voicing and the classification systems of consonants in MSA.

Phonetically, voicing is thought to be the result of the frequent modulation of air that causes the vocal folds to vibrate. According to Daniel Jones (1976) most ordinary speech sounds are produced with either "breath" or "voice". Those which contain "breath" are called voiceless and those that contain "voice" are called voiced.

The first systematic categorization of Arabic sounds was that done in Al-kitaab 'The Book' by the distinguished Arab linguist Sibawayhi (1966) who died late in the eighth century. He was a prominent member of the old school of BaSrah in Iraq. In that classification the consonants were divided into various groups by a system of phonetic features which described, among other things, the place and manner of articulation. One of the most general features was the opposition between /majhuura/ which was assumed to correspond to 'voiced' and /mahmuusa/ to 'voiceless'. The definitions given for each of these categories are as follows:

> " The majhuur is a sound in which pressure at the point of articulation is strong, and breath /nafas/ is confined or prevented from flowing until contact ends, and the voice has occurred. As for the 'mahmuusa', it is a sound in which pressure at the point of articulation is weak or weakened /aʊDʕifa/ so that breath occurs with it. You can tell that this is so if you try to repeat a sound together with the flow of breath, if you attempt this with a /majhuur/ you will not succeed" (Sibawayhi, 1966, Vol. I).

Accordingly, Sibawayhi includes among the majhuura the following sounds "ʔ, alif /a/, g, q,̥dʒ, j, D, l, n, r, T, d, z, δ, ð, b, m, w" and among the mahmuusa group he includes: " h, ħ, x, k, ʃ, S, s, t, θ , f". Since, then, this classification was elaborated and practically re-used by subsequent grammarians without significant modifications.

Before attempting to review the different interpretations of Sibawayhi's definition, it may be useful to consider these points:

1. It appears from Sibawayhi's definitions that voicelessness or breath 'nafas' is correlated with the 'mahmʊʊsa' phonemes, and voice 'Sawt' is correlated with the majhuura phonemes.

2. Sibawayhi regards 'dʒahr' as the result of the strong pressure at the point of articulation and the 'hams' as the weakening of this pressure, but does not mention the source from which it comes.

3. He does not refer to the role of the vocal folds in the distinction between these sounds.

Because of that modern phoneticians have certain reservations about Sibawayhi's classification and the ambiguity of his definition. The source of ambiguity is not only in the inclusion of /q, T, ʔ/ among the majhuura consonants but also for the different semantic interpretations the terms 'majhuura' and 'mahmuusa' may be given. The former, for instance, is described as 'voiced', 'lenis', 'pressed', 'non-breathed', and 'sonorous', and the latter as 'voiceless', 'fortis', 'non-pressed', 'breathed' and 'muffled'. Thus, Blanc (1967) perceives the contrast 'majhuura/mahmuusa' to be between 'sonorous' and 'muffled'. His justification is that "The mahmuusa are thus voiceless and with 'breath' by definition, but the majhuura are not necessarily voiced. They merely lack voiceless breath i.e. may either have voiced breath (voice) or lack of breath altogether".

Garbell (1958), on the other hand, sees the contrast as one between 'breathed' and 'non-breathed'. According to him the breathed group (mahmuusa) will include everything that is voiceless such as voiceless fricatives, and voiceless slightly aspirated plosives. The non-breathed group (majhuura) will include everything that is voiced such as the voiced fricatives and plosives without breath including /q,T/; sonorants and semi-vowels.

This grouping seems to fit Sibawayhi's description, but it is not in terms with modern description. According to Daniel Jones (1976) the term 'breathed' may be appropriate to describe voiceless consonants. He maintains that when describing plosives it is convenient to use the term 'voiceless' instead of 'breathed' since it can hardly be said that during the stop of a plosive consonant there is a current of air passing through the oral tract (p. 19). Another categorisation is that of Jakobson (1957). He describes the above contrasts as 'lenis' and 'fortis' rather than voiced and voiceless. In this way he is breaking the tradition by diverting from the classification of the Old Arab linguists such as that of Sibawayhi (1966).

Another similar description is presented by another Arab scholar, Ibin Sïnä as early as (c. circa. 1353). He is a physician by profession. He mentions in his book *The causes of the production of sounds* that the voiced sounds are characterized by weaker articulation and the voiceless sounds with a stronger articulation. Ibin Sïnä is probably the first Arab scholar who correctly identifies the vibratory nature of sounds in Arabic. Concerning this point he writes "I think that the immediate cause of sound is the strong and fast undulation 'tamawwuj' of the air whatever the cause of this may be" (p. 4-5). The majority of phoneticians nowadays assume that the distinction between 'majhʊʊra' and 'mahmʊʊsa' corresponds to that between voiced and voiceless and that the addition of / ʔ, q, T / to the 'majhʊʊra' is an error on the part of Sibawayhi (Gairdner, 1925; Al-ani,1976; Mitchell, 1990 among others). They believe that the brief voiceless breath 'nafas' associated with /q, T, ʔ/ appears to Sibawayhi to be equal to the lack of frication noise. Therefore, the fact that /q/ and /T/ are associated with a brief period of frication is sufficient for Sibawayhi to group them among the 'majhʊʊra'. This interpretation seems to be more reasonable especially if one considers Sibawayhi's limited knowledge of the larynx and the unavailability of sophisticated experimental techniques in his time as is today. Such modern techniques have shown beyond doubt that /q, T/ are voiceless rather than voiced. As for / ʔ/ it is by definition neither voiced nor voiceless since the vocal folds are momentarily closed during the production of this sound.

To sum up, recent phonological descriptions agree to the following voicing contrasts in MSA (Mitchell, 1990, Al-ani, 1970; Gairdner, 1925): /θ/ vs /ð/, /t/ vs /d/, /s/ vs /z/, /T/ vs /D/, /ʃ/ vs /z/, /ɣ/ vs /x/, /ʕ/ vs /ħ/. It can be seen from the above pairing that each of the sounds /k/, /q/ and /h/ lacks a voiced correlative in MSA and that no voiceless consonant phoneme corresponds to /b, δ, l, n, w, j/. In YSA, however, there is a further voicing contrast namely that between /k/ and /g/ as in the words /kaala/ 'he measured' and the word /gaala/ ' he said'. In other Arab dialects (e.g. Egyptian) additional voicing contrasts can also be found as in /S/ vs /z/ and /k/ and /g/.

1.1.2. Assimilatory features of voicing in MSA and YSA:

So far the consonant categories are considered in isolation. However, certain phonetic phonemena frequently occur and remind us of the fact that what is described as a voiced or voiceless consonant does not necessarily remain so in all contexts. For example, the contextually determined devoicing of voiced consonants is wide spread in YSA and to some extent in MSA. One such process involves the partial assimilation of voiced consonants to voiceless ones (e.g. /sab̥t/ 'sabbath', /wajad̥t/ 'I found', /ħafḁδ̥t/ 'I kept'). The voiced sounds /b,d/ and /δ/ are partially changed to /t/. There are similar cases in which the assimilation is more apparent than in the previous ones to the extent that the voiced sounds are heard as voiceless: (e.g. [birittu] 'I got cold' [quttu] 'I led' and [qabaTtu]

'I received' for [biridtu], [qudtu] and [qabaDtu] respectively). In the first two words the voiced sound [d] is completely assimilated to [t], while in the third one the voiced sound [D] is assimilated to [T].

In addition, the voiceless sound [t] changes to a voiced one when it forms clusters with voiced sounds. The result is geminate voiced consonants as follows: [middaxliin] 'they are entangled' for [mitdaaxliin] , [mizzaʕliin] 'they are not in good terms' for [mitzaaʕliin] and [miDDamniin] 'they are united' for [mitDamniin]. All three cases are examples of regressive assimilation in that the voiceless plosive [t] has assimilated to the following voiced sound. Regressive assimilation is also found in words such as [ðamb] 'sin' or [dʒamb] 'side' for /ðanb/ and /dʒanb/. Although no change of voice has occurred the denti-alveolar nasal sound [n] becomes a bilabial sound under the influence of [b].

The most common of all voicing assimilatory processes in both YSA and MSA is the change of the definite article [ʔal] 'the'. Before nouns, both singular and plurals, it is completely assimilated to the following sound before [ʔal ɧʊrʊʊf ʔaʃ-ʃamsja] 'the solar letters' which include the following phonemes: /t,d, T, D, θ, ð, r, s, z, S, ʃ, δ, l, n/. They are arbitrarily called /ʃamsi/ 'solar letters' because the Arabic [ʃ] sound is the first letter of the Arabic word / ʃams/ 'sun' to which this group of sounds belongs. The points of articulation of these phonemes are quite near to the sound /l/ of the definite article /ʔal/. They are all produced somewhere between the denti-alveolar and the palato-alveolar zones of articulation. The purpose of this kind of assimilation is to ease and facilitate their pronunciation. Some examples are [ʔaʃ-ʃams] 'the sun'; [ʔat-tamr] 'the dates' ; [ʔaD-Dajf] 'the guest'; [ʔaθ-θaani] 'the second'. The change can either be to voiced or voiceless sounds depending on the following sun phoneme.

Other phonemes that have their articulation points outside the above mentioned regions do not show any effect on the definite article. These phonemes include: /b, w, m, f, j, k, g, q, x , ʕ, ʔ, ɣ, ħ, h/. They are called /ħuruuf qamarijah/ 'lunar letters'. They are referred to by this name simply because the /q/ phoneme is the first of the word /qamar/ 'moon' and because /q/ belongs to this group of sounds. As indicated with such sounds the definite article is neither assimilated nor incorporated into the following sound in the words as in these examples: [ʔal-qamar] 'the moon', [ʔal-balad] 'the country', [ʔal-kitaab] 'the book'.

Progressive assimilation of unvoicing and voicing phenomena does exists in YSA. They occur, for example, in these two words: [ʔiztaad] 'increased' for [ʔizdaad] and [ʔiztijaad] 'increment' for [ʔizdijaad].

Furthermore, voiced continuants may become unvoiced under the influence of preceding consonants especially voiceless ones as in these examples: /mat̥n̥/ 'text' ; /sitr̥ / /qism̥/ 'curtain' ; 'division'. Final unvoicing in YSA is not limited to the continuants but may occur with plosives as in these examples [faad̥] 'informed', [baab̥] 'door', [qaSd̥] 'purpose', and [taħmiiD] 'souring; film development'. Ancient Arab speakers in their effort to keep the language unchanged tried to prevent such kind of final voicing in their speech. They inserted a very short vowel of the type [ə] after the final sounds as in [baζd∂] 'after', but this deliberate action of voicing final plosives in MSA requires some effort. The insertion of [ə] is normally neglected in YSA and [ə] has no phonological status in this language. Thus, final unvoicing occurs only pre-pausally.

1.2.Emphasis

1.2.1 Definition:

Phonologically emphasis is simply another opposition feature used by the speakers to generate a further series of phonemes in co-ordination with other established features. It is given the name /tafxiim/ 'dignifying' by Arab linguists and European alike. The term has two derivatives /mufaxxama/ 'dignified consonants' (plur.) and /mufaxxam/ 'dignified consonant' (sing.). Emphasis is an alternative English term for the same phenomenon. Both terms, emphasis and 'tafxiim', however, have been ambiguously used as cover terms with no specific connotations. For consistency and convenience the term emphasis will be used throughout this study.

The ambiguity of the term emphasis may be attributed to the fact that it means different things to different researchers. Most of them believe that emphaticness is realized as either one or two of the following processes: velarization, uvularization, pharyngealization, retraction of the tongue, labialization and lip protrusion, strong articulation, and u-resonance, or heaviness. In what follows phonetic and phonological descriptions of these aspects will be dealt with in some detail.

1.2.1.1 Emphasis as 'velarization':

For many phoneticians emphasis is equivalent to velarization. Thus, Gairdner (1925) states that "we have called it velarization because the most prominent feature of the phenomenon is the raising of the BACK of the tongue towards the soft palate" (p. 192 his emphasis). He also adds that the tongue's blade is tense during its production.

The origin of this theory was Sibawayhi (1966). According to Sibawayhi /ʔiTbaaq/ 'lidding' or emphasis was defined as the raising of the back of the tongue towards the upper palate /hanak ʔaʕlaa/. The difference between the group of /muTbaqa/ consonants and /munfatiħa/ ones (i.e. emphatics and non-emphatics) was that the latter group did not require the tongue to be raised towards the palate.

The muTbaqa consonants were believed to have two points of tongue articulation. Sibawayhi added another three emphatic consonants /ɤ, q, x/ and called them all /mustaʕlia/ 'high consonants' as opposed to /mustafilaa/ 'lowered consonants'. The former were called so because they were described as "consonants which have a raising of the tongue towards the upper palate" (Sibawayhi, p. 285).

In brief Sibawayhi thought that emphasis /ITbaaq/ was a phenomenon involving a secondary articulation and that it was realized as velarization. The problem with such description was the use of vague terms such as /ITbaaq/; /munfatiHa/; whose exact semantic connotations were difficult to determine. It was not clear exactly what the articulatory areas for each of the muTbaqa were. Several phoneticians seemed to follow Sibawayhi (1966) without questioning. Among those who followed his path were Gairdner (1925) Trubetzkoy (1935) O'Connor (1973) Obrecht (1968) NaSr (1959a) Ferguson (1956a).

Trubetzkoy (1935), for example, related emphasis to velarization. His description clarified some points regarding the ambiguity of Sibawayhi's terms when he noted that "the emphatic apicals are not only velarized ... but are also alveolars in contrast with the post-dental non-emphatic apicals' (p. 131). In spite of that he emphasized the point that this shift in the primary place between the two groups of sounds was so slight that it should be disregarded.

* O'Connor (1973) too, claimed that the emphatics and non-emphatics differed in that the former were velarized. They both had the same primary place of articulation, but the emphatic articulation involved an added raising of the back of the tongue. Similarly, Ferguson (1956a) was the first to recognize the presence of a velarized [L] (i.e. emphatic) in Arabic as opposed to the non-velarized [l].

The basic problem for the theory of velarization was in describing the articulatory areas for back emphatic consonants such as /q/. How could such a phoneme have a primary and secondary articulations at the same time and in the same area of articulation? Those who recognized the problem excluded /q/ from the group of emphatics. Further, although some researchers adopted this definition in some occasions they expressed their doubt and uncertainty about it. For instance, inspite of Obrecht (1968) title *Effects of Second Formant on Perception of Velarization Consonants in Arabic*, he noted that emphasis was realized as pharyngealization.

1.2.1.2 Emphasis as 'Pharyngealization':

Other phoneticians believed that the sole realization of emphasis was pharyngealization (Jakobson,1957; Harrell,1957; Al-ani,1970; Delattre,1971; Laufer and Baer,1988, Ali and Daniloff, 1972). Both Jakobson and Harrell attributed emphasis to a pharyngeal constriction. Their descriptions depended on findings reported by MarÇais (1948) from X-ray photographs of emphatic and non-emphatic consonants for an Algerian dialect. In that study the traces showed that the emphatics were produced with a constricted pharynx caused by the retraction of the root of the tongue. Jakobson also indicated that the emphatic/nonemphatic contrast corresponded to the distinctive features 'flat-plain'. The articulatory correlate of the feature flat was stated to be a narrowing which occurred at one end of the vocal tract whether at the front or at the back.

Another different view of emphasis as a pharyngeal constriction was expressed in Laufer and Baer (1988). They disagreed with the statement that it was the root of the tongue which made contact with the pharyngeal wall, rather they believed that it was the epiglottis which constricted with the lower pharynx to form a secondary place of articulation. *Ali and Daniloff (1972) in their cinefluorographic study noted that the term 'velarization' is 'inappropriate' to describe Arabic emphatics because the velum contributed little or nothing to their production (p. 100). In addition, Al-ani (1970) in his physiological investigation of the Arabic phonemes stated that his X-ray films clearly showed that it was the pharynx which was relevant for the emphatic articulation. He, therefore called them 'pharyngealized sounds'.

The third approach believed that emphasis was not particularly related to either velarization or pharyngealization but rather to the changing shapes and sizes of the tongue. 'Lateral expansion' of the tongue was an additional articulatory action associated with emphasis (Mitchell,1990).

To conclude, most of the theories mentioned above seem to equate emphasis with one or two features. This seems to be the reason for their contradictions. In our view emphasis is a more complicated phenomenon than can be described by a single feature (See Lehn, 1963 for a similar view). It is likely that it involves all of the above articulatory processes, though they might not be equally prominent. Velarization and pharyngealization seem to be the most important components. They may differ from one speaker to another. Harrell (1957) has pointed out that women in Egypt tend to use emphasis less in their speech. They also differ depending on the emphatic sound involved. It is claimed by some Arab grammarians that /T/ is the strongest of the emphatic sounds.

1.2.2 Phonological categorization of emphatics and non-emphatics:

Views of phoneticians diverge regarding the identity of emphatics in Arabic. Different linguists have different opinions of what can be considered as an emphatic consonant and what cannot. The differences may be attributed to their different perception of emphasis.

Early Arab grammarians identified two sets of emphatics and non-emphatics. The first involved the contrasts /T/ versus /d/; /δ/ versus /ð/ and /S/ versus /s/. The second set involved four other emphatics but unpaired /D, q, x, ɤ/. This pairing was done by Sibawayhi who stated:

> "If there were no emphasis (iTbaaq) the Taa would become a daal, the Saad, a siin the δaa a ðaal, and the Daad would remain outside the language because there is no other consonant with that place of articulation" (translated from a quotation in Aniis, 1971, p. 59).

From this it could be observed that the original Arabic /D/ as described by Sibawayhi was differently pronounced. It was pronounced laterally against the left side teeth, with strong affrication. The combination of emphasis with affrication and lateralizing doubtless characterized this sound. The ancient Arabs considered its pronunciation so difficult that they denied to foreigners the ability to pronounce it and, therefore, called their language after it (i.e. the language of the Daad).

Nowdays, however, all Arabic dialects including YSA as a result of sound change and analogy aquire some new sounds and new contrasts. Thus, /D/ has changed from a fricative to plosive and paired with /d/. It has lost its affrication and its lateral articulation (cf. Gairdner,1935). /T/ has aquired a new non-emphatic cognate /t/. Similarly, the emphatic sound /q/ is now paired with /k/. The pairs /S/ versus /s/ and /δ/ versus /ð/ preserve their contrasts. Some variations may be found in certain Arab dialects (e.g. Egyptian, Syrian, and Lebanese) where the sound /δ/ versus /ð/ is replaced by or alternatively used with /δ/ versus /z/. In YSA no such change has happened.

Modern classifications of Arabic consonants are often associated with confusing and contradictory inventories. Such a problem arises from the insistance on classifying them soley on the basis of place and manner of articulation with no consideration of the emphatic articulation. Emphatics, in our view, should be classified by taking into account additional criteria to those of manner and place of articulation. Different zones of articulation in Arabic can be found. These include bilabial, dental-alveolar, palatal-alveolar, velar-uvular, pharyngeal and glottal. Before determining the emphatics and non-emphatics in YSA, it is worth examining some of the previous classifications and the criteria used to distinguish them.

Phoneticians dealing with the classification of emphatics and nonemphatics can be divided into three groups. The first are those who follow the old Arab grammarians by reducing the number of emphatics and non-emphatics to the minimum (cf. Mitchell,1990; Gairdner,1925). Mitchell, for instance, includes /T, D, S, δ/ and with certain reservation [l]. Apart from the limitation of being influenced by the Egyptian dialect, as appears from the inclusion of /z/, there is every reason for being cautious about following the exact description of the old Arab grammarians partly because of the changes which have occurred to the language.

The second group are those who broaden the domain of emphatic/nonemphatic contrasts to include pairs that are produced at different zones of articulation. Table 1.1 below compares three of such categorizations. This Table shows that these phoneticians grouped /T, D, S, q/ as the emphatic counterparts of /t, d, s, k/. In this they are in agreement with most of the other descriptions. The contradiction starts when they include pairs such as /x/ versus /ʃ/ and /z/ versus /ɣ/. This type of contrast cannot be accepted because the *primary* places of articulation for the consonants within each contrast are in different zones of articulation. /ʃ/ and /z/, for example, are in the palato-alveolar zone whereas their emphatic counterparts are in the velo-uvular zone. The implications of such random pairing is that more emphatic pairs can be emphatics and non-emphatics at the same time, as it appears from the lists where the consonants /x/ and /ɣ/ are allowed to be emphatics

(Jakobson,1957)		(Trubetzkoy,1935)		(Delattre,1971)	
Emphatics	non-Emphatics	Emphatics	non-Emphatics	Emphatics	non-Emphatics
/T/	/t/	/T/	/t/	-	-
/D/	/d/	/D/	/d/	-	-
/S/	/s/	/S/	/s/	-	-
/q/	/k/	/q/	/k/	/q/	/k/
/ɣ /	/z/	/ɣ/	/g/	/ħ/	/x/
/x/	/ʃ/	/x/	-	/ʕ/	/ɣ/
/ħ/	/h/	/ħ/	/h/	-	-
/ʔ /	/ʕ/	-	-	-	-

Table 1.1 Three phonological categorizations of the emphatics and non-emphatics in Arabic.

by Jakobson and Trubetzkoy and non-emphatics by Delattre. Similarly, /ħ/ is given two different non-emphatic cognates by Trubetzkoy and Delattre respectively.

The third group of phoneticians recognizes two types of emphatics and non-emphatics: primary and secondary (Harrell,1957). The primary ones are /T, D, S, q, δ, Z/ which correspond to the non-emphatics /t, d, s, k, ð, z/. The secondary emphatics are reported to be /R, L,B, M/. They contrast with /r, l, b, m/. The difference between the primary and the secondary ones is in their distribution and frequency of occurrence. The former occur independently and frequently. The latter are found only in the vicinity of emphatics. The primary emphatics occur as distinctive phonemes whereas the secondary ones rarely occur distinctively. They merely occur as allophones.

To recapitulate, most phoneticians seem to accept that /T, D, S, q/ are the emphatic counterparts of /t, d, s, k/ respectively. Although some phoneticians show some reservations about the pair /q/ versus /k/, the vast majority of them classed them as emphatic/nonemphatic cognates (e.g. Jakobson, 1957; Trubetzkoy, 1935; Ferguson,1956a; Delattre, 1971; Ali and Daniloff, 1972; and Laufer and Baer (1988) among others). Their inclusion is justifiable on acoustic and phonetic grounds. From our data and those of others /q/ shares common characteristics with the emphatic sounds. There is a lowering of the second formant and raising of the first one, thus narrowing the gap between F1 and F2 (see our finding in section 3.2). Such effect is equated with the secondary articulation associated with emphatics. It is accepted in acoustic theory of speech production that as the constriction gets further back in the mouth, the value of the F2 gets lower and F1 gets higher (Fant, 1960) Phonologically, Jakobson (1957) in his treatment of the emphatics expresses his approval of Trubetzkoy's remark (1935, p. 212) that ' the pair /q/ vs /k/ carries the same opposition as the pairs /T/ vs /t/ and the like'. Similarly, Laufer and Baer (1988) in their fiberoptic study of emphatics and pharyngeals in Arabic and Hebrew state that their findings support Delattre's claim that in Arabic "the consonant /q/ is the emphatic counterpart of /k/" (1971, p. 133). * Ferguson (1956a) points out that /q/ has similar influence on neighbouring sounds as that observed for /T/ and /D/. For instance, in the word /maSqaT/ 'the capital of Oman', the emphatic sound /q/ causes the nonemphatic sound /s/ to change to the emphatic sound /S/. With regard to the suggestion that /x,ɣ, B,R, M/ are emphatics, no convincing and consistent evidence are found in support of that claim. For these reasons the emphatic sounds in YSA as far as this study is concerned are /T, D, q/, while their nonemphatic counterparts are /t, d, k/.

Zone →		Labial	Dental-alveolar		Palatal-alveolar		Velar-Uvular		Pharyngeal	Glottal
Manner ↓	Voicing		Non-emp.	Emp.	Non-emp.	Emp.	Non-emp.	Emp		
Plosives	voiceless:	-	t	T			k	q		
	voiced:	b	d	D			g			ʔ
Fricatives	voiceless:	f	θ	-			x		ħ	
	voiced:	-	ð	δ			ɣ		ʕ	h
	voiceless:		s	S	ʃ					
	voiced:		z	-	z					
Affricates	voiceless:				-					
	voiced:				dʒ					
trill	voiced		r							
lateral	voiced				l					
nasal	voiced	m	n							
Semi-vowels	voiced	w			j					

Table 1.2 Zones and manner of articulation for the consonants in relation to emphasis and voicing in YSA.

* Based on the above observations a classification of the YSA consonants is presented below. The criteria adopted in this study for accepting or rejecting a certain consonant as emphatic or nonemphatic are the following.

* 1.The emphatics and their non-emphatic cognates should retain their primary articulation at the same zone of articulation.

2. Emphatics are characterized by having, at least, two places of articulation: primary and secondary ones.

3. The primary and the secondary places of articulation for a given emphatic consonant do not co-occur at the same zone of articulation. For the purpose of this study the consonants which fulfill these requirements are the emphatics /T, D, S, δ, q/. Their non-emphatic cognates are /t, d, s, ð, k/ respectively (see Table 1.2 above).

1.2.3 Assimilatory features of emphasis:

Our previous treatment of emphasis deals with it as a feature of the consonant itself. Other approaches regard it as a distinctive feature of the vowel as well. The assumption behind this view is that emphatics exert an influence on neighbouring sound segments, both within and across word boundaries. Such influence includes both vowels and consonants other than emphatics. Thus /a/ is realized as an open front vowel unless it occurs in the vicinity of emphatic consonants. After emphatics it sounds dark and is realized as a back or back-central vowel like /a/ in the English word 'father' (Gairdner,1925). Such a vocalic change and similar ones led some phoneticians (e.g. Ferguson, 1956a) to call for recognizing these vowels as independent emphatic phonemes. Ferguson supports his argument with the example of the [l] in the word /lanba/ 'lamp' which is pronounced by Arab speakers as emphatic because of its association with the back variant of /a/. The original /l/ in the European word is neither emphatic nor dark and yet Arab speakers pronounce it as emphatic. His justification is that emphasis is a feature of the vowel which extends into the consonant. He states that "In view of the extensive qualitative variation of the vowels it is conceivable that emphasis in Arabic should instead be regarded as a distinctive feature of the vocalic system with contrasting sets of non-emphatic and emphatic vowels" (p.163).

Further assimilatory features of emphatic consonants on other vowels are reported by Gairdner (1935) and Harrell (1962). When long or short /i/ succeeds one of the emphatics a glide of the /u/ or /w/ type is heard, as in /Tiib/ 'goodness' and /qiil/ 'it was said'. The influence of emphasis on back vowels is less clear. Although Harrell claims that back vowels are lowered in Moroccan Arabic, no such effect can be found at least in YSA.

Emphasis subjects the Arabic compound vowels /aw/ and /ay/ to the same effect as /a/. Since they both begin with /a/ they become more back as in these examples: /Tawq/ 'power' and /qawm/

'people'. The ancient Arab phoneticians were aware of such vocalic modifications. They formulated some rules by which they could account for these and similar changes. For example, they mentioned /imaalaa/ 'deflection' of the /fatħa/ in the direction of /kasrah/. In other words, /imaalaa/ caused a change of the vowel /a/ into /i/ when it occurred with sounds other than the following eight consonants /S, D, T, δ, r, x , q, ɣ/ (Gairdner,1935). These consonants were thought to be preventative of /imaalaa/. The relevance of such an effect was that it involved a change of the vowel quality.

Another phenomenon described by the ancient Arab grammarians was the /mufaxxama/ 'emphatic or dignifying' /a/ which was said to have the opposite function of the /a/ of /imaalaa/ or 'deflection'. The dignifying /a/ moved towards the /u/ or (Dama). However, it should be admitted that their discussion of these changes were not quite clear and not always in good agreement with modern phonetic descriptions.

✳ Emphasis is also found to extend to neighbouring consonants. Non-emphatic consonants may become emphatics by a process of assimilation as in these examples: /miSSaalħiin/ 'they are reconciled' for /mitSaalħiin/, /miDDaaminiin/ 'they are united' for /mitDaaminiin/ and /miTTaabiqiin/ 'they are identical' for /mitTaabiqiin/. In the three cases the non-emphatic consonant /t/ has assimilated to the emphatic consonants /S, D, T/ respectively. It is a regressive assimilation. Progressive assimilation is also quite common. Some examples are: /?iDTiraar/ 'obligation' for /?iDtiraar/; /TaLab/ 'request' for /Talab/. Examples of de-emphasis are reported to have been heard in some Arab dialects (Gairdner,1925) as in this example from Egyptian Arabic: /middaaji?/ 'annoyed' for /mitDaajiq/, and /sadda?/ 'believed' for /Saddaq/. The loss of emphasis in the consonants /S/ and /q/ is not only confined to Egyptian, but can be found in Syrian, Palestinian and Lebanese Arabic dialects. In all of them the /q/ sound assimilates to the glottal stop /ʔ/. In the Kuwaiti dialect it assimilates to the voiced palatal-alveolar affricate /dʒ/ as in /dʒaasim/ 'a man name', while in the Sudanese Arabic dialect it changes to /ɣ/ as in /ħaɣ/ 'truth'. All these are phonetic colloqail variations and do not apply to MSA.

It should, however, be remembered that the presence of an emphatic consonant in a word is an important component for the identification of emphasis and without it emphasis cannot be located.

What differentiates the word /Tiin/ 'mud' from the word /tiin/ 'figs' is the emphatic consonant /T/ and without it the emphatic contrast collapses.

*Accordingly, it is the consonant which carries emphasis. Ferguson's (1956) generalization of emphasis to all vowels on the basis of the example /Lanba/ 'lamp' is the exception rather than the rule. The only agreed upon emphatic vowel is the back variant of the low open vowel /a/ or /aa/ in the neighbourhood of emphatic consonants. So, the generalization of emphasis to all vowels seems to be unfounded.

1.3.1The linguistic context and distribution of the plosive consonants in MSA and YSA:

The linguistic contexts in which plosives occur and the possible combinations they may form with other classes of sounds will be dealt with in this section. The plosives occur freely in all positions: word-initially, medially and finally. They combine with all short and long vowels in MSA and in YSA. They, also, occur geminated but only word-medially and finally. Gemination is the doubling of consonants. Some examples are /ħatta/ 'until' /sadda/ 'he closed' /ɤaTTa/ 'he covered'. Gemination in word-initial position is observed neither in MSA nor in any other Arabic dialect. In fact no semitic language seems to have this pattern in its morphological system as this quotation from Greenberg (1950) indicates:

> "Thus, while sequences such as *mmd are virtually non-existent in Semitic languages, Arabic madd 'to stretch', farr 'to flee' etc. are representative of a common Semitic type" (p. 431).

Like consonant gemination, plosive clustering with other sounds in initial position is rarely found in MSA. It is confined to medial and final positions. In YSA, word-initial consonant clusters are found to occur in a few words (e.g. /bqara/ 'cow'). In this case, some YSA speakers tend to omit the vocalic element following the first consonant, thus allowing initial consonant clusters to happen.

The non-occurrence of initial consonant clusters in MSA is attributed to a general rule in Arabic stating that no word should start with a /ħarf saakin/ 'an unvocalized consonant'. Consequently, almost all word-initial consonants are followed by a vowel. The usual consonant cluster in MSA and YSA is two elements, but three elements may also be found at word boundaries. Some examples are:

Word-medial clusters	**Word-final clusters**
/saktah qalbijah/ 'heart attack'	/qalb/ 'heart'
/ʔaqdaam/ 'feet'	/ʔarD/ 'earth'
/ʔaTnaan/ 'tons'	/nuTq/ 'pronunciation'

Three elements clusters can be found in word boundaries as in / ɧarbhum/ 'their war' and /ʃaʕbhʊm/ 'their people'. Such kind of consonant patterning is usually broken by the insertion of the intrusive vowel or what is called by some phoneticians the 'anaptyctic vowel' (Gairdner, 1925). It may be /a/ or /i/ but more frequently it is /u/ as in: / harbuhum/ and /ʃaʕbuhum/.

A similar role is performed by /hamzat ʔal-waSl/ 'the linking glottal stop'. It is used when two or more unvocalized consonants occur at word boundary so that it prevents the occurrence of consonant clustering and to make their pronunciation easier as in this name: /ʕamr bin wid/ which becomes /ʕamr ʕibin wid/.

The clustering is also governed by the consonant compatibility of place. The bilabial plosive /b/ occurs freely with both back and front consonants, medially and finally with the exception of clusters such as /bf/, /bm/ and /fb/. It does not combine with consonants which are produced at almost the same place of articulation such as /f/ and /m/ (Al-Makhzuumi, 1960; Al-AnTaaki, (n.d.); Bohas, Guillaume and Koulloughli, 1990 and Al-ani, 1970).

The denti-alveolar plosives, both emphatics and nonemphatics, occur freely with the fricatives particularly with those produced at the back section of the vocal tract and the labials. Examples of such combinations are listed in Appendix 1. It can be seen that the denti-alveolar plosives do not form consonant clusters with each other except on rare occasions as in this example, /ʔaDdaad/ 'contradicting forces'. Otherwise combinations such as /Tt/, /td/, /dt/, are very rare indeed.

There are constraints on the plosives' clustering with /θ, δ, ð, z, s, S, ʃ/. These constraints are more apparent with the emphatics /T/ and /D/ than with /t/ and /d/. The pharyngeal and laryngeal consonants are subject to similar restrictions. According to Al-Makhzuumi (1960) /ʕ/ and /ħ/ never combine as clusters or separately in a word for the same reason of incompatibility. The same remark is said to apply to /h, ħ/; /ʕ, ɤ/; /k, q/; /g, k/; /h, ʔ/. All these pairs are listed by Arab grammarians as incompatible consonants. Other consonants, apart from these back consonants, seem to form two element clusters with other sounds in varied degrees of frequency and freedom depending on how close or far their places of articulation are (See Appendix 1).

1.3.2 Syllable structure in MSA:

As part of our review of contextual changes in YSA this section and the subsequent ones will pursue in some detail syllable-stress in MSA and YSA. In spite of their detailed treatments of certain phonetic features, the old Arab grammarians do not mention the syllable in any of their grammar books. This may be attributed to one of two possible reasons: (1) a lack of interest or (2) a deficiency in their methodological approach or both. Our belief is that the second one is the more probable because of their use of the Arabic orthographic systems as their criterion of reference in describing phonetic and phonological phenomena. For example, the word /kataba/ 'he wrote' would be described by them as consisting only of three consonants /k t b/ without reference to the vowels because in the Arabic writing system the short vowels are not represented by separate phonemes. "It is such concentration on written form that has led in Arabic phonology to an obsessive concern with consonants and a consequent failure to appreciate the vital importance of the syllable, more especially, of the total pattern provided by syllables in combination" (Mitchell,1960 p. 329). Consequently, our treatment will depend on what is written about it by modern Arab linguists and western phoneticians.

The building blocks of the syllable in Arabic or any other language are the consonants and vowels. The consonants in MSA always occur at the margins of the syllable (i.e. beginning and end). Usually the term 'onset' refers to the consonant that initiates a syllable; 'nucleus' to the vowel (s) that follow(s) the 'onset'; and 'termination' to the consonant(s) that end(s) the syllable(s) as in these examples: /baab/ 'door' (CVVC). In this monosyllabic word (C) stands for consonant, and (VV) for a long vowel. Although the double symbols (VV) are used to represent the vowel, it is in fact a single unit. The doubling of long vowels is used merely to indicate length and to distinguish them from the corresponding short vowels. The 'onset' consists of one consonant. No syllable onset in Arabic consists of two consonants. The 'nucleus' may be a short or long vowel. The termination in /baab/ consists of one consonant but two may occur in this position and sometimes no termination is found. Since all consonants including /w/ and /j/ occur only as onset or termination, and since all syllables must have a vowel as their nucleus, the number of syllables in a word is by definition equal to the number of vowels[1]. Thus, in the word /makaatib/ 'offices' there are three syllables.

The syllable types are classified in different ways. They are closed, when they end with a consonant as in /Tiin/ 'mud' and open when they end with a vowel as in /ma/ 'what'

The syllables in Arabic are described according to their length. They are divided into three types:

1. Short syllables consisting of one consonant and one vowel (CV) as in /bi/ 'in'.
2. Medium syllables represented by a consonant and a long vowel (CVV) or a consonant, vowel and another consonant (CVC as in /laa/ 'no' or /min/ 'from').
3 Long syllables. This may be (CVVC), or (CVCC), or (CVVCC) as in these examples: / ɣaab/ 'went away'; /ħarb/ 'war'; /ħarr/ 'hot'. Taking both the above classifications combined, most investigators agree to the following six syllable types in Arabic (Mitchell,1960; Al-Ani, 1970,1979; Harrell,1957; Lehn,1967):

Syllable Pattern	Example	Meaning	Syllable Description
1 CV	/bi/	'in'	short /open
2 CVC	/sin/	'tooth'	medium/closed
3 CVV	/fii/	'in'	medium/open
4 CVVC	/baat/	' to pass the night'	long/closed
5 CVCC	/ɣarb/	'west'	long/closed
6 CVVCC	/ħaarr/	'hot'	long/closed

Table 1.3 Syllable types in MSA

Of the six types only two are open: numbers 1 and 3. The other four types are closed.

Concerning their frequency of occurrence, the first three are the most frequent. They occur in all positions: word-initially, medially and finally but they occur more in word-initial position. (CVVC) and (CVCC) are limited in their occurrence to word-final position. They never occur word-initially. They are less frequent than the previous three types. The least frequent of all is (CVVCC) which occurs in word-final position and in pausal forms with geminates.

---

[1] It should be remembered that the long vowels are represented by the double symbols (VV). In this case, as indicated earlier, the symbol (VV) should be counted as one unit and not as two separate vowels.

Syllable structure in MSA is influenced by the presence or absence of pausal forms. Pausal forms are grammatical terms indicating the omission of inflectional endings such as -a, -u, -i, -an, -un, -in and the replacement of --atan, -atun, -atin by -ah at the end of words. This can be seen by comparing these two groups of words. The examples show that the occurrence of pausal form reduces the number of syllables by one in each case as a result of the elision of the nominative, accusative and the genitive case endings and also the nasalized sound or /tanwiin/ which are all caused by the presence of pausal forms.

Inflectional Endings	**Non-pausal Forms**	**Pausal Forms**	**Meaning**
-a	kataba	katab	he wrote
-u	jaktubu	jaktub	he writes
-i	fil-kitaabi	fil-kitaab	in the book
-an	kitaaban	kitaab	a book
-un	kitaabun	kitaab	a book
-in	kitaabin	kitaab	a book
-atan	kitaabatan	kitaabah	writing
-atun	kitaabatun	kitaabah	writing
-atin	kitaabatin	kitaabah	writing

Table 1.4 Pausal forms and syllable structure in MSA

The omission of certain inflectional endings is determined by the type of grammatical words preceding them and is not optional as this table may suggest.

On the other hand, suffixes and prefixes have the opposite function. Some of these are pronouns /-hum/ 'their'; /-na/ 'us' and the imperative particle /ja-/ or the future indicator /sa-/ 'shall' preceding Arabic verbs. As more of these are found in a word, the number of syllables will be increased. From our observation, it is found that up to ten syllables may occur in certain Arabic words as a result of the additional prefixes, suffixes and particles as shown by the examples in Table 1.5.

The long words (7-10 syllables) are less frequent in speech. In fact most of the Arabic words consist of 3, 4 or 5 syllables.

The relationship between syllable boundaries and gemination is quite close in that phonetically geminate consonants (CC) are one physical component with no clear boundaries between the two consonants. Phonologically, however, and for convenience of analysis the geminate consonants at word-medial position would be regarded as belonging to two different syllables as in this example: /ʃad`da/ (CVC`CVC).

Example	**Equivalent Translation**
1.ʕilm	knowledge
2.ʕaalim	Scientist, knowledgeable
3.jaʕlamuun	They know
4.jataʕallam	He is learning
5.jataʕallamuun	They are learning
6.sajataʕallamuun	They are going to learn
7. jataʕallamuunahu	They are learning it
8. sajataallamuunahu	They are going to learn it
9. sajataʕallamuunahumaa	They will learn the two of them
10. fasajataʕallamuunahumaa	They will certainly learn the two of them

Table 1.5 Affixation and syllable structure in MSA

1.3.3.Stress patterns in MSA:

Sress, like the syllable, was never mentioned by the old Arab grammarians such as Al-Makhzuumi (1960) and Sibawayhi (1966). Description of it started to emerge early this century with work of Gairdner (1925) which shed some light on this important phenomenon. It was, however, concerned with Egyptian Arabic. Work was then delayed to the period after the second world war where studies of stress in different Arabic dialects were conducted. Among them were Birkeland (1954); Mitchell (1960, 1990); Garbell (1958); Harrell (1957); Al-ani (1970). The following treatment of how stress works in Arabic will depend on these sources as well as others.

Stress is defined by Birkeland (p. 6) as an "emphasis on a special syllable of a word of such a nature that emphasis on any other syllable is commonly disapproved of". From a perceptual point of view the stressed syllable is more prominent and stands out above others.

Stress can be identified by one or more of the following cues: pitch, loudness, quality,quantity and possibly emphasis (See, for example, O`Connor,1973; Couper-Kuhlen,1986; Roach, 1983). Thus, pitch may rise, fall or remain level in speech. If one syllable in a word is produced with a rising pitch and one with a falling pitch, then, the two syllables will be perceived differently.

Secondly, the stressed syllables are reported to be heard louder than the unstressed ones, but the relation here, as indicated by O`Connor (1973) is an uncertain one. Consequently, loudness is not enough by itself. It has to be supported by other cues especially pitch, at least for English. Thirdly, syllable quantity is particularly important for the identification of Arabic stress. The unstressed syllable /-ɤir/ in /Saɤir/ 'become smaller' is perceptually shorter than when it is stressed as in /Saɤiir/ 'small'. Emphasis may be a correlate of stress. The emphatic sounds seem to be produced with what Jakobson (1957) describes as 'strong pressure'. Emphatics are also described as being 'heavy'; 'tense'; and 'thick'. These features seem to indicate that emphasis may be a likely correlate of stress in Arabic.

1.3.4. Levels of stress:

In syllabic and polysyllabic words, it is claimed by some phoneticians that there are different levels of stress (NaSr,1959b; Al-ani,1970). Al-ani recognizes three levels of stress in Arabic : primary (` ) ; secondary (′ ) and weak (left unmarked) as in this example: /ra′ʔiisu`hunna/ 'their boss' /CV′CVVCV`CVCCV/. Mitchell (1960), on the other hand, in his treatment of syllabicity and prominence in Arabic does not concern himself with such a division. Instead, he concentrates on which syllables may be accented and which may not. The same approach is followed by Birkeland (1954) and Gairdner (1925). Consequently, in this study we shall consider stress as an entity which either does or does not occur.

Stress location in Arabic is relatively unpredictable. It has a semantic function. Some examples of the existence of such a contrastive function of stress in Arabic can be seen in Harrell (1957) and NaSr (1959b). Although most of the comments they made are applicable to MSA, some of the examples are only true of the dialects they investigate, namely Egyptian and Lebanese respectively. The location of Arabic stress depends on the number, length and sequence of the syllables in a word. It is based on whether the word has one, two or more syllables and whether they are open or close; short or long and on what order they may come. Based on the above facts and on what is stated by

Birkeland (19540) Gairdner (1925) and Harrell (1957) the following rules can be formulated about word stress in MSA:

1. The last long syllable is always stressed (e.g. /qa`riib/ 'near').
2. When a word contains two long syllables or more the nearest to the end is stressed (e.g. /ʕadaa`daat/ 'electric meters').
3. When a word contains no long syllables, but contains one or more medium syllables, the one nearest to the end is stressed (e.g. /qariibu`hunna/ 'their relative').
4. When a word contains no long or medium syllables and only consists of short syllables, the stress falls on the first syllable (e.g. /`katama/ 'he kept the secret'). The exception to this last rule is when the word consists of more than three short syllables. In this case the stress moves one syllable towards the end of the word for each short syllable added (e.g. /ku`tubuhu/ 'his books').

In addition, the link between grammatical forms and stress in Arabic manifests itself in several ways. Pausal forms, for example, cause stress shift as in these examples: /ʕaqlun/ 'mind' for non-pausal forms and /ʕζaql/ for pausal form. The dual and the plural forms attract stress as in these examples:

Singular		**Dual**		**Plural**	
/du`kaan/	'a shop'	/dukaa`naan/	'two shops'	/dakaa`kiin/	'shops'
/mux`liS/	'sincere'	/muxli`Saan/	'two sincere men'	/muxli`Suun/	'they are sincere'

Prefixes, however, seem to have little effect on stress placement. These include among others the definite article as in these examples: /`xatama/ 'he stamped it'; /jax`tim/ 'he stamps it'; /`dars/ 'lesson'; /ʔad`dars/ 'the lesson'.

The above notes apply only to stress in isolated words. In utterances, however, not all words will remain stressed. Some syllables may become shorter and in certain contexts may lose their stress as in these examples: /ʔiktu`buun / 'you write' (pl. imp.) and /ʔiktubud`dars / 'you write the lesson' (pl. imp.).

Clearly, the syllable /-`bʊʊn/ has lost its prominence in connected speech. This leaves no doubt that any treatment of stress in Arabic should be aware of such changes. It is a point which Ferguson (1956b, p. 387) has drawn attention to when he writes:

> "Every consideration of Arabic word-stress which does not take phrasal stress patterns into account is bound to be incomplete and misleading".

1.3.5 Syllable structure in YSA:

As has been done for MSA the syllable structure of YSA will be described next. From our examination and observation of YSA words, it is found that at least nine syllable structures can be identified. Table 1.6 sumarizes them with examples:

Syllable Pattern	Example	Meaning	Syllable Description
1 CV	/wa/	and	short/open
2 CCV	/xʃabah/	wood	medium/open
3 CVV	/maalaa/	diverted	medium/open
4 CCVV	/tfaagad/	check	long/open
5 CVC	/min/	from	medium/closed
6 CVCC	/waʕd/	promise	long/closed
7 CVVC	/maal/	wealth	long/closed
8 CVVCC	/ħaarr/	hot	V. long/closed
9 CCVVC	/bTaaT/	potatoes	V. long/closed

Table 1.6 Syllable types in YSA

A clear difference between YSA and MSA is that the former includes consonant clusters in word-initial position. This can be seen in the patterns (CCV), (CCVV) and (CCVC). Such a differences are caused by the elision of the vowel in the first syllables. For example, the long syllable word /ʃʕiir/ 'oats' consists of two syllables in the MSA word /ʃaʕiir/. Similarly, the YSA word /xʃaba/ 'wood' consists of three syllables in the MSA word /xaʃaba/. In both cases YSA speakers omit the first short syllable of MSA words to form a new long or medium syllable. This phenomenon is not limited to YSA, it exists in some other Arabic dialects such as Iraqi, Syrian-Palestinian and Kuwaiti (Blanc,1953; and Garbell, 1958). The other six patterns are similar to MSA. The list is far from comprehensive and it is likely that other types occur in YSA such as /CCVCC/; /CVCCC/ and /CCVC/. As far as this study is concerned no examples for these patterns can be found.

1.3.6 Stress patterns in YSA:

Most studies of stress concentrate on three Arabic dialects: Egyptian, Syrian-Palestinian and Iraqi. Important as they may be, these studies do not tell the whole story about Arabic dialects. Till now very little is known about the Peninsular and southern Arabic dialects including YSA. With regard to this particular point Lehn (1967) writes "Until more is known about these dialects, one should regard with scepticism generalizations about all Arabic dialects" (p. 321).

Another reason for exploring these varieties of Arabic is related to the fact that MSA has acquired its stress rules from the different Arabic dialects, a fact that is echoed by many linguists. Referring to the importance of colloqials in describing Arabic stress Mitchell (1990) writes "Vernacular influence is nowhere more clearly in evidence than in the matter of accentuation". Birkeland (1954), too, emphasizes the same point when he says that "the fixed word stress, used in most parts of the Arab world and by most European scholars when reading classical Arabic, is secondarily introduced into this language from the colloqial". For these reasons it is considered necessary to give brief notes of how stress in YSA works and by what rules it is governed.

Unfortunately, no detailed treatment of stress in YSA could be found. The only reference which touches upon this important aspect of YSA is Qafisheh (1990). Based on this and on our own knowledge as a native speaker of YSA, the following word stress rules can be formulated:

1. Monosyllabic words are stressed (e.g. /`siin/ 'the letter 's')
2. In two-syllable-words, stress falls on the last syllable if it is long (e.g. / ʕa`δiim/ 'great').
3. In two-syllable-words, the stress falls on the last syllable if it is medium, and the first one is short (e.g. /ra`dʒaaʔ/ 'a wish'). Otherwise, the first syllable is stressed (e.g. /`ʔismi/ 'my name'; and /`mara/ 'a woman').
4. In three-syllable words, if the second and third syllables are short stress falls on the first syllable (e.g. /`waahida/ 'one' (f) and /`kataba/ 'wrote').
5. In three-syllable words, if the second syllable is medium, and the last and first syllables are not long, the second syllable is stressed (e.g. /ɧʊ`kʊʊmah/ 'government').

6. In four-syllable words if the last syllable is long it is stressed as in /muʕalli`miin/ 'teachers'. If it is not long and the second to the last is long or medium, then the second to the last is stressed (e.g. /muta`naffas/ 'breathing space' and /jatabaaħaθ/ 'he discusses with').

Thus, in YSA if the last syllable is long it is stressed no matter how many syllables the word may have. The stress placement also depends on the number and type of syllables found in a word. Vowel length also plays an important role in the identification of stressed syllables in YSA as in these examples, /`qadim/ 'he arrived' and /qa`diim/ 'old'.

Stress and vowel length are usually concomitant in YSA. If there is only one long vowel in a word, that long vowel appears to always carry the stress. Where there is more than one long vowel in a word, normally, the one nearest to the end is stressed as in this example, /ʃaɤaa`laat/ 'female workers'.

Chapter 2: Literature Survey on Plosives

In this section we will attempt to review critically previous work on the plosives in a wide range of languages. They will be dealt with at three levels: acoustic, perceptual and aerodynamic in that order. The various mechanisms, findings and hypotheses related to voicing and emphasis will be discussed in this part. Some of the assumptions raised by this review will be subjectd to a detailed investigation later on (chapters 3, 4 and 5 of this study). The three sections 2.1, 2.2 and 2.3 are meant to correspond to the following three chapters in part two of this study. For example section 2.1 corresponds to chapter 3, section 2.2 to chapter 4 and section 2.3 to chapter 5. Our findings will also be discussed and compared in relation to remarks made in this section.

2.1. Survey of acoustic studies on plosives with reference to emphasis and voicing.

2.1.1 Acoustic studies on voicing.

2.1.1.1 General background on plosive production:

The production of plosive consonants is influenced by three articulatory processes: (a) the supraglottal actions caused by the movements of the articulators in the pharyngeal, oral, and nasal cavities; (b) the glottal actions; (c) and the subglottal (respiratory) actions. Thus in the vocal tract they are produced with the nasal cavity closed, a rapid closure somewhere in the vocal tract and a release of the air pressure which is built up behind the constriction point. The closure occurs when one articulator moves against another, thus preventing any flow of air through the vocal tract. Eventually, the two articulators part resulting in the release of the plosive consonant.The constriction location varies from one consonant to another. For /b,p/ the occlusion occurs when the upper and lower lips come together and the release of air occurs as they separate. The /t,d/ occlusion is formed when the tip of the tongue is raised against the alveolar ridge, while for /k,g/ it occurs when there is contact between the back of the tongue and the velum particularly in the context of back vowels. When they are followed by front vowels the contact occurs between the back of the tongue and the hard palate (Abercrombie,1967 p. 51). Thus, there are three phases in the plosive production at the supraglottal cavity: closure, hold and release (Gimson,1989; Roach,1983).

2.1.1.2 Voicing:

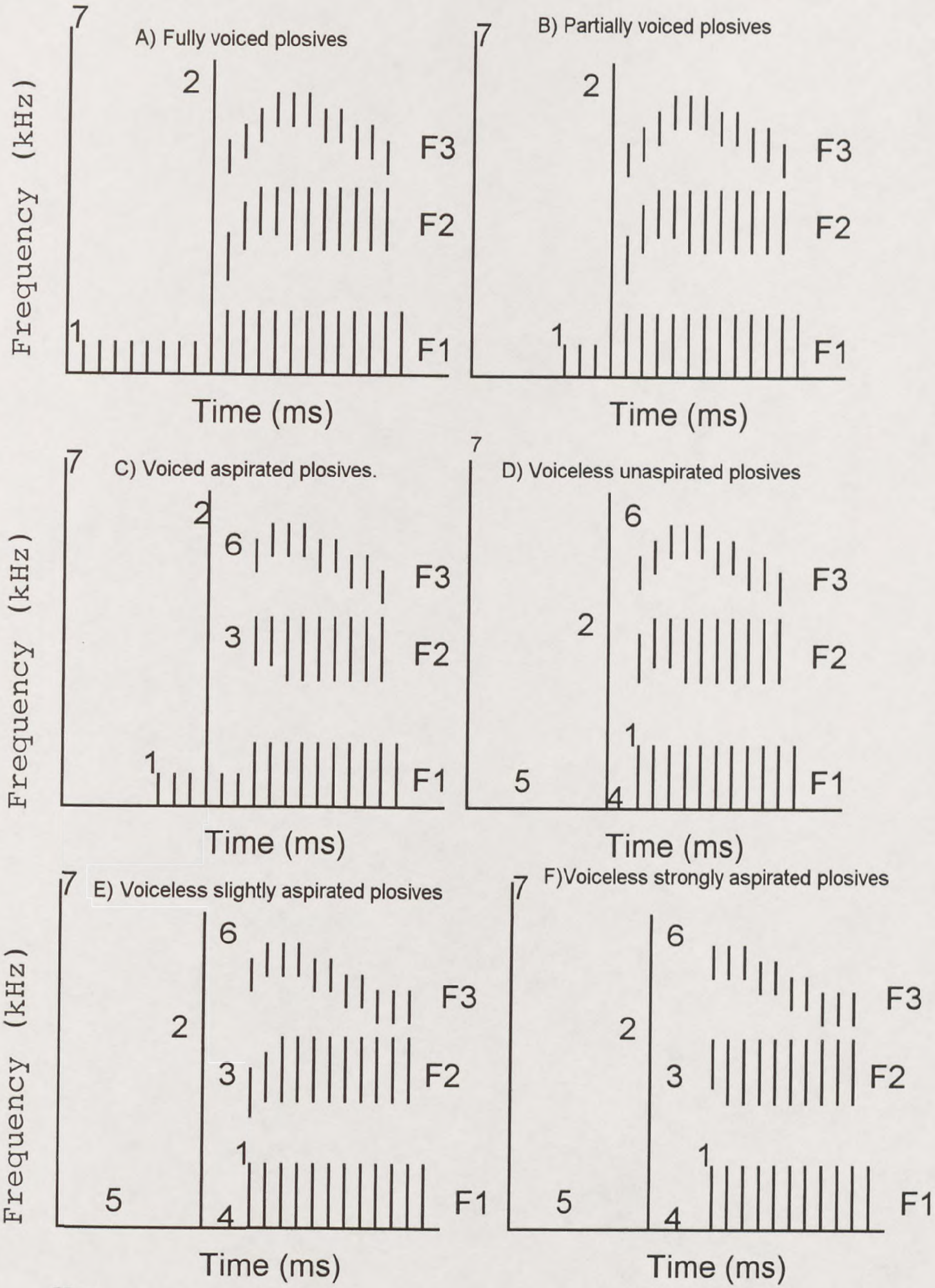
The second aspect influencing the production of plosives is what happens in the larynx. The state of the glottis is important because it , together with the stiffness of the vocal folds, determines the sound sources for the different consonants (Ladefoged,1982; Pickett,1980). The glottis assumes different shapes and sizes. It can be very small, small or large. In voiced sounds the vocal folds are nearly or completely closed and vibrating, while for the voicelesss ones they are expected to be apart. Consequently, voicing or lack of it is determined by the presence or absence of the vocal folds vibration which in its turn depends on the size of the glottis, the stiffness of the vocal folds and also on aerodynamic conditions at the glottis.

Acoustically voicing is correlated with periodic waveform reflecting the vocal folds vibrations. Phoneticians recognized these voicing characteristics many years ago and developed systems to describe the different consonants accordingly. For example, Chomsky and Halle (1968) use the voicing feature in a binary form to denote 'voiced' or 'voiceless' whereas Ladefoged (1982) uses it as a 'physical scale' with various degrees of voicing depending on the opening size of the glottis.

The other glottal feature associated with the plosive production is that of aspiration. Aspiration is used here in two different senses. The first refers to the aspiration noise source, that is turbulence noise arising at the glottis when the vocal folds are slightly apart. The second refers to the traditional phonetic terms in which an aspirated consonant is opposed to an unaspirated one. These two senses are not in conflict (Scully, 1983). On the contrary they complement each other. Thus an aspirated plosive is produced with aspiration noise in its release phase . Aspiration is a property of voiceless sounds but may occur with voiced ones as in Hindi and Marathi (Lisker and Abramson, 1964). Traditionally, aspiration is defined as a 'puff of air'. Experimental evidence, however, has challenged such a simple view of aspiration and suggests that it involves "the glottal opening at the time of release of the oral closure of a stop" (Kim, 1970). According to him the more open the glottis at the time of release, the more aspiration there will be. Aspiration, however, does not occur with voiceless sounds in all languages. It is particularly associated with Germanic languages such as English. In acoustic terms, it occurs after the release burst and is shown as random markings on spectrograms at the range of higher formants.

2.1.1.3 Voice Onset Time defined:

The combination of both the aspiration noise and voicing constitutes what is known as voice onset time (VOT). VOT is defined as the relative timing between the release of a plosive and the start of voicing (Lisker and Abramson, 1964).



Key:
1= Onset of voicing. 2= Release transient. 3= Aspiration noise. 4= Voice onset time. 5= Silence.
6= Frication noise. 7= Onset of 'acoustic closure'.

Figure 2.1.1 Sketches of different types of plosive-vowel syllables as they are expected to appear on wideband spectrograms. The ordinate shows frequency and the axis line defines the moment of 'acoustic closure' where this is observable.

It is measured with respect to the instant of release. If voicing begins before the release VOT is given a negative value, if voicing begins after it VOT is assigned a positive VOT value. The terms used to describe the VOT in these two regions are called voicing 'lead' and voicing 'lag' respectively.

On the usefulness of this measure Lisker and Abramson state that 'a difference of voicing not only separates voiced from voiceless stops, but that it equally well distinguishes aspirated from unaspirated stops, where the latter are both commonly called voiceless'. They observe that aspiration and voicing are mutually exclusive. Each one tends to appear relatively clearly on the spectrograms only when the other is absent or weak. In other words they do not occur simultaneously in the same position on the spectrograms. This statement is not always true in all languages. For example, voiced aspirated plosives of Hindi are produced with simultaneous weak breathy voicing and aspiration noise occurring together. When this happens it poses a measurement problem of VOT for the voiced and voiced aspirated plosives. Figure 2.1.1 shows sketches of six different types of plosive voicing ranging from fully voiced plosives (a) to voiceless strongly aspirated plosives (f). The periodic pulsings at lower frequency region and aspiration noise are drawn as they are expected to appear on wideband spectrograms. If we consider the phonetic features of voiced/voiceless and aspirated/unaspirated we will see that the four possible combination of these features are accounted for: 1.voiceless aspirated. 2. voiceless unaspirated. 3. voiced aspirated. 4. voiced unaspirated.

2.1.1.4. Cross-language studies of VOT:

Different languages have different plosive voicing categories. In their cross-language investigation of various plosive types of 11 languages, Lisker and Abramson (1964) found that the VOT together with aspiration separated the plosives into two, three or four opposition types. Four of the languages investigated (Dutch, Spanish, Hungarian and Tamil) had their VOT values in the voicing lead and the short voicing lag. Cantonese and to some extent English had also two opposition types like the previous four languages, but theirs occupied different VOT ranges. Their VOT values were distributed in the short and the long voicing lags.

Although the English plosive phonemes /bdg/ and /ptk/ are described as voiced and voiceless, this distinction may not hold true in all positions. /bdg/ are generally voiced in intervocalic position, but they are usually not so in initial and final positions (Jones, 1976; Gimson, 1989). Thus, it is not surprising that English shares the same regions of VOT values as Cantonese. This similarity is not without exception because one of the four English informants in Lisker and Abramson's experiment tends to produce VOT in the voicing lead region for the voiced plosives, and in the short lag for the voiceless ones. This may have been caused by dialectal variation among the individual subjects. Indeed, Painter (1979) observes that because of dialectal variation he frequently produces voicing before the release in voiced plosives during the closure phase even in initial position.

In brief Lisker and Abramson (1964) found that the plosives with two opposition types were associated with two average median VOT values -100 to 10 ms and 10 to 75 ms. Dutch, Spanish, Hungarian and Tamil occupied the median VOT range -100 and 10 ms, whereas Cantonese and English occupied the median VOT range from 10 to 75 ms.

Languages with three opposition types were Eastern Armenian and Thai. They had their median VOT values within the ranges -100, 10, 75 ms. Korean which employed three opposition types, however, was reported to have all its plosive categories in the voicing lag. They were differentiated in part by the amount of aspiration noise they had (Lisker and Abramson (1964); Kim, 1965; Hardcastle, 1973). The centering regions for the VOT values were reported to be 10, 35 and 90 ms for the voiceless unaspirated (tense), voiceless slightly aspirated (lax), and the voiceless heavily aspirated plosives respectively (Kim,1965). Korean, however, presents special problems in that these three opposition types cannot always be found in the speech of all speakers.

Languages associated with four opposition types were represented by Hindi and Marathi. In these languages aspiration was not only found with the voiceless plosives but also with the voiced ones. It was in these languages that the measure of VOT was not by itself sufficient to separate all plosive types. The problem arose from the fact that overlap in the VOT values were found between the voiced unaspirated and the voiced aspirated plosives. The overlap was, however, limited to these categories.

2.1.1.5. Voice onset time in Arabic:

One important study of the voicing feature in Arabic was that of Yeni-komshian, Caramazza, and Preston (1977). Motivated by Lisker and Abramson's (1964) findings regarding the success of VOT in separating the voicing types in several languages, they decided to test this measure in another language (namely Lebanese Arabic). Their spectrographic investigation focussed on the word-initial plosives in the context of the three short vowels /i,a/ and /u/. The results showed that the VOT separated the plosive types into well defined ranges: voiced sounds having their VOT in the voicing lead, and voiceless ones having their VOT in the short voicing lag. In spite of that, to their surprise, there were some occurrences of VOT overlap for the pairs /t,d/ and /T,D/. The region of overlap was said to be between zero and 30 ms voicing lag.

Another study of the VOT in Arabic was a comparison of the voiceless plosives in the Saudi Spoken Arabic and the American English (Flege and Port, 1981). They concluded that 'word-initial Arabic voiceless stops /t,k/ seem to be produced with somewhat shorter VOT values than similar stops in English'.

To my knowledge, there appears to be no experimental study which is specifically designed to examine voicing and VOT in YSA. Although, the previous two studies are concerned with two spoken Arabic varieties, it can hardly be said that all their findings are true for YSA. There are some phonetic differences which cannot be ignored, regardless of how minor they may be. For example, Lebanese speakers of Arabic replace the uvular plosive consonant /q/ with the glottal stop / ʔ/ in their speech, whereas in YSA this tendency does not exist. Instead the /q/ phoneme has two variants [q] as it is the case in Modern Standard Arabic (MSA) and [g] which is a voiced velar plosive.

Other investigators expressed their doubt about the efficiency of the VOT measure and found that it was not always successful in separating the different types of plosives. Thus, Caramazza, Yeni-Komshian, Zurif and Carbone (1973) indicated that the VOT in Canadnian French was a sufficient cue neither at the production nor at the perception levels. It failed to distinguish /p,b/, /t,d/ and /k,g/ in that dialect. Kim (1965) also found an overlap in the production of one subject in his study of Korean plosives.

In view of such variations between different forms of spoken Arabic and the overlap in the VOT values between homorganic plosives of some languages, it was important to determine the extent of the separation of the plosive types in YSA by means of VOT and how the results compared with those from other languages.

Lisker and Abramson (1967, 1970 and 1971) have repeatedly shown the reliability of the VOT alone for differentiating plosive categories in different languages. In one of their studies (1971, p.776) they state that 'VOT is the single most effective measure' for sorting plosives into different phonemic categories with respect to voicing.

It will be interesting to see if the vocalic context affects the VOT values of the preceding plosive consonants. Some studies have demonstrated that the vocalic context affects consonant duration and VOT. Thus, Klatt (1975) has reported a statistically significant difference in the VOT values for the voiceless plosives depending on whether a high or mid vowel follows the plosives. Similarly, Suomi (1985) has found for Finnish, that the vowel exerts a reliable influence on the plosive-vowel waveform from the very beginning of the syllable.

2.1.1.6 Force of articulation:

A reference should be made to two more relevant features, namely force of articulation and emphasis. It is claimed that the pairs /p,b/; /t,d/ and /k,g/ can be distinguished not only by the presence or absence of voice, but also by the force of breath and the muscular effort involved in their production (Halle, Hughes, and Radley (1957); Malecot, (1970). It is suggested that /ptk/ are produced with stronger energy than /bdg/. For this reason the former group of sounds are described as 'fortis' or 'tense' whereas the latter ones as 'lenis' or 'lax'.

In his treatment of these sounds O'Connor (1973) preferred the opposition feature 'fortis' vs 'lenis' to the others (i.e. voice vs voiceless or aspirated vs unaspirated). His justification for that choice was that although aspiration served to distinguish the plosives, it failed to do so with the fricative consonants. This view seemed to be shared by Halle et. al. (1957). They believed that the role of the vocal folds vibration was relatively less important. They stated that '...the traditional terms 'voiced' and 'voiceless' seem somewhat inappropriate. ... Instead we shall refer to /ptk/ as tense and to /bdg/ as lax stops'(p. 107). On the other hand, other phoneticians strongly objected to such views and openly questioned the nature of this feature. In this respect Roach (1983) commented by saying that nobody had really proved it. He added that force of articulation is very difficult to define and measure. For this reason we will rely on VOT in the study of voicing in YSA (See section 3.1). It is well defined with clear acoustic boundaries on the spectrograms.

#2.1.1.7 Emphasis:

Another general phonological opposition feature in Arabic is emphasis. It distinguishes between the emphatic sounds /T,D,q, S/ and the nonemphatic ones /t, d, k, s/. Although, this feature does not exist in many languages, it is of particular importance in Arabic and probably in other semitic languages such as Hebrew and Ethiopic Ge'ez.

2.1.2 Acoustic studies on formant transitions:

When a plosive consonant is uttered in combination with a following vowel the vocal tract configuration undergoes several changes from the one appropriate for the plosive to that of the vowel. These changes correspond to and are reflected in the formant pattern variations. Fant (1973, p. 24) summarised this point when he said that "The resulting continuous variations in the dimensions of the vocal cavities determine uniquely the variations of the vocal tract resonance frequencies, the F-pattern".

The extent and direction of the formant transitions was one of the aspects that received great attention by a number of investigators. Most of the investigators who worked on this feature assumed the relevance of these variables for the consonantal place of articulation.

One of the early experiments on this aspect was that by Cooper, Delattre, Liberman, Borst and Gerstman (1952) in which it was suggested that the /p/ sound was characterised by a low frequency burst and a rising transition, /t/ by a high frequency burst and a falling transition; while /k/ was associated with a low frequency burst and a falling transition. These findings implied that the transition direction of each consonant remained the same regardless of the target position for the vocalic contexts. These predictions were examined in English and other languages and most findings appeared to contradict and argue against them. Thus, Fischer-Jorgensen (1954), from data on Danish plosives, rejected this classification because it failed to take into account other factors such as the vocalic contexts. For example, it was assumed that the F2 transitions for /t/ were rising to front vowels. she added that the direction of the transitions was not by itself a relevant factor as much as was the starting point of the formant transition. The latter feature was, for instance, responsible for separating /p/ from /t/. The transitions began lower for the former than for the latter and the direction of the transition was, therefore, determined by the target position of the formant in the following vowel. To that extent the transitions from a certain consonant were invariant for each place of articulation.

On the other hand, Halle et.al. (1957) found falling F2 transitions for /d/ before front vowels, a finding that was similar to that of the Haskins Laboratories. However, they could not specify a locus. From their description of other consonants in the vicinity of the different vowels one could see that the formant transitions were indicative of the place of articulation but only in the same vocalic context. Their observations confirmed Fischer-Jorgensen's point that the transition was not unique for the same consonant in different vocalic contexts.

Earlier research about the interaction between consonants and vowels resulted in the 'hub' and the 'locus' theories (Potter, Kopp and Green,1947; Delattre, Liberman and Cooper, 1955). The 'hub' theory stated that each consonant was associated with a fixed second formant frequency position. The 'locus' was similar to the 'hub' theory but it was more detailed. The locus was regarded as a common point on the frequency scale from which all formant transitions for the same consonant to all possible vocalic contexts should begin. This point was reported to be about 50 ms ahead of the release for English voiceless plosive consonants. The articulatory significance of the locus was claimed to be an invariant vocal tract configuration. The 'locus', in the Haskins Laboratory sense, was believed to correspond to the limiting position by the articulators at the closure state for consonants.While the 'hub' theory was only concerned with the second formant, the locus theory went byond that to include the direction of the movement of the transitions. It also extended the concept to all formants. These characteristics were summarised by Delattre et.al. (1955) when they state that "... there are, for each consonant, characteristic frequency positions, or loci, at which the formant transitions begin, or to which they may be assumed to point. On this basis, the transitions may be regarded simply as movements of the formants from their respective loci to the frequency levels appropriate for the next phone, wherever, those levels might be".

Both theories assumed that the steady state positions of the vowel formants remained the same regardless of the preceding consonants. Some examples of the second formant loci reported by Delattre et. al. (1955) were the following: 700 Hz for labials and 1800 Hz for alveolars. The velar consonants were found to have two varied second formant loci, one at 3000 Hz and the other in a lower frequency region.

Since then the findings and the implications of these theories have been the subject of intensive investigation in order to check their validity and applicability to different vocalic contexts, and to variations in manner of articulation, subjects and language. In English, for example, Stevens and House (1956) examined the transition patterns associated with the different consonant vowel combinations. Their results supported the claim of Delattre et. al. (1955) with several reservations. For example, the F2 locus for the alveolar consonants was relatively fixed. For the labials, however, it varied from 700 Hz to 1500 Hz while for the velars it ranged between 600 Hz and 2500 Hz . The third formant locus was not found to vary from one consonant place to another. Such findings led them to conclude that the locus hypothesis as presented by Delattre et.al. (1955) was not entirely accurate and that it should be modified.

A more extensive investigation of the transitional characteristics in initial and final consonants produced at different places of articulation was conducted by Lehiste and Peterson (1961). Their study differed in that it included almost all of the English consonants and not only the plosives as it was the case with those studies which preceded it. Again their data showed, with the exception of the consonants in the alveolar place of articulation, that the F2 loci were very much dependent on the vocalic context. It was not possible to assign each of the labials and velar places fixed F2 loci. They also emphasized the importance of formants direction.

The picture became more complex as the vocalic environment was considered. From data on Swedish plosives, Öhman (1966) found that the /b/ and /d/ loci in CV-transitions were not independent of the vowel that preceded the CV-sequence. His data showed that the F2 loci were lower when preceded by back vowels and higher when preceded by front vowels. The data also showed overlapping in the F2 locus suggesting that the interpretation of the locus depended upon the preceding and following vowels. He also noticed a coarticulatory influence of the consonant on the formant frequencies at the steady state position of the vowel. The F2 and F3 frequencies of any vowel before /g/ were somewhat different from those of the same vowel before /b/ or /d/. Thus, lower F2 steady state frequency values of the vowel were seen to occur before /d/ than before /g/ although the variations were quite small.

Other findings in line with the above studies were reported by Kewley-Port (1982). In this study fixed loci were not obtained for /b/ and /g/. As a result all formant transitions could not serve as acoustic correlates of place of articulation for these consonants. For the alveolar sound /d /, however, fixed F2 and F3 loci were found. The average F2 locus occurred at 1797 Hz.

To sum-up, most of these studies on English and other languages seem to agree on the conclusion that, with the exception of the alveolar consonants, the F2 loci appear to have failed to characterize the consonantal place of articulation in all vocalic contexts.

Acoustic studies on the role of formant transitions in characterising the phonetic features in Arabic consonants are relatively few as compared with those of English. Some of the studies that dealt with such aspect were Odisho (1973) Giannini and Pettorino (1982) and Al-ani (1970). Both studies were concerned with the transitions as distinctive correlates of the emphatic/non-emphatic categories in the spoken Arabic of Iraq. Odisho (1973) measured the transition frequencies in the first and second formants for both emphatics and non-emphatics. His results showed the emphatics to be associated with a downward shift of F2 by a maximum value of 700 Hz. On the other hand, the F1 was associated with an upward frequency shift reaching up to a value of 250 Hz.

Similar findings for the F1 and F2 loci values associated with each consonant category were also reported by Giannini and Pettorino (1982). They found the non-emphatic consonants to have loci values located at 250 and 2000 Hz for the F1 and the F2 respectively; whereas the emphatics were associated with loci values at 600 and 1000 Hz for the same formants .

The findings from both studies supported the prediction of Klatt and Stevens (1969) when they regarded the acoustic changes in back consonants as a corresponding behaviour to the pharyngeal constriction. Klatt and Stevens used an electrical analog to simulate the vocal tract shape in order to examine the acoustic properties of certain Arabic consonants produced at the pharyngeal cavity. Their main objective was to see the effect of creating a relatively narrow constriction in the pharyngeal area on the shape of the formant patterns. One of their findings was that the F1 and F2 were rather close. F1 was higher and F2 was lower in frequency than consonants produced with a front constriction. They, therefore, hypothesised that similar trends might be expected to occur for realistic vocal tract shapes keeping in mind the influence of changes caused by the vowels that are adjacent to the consonants.

✳All three studies indicated that the main feature distinguishing the two groups was the associated changes in the lower two formants which appeared to reflect the presence of pharyngeal constriction in the emphatic articulation. These findings also agreed with Al-ani (1970) who from X-rays and spectrographic data confirmed that the Arabic emphatics /T/ and /D/ were associated with a pharyngeal constriction. He also maintained that the vowels / i, ii, ʊ,ʊʊ / underwent a general phenomenon of retraction of the back of the tongue suggesting the presence of emphatic coarticulation on neighbouring vowels.

2.1.3..Acoustic studies on the temporal relationships between plosive consonants and vowels:

2.1.3.1 Contoid duration and the voicing distinction:

Since the publication of Lisker's paper (1957) on the closure duration of certain English plosives it was generally assumed that one of the acoustic realizations of the voiced/voiceless contrasts was a difference in the closure duration. The evidence presented in that study was a set of spectrographic measurements showing that intervocalic /p/ within bisyllabic words was regularly longer than /b/. The average closure durations found for /p/ and /b/ were 120 ms and 75 ms respectively. Slightly lower closure duration values in poststressed intervocalic plosives were also reported by Sharf (1962) for English. He obtained mean duration values of 92 ms and 60 ms for the same voicing contrast. Although both studies showed similar patterns with regard to voicing, the extent of the differences displayed by the latter one were relatively small compared with the former one. It was not clear whether such reduction was due to contextual factors or to measurement conventions employed by the two studies. The relation between closure duration and voicing was also examined in some languages other than English. In Japanese, for example, Homma (1981) found longer closure durations for the voiceless plosives as compared to the voiced ones. French, Russian and Korean were observed to display similar behavior with regard to this phenomenon (Chen, 1970). Similarly, from data on connected speech O'Shaughnessy (1984) presented average durations for the voiceless plosives in a read French paragraph that showed them to be consistently longer than the voiced ones by a difference of about 10-15 ms (i.e. 76 ms and 63 ms). Earlier Wajskop (1977) examined French voiced and voiceless intervocalic plosives in natural sentences and provided results that undermined any difference in closure durations as a function of voicing. No one acoustic interval investigated in that study was independently capable of separating French /ptk/ from /bdg/. Nevertheless, when voicing and stress interacted a greater influence on the closure duration was seen.

Further evidence by Lisker (1957) from data based on tape-cutting experiments indicated that such reported difference is of some perceptual significance. If the closure duration for plosive voicing contrasts was appropriately lengthened for the /b/ in *ruby* or *rabid*, the words were heard as *rupee* and *rapid* respectively. Similarly, Port and Dalby (1982) provided evidence supporting the perceptual relevance of the consonant/vowel ratio (i.e. ratio of the closure duration to the vowel duration) for the voicing distinction in English. They demonstrated that when other cues to voicing were ambiguous in words such as *dibber-dipper* or *digger-dicker*, this ratio proved to be the primary cue to voicing for syllable-final plosives. The foregoing were examples of phonemic distinctions cued by the duration of a sound segment.

Other perceptual work tried to establish the durational limits for a certain consonant category regardless of its phonological attribute. Such studies concentrated on the maximum/minimum durational values and the just noticeable differences (JND) each segment might have (Lehiste,1970; Huggins,1972; Klatt, 1976, 1977). For example, Huggins investigated, among other things, the acceptable durational ranges for the English sound /p/ in a variety of contexts. The longest acceptable /p/ closure in intervocalic position was about 85 ms, while the shortest one was in the range 34-45 ms. On the basis of this finding he advanced the hypothesis that the JND for a segment might be determined by the articulators that produced it. He predicted that segments which were produced by the same articulators might have similar durations.

On the same line Klatt assumed that each segment type had a certain minimum duration required for the satisfactory production of that segment. He determined the inherent durations for the English plosives /t,k,d,g/ to be 65 ms while the minimum values varied according to voicing. It was 40 ms for /d,g/ and 50 ms for /t,k/. From these as well as other findings, he proposed some durational rules to be used as indicators for capturing the segmental durations and the timing control in English speech.

On the other hand, acoustic evidence from data based on connected speech showed the closure duration to yield inconsistent and unreliable differences between the voicing categories. In Umeda's (1977) study contradictory results to earlier ones were presented for one subject. Longer durations for the voiced alveolar and velar consonants than for their voiceless cognates in the intervocalic prestressed environment were shown. Another such example could be seen in Stathopolous and Weismer (1983). Although the results presented for plosive voicing in word-final position confirmed Lisker's (1957) conclusion, in initial position the reverse effect was observed, with the voiced plosives having longer closure duration than the voiceless ones. Sharf (1962), too, who measured the closure interval of poststressed intervocalic plosives, found insignificant difference but only for the alveolar voiced and voicelesss contrast /t/ vs /d/. The previously mentioned differences in the closure duration for voicing in the velar and the bilabial cosonants were virtually eliminated in words such as latterladder in American speech. In Danish, too, the set /bdg/ were associated with longer closure than /ptk/ (Jorgensen,1977). In this particular language both /bdg/ and /ptk/ sets were described as voiceless differing only in the presence and absence of aspiration. A relatively recent report by Lisker (1972), on a single speaker, suggested that when the closure duartion was examined in a variety of positions the voiceless plosives had greater durations than the voiced ones only in the intervocalic poststressed positions. This was a clear indication that what was assumed to be an established feature of voicing turned out to be far from being a regular one.

Additional evidence in support of this statement could also be found in the more recent and detailed studies by Crystal and House (1982; 1988) and by Luce and Charles-Luce (1985). These studies showed evidence that questioned the utility of the closure duration as an acoustic correlate of voicing. The former authors found the difference between the voicing cognates to be on average less than 10 ms. Similarly, Luce and Charles-Luce (1985) concluded that the closure duration failed to invariably distinguish plosive voicing, while the preceding vowel duration was proven to be "the most reliable correlate of voicing".

2.1.3..2.Vocoid duration and voicing:

In addition to their influence on the consonant duration the vowels were themselves affected by the consonantal environments. The vowel duration was determined by a number of factors some of which were the consonant voicing and the place of articulation (Denes,1955 ; Zimmerman and Sapon,1958; Mack,1982; Delattre,1962; Peterson and Lehiste,1961; Jorgensen,1964,1977; Fant,1973, Chen,1970; Wajskop,1977; Homma,1981). In English, for example, it was shown that the vowel duration varied significantly as a function of the voicing characteristics (House and Fairbanks, 1953). Vowels in a voiced environment were longer than those in a voiceless one (cf. Crystal and House,1988). If voicing was kept constant, the consonants that differed in the manner of production influenced the vowel duration in different ways. The vowel lengthening effect in the fricative environment was greater than that in a plosive environment. In his cross-language study of vowel length, Chen (1970) concluded that the vowel was invariably longer before a voiced consonant than before a voiceless one in all four languages investigated: English, French, Russian and Korean. Although the significance of the difference depended on the language, he argued for the incorporation of vowel length as a 'universal' phonological feature of voicing.

Various interpretations were advocated for the occurrence of such lengthening. Raphael (1981) concluded that the durational differences between vowels preceding voiced and voiceless consonants were controlled physiologically. He believed that the articulators were maintained somewhat longer for vowels when they preceded voiced consonants. On the other hand, Zimmermann and Sapon (1958) preferred to explain it by the inherent characteristics within each language. The vowel length variability was attributed to the differences in the speed of the transition from the vowel to the consonant closure (Chen, 1970). He predicted that the transition from a vowel to a voiceless consonant closure would be faster than the transition from a vowel to a voiced plosive consonant.

On the other hand Chen's assertion, that it was only the preceding vowel rather than the following one which exhibited a lengthening effect, was contradicted by evidence from several other languages. For French an average difference of 17 ms for the word final vowel in voiced and voiceless environments was reported by Wajskop (1977). This effect, however, might be a language specific phenomenon.

In addition, in Japanese both the preceding and the following consonants were also seen to have a voicing effect on the vowel length (Homma, 1981). He noted that the voicing of the preceding plosive environment inflected a stronger influence on the vowel duration than that of the following plosive environment. Other languages such as Danish (Jorgensen, 1964, 1977) and Swedish (Fant, 1973) showed similar influence on the vowel duration by the preceding consonants. Therefore, vowel length did not vary only as a function of the following consonant voicing but as a function of the preceding one as well.

2.1.3..3 The Contoid /Vocoid ratio and voicing:

In addition to the consonant and vowel durations as independent correlates of voicing, some researchers expressed their desire for combining them into one measure by relating the duration of one of them to the other. The concept was first introduced by Denes (1955) who demonstrated that the subjects' ability to detect the difference between the voicing categories depended on the relative duration of both the vowel and the final fricative. For example, the final fricative sound in the words 'his' and 'hiss' was heard as [s] when it had the same duration as the vowel that preceded it or longer. It was heard as [z] when the fricative was much shorter than the vowel. The most explicit form of this hypothesis was introduced by Port (1981). In that paper he predicted that the durational changes in the voicing of a syllable final plosive could be expressed as a ratio between the vowel duration and the closure duration. His conclusion from the results he obtained was that the C/V ratio "...is an extremely powerful cue for phonological voicing in English". The concept was also tested perceptually by Port and Dalby (1982) who systematically varied the consonant and vowel durations in the words 'dibber/dipper' and presented the stimuli for identification at different rates of speech. Again the responses confirmed that the main ratio of the consonant/vowel remained constant in spite of changes in rate.

On the other hand, not all investigators who used the consonant/vowel ratio shared Port's conclusion. In fact there was evidence that argued against it (Massaro and Cohen, 1983). One of the arguments they presented was that decreasing the rate of speech did not automatically lengthen the vowel and the consonant by the same duration. Similarly, increasing it did not necessarily imply equal shortening of both segments. They indicated that the two components should be treated as independent correlates of voicing. Other attempts to test the consonant/vowel ratio were made by Luce and Charles-Luce (1985) and Crystal and House (1988). They, too, failed to provide any evidence that might be regarded as a confirmation of that hypothesis.

2.1.3.4 Contextual factors and segment duration:

It was also expected that both the closure and vowel duration would vary according to changes in certain factors such as stress, the number of syllables, rate of speech, position, place and manner of articulation and gemination. For example, segments in stressed syllables were longer in duration than those in unstressed ones. (Stathopoulos and Weismer, 1983; Klatt,1976; Fry 1958; O'Shaughnessy, 1984). The stress effect varied depending on whether the segment was consonant or vowel and also on the structural properties of the language itself. It was greater for vowels than for consonants. The loss of stress reduces the length of the following vowel by about 70 ms in French (Wajskop, 1977). For English, the average duration for stressed vowels was approximately 130 ms as compared with 70 ms for the unstressed vowels in connected speech (Klatt, 1976). On the other hand, very small differences in vowel duration in stressed and unstressed contexts were found in Japanese. The mean values reported were 91 ms and 95 ms respectively (Homma, 1981). However, in a cross-language study of duration Fant, Kruckenberg and Nord (1991) investigated the durational correlates of stress in three languages: Swedish, English and French. In particular they tried to find the extent to which language specific patterns of syllable structure can affect the overall duration of stressed and unstressed syllables. Their results showed that stressed syllables in Swedish and English, which were both stress-timed languages, attained lengthening of the order 100-150 ms in positions other than before a pause. In French, which was a syllable-timed language, the syllable lengthening is of the order 50 ms.

Most investigators agreed, when other factors were kept constant, that closure duratiom for labials was longer than that for alveolar and velar consonants (Umeda, 1977., Wajskop, 1977, Lehiste, 1970; Stathopoulos et.al. op.cit., Smith, 1978; O'Shaughnessy, 1981; Jorgensen, op.cit., Port, 1979; Luce and Charles-Luce 1985, and Homma, 1981). This appeared to be true for all positions and voicing-stress combination. The extent of the difference was very much dependent on the language. A difference of about 20 ms between labial and alveolar plosives was found in French (O'Shaughnessy, 1981), while in Japanese the difference was very small indeed (Homma, 1981). Concerning the difference between the alveolars and velars, the available findings were inconsistent and in many cases insignificant. Some of the studies reported the alveolars to be associated with longer closure than the velars (Umeda, 1977., and Fiscer-Jorgensen, 1977) while others found the reverse order (Port, 1979, Wajskop, 1977, Luce and Charles-Luce 1985).

Previous research also suggested that the closure duration was strongly correlated with the single-geminate contrast. Thus Lisker (1958) from a cross-language study of the phenomenon in four languages (Italian, Swedish, Marathi and Telugu) found the geminates to be considerably longer than the single forms. This conclusion was also confirmed by the responses he obtained from perceptual tests for Tamil words. In Japanese the closure duration ratios for single and geminate plosives was 1:3. The influence of voicing was greater on single forms than on geminates.

A survey of the literature showed that most research on the gemination feature was in the form of perceptual tests. For example, in order to locate the phoneme boundary Pickett and Decker (1960) lengthened the closure for /p/ in the word *topic* from 60 ms to 585 ms. Their results showed that closure longer than 250 ms was heard as /pp/, while a closure duration lower than 150 ms was heard as /p/. Similarly, Obrecht (1965) manipulated synthetic stimuli to localize the perceptual boundary between single and geminate consonants in Arabic. Varying the closure duration for intervocalic /b/, duration of the nasal resonance for intervocalic /n/, and duration of frication noise for initial /S/, the perceptual boundaries were determined between 140-160 ms for /b-bb/ and /S-SS/ contrasts and 90-110 ms for /n-nn/. He concluded that duration was a strong cue for discriminating between geminates and single forms in Arabic. Commenting on the sharpness of the location of the phoneme boundaries in his study, as compared with that of Pickett and Decker (1960), he indicated that such an action might have been a reflection of the differences in the structure of the two languages and also the familiarity of the subjects with that distinction. According to him the sharpness of responses displayed by the Arabic subjects might be related to the frequent occurrence of gemination in their language.

More recently, Lahiri and Hankamer (1988) examined the consonantal length opposition in two unrelated languages: Turkish and Bengali. They, too, concluded that the primary variable distinguishing the opposition was the closure duration while VOT and vowel duration play only a secondary and perhaps unimportant role in that distinction.

2.2. Survey of the perceptual studies on plosives with reference to emphasis and voicing.

2.2.1 Perceptual Studies on Voicing.

Some acoustic features are important for the voicing distinction in word-initial position. Of these acoustic cues are voice onset time (VOT), the fundamental frequency of voicing (F0) and the first formant transition. The perceptual importance of these acoustic features can be examined through use of synthetic speech as demonstrated, for example, by Liberman, Delattre and Cooper (1958) Stevens and Klatt (1974) Abramson and Lisker (1968) Lorge (1967) Zlatin (1974) and Fujimura (1971).

Liberman et. al. (1958) synthesized CV-syllables using the Voback Synthesizer of the Haskins Laboratories. In their experiment they showed that by cutting back and blanking out the initial part of the first formant transitions by various amounts the voiced response was reduced. When the cutback reached the onset of voicing the perceived responses were voiceless. Abramson and Lisker (1968) conducted identification and discrimination tests of voice timing in two languages: English and Thai. They synthesized three continua of VOT variant stimuli representing the contrasted pairs /p,b/, /t,d/ and /k,g/ on the parallel formant synthesizer in the Haskins Laboratories. Their conclusion was that "There is quite a good match between the perceptual crossover zones and the production boundaries even though the test subjects were not the ones who provided the utterances for measurements". Other relevant perceptual and productive studies of the VOT in English were done by Lorge (1967) and Zlatin (1974). The findings from both studies confirmed the results obtained by Abramson and Lisker (1968). The judgements they got from their listeners indicated that there was a sharp perceptual boundaries between the voiced and the voiceless responses.

On the other hand, Stevens and Klatt (1974) in their turn emphasized the role of the first formant transitions in the voiced-voiceless distinction of the alveolar place of articulation. They used synthetically generated syllables. While they did not rule out the importance of the VOT, they showed that with greater transition durations there was an increase in the VOT value at the boundary between synthetically generated consonant-vowel syllables /da/ and /ta/. They concluded by suggesting that there were two related cues underlying the listeners' judgements. The long VOT usually indicated a voiceless consonant, whereas the rapid first formant transition at the onset of voicing was a characteristic of voiced ones. They also added that if the transition was completed before the onset of voicing the consonant was identified as voiceless and if the VOT was short (less than 20 msec.) the consonant was perceived as voiced.

Another cue which was reported to be effective in the perception of voicing was the fundamental frequency. Fujimura (1971) noted that a change from /ka/ to /ga/ or /ta/ to /da/ could be obtained just by varying the fundamental frequency inflection, particularly in the phoneme boundary region. Haggard, Ambler and Callow (1970) used stimuli differing only in the F0 curve at voicing onset. They, too, showed that most subjects were able to categorize labial plosives as voiced and voiceless depending on whether there was a low rising F0 contour or a high falling one. They explained this by suggesting that a low rising fundamental frequency contour indicated a closed glottis, whereas a high falling one indicated a glottis that was partly open. From these studies it could be seen that listeners responded to different acoustic cues in judging the plosive voicing categories and that no acoustic cue by itself could be claimed to be responsible. However, it was reported in a number of studies that the VOT was the primary cue of voicing as compared with F0 or F1 transitions. In this respect Fujimura (1971) indicated in his treatment of the relative importance of these cues that the VOT was a primary one for the plosive voicing distinction and that the effect of F0 as a secondary cue was easily overridden. Similarly, Lisker (1975) in a study entitled "Is it VOT or a first formant transition detector?" stated that the presence of a voiced F1 transition was not a requirement for the plosive sound to be heard as voiced. Its absence was not a requirement for the perception of the English voiceless plosives /ptk/. Later on, Lisker, Liberman, Erickson and Dechvitz (1977) compared the relative importance of the two measures: VOT and the voice transition duration (VTD) of the first formant. The conclusion drawn from their study was that VOT was more effective and superior than the VTD in distinguishing the plosive voicing in English. Consequently, our investigation will focus on the role of VOT in the perception of the voicing categories in YSA (See section 4.1).

2.2.2 Categorical perception and the phoneme boundary:

Categorical perception was defined as that process by which listeners responded to each category in absolute terms and all the stimuli in a certain continuum were perceived to be in either one category or the other. Studdert-Kennedy, Liberman, Harris and Cooper (1970, P.234) described it as "a mode of perception by which stimuli are responded to, and can only be responded to in absolute terms". In this detailed reply to Lane's critical review (1965) they set certain criteria and requirements for the occurrence of categorical perception. First, they stated that the identification function between phonemes should be sharp and abrupt. Secondly, distinct peaks in the discrimination curve ought to be seen at the phoneme boundary. Thirdly, points of lower regions (or "troughs" as they preferred to call them) should also be seen on the discrimination curve at the chance level on either sides of the phoneme boundary.

2.2.3 Perception studies on formant transitions:

Researchers in the last four decades have been trying to uncover the various processes and stages in speech perception. Serious scientific work began in the Haskins Laboratories with the invention of a Synthesizer called 'the Pattern Playback' that converted simplified hand drawn spectrographic patterns into sounds. These patterns were prepared from original spectrograms first by identifying the important acoustic features and then by reproducing them using that machine. The drawing of the patterns involved some approximation as indicated by Cooper, Delattre, Liberman, Borst and Gerstman (1952).

The main objective of their studies was to examine as many perceptual attributes of speech as possible using systematically varied sythetic stimuli of certain acoustic parameters in the hope that eventually a coherent picture would emerge. Their long term goal was to build a model which could explain the different perceptual processes on the segmental level and when appropriate on higher level processing (i.e. at the syntactic and semantic levels).

Consequently, in order to achieve their aim they started investigating the perceptual properties of different portions of the speech signal. One of these was the rate and direction of the formant transitions. The term transition as described in their studies was equivalent to the rapid changes at the boundaries between consonant and vowel. For English it was shown experimentally that these transitions carried a great deal of the information leading listeners to the identification of the adjacent consonant. For example, Cooper et. al. in their (1952) study predicted that since the points of articulation for the consonants and vowels occurred in different positions along the vocal tract, it followed that the transitional movements from one articulatory position to another would result in corresponding changes in the acoustic pattern. Based on two formant stimuli of synthetic speech they systematically varied the extent and rate of the second formant transition in the voiced plosives /bdg/ in the environment of seven different vowels. The identification curve showed that all subjects were able to hear a rising second formant transition as /b/ and that falling transitions might be heard as /g/ or /d/ depending on the following vowel.

Later on Liberman, Delattre, Cooper and Gerstman (1954) found that rising transitions were in general heard as /p/ or /b/ whereas falling transitions to the steady state of the second formant were heard as either /t, d/ or /g, k/ depending on the size of the transition and the vowel with which it was paired. In the case of /d, t/ the F2 onset frequency was found to be near the second formant of the vowel /e/. From this position the second formant transition was expected to rise or fall to the steady state of the vowel. The stimuli were perceived as /t,d/ if the F2 transitions were slightly rising for front vowels or sharply falling for back vowels. /k,g/ were perceived if the F2 transitions were sharply falling for front vowels or slightly falling for back vowels. According to Liberman et. al. (1954) no plosive consonant was perceived without some degree of second formant transitions particularly when the first formant was straight. Based on these findings they went on to predict the existence of a second formant locus[2] for certain consonants. The subjects' labelling function showed that the frequency positions at which the plosive consonants /g/, /d/ and /b/ were best heard at approximately 3000 Hz, 1800 Hz and 700 Hz respectively.

---

[2] For details on this theory see section 2.2

Recently, the locus theory was supported not only from CV- but also from VC- transitions. Using natural CV- and VC- syllables representing plosives and fricatives in the context of the neutral vowel /ə/, Sharf and Hemeyer (1972) tried to determine the importance of formant transitions in the perception of place of consonant articulation. The aspiration portion was eliminated from transitions. Results showed that the subjects were able to identify voiced and voiceless in both CV- and VC- transitions. They were more successful with the latter type than with the former. This was explained by what they called 'the transition sufficiency hypothesis' which assumed that the extent of the articulatory movements during the CV- transition was not as sufficient as that of the VC- transition. They added that this lack of transition sufficiency in the CV- syllables might have been caused by the elimination of the aspiration noise.

To sum up, the previous studies demonstrated that formant transitions and particularly those of the second formant were generally shown to be sufficient cues for the plosive consonant recognition. They emphasized the role of context-conditioned cues within the consonant-vowel syllable.

In spite of that other researchers believed that plosive consonant perception was independent of the following vowel (Cole and Scott, 1974a, 1974b; Blumstein and Stevens, 1980). Those researchers believed that there were invariant perceptual cues for each plosive consonant corresponding to its place of production. They believed that there was a one to one relationship between the phonetic features of sounds and the acoustic signal, a claim strongly rejected by the advocates of the Motor Theory. Thus, Cole and Scott (1974a) showed that a burst excised from its original vowel context and transposed onto a different vowel context was still identified accurately for place of articulation. These results were obtained for labial and alveolar consonants, although velar consonants were not identified consistently across such transformation. They came to the conclusion that the perception of consonants was characterised by invariant properties in the brief onset spectra of the CV-syllable. Similar findings were reported by LaRiviere, Winitz and Herriman (1975) and Winitz, Scheib and Reeds (1972). Both studies suggested that the burst noise was a primary cue and that the vocalic transition was neither sufficient nor a necessary cue for the recognition of initial voiceless plosives. A more recent and extensive work representing this view can be found in Stevens and Blumstein (1978); Blumstein and Stevens (1980). The outcome of these studies was that the brief onset spectra particularly at the first 10 to 20 msec provided invariant properties for place of articulation of the plosive consonants studied regardless of the following vowel contexts. These spectral invariant cues were described as 'diffuse-rising' for /d, t/ 'diffuse-falling' or 'diffuse-flat' for /b,p/ and 'compact' for /g,k/. Comparing the relative role of the noise burst and the formant transitions they remarked that the latter play only a secondary role in consonant identification.

To sum up, there appears to be two conflicting views regarding the perceptual importance of the transitional cues. The first emphasized the context conditioned cues and that no plosive consonant could be perceived without the inclusion of some portions of the vocalic transitions. The other view believed that these cues played a small marginal role and that the primary invariant cues could be found in the brief onset spectra of the consonant-vowel syllable.

2.3 Survey of articulatory and aerodynamic studies of voicing and emphasis in plosives:

2.3.1 Aerodynamic factors and voicing:

Speech is often described as "modified breathing". The respiratory process consists of air coming into and out of the lungs. This is achieved by varying the volume and the pressure inside and around the lungs by muscular and other forces acting on the thoracic cavity. Speech is almost always produced in the expiratory phase of the respiratory process. It is a process of transformation of airflow into sounds through the movements of the articulators and the contraction of certain relevant muscles in the chest, laryngeal, pharyngeal and oral tracts. In addition, the air coming from the lungs is obstructed at several points along the vocal tracts. This action helps to convert the airstream into intelligible speech (Hardcastle, 1976; Borden and Harris,1984; Ohala,1980).

Aerodynamic data are considered to be useful because measurements of oral air pressure and oral flow during speech production can provide valuable information from which inferences about the function of the articulatory and the physiological mechanisms can be drawn. Simultaneous recording of pressure and flow should provide a more complete understanding of processes of speech production. Aerodynamic conditions near the constriction combine with articulatory shapes and movements to generate all the acoustic sources. There are several stages in the speech communication process. The aerodynamic stage is an important one in that chain. It "provides the link between the articulatory and the acoustic stages of speech communication" (Scully, 1974). Some, perhaps a lot, of the complexity of speech acoustic patterning arises from the processes of mapping from articulation to acoustics, in which aerodynamics plays an important part. If that is the case, then the aerodynamic variables would provide us with further insights to explain and interpret our acoustic results.

The importance of aerodynamic studies in characterising certain phonetic phonemena as recognised earlier this century with the work of Stetson (1951) and Hudgins and Stetson (1937). These researchers among others laid the foundations for experimental techniques. The use of the Kymograph and other associated instruments enabled them to study air pressure and airflow and to reach conclusions never found before in spite of the obvious limitations of their instruments. Some of their findings were helpful and pioneering, others, however, were controversial. As far as the plosives are concerned the two sets /ptk/ and /bdg/ are differentiated by air pressure variations above and below the glottis. According to Stetson (1951) this difference was unrelated to glottal actions in anyway. Further he claimed that in rapid speech the voicing feature failed to distinguish between the two sets /ptk/ and /bdg/, whereas the force of articulation proved to be more powerful in separating them into 'fortis' and 'lenis' (p. 93). Since then a number of aerodynamic studies have been conducted on English and other languages. Some gave further support for Stetson's claim; others were against it. In what follows, the results of these studies are reviewed and examined critically in the lights of the recent developments and the improvemnts of experimental phonetic methodology.

The aerodynamic parameters are found to be useful and powerful indicators of the voicing feature. Black (1950) reported systematic variations in oral pressure between the English voiced and voiceless consonants with the latter showing greater air pressure values than the former. In most cases where flow rate was measured with a mask air pressure was measured with an equipment which was cumbersome and less accurate. He, therefore, urged caution in the interpretation of his data due to the apparent limitation in his methodological approach. Isshiki and Ringel (1964) investigated the airflow rate during the production of some English consonants in (CV-) and (-VC) syllables. They concluded that variations in the peak airflow values occurred as a function of voicing. Greater airflow values were associated with the voiceless consonants than with the voiced ones.

Unsatisfied with the methods used in the previous studies, Subtelny, Worth and Sakuda (1966) conducted an investigation on both the air pressure and airflow variables with slightly different instrumentation. They used a 'hot wire anemometer' instead of a mask. They found lower intraoral air pressure amplitudes, shorter durations and lower airflow rate for the voiced plosives as compared with the vioceless ones. They also claimed that their data showed lower pressure values for the English consonants than in other studies (e.g. Black (1950) and Isshiki and Ringel (1964). They attributed this lack of correspondence to the differences in the reseach instruments used in both studies. They indicated that the higher air pressure and airflow values obtained by Black (1950) and Isshiki and Ringel (1964) may have been caused by an artefact of their instrumentation.

The limitations of many air pressure and flow sensing systems were addressed by some researchers (e.g. Scully, 1969). In her report she listed some of the problems associated with the use of the aerometer and the interpretations of pressure and flow data. Of the many practical and theoretical related problems she mentioned the following:

1- The adjustment of the mask to fit different faces.

2- The mask might restrict jaw movements.

3- The mask reduced auditory feedback to the speaker which made him speak in an unnatural way.

4- Oral airflow in and out were combined in one channel. The baseline was then not visible during speech. But this problem could be solved by asking the speaker at the start of recording to hold his or her breath.

In spite of its limitations the face mask flow unit would appear to offer fewer problems than any other available sensing system. For instance, the hot-wire flow meter has its own inherent problems. It always gave positive indications for both inspiratoy and expiratory flows (Gilbert,1973). In addition, its frequency response to airflow variations associated with speech production was not adequate (Rothenberg, 1977). Distortion to the recording could occur as a result of the environmental turbulence. On the other hand, improvements to the pneumotachometer flow unit were demonstrated by Rothenberg (1977). This solved many of the problems discussed above particularly those related to the frequency response and the speech distortion of the acoustic output.

Gilbert (1973) found for English that the voiceless plosives were associated with greater oral airflow rates than their voiced cognates regardless of phonetic context. Other airflow studies showed results which were consistent with the above studies. Thus Stathopoulos and Weismer (1985) reported that adults produced significantly higher airflow values in voiceless plosives than in their voiced cognates. Warren and Wood (1969) also found the total air volume to be considerably higher during the voiceless sounds than during the voiced ones.

On the other hand, some aerodynamic studies failed to find any differences in the airflow rates as function of voicing. Nihalani (1975a) based on an airflow study of Sindhi, a four category language, claimed that his results showed similar airflow values not only for the voiced aspirated versus the unvoiced aspirated but also for voiced unaspirated versus unvoiced unaspirated plosives. This unexpected finding was criticized by Dixit and Brown (1986). For Hindi plosives they observed significant differences in the peak oral airflow values for the voiced aspirated and the unvoiced aspirated plosives. In an earlier study of the laryngeal behaviour (Dixit, 1975) claimed that the size of the glottis during the production of Hindi voiced aspirated plosives is almost half that for the unvoiced aspirated. What was unexpected in their results, however, was the lack of variations in the flow rate across voiced and voiceless unaspirated plosives.

On the other hand, aerodynamic studies based on the supraglottal air pressure measurements have repeatedly shown distinct values for the voicing category. For example, Arkebauer, Hixon and Hardy (1967) performed some measurements of oral air pressure for ten English adults and ten children. Their data revealed that greater air pressure values were associated with the voiceless plosives than with their voiced counterparts. These results were consistent for both children and adults. Similarly Miller and Daniloff (1977) measured the peak oral pressure, closure durations and the peak sound pressure for the English plosive consonants in both citation and conversational contexts. Their data revealed that the voiceless plosives were systematically produced with greater peak oral pressure values in all contexts than their voiced counterparts. They also reported significant differences as a function of context. Commenting on the systematic patterns observed for these variables, they stated that " ..these measures may be viewed as more robust, compared to a burst-oriented VOT measure".

So far these studies partly agreed with Stetson (1951) in that the two sets /ptk/ and /bdg/ involve variations in the peak oral pressure values. However, none of them supported his claim that the glottal action was irrelevant for the voicing distinction. On the contrary there is a constant reference to the importance of the role of glottal and supraglottal obstructions to the pulmonary airstream. According to Warren (1976, p.108) the difference between the airflow patterns in breathing and in speech results from an orifice type of airflow which develops as constrictions from within the respiratory tract during phonation. This type of airflow might exist at the glottis, the pharyngeal orifice or at various points along the oral air passage. Similarly, Van den Berg (1957) believed that the vocal folds vibrations were to a great extent the outcome of aerodynamic and muscular forces which interacted in an appropriate way. Of these aerodynamic forces he mentioned subglottal air pressure, the Bernoulli force produced by negative pressure created transglottaly by high velocity airflow and supraglottal pressure produced by articulatory constrictions in the upper airway.

Scully (1969) wrote that the vocal tract might be thought of as an "electric circuit" with the volume flow rate as the acoustic equivalent of current and air pressure as the equivalent of voltage. Anything which obstructed the airflow might result in a pressure drop and might be called resistance. She listed three types of major vocal tract resistances to the airflow rate:

(1) Glottal resistance during voicing.

(2) Oral constriction caused by the raising of the tongue or the lips against the passive articulators.

(3) Velopharyngeal constriction.

Of these three the first was thought to be the one which accounted for most of the airflow resistance at least for vowels.

The relationship between the cavity size and the peak oral air pressure and airflow was investigated by Subtelny et. al. (1966). In particular they tried to determine the effect of vocal tract size on the intraoral pressure amplitude. Their results confirmed their predictions that larger pressure amplitudes as more associated with children than adult females. The lowest were observed with adult males. Based on these results they concluded that higher oral pressure values were associated with smaller cavity sizes and lower ones were associated with larger cavity sizes. On the other hand, Brown and McGlone (1969) observed no relationship between intraoral pressure and oral cavity size.

2.3.2 Aerodynamic factors and the 'fortis' vs 'lenis' distinction:

Some researchers (e.g. Malècot, 1955, 1966a, 1966b, 1968) believed that variations in the oral air pressure might be regarded as an evidence in favour of the 'fortis' vs 'lenis' distinction in the English plosives /ptk/ and /bdg/. In his (1966) study Malècot tested the effectiveness of the peak oral air pressure value in the distinction between plosive cognates in nonsense syllables. His main concern was to see if the peak air pressure values could be used as an 'index' for the articulatory energy associated with the production of plosives. Higher peak pressure values were observed for the plosives /ptk/ than for /bdg/. Accordingly, he classified the two groups /ptk/ and /bdg/ as 'fortis' and 'lenis' respectively. It seems that the higher pressure values obtained for the 'fortis' plosives were assumed to be caused by the strong energy of the articulators while the reduced oral air pressure values found for the 'lenis' were explained by a weaker energy of the articulators. However, he seems to acknowledge at least in part the relevance of the laryngeal action to the pressure variations for the two sets.

Other researchers (e.g. Lisker, 1970) argued against the inclusion of such terms in phonetic description. Having expressed his doubt about the relevance of these terms, Lisker investigated the peak oral pressure values for one American subject. His concern in that study was to determine whether the two sets of plosives /ptk/ and /bdg/ could be distinguished solely on the basis of oral air pressure as 'fortis' and 'lenis'. His data showed a great deal of overlap in oral pressure values between the two sets of plosives. This led him to conclude that "the measure of peak pressure provides a less than completely adequate basis for separating /p, t, k/ and /b, d, g/ in all positions in which both sets are found"(p. 223). He emphasized the point that differences in oral air pressure did not constitute evidence for the existence of a 'fortis' versus 'lenis' dimension independent of voicing. He argued that this evidence invalidates Stetson's (1951) claim and that his results lend further support for the acoustic measure of VOT which he described as "no less and quite possibly more effective in separating the categories of English stops'(p. 226).

In addition, Chomsky and Halle (1968) claimed that the feature 'tense' and the aspiration associated with it are positively related to the subglottal pressure. This claim was investigated by Kim (1965), Dart (1987) and Lee and Smith (1972) who used data based on Korean plosives. Kim (1965) reported higher oral pressure values for the heavily aspirated plosives than for the other two types of voiceless plosives. Similarly Dart (1987) found higher oral pressure for the 'fortis' as opposed to the 'lenis'. It should be remembered that the 'fortis' and 'lenis' distinction in these studies as used in quite a different sense from that in either Stetson (1951) or Malècot (1968). In these studies the tense-lax contrast is equated with the presence versus absence of 'heightened subglottal pressure and aspiration'. Hence, these results are not cmparable with the above studies.

Lee and Smith (1972), on the other hand, reported results for Korean plosives showing heightened subglottal pressure for heavily aspirated plosives. This finding might be taken as an evidence in favour of Chomsky and Halle (1968). More importantly, however, Lee and Smith did not find any supporting evidence linking the tense-lax contrast with the value of subglottal pressure. Their data were obtained from measurements of oral and direct subglottal pressure. Their conclusion was that "there was no systematic difference in subglottal pressure between the tense and the lax stops".

To our knowledge the only study whose results showed positive relation between the subglottal pressure and aspiration was that by Nihalani (1974) for Sindhi plosives. The subglottal pressure values were obtained through use of an oesophageal pressure sensor. He reported higher subglottal pressure values not only for the aspirated as opposed to the unaspirated plosives but also for the voiced as compared with the voiceless ones. Caution, however, is needed in the interpretation of these results because of methodological limitations associated with this study particularly in the technique used for obtaining the subglottal pressure values in his study. (See Kunze, 1964 for a detailed review of the various techniques used for obtaining subglottal pressure values).

The results of the above studies seem to be inconsistent. No sufficient and reliable evidence in favour of Chomsy and Halle's claim was found. On the contrary, Scully (1976) based on modeling of some English consonants reached the conclusion that variation in the subglottal pressure was perhaps "...the least likely to be employed in the production of aspirated plosives" (p.147).

With regard to the relation governing the subglottal pressure and voicing, there appeared to be a general agreement among most researchers that the respiratory system in speech seemed to generate a constant subglottal pressure regardless of the voicing of the plosives or the presence or absence of aspiration (McGlone and Shipp 1971; Netsell, 1969; Lofqvist, 1975 and Ohala,1980). If some subglottal pressure differences exist between the voicing cognates then it should be explained with reference to the glottal and supraglottal resistance (Ladefoged, 1963).

In brief, most studies seem to agree on the conclusion that the plosives /ptk/ are associated with higher peak oral air pressure than /bdg/. The labels 'fortis' versus 'lenis' or 'tense' versus 'lax' given to them by some phoneticians on the basis of these definitions have not yet been proven experimentally.

2.3.3 Suggestions by various researchers about possible voicing mechanisms

Several interpretations were suggested as possible mechanisms for the occurrence of air pressure and flow variations and their reliable role in the distinction between the two sets of plosives. Lisker, Abramson, Cooper and Schevy (1969) suggested that this voicing distinction was caused by the closing and opening of the glottis rather than by an independent fortis-lenis contrast. They used a transillumination of the larynx in running speech. In this technique "a miniature incandescent bulb is introduced into the laryngeal vestibule through the nose, while a photocell placed below thyroid cartilage registers the variable light transmitted through the glottis and the tissues of the neck (p. 1544)". They observed that during the voiced plosives the vocal folds were approximated and the glottis was virtually closed and remained closed at and after the release of oral closure. According to them the air from the lungs , in most cases, met with a more severe vocal folds obstruction during voiced plosives (even after the oral release) than was the case with voiceless plosives.

Other studies seem to indicate that the variations in supraglottal pressure between the voiced and the voiceless plosives could perhaps be interpreted with reference to the voice maintaining mechanisms during the closed phase of the voiced plosives. Various interpretations of how this might be achieved were advocated by direct and indirect observations of the glottal and supraglattal structures. Rothenberg (1968) pointed out that for the voicing to be maintained during the closed phase of voiced plosives the subglottal pressure needs to be higher than the supraglottal pressure. He added that the transglottal airflow during the closure equalized the subglottal and the supraglottal pressures. Once this equalization has occurred the voicing ceases to exist. Accordingly, for voicing to be maintained during the closed phase of the voiced plosives, some mechanisms for absorbing the transglottal airflow are necessary to reduce the rate of rise in supraglottal presssure during the production of these voiced sounds. For the achievement of that purpose Rothenberg suggested that one or both of the following two possible mechanisms might be used: (1) passive enlargement of the supraglottal cavity as a result of overpressure and (2) active enlargement caused by muscular contraction. Consequently, the supraglottal air pressure rise might be reduced by one or both of these mechanisms; although he added that the latter is more likely to be used for that purpose than the former.

Kent and Moll (1969) based on a cinefluorographic film of the vocal tract during the production of American English plosives observed that "the voiced stop is produced with a larger supraglottal volume than its voiceless cognate". Like Rothenberg, they believed that the pharyngeal expansion was probably caused by active movement of the solid structure and appropriate muscular contraction rather than invoked by a passive response to overpressure. Such pharyngeal cavity enlargement during the voiced plosives was associated with a depression of the hyoid bone and consequently a lowering of the larynx. The laryngeal lowering was found by Hudgins and Stetson (1937) who reported that this action resulted in an oral air pressure reduction of 7 cm $H2O$ in voiced plosives.

In studying the relation between the transglottal airflow during the closure of voiced plosives and the supraglottal pressure drop Lubker (1973) observed that the velopharyngeal leak might be regarded as another voice maintaining mechanism. He agreed with Rothenberg (op. cit.) that active enlargement is more likely to be responsible for the absorption of the transglottal airflow. Both actions would result in a pressure drop during the closed phase of the voiced plosives. He felt that the velopharyngeal leak may be the result of an active palatal movement rather than a response to an increased oral pressure. When comparing the velopharyngeal leak and the active enlargement of the supraglottal cavity and their relative roles in reducing the rise in supraglottal pressure during the closed phase of the voiced plosives, he seemed to indicate that the former played only a secondary role. He stated that "a velopharyngeal leak represents the least likely single mechanism, while active cavity enlargement appears to be the most likely". Similarly, Nihalani (1975b) suggested that the supraglottal pressure drop during voiced and voiced aspirated plosives in Sindhi might be attributed to the velopharyngeal leak. He found that the non-nasal voiced and voiced aspirated plosives were associated with a slight amount of nasal airflow. In spite of that, it was not clear from his data whether this velopharyngeal leak was caused by a passive response to overpressure in the supraglottal cavity or to active palatal movements as suggested by Lubker (1973).

Perkell (1969) based on a cineradiographic study showed changes in pharynx volume during the closed phase of the voiced and voiceless plosives. It was larger during the voiced than during the voiceless plosives production. He predicted that the reduced oral pressure values obtained in aerodynamic studies for the voiced plosives might be the result of a passive response of the supraglottal cavity to overpressure. This led him to describe the two sets of plosives /ptk/ and /bdg/ as 'tense' and 'lax' respectively.

To sum up, the above studies seem to differ regarding the point of whether there are voice-maintaining mechanisms needed for fully voiced plosives or not. Some of the studies include mechanisms such as enlargement of pharynx and cheek walls and slight escape of air via the velopharyngeal port into the nose. Another question raised by these studies was whether the enlargement of the supraglottal tract was caused by an active and deliberate movement of the muscle forces or by a passive one due to air pressure pushing the pharynx walls. Most of the available evidence seems to suggest that it is more likely to be caused by the former than the latter.

2.3.4 Aerodynamic factors and phonetic environment:

A consistent reference was made by the individual studies to the point that the pressure and flow variations depended not only on whether the plosives belonged to the /ptk/ or the /bdg/ sets but also on some aspects of the context in which they were produced. The phonetic environment and stress were some of the important aspects of this context (Lisker, 1970; Brown , McGlone , Tarlow and Shipp, 1970; Dixit and Brown, 1978 among others). Brown et. al. (1970) suggested that oral air pressure variations reflected the changes of phonetic position of consonants even when the words were in sentences. They found, for example, that the English plosives /t/ and /d/ were consistently produced with greater oral pressure when they occurred in word-initial position than in word-final position.

The results of various studies on this aspect were contradictory. No widespread or commonly accepted explanation can be found. Thus, Black (1950) reported diminishing oral pressure values for the English consonants in initial, medial and final positions in that order. Other studies, however, showed higher peak oral air pressure values with either the word-medial plosives or the word-final plosives than with the word-initial ones. For example, Arkebauer, Hixon and Hardy (1967) presented higher peak oral air pressure values for English plosives in the intervocalic position than in either the post- or the pre-vocalic positions. Similarly, Malècot (1966), on the other hand, found that the word-medial and word-final plosives were produced with higher oral air pressure than the word-initial ones.

Airflow studies showed similar contradictory results. Thus, Gilbert (1973) reported higher oral airflow rate for pre-vocalic than either inter-vocalic or post-vocalic English plosives. Stathopoulos and Weismer (1985) also revealed similar results although they emphasized that this trend interacted with the position of stress. On the other hand, Isshiki and Ringel (1964) found that English consonants in final position were associated with higher airflow rate than those in initial position. Therefore, further studies on other languages are needed.

2.3.5 Aerodynamic factors and the place of articulation:.

Description of the relationship governing the plosive consonants occurring in various articulatory places and the differences in the rate of oral airflow or oral air pressure are not yet clearly understood.

Most of the studies which dealt with this phonetic variable did not report significant differences in oral pressure values. Dixit and Brown (1978) investigated peak oral pressure for Hindi plosives in CVCVC nonsense utterances and found significant differences in peak oral pressure as a function of place of articulation. The place of articulation and the volume of the oral cavity were inversely related. Such a finding has not been supported by other data. Indeed Lisker (1970) and Miller and Daniloff (1977) found nonsystematic patterns for the place of articulation and the peak oral air pressure.

Similar results regarding the relation between oral airflow rate and the place of articulation were reported by Gilbert (1973), Nihalani (1975a) and Dixit and Brown (1986). All observed non-systematic effects of the articulatory place on oral airflow rate in the production of the plosives consonants.

2.3.6 Aerodynamic and articulatory descriptions of emphasis:

To our knowledge there are not many detailed aerodynamic studies on the Arabic emphatics and non-emphatics. Odisho (1973) investigated the pharyngeals and pharyngealized (emphatic) consonants in Baghdadi Arabic. The mingograms were used to illustrate airflow and air pressure values in the emphatics and the non-emphatics and no actual measurements were derived from these traces. Consequently, it was not possible to make general conclusions about them.

However, there have been some physiological studies on emphasis based on direct observation of the articulators using various techniques. Panconcilli-Calzia (1924) and reviewed in Giannini and Pettorino (1982) suggested that the emphatics were accompanied by a pharyngeal constriction caused by contraction of the pharyngeal muscles and a retraction of the hyoid bone as well as a raising of the larynx. He claimed that such an articulation was characterized by a lowering of the epiglottis towards the glottis. Similarly, Marçais (1948) based on X-ray and palatographic data for an Algerian subject agreed with Calzia that the emphatics were distinguished from the non-emphatics by a pharyngeal constriction. He believed that the dorsum of the tongue was lowered and the root of the tongue was retracted so much as to cause a considerable constriction in the pharyngeal cavity. For Calzia it was the retraction of the hyoid bone and the contraction of the pharyngeal muscles that cause the secondary constriction while for Marçias it was the retraction of the tongue root. Marçias also felt that the hyoid bone and the larynx were raised during the emphatic articulation. According to him this pharyngeal constriction was almost complete.

The role of the back of the tongue as the secondary articulator was confirmed by Ali and Daniloff (1972) from high speed lateral cineradiographic data for three speakers of the Baghdadi Arabic. They concluded that the active secondary articulator involved in the production of these sounds was the root of the tongue which as a result of its retraction forms a constriction in the pharynx. They emphasized that the posterior pharyngeal wall as well as the velum were not observed to have any role in the production of such sounds and that the constriction with the velum was only observed for the emphatic sound / k /. They did not confirm the retraction nor the raising of the hyoid bone. No raising of the larynx was observed (op. cit. p100).

Al-ani (1970) showed with X-ray pictures that for the emphatics there was a constriction in the pharynx. A slight retraction of the main articulators was also reported. This last finding was rejected by Laufer and Baer (1988) and Giannini and Pettorino (1982). They believed that the emphatics and nonemphatics had similar main articulation place and that the difference between them was that the emphatics were characterized by this additional secondary articulation. The secondary articulation was less constricted than the simultaneous primary articulation. Concerning this point Laufer and Baer (1988) stated that "In the pair /T t /, for instance, there is one main articulation: alveolar articulation to the degree of 'stop'; the emphatic sound is differentiated from its non-emphatic cognate in that its production involves a secondary articulation in another place and with a wider constriction than for 'a stop'"(p. 183).

A similar view to Al-ani's was reported by Gairdner (1935, p.35) who believed that there was a slight difference in the primary place of articulation between /t, d, / and /T, D, /. This difference was believed to be one of dental versus alveolar places of articulation respectively. Apart from these two references there appears to be no experimental evidence concerning this emphatic/nonemphatic difference in the primary place of articulation.

Giannini and Pettorino (1982) examined one Arabic speaker from Baghdad and presented acoustic and X-ray data. The radiographic data showed that emphatic sounds had a constriction in the pharynx while nonemphatics did not.

Laufer and Baer (1988) who used a fiberscope film recording of ten Arabic and Hebrew speakers concluded that emphasis was realised as a secondary articulation. They believed that the epiglottis, which moved back during emphatic articulation, formed a constriction with the pharygeal wall. They also believed that the tongue root was to some extent involved in this constriction. The answer to the question of whether the epiglottis independently moved backward or was caused to move by the tongue root retraction could not be determined from their data. Contrary to previous claims, they argued that the pharyngeal constriction occurred at the lower part rather than at the upper or middle parts of the pharynx. Ghazeli (1977) based on cineradiographic study of Tunisian Arabic speakers showed that the emphatics were produced with a pharyngeal constriction somewhere between that for the uvular sounds /x ,ɤ/ and the pharyngeal sounds /ħ, ʕ, h /. Delattre (1971), too, observed that during the emphatics there was a constriction in the upper pharynx achieved by movements of the lateral pharyngeal wall.

Elgendy (1992) examined the jaw and lips movements for the pharyngeals and emphatic articulation in the speech of three Egyptian subjects. The measurement systems were a set of transmitter and receiver coils. He assumed that the mandible movement was coordinated with the muscular activities in the lower jaw causing constriction in the pharynx as well as the movement of the tongue. His main finding was that the mandible position during the emphatics was higher than that during the non-emphatics. It manifests less degree of displacement relative to the rest position established by the experimenter. The distance between the upper jaw and the mandible depended on the location of the constriction in the pharyngeal cavity. It was found to be wider as the constriction location became lower in the pharynx wall.

Almost all of these studies referred to the existence of co-articulation of emphatics with adjacent vowels. Some of the above investigators observed that the pharyngeal constriction during the emphatic articulation began within the preceding vowel and persisted during the following vowel Ali and Daniloff (1972); Giannini and Pettorino (1982) Laufer and Baer (1988) Elgendy (1992).

Since most of the above studies were not concerned with the aerodynamic characteristics of the emphatic nonemphatic distinction, it would be of interest to us to examine these variables in the plosives of YSA (See sections 5.1 and 5.2).

PART TWO: EXPERIMENTAL INVESTIGATION

Chapter3: Acoustic analysis of voicing and emphasis in the plosives of Yemeni Spoken Arabic

3.1Voice Onset Time in Yemeni Spoken Arabic

3.1.1 Aims:

The aim of this study is to make some statements about the range of VOT distribution for the plosive consonants in YSA and to investigate some of the claims or hypotheses made about the VOT (See section 2.1.1 for details about them). It will be interesting to see if the VOT acts as a sufficient acoustic correlate of the plosive voicing in YSA.

Another related aim is to see if the vocalic context affects the VOT values of the preceding plosive consonants. Some studies have demonstrated that the vocalic context affects consonant duration and VOT. This experiment will try to find how this effect may be manifestied in YSA. (Further details on the objective of this investigation are mentioned in section 2.1.1).

3.1.2 Method:

3.1.2.1 Sample material :

Eight YSA plosive consonants were chosen for examination , four of them were voiced /b, d, D, g/ and the other four were voiceless /t, T, k, q/. They were put in monosyllabic words (consonant-vowel-consonant) in initial position. The context was represented by six vowels: three short and three corresponding long vowels /i, a, u, ii, aa, uu/ . Table 3.1.1 shows YSA meaningful words containing the target plosives in combination with these different vocalic contexts. The words were manipulated in such a way that they allowed for comparison between the voiced and the voiceless consonants, followed by the same vocalic contexts, with the final consonant for each pair of words was kept similar where possible.

All words were written as one list. They were randomized in six different orders with a dummy word at the beginning and another at the end. They were intended to avoid different prosodies for first and last word in a list. The 52 words including the two dummy ones at the beginning and end of the list were written on a size A4 card to avoid any noise created by the paper rustling. They were randomised in six different orders.

Words	Equivalent Translation	Words	Equivalent Translation	Words	Equivalent Translation
baat	'to pass the night'	biir	'a well of water'	buuq	'horn'
taab	'to repent'	tiin	'figs'	tuuq	'desire'
daab	'creeping animal'	diin	'religion'	duuq	'Duke'
Taar	'to fly'	Tiiq	(imp.) 'tolerate'	Tuur	'a mountain name'
Daar	'harmful'	Diiq	'tiredness'	Duur	'starvation'
kaad	'to be on the point of '	kiif	'comfort'	kuub	'cup ; glass'
gaad	'has led'	giis	(imp.) 'measure'	guud	(imp.) ' lead'
qaad	'has led'	qiis	(imp.) 'measure'	quud	(imp.) ' lead'
bat	'completed'	bir	'wheat'	buq	'mouth'
tab	'destroyed'	tim	(imp.) 'finish'	tub	(imp.) 'repent'
dab	'to walk'	dim	'cat'	dub	'bear'
Tar	'turned out'	Til	(imp.) 'visit'	Tur	(imp.) 'turned out'
Dar	'caused harm'	Dil	shadow'	Dur	'harm'
kad	'hardwork'	kif	(imp.) 'stop'	kum	'sleeve'
gad	'just'	gis	'priest'	gum	(imp.) 'stand up'
qad	'just'	qis	'priest'	qum	(imp.) 'stand up'

Table 3.1.1 Sample material containing the word-initial plosives in the context of long and short vowels in YSA.

3.1.2.2. Recording:

For consistency and for minimizing the effects of suprasegmental variables the words were rehearsed before recording. The subject, who was the writer of this study, recorded the six repetitions for each word with normal tempo. He read them in their citation forms with a falling pitch. They were recorded in sessions of three runs each. Each session was separated from the other by a five minutes interval to allow the subject to keep a constant mode of speech and in order not to be affected by tiredness. Recording took place in a sound treated room in the Department of Linguistics and Phonetics. The material was recorded on a high quality tape recorder (type Sony PCM F1) with a Sony HI FI tape. The microphone used was B & K ½ inch condenser (MIC) with the B & K portable sound level meter. The microphone signal was recorded on track one of the tape, with a laryngograph (Lx) signal recorded on track two. The distance between the microphone and the mouth during this recording was 110 mm/100 mm. The tape recorder setting was as follows: The recorder level volume was 3¾ for (Lx) and 4¾ for (MIC). The peak sound pressure level at the microphone was 101.09 dB as measured by the B & K portable sound level meter. The Lx signal was recorded by means of two electrodes which were strapped on the subject's neck.

Before recording the material a 1 kHz reference tone and a square wave were recorded so that whenever another recording session was needed, this reference tone could be used to match the conditions under which previous recording sessions were done.

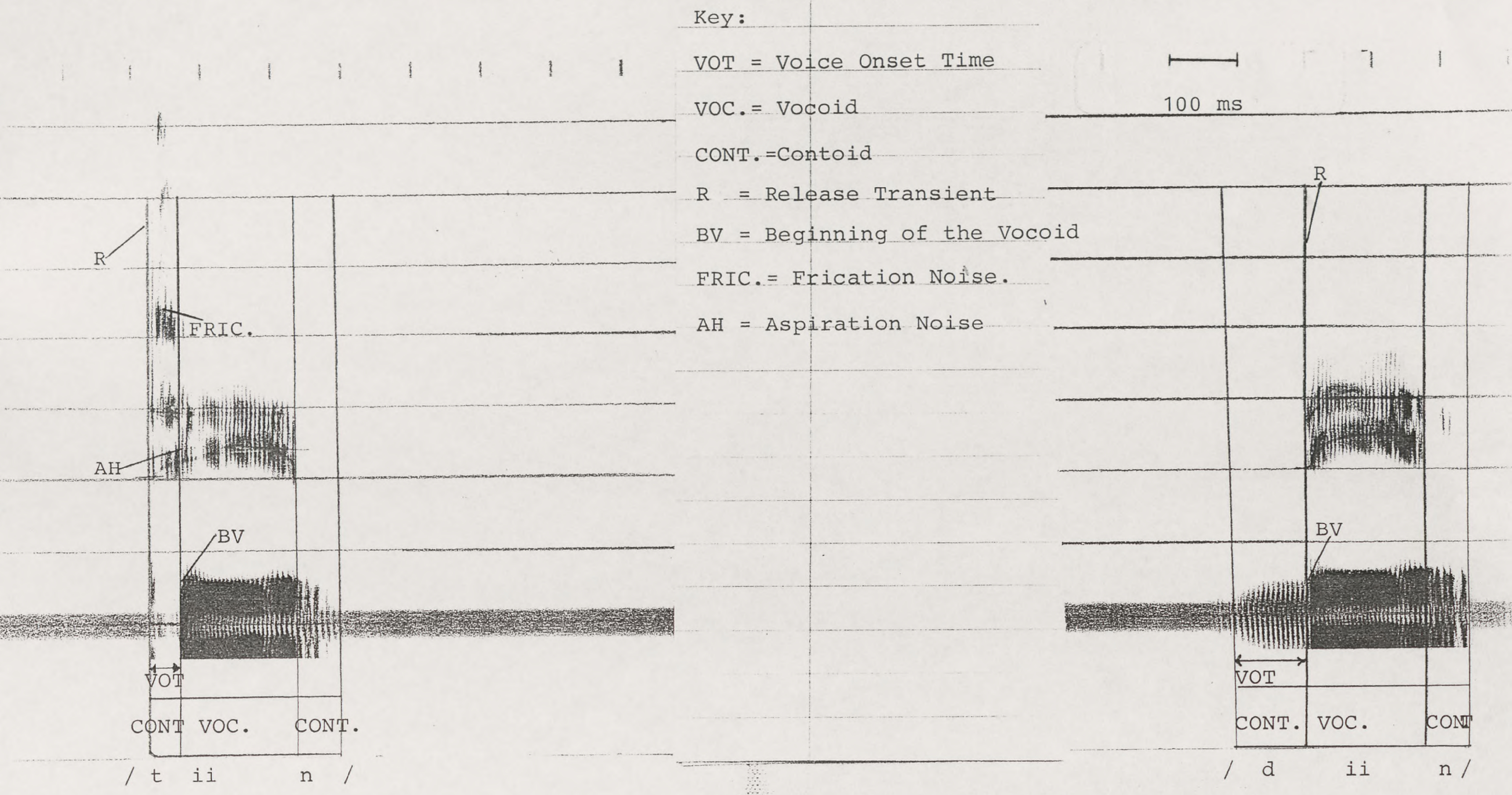
After the recording was done an auditory check was made to ensure that the recording was of high quality and that the words were said as intended. The experimenter played them back on a tape recorder and wrote down what he heard. All words were correctly identified and were good versions of the original ones.

3.1.2.3 Measurements:

Wideband spectrograms were made for all the words. They were made using the 700 spectrograph (Voice Identification Inc). The abssisca of the spectrograms shows the time in (ms), and the ordinate displays frequency in (Hz) with a linear calibration. Each calibration line is 1000 Hz apart from the next one up to 8000 Hz. The 'Hi' setting was used. This boosts high frequencies. (This is done because we perceive high frequencies about 3 to 6 kHz as loud; the spectrogram shows these higher frequencies more clearly, more like the way we perceive them- than a strictly 'Hi-Fi' 'level' setting would do). Figure 3.1.1 presents typical examples of the measurement procedure and of the segmentation criteria. VOT measurement is done in accordance with the usual procedure established by Lisker and Abramson (1964). VOT is defined as the interval between the first clearly identified striation and the plosive burst. In YSA the burst usually comes after the onset of voicing with voiced plosives, while with voiceless ones voicing begins 20-50 ms after the burst.

However, it is not implied by this segmentation that each segment has clear cut boundaries with an exact beginning and end. We are fully aware that perfect correspondence between the portion given for a segment on the spectrogram and how it will actually sound is almost impossible.

Figure 3.1.1 Wideband spectrograms ilustrating the segmentation criteria and the VOT measurement procedure employed in the analysis of voicing in YSA.



Key:

VOT = Voice Onset Time

VOC.= Vocoid

CONT.=Contoid

R = Release Transient

BV = Beginning of the Vocoid

FRIC.= Frication Noise.

AH = Aspiration Noise

Therefore our method is an approximate and convenient way of analysing the VOT on the wideband spectrograms.

3.1.3. Subject background:

The writer of this study served as the sole subject in this experiment. He is a native speaker of Yemeni spoken Arabic. He was brought up and educated in Yemen. His knowledge of that language did not only qualify him to act as a subject for this study, but also enabled him to choose suitable Yemeni currently occurring words. His movements between different Yemeni cities gave him also sufficient awareness of the slight variations in speech styles within the Yemeni population. The words chosen are quite common and widely used by all native speakers of Yemeni spoken Arabic.

3.1.4 Results

Results of measurements of VOT are displayed in Table 3.1.2. They show the range and median VOT values for the different plosive consonants. It can be seen that the voiced sounds occupy the voicing lead whereas the voiceless ones have their VOT values in the voicing lag region. Even those plosives which do not have corresponding cognates occupy their own ranges on the VOT dimension as if they had one. Thus, the /b/ sound occupies the voicing lead region regardless of the vowel that follows. This is in marked contrast with what Yeni-Komshian et. al. (1977) find in their study. They note that some of the occurrences of /b/ have VOT values falling between zero and +20 ms. The voiceless uvular plosive /q/ has short voicing lag. The homorganic pairs /t,d/ , /T,D/ and /k,g/ are also separated into different well defined ranges though not without exception. In particular /T/ and /D/ show some overlap in their VOT occurrences especially when they are followed by the long close vocalic contexts /ii/ and /uu/. The degree of overlap, however, is limited and does not exceed the +20 ms lag. The overlap has also been observed in the production of the Lebanese speakers, but in that study it occurs on a larger scale to the extent that it includes /t,d/ as well.

In our data the VOT value for each sound seems to depend not only on voicing but also on the presence of other features such as emphasis and the amount of frication noise available. The * voiceless plosives /t/ and /k/ are found to have slightly higher VOT than /T/ and /q/. The VOT values for /t/ range from 15 ms to 55 ms, whereas for /T/ it ranges from -30 ms to 25 ms. Similarly, for /k/ the VOT value ranges between zero and 80 ms, while for /q/ it is between 5 and 75 ms. These VOT ranges for /t/ and /k/ appear to go beyond the short lag period determined by Lisker and Abramson

(1964). They are in agreement with results from a study of voicing in spoken Saudi Arabic (Flege and Port, 1981) where the VOT values presented for /t/ and /k/ are fairly consistent with the ones shown by this study.

Plosive	**N**	**Value**	**/aa/**	**/uu/**	**/ii/**	**/a/**	**/u/**	**/i/**	**Overall Range**
/b/	6	range:	-90: -15	-90: -55	-110: -50	-90: -20	-100:--50	-120: -55	-120: -15
		median:	-57	-60	-85	-63	-78	-70	
/t/	6	range:	20: 25	35: 50	40: 55	15: 35	20 : 40	30: 40	15: 55
		median:	20	40	50	20	35	40	
/d/	6	range:	-90: -60	-70: -40	-115: -80	-130: -40	-120: -55	-110: -60	-130: -40
		median:	-80	-55	-100	-75	-60	-77	
/T/	6	range:	5: 20	5: 20	-30: 25	0: 10	0: 15	0: 20	-30: 25
		median:	10	10	10	10	10	10	
/D/	6	range:	-70: -40	-40: 20	-75: 20	-80: 10	-60: 10	-120: -25	-120:20
		median:	-45	-23	-27	-60	38	-65	
/k/	6	range:	31: 45	0: 70	20: 80	35: 50	40: 60	40: 55	0: 80
		median:	40	55	58	40	48	40	
/g/	6	range:	-80: -30	-130:-20	-120: -10	-120: -60	-100: -30	-110: -10	-130: -10
		median:	-55	-55	-70	-80	-48	-50	
/q/	6	range:	10: 20	5: 75	5: 20	10: 20	5: 20	5: 15	5: 75
		median:	18	13	13	20	10	10	
Overall Range									**-130: 80**

Table 3.1.2 Ranges and median VOT values (in ms) for the plosive consonants in combination with the six vocalic contexts as produced by subject AA.

However, /t/ and /k/ for this speaker of YSA are associated with higher VOT values than the Lebanese corresponding plosives. In that study the mean VOT values for each of the two sounds does not exceed 30 ms. It is even lower in some vocalic contexts.* Thus, the voiceless plosives in YSA are separated into two groups on the basis of VOT: voiceless plosives whose VOT values sometimes go beyond the short voicing lag (these include /t/ and /k/ in the context of front close and back vowels), and those whose VOT values are within the short voicing lag region (these include /T/ and /q/).

On the other hand, some voiced sounds are associated with a big voicing lead reaching a maximum value of -130 ms for /d/ and /g/ and -120 ms for /b/ and /D/. Figure 3.1.2 shows

Comparisons	N	t-test value	P-value[1]
tab vs dab	6	8.02	0.0005**
tim vs dim	6	14.26	0.0000**
tub vs dʊb	6	9.74	0.0002**
taab vs daab	6	23.46	0.0000**
tiin vs diin	6	28.00	0.0000**
tuuq vs dʊʊq	6	19.11	0.0000**

Table 3.1.3

Comparisons	N	t-test value	P-value
Tar vs Dar	6	4.3	0.0076**
Til vs Dil	6	6.06	0.0018**
Tur vs Dʊr	6	3.49	0.018*
Taar vs Daar	6	11.09	0.0000**
Tiiq vs Diiq	6	2.53	0.039*
Tuur vs Dʊʊr	6	3.20	0.023*

Table 3.1.4

Comparisons	N	t-test value	P-value
kad vs gad	6	13.69	0.0000**
kif vs gis	6	6.43	0.0014**
kum vs gʊm	6	9.24	0.0002**
kaad vs gaad	6	11.49	0.0001**
kiif vs giis	6	6.85	0.0002**
kuub vs gʊʊd	6	5.95	0.0003**

Table 3.1.5

Tables 3.1.3 - 5 t-tests comparing the VOT values for the homorganic pairs /t,d/ ; /T,D/ and /k,g/ in various vocalic contexts.

---

[1] The two asterisks (**) indicate that the difference between the two sets of VOT values for each pair is **highly significant** at the 1% level, whereas one asterisk (*) indicates that the difference between the two sets of VOT values for each pair is statistically **significant** at the 5% level. The absence of asterisk indicates that the difference is **not significant** at the 5% level.

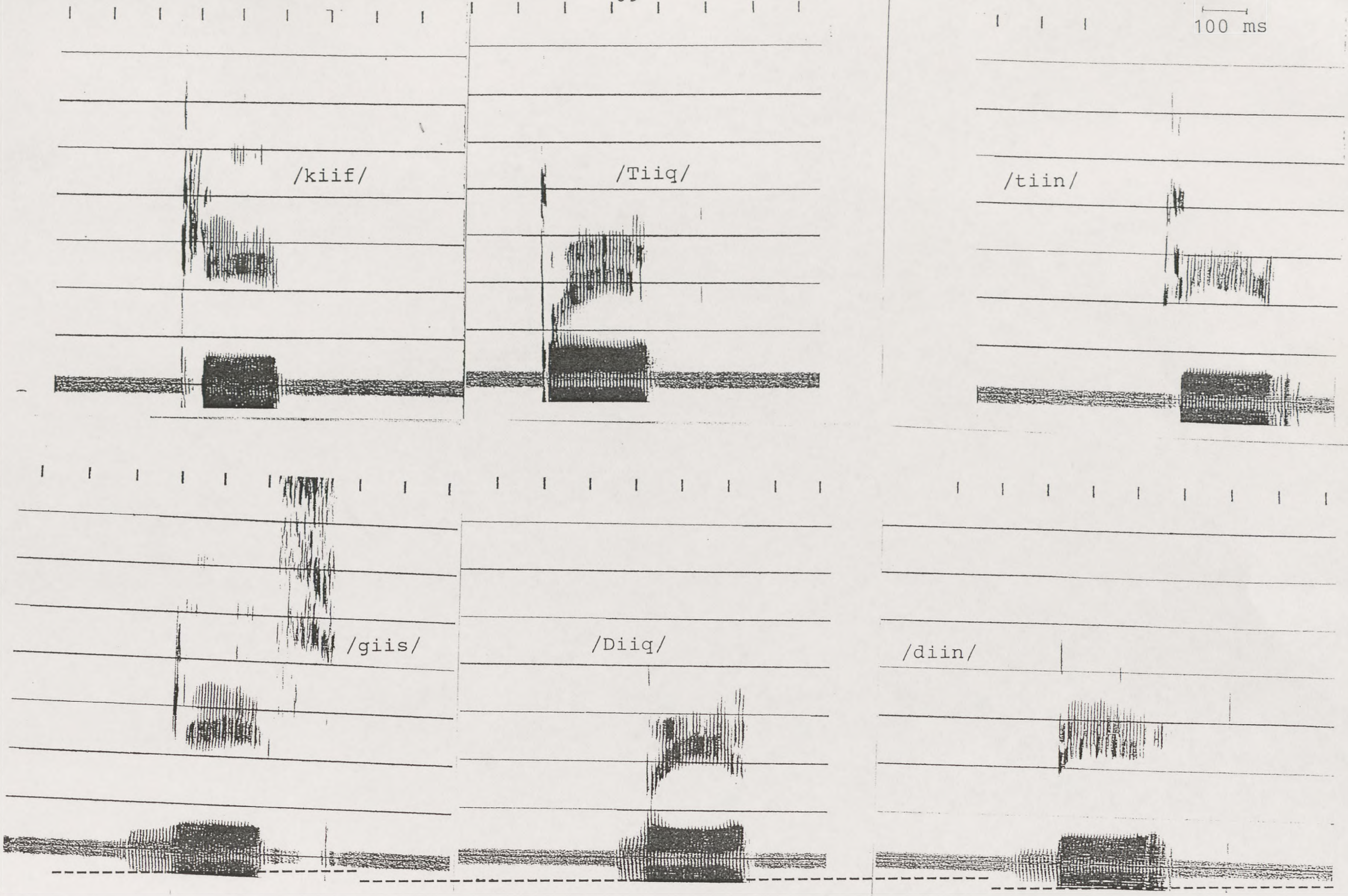


Figure 3.1.2 VOT for three pairs of homorganic voiced and voiceless plosives in citation forms.

spectrograms of VOT values associated with each plosive consonant in combination with the vocalic context /ii/.

To test the significance level of the VOT differences for the homorganic plosives in YSA a series of t-tests (two-tailed) for groups are conducted. The results show the differences between all groups compared to be highly significant at the 1% level for the pairs /t/ vs /d/ and /k/ vs /g/ as it appears from Tables 3.1.3-5. The differences between /T/ and /D/ is also found to be significant at the 5% level.

3.4.1.2 The effect of vocalic contexts on VOT:

There is a slight indication that the VOT value undergoes some variation from one vowel to another for the same sound. It is interesting to see if the vocalic context has a role to play in this variation. Table 3.1.6 gives the mean VOT value, standard deviation and the results of oneway analysis of variance for each consonant across the six vocalic contexts (F-ratio). The mean VOT values differ for the same plosive consonant depending on the vowel that follows. The differences between means for each sound across the vocalic contexts are only significant with two sounds at the 5% level. These are /t/ and /d/ as it appears from the F-ratio for each sound across the six vocalic contexts. Although the other plosive consonants show some differences, these are not statistically significant at the 5% level.

As far as variation caused by the vocalic context is concerned there appears to be no systematic pattern affecting VOT. This finding contradicts the remark made by Yeni-Komshian et. al. (1977) with respect to vocalic influence on the VOT. They observed a tendency for shorter leads and longer lags in the production of plosives before /i/ than before /a/ or /u/ in Lebanese Arabic. This tendency does not seem to be the case in YSA. In fact results for the voiced plosives suggest the opposite. They have a longer voicing lead before /ii/ and /i/ than before /aa/ or /uu/. This generalization is not without exception. /Daar/ , for example, has a higher voicing lead than /Diin/. Similarly, in the context of long vowels /gaad/ has a higher voicing lead than /giis/.

The tentative conclusion which may be drawn from this investigation is that voiced sounds have a longer voicing lead before close front vowels /ii/ and /i/ than before back vowels /uu/ and /u/ in YSA. With regard to the voiceless sounds this relationship disappears completely. Their VOT values do not consistently show longer lags before /i/ and /ii/ than before the open front vowel /aa/, /a/ and the back vowels /uu/ and /u/. Instead, their VOT values are shown to vary from one vocalic context to the other with no consistent pattern affecting their occurrences. Thus, /T/ has longer lags with back vowels /uu, u/ than with front vowels /i, ii, a, aa/ while /t/ has longer lags before close front vowels /i, ii/ than before the open front vowels /a, aa/ and back vowels /u, uu/. This relationship becomes even more complicated with /k/ and /q/. With these sounds there is further nonsystematic variations in VOT values as function of short and long vowels. /k/, for instance, has longer lags before /u/ and /ii/ than before /uu, i, a, aa/.

Plosives	**N**	**Value**	**/aa/**	**/uu/**	**/ii/**	**/a/**	**/u/**	**/i/**	**F-ratio**	**P-value**	**DF**
/b/	6	Mean	-56.7	-65	-80.50	-58.3	-76.67	-80.8	1.38	0.257	5,3
		S.dev.	(29.3)	(12.65)	(20.92)	(30.8)	(16.63)	(27.3)			0
/t/	6	Mean	21.67	41.67	50	22	33.33	37.50	23.17	0.000**	5,3
		S.dev.	(2.58)	(5.16)	(5.48)	(6.78)	(8.16)	(4.18)			0
/d/	6	Mean	-79.17	-54.17	-100.83	-77.5	-74.2	-79.17	3.46	0.014**	5,3
		S.dev.	(10.21)	(11.14)	(12.01)	(29.6)	(25.8)	(19.60)			0
/T/	6	Mean	10.83	11.67	9.17	8.33	10	9.50	0.24	0.943	5,3
		S.dev.	(5.85)	(5.16)	(8.01)	(4.08)	(5.48)	(6.75)			0
/D/	6	Mean	-48.33	-25	-34.2	-50	-31.7	-70	2.21	0.079	5,3
		S.dev.	(11.69)	(20.98)	(32.2)	(30.3)	(27.3)	(32.7)			0
/k/	6	Mean	39.67	46.7	55.83	40	47.50	44.17	1.07	0.399	5,3
		S.dev.	(5.54)	(24.8)	(20.60)	(5.48)	(7.58)	(6.65)			0
/g/	6	Mean	-55.83	-63.3	-71.7	-85	-53.3	-57.5	0.85	0.524	5,3
		S.dev.	(19.60)	(38.3)	(40.7)	(21.68)	(25.6)	(38.2)			0
/q/	6	Mean	16.67	21.7	13.33	17.50	10	10	0.95	0.466	5,3
		S.dev.	(4.08)	(26.4)	(6.06)	(4.18)	(5.48)	(3.16)			0

Table 3.1.6 Mean VOT values in (ms), standard deviation and F-ratio for the plosive consonants in combination with the six vocalic contexts.

As a result of this nonsystematic relationship between the mean VOT values for the plosives and the following vowels, it can hardly be said that the variations of mean VOT is the result of the vocalic effect alone. Although the variation exists, it could be the result of a combination of factors, one of which may be the vocalic context. A similar finding has been reported by Caramazza et.al. (1974) from data obtained for two French dialects. In that study the authors fail to find any systematic variation in VOT for the voiced plosives in the vocalic contexts /i, o, a/. Their conclusion is that "there was no pattern as function of vocalic context for the voiced consonants, VOT values for the three voiced stops in the context of the three vowels were roughly the same".

To sum up this discussion about possible systematic effect of vowel contexts on VOT a reference should be made to vowel height (i.e. close and open vowels). The mean VOT values for //t, T/ and /k/ are higher before the close vowels /i/ and /u/ than before the open vowel /a/. Table 3.1.7 presents the mean values for these consonant before close and open vowels. It shows also the differences between the VOT before close vowels and open ones as well as the percentages of the difference between means.

	N	Mean VOT before close vowels	Mean VOT before open vowels	Difference (ms)	% of difference between means	t-test value	P-value
/t/	6	35.41	22.00	13.41	23.35	-4.17	0.0032**
/T/	6	9.75	8.33	1.42	7.85	-0.49	0.64
/k/	6	45.80	40.00	5.80	6.75	-2.09	0.066

Table 3.1.7 Variations in VOT values for some voiceless plosives as a function of vowel height.

The difference between mean VOT values for /t/ before close vowels and open ones represents about 23.35 %. This means that VOT for /t/ is 23.35% longer before close vowels than before open ones. Similar differences in the VOT values before close and open vowels are found with the sounds /T/ and /k/, though the percentages of the difference between the means are not as high as with /t/.

This finding is in agreement with Klatt's (1975) study. He finds that "the VOT is 15% longer before the 'high' vowels /i, u/ than before /ay, ɛ/ ". Klatt tries to explain his finding in terms of the possible effect the close vowels may impose on the larynx. According to him the close vowels seem to influence the behaviour of the larynx in such a way that "the laryngeal fundamental frequency is higher (House and Fairbanks, 1953) and voicing is less easy to initiate or sustain than in other vowels". However, if this explanation is applied to /q/, it becomes less appropriate because this sound has a higher mean value before the open vowel /a/ than before the close vowels /i/ and /ʊ/. The effect occurs with both short and long vowels. It is the reverse effect on VOT of what we have seen with /t,T/ and /k/. With voiced plosives, there is no systematic VOT variation as function of close and open vowels.

To compare the VOT values for the voiceless plosives /t, T, k/ before close and open vowels t-test for independent samples are used. The results are shown in Table 3.1.7. Only the sound /t/ shows highly significant differences before the two types of vowels at the 5% level. The differences associated with the other sounds are found to be insignificant.

3.1.4.3 The effect of place of articulation on VOT in YSA:

It has been indicated by some researchers that the VOT value increases for voiceless plosives as the consonant place moves further back from the lips towards the velum and uvula (Lisker and Abramson 1964; Yeni-Komshian 1977.; Hardcastle, 1973). In order to investigate the extent to which these statements are applicable to YSA, it may be useful to clarify some points.

In Yemeni Spoken Arabic there is no voiceless bilabial plosive cognate for /b/. Some of the consonants are defined by more than one phonological feature. When such features interact with voicing in the same contrast, they will certainly have their effect on the VOT values of these sounds. There are, for example, the emphatic sounds which have a secondary place of articulation at the pharynx, in addition to their primary one (Laufer and Baer, 1988). These will be discussed later on as a group and contrasted with their emphatic counterparts. Moreover, our data differ from those of previous studies in that they include the uvular place of articulation. Thus, in our treatment of the place effect on VOT such factors should be kept in mind.

First, the voiceless denti-alveolar /t/ in combination with different vowels shows lower mean VOT values than the voiceless velar /k/. This finding gives further support for the finding presented by the above researchers.

In an attempt to explain the mechanisms underlying such systematic variations, Hardcastle (1973) gives two possible reasons. The first is that the back of the tongue moves slower than the tip of the tongue or the lips. This means that VOT during the release of /k/ will be delayed more than that during the release of /t/ or /p/. The second possibility is related to the point of constriction and the amount of supraglottal pressure behind it. It is said that 'the supraglottal cavity' is smaller behind the constriction point at the velum than it is behind that of the alveolar ridge or the lips. The pressure would be greater behind the velum than behind the lips or the alveolar ridge. Accordingly, velars would have longer VOT than labials.

These two explanations should hold true not only for Korean but also for YSA, particularly for /t/ and /k/. For /q/, however, they apply only in part. Table 3.1.8 below indicates that as the place of articulation for the plosives moves back beyond the velum the mean VOT for the voiceless plosives decreases.

Place of articulation	**N**	**/aa/**	**/uu/**	**/ii/**	**/a/**	**/u/**	**/i/**
/t/	6	21.67	41.67	50.00	22.00	33.33	37.50
/T/	6	10.38	11.67	9.17	8.33	10.00	9.50
/k/	6	39.67	46.70	55.83	40.00	47.50	44.17
/q/	6	16.67	21.70	13.33	17.00	10.00	10.00

Table 3.1.8 Relation between the mean VOT values (ms) and the place of articulation.

Thus the VOT values for the voiceless plosives show /q/ to have consistently lower mean values than /k/. The point of constriction is expected to be further back than /k/ and the volume of the 'supraglottal cavity' behind it is even smaller than that for /k/. Therefore, the pressure might be expected to be greater behind /q/ than /k/. Other possible factors affecting these VOT variations are the timing of maximum glottal area relative to the release, the size of maximum glottal area and spread of return of vocal folds to voicing position.

3.1.4.4 Emphasis and VOT:

On the other hand, the effect of place on the VOT values for emphatics can be understood with reference to their points of constriction and how they differ from their nonemphatic counterparts. Having found that the emphatics are to some extent produced differently from their nonemphatic counterparts, it seems reasonable to consider the corresponding differences in VOT separately, rather than to treat the two groups as if they were similar (See section 1.2.2).

Table 3.1.9 below presents the mean VOT values for the two categories. Each voiceless emphatic sound is found to have a lower voicing lag than its nonemphatic counterpart.

	N	/a/	/i/	/u/	/aa/	/ii/	/uu/
/T/	6	8.33	9.50	10.00	10.83	9.17	11.67
/t/	6	22.00	37.50	33.33	21.67	50.00	41.67
T/t ratio	6	0.38	0.25	0.30	0.49	0.18	0.28
/D/	6	-50.00	-69.20	-31.70	48.33	-34.20	-25.00
/d/	6	-77.50	-79.17	-74.20	-79.17	-100.83	-54.17
D/d ratio	6	0.64	0.87	0.42	0.61	0.33	0.46
/q/	6	17.50	10.00	10.00	16.67	13.33	21.70
/k/	6	40.00	44.17	47.50	39.67	55.83	46.70
q/k ratio	6	0.43	0.22	0.21	0.42	0.23	0.46

Table 3.1.9 VOT values (ms) and ratios of the means for homorganic emphatics and non-emphatics in YSA.

Thus, /T/ and /q/ have lower mean VOT values than their emphatic counterparts /t/ and /k/ across all vocalic contexts investigated. Similarly, the voiced emphatic sound /D/ has a shorter voicing lead than /d/. As it appears from their ratios the mean VOT of the nonemphatic one is almost three times as long as the emphatic in the case of the pairs /t, T/ and /k, q/. It is twice as long in the case of the pair /d, D/. It seems likely that the lower VOT values with /T, q/ and the shorter voicing lead with /D/ are reflections of what happens in the pharynx. The constriction in the pharynx may have the effect of reducing the VOT values for all these sounds.

3.1.5 Conclusions:

This study is an investigation of the voicing distinction in YSA. It is found that the VOT, to a large extent, is capable of separating the plosive consonants into different types, though not without exception. The distribution of the VOT values for word-initial plosives has occurred along three relatively specific ranges:

(1) The voicing lead region where voicing precedes the release of the plosive. This applies to all voiced plosive consonants except for some occurrences of /D/.

(2) The short voicing lag range where voicing begins at the release or sometimes after it. This is exemplified by voiceless plosives with VOT values between zero and 20 ms. It applies to /T/ and /q/.

(3) The moderate lag region, where the VOT value sometimes exceeds the 55 (ms) as with /k/ and /t/ but only in the environment of front and back vowels.

To examine the effect of some factors on the VOT variations, the vocalic context, place of articulation and emphaticness are all considered. The results show that there is no systematic effect on VOT asa function of vocalic contexts. The voiceless plosives /t, T, k/, however, exhibit some effect as a function of close and open vowels but the differences are not found to be statistically significant except with /t/.

The place of articulation has also shown some influence on VOT values in different ways and with different plosives. There is an increase in VOT from /t/ to /k/. Then a reduction is noted in the VOT values as the place of articulation goes beyond the velar place of articulation from /k/ to /q/.

There are some tokens with which VOT is not completely successful in separating homorganic pairs particularly in the presence of emphatic voiced and voiceless sounds such as /T/ and /D/. These two sounds are found to overlap in some of their VOT occurrences. This overlap cannot be ignored no matter how small it is. It may be that other features are needed besides VOT to distinguish these sounds. Apart from that VOT proves to be effective in separating almost all plosive voicing opposition types in YSA.

3.1.7 VOT in connected speech:

One of the limitations of the above study is that it focuses only on citation forms which lack the naturalness found in connected speech and removes a lot of its complexities. Consequently, an investigation of the VOT in word-initial plosives in connected speech is conducted to compensate for this limitation. The aim is to see whether the efficiency of the VOT measure found in citation forms will be carried over to connected speech. In view of the fact that connected speech is characterised by redundancy of cues that contribute to its intelligibility, it is interesting to see if the voicing contrast is still maintained and to what extent the values obtained differ from those in citation forms.

The investigation of VOT in connected speech is complicated by the fact that it involves a number of variables which may not be easily controlled. Among these are rhythm, tempo, coarticulation and stress. Thus, Lisker and Abramson in their 1967 study of the contextual factors, find the homorganic voiced plosives to vary in their VOT values depending on whether they are in stressed or unstressed syllables and whether they are in citation or in connected speech. Their conclusion is that "the two categories /ptk/ and /bdg/ are characterized by significantly different distributions of VOT values. This relation holds true quite independently of several contextual factors investigated" (p.24).

Similarly, Baran et. al. (1977) examine VOT in the speech of three adult females under three conditions: adult directed conversation, child directed conversation and recitation. Their system of VOT measurement differs slightly from that of Lisker and Abramson (1967) in that it contains seven descriptive categories in order to account for all possible VOT patterns observed in that study. They, too, conclude that 'the voicing contrast based on VOT is maintained in conversational speech' (p.344).

In our study, words used in citation forms were embedded in connected sentences (See Appendix 2). For ease of segmentation and to help identify these target words they were all followed by fricatives. They were rehearsed before recording. They were read in a casual conversational style. The recording procedure remained the same as that in citation forms. The measurements were also made from wideband spectrograms. The measurement criteria remained the same as in citation forms. All the plosives were in word-initial position.

3.1.7.1 Comparisons of VOT in citation forms and connected speech.

Results of VOT measurement in connected speech are presented in Table 3.1.10. It is evident from these results that the VOT is still an important cue for distinguishing the voiced and voiceless plosives in YSA. The voiceless plosives /t, T/ and /k/ have quite different mean VOT values from their voiced cognates /d, D/ and /g/. The voiceless plosives occupy the voicing lag region, whereas their voiced counterparts occupy the voicing lead region. Most VOT values for /b/ occur between -60 ms and -40 ms. The pair /t, d/ is well separated.

Word-initial plosives	**N**	**Citation Forms**	**Connected Speech**	**Difference**	**Word-initial plosives**	**Citation Forms**	**Connected Speech**	**Difference**
tab	6	22.00	35.00	-13.00	**dab**	-77.50	-78.33	0.83
tim	6	37.50	32.50	5.00	**dim**	-79.17	-75.00	- 4.17
tub	6	33.33	30.83	2.50	**dub**	-74.20	-76.67	2.47
taab	6	21.67	21.67	0.00	**daab**	-79.17	-85.83	6.66
tiin	6	50.00	43.33	6.67	**diin**	-100.83	-61.67	-39.16
tuuq	6	41.67	40.00	1.67	**duuq**	-54.17	-71.67	17.50
Tar	6	8.33	9.17	-0.84	**Dar**	-50.00	-45.00	-5.00
Til	6	9.50	3.33	6.17	**Dil**	-69.20	-71.67	2.47
Tur	6	10.00	13.33	-3.33	**Dur**	-31.70	-74.17	42.47
Taar	6	10.83	5.00	5.83	**Daar**	-48.33	-60.83	12.50
Tiin	6	9.17	10.00	-0.83	**Diiq**	-34.20	-65.80	31.60
Tuur	6	11.67	13.33	-1.66	**Duur**	-25.00	-65.83	40.83
kad	6	40.00	39.17	0.83	**gad**	-85.00	-57.00	-10.00
kif	6	44.17	50.00	-5.83	**gis**	-57.50	-58.33	0.83
kʊm	6	47.50	43.33	4.17	**gum**	-53.30	-75.00	21.70
kaad	6	39.67	34.17	5.50	**gaad**	-55.83	-66.67	10.84
kiif	6	55.83	60.00	-4.17	**giis**	-71.70	-83.33	11.63
kuub	6	46.70	15.67	-4.97	**guud**	-63.30	-85.83	22.53
qad	6	17.50	13.33	4.17	**bat**	-58.30	-68.33	10.03
qis	6	10.00	8.33	1.67	**bir**	-81.30	-39.17	-42.13
qum	6	10.00	12.50	-2.50	**buq**	-76.67	-72.50	-4.17
qaad	6	16.67	14.17	2.50	**baat**	-56.70	-57.50	0.80
qiis	6	13.33	8.33	5.00	**biir**	-82.50	-73.33	-9.17
quud	6	21.70	4.17	17.53	**buuq**	-65.00	-71.67	6.67

Table 3.1.10 Mean VOT values "ms" (six repetitions) in citation forms and connected speech as produced by subject AA.

The majority of the VOT values for /d/ occur within the range -60 ms and -80 ms, while those for /t/ are mostly within 15 and 55 ms. Similar ranges are obtained for /k, g/.

Those plosives which do not have a voiced cognate (e.g. /q/) have some of their VOT values extended to the voicing lead region in connected speech. There is only one occurrence of overlap between /T/ and /D/. This is surprising, since it is unexpectedly less than that in citation forms. One difference between these results and those of English as reported by Lisker and Abramson (1967) and Baran et. al. (1977) is that the voiced plosives in this study, as compared with those of English, are shown to have a big voicing lead, while the voiceless ones are found to have a short voicing lag.

On the other hand, the VOT values show apparent differences in the two conditions: citation forms and connected speech. In 32 comparisons out of the 48 those in citation form (CF) have higher absolute mean VOT values than that of connected speech (CS). The voicing lead for voiced plosives in CS is not as great as in CF.

By contrast, the voiceless plosives are characterized by greater VOT variations in CS than in CF particularly the /t/ and /k/ sounds. An overall t-test for matched pairs comparing the mean VOT values in CF and in CS shows the difference in the two groups of means is statistically significant ($P = 0.029$). This finding is consistent with that of Lisker and Baran et. al. (1977).

Figure 3.1.4 (a) to (e) shows the VOT distributions along the voicing dimension for the different plosive consonants in the two conditions. It is evident from these graphs that the voiceless plosives /t, T, q/ are all associated with higher VOT values in citation forms than in connected speech. Similar remarks can be said about the voiced plosives /b/ and /d/. Both show lower absolute VOT values in CS than in CF. The only exception noted is the voiced plosive consonants /D/ and /g/ for which the situation is reversed.

3.1.7.2. Discussion and Conclusions:

This experiment is concerned with the VOT in connected speech. It shows that VOT is still an important correlate of voicing in YSA. This is consistent in all the vocalic contexts. A few occurrences of overlap for /T/ vs /D/ contrast are, however, found. Although the general usefulness of the voice onset time measure was established, it does not invariably separate voiced and voiceless plosives. Instances of failure are also found by Lisker and Abramson (1964) in the case of two phonetically different categories of voiced plosives in both Hindi and Marathi which have similar

Figure 3.1.4a Percent /d/ and /t/ production occurrences as a function of voicing (VOT) in citation forms (CF) and connected speech (CS) for one subject, AA (N= 36).

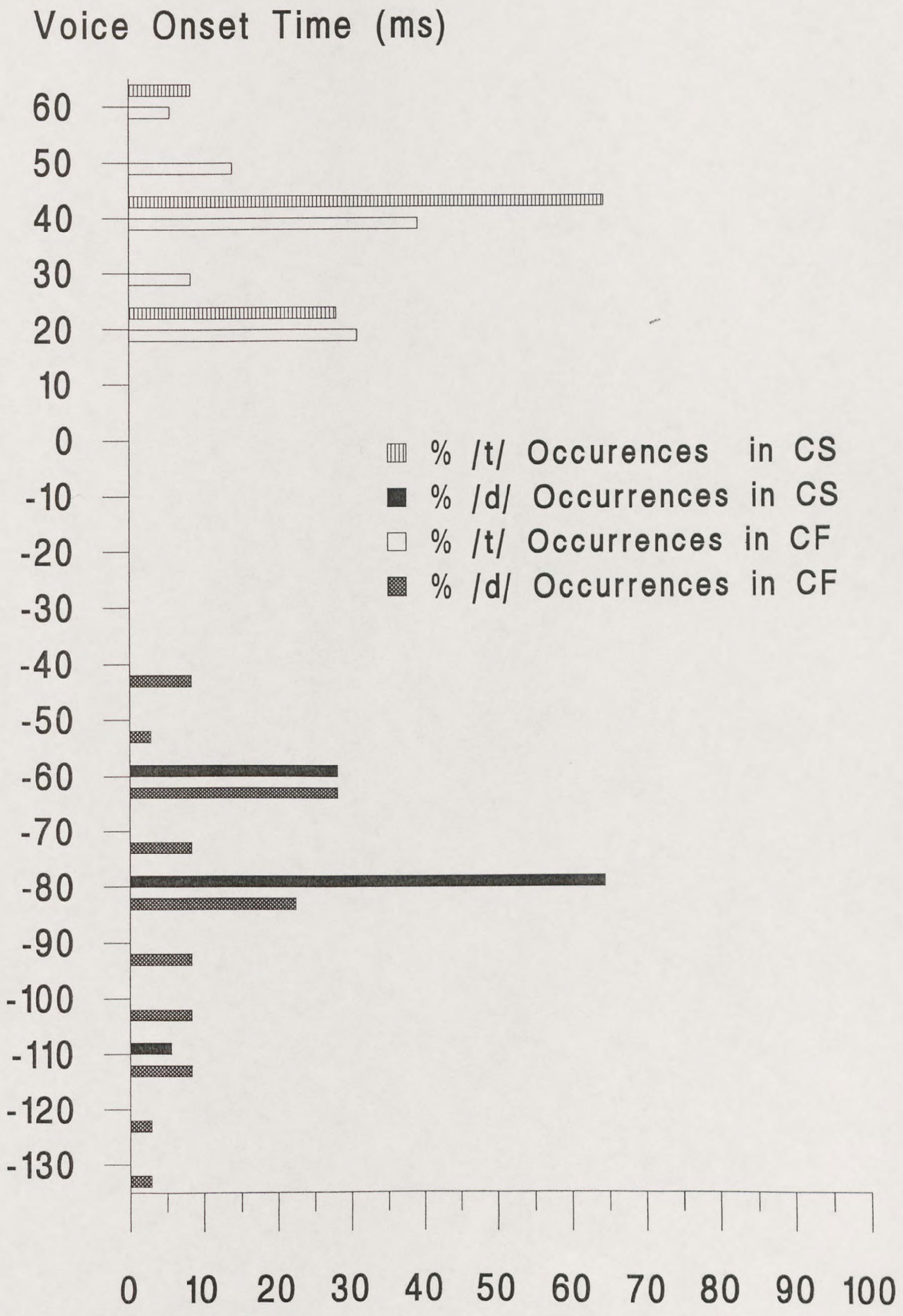


Figure 3.1.4b Percent /D/ and /T/ production occurrences as a function of voicing (VOT) in citation forms and connected speech for one subect AA (N= 36).

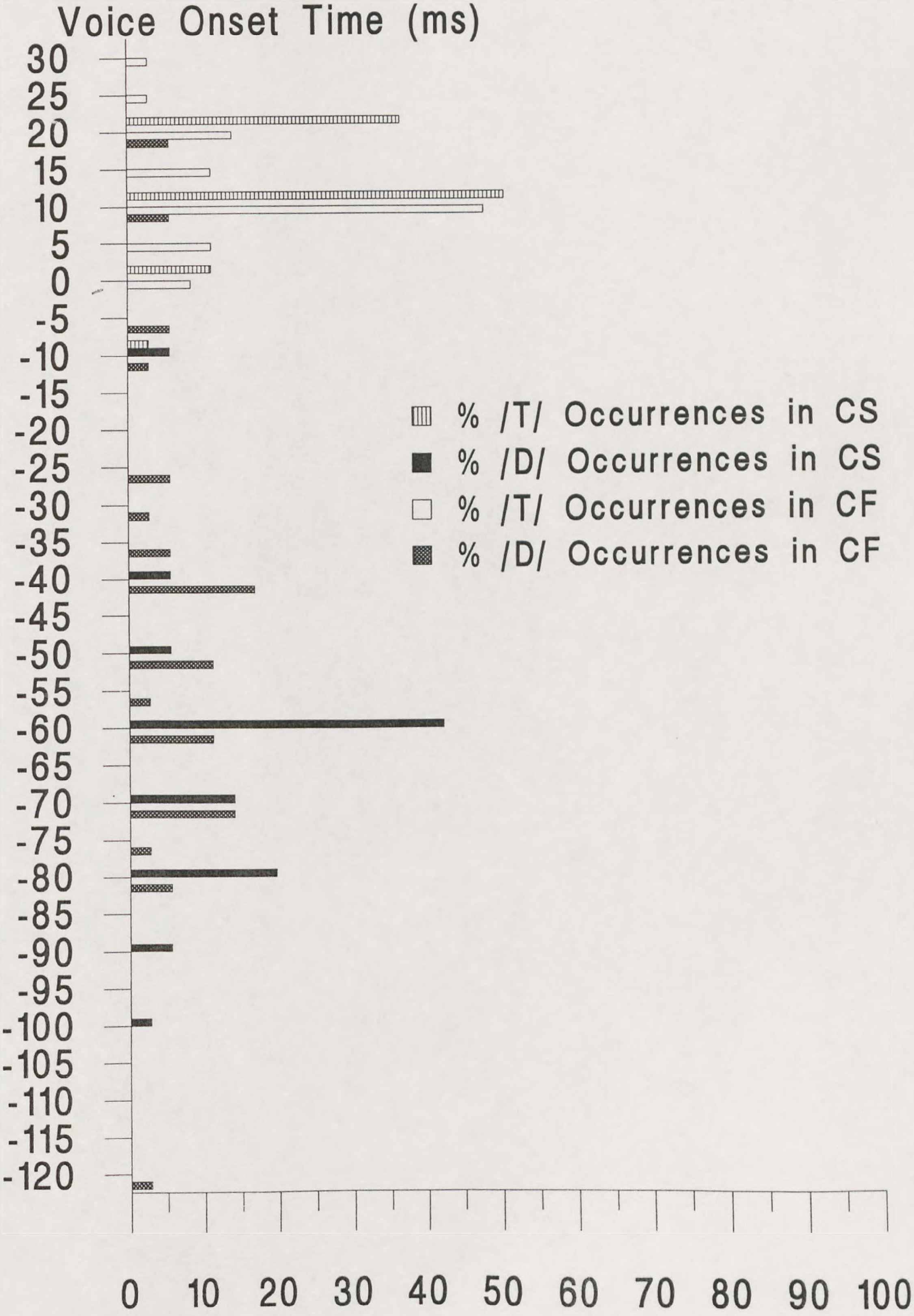


Figure 3.1.4c Percent of /g/ and /k/ production occurrences as a function of voicing (VOT) in citation forms and connected speech for one subject, AA (N= 36).

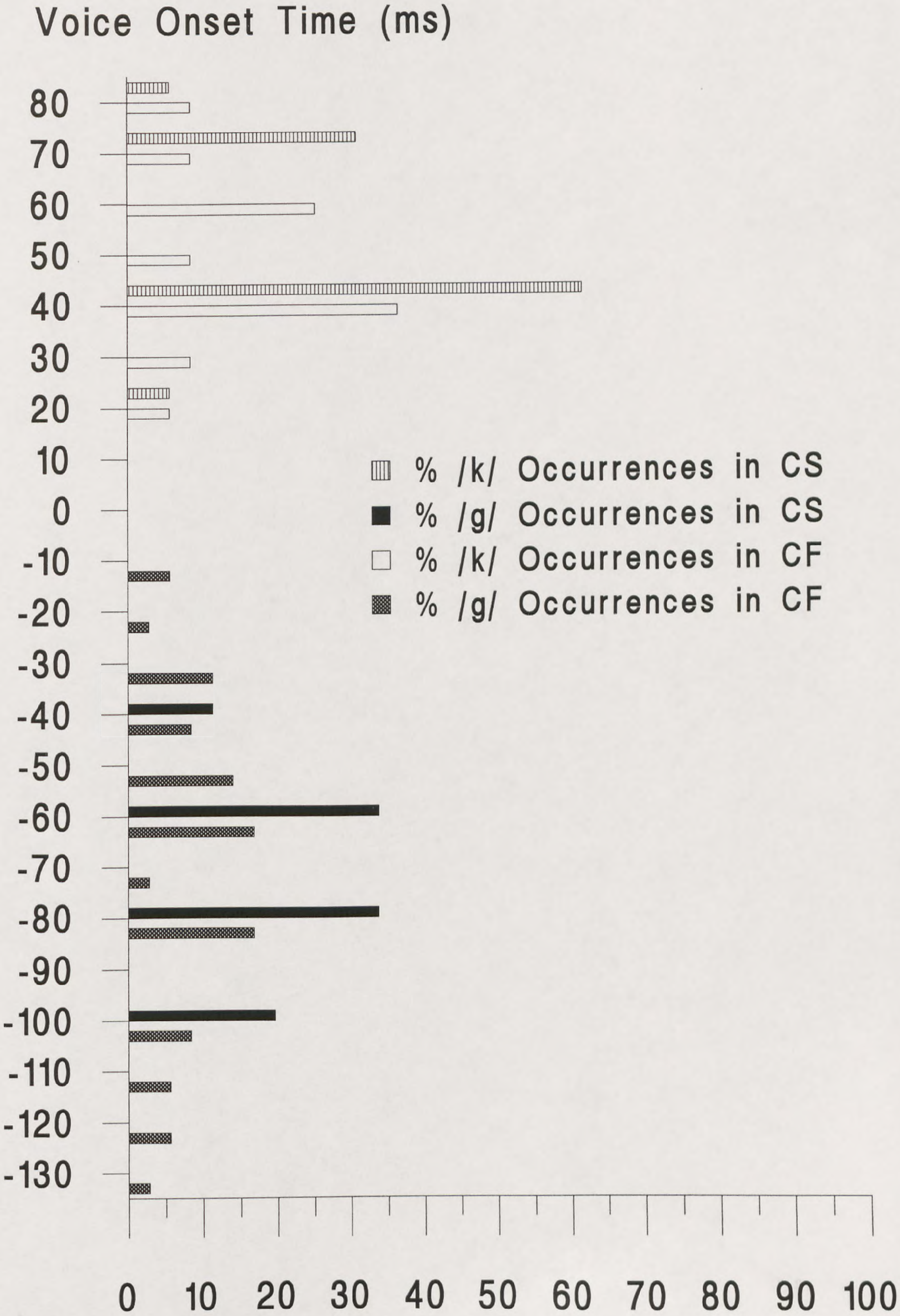


Figure 3.1.4d Percent /b/ production occurrences in citation forms and connected speech for one subject, AA (N= 36).

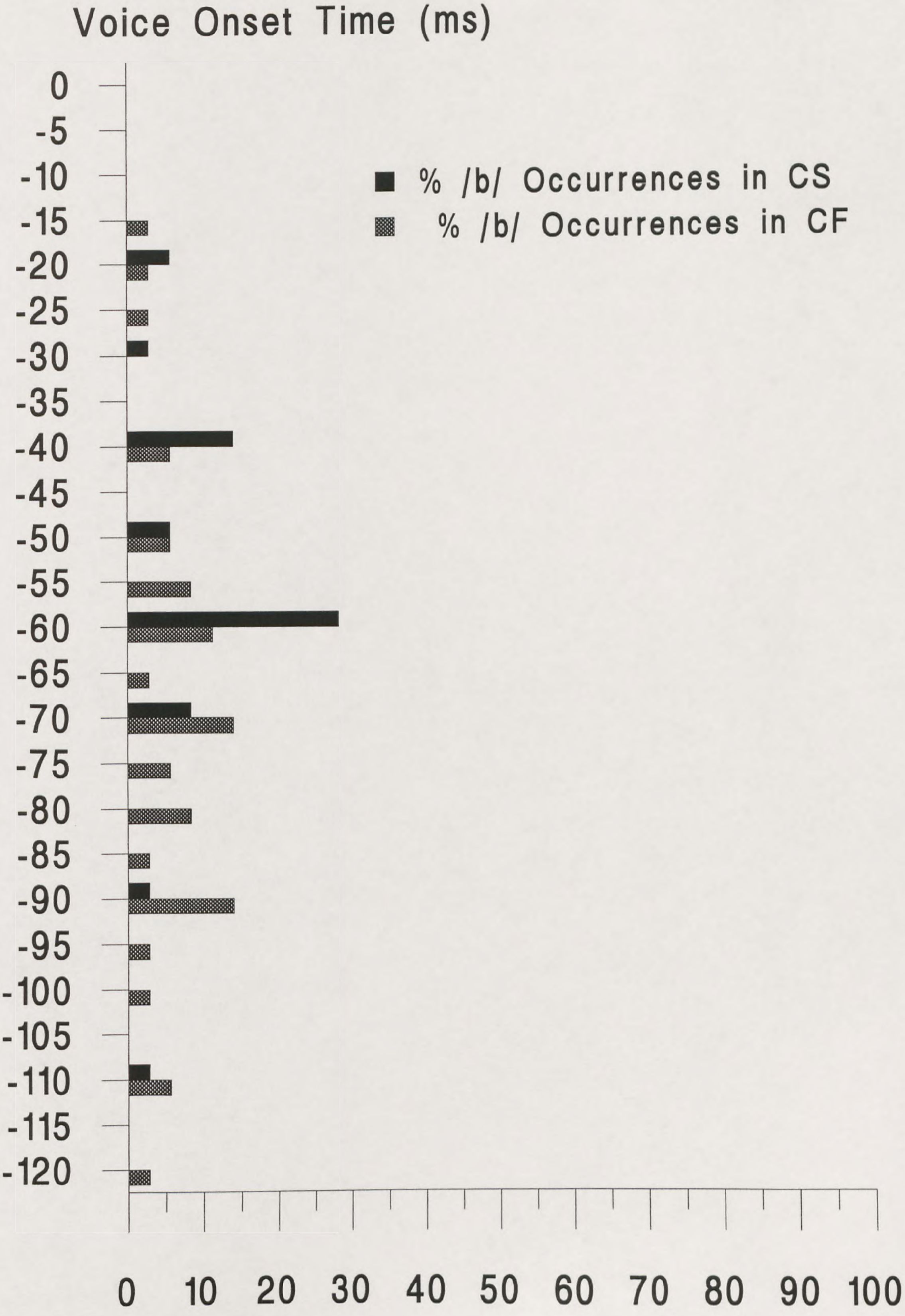
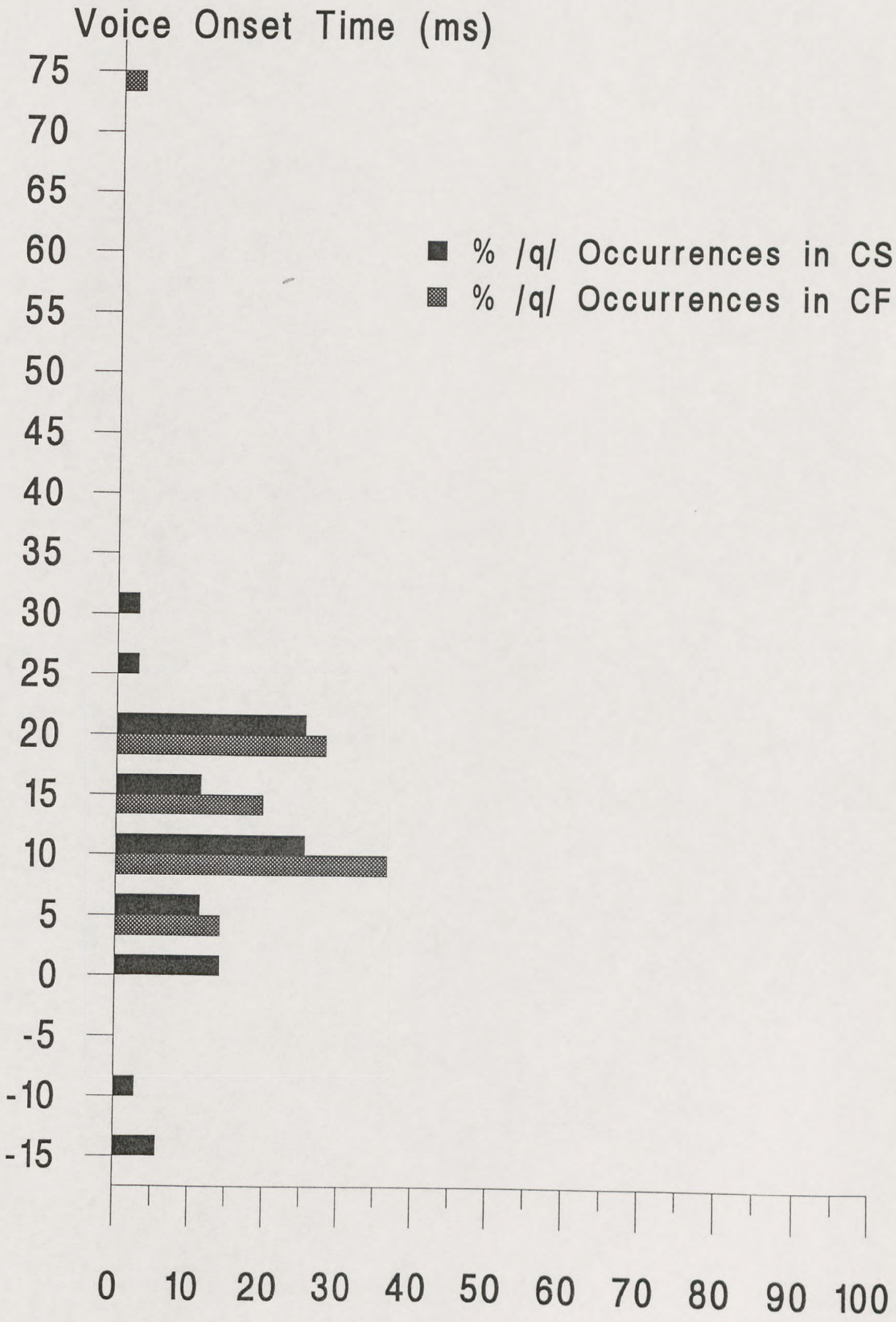


Figure 3.1.4e Percent /q/ production occurrences in citation forms and connected speech for one subject, AA (N= 36).



distributions of voice onset time, but differ from YSA by the presence of aspiration superimposed on voicing in the case of one of the plosive categories.

The absence of the burst in some of the occurrences for the voiced plosives in connected speech is another problem since its presence is a requirement for the measuremet of the VOT. Most word-initial plosives are associated with lower absolute VOT values in connected speech than in citation forms. This may have been caused by the reduction imposed by the rules governing speech in context. Connected speech is controlled by "linguistic and pragmatic constraints which operate to increase redundancy. This condition allows for a greater degree of articulatory imprecision" Baran et. al. (1977). If that is the case then it is likely that this 'articulatory imprecision' , which appears to be a consequence of coarticulation and reduction of sounds and perhaps words may account for this VOT variation in both contexts. Although some of these 'constraints' or contextual factors have been examined there are still others which need to be investigated before any firm conclusion can be drawn. Some of these are syllable rate of an utterance, interspeaker variations, stress and gemination among others. One way such contextual factors may be examined is by conducting a series of cross-language studies on VOT in connected speech each dealing with a certain variable at a time. It has to be admitted that segmentation and measurement procedures in connected speech are relatively less accurate than those in citation forms. Nevertheless a study of these contextual factors in connected speech seems to be worthwhile as it complements the study of citation forms. In our view these kinds of studies could be very useful if they uncover some of the redundant acoustic features which are not necessary for the distinction between homorganic plosives in citation forms, but may be indispensible in connected speech.

3.2 Acoustic study of formant transitions in the emphatic and non-emphatic distinction in YSA

3.2.1 Aims

This experiment will look at the acoustic properties and trends associated with the formant patterns for the emphatic and non-emphatic consonants of the YSA in combination with long vowels /aa, ii, uu/. In traditional phonetic description the vowel context /uu/ and the back variant of /aa/ following emphatics are described as back vowels.

The main objective of this experiment is to explore how the F1, F2 and F3 values at the release of a plosive vary with respect to the associated vowel . In particular, we want to consider if the consonant loci are distinctive for the two categories at different places of articulation and to see if the emphatics and non-emphatics display different formant loci as a result of the secondary articulation associated with this group of sounds (See section 2.1.2).

In addition, the transitional characteristics associated with each category of the contrast will be examined. It is assumed that the transitional change will vary for different consonant-vowel combinations and will consequently display characteristically distinct transition information. If this assumption was proved to be valid and the extent and direction of change between the formant loci and the vowel formant target position were significant, then it might be that such differences could be correlated with the linguistic contrast.

Another related aim is the possibility of inferring the coarticulatory features from the F-pattern analysis. This may enable us to decide whether or not there are consistent rules that explain the interaction of the vowel and the plosive consonant gestures in CVC words. Specifically, we will attempt to describe the patterns that may reflect the coarticulatory effect of the consonant on the vowel and vice versa in the above contexts. If such coarticulation is observed, then it may be relevant to discover the limit of its spread in the word. The vowel target formant values were estimated in the environment of both emphatics and non-emphatics. It is likely that the changes in the formant values for the same vowel in both consonant environments (i.e. the emphatics and the non-emphatics) will help us discover the extent of the effect of emphasis on neighbouring vowels.

3.2.2 Material

The material comprised monosyllabic words. It contained a set of meaningful YSA words initiated with either emphatic or non-emphatic consonants and combined with the three vocalic contexts /aa/, /ii/ and /uu/ in all possible ways. There were 18 different words. The three vocalic contexts were chosen because they represented three important points on the vowel triangle characterising Arabic. All vowels were phonemically long and their length was symbolised by doubling the vowel concerned. Table 3.2.1 shows the sample material used in this study.

Emphatic Consonants	Non-emphatic Consonants
Taab	taab
Daar	daar
qaad	kaad
Tiin	tiin
Diin	diin
qiis	kiis
Tuub	tuub
Duur	duur
quut	kuut

Table 3.2.1 Minimal pair words of emphatic and non-emphatic consonants in the three vocalic contexts.

All words were recorded[2] in six randomised orders and were embedded in an identical sentence frame, (/ ʔiqraʔaa--------sariiʕan/ 'read (you two)..........soon'), with similar stress and pitch patterns. Two native speakers of the same general dialect participated in this recording. An auditory check was done by playing back this recorded material to other YSA subjects. Almost all words were heard accurately. Only those words which were correctly identified by the listeners were selected for acoustic analysis. Words in which there was some doubt about their auditory identity were excluded from our analyses. Wideband spectrograms were produced of all the correctly identified words.

---

[2] The recording technique is the same as that of the VOT study. Consequently, no details will be given here. Instead the reader is referred to section 3.1.2.2.

3.2.3 Measurement Procedure

The focus was on the frequencies of the first three formants. All words containing the emphatic and non-emphatic consonants within the frame sentence were identified. Guidelines were developed for measuring the formant transitions. A formant transition was regarded as that portion of a formant beginning at the consonant release and ending at the first point of the steady state portion of the vocoid (see Figure 3.2.1).

Two measurements were made: one for formant locus and the other for obtaining the formant target position for the vowel. The term locus was used in the same sense as that of Delattre et. al. (1955). As is already known, although in some cases the onset of formant transitions originated visibly from the release transient, in some instances they were delayed for about 20 to 50 ms depending on the consonant voicing type or place and whether the sound segment is emphatic or nonemphatic. In many cases the formant transitions were visible and could be traced all the way back to the release or very near to it. This applied to voiced emphatics and non-emphatics. For voiceless emphatics and non-emphatics it was necessary to extend the formant transition back from the vocoid onset to the release. It was extended in the same transition direction as at vocoid onset. The onset of the vocoid was identified by the presence of regular striations on spectrograms and the increased darkening of the frequency bands through out the whole vocoid as compared with that of the preceding contoid. All formants loci were measured at the origination point of the transition at the consonant release.

The second formant target was defined as that point of maximum or minimum frequency reached. Once that point was identified a line was drawn at the middle of that formant and continued all the way through till the end of the vocoid. We are aware of the fact that in actual speech there is hardly a fixed formant target or steady state position because of the continuous variations in the formant frequencies. In all cases examined there was always a point at which the second formant frequency gradient with respect to time was zero.

On the basis of the above two measurements, the extent and direction of the formant transitions from the formants' loci to their targets in the following vowel were derived as a difference in frequency between the two points.

Key
R = Release burst.
VOC. = Vocoid Onset.
TG = Point of formant targets.

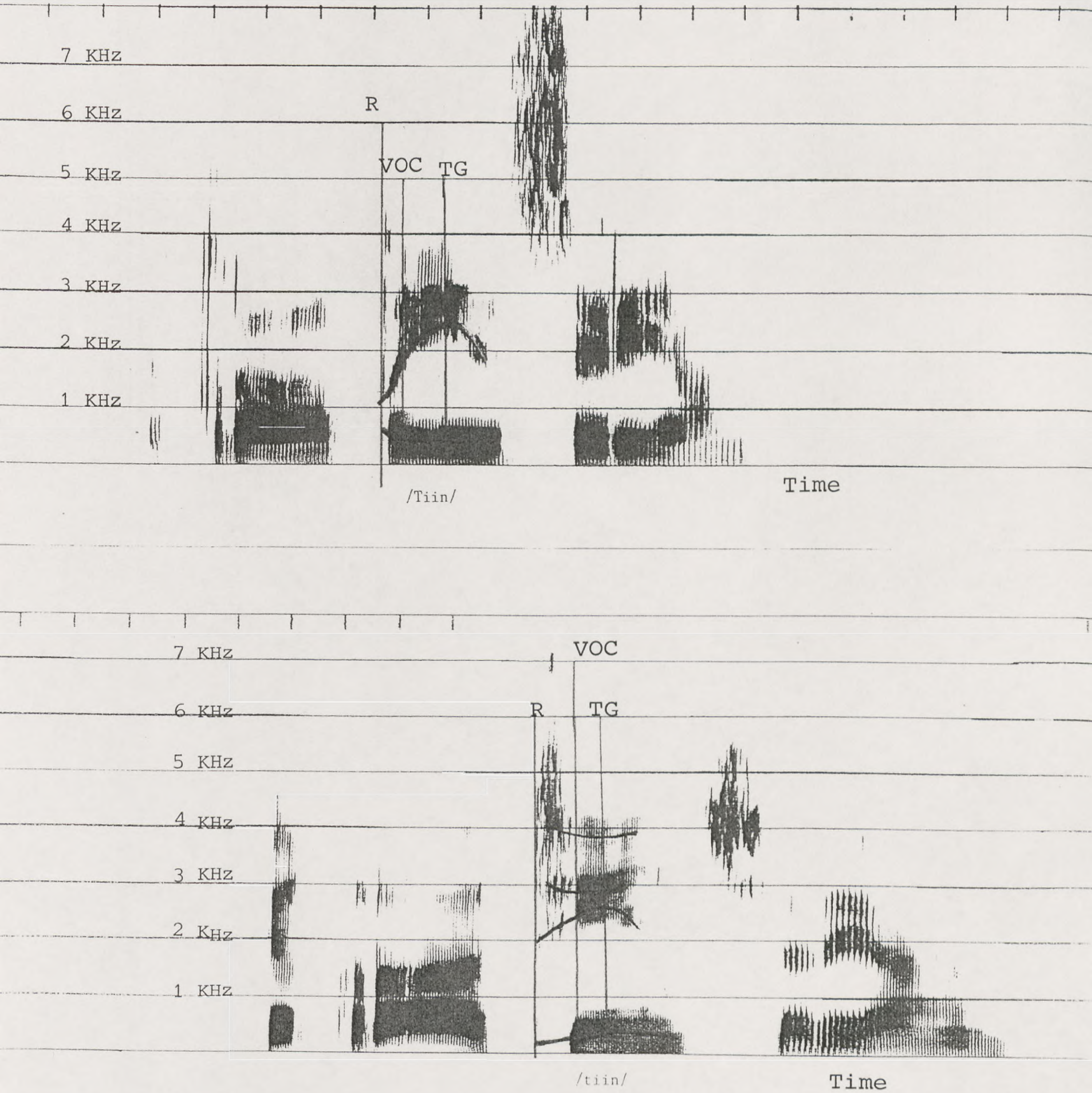


Figure 3.2.1 Wideband spectrograms of the words /Tiin/ and /tiin/ in the frame sentence /?iqra?aa -----sariiɛan/. They show the measurement procedures.

An example of the method used for obtaining the two types of measurement can be seen in Figure 3.2.1. Each of the six repetitions was measured three times to assess reliability and the mean of these three measurements was regarded as the value to be used in our analyses. The three measurements agreed to within ± 50 Hz. This value, therefore, would be our estimation of the accuracy of measurement in this study.

The measurements of the F3 at the target and the locus position could not be obtained for consonants in the environment of back vowels. In such context this formant did not show clearly.

3.2.4. **Results**

Table 3.2.2 presents the F1, F2 and F3 loci for the emphatic and non-emphatic consonants. For subject AA the F1 loci for the emphatic category /T/ range between 500 and 600 Hz while for /t/ they vary between 320 and 360 Hz. Similar ranges for the F1 loci are obtained for subject WS. The /q/ and /k/ contrast shows similar ranges of the F1 loci as that displayed by /T/ and /t/. Thus, the values for /q/ were between 530 and 610 Hz while for /k/ they range between 280 and 325 Hz. Although the pattern is the same for /D/ and /d/ they are shown to have lower F1 loci than in the above pairs. F1 locus is between 205-260 Hz for /d/ and 310-520 Hz for /D/.

One observation which may be made about these results is that, although these ranges are not large, they indicate that the F1 locus is not fully fixed across different vocalic contexts. The F1 locus in some cases is slightly higher in the /ii/ environment than in the other ones. If the F1 loci for the emphatic and non-emphatic categories are considered as groups and in the same vowel context, one can see that they occupy two clearly distinct frequency regions. The former group of consonants occurs at higher frequency level (above 500 Hz) than the other group which shows an average F1 locus of 290 Hz. The exception to this generalization is /Duur/ for subject AA and /D/ in the contexts of all three vowels for subject WS.

The F1 target positions, Table 3.2.3, for the vowels are associated with similar differences in the environments of the emphatics and non-emphatics. They show higher values in the emphatic environment than in the non-emphatic environment. The extent of the separation is closely linked to the vowel quality. For example, the F1 target positions are 690 Hz, 470 Hz and 395 Hz for /aa/, /uu/ and /ii/ respectively in the /q/ enviroment. In the /k/ environment the target positions are 640 Hz, 420 Hz and 355 Hz for the same vowels Hz.

		Subject AA				Subject WS		
	N	/aa/	/ii/	/uu/	Formant loci	/aa/	/ii/	/uu/
/T/	6	600	500	590	**F1**	560	565	560
	6	1170	1075	1130	**F2**	1150	1240	990
	6	2850	2985	?	**F3**	3070	3120	?
/t/	6	360	265	320	**F1**	415	305	300
	6	1890	1950	1890	**F2**	1785	1885	1825
	6	2850	3025	?	**F3**	3000	2940	?
/D/	6	520	520	310	**F1**	470	410	495
	6	975	1000	950	**F2**	960	1015	800
	6	2790	2945	?	**F3**	3090	3035	?
/d/	6	255	205	260	**F1**	220	230	210
	6	1825	1885	1845	**F2**	1800	1965	1735
	6	2740	2685	?	**F3**	2940	2870	?
/q/	6	570	610	530	**F1**	505	575	520
	6	1185	1200	1000	**F2**	1140	1270	700
	6	2745	3040	?	**F3**	3020	3070	?
/k/	6	315	280	325	**F1**	385	280	310
	6	2320	2300	1080	**F2**	2100	2200	1000
	6	2885	3125	?	**F3**	2725	2920	?

Table 3.2.2 Mean frequency values (Hz) of six repetitions for F1, F2 and F3 loci for the emphatic and non-emphatic consonants in YSA.

		Subject AA				Subject WS		
	N	/aa/	/ii/	/uu/	Formant Target Position	/aa/	/ii/	/uu/
/T/	6	700	455	440	**F1**	670	425	465
	6	1275	2445	875	**F2**	1250	2400	860
	6	2650	2780	?	**F3**	3000	3000	?
/t/	6	690	370	405	**F1**	685	405	450
	6	1635	2460	910	**F2**	1430	2570	900
	6	2550	2980	?	**F3**	2740	2950	?
/D/	6	680	455	400	**F1**	670	450	490
	6	1300	2450	810	**F2**	1185	2420	890
	6	2550	2870	?	**F3**	3070	3000	?
/d/	6	630	360	420	**F1**	655	405	455
	6	1585	2475	975	**F2**	1510	2520	1000
	6	2500	2930	?	**F3**	2900	2930	?
/q/	6	690	395	470	**F1**	660	425	430
	6	1375	2450	840	**F2**	1310	2360	820
	6	2590	2965	?	**F3**	2885	2900	?
/k/	6	640	355	420	**F1**	670	370	410
	6	1775	2635	880	**F2**	1640	2390	875
	6	2610	3100	?	**F3**	2490	2980	?

Table 3 2.3 Mean frequency values (Hz) of F1, F2 and F3 at the vowel target positions.

The strongest acoustic variable operating on the emphatic/non-emphatic contrast is the second formant locus frequency. The F2 loci for the two categories show the ranges 1075 and 1170 Hz for /T/ and 1890 and 1950 Hz for /t/ for AA. This noticeable gap separating the two sets of F2 loci seems to be a common attribute of all the emphatic and non-emphatic consonants in this study. F2 locus values are more or less similar for each of the different denti-alveolar plosives (/T, t, d/) across the different vocalic contexts with the exception of /D/. A oneway analysis of variance shows the variations in the F2 loci across the three vocalic environments to be insignificant at the 5% level of confidence for /T, t/ and /d/. The variations are significant for /D/, /q/ and /k/ as can be seen from Table 3.2.4 below.

	N	F-value	P-value
/T/	6	3.14	0.073
/t/	6	2.93	0.085
/D/	6	4.09	0.038 *
/d/	6	2.30	0.135
/q/	6	100.16	0.000 **
/k/	6	1192.50	0.000 **

Table 3.2.4 Oneway analysis of variance of the F2 loci across the three vocalic environments.

If these loci values are used as a way of characterising the emphatic and non-emphatic categories, then the two groups were clearly seen to occur at completely different regions. The differences were all statistically highly significant ($p \leq 0.01$) as is seen in Table 3.2.5 below.

Although this statement applies to all emphatic and non-emphatic contrasts including /q/ vs /k/ in the environment of /ii/ and /aa/ the extent of the difference of this last contrast appears to be not as great for /quut/ vs /kuut/ as for the others. This may be explained by the fact that /k/ occurs with varied F2 loci before front and back vowels. The lowering of the F2 loci for /k/ before back vowels to an average value of about 1080 Hz seems to neutralise the contrast found between them in the context of front vowels.

Emp.	vs	Nonemp.	N	T-value	P-value
Taab	vs	taab	6	-21.55	0.0000**
Daar	vs	daar	6	-41.64	0.0000**
qaad	vs	kaad	6	-31.19	0.0000**
Tiin	vs	tiin	6	-24.37	0.0000**
Diin	vs	diin	6	-26.99	0.0000**
qiis	vs	kiis	6	-132.63	0.0000**
Tuub	vs	tuub	6	-18.39	0.0000**
Duur	vs	duur	6	-44.43	0.0000**
quut	vs	kuut	6	-4.74	0.0052**

Table 3.2.5 t-test results of the F2 loci for the emphatic and non-emphatic consonants.

As far as the place of articulation is concerned, variations within the members of the same category are quite common. Thus, the /t/ and /k/ consonants are differentiated by the fact that the F2 loci associated with /k/ is higher than that of /t/ or /d/ in both the /ii/ and the /aa/ environments. Parallel differences are repeatedly shown in the emphatic class where /q/ and /T/ are associated with higher F2 loci than that of /D/ before the front vowels.

Figure 3.2.2 shows spectrogram illustrations of these F2 loci differences in the two categories. The spectrograms of /T/ and /t/, /q/ and /k/ are characterized by the absence of signal before the release phase. During the release, the non-emphatic contrasts /t/ and /k/ are followed by a slight portion of frication noise when occuring before the vocalic context /ii/.

In addition, the F2 target values are to a great extent distinct for the emphatic and non-emphatic categories. In Figure 3.2.4 below, which is constructed using F2 values from Tables 3.2.2 and 3.2.3, the measured F2 steady state frequencies display the opposite effect to that of F1. In all vocalic contexts, the F2 steady state frequencies in the emphatic environments undergo some lowering relative to those in the proximity of the non-emphatics. This behaviour is more apparent with /aa/.

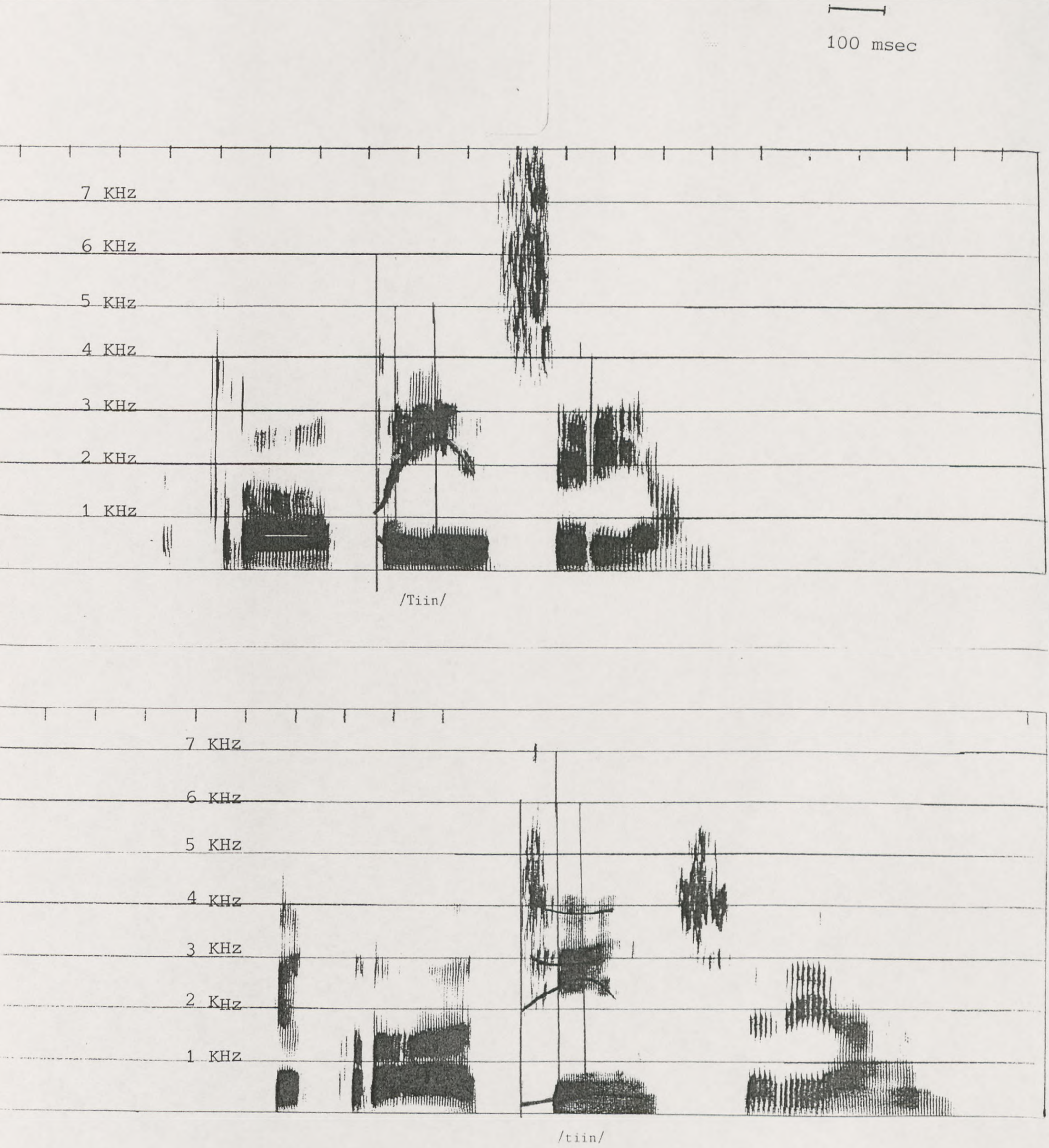
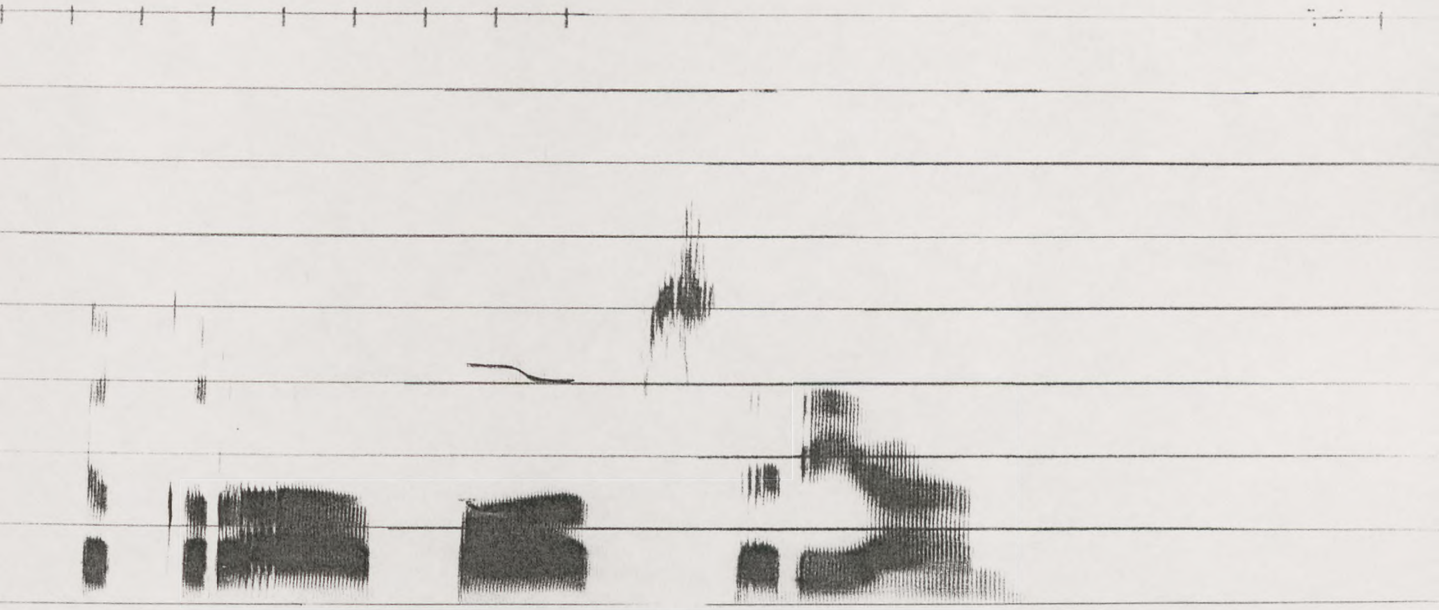
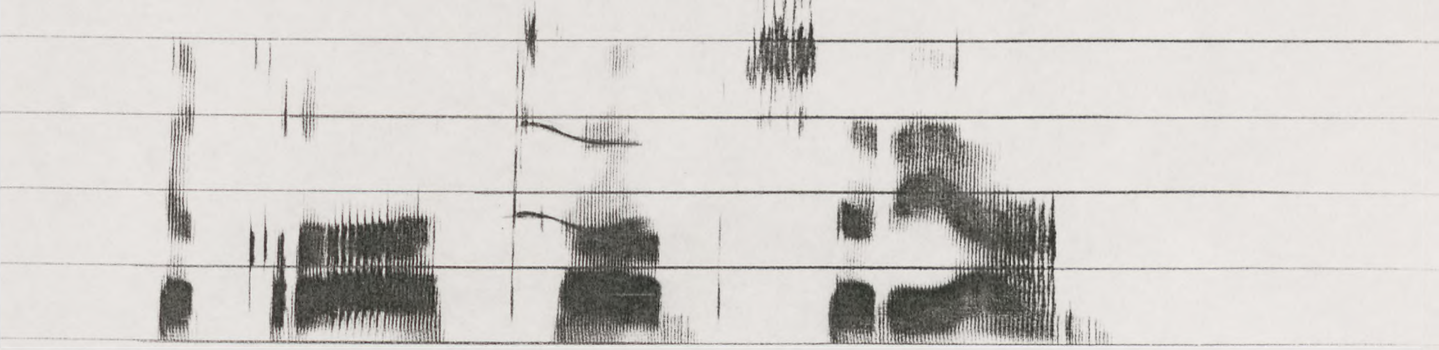


Figure 3.2.2 Wideband spectrograms for word-initial emphatics and non-emphatics uttered by informant WS. The words illustrate the differences in formants frequencies at the consonantal loci and the formant targets for the vowel. All words are uttered in the frame sentence /?iqra?aa ------sarii an/.



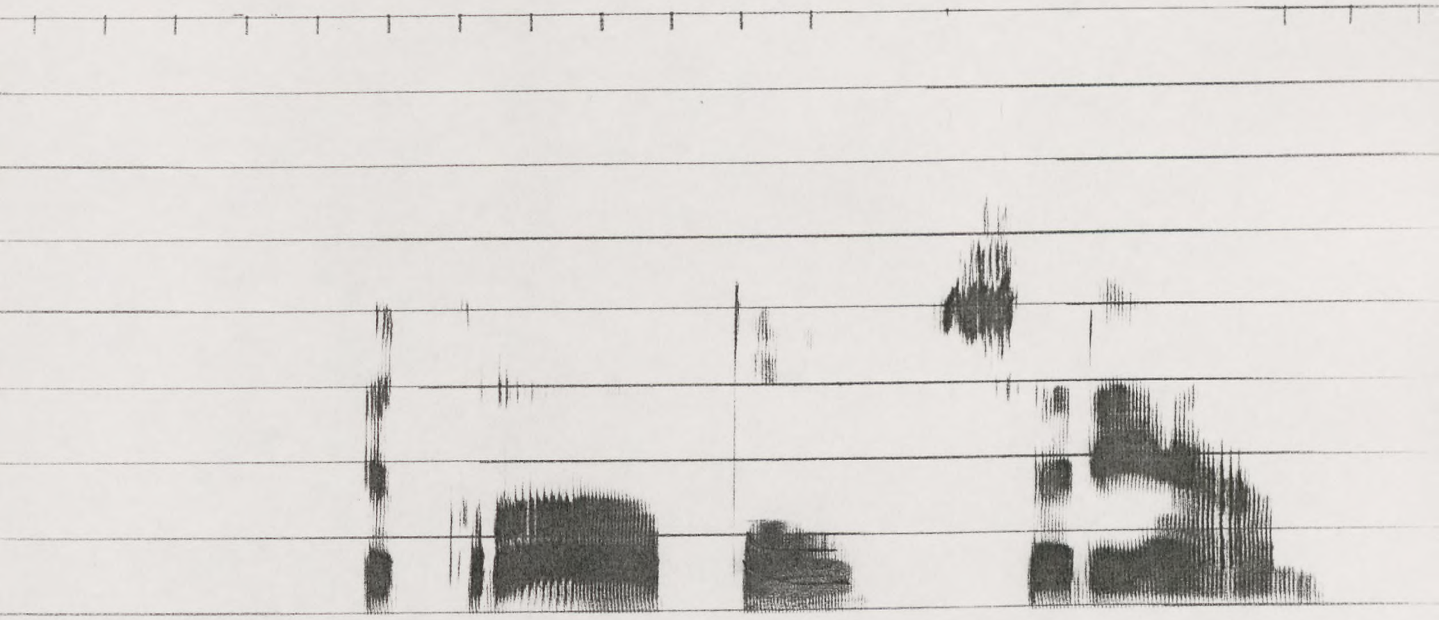
/Taab/



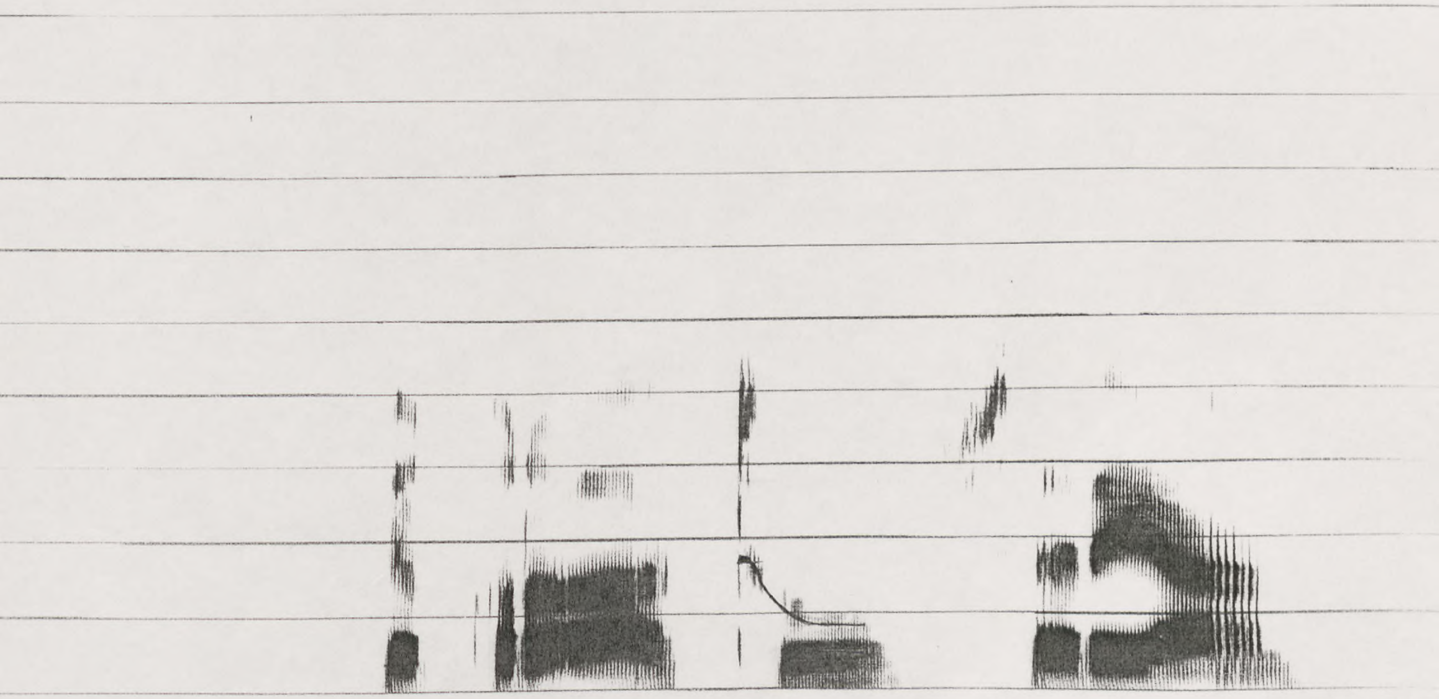
/taab/

Figure 3.2.2 (Continued).

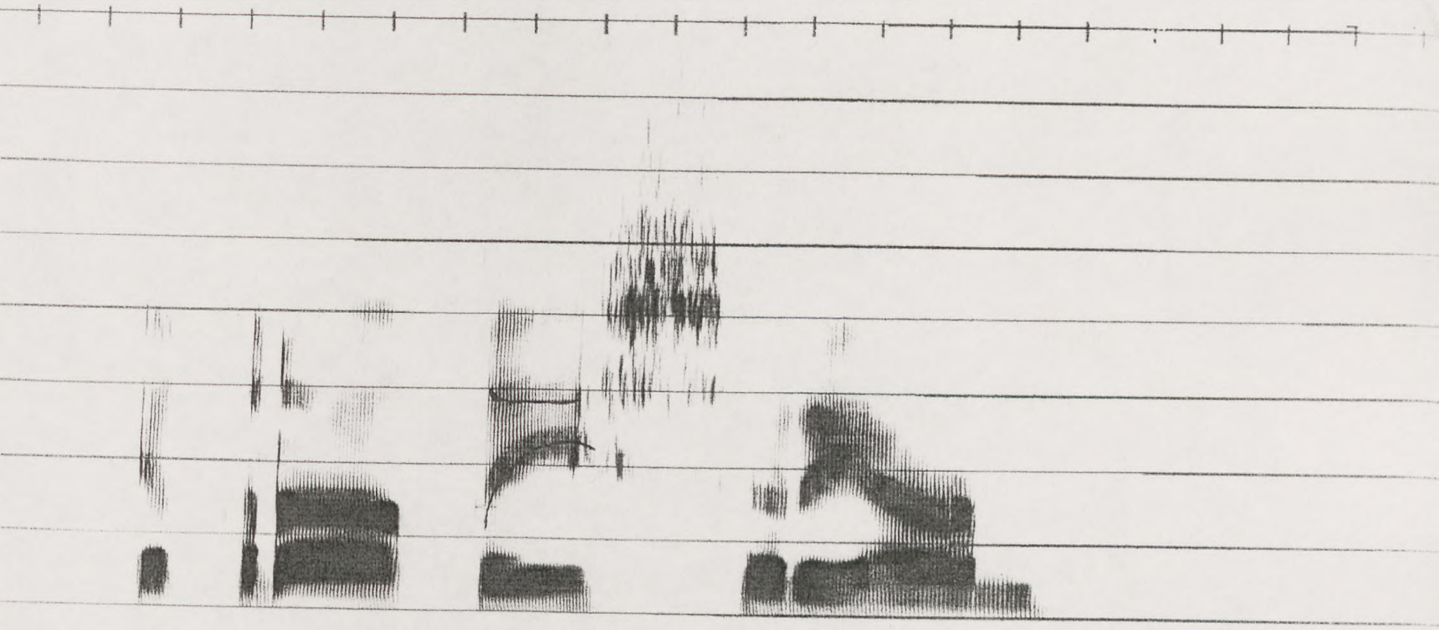
Figure 3.2.2 (Continued).



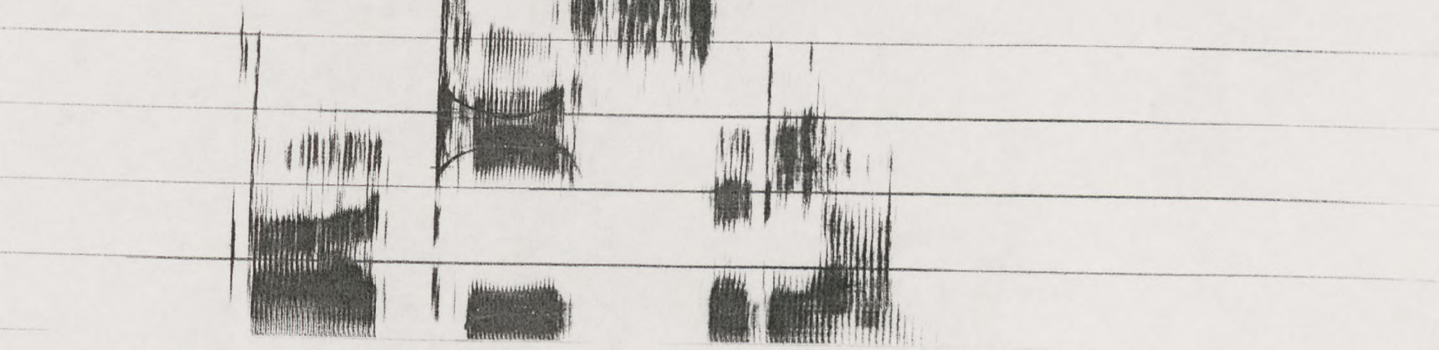
/Tuub/



/tuub/



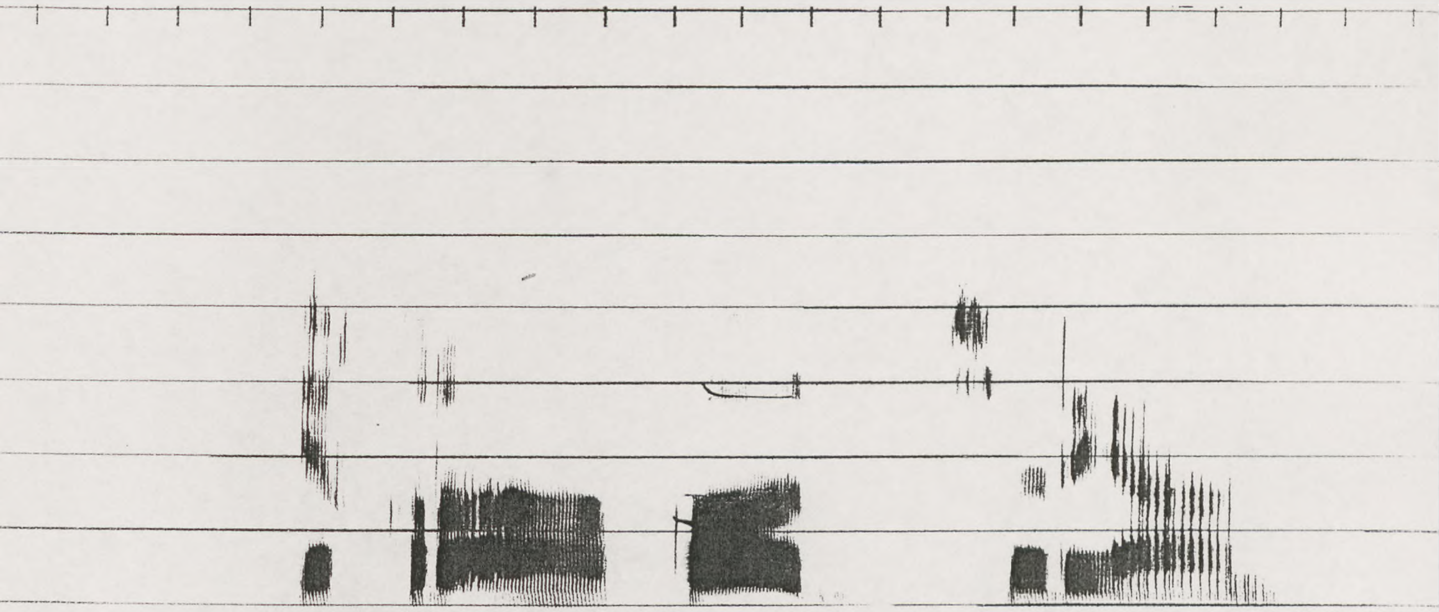
/qiis/



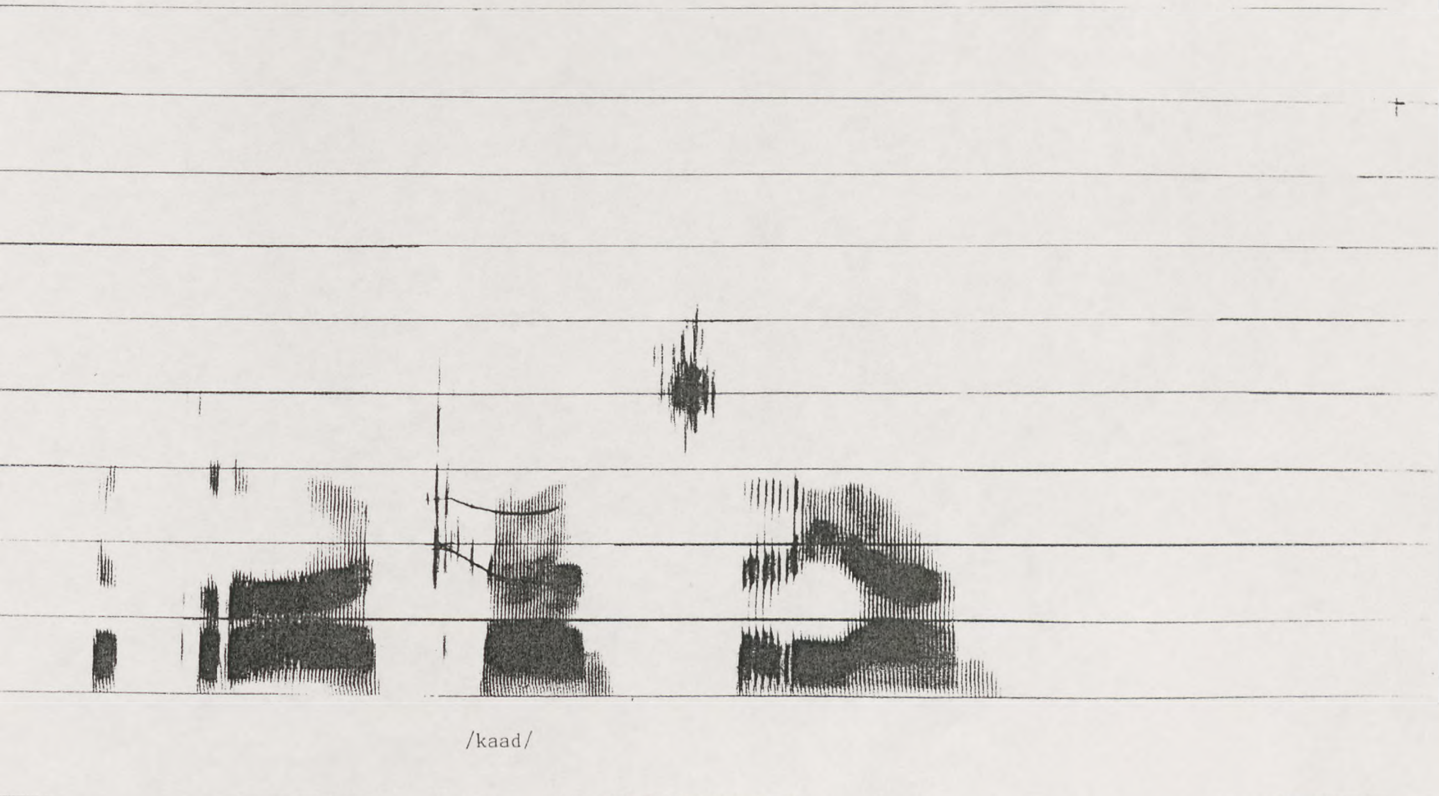
/kiis/

Figure 3.2.2 (Continued).

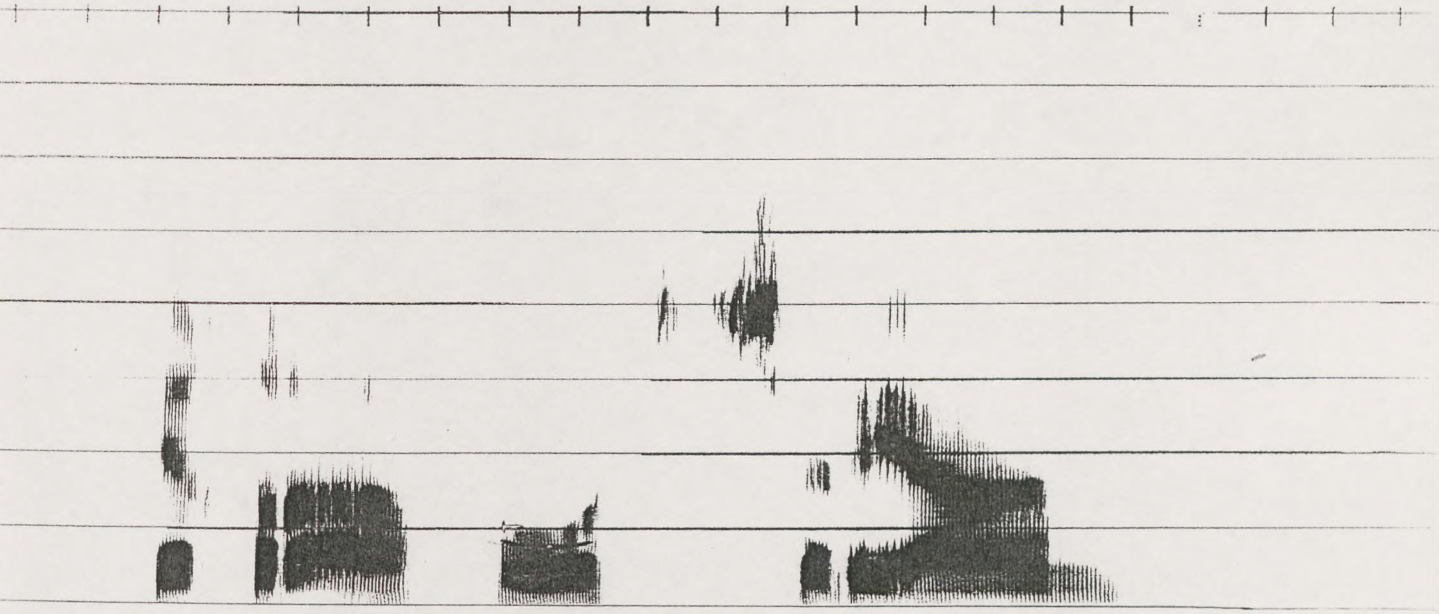
Figure 3.2.2 (Continued).



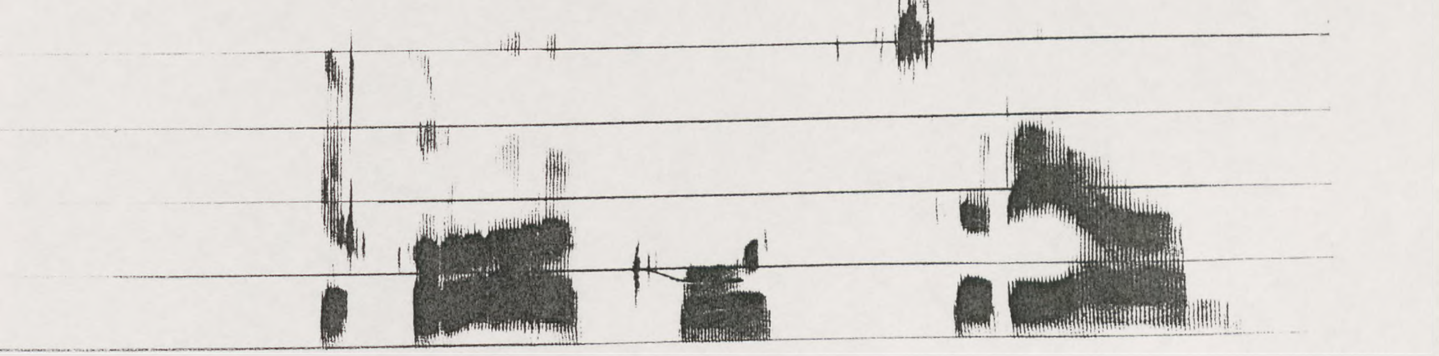
/qaad/



/kaad/



/quut/



/kuut/

Figure 3.2.2 (Continued).

Emp vs Non-emp	N	T-value	P-value
Taab vs taab	6	-8.21	0.0004 **
Daar vs daar	6	-2.68	0.044 *
qaad vs kaad	6	-16.56	0.0000 **
Tiin vs tiin	6	-0.06	0.95
Diin vs diin	6	0.00	1.0
qiis vs kiis	6	-5.97	0.0002 **
Tuub vs tuub	6	-0.98	0.36
Duur vs duur	6	-4.49	0.0030 *
quut vs kuut	6	-1.21	0.26

Table 3.2.6 t-test results for the F2 target positions of the vowel in the emphatic and non-emphatic environments.

As regards the domain of emphasis the spectrograms show that the F1 and the F2 values in the vowels following the emphatics undergo definite modifications as compared with those following the non-emphatics. The F1 approaches F2 in an emphatic environment. This effect is not observed in the same vowels following non-emphatic consonants. This by itself is a sufficient evidence that emphasis is not confined to the emphatic consonant but rather extends its influence to the following vowels. The coarticulatory effect of emphasis is also seen to extend to the preceding vowel even if this vowel is at the word boundary. The second formant transition of /aa/ in / ʔiqraʔaa /, as it appears from the spectrograms, is lower in value when it occurs before the emphatics than when it precedes the non-emphatics. It is pointing downwards to the F2 loci of /D/, /T/ and /q/ while for the non-emphatic consonants it is pointing upwards in the direction of the F2 loci for /d, t, k/. To sum up, the spectrographic investigation of the emphatic and non-emphatic consonants shows an appreciable effect of emphasis on the neighbouring sounds in that it spreads to the vowel formants at both sides of the consonant.

		Subject AA				Subject WS		
	N	/aa/	/ii/	/uu/	**Transition Magnitudes**	/aa/	/ii/	/uu/
/T/	6	-100	45	150	**F1**	-110	140	-95
	6	-105	-1370	255	**F2**	-100	-1160	130
	6	200	205	?	**F3**	70	120	?
/t/	6	-330	-105	-85	**F1**	-270	-100	-150
	6	255	-510	980	**F2**	355	-685	925
	6	300	45	?	**F3**	260	-10	?
/D/	6	-160	65	-90	**F1**	-200	-40	5
	6	-325	-1450	140	**F2**	-225	-1405	-90
	6	240	75	?	**F3**	20	35	?
/d/	6	-375	-155	-160	**F1**	-435	-175	-245
	6	240	-590	870	**F2**	290	-555	735
	6	240	-245	?	**F3**	45	-60	?
/q/	6	-120	215	60	**F1**	-155	150	90
	6	-190	-1250	160	**F2**	-170	-1090	-120
	6	155	75	?	**F3**	135	170	?
/k/	6	-325	-75	-95	**F1**	-285	-90	-100
	6	545	-335	200	**F2**	460	-190	120
	6	275	25	?	**F3**	235	-60	?

Table 3.2.7 The magnitude of formant transitions[3] (Hz) for the two categories in combination with the different vowels.

---

[3] Following common usage (e.g. Halle, et. al., op. cit.) the term 'negative' is used to denote a rise from a low locus frequency to a higher formant target position. Conversely, the term 'positive' is used to describe a shift from a higher locus to a lower formant steady state frequency.

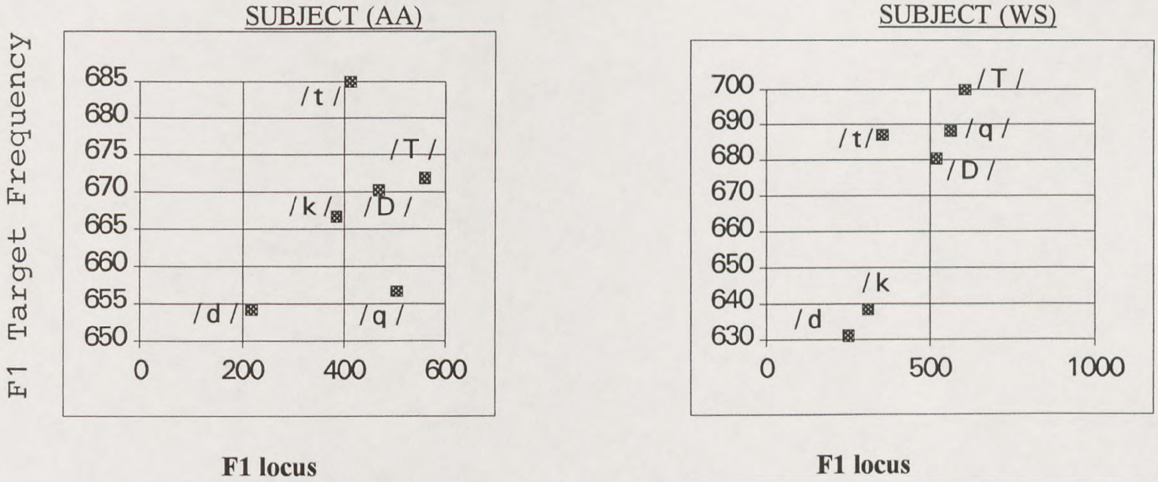


Figure 3.2.3a A plot of F1 values at the consonantal locus and the target position (Hz) in the vowel /aa/.

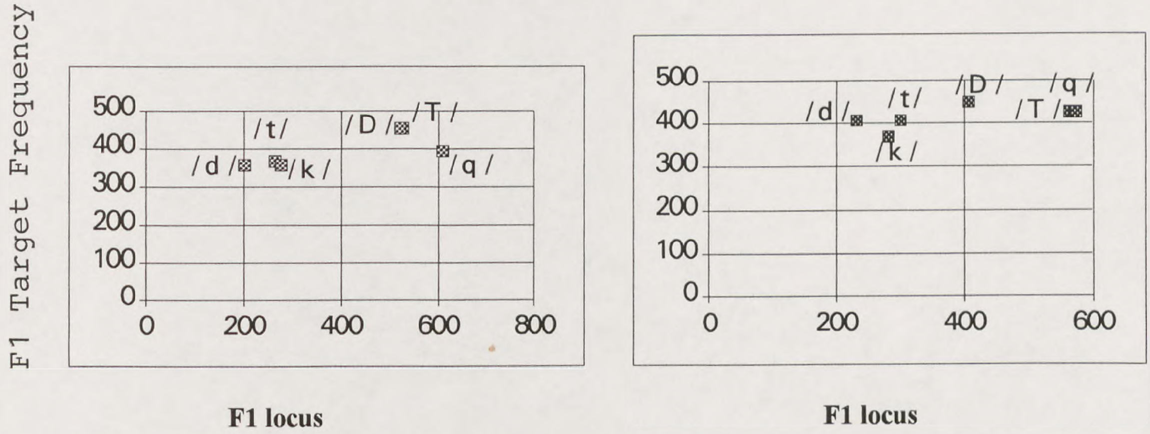


Figure 3.2.3b A plot of F1 values at the consonantal locus and the target position (Hz) in the vowel /ii/.

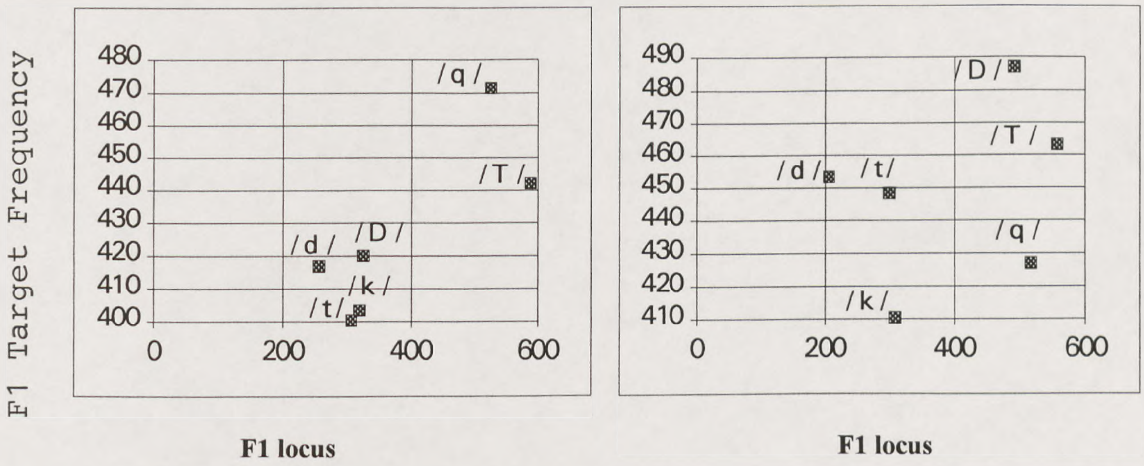
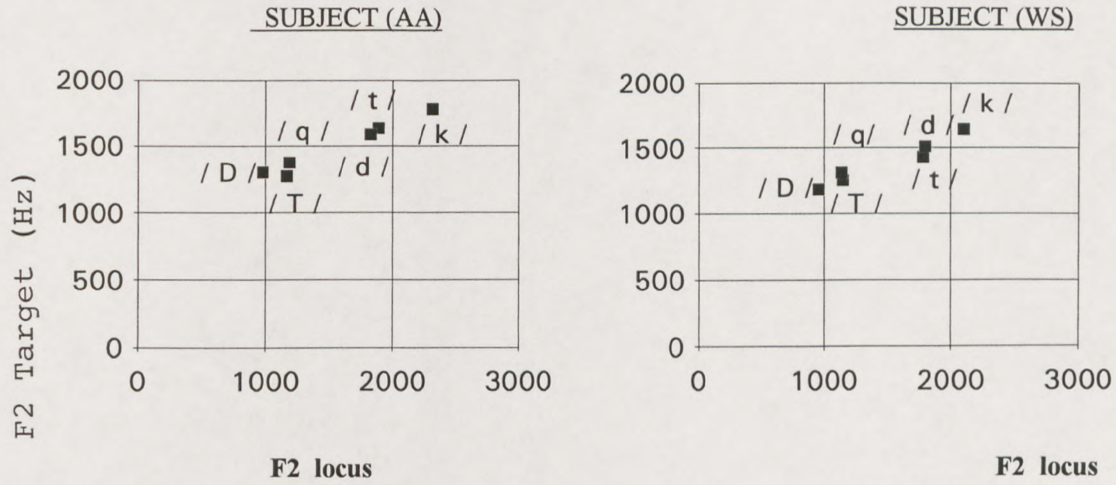


Figure 3.2.3c A plot of F1 values at the consonantal locus and the target position (Hz) in the vowel /ʊʊ/.


SUBJECT (AA)

SUBJECT (WS)

Figure 3.2.4a A plot of F2 values at the consonantal locus and the target position (Hz) in the vowel /aa /.

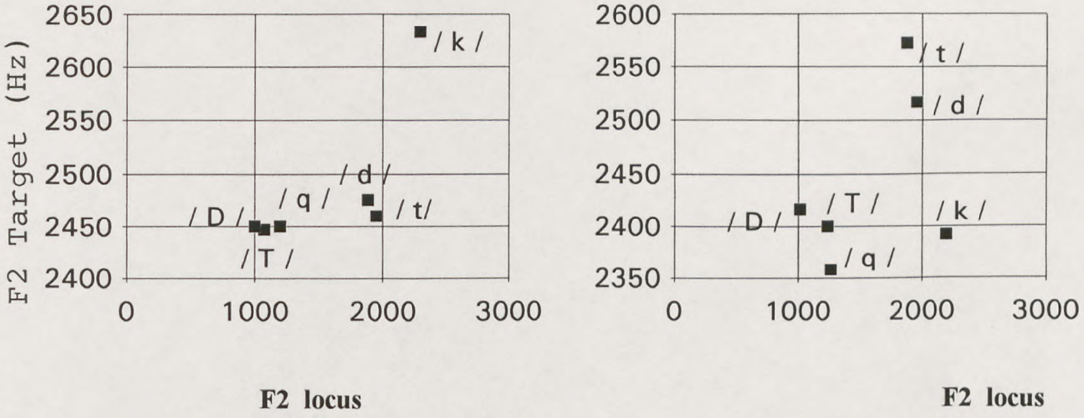


Figure 3.2.4b A plot of F2 values at the consonantal locus and the target position (Hz) in the vowel /ii /.

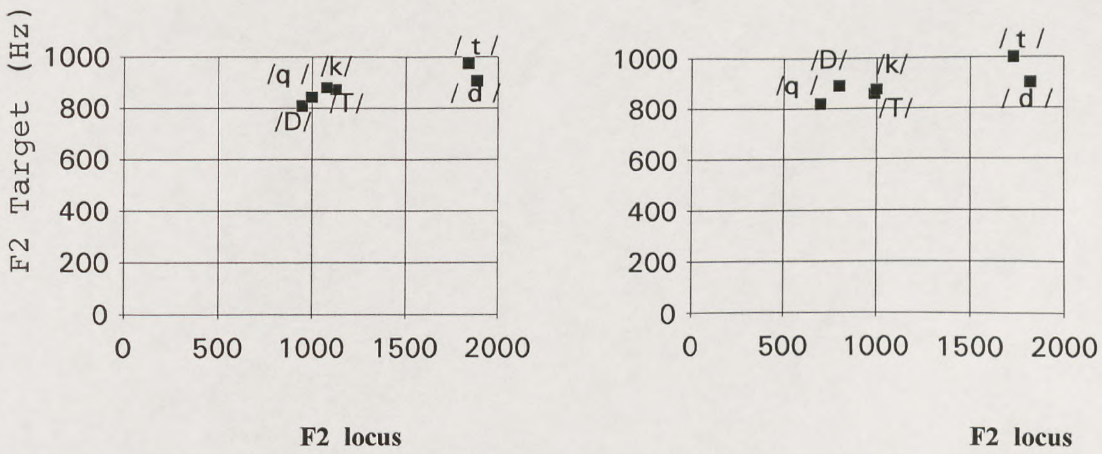


Figure 3.2.4c A plot of F2 values at the consonantal locus and the target position (Hz) in the vowel /ʊʊ /.

Details of the magnitude of formant transitions are presented in Table 3.2.7. This table shows considerable variations and certain acoustic patterns in formant transitions not only for the consonants but also for formant frequencies of the same vowels associated with the different consonants. The extent of formant transitions are dependent on the consonant place as well as the vowel points of articulations. For instance, the first formant transition associated with the non-emphatics is always negative[4]. The magnitude of F1 transition is determined by the position of the steady state frequency of the following vowel. The consonant transition to /aa/, which has a higher target position, is large compared with that to /ii/ and /uu/ whose target positions occur at lower frequencies.

On the other hand, the F1 transitions associated with the emphatics are negative (i.e. rising to /aa/) while with /uu/ and /ii/ they are either falling or flat. Therefore, the transitions to /aa/ are noticeably shorter as compared with that in the non-emphatic category.

Figures 3.2.3a-3.2.3c , which are constructed using data presented in Tables 3.2.2 and 3.2.3, show the extent and direction of the consonant-vowel transitions

The second formant, in its turn, displays similar characteristic changes in transitions for the two sets of consonants. Each consonant category is associated with a second formant transition that distinguishes it from its opposite counterpart. Thus for the emphatics F2 is always rising to /ii/ and /aa/, while for the back vocalic context it is falling or flat. The only rising F2 transitions displayed by the non-emphatics are those in the /ii/ context. With the /aa/ and /uu/ contexts they are falling. As can be seen from Figures 3.2.4a-3.2.4c the main difference in the direction of the F2 transition between the emphatics and non-emphatics is in the /aa/ context. In the former it is rising, while in the latter it is falling. The striking feature of these contrasts is that the emphatic transitions to /ii/ are larger than those of the non-emphatics to the same vowel context. The F2 transitional shift from /T/ to the target position of /ii/ is about -1370 Hz while the corresponding shift of /t/ is about -510 Hz. Similar F2 frequency shifts of transitional patterns are found for /qiis/ (-1250 Hz) and /kiis/ (-335 Hz). Thus one may conclude that the transitional information provided by the first two lower formants are correlated with the emphatic/nonemphatic distinction. They seem quite likely to be able to provide quite adequate cues for their perception.

---

[4] See footnote 3 on page 98.

Even inside each category the extent of rise and fall of formant transitions differs as a function of place of articulation. The extent of F2 transitional fall for /taab/ and /daar/ is about 255 Hz and 240 Hz respectively while for /k/ it is about 545 Hz in the same vowel context.

The F3 transitions are not much different for the two categories. The frequency shifts are quite small compared with that of the F1 and the F2 transitions, except perhaps in the /aa/ context for which some transition variations can be observed. No rising F3 transitions are obtained with any of the two categories except with the denti-alveolar non-emphatics in the environment of /ii/.

It is also possible to describe the emphatics and non-emphatics in terms of frequency position of the second and third formant loci in relation to the second and third formant steady state frequencies. The discriminative power of the second and third formant frequencies is illustrated in Figures 3.2.5a and 3.2.5b.

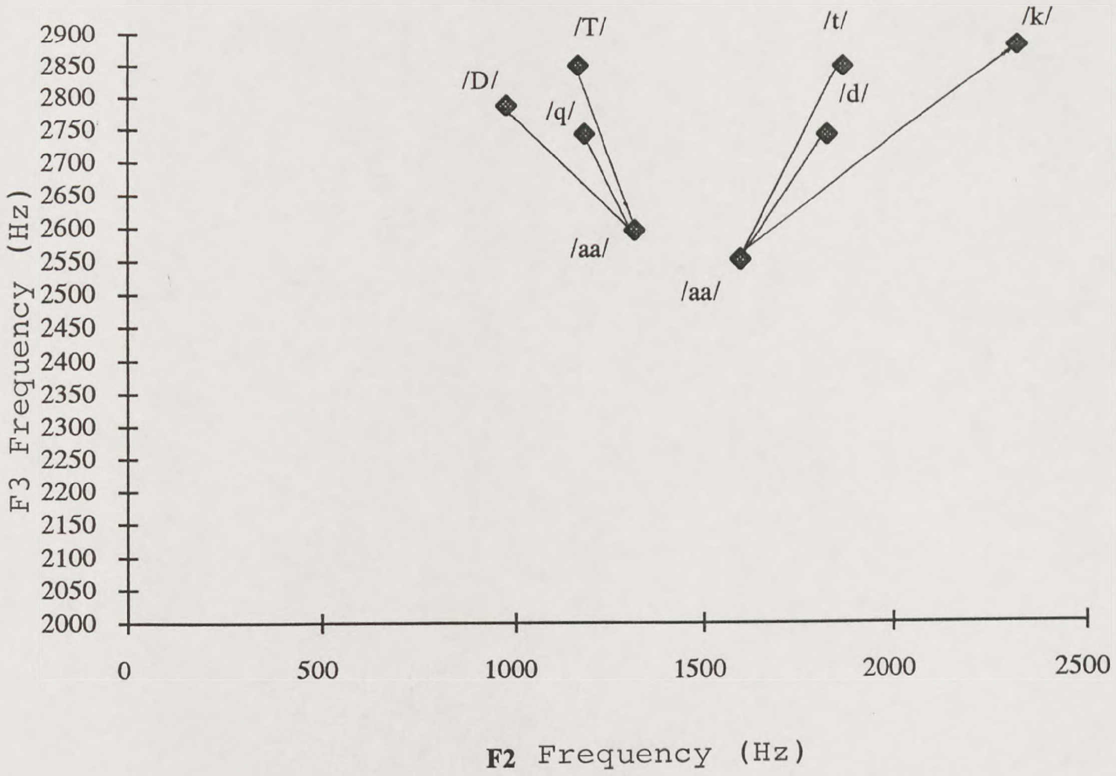
In Figure 3.2.5a both consonant categories show more or less similar F3 loci. On the other hand, the F2 loci in the two sets of consonants occupy quite separate regions. Regarding their vocalic environments, the two categories move to quite similar F3 target positions. For the F2 the average target position across all vowels is slightly lower in emphatic environments than those in the non-emphatic environments.

Similar remarks can be said about Figure 3.2.5b with some reservations. There is no consistent F3 pattern separating the emphatics and the non-emphatics although there is a slight indication of a voicing effect. Thus, the voiced consonants /d/ and /D/ show lower F3 and F2 loci than all others. In this context also the F2 loci are consistently higher for the non-emphatics than for the emphatics.

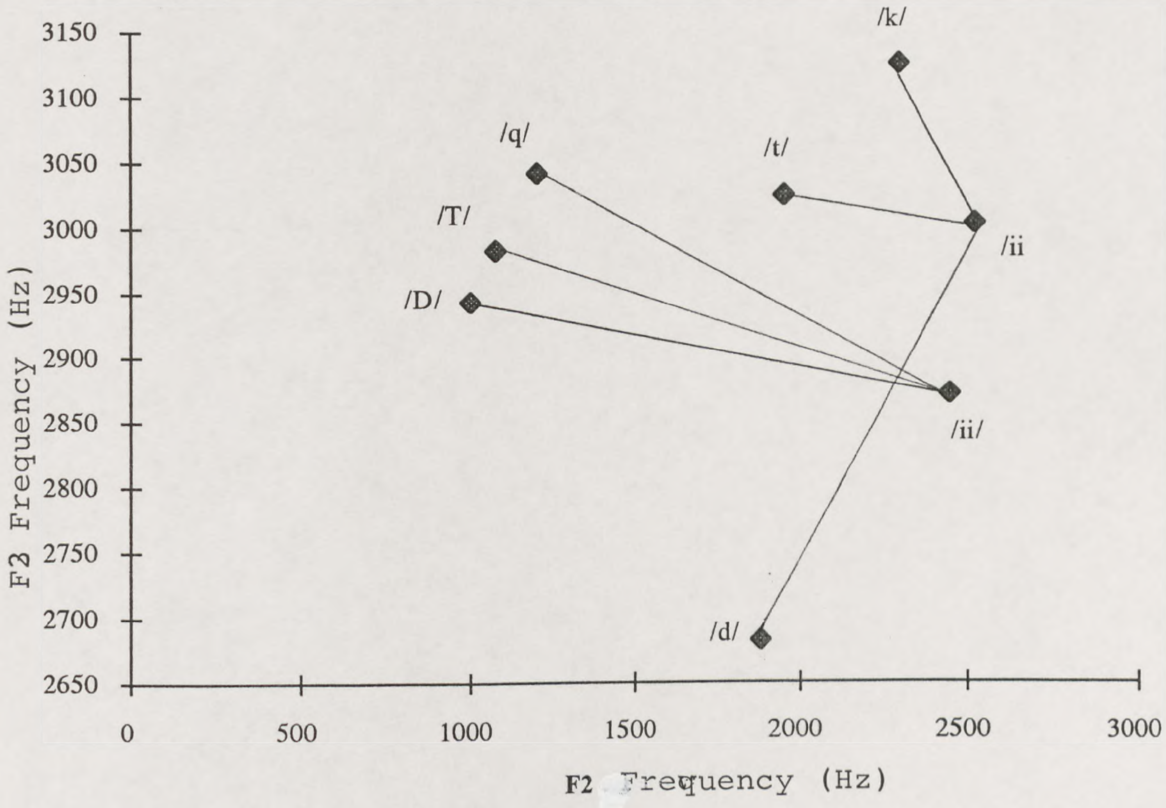
3.2.5 Discussion

The results of this study demonstrate that the major acoustic features separating the emphatics and non-emphatics in YSA are the location of the formant loci and of the formant transitions. With regard to the loci position, the extent of the frequency shifts are very great and marked. The size of shift of the F2 loci, for instance, correlated well with the emphatic/nonemphatic categories and the extent to which they are linguistically utilized as a contrastive device.

The findings on the F1 and F2 loci for denti-alveolar emphatics and non-emphatics are in agreement with those of Odisho (1973) who investigates both sets in nonsense disyllables. His data show the /t/ and /T/ to have F2 loci of 1800 and 1100 Hz respectively; while for /d/ and /D/ the



Figures 3.2.5a A plot of the second and third formant loci for the emphatics and non-emphatics in the context of /aa/.



Figures 3.2.5b A plot of the second and third formant loci for the emphatics and non-emphatics in the context of /ii/.

reported values are 1700 and 1100 Hz in that order. These figures of the F2 loci are very close to the ones given in Table 3.2.2. Similar relations in the two studies for the F1 loci may be seen in the two categories. Contextual dependency of F1 loci is highlighted in our material. However, the generalizability of his results is constrained by his use of nonsense syllables. Although nonsense syllables are convenient for arriving at phonetic conclusions, one might still argue that such syllables lack some linguistic constraints that prevent their extensively different phonetic manifestations by different individuals.

The findings of Giannini and Pettorino (1982) with regard to the assigned F1 and F2 loci are consistent with most remarks made in this study. Their allocation, however, of fixed F1 and F2 loci for the emphatic and nonemphatic categories involve some oversimplification. For instance, it is implied that all emphatics are associated with fixed F1 and F2 loci and all the non-emphatics have one single locus regardless of their place of articulation or vocalic contexts. Thus, the determined F1 and F2 loci for the former group are at 600 and 1000 Hz respectively, while for the latter group they are 250 and 2000 Hz. Our materials have emphasized that some variations occur for the consonants of each category. These may have been caused by one of three factors: place of articulation, vocalic context, and voicing. The variations that we observed are more complex than the fixed loci values reported by Giannini and Pettorino (1982). The predictions that invariant F1 and F3 loci for each consonant may be found is not supported by the findings of this study. The ranges shown for the F1 loci in the nonemphatics are 200 to 350 Hz while for the emphatics they are between 500 and 600 Hz. These frequency ranges are higher when the emphatic/nonemphatic articulations are in the context of vowels articulated at the back section of the vocal tract as compared with those produced at the front.

As far as the context dependency is concerned, the results support the claim that gives one second formant locus only in the denti-alveolar region to each consonant in different vowel contexts with some exception. Oneway analysis of variance has shown insignificant differences in F2 loci for a certain denti-alveolar consonant across the three vocalic contexts, except for /D/. The F2 loci for the /q/ and /k/ vary significantly over a range of frequencies. One possible explanation for this is that the place of articulation for /k/ is more variable while for the denti-alveolars it is relatively stable. There is general agreement among phoneticians (e.g. Abercrombie, 1967; Jones,1976;

Gimson,1989) that /k/ production in English is characterized by shifts in the points of articulation depending on the neighbouring vowels. Thus, its articulation is relatively advanced (i.e. palatal) when it occurs in the vicinity of front vowels and moves back to the velar region when it is in the proximity of back vowels. One would, therefore, expect the change in the F2 loci following front and back vowels to reflect these articulatory adjustments in vocal tract configurations associated with each one of them (Fant, 1960).

In terms of articulation such acoustic features should be considered together with the presence of a secondary constriction at the back section of the vocal tract probably at the pharyngeal region. This correlation between articulation and the acoustic features is confirmed by X-ray data for Arabic emphatics and non-emphatics (Giannini and Pettorino (1982). For the emphatic articulation they observe that the root of the tongue is brought back towards the wall of the pharynx forming a 3 mm constriction. It is also in line with the predictions made by Klatt and Stevens (1969).

On the other hand, the extent and direction of the F1 and F2 transitions appear to provide reliable acoustic information for the emphatic/nonemphatic contrasts. The F3 transitions are the same in both consonants categories. However, a combination of the second and third formant transitions may provide additional acoustic information for the characterization of these consonants.

The formants bendings of the YSA plosives are to a great extent similar to those displayed by Danish, Swedish and English plosives with some exceptions. In YSA, for example, front vowels following denti-alveolars show more bendings. This may be attributed to the relatively short interval following the release of the plosives in YSA which may help the early detection of the formants' onsets. In addition the strong frication and aspiration noise associated with the Danish and Swedish voiceless plosives may cause some delay in the formant identification. Concerning the voicing effect, the consonant vowel transitions, particularly those of F2 and F3, are longer for the voiced sounds /d, D/ than for the voiceless counterparts /t, T/. A possible interpretation of this type of variation is provided by Fant (1973). According to him the greater range of F2 transitions associated with the voiced as compared with the voiceless ones is due to a difference in the relative timing of tongue movements from the position for the consonant to that of the vowel. Such actions are reported to occur earlier in the former as compared with the latter type of sounds. Supporting evidence for this explanation is also published by Gay (1979) from EMG data in which the tongue movement for the following vowel is seen to occur earlier in a voiced sound as opposed to an unvoiced one.

The emphatic transitions are distinct from those of the non-emphatics not only in that they are larger but also they are observed to extend, at least, up to the middle of the following vocoids. In this respect progressive and regressive coarticulations are quite common in this study. The former type are described by MacNeilage and DeClerk (1969) as a phenomenon which consists of "...articulator positions being modified in the direction of the preceding phoneme representation". They suggest that it may be caused by 'mechanical constraints' which may have prevented the articulators from reaching the articulatory position. From our spectrograms it seems that the production of the vowels involves concomitant articulatory adjustments more like that of the emphatic consonants.

The formant steady states are influenced not only by the position of the constriction but also by the presence or absence of lip rounding in a specific vowel. Its acoustic effect resembles that of a retraction of the tongue to a back location. This may explain why /uu/ following emphatics is associated with even lower F2 and relatively higher F1 target positions than those of the same vowel in non-emphatic environment. In the former category the acoustic realization seems to be a function of at least two articulatory gestures, pharyngealization and lip rounding, whereas in the latter one it is assumed to be caused by lip rounding.

Regressive coarticulation is quite visible on our spectrograms as shown by the influence of the emphatic consonants on the formants of the preceding vocoids. In such a case, it is likely that some aspect of the articulatory control for the earlier syllable component is influenced by the following one, particularly the consonantal element.

Such coarticulatory effects are not pecular to YSA. Vocalic influence on the preceding or the following consonants has been documented for Swedish (Öhman, 1966) and for English (Lehiste and Peterson, 1961). Since the present acoustic data for YSA indicate that emphasis spreads to the surrounding vowels, it becomes necessary to treat this feature as one that involves, at least, the CV- syllable or the -VCV- portion of an utterance rather than being confined to the consonant only.

3.3 Other temporal characteristics of voicing and emphasis in YSA

3.3.1 Introduction and aims:

Previous research emphasizes the multiplicity of acoustic cues operating in support of one linguistic feature (See section 2.1.3). Only one of them will be a salient feature in a certain contexs while the others may be redundant. An example of this may be illustrated by the phonological feature of voicing in English where the consonants /p,b/ , /t,d/ , and /k,g/ are differentiated by a number of acoustic correlates such as VOT, formant transitions, aspiration noise and the amplitude of the burst. Yet, these correlates do not have equal importance in all positions and contexts. As far as YSA is concerned the relation between VOT and the voicing categories has been examined in section 3.1. In the present section an attempt to investigate yet another acoustic variable, namely duration, will be described.

The motive behind this investigation is to see the extent to which phonological variables are found to control timing. As the discovered durational patterns of segments in YSA are related to those of other languages the chance for increasing our understanding of the 'universal' or the global rules will also be enhanced. Although this is not the main aim of this study, in our view, the success of any future duration model will be dependent on a sound knowledge of the various temporal relations across a wide range of unrelated languages. It is by now recognised that duration has different linguistic functions in different languages. This implies that a lot of the durational variations may be attributed, in part, to the phonological structure (i.e. the nature of the phonemic contrasts) employed by the individual languages. In certain languages, for example, a difference in the meaning of two words may be associated with just a change in the consonant duration as in the Turkish words */bati/* 'west' and */bat:i/* 'sink' or a change in the duration in the vowel as in the YSA words */sad/* 'dam' and */saad/* 'spread'.

Three durational variables will be considered in this study: the duration of the 'acoustic'[5] closure, the contoid duration and the vocoid duration. It could be that any of the phonological features, emphasis or voicing, may have as one of their acoustic manifestations a difference in any of these three intervals. For that purpose the three temporal intervals are investigated under a variety of

---

[5] The 'acoustic' closure is here differentiated from the 'articulatory' closure in that the latter is assumed to reflect the actual behaviour of the articulators in the vocal tract at the time of constriction, whereas the former represents only an indirect estimate of that behaviour.

manipulations that are predicted to affect their absolute durations. We hope that our data will not only contribute to the description of the temporal relationships governing the phonological features of emphasis and voicing in YSA, but also that they will help us discover the influence exerted by other contextual factors such as stress, gemination and place and manner of articulation on duration.

Since from the above survey of the published sources consonants exhibited readily discernible patterns of environmental influence on the duration of neighbouring vowels, we will attempt to examine the vowel length variation as a function of the presence or absence of voicing.

In addition to that we are interested in testing a claim recently made by Port (1981) which suggests that the consonant/vowel ratio is correlated with voicing and may be independent of certain contextual factors. One of the advantages of this measure, he claimed, is that it is relational and therefore not dependent on the absolute durations of the consonant and vowel. This advantage may make it strong enough to resist any contextual changes. In view of the considerable changes in consonant and vowel duration in this study, due to gemination and stress, it is interesting to see if the relation between voicing and the consonant/vowel ratio will still be unaffected by context.

3.3.2 Sample material and procedure:

Table 3.3.1 includes sequences of words containing intervocalic consonants. The test consonants are systematically varied with respect to voicing, emphasis, stress, gemination and place of articulation. The voicing categories include /t, d/, /T, D/, /k, g/ and /s, z/, whereas the emphatic/nonemphatic consonants are represented by the contrasts /T, t /, /D, d / and /q, k /. The bilabial sound /b/ is only included for the investigation of the place of articulation effect on duration by comparing it with the denti-alveolar and the velar consonants /d/ and /g/ respectively.

All the above consonants are in disyllabic words. They are preceded and followed by the same vocalic context. The word list in each context consists of minimal or near minimal pairs. They are varied in four different conditions: stressed, unstressed, geminated and ungeminated. In the unstressed condition they are all preceded by the vocalic context /a/ and followed by /ii/. In YSA the shift of stress from first to second syllables results in a change in the phonological structure of the word and consequently a difference in word meaning.

In the geminated/ungeminated conditions all consonants are preceded by the vocalic context /a/ and followed by /aa/. In order to investigate the vowel in different consonant environments, an attempt has been made to control all factors which are known to influence its duration. For this reason the following vowel (i. e. /aa/) in many test words is surrounded by the same consonant in the same word. For example, the /aa/ in /fa`taat/ and /sa`daad/ is surrounded by /t/ in the former and by /d/ in the latter. It is hoped that any variation in its duration in the two words would be attributed to those consonantal environments.

Context	**Test words containing the intervocalic consonants**	**Context**	**Test words containing the intervocalic consonants**
Unstressed:		**Ungeminated :**	
	`qatim		fa`taat
	`qadim		sa`daad
	`xaTir		ma`Taal
	`xaDir		qa`DaaD
	`rakib		ma`kaan
	`ʕagib		ma`gaal
	`ʕabiq		ma`saas
	`ʕaqib		ħa`zaaz
			ʃa`baab
			ma`qaal
Stressed:		**Geminated:**	
	qa`tiil		fat`taat
	qa`diim		sad`daad
	xa`Tiir		maT`Taal
	na`Diir		ħaD`DaaD
	ra`kiik		mak`kaar
	ra`giib		ħag`gaad
	ʕa`biiq		ħas`saas
	ra`qiib		ħaz`zaaz
			ʃab`baab
			ħaq`qaad

Table 3.3.1 Sample material used for study of the durational characteristics in emphasis, voicing and the consonant place of articulation in YSA.

Within each of the above four conditions prosodic patterns for every member of the pairs are similar. This will help to minimize variability in duration due to suprasegmental factors. Since the recorded utterances in each context involved a constant stress and intonation patterns, it seems reasonable to assume that stress affects each word in essentially the same manner.

All test words are embedded in the frame sentence /ʔiqraʔaa ---------- sariiʕan/ 'read (you two) --------soon' and each one of them is recorded six times by each of the two native speakers of YSA. Our subjects are required to keep a normal loudness with the additional instruction to stress the marked syllable in the target words. They are also asked to follow a normal and constant rate of speech. For further details on the recording procedure see section 3.1.2.2

Spectrograms are produced using the 700 conventional spectrograph. Segmentation criteria are adopted and duration measurements are obtained.

3.3.3 Segmentation & measurement criteria:

Measurements of the 'acoustic' closure, the contoid duration and the following vocoid durations are made from the speech of two subjects, AA and WS. The criteria for segmentation agree substantially with those established by Fant (1973), Jorgensen (1964), Crystal and House (1982) and Peterson and Lehiste (1961). Slight changes are, however, made to suit our purposes and needs. Peterson and Lehiste have reckoned the vocoid from the release after /bdg/ and from the start of voicing after /ptk/. In our study the beginning and end of the vocoid correspond to the beginning and end of the vowel without including any portion of the release as part of the vocoid (See Figure 3.3.1). The intervals for frication and aspiration noise (if present) are considered part of the preceding contoid. The identification of the termination points of the vocoid is aided by the relative clarity with which they appear on the spectrograms as sudden reduction in intensity level. The vocoid is also determined by the presence of voice in combination with the formant frequencies and the increased darkening intensity bands associated with it.

The 'acoustic' closure is defined as the interval between the offset of the preceding vocoid and the onset of the plosive release marked by the sudden increase in acoustic energy. All closure duration measurements are made for the intervocalic plosives. Word-initial plosives are avoided because of the possibility that a pause may occur simultaneously with the closure, thus invalidating any comparison which may be made.

The whole consonant duration corresponds to the contoid duration. It is regarded as the interval between the termination point of the preceding vocoid and the onset of the following one. By adopting this criterion, most consonants investigated, plosives or fricatives, are clearly identified. A few cases with unclear boundaries are found and are discarded. The fricatives are identified by the visible random frication noise. All measurements are made in milliseconds.

Based on this definition the plosive duration includes not only the closure but also the release interval. In acoustic terms, this means the inclusion of the release transient, the frication and

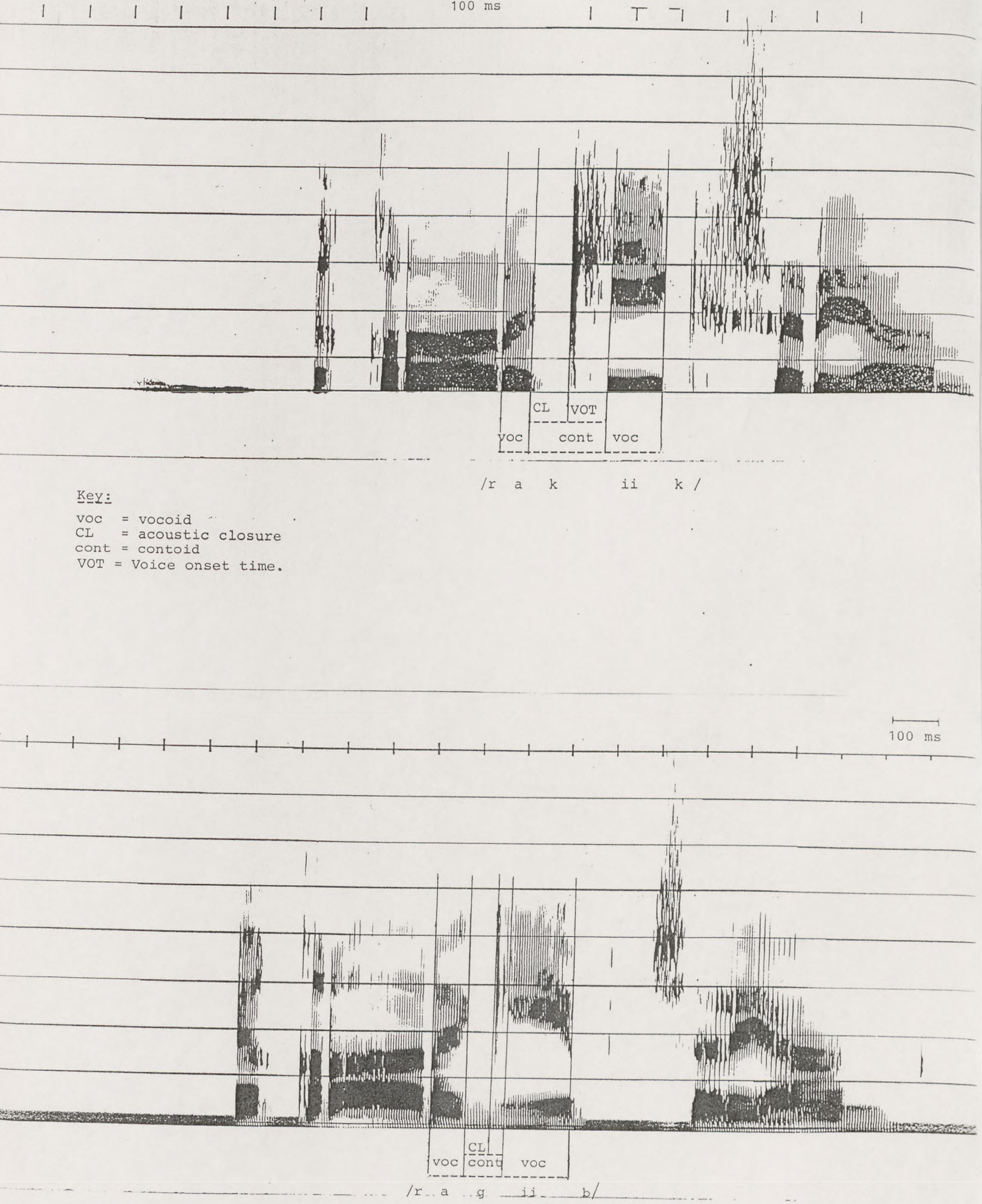


Figure 3.3.1 Wideband spectrograms of the words /rakiik/ and /ragiib/ in the carrier sentence /?iqra?aa ----sariiƐan/. The segmentation criteria used for the duration measurements of the acoustic closure, th contoid and the following vocoid are indicated. The abscissa is time; The ordinate is frequency in KHz.

aspiration portions as part of the consonant. This procedure agrees with Crystal and House (1988) who define the 'complete' plosive duration as "one that includes an identifiable hold portion, or occlusion, followed by a plosion (aspiration and/or frication) release" (p. 285).

Measurement of the contoid duration is motivated by our desire to provide data for comparing plosive and fricative consonant durations and to get an idea of the influence of the manner of articulation on segment duration. Some investigators have assumed that the closure represents the duration of the whole plosive (e.g. Klatt 1976). This is not the view adopted in this study.

To end this section a remark should be made about the problem of segmentation. Complete accuracy and distinctiveness of segments is unattainable particulary when measuring sound duration. Nevertheless, consistency in performing all our measurements is believed to be achieved.

3.3.4 Results

3.3.4.1 Consonant duration and the voicing distinction:

The mean closure duration for the intervocalic voiced/voiceless consonants and their ratios are presented in Table 3.3. 2. The standard deviation for each category is included within brackets. In most cases the voiceless plosives are associated with longer closure duration than the voiced ones. This is true regardless of stress condition and subject variation. For example, the average difference between /k/ and /g/ in the unstressed condition is 13 ms while for /T/ vs /D/ it reaches 20 ms as produced by subject WS. Subject AA produces only a narrow and perhaps negligible difference. It is not more than 6 ms in both stressed and unstressed contexts. In spite of that the obtained measurements for WS are consistently shorter than those shown by AA. On many occasions subject WS provides closure duration values for the voiceless plosives that are shorter than those obtained by the voiced plosives for subject AA. As far as voicing is concerned this observation underlines the fact that the timing control is subject to individual modifications. Each of the two subjects, however, displays a consistent durational pattern. Evidence of that is the constant voiced/voiceless ratios for each subject in different stress conditions.

			Subject AA		Subject WS	
Context	**Intervocalic Voiced & Voiceless consonants**	**N**	**Mean (ms) & (s.dev.)**	**Voiced/ Voiceless Ratio**	**Mean (ms) & (s.dev.)**	**Voiced/ Voiceless. Ratio**
Unstressed:						
	`qatim	6	78 (4)	0.81	65 (15)	0.92
	`qadim	6	63 (14)		60 (0)	
	`xTir	6	96 (14)	0.97	78 (7)	0.74
	`xaDir	6	93 (16)		58 (4)	
	`rakib	6	78 (16)	1.00	73 (8)	0.82
	`ʕagib	6	79 (5)		60 (9)	
	Mean :					
	Voiced		78	0.93	59	0.82
	Voiceless		84		72	
Stressed:						
	qa`tiil	6	86 (8)	0.97	75 (8)	0.93
	qa`diim	6	83 (9)		70 (8)	
	xa`Tiir	6	93 (14)	1.00	92 (7)	0.78
	na`Diir	6	94 (11)		73 (15)	
	ra`kiik	6	100 (6)	0.89	93 (12)	0.73
	ra`giib	6	89 (13)		68 (11)	
	Mean :					
	Voiced		87	0.94	70	0.80
	Voiceless		93		87	

Table 3.3.2 Mean closure durations, standard deviations and duration ratios for the voiced and voiceless consonants in stressed and unstressed contexts.

Both subjects show longer closure duration in stressed than in unstressed conditions. For example, subject AA obtained an average difference of 9 ms for the voiced plosives. The same influence of stress is seen in the voiceless plosives. The closure interval for the voiced plosives is characterized by the presence of voicing in the lower frequency region, whereas for the voiceless ones there is no such presence of acoustic energy.

The voicing contrasts in ungeminated consonants show little variation in the closure duration for subject AA. This may be exemplified by the mean ratio displayed by the voicing cognates in the ungeminate consonants (1.04) see table 3.3.3. This ratio suggests that roughly similar durational values are obtained for the voiced and the voiceless plosives. In the geminated contexts the difference is quite large (36 ms for WS and 22ms for AA) but the ratio is slightly higher than that in the ungeminated context. Clearly the mean ratio for the voiced/voiceless distinction does not remain constant in both contexts. This may be attributed, in part, to the longer closure duration obtained for the geminated consonants as compared with the single forms. The ratio is approximately (2:1).

Since the closure duration does not yield a decisive difference between the voicing categories, our attention is turned to the whole consonant duration in the hope that it will peform better than the closure duration. Tables 3.3.4 and 3.3.5 present the mean durations for the voicing cognates in the four conditions: stressed/unstressed and single/geminate plosives.

The results in all cases confirm our prediction that the voiceless plosives are associated with longer duration than the voiced ones. The difference is quite large and consistent for both subjects. For example, the ratio obtained for /d/ and /t/ in the words /ˋ*qadim*/ and /ˋ*qatim*/ is (0.62) as produced by subject AA and (0.66) by subject WS. These ratios indicate that the voiceless plosives are almost 35% longer than the voiced ones. The stress effect on duration is consistent with the remarks made earlier (namely that consonants in stressed syllables are associated with longer duration than those in unstressed syllables).

To illustrate these differences the duration values for the voicing cognates are plotted as function of stress in Figure 3.3.2. In the unstressed condition the voiced plosives are approximately within the range 83-106 ms while the voiceless plosives are located within 116 - 133 ms.

			Subject AA		Subject WS	
Context	**Intervocalic Voiced & Voiceless Consonants**	**N**	**Mean (ms) & (s.dev.)**	**Voiced/ Voiceless. Ratio**	**Mean (ms) & (s.dev.)**	**Voiced/ Voiceless Ratio**
Ungem.						
	fa`taat	6	67	1.01	85	0.80
			(10)		(6)	
	sa`daad	6	68		68	
			(8)		(8)	
	ma`Taal	6	91	1.02	106	0.59
			(8)		(14)	
	qa`DaaD	6	93		63	
			(10)		(8)	
	ma`kaan	6	72	1.09	90	0.93
			(9)		(9)	
	ma`gaal	6	79		84	
			(13)		(11)	
	Mean :					
	Voiced		80	1.04	72	0.77
	Voiceless		77		94	
Gem.						
	fat`taat	6	169	0.86	202	0.86
			(10)		(19)	
	sad`daad	6	146		173	
			(15)		(5)	
	maT`Taal	6	173	0.91	203	0.77
			(17)		(5)	
	ħaD`DaaD	6	158		156	
			(23)		(11)	
	mak`kaar	6	150	0.81	192	0.83
			(10)		(10)	
	ħag`gaad	6	122		159	
			(12)		(3)	
	Mean :					
	Voiced		142	0.87	163	0.82
	Voiceless		164		199	

Table 3.3.3 Mean closure durations, standard deviations and duration ratios for the voiced and voiceless consonants in geminated and ungeminated contexts.

			Subj AA		Subj WS	
Context	**Intervocalic Voiced & Voiceless consonants**	**N**	**Mean (ms) & (s.dev.)**	**Voiced/ Voiceless Ratio**	**Mean (ms) & (s.dev.)**	**Voiced/ Voiceless. Ratio**
Unstressed:						
	\`qatim	6	133 (7)	0.62	114 (9)	0.66
	\`qadim	6	83 (9)		75 (6)	
	\`xaTir	6	116 (14)	0.91	99 (11)	0.82
	\`xaDir	6	106 (18)		81 (10)	
	\`rakib	6	133 (27)	0.74	128 (14)	0.59
	\`ʕagib	6	99 (8)		76 (14)	
	Mean :					
	Voiced		96	0.75	77	0.68
	Voiceless		127		114	
Stressed:						
	qa\`tiil	6	153 (9)	0.71	153 (15)	0.82
	qa\`diim	6	109 (11)		125 (8)	
	xa\`Tiir	6	128 (19)	0.88	123 (12)	0.75
	na\`Diir	6	113 (16)		92 (13)	
	ra\`kiik	6	178 (10)	0.66	163 (13)	0.72
	ra\`giib	6	117 (20)		117 (19)	
	Mean :					
	Voiced		113	0.74	111	0.76
	Voiceless		153		146	

Table 3.3.4 Mean contoid durations, standard deviations and duration ratios for the voiced and voiceless consonants in stressed and unstressed contexts.

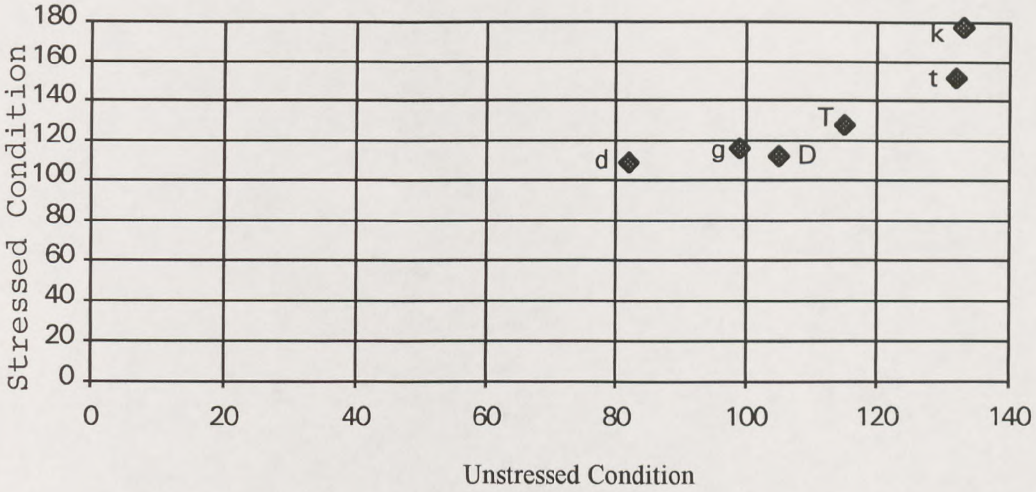


Figure 3.3.2 The relation between mean contoid duration (ms) and voicing in stressed and unstressed contexts .

In the stressed condition the voiced plosives are roughly within the range 109- 117 ms whereas the voiceless ones are extended over a wide range (128 - 178 ms). Clearly the contrasts /t, d/ and /k, g/ are well separated by this measure whereas /T, D/ show only a very narrow difference indicating that the whole consonant duartion may not be an important acoustic correlate of this pair.

The voicing contrast for the fricative consonants /s/ and /z/ in both geminated and ungeminated consonants shows similar durational patterns to that of the plosives with some minor modifications. The fricatives, both voiced and voiceless, are in almost all contexts longer than the plosives. However, the increase is small as it appears from Table 3.3.5. The same remark applies to geminate consonants. Almost all voiced and voiceless geminates are within the range 158-213 ms with the exception of the fricative consonants /ss/ and /zz/ which are observed to occur a few milliseconds above that range.

It can, therefore, be concluded that the contoid durational measure is more robust and efficient in distinguishing voicing in YSA than the closure duration.

The other variable examined in connection with voicing is the duration of the vocoid . Table 3.3.6 displays the contoid and vocoid durations and their ratios for both subjects. A close examination of the results reveals minimal variations in the duration of the vocoid following voiced and voiceless environments. The mean vocoid duration following voiced and voiceless consonants in ungeminated context is 102 ms for both categories, while the values obtained for subject WS are 188 and 186 ms for the two categories respectively. This indicates that the effect of voicing on the following vocoid duration in YSA is negligible. Gemination effect, though quite small, is more apparent than that of the vocoid duration.

			Subj. AA		Subj. WS	
Context	**Intervocalic Voiced & Voiceless Consonants**	**N**	**Mean (ms) & (s. dev.)**	**Voiced/ Voiceless. Ratio**	**Mean (ms) & (s. dev.)**	**Voiced/ Voiceless Ratio**
Ungem.:						
	fa`taat	6	98 (14)	0.89	133 (8)	0.62
	sa`daad	6	87 (8)		83 (4)	
	ma`Taal	6	113 (9)	0.90	128 (12)	0.60
	qa`DaaD	6	102 (18)		77 (10)	
	ma`kaan	6	115 (6)	0.96	120 (8)	0.91
	ma`gaal	6	110 (16)		109 (22)	
	ma`saas	6	134 (10)	0.92	- -	- -
	ħa`zaaz	6	123 (10)		- -	- -
	Mean : Voiced Voiceless		 105 115	 0.91	 90 127	 0.71
Gem.:						
	fat`taat	6	209 (35)	0.81	244 (15)	0.81
	sad`daad	6	169 (28)		197 (10)	
	maT`Taal	6	194 (16)	0.92	223 (4)	0.79
	ħaD`DaaD	6	178 (21)		176 (11)	
	mak`kaar	6	193 (11)	0.82	258 (18)	0.76
	ħag`gaad	6	158 (13)		195 (10)	
	ħas`saas	6	213 (10)	0.93	- -	- -
	ħaz`zaaz	6	198 (20)		- -	- -
	Mean : Voiced Voiceless		 176 202	 0.87	 189 242	 0.78

Table 3.3.5 Mean consonant durations, standard deviations and duration ratios for the voiced and voiceless consonants in geminated and ungeminated contexts.

			Subject AA			Subject WS		
Context	The vowel /aa/ in different consonant environments	N	Mean Contoid duration (ms)	Mean Vocoid duration (ms)	Contoid/ Vocoid Ratio	Mean Contoid duration (ms)	Mean Vocoid duration (ms)	Contoid/ Vocoid Ratio
Ungem.:								
	fa`taat	6	98	97	1.01	133	180	0.74
			(14)	(10)		(8)	(25)	
	sa`daad	6	87	102	0.85	83	187	0.44
			(8)	(14)		(4)	(10)	
	ma`Taal	6	113	104	1.09	128	197	0.65
			(9)	(16)		(12)	(5)	
	qa`DaaD	6	102	93	1.09	77	185	0.42
			(18)	(10)		(10)	(12)	
	ma`kaan	6	115	117	0.98	120	190	0.63
			(6)	(15)		(8)	(10)	
	ma`gaal	6	110	118	0.93	109	193	0.56
			(16)	(9)		(22)	(10)	
	ma`saas	6	134	91	1.47	-	-	-
			(10)	(13)		-	-	-
	ħa`zaaz	6	123	95	1.29	-	-	-
			(10)	(29)		-	-	-
	Mean :							
	Voiced		105	102	1.03	90	188	0.48
	Voiceless		115	102	1.13	127	186	0.68
Gemin.								
	fat`taat	6	209	93	2.24	244	184	1.33
			(35)	(12)		(15)	(34)	
	sad`daad	6	169	89	1.90	197	176	1.12
			(28)	(16)		(10)	(17)	
	maT`Taal	6	194	98	1.98	223	176	1.27
			(16)	(28)		(4)	(10)	
	ħaD`DaaD	6	178	82	2.17	176	182	0.97
			(21)	(16)		(11)	(8)	
	mak`kaar	6	193	115	1.68	258	178	1.45
			(11)	(20)		(18)	(8)	
	ħag`gaad	6	158	100	1.58	195	188	1.04
			(13)	(15)		(10)	(15)	
	ħas`saas	6	213	88	2.42	-	-	-
			(10)	(13)		-	-	-
	ħaz`zaaz	6	198	89	2.22	-	-	-
			(20)	(19)		-	-	-
	Mean :							
	Voiced		176	90	1.95	189	182	1.04
	Voiceless		202	99	2.04	242	179	1.35

Table 3.3.6 Contoid/following vocoid duration ratios and the voicing distinction in geminated and ungeminated forms.

The difference at its maximum level does not exceed 12 ms. The other observation is concerned with the subject variation where the mean vowel duration displayed by subject WS after voiced plosives is 188 ms as compared with 102 ms by subject AA.

Table 3.3.7 presents further evidence of the vocoid duration, this time in stressed and unstressed conditions. They confirm the above remarks that voicing is not a strong determinant of vowel length. It is worth mentioning, however, that the vocoid duration undergoes significant changes as a consequence of the stress shift. The presence of stress almost doubles the vocoid duration. After voiceless plosives it is around 117 ms in stressed context as compared with 53 ms in unstressed context. Again these data provide little or no evidence of the assumed relation between voicing and vocoid duration.

Word-medial plosives in unstressed context	**N**	Subj. AA **Mean vocoid duration (ms) & (s.dev.)**	Subj. WS **Mean vocoid duration (ms) & (s.dev.)**	**Word-medial plosives in stressed context**	**N**	Subj. AA **Mean vocoid duration (ms) &(s.dev.)**	Subj. WS **Mean vocoid duration (ms) & (s.dev.)**
`qatim	6	51	79	qa`tiil	6	94	130
		(2)	(7)			(11)	(13)
`qadim	6	48	90	qa`diim	6	109	151
		(9)	(9)			(8)	(29)
`xaTir	6	60	81	xa`Tiir	6	158	152
		(11)	(11)			(30)	(19)
`xaDir	6	59	95	na`Diir	6	125	155
		(10)	(12)			(8)	(20)
`rakib	6	49	77	ra`kiik	6	99	121
		(8)	(13)			(17)	(7)
`ʕagib	6	58	88	ra`giib	6	104	132
		(4)	(17)			(20)	(18)
Mean :				Mean:			
After voiced:		55	91	After voiced:		113	146
After voiceless:		53	79	After voiceless:		117	134

Table 3.3.7 Mean vocoid durations (ms.) and standard deviations in stressed and unstressed contexts.

On the other hand, there is a slight indication that emphasis may contribute to the vocoid length variation (See Figure 3.3.3). It can be seen that the vocoid duration in the environments of /T/and /D/ is longer than in the other voicing environments. In addition, the vocoid duration following different consonantal environments in the unstressed context is within a range of 48 - 60 ms whereas in the stressed context it is between 94 ms after /t/ and 158 ms after /T/.

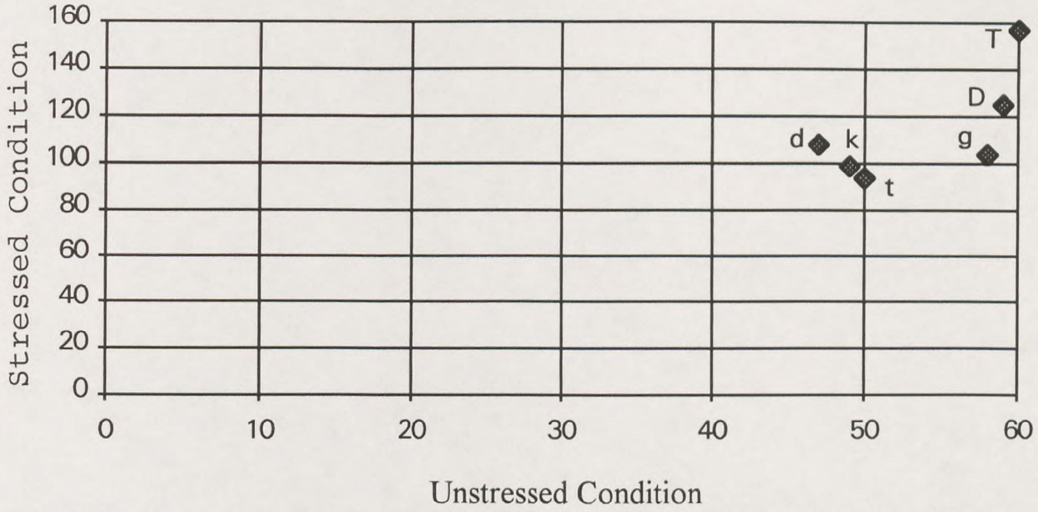


Figure 3.3.3 The duration of following vocoid in stressed and unstressed conditions. The consonant environment is also indicated.

An important aim of this study is to examine the relation between voicing and the contoid/vocoid ratio. If the data in Table 3.3.6 are re-examined particularly those concerning the contoid/vocoid ratios, one may be able to get an idea of that relation. The mean contoid/vocoid ratio for the voicing categories differs greatly. In the ungeminated contexts, for example, SW obtains a mean contoid/vocoid ratio of (0.48) for the voiced plosives, and (0.68) in the voiceless ones. In the geminated context, these mean ratios are increased to (1.04) and (1.35) ms for the two categories respectively. The mean contoid/vocoid ratio obtained by AA are quite high as compared with that of SW. This is a clear indication that the contoid/vocoid ratio does not remain constant in different contexts. The change is caused by the durational increase or reduction in one of its components (i.e. the contoid or the vocoid). For instance, the lower contoid/vocoid ratio displayed by WS may be explained by the increase in the vocoid duration without a parallel increase in the contoid duration. Similarly, the variations in the contoid/vocoid ratio in geminated and ungeminated consonants can be explained in the same terms. The considerable increase in the closure duration in geminated forms without a corresponding lengthening in the vocoid may account for this modification. To this extent it can be said that the contoid/vocoid ratio is not independent of the absolute duration shown by its components. If one or both of them could successfully separate the voicing distinction, then the contoid/vocoid ratio should reflect that and should differentiate the two cognates. If, however, both of them could not separate voicing, then the contoid/vocoid ratio will also fail to do so.

In this study, the relative success of the contoid/vocoid ratio in distinguishing the voicing pairs is partly a reflection of the differences found earlier in the contoid duration.

3.3.4.2 Emphasis and the closure duration:

It has been indicated in section 2.1.3 that voicing is a feature that exists in almost all languages with some variations determined by the characteristics of the individual languages. Emphasis, on the other hand, is a rare feature which carries a high information load in Semitic languages. This section, therefore, will aim at discovering if there is any relation between the phonological feature of emphasis and duration.

Table 3.3.8 presents data that show the mean closure durations, the standard deviations and the ratios for the emphatic and non-emphatic consonants. In both stressed and unstressed conditions, the closure duration for the emphatics is invariably longer than that of the non-emphatics. The difference is quite great in the latter context and particularly for subject AA. The closure duration for the emphatics are 42% longer than that of the non-emphatics as appears from the average ratio. This difference is reduced to 16% in the stressed context. Unlike its performance in voicing, the influence of stress on the emphatic/nonemphatic contrast is negligible and inconsistent.

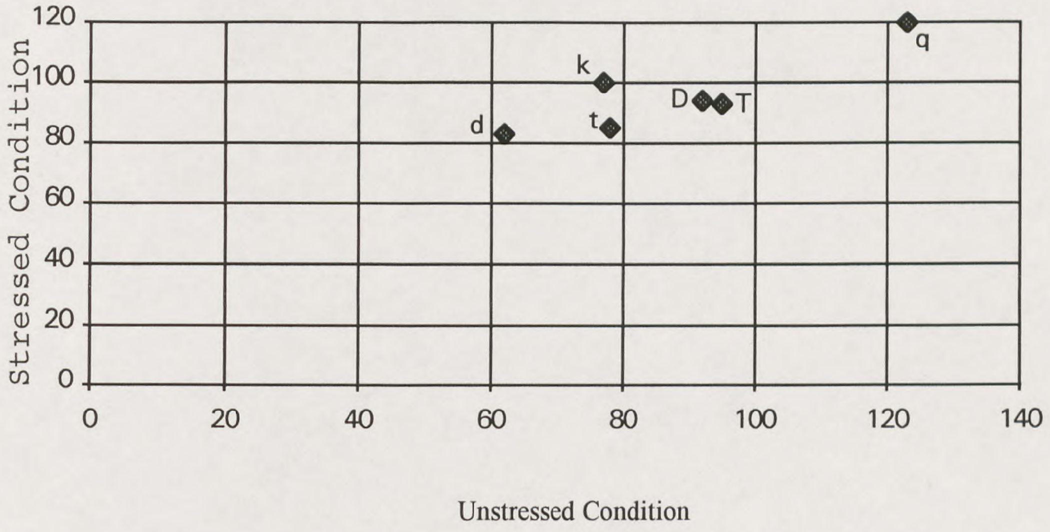


Figure 3.3.4 The relation between closure duration and emphasis in stressed and unstressed conditions.

For example, /T/, /D/ and /t/ are shown to have roughly the same closure duration values in stressed and unstressed contexts (See Figure 3.3.4), while for /q/ longer closure durations in the unstressed condition are found. Each emphatic consonant regardless of context is well separated from its nonemphatic consonant.

The results from Table 3.3.9 show similar increase in the duration of emphatics over the non-emphatics except for WS in the geminated context. In this context the mean ratio is (0.97).

			Subject AA			Subject WS	
Context	Intervocalic Emphatic & Nonemphatic consonants	N	Mean (ms) & (s.d.)	Emphatic/ Nonemphatic Ratios	N	Mean (ms) & (s.d.)	Emphatic/ Nonemphatic Ratios
Unstressed:							
	`xaTir	6	96 (14)	1.23	6	78 (7)	1.20
	`qatim	6	78 (4)		6	65 (15)	
	`xaDir	6	93 (16)	1.47	6	58 (4)	0.96
	`qadim	6	63 (5)		6	60 (0)	
	`ʕaqib	6	123 (12)	1.58	6	88 (8)	1.21
	`rakib	6	78 (16)		6	73 (8)	
	Mean :						
	Emphatics		104	1.42		74	1.12
	Nonemphatics		73			66	
Stressed:							
	xa`Tiir	6	93 (14)	1.08	6	92 (7)	1.23
	qa`tiil	6	86 (8)		6	75 (8)	
	na`Diir	6	94 (11)	1.13	6	73 (15)	1.04
	qa`diim	6	83 (9)		6	70 (8)	
	ra`qiib	6	121 (10)	1.21	6	103 (10)	1.10
	ra`kiik	6	100 (6)		6	93 (12)	
	Mean :						
	Emphatics		103	1.16		89	1.13
	Nonemphatics		89			79	

Table 3.3.8 Closure durations, standard deviations and duration ratios for the emphatics and non-emphatics in stressed and unstressed contexts.

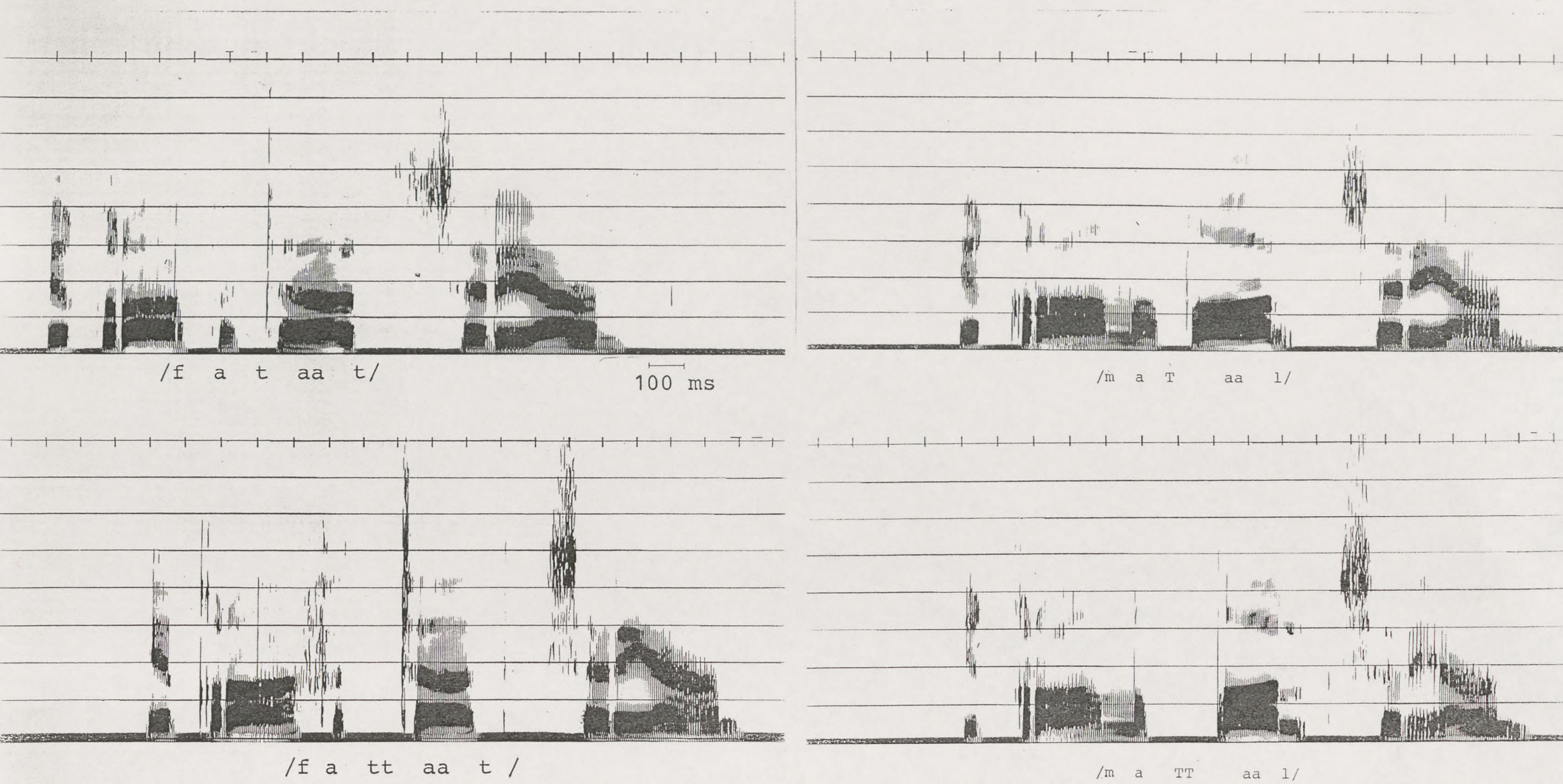
It indicates that the closure duration for the emphatic/nonemphatic contrasts is roughly the same. In view of the clear-cut difference observed in other contexts, it is not immediately clear to us why such a contradiction does happen to such a context. One possibility for that action may be the interaction between emphasis and gemination which appears to reduce the difference to a minimal level.

			Subject AA			Subject WS	
Context	**Intervocalic Emphatic & Nonemphatic consonants**	**N**	**Mean (ms) & (s.d.)**	**Emphatic/ Nonemphatic Ratio**	**N**	**Mean (ms) & (s.d.)**	**Emphatic/ Nonemphatic Ratio**
Ungeminated:							
	ma`Taal	6	91 (8)	1.36	6	106 (14)	1.25
	fa`taat	6	67 (10)		6	85 (6)	
	qa`DaaD	6	93 (10)	1.37	6	63 (8)	0.93
	sa`daad	6	68 (8)		6	68 (8)	
	ma`qaal	6	102 (15)	1.42	6	100 (0)	1.11
	ma`kaan	6	72 (9)		6	90 (9)	
	Mean :						
	Emphatics		95	1.38		90	1.11
	Nonemphatics		69			81	
Geminated:							
	maT`Taal	6	173 (17)	1.02	6	203 (5)	1.00
	fat`taat	6	169 (10)		6	202 (19)	
	ħaD`DaaD	6	158 (23)	1.08	6	156 (11)	0.90
	sad`daad	6	146 (15)		6	173 (5)	
	ħaq`qaad	6	154 (9)	1.02	6	194 (26)	1.01
	mak`kaar	6	150 (10)		6	192 (10)	
	Mean :						
	Emphatics		162	1.05		184	0.97
	Nonemphatics		155			189	

Table 3.3.9 Mean closure durations, standard deviations and duration ratios for the emphatics and non-emphatics in geminated and ungeminated contexts.

A further feature of this data is that /T/ is shown to have the longest closure interval, whereas the shortest one is seen for /d/. Spectrograms illustrating the emphatic/nonemphatic contrast /T/ vs /t/ in single and geminated words are shown in Figure 3.3.5.

Figure 3.3.5 Wideband spectrograms illustrating the closure duration for the word-medial emphatics and non-emphatics in geminated and single forms.



3.3.4.3 Place of articulation and the closure duration:

The last variable examined in connection with duration is the place of articulation. Closure duration measurements for the voiced consonants /b/, /d/, and /g/ are done on the assumption that they reflect three distinct places of articulation: bilabial, denti-alveolar and velar (See Table 3.3.10).

Context	**Place of Articulation Test words**	**N**	Subj. AA **Mean (ms) & (s.dev.)**	Subj. WS **Mean (ms) & (s.dev.)**	**Context**	**Place of Articulation Test words**	**N**	Subj. AA **Mean (ms) & (s.dev.)**	Subj. WS **Mean (ms) & (s.dev.)**
Unstressed:					Ungeminated:				
	\`ʕabiq	6	71	55		ʃa\`baab	6	78	78
			(9)	(8)				(13)	(15)
	\`qadim	6	63	60		sa\`daad	6	68	68
			(5)	(0)				(8)	(8)
	\`ʕagib	6	79	60		ma\`gaal	6	79	84
			(5)	(9)				(13)	(11)
Stressed:					**Geminated:**				
	ʕa\`biiq	6	87	62		ʃab\`baab	6	154	175
			(17)	(15)				(17)	(6)
	qa\`diim	6	83	70		sad\`daad	6	146	173
			(9)	(8)				(15)	(5)
	ra\`giib	6	89	68		ħag\`gaad	6	122	159
			(13)	(13)				(12)	(3)

Table 3.3.10 Mean closure duration and place of articulation in different contexts.

The results prove that the order effect of place varies according to the context in which they are found and the subject who produces them. In most contexts subject AA, for example, displays the order /g/ > /b/ > /d/ except in the geminated context where the order is changed to /b/ > /d/ > /g/. The difference in durations among the three places are quite small. Subject WS, on the other hand, does not show any consistent pattern for place, at least in the stressed and unstressed contexts. These findings emphasize the believe that the closure duration may not be an important indicator of the place of articulation.

3.3.5 Discussion

Of the three durational intervals examined in this study, only the closure and the contoid durations prove to be capable of separating the voicing and the emphatic categories with different levels of efficiency depending on which of these features is involved and their interaction with other contextual factors. Thus, the closure duration performed poorly with voicing. The difference for the voicing cognates is not large enough nor is it consistent in various contexts. In one of the contexts examined the mean closure durations reported are 78 and 84 ms respectively. Although durational differences as small as 10 ms are perceptable (Huggins,1972), a difference such as the one displayed by the voicing contrasts in YSA is not likely to be linguistically significant. These results are in agreement with those of Wajskop (1977), Umeda (1977), Crystal and House (1982, 1988), who find very small differences in closure duration as function of voicing. On the other hand, the findings do not support the generalization which claims that the closure duration is correlated with voicing (Lisker, 1957; Chen, 1970). There seems to be, at least, two reasons for this contradiction. The first may be related to the amount of context available and whether the words containing the voicing contrasts are recorded in isolation or in a frame sentence. Thus, in citation forms, (e.g. Chen, 1970) the mean closure durations presented for the two categories in English are quite high (88 ms and 140 ms for the voiced and voiceless plosives respectively). On the other hand, those studies which employed a frame sentence (e.g. this study or that by Wajskop, 1977) are associated with lower closure duration values for the two voicing categories. Similar results are obtained for data in connected speech (e.g. Luce and Charles-Luce,1985). The mean duration values reported for the voiced and voiceless categories are 68 ms and 86 ms respectively.

The second possible reason for the relative failure of the closure duration to yield large differences between the voicing categories in YSA may lie in Lisker's comment (1972) that the voiced and voiceless bilabials are not found to differ significantly "Except when preceded by a stressed syllabic and followed by an unstressed one". A re-examination of the material in Table 3.3.1 shows that the intervocalic plosives are not manipulated in accordance with the requirement mentioned by Lisker (1972). It is not surprising, therefore, that they do not display linguistically relevant differences.

In addition, variability in the closure duration is affected by one or a combination of other factors (e.g. emphasis, gemination, stress and the place of articulation). While the relation between the closure duration and voicing is not evident, its relation with emphasis is quite strong.

On the other hand, for certain consonants durational modifications in geminated and single forms are observed. The articulatory closure time for the geminated consonants is greatly lengthened in comparison with that in single forms. For instance, the mean acoustic closure duration values found for the single and geminate bilabial plosives in YSA are 78 ms and 158 ms. This finding is consistent with the perceptual boundary determined for the same contrast by Obrecht (1965). In view of the large differences observed for this contrast it seems reasonable to suppose that the closure duration is likely to be an important acoustic correlate for the perception of gemination and perhaps a primary one. The other related issue concerns the durational relationship governing the contoid and vocoid. The increase in the closure duration for the geminate consonants is also accompanied by a shortening in the duration of the following vocoid. Although at first impression one may be inclined to describe it as a compensatory process, there are certain points which need to be kept in mind. The compensation is not complete in the sense that the closure and vowel durations do not increase and decrease respectively by the same amount of duration. The increase in closure duration for the geminate consonant is much greater than the decrease in the vowel duration. The hypothesis of compensatory temporal adjustment (MacNeilage, 1968) requires the syllable units to have equal duration. If one segment is lengthened or shortened, then that will be compensated for by a neighbouring one. Clearly this requirement is not fulfulled by our data. On the contrary, they may argue against it. Some other investigators (e.g. Homma, 1981) maintain that the compensatory process is not only confined to the syllable units, but also it operates over whole words. Since our measurements are only made for particular segments rather than whole words, this claim cannot be rejected or accepted until it is tested on other appropriate material.

It is also interesting to note that the order effect of the place of articulation is to some extent symmetrical. Although the variations are not large its relative consistency in some contexts suggests that duration may be used as a cue for the consonant place of articulation. It follows that any interpretation of the durational effect should take into account the relative mobility of the associated articulators. Since the speed of the articulators (the lips, the back of the tongue and the tip of the tongue) has been suggested as playing an important role in determinig duration (Lehiste, 1970) it may be expected that a shorter closure duartion is required by the more mobile articulators to reach their articulatory target positions. This may explain why /g/ and /b/ are in most cases associated with longer closure durations than /d/. No simple explanation can be provided for /d/ showing longer closure duration than /g/ in geminated context. This last remark highlights the fact that the order effect of place on duration is more complex and that its interaction with other factors such as voicing and/or emphasis may have its contribution.

The strongest indicator of voicing in this study is the contoid duration. The inclusion of the release portion as part of the contoid greatly enhances the distinction. This conclusion is also supported by evidence from aerodynamic data which will be presented later in section 5.1 of this thesis, and by data from Zue (1976). Both studies show the release duration of the plosive consonants to vary greatly as function of voicing.

The contoid duration is also efficient in identifying the consonant manner of articulation for fricatives. The single and geminate fricatives are characterized by longer duration than that of the plosives.

With regard to the predicted connection between the contoid/vocoid ratio and voicing our results show that this ratio changes from one context to another. This may be due to the unequal increase or reduction in the duration of both of its components: the contoid and vocoid durations. It does not take account of subject variation. In some instances it maintains an overlap between the voicing contrasts. These findings are in agreement with Massaro and Cohen (1983); Luce and Charles Luce (1985) and Crystal and House (1988). All agree that the contoid/vocoid ratio is an unstable and redundant measure of voicing.

The third measure, vocoid duration, does not yield reliable differences between any of the opposing categories. Contrary to the findings of many studies (Chen, 1970; Luce and Charles Luce (1985), 1985; House and Fairbanks, 1953 among others) our results show no systematic relation between following vocoid length and voicing. One possible explanation for this discrepency may be attributed to the position of the word in which the vowel is found. Indeed, available evidence (e.g. Klatt, 1976) suggests that the effect of voicing on vowel lengthening is confined to the phrase-final position. He states that "A large difference in vowel duration is only seen in phrase-final environments, so it is only in these cases that the durational cue has primary importance" p. 1219. Thus, the lack of variation in our results may be, in part, due to the fact that the words containing the vowels are not in phrase-final position.

Another reason for that contradiction may be due to the differences in measurement conventions. In our study the vowel is measured from the vocoid onset after both voiced and voiceless consonants. On the other hand, in some studies (e.g. Peterson and Lehiste, 1960 and Klatt, 1976) the vowel is measured from the onset of voicing for /ptk/ and from the start of release for /bdg/. It is, therefore, likely that the reported increase in vocoid duration after /bdg/ is caused by the inclusion of the short release interval as part of the vowel. Had the vocoid been measured from the onset of voicing for both voiced and voiceless plosives, the vocoid length variability may have disappeared in their study.

Although stress appears to have contributed to the voicing distinction its primary influence is mainly on the vowel duration. The lengthening of vocoids in stressed as compared to unstressed syllables is observed to occur regardless of voicing. This evidence lends further support for Fry's (1958) and Fant's et. al. (1991)conclusion that stressed and unstressed syllables are correlated with longer and shorter vocoid durations respectively.

To conclude: the results of this investigation have shown that the three durational intervals in YSA change with changes in phonological features that serve to distinguish words. Systematic variation in closure duration, the whole consonant duration and vowel duration depends on linguistic factors such as voicing, emphasis, gemination, stress and place of articulation. While the closure duration is less than satisfactory in distinguishing voicing, the consonant duration is a good indicator of that contrast. With emphasis and gemination, however, the closure is the primary interval separating the categories of those features. This study also fails to find any relationship between vocoid length and voicing . Most of the vocoid length variability in this investigation is attributed to stress. Duration variation is also determined by the segment type (i.e. whether it is plosive or fricative, consonant or vowel). In addition, various characteristics of the data suggest that intersubject differences may be greater than the differences caused by some contextual factors.

Chapter 4: Perceptual characteristics of voicing and emphasis in YSA

4.1 Perceptual Study of voicing in Yemeni Spoken Arabic

4.1.1 Introduction and Aims:

This study is motivated by the following reasons. As it appears from the review in section 2.2.1 and 2.2.2 most studies seem to focus on English and only a few on other languages. Cross language studies on speech research show that if the acoustic properties of speech are to be related to the theory of distinctive features they must be examined on the acoustic, perceptual and articulatory levels not only in English but also in a wide range of languages (Lisker and Abramson, 1964; Abramson and Lisker 1965, 1970). To our knowledge there appears to be no systematic and detailed perceptual study based on synthetic speech of VOT in YSA or Modern Standard Arabic. This shows clearly how important it is to start examining these features in other languages. Therefore, this study will be one step further in that direction.

Another aim is to compare the results of the acoustic analysis of VOT in section 3.1 with those obtained in this study to discover how close the match between the two results is. This may shed some light on the nature of the relationships between perception and production.

It may also be interesting to find where the individual listeners assign the phoneme boundaries and whether it (i.e. the 50% crossover point) is fixed or that it varies from one listener to another. If it is not fixed, then it might be important to find the range of this shift on the VOT dimension.

It is shown that categorical perception is a "language specific" phenomenon. In the cross-language study of voicing by Abramson and Lisker (1968), they look at the effect of the systematic variations in VOT on the plosives' perception by native listeners of English and Thai. The two languages vary in the numbers of phonological categories with regard to voicing. Their study shows that the subjects are sensitive to the language specific characteristics in that the phoneme boundaries are different in the two languages. The language experience of the listener will have its effect on the perception of the category boundary. In addition, categorical perception is said to be an acquired one (Liberman et. al., 1957). This implies that voicing categorization can be improved by training, particularly when learning foreign languages. For this reason, the study of various perceptual cues for a number of acoustic patterns, and not only VOT, in different languages will be invaluable in the specification of phonological contrasts. On the basis of that, a training scheme of how to acquire a certain voicing contrast may be developed.

4.1.2 The instrumental set-up:

Prior to synthesis, recording of natural CV-syllables of a YSA subject was done. Wideband and narrowband spectrograms were made on the 700 spectrograph and were used to derive estimated values for the formant frequencies, segment durations, the fundamental frequency and the intensity level of the sound signal. These values formed the basis on which the synthesized stimuli were generated. Further analyses of the natural CV-syllables were performed by the ILS package on the Masscomp computer (5500) partly to serve as a check up and partly to get more details about certain acoustic features, e.g. fundamental frequency (F0); amplitude of voicing (av); formant amplitudes and bandwidths as they changed with time.

The speech stimuli were synthesized at the Department of Psychology, University of Leeds using a cascade-parallel formant synthesizer developed by Dennis Klatt. It contains 47 parameters with two voicing sources: natural and impulse train. It has also an aspiration source which, together with the voicing source, is connected to the cascade branch. The frication noise source is connected to the parallel branch. More details about the synthesizer can be seen in Klatt's paper, (1980).

The prepared values for the stimuli's parameters as they changed across time were entered into the synthesizer. They were then digitized by the Masscomp computer and output via a D to A conventer. The analog signal was low-pass filtered at 5.9 KHz, with an anti-aliasing hardware filter (elliptical, designed by Celia Scully and Eric Brearley in the Department of Linguistics and Phonetics). The output signal was played back using a Ferrograph tape recorder unit together with a headset. Figure 4.1.1 shows schematically the order of this process.

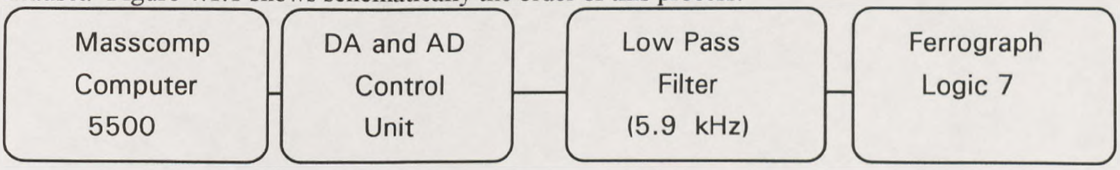


Figure 4.1.1 Block diagram of the instrumental set-up used in the perceptual experiments of voicing and emphasis in the plosives of YSA .

4.1.3 Stimuli:

After comparing the advantages and disadvantages of both synthetic and natural speech we prefer to choose the former. Although synthetic speech lacks the richness of natural speech its usefulness is that it allows the experimenter to investigate and manipulate a particular variable in the speech signal. It also helps in achieving a higher degree of precision and control of variables during the creation of sounds. It simplifies the acoustic pattern of actual speech by deleting, isolating or neutralising certain other features, so that listeners' judgements are correlated with only the systematic changes made to the variable investigated in the synthetic speech signal.

Natural speech is normally characterized by many variables which need to be controlled. This task may prove to be rather difficult particularly if the variable manipulated is the second formant. In fact it is impossible to isolate or manipulate such a variable in natural speech.

The synthesis strategies used in this study were as follows: Each synthetic stimulus was 405 msec in duration. The fundamental frequency of voicing (F0) started at different frequency points for different continua. For the first continuum it started at 165 Hz and remained so till 100 msec, then it fell linearly to 100 Hz at 405 msec. For continuum number two it started at 150 Hz and fell to 100 Hz at the end of the stimuli. The amplitude of voicing (av) was set to a constant value of 60 dB where needed from the start till 350 msec, then it fell gradually to zero at 400 msec. Where the voicing was not needed it was set to zero. The spectral tilt of voicing (tl) contributed to the generation of the voicebar for the voiced sounds /daa/ / and /Daa/. It was set to a value of 28 dB where the prevoicing was needed and to zero where it was not. A large spectral tilt of 28 dB was suitable for prevoicing (i.e.voice during the closure) whereas a value of zero was given where high frequencies showed up a lot. The plosive release consisted of a burst and frication noise. The amplitude of frication (af) and the burst amplitude (ab) were the two parameters used to generate it. They were both given values of 60 dB at the time of the release burst and a zero value otherwise. The first formant started with low but constant frequency value for all stimuli. At the plosive release and the start of the vocoid it rose abruptly to a steady state of 650 Hz. The course of the next two formants (F2 and F3) differed from one continuum to another. The contrasted pairs were given appropriate values depending on whether they were denti-alveolar emphatics, or denti-alveolar nonemphatics. The onset frequencies for F2 and F3 in the first continuum were 1800 Hz and 3100 Hz respectively and they fell to steady states of

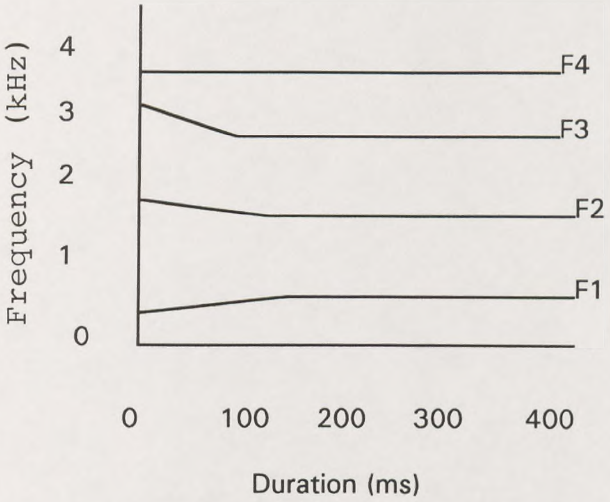


Figure 4.1.2 A plot of the first three formants (their onset frequencies and steady states) in the pattern used for the study of VOT in the /daa-taa/ continuum. The ordinate shows frequency in kHz.

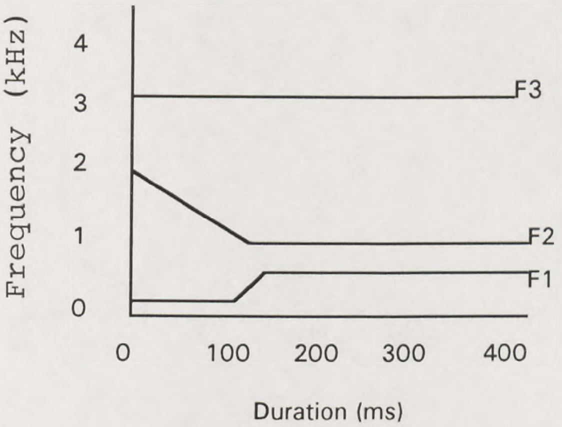


Figure 4.1.3 A plot of the first three formants showing the onset frequencies and the steady states in the pattern used for synthesizing the VOT Continuum /Daa-Taa/. The ordinate shows frequency in kHz.

1650 Hz and 2833 Hz respectively. For continuum two these same formants had onset frequencies of 1600 Hz and 2500 Hz repectively. The F2 fell to a steady state of 900 Hz whereas the F3 was kept unchanged. F4 and F5 if present at all were kept constant at the default values for all stimuli. The same thing was done with the formants' bandwidths (See the Klatt parameter values in Appendices 3 and 4). The first five formant amplitudes were kept constant at 70, 55, 50, 40 and 30 dB respectively for all stimuli in this test (i. e. 2) and 60, 50, 40, and 30 in test one. Schematic diagrams showing the course of the first three formants are presented in Figures 4.1.2 and 4.1.3 above.

4.1.4 Subjects and procedure

The subjects who participated in both tests were eight native speakers of YSA. Seven of them were males and one was female, HD. They all reported having normal hearing with no obvious auditory illness or hearing disorders. All of them are living in Yemen at the time of writing this study except AA. Subjects MBA and RA who participated in test one decided not to continue in test two. Those were replaced for test two by two additional male speakers.

The first continuum contained thirteen stimuli with VOT varied in 20 msec steps from -100 msec to 140 msec. The second one was varied in 20 msec or 10 msec steps depending on whether it was at peripheries or near the phoneme boundaries. The range was between -70 msec and +80 msec giving a total of 15 VOT variant stimuli. The range of each of these continua encompassed the phoneme boundaries established for each contrasted pair in our acoustic analysis of VOT. Table 4.1.1 below presents the number of stimuli in each test and the designated VOT values.

Wideband spectrograms of the synthesized stimuli were made and their VOT values were measured. No differences between the calculated and the measured VOT values were observed. Wideband spectrograms of the natural and synthetic CV-syllables can be seen in Figures 4.1.4 (a-b) and 4.1.5 (a-b). It is also important to add that all stimuli within each continuum were similar in all respects (e.g. durations, burst amplitudes, formant frequencies etc.) except for VOT. The synthesized stimuli were then recorded in ten randomised orders using a Ferrograph logic 7 tape recorder. The line input and master record of the unit were at six and a half and seven respectively during the recording of

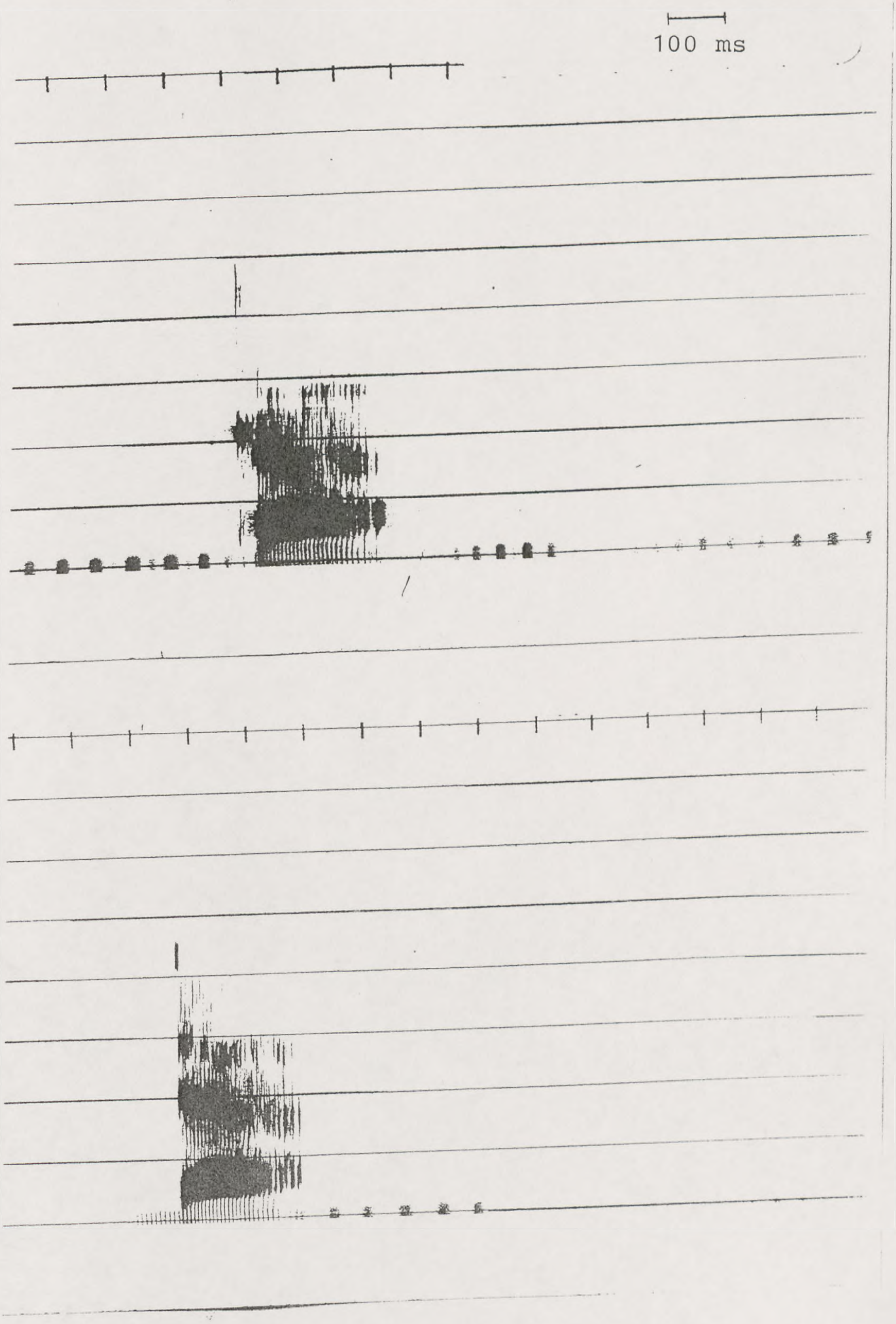


Figure 4.1.4a Wideband spectrograms of natural [taa] above and [daa] below as produced by a YSA speaker, AA.

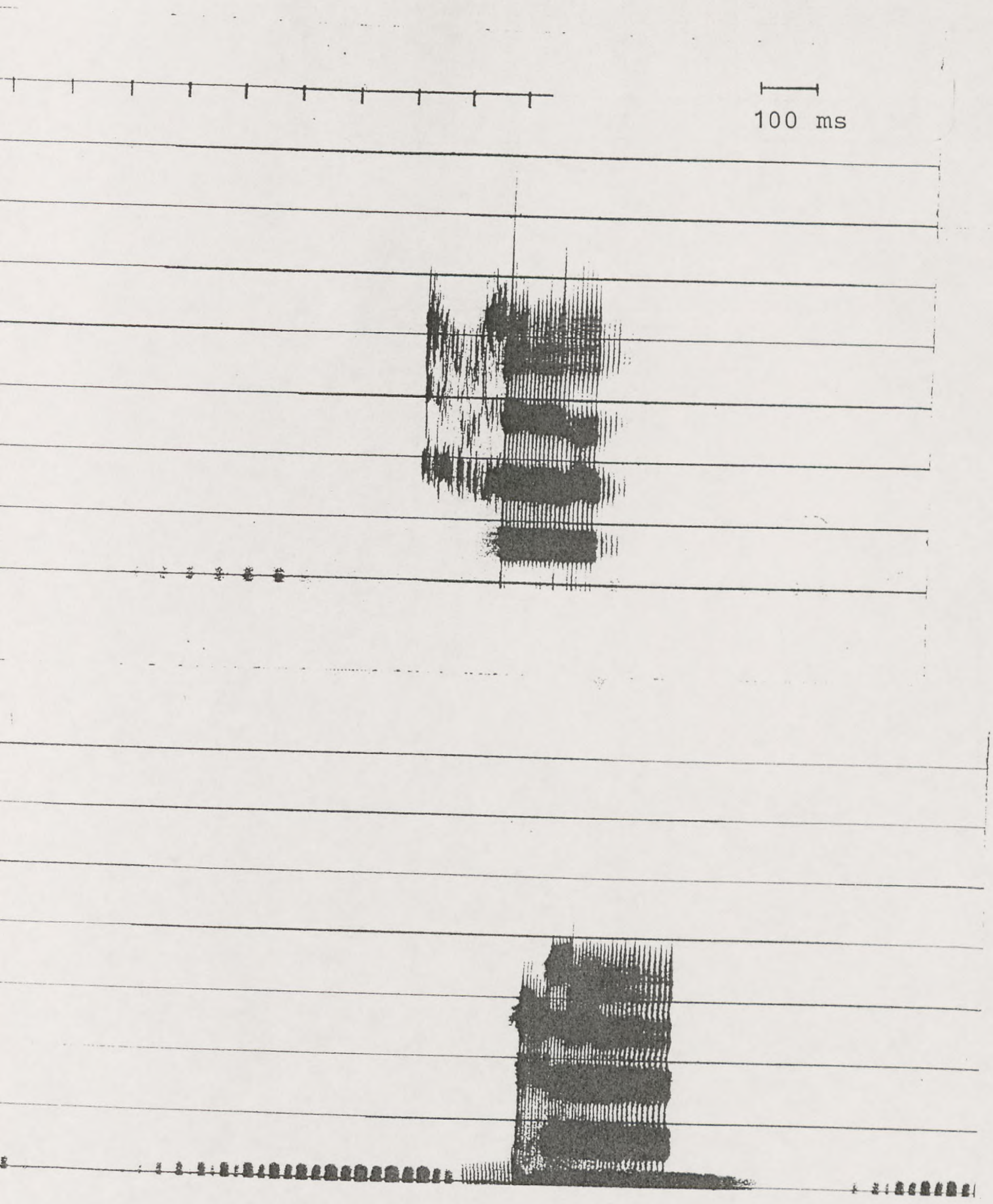


Figure 4.1.4b Wideband spectrograms of the endpoints for the synthetic continuum presented for the identification of the categories in test one: [taa] vs [daa].

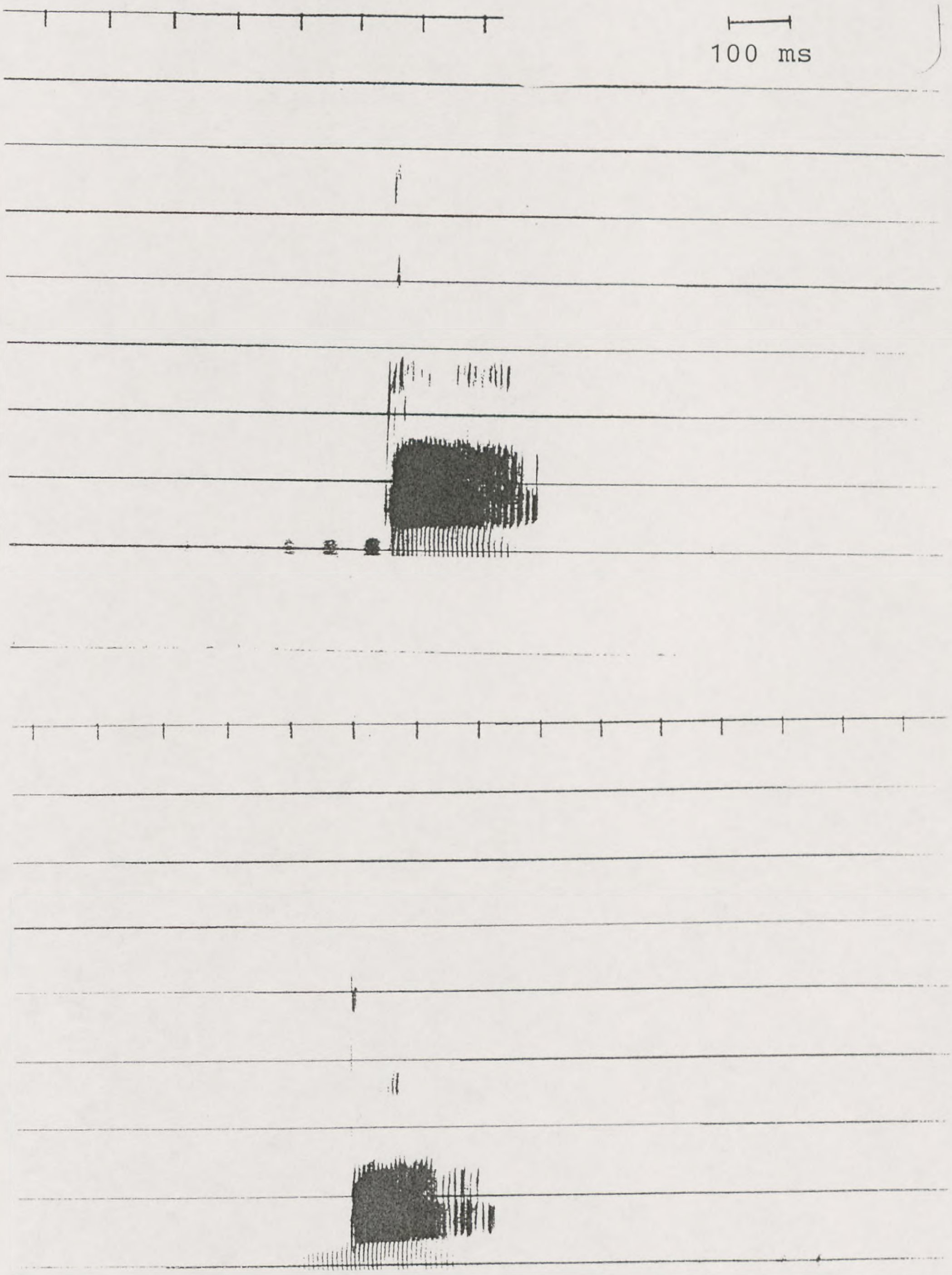


Figure 4.1.5a Wideband spectrograms of natural [Taa] above and [Daa] below as produced by a YSA speaker, AA.

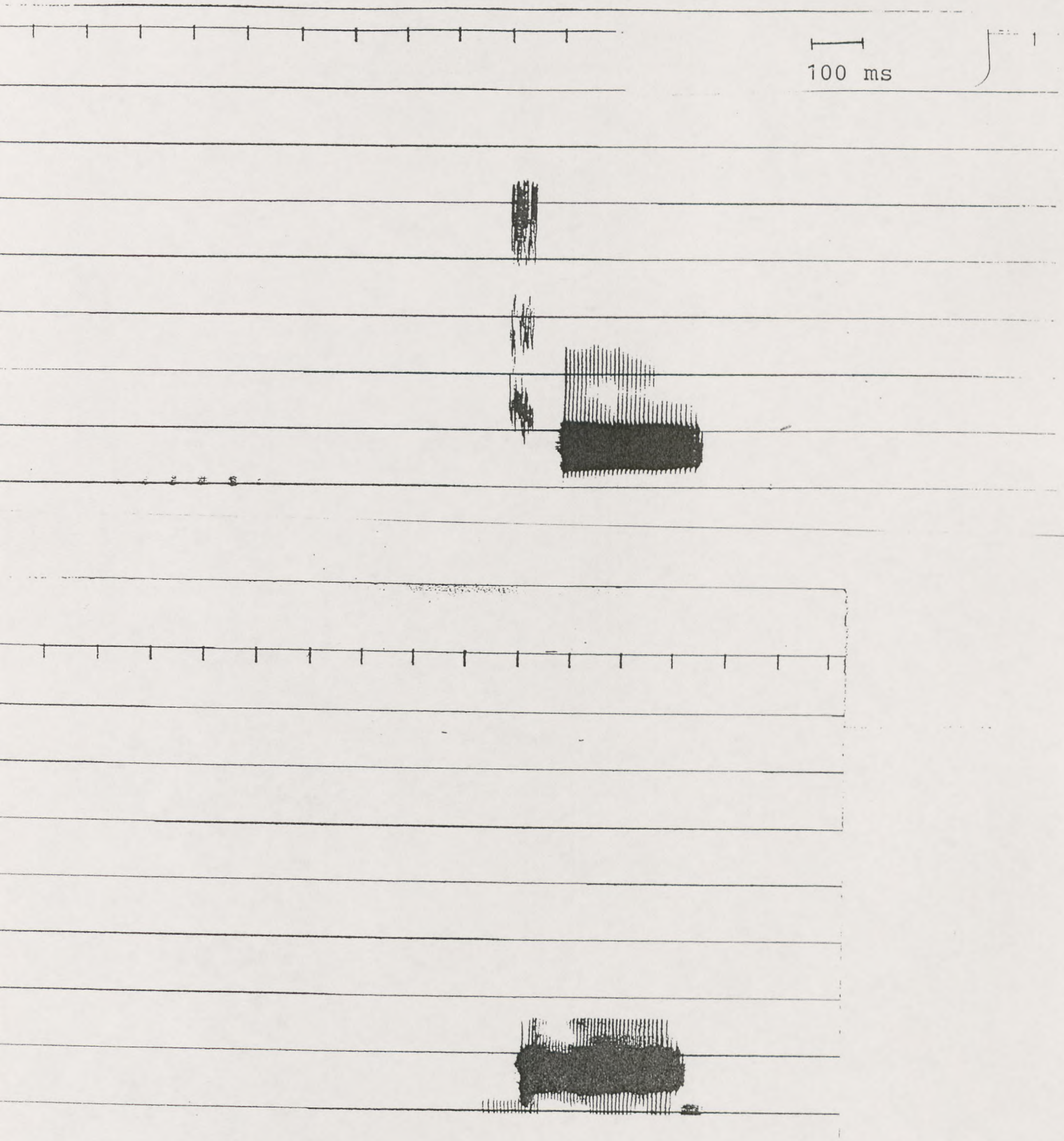


Figure 4.1.5b Wideband spectrograms of the endpoints for the synthetic continuum presented for the identification of the voicing categories in test two: [Taa] vs [Daa].

the two continua. The randomization process was automatically done using a computer programme called "mtape". It was one of those programs designed by Klatt to be utilised by the synthesizer.

Stimulus No.	**Determined VOT (ms) in Cont. one**	**Determined VOT (ms) in Cont. two**
1	-100	-70
2	-80	-60
3	-60	-50
4	-40	-40
5	-20	-30
6	0	-20
7	20	-10
8	40	0
9	60	10
10	80	20
11	100	30
12	120	40
13	140	50
14	-	60
15	-	80

Table 4.1.1 The designated VOT values used in the preparation of the two tests.

It produced an answer sheet for the randomised stimuli as well as a response sheet for the listeners to write in their responses. There was a two and half seconds interval between stimuli. Ten extra stimuli were added to each continuum in addition to the first and last two stimuli. It was hoped that the first five stimuli would be used for familiarisation. At the end the listeners might be distracted or might start to get tired. These extra stimuli were intended to minimise this problem. The stimuli were recorded onto an Ampex reel tape and were copied into a cassette tape for playing to the subjects.

The subjects were provided with response sheets and were asked to write their responses as they listened to the tape. It was a forced choice task between /taa/ or /daa/ in continuum one, and /Taa/ or /Daa/ in continuum two. In addition, they were instructed to write the perceived stimuli as soon as they heard them and not to make any changes after the listening test had finished.

4.1.5 Results:

4.1.5.1 Test One:

The subjects judgements for test one are displayed graphically in Figures 4.1.6a and 4.1.6b. The line graphs show the percentage of voiced identification for 13 stimuli in test one. These percentages are related to varying VOT values along the abscissa. The plosive voicing boundary (i.e. the 50% crossover point) is generally observed to occur at the voicing lead as shown in Table 4.1.2. The 50% crossover point is defined as that point along the VOT dimension where each of the voicing categories is heard 50 % of the time. Here subjects cannot decide whether the perceived stimuli belong to the voiced or the voiceless category. Its range is between -15 ms and 14 ms.

Subject	**The 50 % Crossover Point (msec)**
AA	14
MB	-4
HD	0
RA	-12
SG	-12
FN	-8
MBA	-15
AD	-5
MEAN	-5.25
S. DEV.	9.21
RANGE	-15 : 14

Table 4.1.2 The mean, standard deviation and the range (ms) of the perceived 50 % crossover points by the subjects in the /daa/ vs / taa/ continuum.

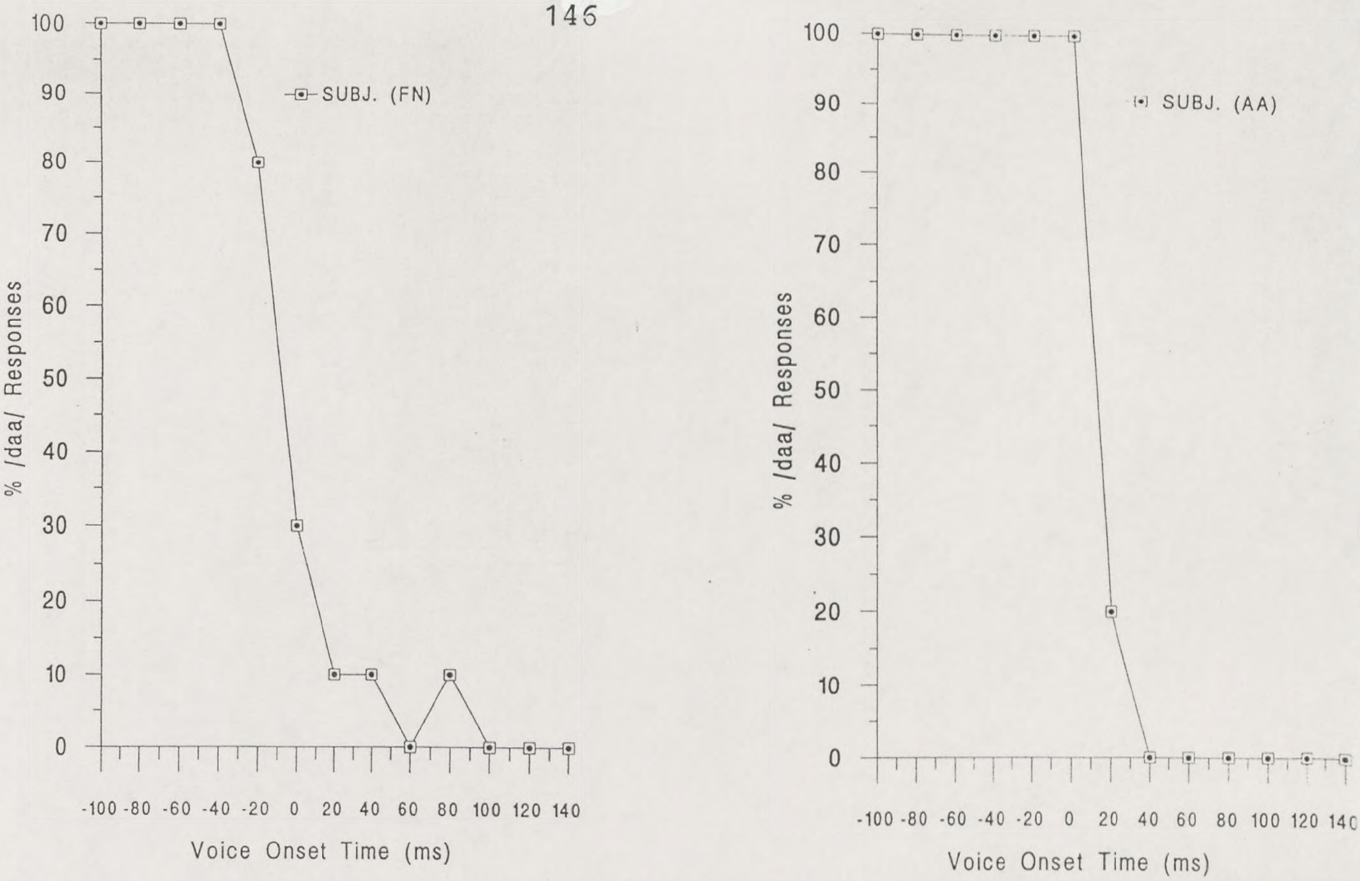
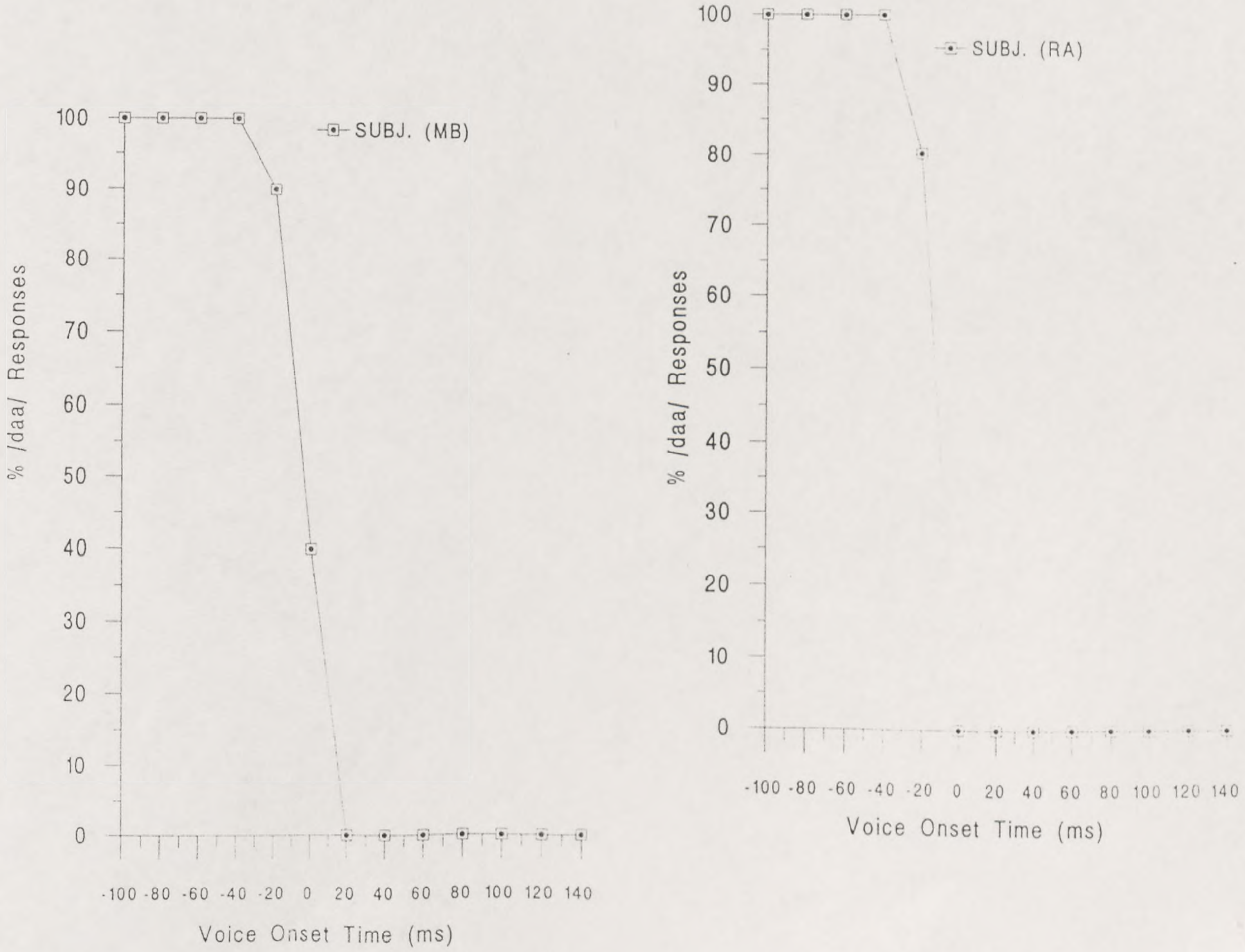
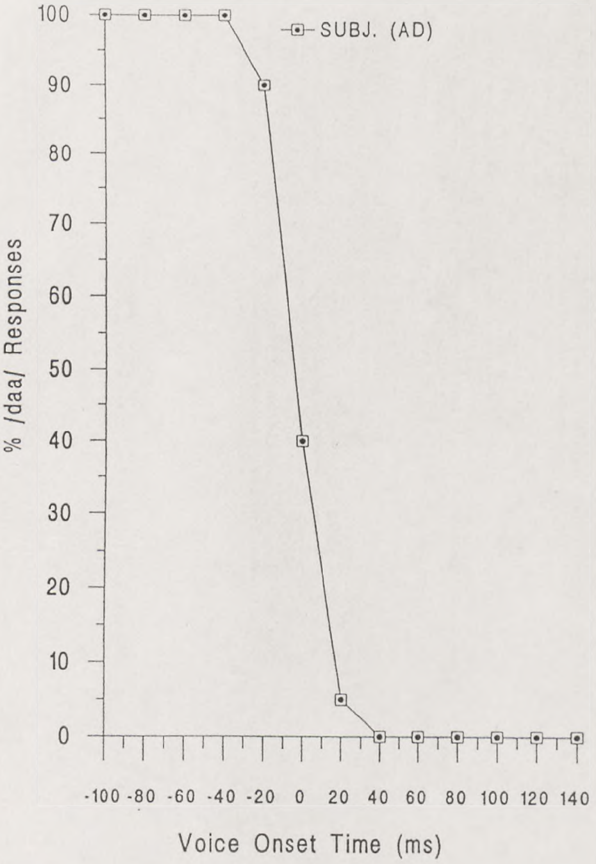


Figure 4.1.6a Percent of [daa] responses by each of the eight subjects for test one stimuli.





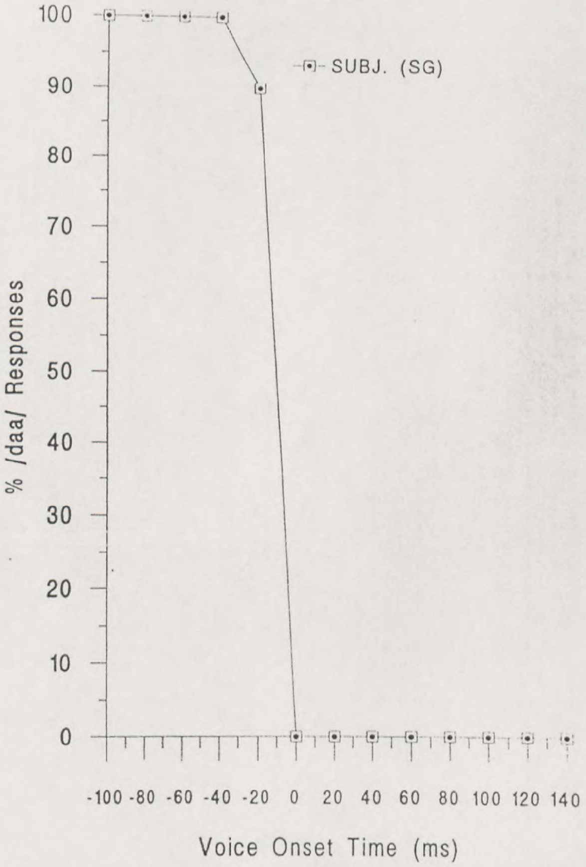
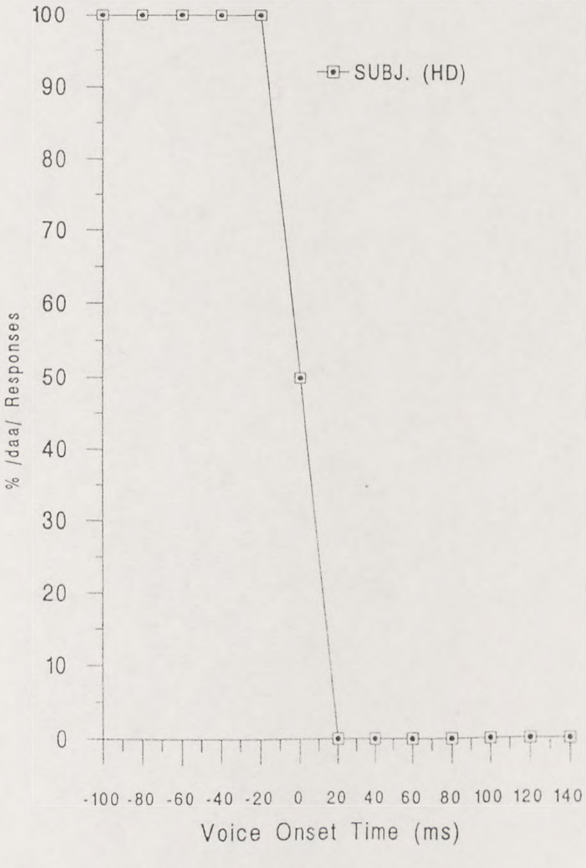


Figure 4.1.6a (Continued).



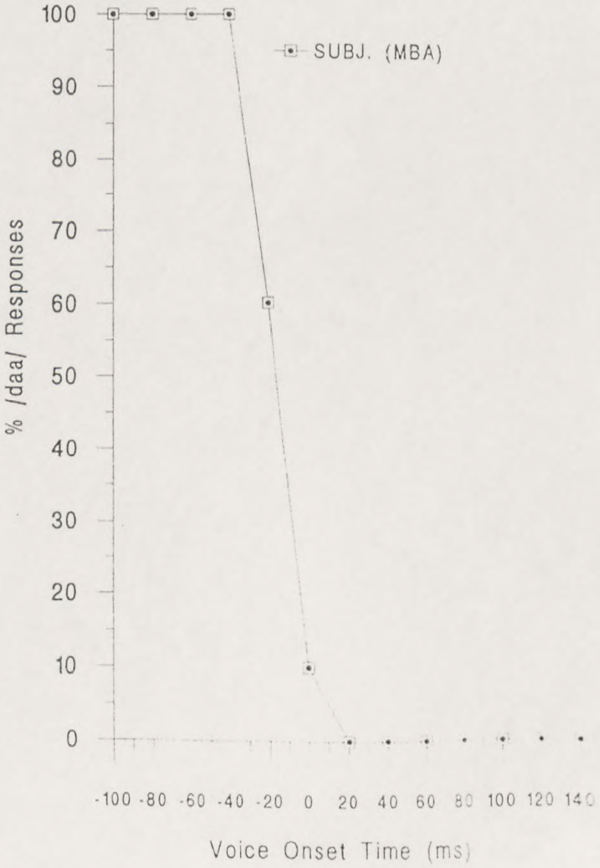
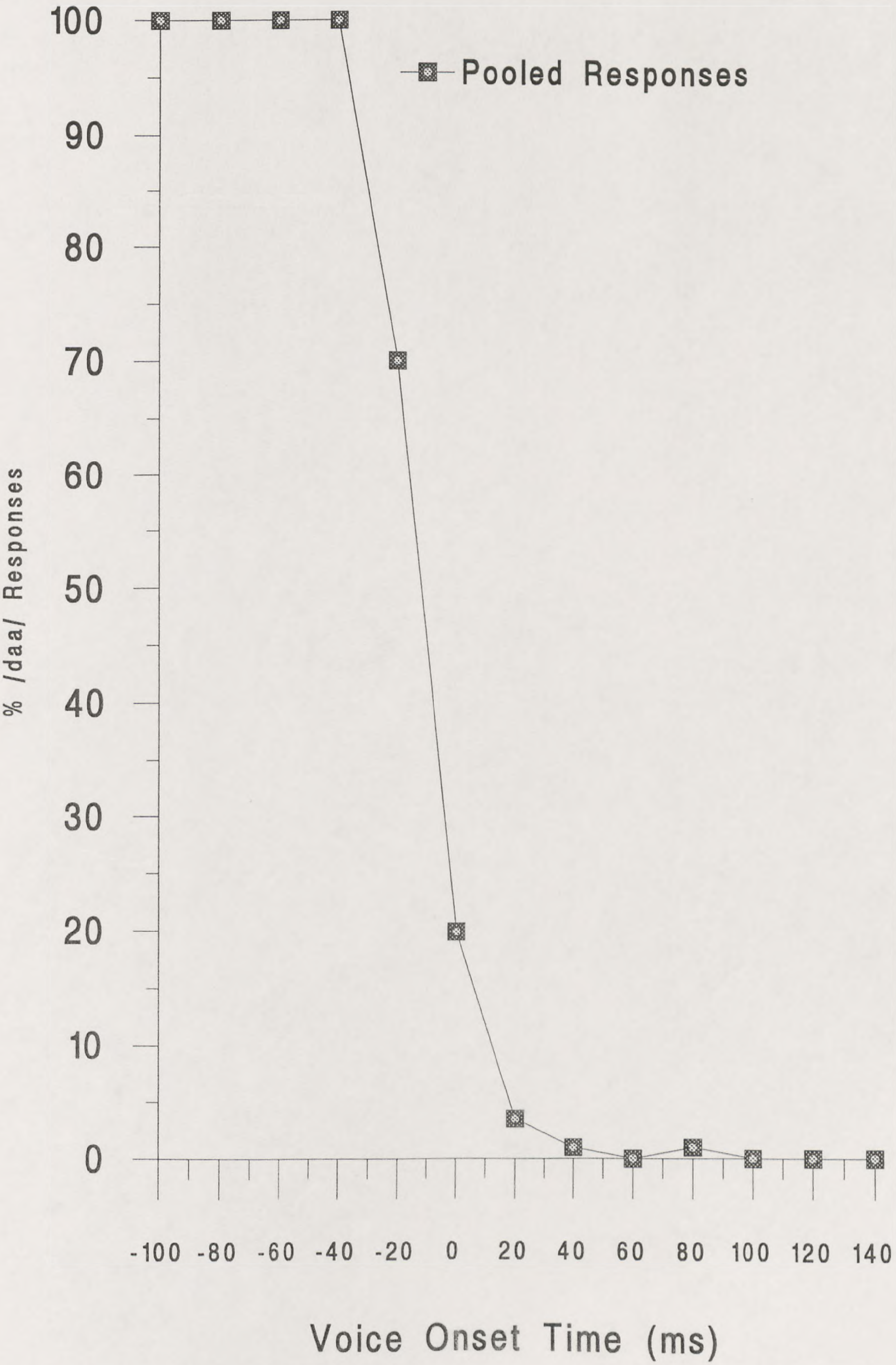


Figure 4.1.6b Pooled responses to test one stimuli by the 8 subjects.



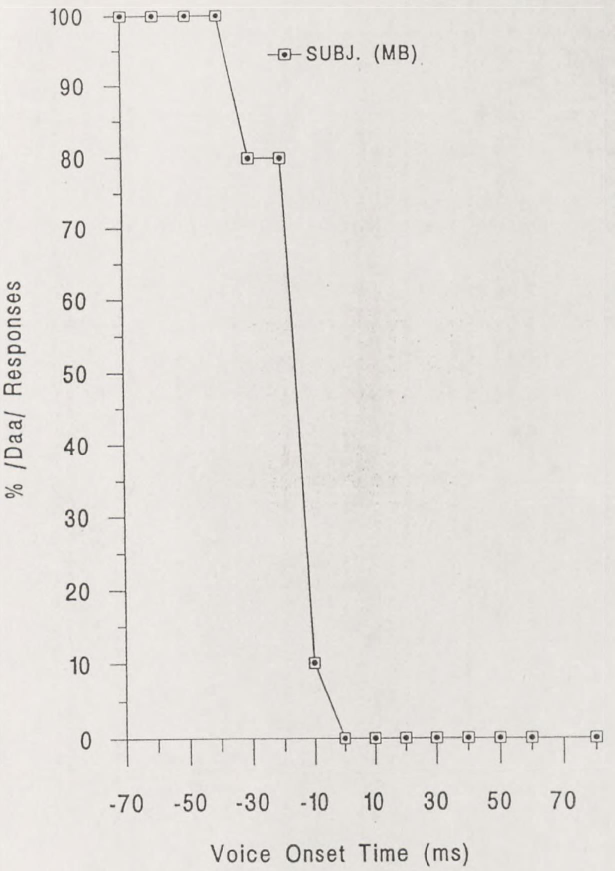
The only subject who assigns the VOT category boundary at the voicing lag is subject AA.

On average the 50 % crossover point occurs at - 5.25ms. Although the individual subjects slightly vary in assigning the category boundary, their responses suggest that they are sensitive to variations in the VOT values and that they are not responding randomly.

4.1.5.2 Test Two:

In test two, as can be seen from Figures 4.1.7a, 4.1.7b and Table 4.1.3 the perceived stimuli are separated into two distinct voicing categories: / Daa / and / Taa /. Like that of test one, the VOT boundary occurs mostly in the voicing lead. The level of perceptual confusion at the category boundary in this test appears to extend considerably for some subjects while for others it is quite small. The examples of the former group are shown by subjects AA, FN and NAM who assign the VOT boundary at 4 msec, 3.5 msec and 7.5 msec respectively. What is surprising about these results is not the occurrence of the category boundary in the short lag region as much as it is the presence of more than one 50 % crossover point for each of the above subjects. It occurs, for example twice for subject FN at VOT -10 ms and 17 ms. For the other three subjects it is perceived at three different locations: at VOT -5 ms, 3 ms, and 14 ms for subject AA, and at -7 ms, -5 ms, and 10 ms for subject HD. For subject NAM it occurs at -5, 3.5 ms and 24 ms of the VOT dimension. In each case the three crossover points are averaged so that the mean 50 % crossover can be used for any required statistical analysis. The judgements of the remaining subjects are all associated with one crossover point. For this latter group the voicing cue appears to be quite sufficient for the /Daa-Taa/ contrast.

In an earlier pilot study of VOT these same stimuli are presented to speakers representing different varieties of spoken Arabic. Similar findings to the ones mentioned above are found. It is expected that the listeners' perceptual performance for these same contrasts based on natural speech will be maximal with little ambiguity or confusion. The /Daa-Taa/ contrast in natural speech is not only cued by the relative onset of voicing but also by the acoustic feature of emphasis. It is assumed that these cues and probably others interact in such a way that makes the distinction quite sharp. In our synthetic speech, on the other hand, a certain amount of ambiguity may have happened for two reasons. First it may be due to the methodological requirements and control over parameters, with the need to keep them all constant except VOT. Secondly, since the only varied parameter in these



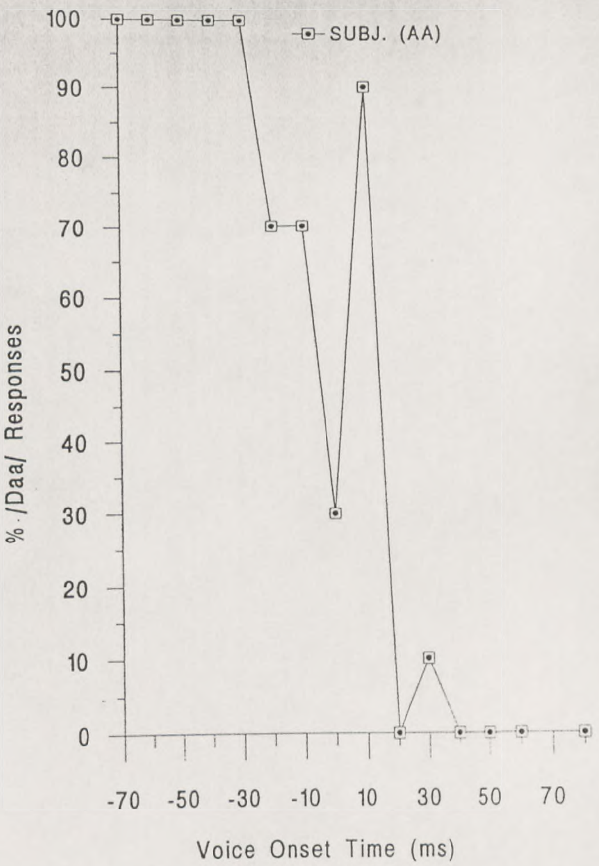
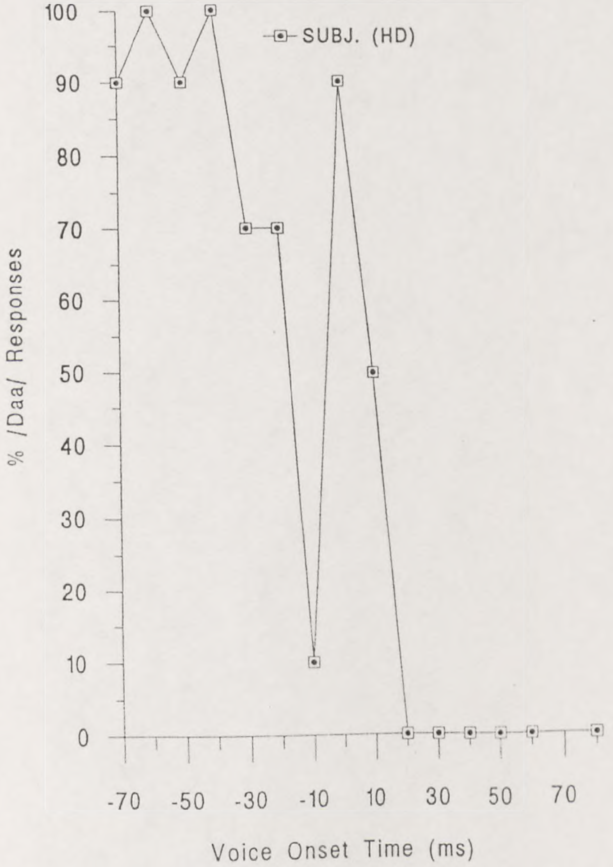
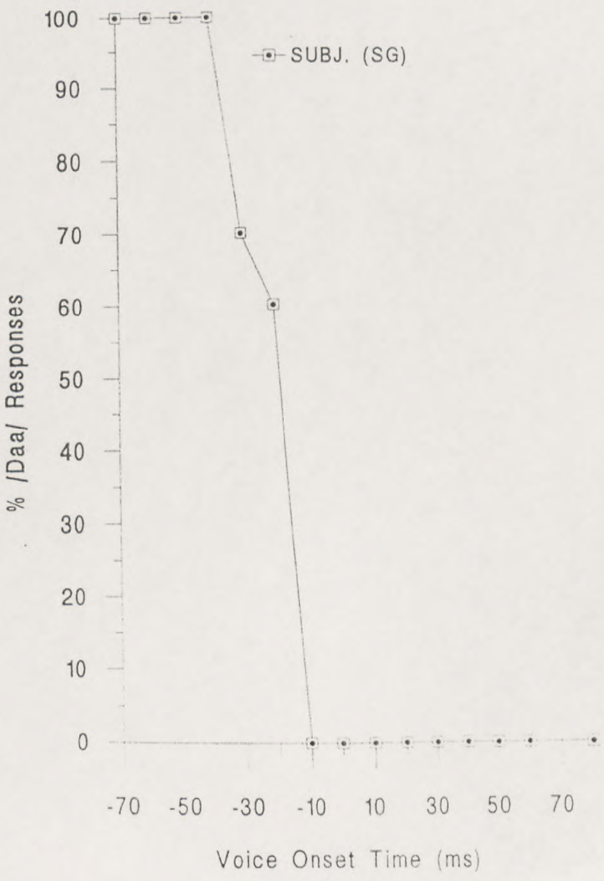


Figure 4.1.7a Percent of [Daa] responses by each of the eight subjects for test two stimuli.





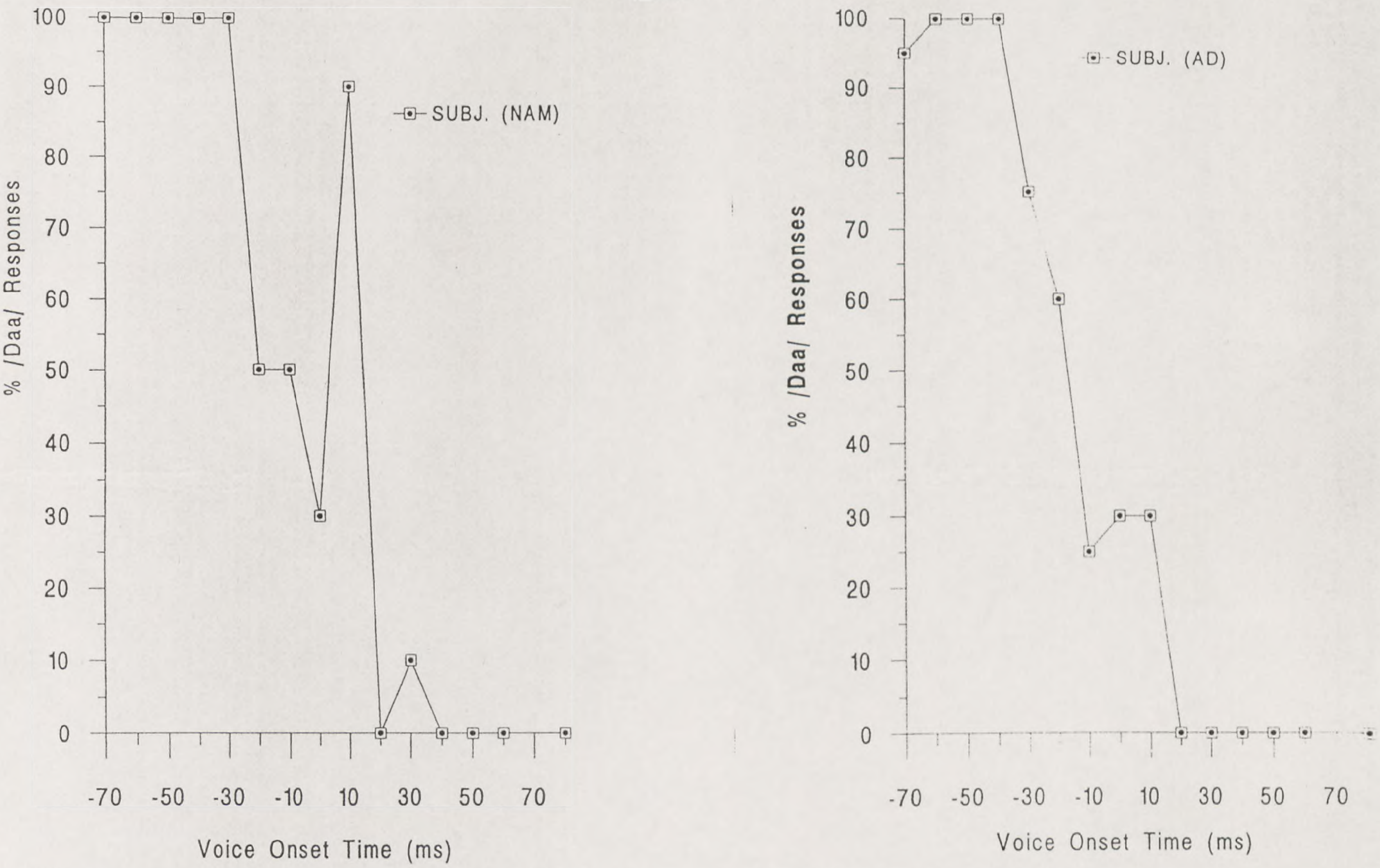


Figure 4.1.7a (Continued).

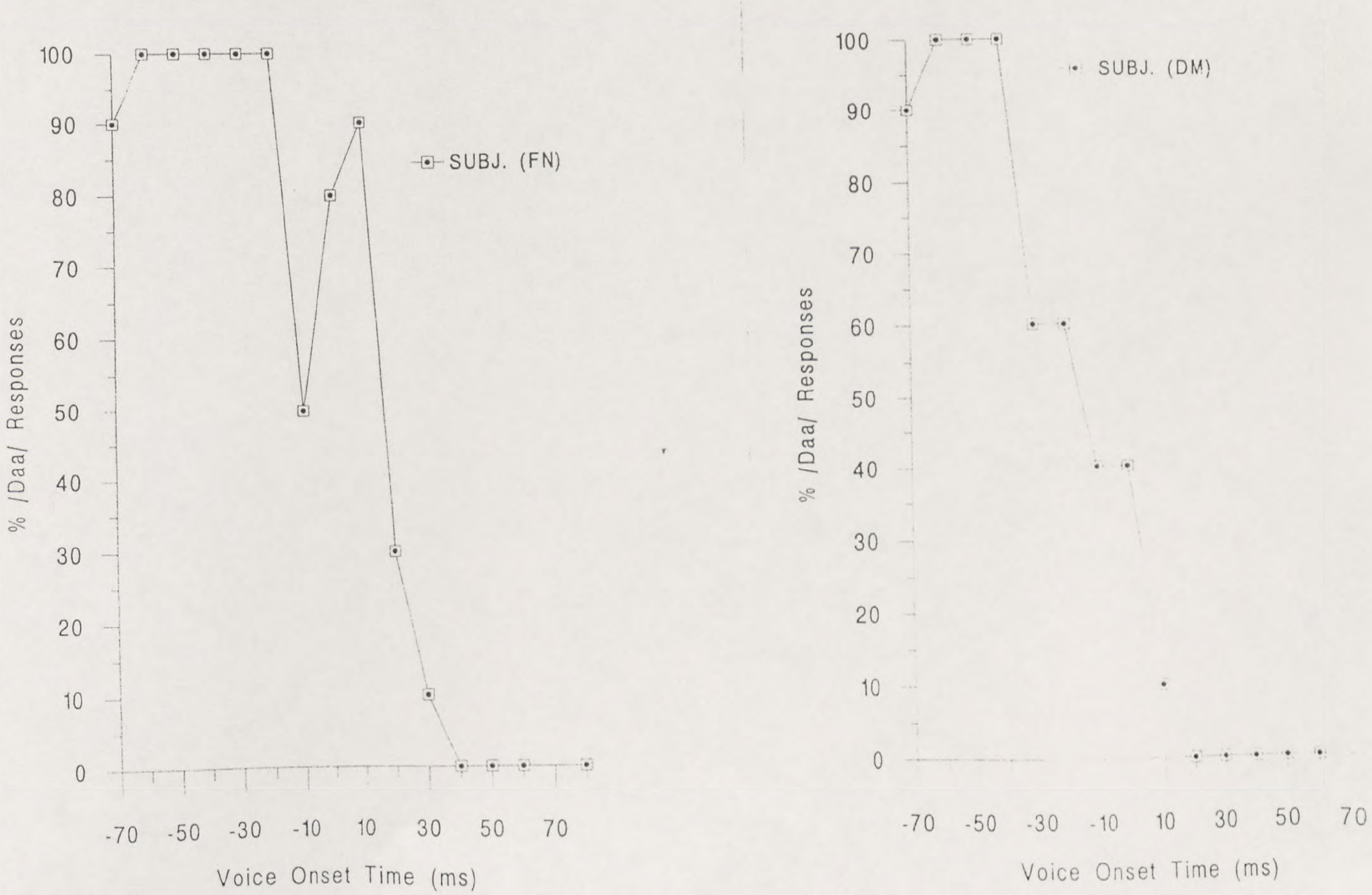
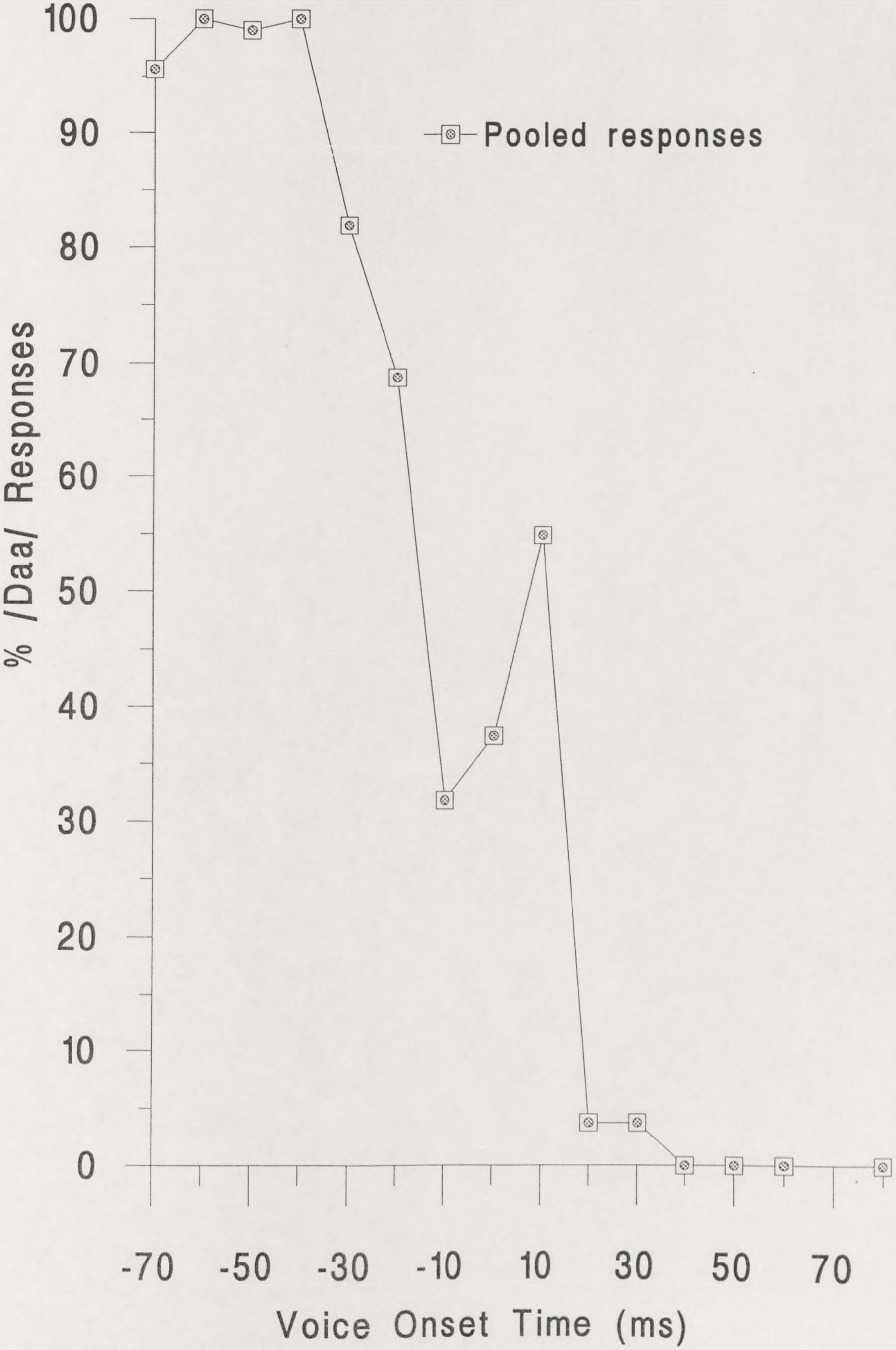


Figure 4.1.7b Pooled responses to test two stimuli by the 8 subjects.



synthetic stimuli is VOT, the assumed neutralised effect of emphasis may be, in part, responsible for the increased random responses at the category boundary. What makes us contemplate this prediction is the fact that the majority of listeners appear to face a perceptual difficulty for voice onset at or immediately after the plosive release and in particular between VOT values of zero and 15 msec. This is the region which corresponds to the rapid changes of the F2 from its onset frequency observed in our acoustic analysis.

In the present test the minimum 50 % crossover point is identified at -17 ms VOT while the maximum one is at 7 ms VOT. This may be taken as a further evidence that the / Daa-Taa/ boundary is not fixed for the individual subjects.

In addition, the averaged perceived category boundary at the 50 % crossover point occurs at approximately the same location as that of test one. It occurs at -6.33 ms VOT.

Subject	**The 50 % Crossover Point (msec)**
AA	4
MB	-16
HD	-o.67
SG	-17
DM	-15
FN	3.5
AD	-17
NAM	7.5
MEAN	-6.33
S. DEV.	10.66
RANGE	-17 : 7

Table 4.1.3 The mean , standard deviation and the range (ms) of the perceived 50 % crossover points by the the ten subjects in the /Daa/ vs / Taa/ continuum.

A comparison of the listeners' peceptual performance shows that their positioning of the 50 % crossover point is quite consistent in both tests. In spite of the slight variability in the behaviour of some individuals, a t-test of matched pairs shows the differences to be insignificant (t = 0.17, P= 0.87). This value indicates that they are responding systematically, regardless of the variations in the perceptual behaviour of some subjects observed at the boundary region in the two tests.

On the other hand, certain stimuli at the far left of the identification curves and in particular at the voicing lead beyond the -70 ms VOT start to show a gradual reduction in the level of /Daa/ response. This may indicate that the voicing lead has a certain limit in this contrast and if the designed VOT value exceeds this limit the intelligibility level of that item may pose probable perceptual difficulty.

4.1.5.3 Perception and production:

As regard the relationship between perception and production of the plosive consonants /daa/ vs /taa/ or /Daa/ vs / Taa/ there appears to be a close match between the two variables. An inspection of the mean 50 % crossover point in the labelling task shows that the two tests occur at -5.25 ms and -6.33 ms VOT respectively. These points as it appears from Figures 4.1.8 and 4.1.9 match well the VOT boundary obtained in the production results which are in fact based on one speaker, AA. There appears to be consistency between this subject's behaviour both as a speaker and as a listener. This evidence adds further support for Liberman's et. al. claim (1967) that the subjects' behaviour both as a speaker and as a listener is related and that production and perception are two interrelated processes.

This apparent relationship between perception and production shown by the results of this study is also consistent with that of Lorge (1967) for English /t/ and /d/ and Abramson and Lisker (1965, 1968). However, Abramson and Lisker (1965) find that the match between perception and production for the voicing categories for both Spanish and Thai listeners to be less than perfect. They explain this by suggesting that their synthetic stimuli may not have contained the appropriate information required by the Spanish and the Thai listeners. They mention for example, that the Thai speakers are particularly sensitive to the presence or absence of the voicing lead which they say is necessary for the differentiation between the three voicing categories in that language.

As far as categorical perception is concerned this study does not make use of any discrimination tests. Instead the lower and upper limits of the phoneme boundary are identified for each subject. These are defined as points "...along the VOT continuum where an individual's responses were associated with voiced or voiceless plosive respectively, 75% of the time" (Zlatin, 1974, p.983). The aim is not only to find the endpoints of the phoneme boundary at 75 % level but also to estimate its

Figure 4.1.8 Percent of /d/ and /t/ production occurrences and percent of /d/ responses as function of VOT for one subject, AA.

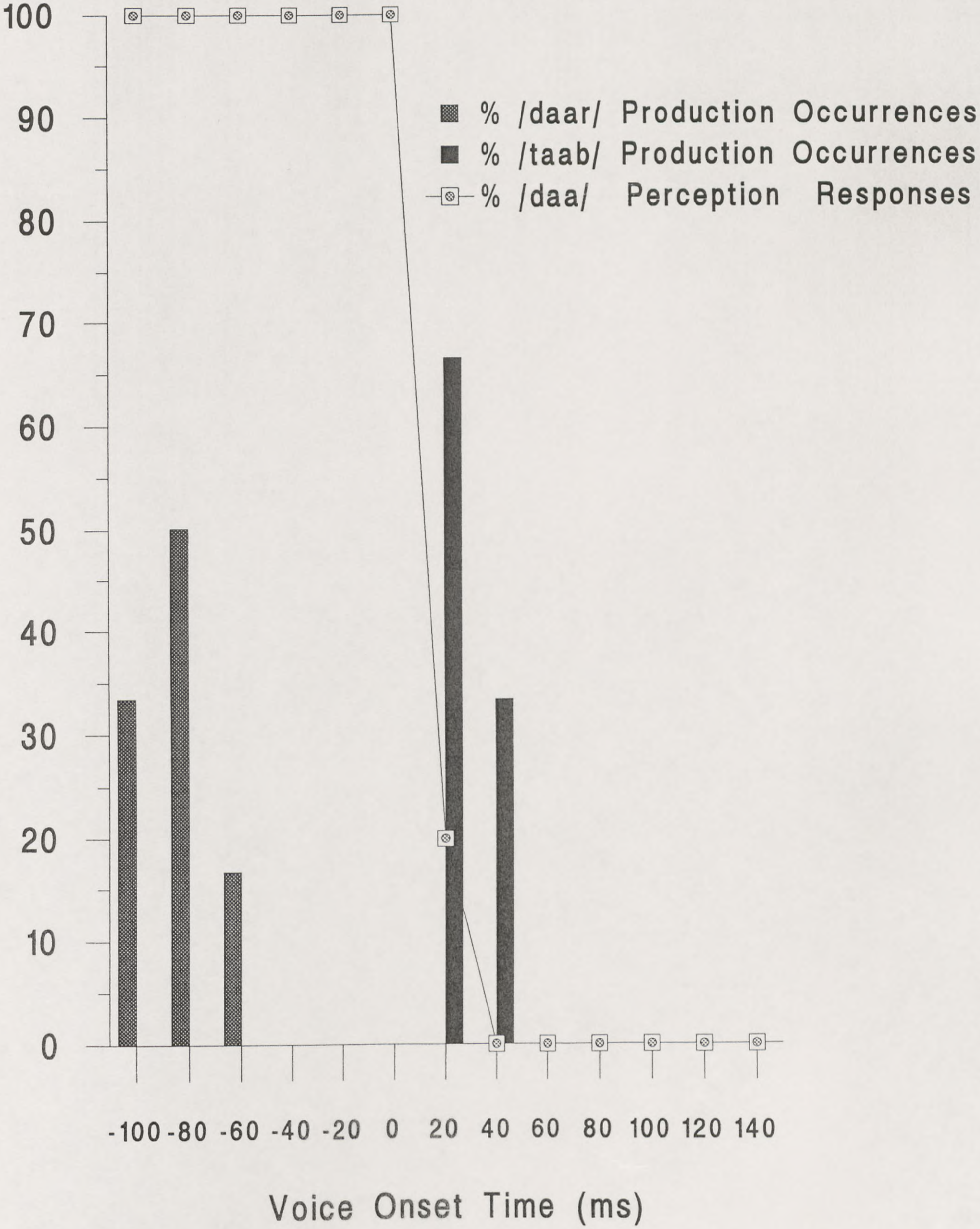
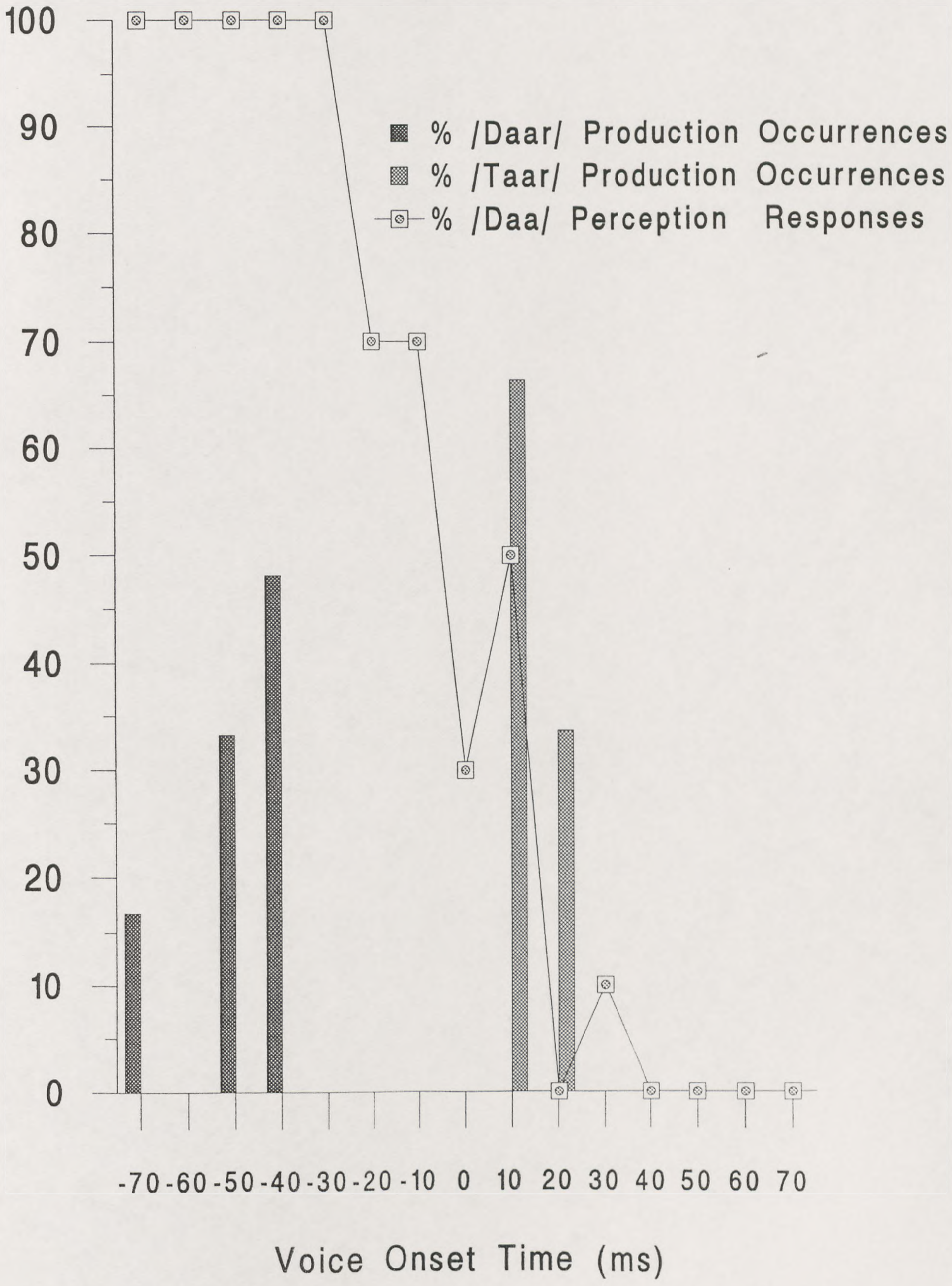


Figure 4.1.9 Percent of /D/ and /T/ production occurrences and percent of /D/ responses as function of VOT for one subject, AA.



width. This latter variable is defined as the difference in ms between the upper and the lower limits. It is hoped that specification of these parameters will enable us to get an idea of how fast and sharp the change is from 75 % correct voiced responses to 75 % correct voiceless responses.

Table 4.1.4 displays the VOT boundary values at the lower and upper limits for the voicing contrast /daa-taa/. It appears that these values vary from one subject to another. The lower limit for subject AA, for instance, is at 10 ms VOT, while the upper limit is at 18 ms VOT. Consequently, the shift from 75 % /daa/ responses to 75 % /taa/ responses can be estimated to be approximately 8 ms. Subject FN, on the other hand, displays the biggest phoneme boundary width, 23 msec. Generally, the VOT range of variations in the perceived lower limits is between -27 msec and 10 msec, while the VOT range for the upper limits is between -6 ms and 18 msec. The average perceived values for these two extremes is at -13 msec VOT for the lower boundary points and 4 msec for the upper ones.

The Phoneme Boundary **

Subject	Lower limit	Upper limit	Width
AA	10	18	8
MB	-14	7	21
HD	-8	8	16
RA	-18	-6	12
SG	-16	-6	10
FN	-18	5	23
MBA	-27	-5	22
AD	-14	8	22
MEAN	-13	4	17
S. DEV.	10.76	8.60	6.07
RANGE	-27 : 10	-6 : 18	8 : 23

Table 4.1.4 The lower limit, upper limit and the width of the phoneme boundary for /daa/ versus /taa/ at the 75 % perception level .

** All values in this table and subsequent ones are in (ms).

On average, each subject needs about 17 msec for the transition from 75 % voiced responses to 75 % voiceless responses. This period indicates that the distinction is quite sharp and categorical. It is true that discrimination tests as described by Studdert-Kennedy et. al. (1970) is not performed in this study, nevertheless the boundary width seems to give a fairly good estimation.
These results are fairly consistent with that obtained by Zlatin (1974) who finds a mean category boundary width of 5.80 msec for the alveolar voicing contrasts in English.

In the second test, Table 4.1.5, the boundary width is found to have a bigger value than that of test one. Thus, the average perceived phoneme boundary at lower limit is at 23.38 msec. The mean upper limit occurs at VOT boundary 6.125 msec. Some examples of the phoneme boundary width are AD = 39 ms, NAM = 38 ms and FN = 38 ms. In spite of that the mean boundary width is found to be 29.5 ms. Clearly it is quite extensive particularly if compared with that obtained for English /da/ vs /ta/ by Zlatin (1974). It is almost double that of test one. The significance of this statement is confirmed when the perceived phoneme boundary width in the two tests are subjected to a statistical test of significance. The related (2-tailed) t-test result is ($t = -2.90$; $p = 0.02$). It indicates that the same subjects responded differently to the lower and upper points of the phoneme boundary in the two tests. This difference may be explained by the perceptual difficulty exhibited by some subjects at the category boundary in test two. In the case of those particular subjects, categorical responses are not obtained. The perceptual difficulty associated with this particular voicing contrast is not entirely unexpected at least for subject AA who has shown VOT overlap for /D/ and /T/ in the acoustic study of VOT.

	The Phoneme Boundary		
Subject	**Lower limit**	**Upper limit**	**Width**
AA	-20	7	27
MB	-20	-13	7
HD	-20	14	34
SG	-30	-14	16
DM	-33	4	37
FN	-15	23	38
AD	-28	11	39
NAM	-21	17	38
MEAN	-23.38	6.125	29.5
S. DEV.	6.19	13.99	12.21
RANGE	-33 : -15	-14 : 23	7 : 39

Table 4.1.5 The lower limit, upper limit and the width of the phoneme boundary for the voicing contrast /Daa-Taa/ at the 75 % perception level.

4.1.6 Discussion:

These results are consistent with those reported by Yeni-Komshian et. al. (1977) regarding voicing in Lebanese Arabic. They use simple imitation responses instead of forced choice identification test to investigate whether the subjects are sensitive to the systematic changes in their synthetic stimuli of VOT. Their results show quite similar ranges to the ones established in this study. In their study the shift from the voiced category to the voiceless one is bounded roughly by the 30 msec voicing lead and the 35 ms voicing lag. Their best estimate of the 50 % crossover point is between zero and 15 msec (i.e. 5 to 15 msec to the right of the one obtained in YSA).

The findings are also in agreement with that of Watson (1992) for French, Lee Williams (1977) for Spanish. Watson finds adult French listeners to assign the category boundary at the range of zero to 15 msec. Williams (1977) who investigates the perceptual effect of the voicing lead on the voiced and voiceless distinction by Peruvian Spanish listeners concludes that all the listeners who participated in that test have their 50 % crossover point at the voicing lead region. Furthermore, she finds that the presence of voicing lead proves to be ".. a sufficient positive cue for voicedness (as opposed to voicelessness) for most of the Spanish listeners". Thus, such languages as YSA, French and Spanish appear to share a common characteristic (i.e. the need for a voicing lead of some duration for the accurate identification of voiced plosives). The saliency of this cue particularly for the voiced plosives is increased by the fact that it is the only variable present during the closure interval and prior to the articulatory release. It is our assumption that the absence of the voicing lead from the voiced plosives will seriously reduce their intelligibility at least for YSA. This is because YSA, unlike English does not utilize voicing lag greater than about 55 msec for the denti-alveolar plosives. The range for YSA Arabic /t/ corresponds roughly to the range for English /d/. A number of researchers have indicated that the VOT category boundary in English falls between 25 msec and 40 msec (Abramson and Lisker, 1965, 1968; Lorge, 1967; Zlatin, 1974). They all agree that the voicing contrast in English cannot be described in terms of the presence versus absence of voicing during the closure, but rather as a difference in the degree or the relative timing of voicing along the VOT dimension. Thus English /d/ is said to be perceived in the short lag region whereas /t/ is perceived in the long lag region.

Although in YSA most subjects assign the crossover point at the voicing lead region, as indicated above, there are some who assign it in the short voicing lag region. This is a clear indication that the voiced response is not entirely cued by the voicing lead. Other acoustic cues may have a role in shifting the VOT boundary in YSA. Among these possible cues is the F1 transition. Some researchers suggest that the VOT boundary in English is not fixed and that it varies directly with the transition duration (Stevens and Klatt, 1974). The VOT boundary value is higher with increased transition duration. Other researchers have argued against this conclusion and suggest that the voiced plosives judgement is not so much determined by the rising F1 transition as much as it is related to the low F1 onset frequency (Lisker,1975). Thus, a VOT value at the phoneme boundary will be heard as voiced if the F1 onset frequency is low and as voiceless if it has a higher onset frequency.

Evidence from languages other than English suggests that the use of this parameter as a voicing cue in a contrastive way depends on whether the language relies on the voicing lead or on the long lag in its voicing categories. According to Simon and Fourcin (1978) the role of F1 in shifting the VOT category boundary will be limited to languages which make use of the short versus long lag contrast in their voicing categories, rather than those which contrast the presence versus absence of the voicing lead. In the latter type both the voiced and the voiceless plosives are believed to have similar low F1 onset frequency. Their claim is supported by data from French in which they show that French children do not respond to variations in the F1 transitions.

Thus, since YSA like French is characterised by an extensive voicing lead in the production of its voiced plosives, it may be predicted that the F1 onset frequency will be similar for both voicing categories and that the YSA listeners will not be sensitive to the perceptual effect of this cue.

Similarly, the level of the burst and amplitude of the aspiration noise, though necessary in some languages, is less likely to be a relevant factor affecting the VOT category boundary in YSA. Repp (1979) based on data from English finds a trading relation between what he calls "A/V ratio" and the VOT boundary. By the terms "A/V" he means the ratio of the overall amplitude of aspiration noise (A) to that of the voicing portion (V). He finds that an increase in the "A/V ratio" results with a lowering of the VOT boundary.

Nevertheless, it is our feeling that the aspiration feature will have little effect on the VOT phoneme boundary in YSA. In such a language neither the voiced nor the voiceless plosives have aspiration noise, and if it occurs particularly with the velars in the context of front vowels, as shown in our acoustic analysis, it will not be expected to be sufficient to give this feature any perceptual weight. More recent evidence of this assumption is given by Watson (1992) from a language whose voicing distinction has some resemblance in the voicing feature to that of the YSA. He finds little effect of the presence or absence of aspiration in shifting the VOT boundary by French adult subjects, while the English listeners are sensitive to such a cue. He attributes this difference between the speakers of the two languages in response to this cue to the language experience.

In our view, the more likely cue to be used by the YSA listeners in modifying the VOT boundary is the fundamental frequency of voicing. The reduced chance of effectiveness of such cues as aspiration amplitude and the F1 onset frequency in YSA, gives the fundamental frequency of voicing a particular importance. This prediction is strengthened by the remark of Haggard, Summerfield and Roberts (1981) that the F0 cue tends to be stronger in languages characterised by the voicing lead, while in languages characterised by the occurrence of the long lag in their voicing system it tends to be less important.

Another issue related to the shift of the VOT boundary bears on the vocalic context. One limitation of this study is that it makes use of only one vocalic context namely the long vowel /aa /. A future study of the perceptual characteristics of VOT in YSA should include plosives in all vocalic contexts.

To conclude: one of the findings of this study is that the voicing lead is crucial for the voicing distinction in YSA. The mean VOT boundary is found to occur in the voicing lead. The shift from the voiced category to the voiceless one is sharp for /daa/ vs /taa/. These findings confirm our prediction that the relative time between the plosive release and the voicing onset is a quite reliable indicator of the voicing distinction in YSA not only acoustically but also perceptually. The results have also shown a fairly close correspondence between perception and production.

4.2 Perceptual Study of Emphasis in Yemeni Spoken Arabic

4.2.1 Introduction and Aims

4.2.1.1. Test One: Formant Transitions and Emphasis.

The aim of this test and the next one is to investigate the role played by the formant transitions, and in particular that of the second formant, in the distinction between emphatics and non-emphatics in YSA. There are not as many studies on the transitional cues and consonant perception in Arabic and YSA as there are in English.

To our knowledge, the only study that dealt with this aspect was that of Obrecht (1968) in which synthetic consonant vowel syllables were used and the F2 onset frequency was varied to determine its effect on emphasis. The listeners responses showed that they were able to identify both emphatic categories accurately from the F2 onset frequency and the associated change in transitions.

To further investigate the perceptual effect of systematic variations in onset frequency and transitional cues on the emphatic distinction in YSA the present study is conducted. The aim is to determine the range of the phoneme boundary at its lower and upper limits for another Arabic dialect and with stimuli covering a wider range on the frequency scale from 700 Hz to 2100 Hz with equal increments of 100 Hz. This may also help us to find the F2 onset frequency values beyond which the emphatics cannot be heard. In addition, the obtained boundary point will be compared with those results from the acoustic analysis. This may add more information on the relation between perception and production and in particular how the individual subjects, particularly AA and WS, react to that relation. In our acoustic analysis, the emphatic and non-emphatic consonants have clearly shown differences in the nature and extent of the formant transitions. It is found, for example, that the transition duration associated with the nonemphatics is approximately 50 msec whereas for the emphatics it extends up to 80 msec.

4.2.1.2 Test two: the role of context in the distinction between emphatics and non-emphatics:

From our acoustic analysis it is also felt that vowel formant structure may have an effect on the consonant-vowel intelligibility. In particular, the location of the F2 steady state frequency for the vowel and the particular consonant with which it is combined vary significantly from one context to another. This variable appears to be more crucial in changing the whole visible pattern of the CV-syllable than does any other formant feature. It is evident from our acoustic analysis that the emphatics and non-emphatics do not only vary in their F2 onset but also in their F2 steady state frequencies. Results show that the emphatics are associated with lower F2 steady state as compared with that of the non-emphatics. The role of context in cueing the emphatic/nonemphatic categories has also been investigated by Ali and Daniloff (1974). They tried to determine whether listeners could identify the presence of an emphatic or non-emphatic consonant in a word even if the contoid concerned is removed. They predicted that the contextual effect of emphasis as a distinctive feature would extend to the neighbouring vowels and that this effect would be great enough to enable the listeners to perceive the words even if the emphatics and non-emphatics including their transitions had been removed from the test words. The Iraqi listeners, to whom these stimuli were presented were able to identify the emphatics and non-emphatics above chance level. It was therefore concluded that emphasis has an effect on the neighbouring vowels. A similar vocalic influence was reported for English by Schatz (1954). He found that when the consonant burst was paired with the vowel from which it was removed the consonant was heard correctly. However, when the consonant burst was combined with a different vowel context, the identity of the consonant was completely changed. Again, it was concluded that the vocalic context had an important role in characterizing the plosive consonant. Not all researchers, however, seemed to accept that conclusion. Grimm (1966) investigated the influence of the consonant on the neighbouring vowel and vice versa by gradual deletion of the consonant. He found that the consonant responses were below chance level unless the stimuli presented to the listeners contained some portions of the consonant indicating that the influence of the consonant in Grimm's investigation was not perceptually evident on the vowel.

The aim of this perceptual test is to investigate the assumption that vowel allophones can be used as perceptual cues for the emphatic and non-emphatic consonants in YSA and in what way the vowel context affects the perception of emphasis. In particular the study will center on the question of whether the emphatics and non-emphatics may be accurately identified by certain changes in the vocalic context. As indicated earlier (See section 1.2.2) these two categories are shown to be followed by two clearly distinct variants of the vowel /aa/. The emphatics are normally associated with a backed allophone of /aa/ whereas the non-emphatics are followed by a fronted variant. It is thought that this particular context has larger modifying influence on the emphatics than on the non-emphatics and it may serve, therefore, as an adequate cue. If this prediction proves to be right then this study will have some implications for the domain of emphasis and perhaps for the interaction of different acoustic cues for the perception of emphasis.

4.2.2. Stimuli:

To find whether or not the F2 onset frequency and steady state when manipulated separately can in fact be used as effective cues synthetic stimuli for the emphatic and non-emphatic consonants were varied and the resulting stimuli were presented to a group of YSA listeners for identification. Two continua of emphatic and non-emphatic CV-syllables are generated using the same terminal-analog Klatt Synthesizer (type KLSYN version 1.4, 1984) used for the voicing study. The sampling rate for the Synthesizer output was set at 10 KHz and the output was low-pass filtered at a cut-off frequency of about 5900 Hz (See 4.1.2 for more details on the instrumental set-up used in the perception studies).

Initially, it was planned that actual word pairs should be synthesized with only F2 varied and all other parameters to be kept constant. Monosyllablic utterances were eventualy adopted because it was felt that the syllable approach would be more appropriate especially with regard to the control of variables. This approach was more convenient in view of the time required for synthesizing a more satisfactory stimulus. The two continua had different vocalic environments: the first was in the context of /ii/ while the other was in the context of /aa/.

The specifications of the stimuli parameters were based on estimated values from naturally spoken words by speaker AA: /Tiin/ 'clay' and /tiin/ 'figs' for test one and /Taab/ 'ripe' and /taab/ 'repented' for test two. The basic stimulus set consisted of a burst duration of 50 msec in test one and 15 ms in test two. It consisted mainly of frication noise with an amplitude of 60 dB generated from the parallel branch of the synthesizer. The burst was followed by the formant transitions and the steady state frequencies of the formants for the following vowels. The transitions were moving from the burst towards frequency values appropriate for the following vowel. The fundamental frequency (F0) in both continua was falling. In the first test it started at 123 Hz and fell gradually to 65 Hz at the end of the stimulus. In test two it began at 150 Hz and fell linearly to 100 Hz at the end of the stimulus. The amplitude of voicing (av) in both tests was set to zero where it was not needed and to 60 dB where the voicing source was required. At the last 50 msec of each stimulus, the amplitude of the voicing source, which was derived from the cascade branch of the synthesizer, was lowered gradually till it reached zero at the end. The formant bandwidths (p1-p5) were left at the default values, whereas the formant amplitudes (a1-a5) were given appropriate values (See Appendix 6 for more details on the specification of the parameter values).

The formants frequencies varied in both continua according to the vocalic contexts and the task required by each test. Consequently, the formant structure and in particular F2 in both sets need to be described separately.

The basic stimulus in test one from which others were varied had a total duration of 300 msec. The extent of the F2 transition was between 50 msec and 80 msec and the steady state duration was generally 200 msec. There were four formant resonators contributing to this stimulus. The first three formants had rising transitions, while the fourth one had a flat transition. The F1 transition was given a short duration of 30 msec. Its onset frequency began at 180 Hz and continued to rise till it reached the value appropriate for this formant steady state at 330 Hz. F3 had a transition duration of 40 msec. Its steady state was set at 3000 Hz. The addition of the third formant to these stimuli was according to recommendations of Liberman et. al. (1954). They indicated that with the first two formants of certain plosive consonants in the context of /i/ they were not able to get a clearly dominant /t/ or /d/ responses. They therefore suggested that an increase in the identifiability of these plosives would result from the inclusion of appropriate transition of the third formant, though they added that its effect is less important as compared with that of the second formant. Figure 4.2.1 shows a schematic diagram of these formants.

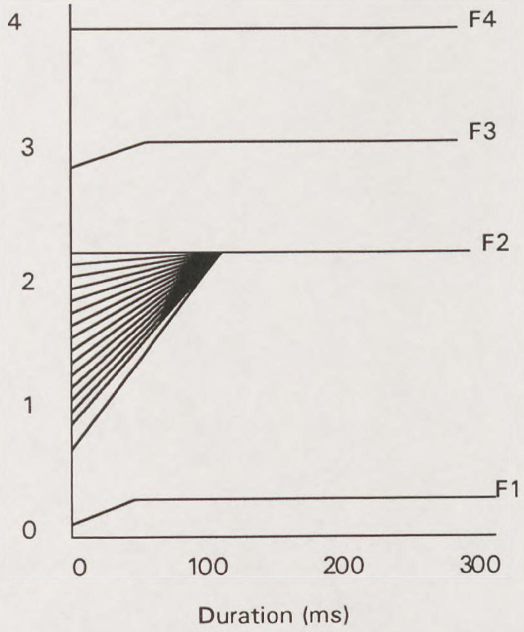


Figure 4.2.1 Schematic representation of varied F2 onset frequencies and the other formants used in the synthesis of the /Tii-tii/ continuum for the study of the emphatic/nonemphatic distinction. The ordinate shows frequency in kHz.

The F2 transition which constituted the main variable in this study had a transition duration ranging between 50 msec at nonemphatic end of the continuum to 80 msec at the emphatic end. In all cases the F2 transition was moving towards frequency appropriate for the steady state of the vowel which was at 2200 Hz. The F2 onset frequencies were varied in steps of 100 Hz from a point 700 Hz below the second formant steady state to a point 2100 Hz below that frequency. There were 15 such starting points. These designated values for each stimulus were listed in Table 4.2.1 below:

After several attempts at getting intelligible emphatic and nonemphatic consonants, the F2 transition was varied from 80 msec at the emphatic end of the test to 50 msec at the non-emphatic one. The transition duration was increased in 2 msec steps. Longer transitions were found to result with some improvement in the identification of the emphatic category. Apart from this all other parameters were kept unchanged in all the stimuli. Wideband spectrograms of the end and middle points for the synthetic continuum are shown in Figure 4.2.3 below.

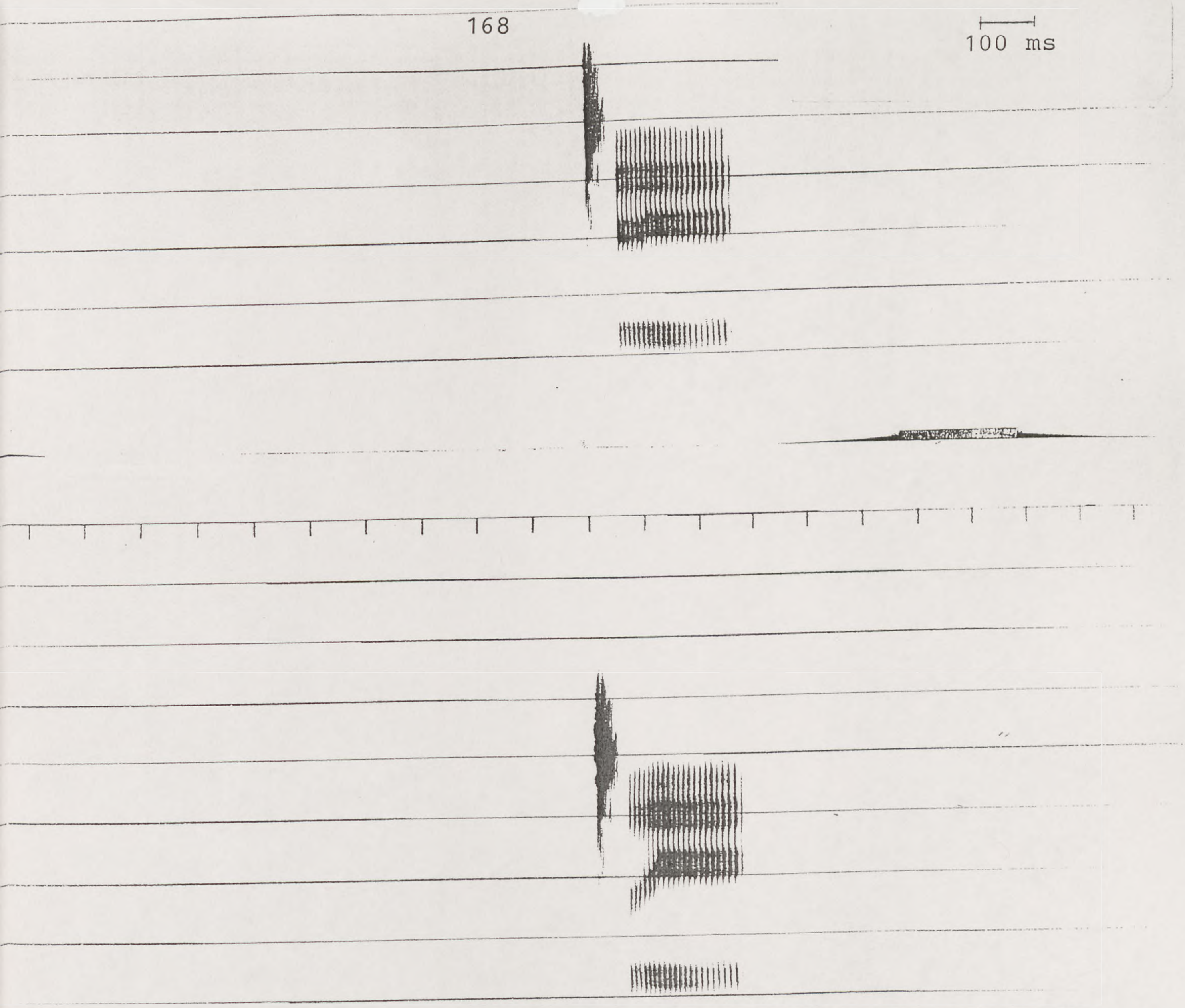
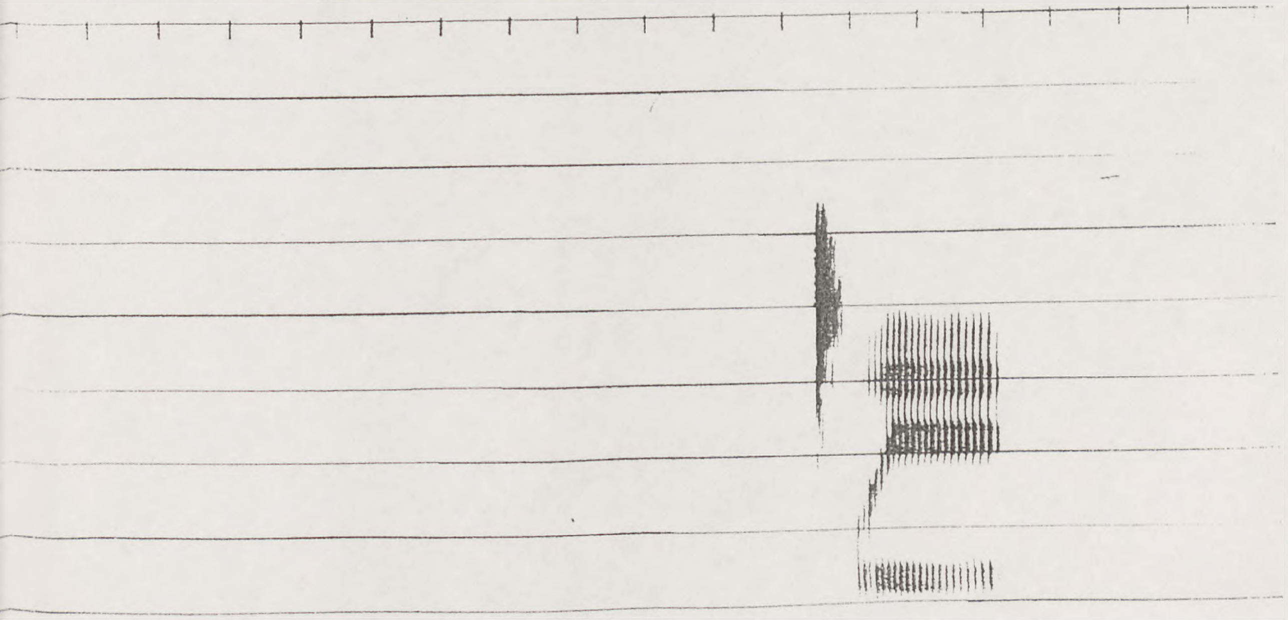


Figure 4.2.3 Wideband spectrograms of the end and middle points for the synthetic continuum presented in the identification of the emphatic and non-emphatic categories in test one stimuli: [Tii] vs [tii].



Stimulus No.	The F2 Onset Frequency (Hz)
1	700
2	800
3	900
4	1000
5	1100
6	1200
7	1300
8	1400
9	1500
10	1600
11	1700
12	1800
13	1900
14	2000
15	2100

Table 4.2.1 Calculated values of F2 onset frequencies (Hz) for the /Tii/ versus /tii/ continuum.

In the second test, on the other hand, only three formants were used in the generation of the basic synthetic stimulus. The frequency of the F2 steady state for the vowel allophone was the main variable in this test. Systematic manipulation of this variable resulted in a continuum in which at one end the F2 steady state frequency was appropriate for the back variant of /aa/ while at the other end of the continuum it was more appropriate for the open front variant of /aa/. It was assumed that a change from one of these vowel allophones to the other would cue the identity of the emphatic and non-emphatic categories. The F2 onset frequency was constant in this test. All F2 stimuli had an onset frequency of 2000 Hz. The first formant onset frequency began at 490 Hz and rose to a steady state of 600 Hz (See the accompaying parameter values in Appendices 6 and 7). The transition duration was 35 msec. The F3 was slightly falling. Figure 4.2.2 presents these various F2 steady states schematically.

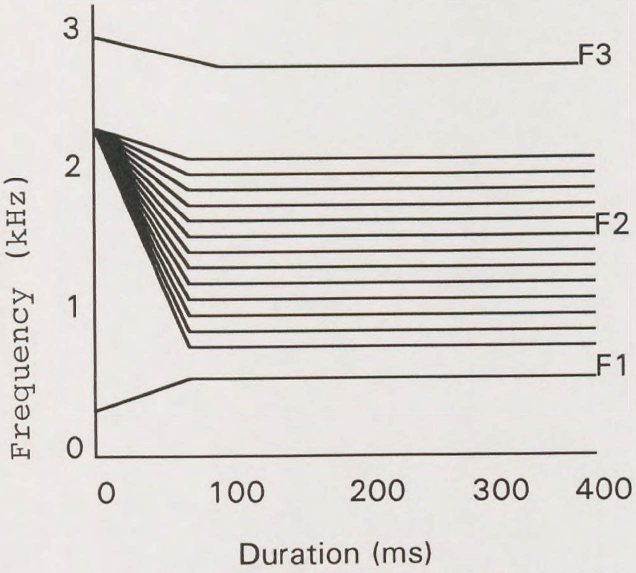


Figure 4.2.2 Schematic representation of the varied F2 steady state frequencies used in the synthetic pattern for the study of emphasis in the /Taa-taa/ continuum. The ordinate shows frequency in kHz.

Stimulus No.	F2 Steady State Frequency (Hz)
1	700
2	800
3	900
4	1000
5	1100
6	1200
7	1300
8	1400
9	1500
10	1600
11	1700
12	1800
13	1900

Table 4.2.2 Calculated values of F2 steady state frequencies (Hz) for the /Taa/ versus /taa/ continuum.

The F2 steady state was varied in equal steps of 100 Hz between 700 Hz, the frequency value appropriate for producing a quite intelligible emphatic plosive, to 1900 Hz, the level of frequency appropriate for producing a clear non-emphatic consonant. The F2 steady states representing the thirteen stimuli in this test are given in Table 4.2.2 .

To check the varied parameters, wideband spectrograms of all stimuli in the synthetic stimuli for the two tests were produced. The F2 onset frequency in test one and its steady state in test two were measured. The measured values were compared with the calculated ones and good correspondence between the two set of values was obtained. Wideband spectrograms of the end and middle points of the synthetic stimuli are shown in Figure 4.2.4.

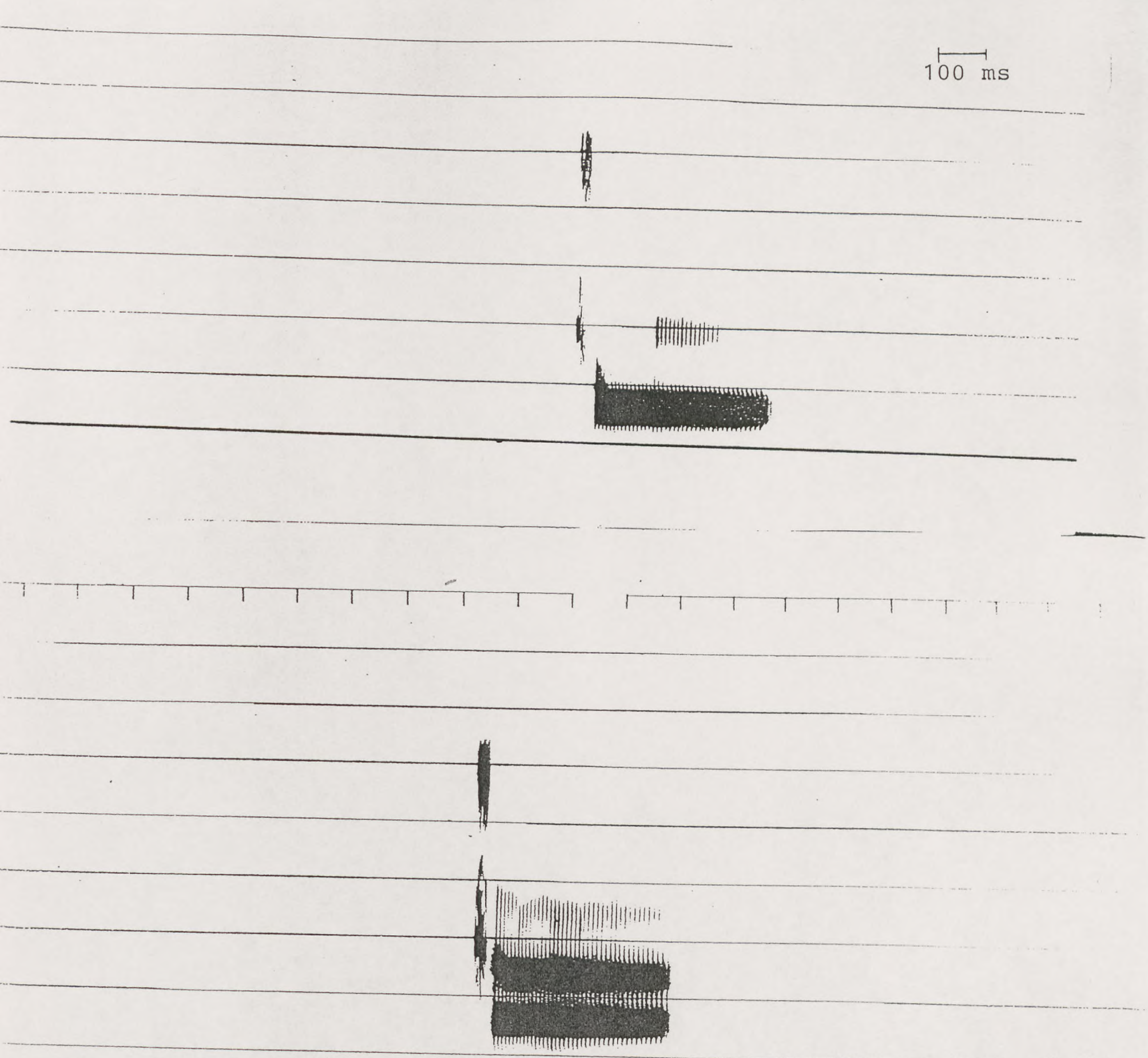
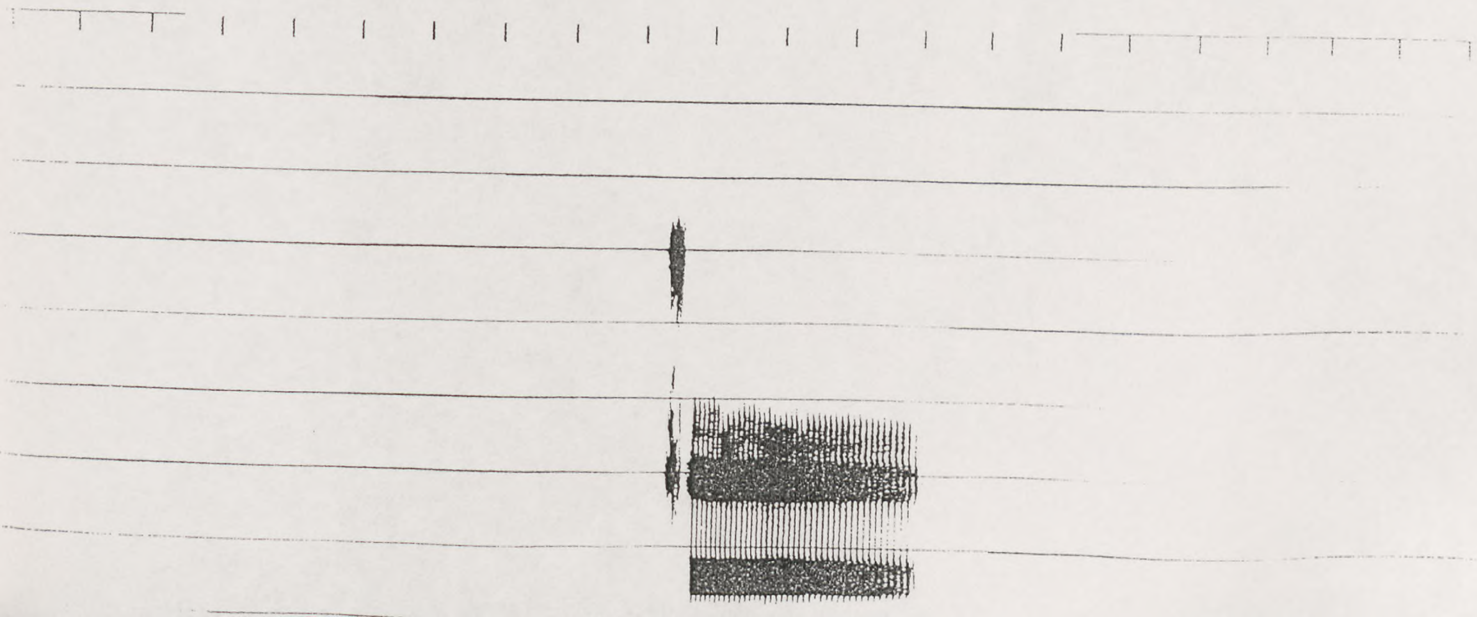


Figure 4.2.4 Wideband spectrograms of the end and middle points of the synthetic continuum presented for the identification of the emphatic and non-emphatic categories in test two: [Taa] vs [taa].



4.2.3 Procedure:

The tests preparation consisted of randomization and recording of synthetic material. The stimuli in each test were organised in ten randomized orders using the same program utilised for the preparation of the VOT continua. To avoid repetition, the reader is referred to section 4.1.4 for details on the procedure. Ten extra stimuli were added to test one and thirteen to test two. It was hoped that these additional stimuli would be used as practice items. They would help the listener to get aquainted with the nature of the task they were required to do and to familiarize themselves with the synthetic utterances. The subjects were given response sheets with the accompanying instructions in Arabic. The instructions read to the listeners were as follows:

"T*his is an experiment in speech perception. You are going to hear a sequence of consonant-vowel utterances. In the first test you are asked to identify the 160 syllables as /Tii/ in /Tiin/ or /tii/ in /tiin/. In the second test there will be 143 stimuli. You are asked to write your response on the sheet provided as /Taa/ in /Taab/ or /taa/ in /taab/. You should write the sound as soon as you hear it. Do not add or change any item after the test has finished. There will be 2500 msec. interval between successive stimuli. Each group of ten stimuli within each test is separated by a 5000 msec interval*".

Examples of translated copies of the answer and reponse sheets for both tests can be found in Appendix (5). The response sheet was prepared with the appropriate blanks for writing the answers. It was a forced choice task. The administration of the two tests was separated by a ten minutes break.

After the administration of the tests the subjects were invited to comment on the intelligibility of the synthetic utterances. One of the points they mentioned was that some stimuli were quite intelligible and easy to identify while some others were quite ambigious because they could not be grouped as belonging to one category or the other. In fact, one listener HD indicated that not all the stimuli had the same degrees of intelligibility . Some were very intelligible /T/ or /t/ while others were in between. This might be an indication that this listener was quite aware of the controlled variations in each test and that she was responding positively to them. They also commented that at the beginning of each test their responses relied on guessing but that as the test went on they became more familiar with them. This was expected because almost all of them were not exposed to synthetic utterances before. On the whole they were satisfied with the quality of the synthetic utterances. All of the subjects participating felt that there were two distinct categories within each test and none of them complained about serious intelligibility problems in these utterances.

4.2.4 Subjects:

The subjects were ten native speakers of Yemeni Spoken Arabic. Eight of them were males and two were females, HD and WR. They all reported having normal hearing with no obvious auditory illness or hearing disorders. Eight of them participated in the voicing experiments while others were new ones. Their age ranged between 22 and 36 years old. All of them were living in Yemen except, AA and WS who were studying in the U. K. The same subjects participated in both tests.

4.2.5 Results:

4.2.5.1Test One: Formant transitions.

All the responses were calculated for the emphatic and non-emphatic categories. This was done for each of the ten subjects. The responses were then transferred into percentages and presented graphically in Figure 4.2.5a. The ordinate shows the percentages of /Tii/ responses presented as a function of the varied F2 frequency in the abscissa. The figure displays the results for each subject. The emphatic and non-nemphatic consonants occupy clearly distinct regions. The variations in the F2 onset frequency prove to be a sufficient cue enabling the listeners to categorize the consonant-vowel syllables /Tii/ from /tii/. This kind of identification is achieved by all subjects without exception. Nevertheless, they are not all similar in their allocation of the 50 % crossover point along the frequency scale. For example, subject AA heard stimuli 1-6 as /Tii/ whereas stimuli 8-15 are identified as /tii/. At the region of 1200 Hz and 1300 Hz (i.e. stimuli 6 and 7) he appears to respond randomly. For this particular subject, this is the region of perceptual confusion at which his responses changed suddenly from emphatic to non-emphatic. Specifically this change occurs at approximately 1274 Hz. At this point both /Tii/ and /tii/ are perceived 50 % of the time. Subject WS has a slightly higher crossover point at approximately 1500 Hz. All stimuli in each side of this point are perceived as belonging to the same category and in contrast to the other. Subject MB identifies the stimuli in a similar manner to WS. They both find difficulty perceiving the stimulus at the region of 1500 Hz. Generally, the phoneme boundary points displayed by most subjects is between 1450 Hz and 1550 Hz. The actual range of emphasis as shown by all subjects is between 1620 Hz and 1274 Hz. The perception of emphatics is characterised by lower F2 onset frequencies whereas the non-

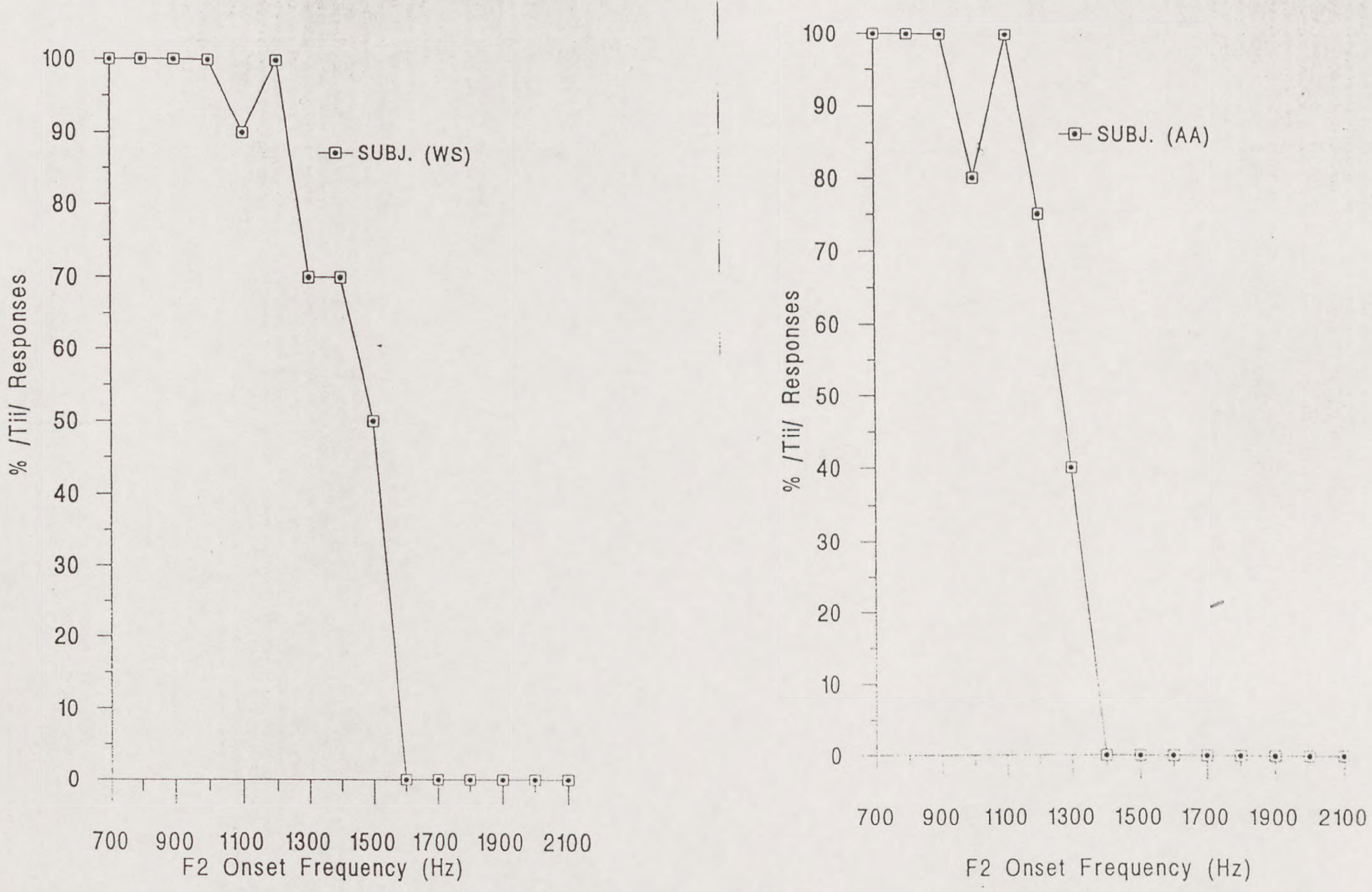
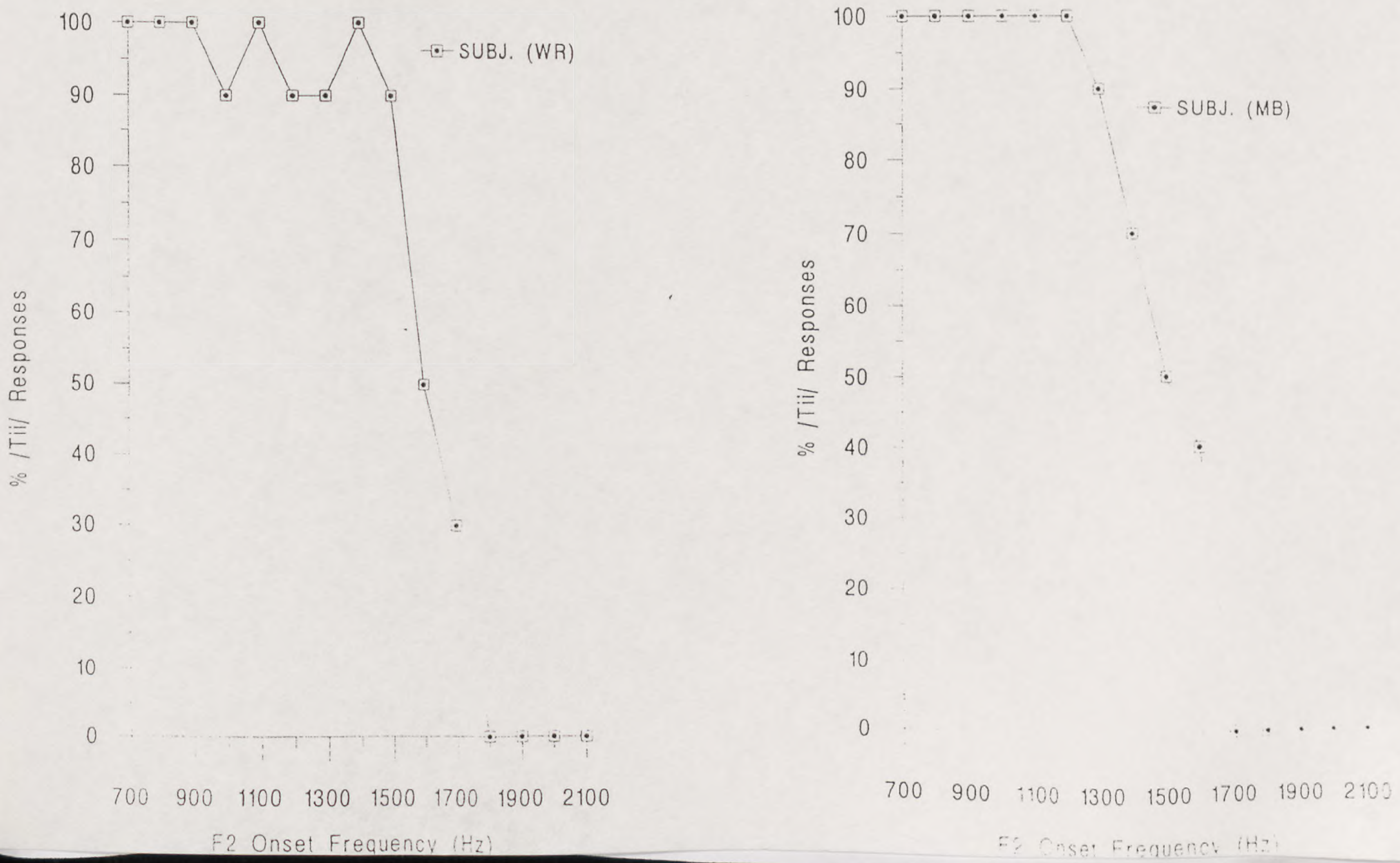


Figure 4.2.5a Percent of the [Tii] responses by the individual subjects for test one stimuli.



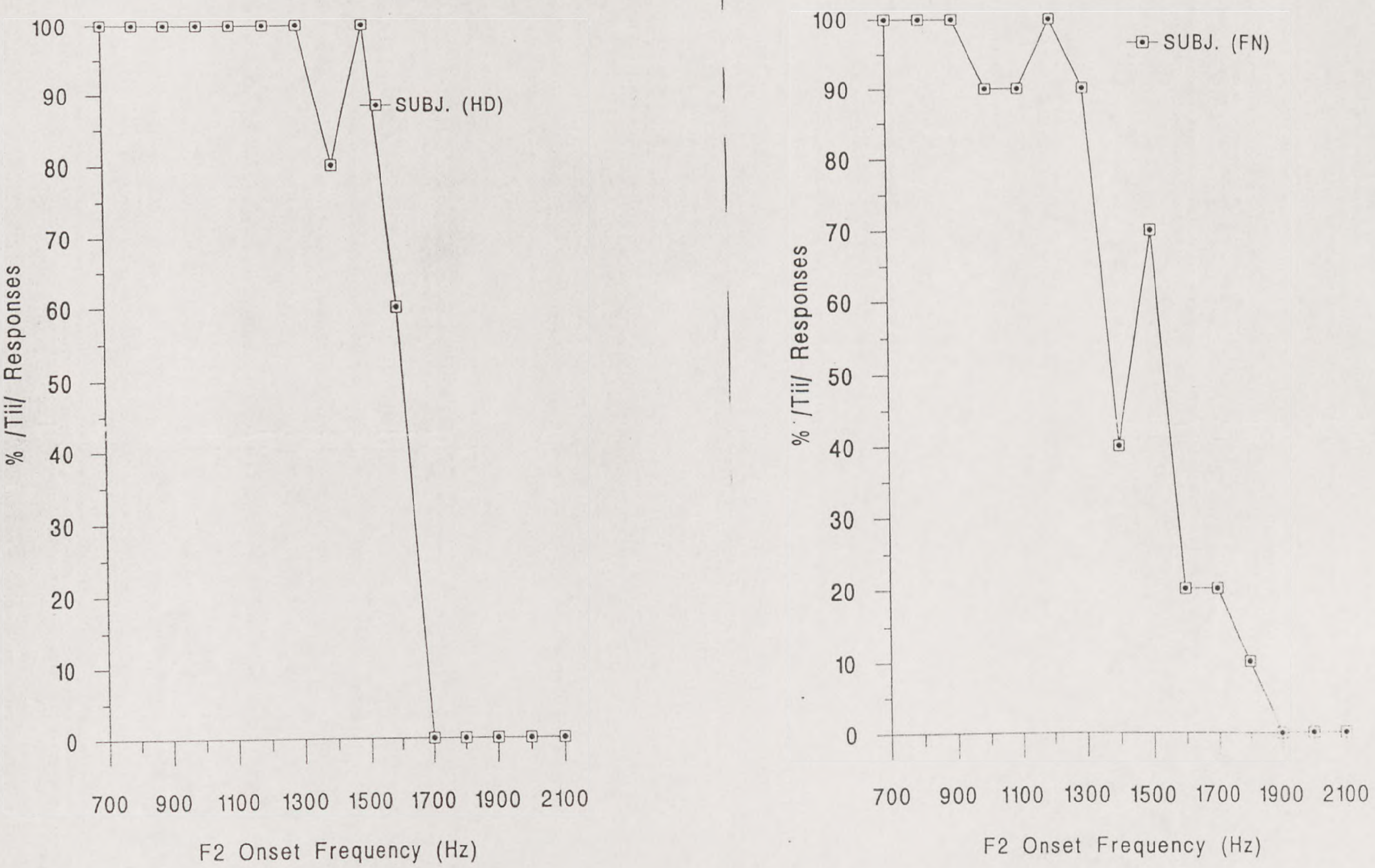
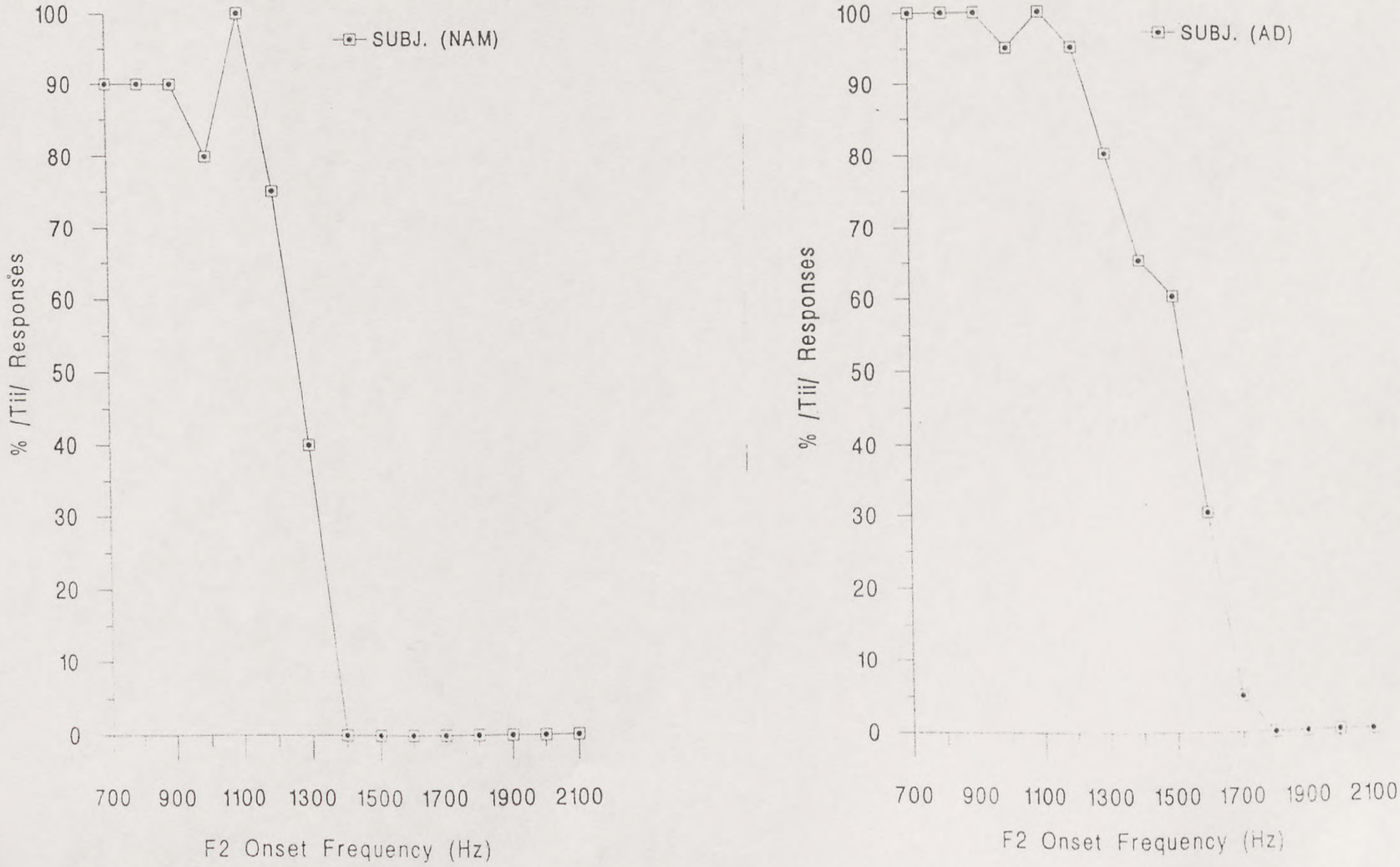
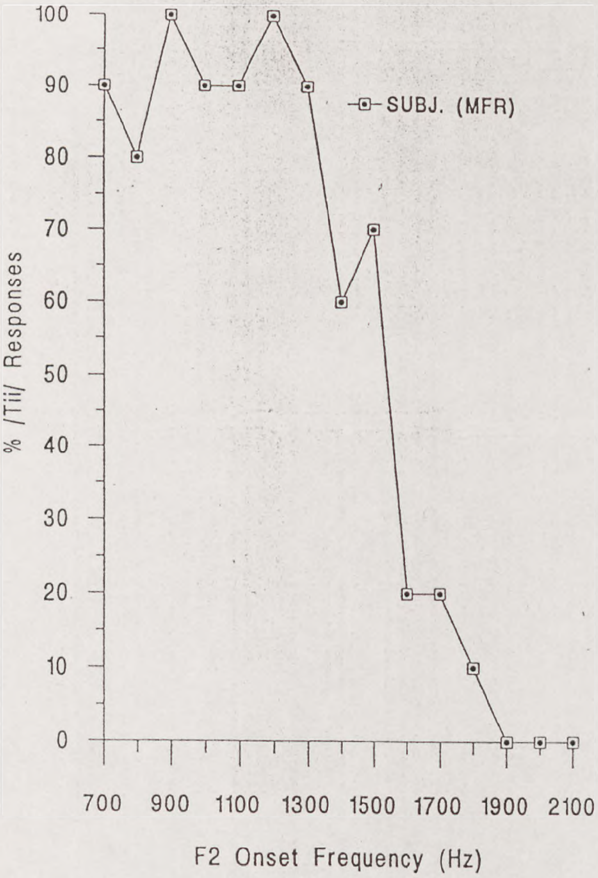


Figure 4.2.5a Percent of the [Tii] vs [tii] responses by the individual subjects for test one stimuli.





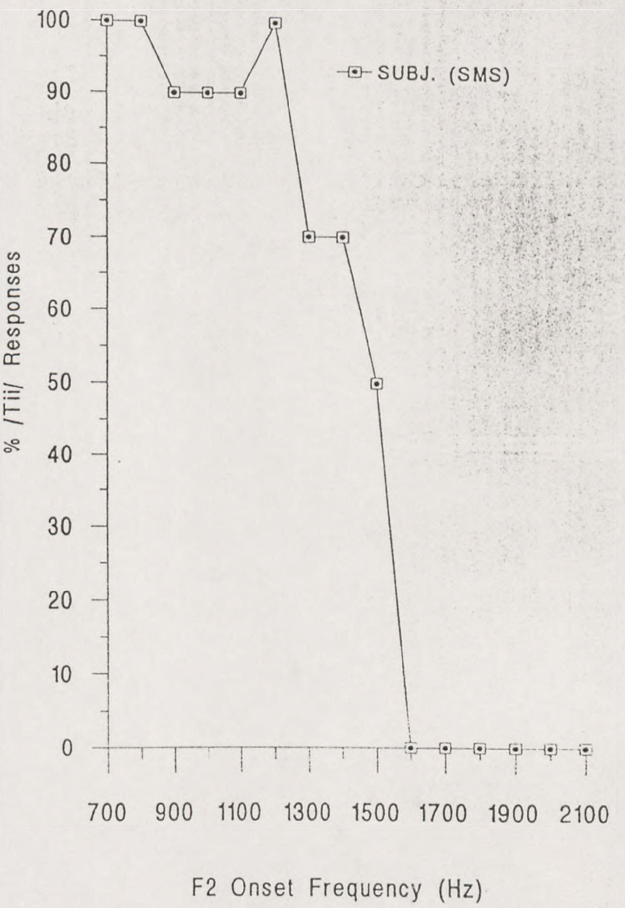


Figure 4.2.5 a (Continued).

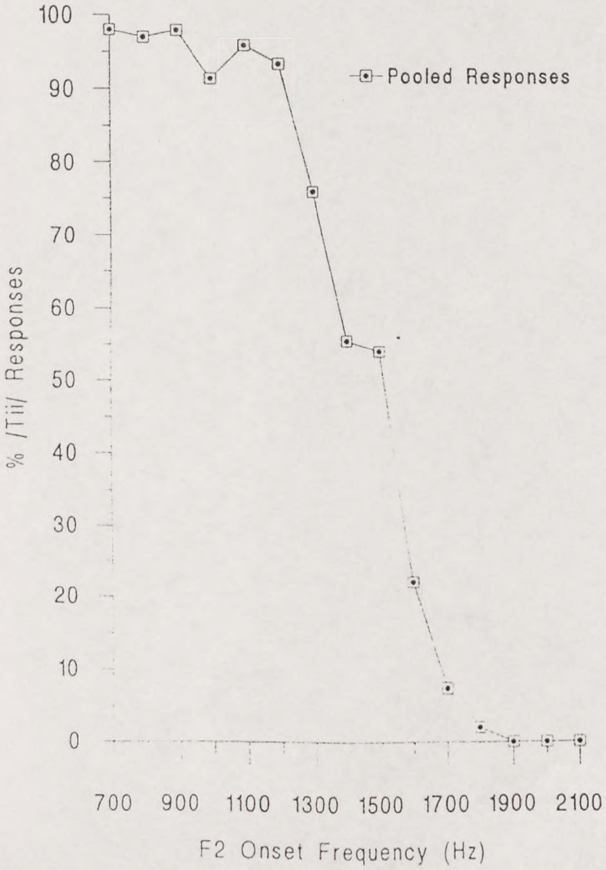


Figure 4.2.5b Percent of the pooled responses to test one stimuli by the ten subjects.

emphatics are perceived at higher ones. There are two possible reasons for such a behaviour: The first is that the region of emphasis is quite wide extending over an extensive range on the frequency scale. The second is that there may be other perceptual cues interacting with the present one and only some subjects seem to be responding to these additional cues. Hazan and Rosen (1991) relate the amount of variability found in the labelling of speech contrasts to cue salience which depends on the speech pattern complexity of the stimuli and on the vowel environment.

Subjects WR and HD who are both females have the highest crossover points in this test. The category boundaries for WR occurred at 1605 Hz and for HD it lies at 1620 Hz. It is not clear why the 50 % crossover points for these two females are higher on the frequency scale than for the male subjects. A possible explanation for this is that female listeners are very likely to produce higher formant frequencies than males. Thus if one accepts that an individual's perception of speech is related to his/her own production then that may give partial explanation to that. To find the average phoneme boundary point for the emphatics and non-emphatics in this test all the responses by the ten subjects are pooled and displayed in Figure 4.2.5b. Table 4.2.3. shows more details of the phoneme boundary values associated with each subject in the /Tii/ vs /tii/ continuum.

Subjects	**The 50% Crossover Points (Hz)**
AA	1274
WS	1500
WR	1605
MB	1500
HD	1620
FN	1451.6
AD	1530
NAM	1275
SMS	1500
MFR	1545
MEAN	1480
S. DEV.	119.5
Range	1274 - 1620

Table 4.2.3 The phoneme boundary at the 50 % crossover points for the individual subjects in the /Tii/ vs /tii/ continuum.

4.2.5.2 Test two: the role of context in the emphatic/non-emphatic distinction.

The responses for this test are treated in the same way as in test one. The results are presented in Figure 4.2.6a in the form of percentages against varied values of F2 steady state frequencies. One observation about these graphs is that the transitional period from one response category to the other is rather sharp. The responses of almost all subjects are associated with only one 50 % crossover point on the frequency scale. In this respect the subjects are quite decisive in their responses. They seem to be more systematic in their categorisation of this contrast than they were in test one. In test one some subjects experienced some difficulty. These can be observed in the responses of subjects HD, FN and MFR. Subject FN, for example, has three crossover points at 1377 Hz, 1427 Hz and 1550 Hz.

In this test the category boundary for the /Taa/ vs /taa/ contrast is found to occur at 1100 Hz for subject AA. For subject WS that point is assigned at an F2 steady state frequency of 1185 Hz. The other subjects show also similar values in the positioning of the 50 % crossover points. The subjects judgements are within the range of 1255 Hz and 1100 Hz. All other points are within these two extremes. Table 4.2.4 presents the 50 % crossover points for all the listeners.

Subjects	The 50 % Crossover point (Hz)
AA	1100
WS	1185
WR	1230
MB	1175
HD	1255
FN	1250
NAM	1100
AD	1205
SMS	1185
MFR	1250
MEAN	1193.5
S. DEV.	57.4
Range	1100 - 1255

Table 4.2.4 The phoneme boundary at the 50 % crossover points for the individual subjects in the /Taa/ vs /taa/ continuum.

These results indicate that the F2 steady state frequency is a sufficient cue for distinguishing the emphatic/non-emphatic categories. All subjects' responses are pooled and displayed in Figure 4.2.6b

4.2.6 Discussion of both perceptual tests on emphasis:

The average F2 onset frequency at the 50% crossover point for the emphatic and non-emphatic categories lies at approximately 1480 Hz for. The range of variation in assigning the 50% crossover point along the F2 onset frequency scale is between 1274 Hz and 1620 Hz. Thus one would limit the region of emphasis between these two extreme values. In real speech the mean F2 onset frequency for the

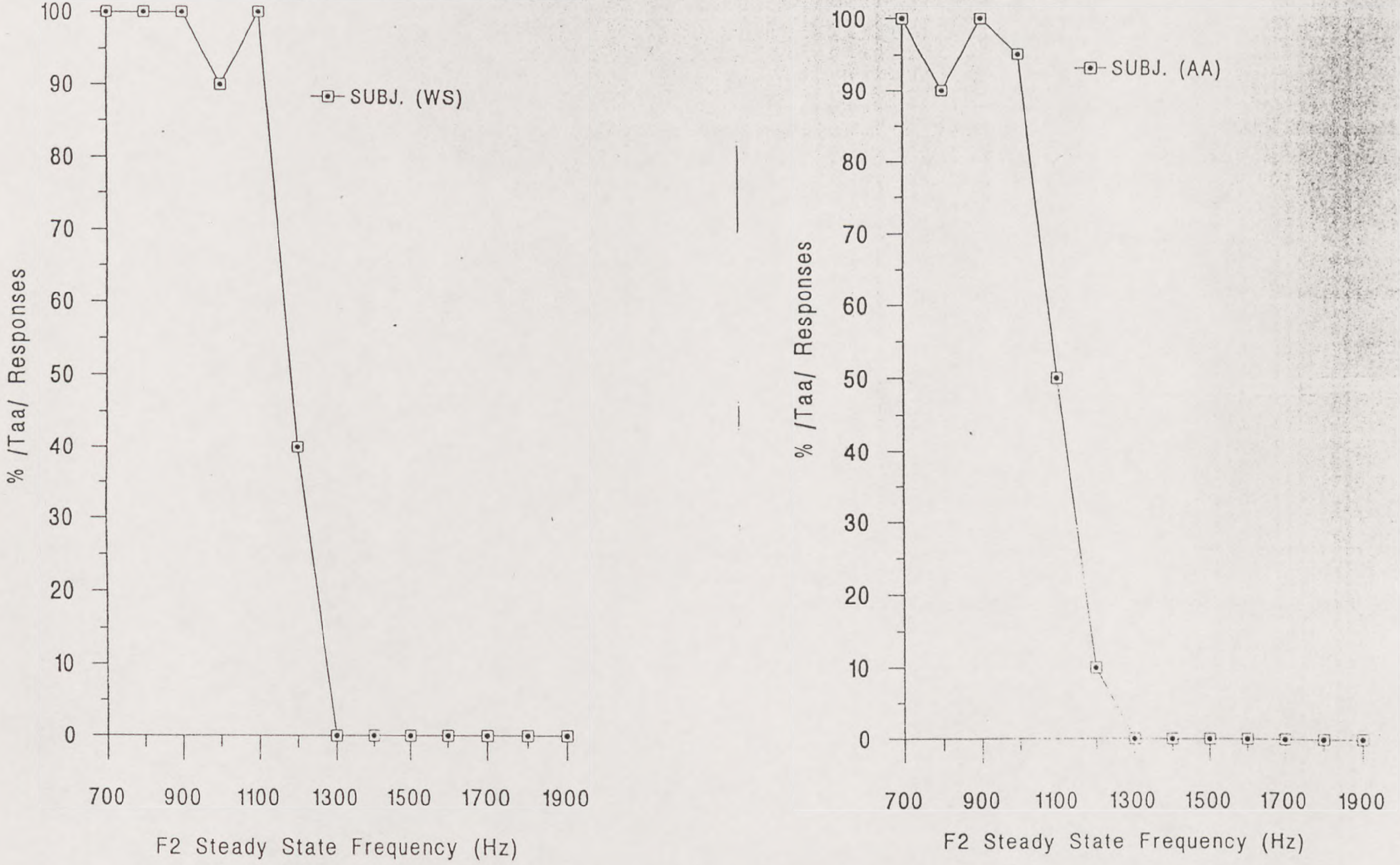
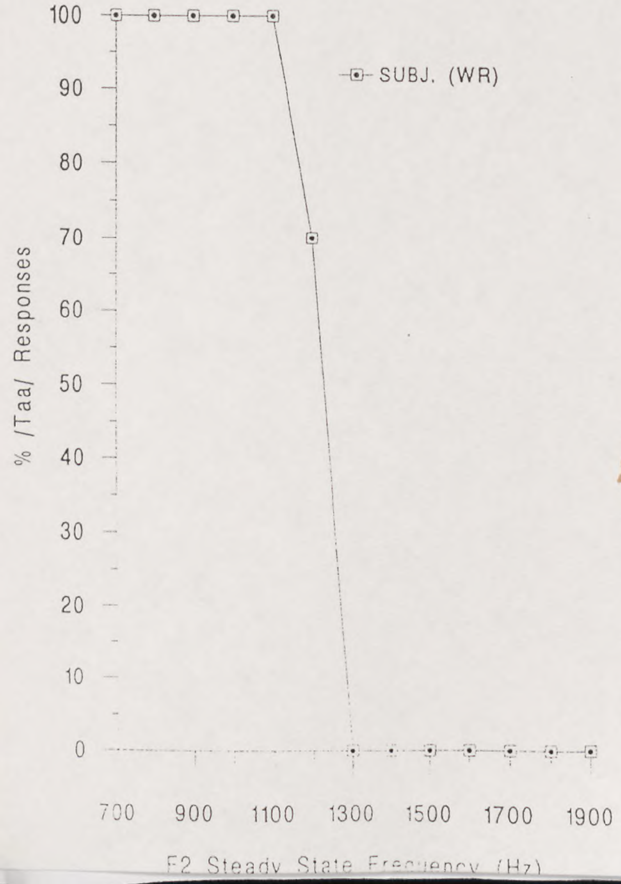
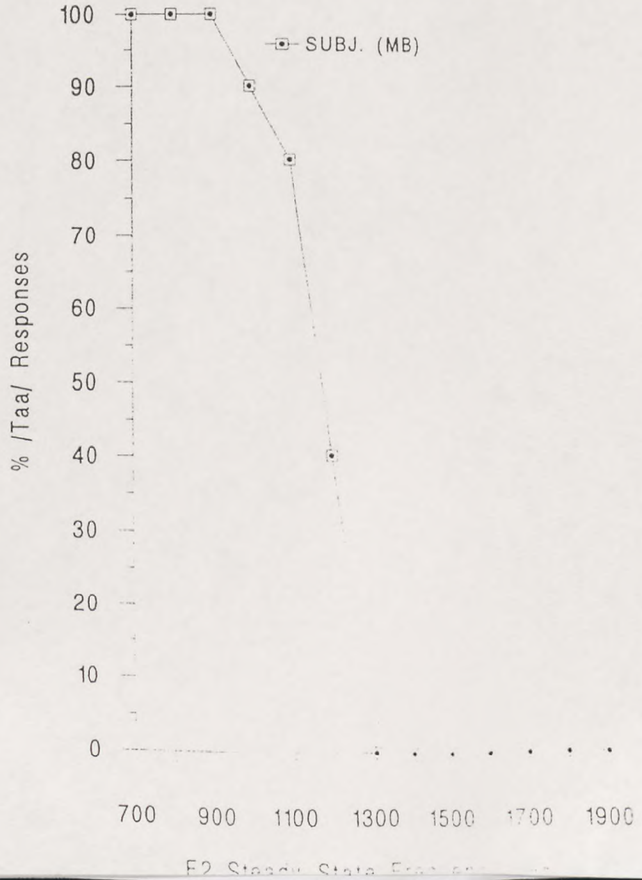


Figure 4.2.6a Percent of the [Taa] vs [taa] responses by the individual subjects for test two stimuli.





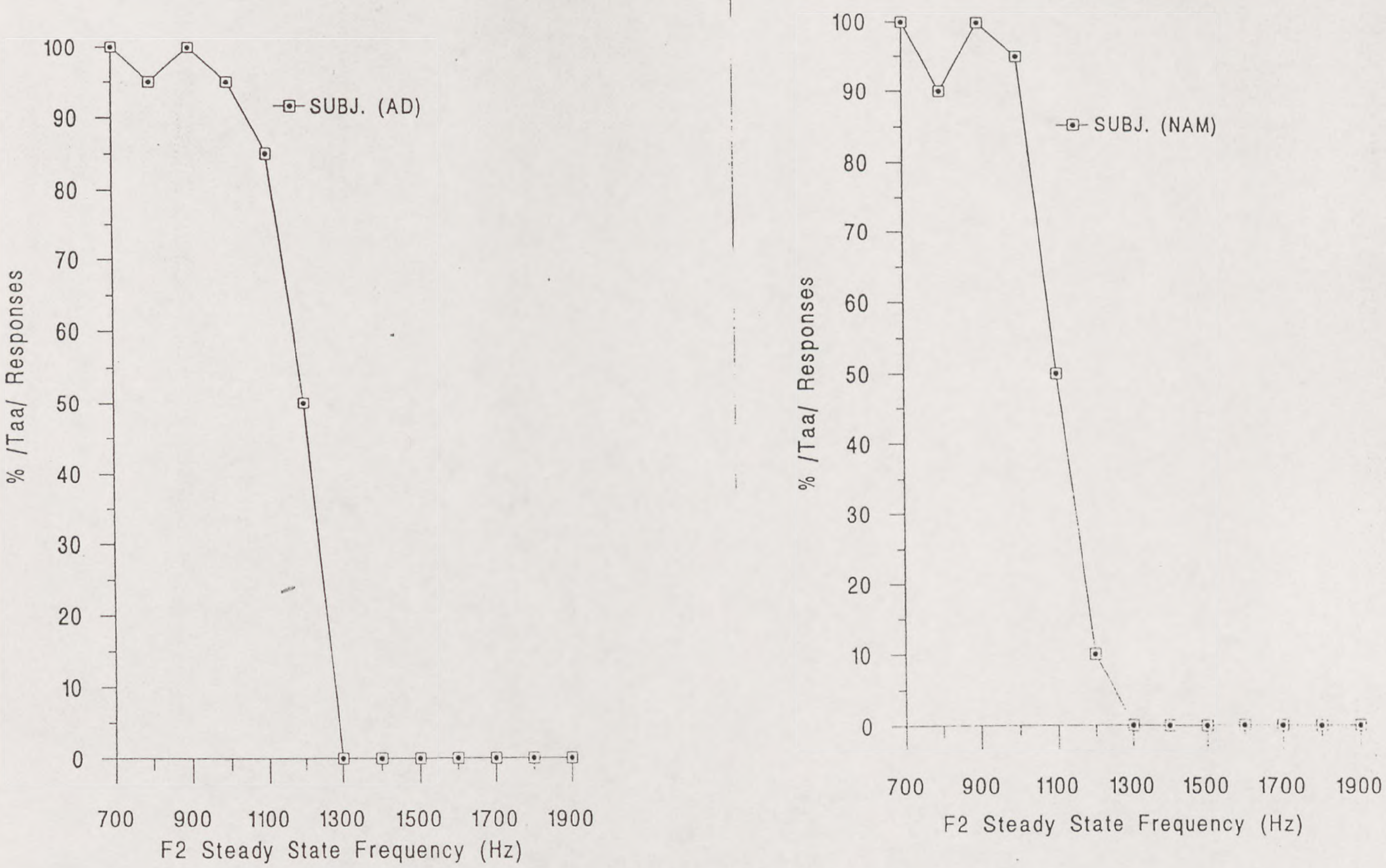
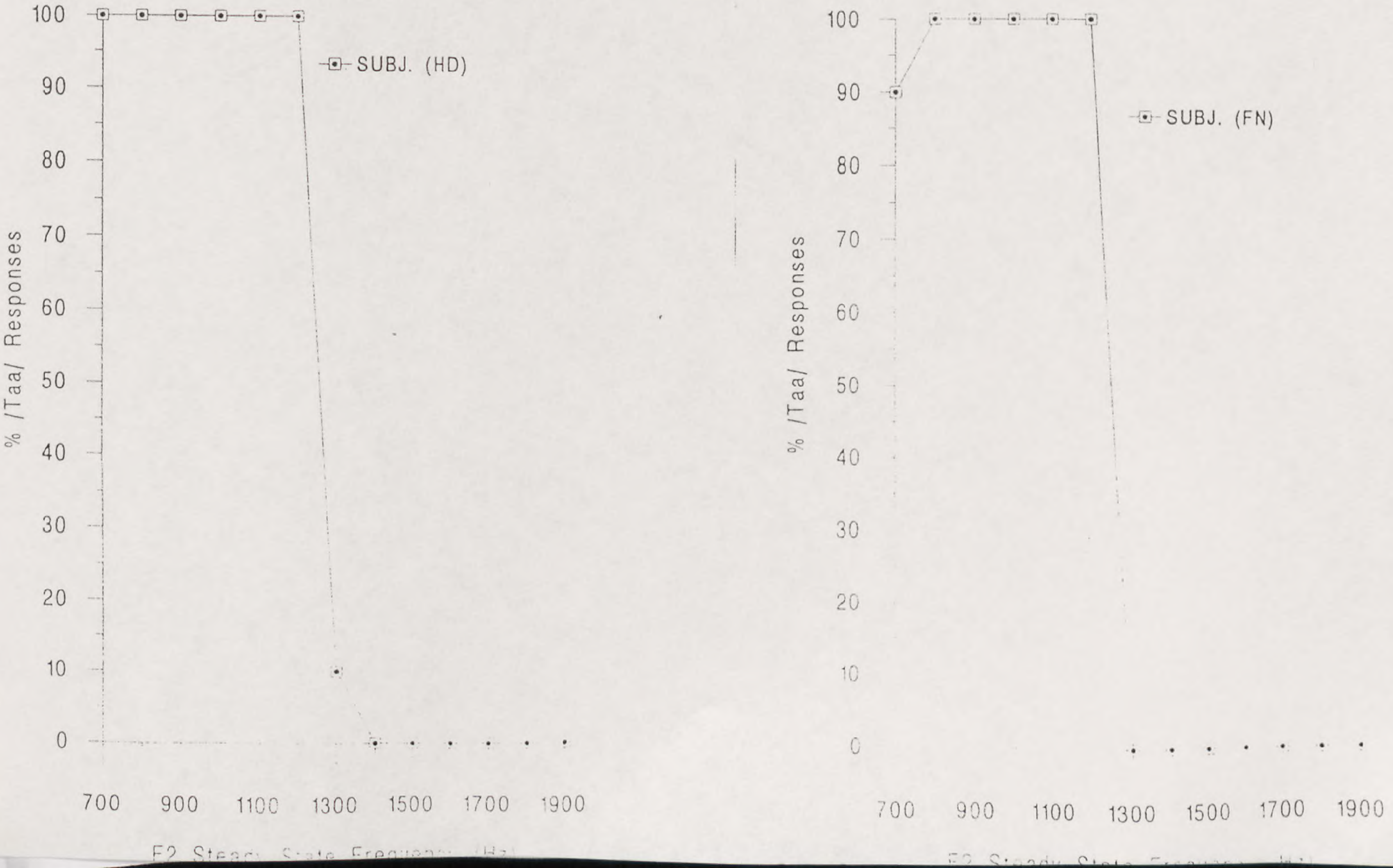


Figure 4.2.6a (Continued).



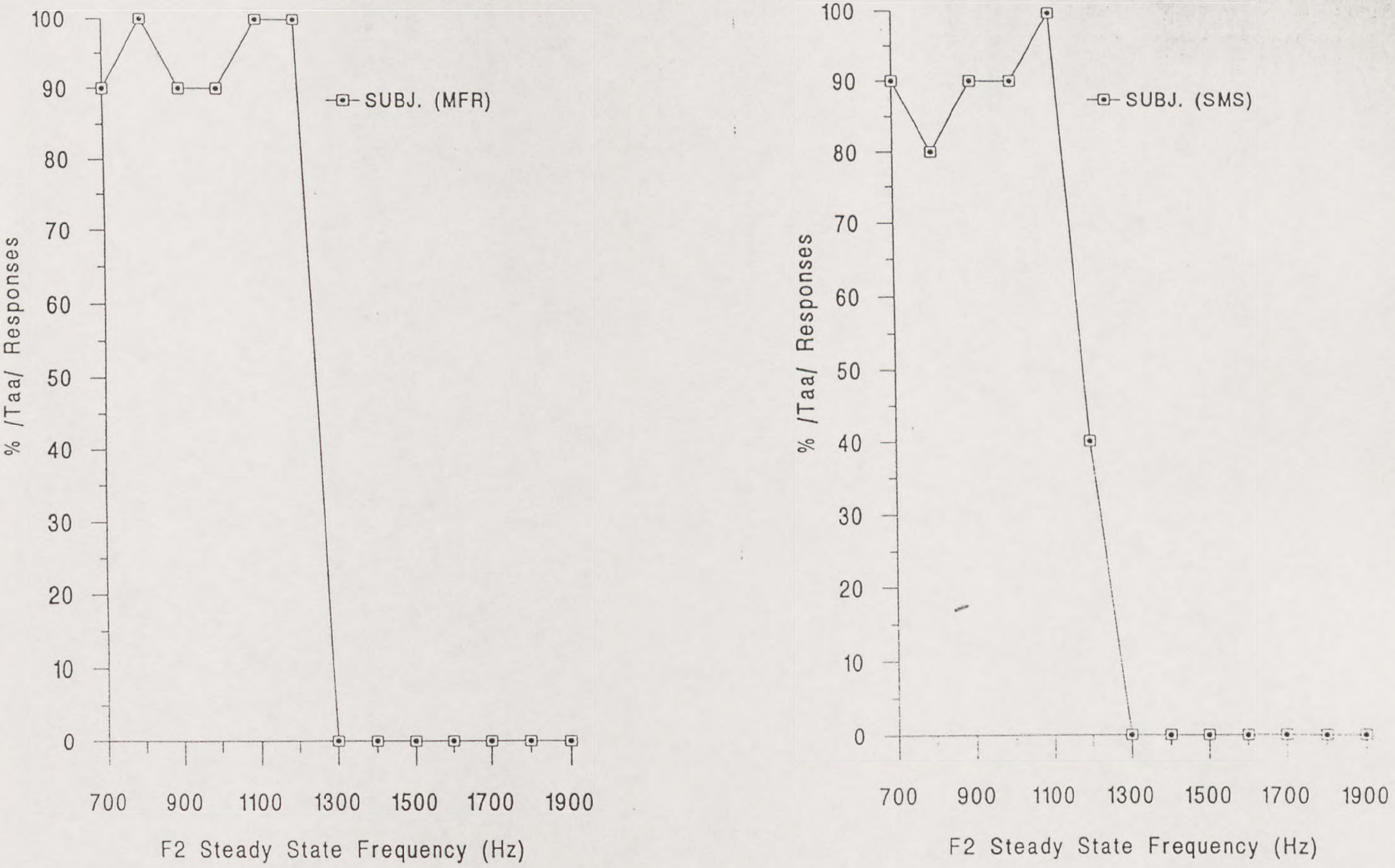
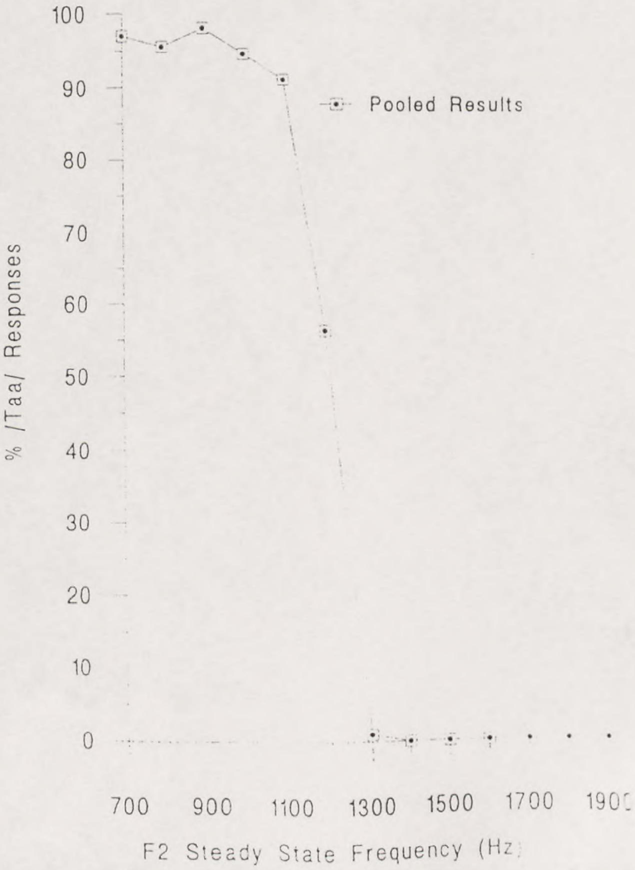


Figure 4.2.6a (Continued).

Figure 4.2.6b Percent of the pooled responses to [Taa] vs [taa] in test two stimuli.



emphatic consonant /T/ lies at 1500 Hz as opposed to a non-emphatic onset frequency value at 2050 Hz (See section 3.2). This suggests a fairly good correspondence between the two values.

Similarly, the average F2 steady state frequency for the emphatic consonant /T/ in the context of /aa/ is heard at 1193 Hz. In actual speech the corresponding mean value for the same consonant vowel syllable is found to be at the vicinity of 1280 Hz. With the non-emphatic counterpart it is found to lie at an average F2 steady state frequency value of 1565 Hz. Again the perception and production results match well, although the relation is not as marked as that of the transitional cue.

These perception results are consistent with that of Obrecht (1968). He found the perceived phoneme boundary for /Tii/ vs /tii/ based on the F2 onset frequency to be at 1450 Hz while the region of emphasis is determined to be between 1000 Hz and 1400 Hz. One difference between the findings of the two studies is that the zone of emphasis occupies a slightly higher region on the frequency scale in this study of YSA than in his study of Lebanese Arabic. Thus, the upper, lower as well as the mean values are all associated with a slight increase in the YSA study as compared with the Lebanese one. These differences may be attributed to subject or dialectal variations. If that is the case then there is a real motive for further examination of another Arabic dialect for comparison and confirmation of the available evidence. Another reason for that difference may well be due to the different synthetic stimuli.

In addition, the results are in agreement with Obrecht's (1968) conclusion that the F2 transition and the F2 steady state work in the same range and produce similar results when manipulated separately. Although the mean 50 % crossover point for the F2 steady state of the emphatic category here is perceived at a lower frequency region than that of Obrecht (1968), it is almost within the identified range of the F2 onset frequency. It is true that the F2 onset frequency occupies a wider range than that of the F2 steady state but the two cues overlap to a great extent.

On the other hand, systematic variations in the consonantal context have demonstrated that the F2 steady state is not less effective and reliable in cueing the emphatic/nonemphatic categories than that of the transitions. With all other parameters kept constant, the change in the F2 steady state is by itself sufficient for the accurate identification of the emphatic and non-emphatic consonants.

Since the creation of an intelligible emphatic consonant requires the lowering of the F2 steady state, it seems that emphasis is not limited to the consonant itself, rather it extends to neighbouring vocalic contexts. This F2 lowering may be attributed to the secondary articulation. One of the correlates of this phenomenon is the narrowing of the tongue passage which in its turn causes a down shift of the F2. Since the F2 lowering has occurred for both consonant and vowel one may speculate that the narrowing of the tongue passage associated with the emphatic consonant may be present during the realisation of the vowel as well. The mutual consonant-vowel influence is a well-known phenomenon. Fant (1960) draws a distinction between two such effects: those caused by the tongue movements during consonant-vowel or vowel-consonant articulation and those related to the release of the tongue during a consonantal closure or the movement towards closure. As far as the results of this study are concerned the first type seems to be more responsible for such an effect than the second type particularly in the vowel context /aa/ which has two distinct variants each occurring with only one category to the exclusion of the other. It has a backed variant following emphatics and an open fronted one following non-emphatics.

The perceived effect of context in YSA is in agreement with that of Ali and Daniloff (1974). Although the two studies have approached the phenomenon differently, they emphasize the role of context in that distinction. They share the conclusion that the domain of emphasis is not limited to the emphatic consonant but rather extends to the following vowel. In our view, these findings are quite relevant for the theory of distinctive features. The feature of emphasis appears to operate on both consonant and vowel. According to the distinctive features' formulisation by Jakobson and Halle (1956) regarding the binary classification, the consonants and vowels are regarded as belonging to separate features 'consonantal/non-consonantal' or 'vocalic/non-vocalic'. However, the above finding suggests that both the emphatic consonant and its vocalic context are dominated by one and the same feature (i.e. emphasis). Thus a specific acoustic feature may exist in more than one phonetic segment.

It is our assumption that the addition of a component of emphasis to the normal YSA non-emphatic plosive will result in some articulatory and perceptual adjustments. The preparation of the speech stimuli has provided some proof of that. It is found that they require a slightly longer time for realization. The non-emphatic plosive /t/ is heard at a transition duration of 50 ms while its emphatic counterpart /T/ is not detected by the listeners below 65 msec. The more the transition duration value is increased up to a value of 80 msec, the better the intelligibility of the emphatic sound becomes. This is particularly true of those emphatics in the context of /ii/ which display a very extensive transitional movement. A possible reason for the occurrence of such a phenomenon may be attributed to the long distance separating the lower onset frequency associated with the emphatics and the high F2 steady state frequency of the vowel. Thus, the value of formant transitions as perceptual cues for consonants depends on the magnitudes of the formant movements, which in turn depends on the articulatory positions of both the consonant and vowel.

In spite of what has been said about the efficiency of the F2 transitions and steady state frequencies, it is not implied that either of these features alone can be a sufficient cue for the emphatic/nonemphatic distinction in isolation from the other components of the speech signal. On the contrary, it is the interaction between these different aspects, but with the F2 component playing the dominant role. This opens up the possibility for contemplating the possible effect of other variables such as the release burst, and the relative interval between the burst and the onset of voicing. The likelihood of this intervening effect being important seems to be increased by the fact that some subjects show more than one crossover point particularly in test one. Consequently, to enhance maximum intelligibility in the emphatic/nonemphatic distinction these additional cues need to be integrated with other cues in the synthetic stimuli and in the appropriate manner. It has been shown, for example, that the integration of cues in the perception of phonological contrasts is a requirement for the accurate identification of plosives (Hazan et. al. op. cit; Cooper et. al. (1970) Dorman, Studdert-Kennedy and Raphael, 1977).

To end this disscusion a remark about the limitations of this study should be made. The region of emphasis mentioned earlier is based on the examination of the plosive consonants /T/ and /t/ in the context of /aa/ and /ii/. Although these consonants are assumed to be representative of the contrast under investigation, there are other consonants which if tested may add further information. A comprehensive perceptual study should include more emphatic/nonemphatic consonants (e.g . /D, T,q/ vs /d, t, k/) in the context of long and short vowels. Such a study may add further information regarding the region of emphasis. More importantly, the domain of emphasis may not only extend to the following vowel, as found in this study, but also may go beyond to the following syllable and perhaps the whole word.

Chapter 5: Articulatory and aerodynamic study of emphasis and voicing in YSA

5.1.1 Introduction:

While emphasis and voicing as represented by the plosive categories of YSA are phonologically important there appears to be several uncertainties in the articulatory description of this group of sounds particularly the emphatic ones. The need for this kind of study is also increased for the reason that most of the aerodynamic studies, as shown in section 2.3, are concerned with English. Very few studies can be found on MSA and none at all on YSA. Thus in an attempt to increase the inventory of this type of study on other languages an investigation of the plosives of YSA will be conducted. The study is divided into two parts aimed at examining the aerodynamic characteristics of both features (i. e. voicing and emphasis).

5.1.2 Aims:

5.1.2.1 Emphasis:

The main assumption underlying this work is that the term emphasis describes a linguistic phonemic feature. There is evidence (see section 2.3) for an articulatory difference between the emphatics and nonemphatics. This aritculatory difference may give rise to aerodynamic differences which we can observe. It is hoped that any systematic variations which we may observe will clarify our understanding of the articulatory movements involved.

Aerodynamics links articulation to sound sources, so by studying aerodynamic patterns we may be able to explain some of the acoustic differences found.

Another aim of this study is to determine the effect of the vocalic contexts on the peak oral air pressure (Po) and the peak oral airflow (Uo) values for the emphatic/non-emphatic distinction. It is predicted that the emphatics in the environment of /aa/ might show lower rates of airflow and probably longer articulatory closure duration than those in the environment of /ii/.

In addition, the effect of gemination on the peak oral airflow and air pressure values for the emphatics and the non-emphatics will be examined. Material prepared for this purpose has been included as shown in Table 5.1.2. It is expected that the geminated emphatic and non-emphatic contrasts will show similar patterns to those found in single forms.

The study also contained material for the investigation of the effects of segmental phonetic position (i. e. word-initial and medial position). Previous studies showed apparent disagreements regarding the effect of this contextual aspect. However, none of them was conducted with emphasis in mind. The assumpiton behind including this variable was motivated by the claim of many researchers that pressure variation was not only influenced by voicing, but also by other contextual factors the most important of which was phonetic position.

5.1.2.2 Voicing:

The main aim of the present experiment is to investigate the aerodynamic patterns and the durational relationships for the voicing distinction in YSA. In view of the apparent contradictions between the findings of different studies regarding the effectiveness of oral pressure in distinguishing this contrast it is believed that a study of these aerodynamic patterns should be quite useful in characterising the articulatory behaviour of the speakers for this distinction. The contradictory results of the previous studies suggest that the language variations as well as the speaker variations cannot be ignored and that the same linguistic feature may show different phonetic characteristics in different languages.

Another related aim of this study is to look at the effect of context on the air pressure and airflow patterns for the voicing categories. This includes position (i.e.initial and medial), vocalic contexts (/ii/, /aa/ and /ʊʊ/) and gemination. It is hoped that such treatment of phonetic context will enable us to establish the relative discriminatory power of each of those factors on the aerodynamics of the voicing cognates. It may or may not be that those patterns will show any difference for the voicing distinction in the two positions, the three vocalic contexts or in geminated forms.

In addition, this study will attempt to investigate the place of articulation effect on the oral pressure and the oral flow in YSA. As indicated in section 2.3 there are not many studies on this phonetic aspect. It should be interesting to see how the place of articulation interacts with the aerodynamic variables as well as with the articulatory duration, particularly that of the closure in YSA.

5..1.3 Instrumental set-up

The experimental set-up is shown in Figure 5.1.1.

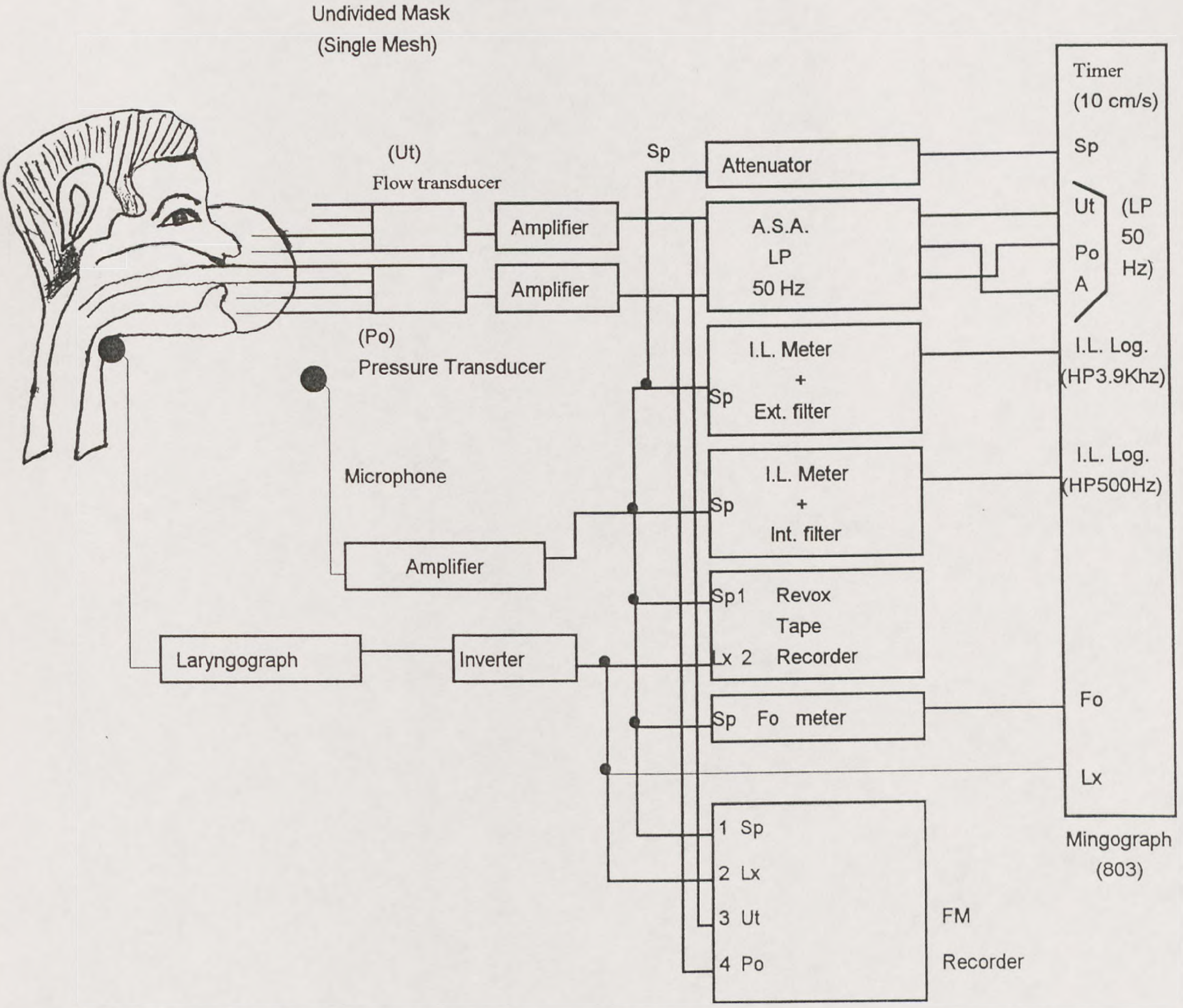


Figure 5.1.1 Experimental set-up used for the study of aerodynamic data of emphasis and voicing in Yemeni Spoken Arabic.

The apparatus used included the following:

1. Rothenberg mask (an undivided mask, single mesh, Glottal Enterprises).

2. Aerodynamic Speech Analyser (A.S.A.).

3. Rothenberg flow transducer.

4. Mingograph (Siemens Elema, 803).

5. FM tape recorder.

- A back up copy of the airflow, air pressure , Lx and the sound pressure signals were kept on tracks three, four, two and one respectively. The input levels to these tracks were at two for the airflow signal and at one for the other three signals.

6. Polyethylene tube.

7. Gaeltec pressure transducer

8. Gaeltec pressure amplifier

9. Reference tube.

10. Laryngograph.

The laryngograph's input level was set at three o'clock.

11. B & K condenser microphone (type 4134) with the B & K sound level meter (type 2231) used as a pre-amplifier.

12. Revox tape recorder (Dep. no 277):

- An audio version of the Lx signal was recorded on track two of a high quality open reel tape (type Ampex 1200). The auxiliary input to this track was set at five and the output level at eight. The balance was at twelve o'clock.

-The sound pressure signal (Sp) was recorded on track one. The input to this track was set at four. The output level was at eight.

13. Step attenuator (type A64 600 Ohms Impedence).

14. Fundamental frequency meter (type FFM 650 linear scale 5 volts; F-J Electronics).

15. I.L. meter with external filter for I.L. trace High Pass filtered at 3.9 kHz (F-J Electronics).

16. I.L. meter with internal filter for I.L. trace High Pass filtered at 500 Hz (F-J Electronics).

5.1.4 Method and procedure

5.1.4.1.Oral Airflow:

Traces for airflow were obtained by fitting a mask to the subject's face. The speakers' face was pressed against the mask with the aim of preventing air leakage. Initially, a divided mask was used. The idea behind using this separated mask was to check whether the two subjects were separating nasal and oral airflow in their speech.

The separated mask was constructed of relatively hard rubber constructed from two British oxygen anaesthetic masks. Two flowheads were fitted to the mask: one for oral airflow and the other for nasal airflow. Each of the flowheads was connected to Gaeltec pressure transducers type (8T ± 2 cm H2 O). The two parts combined formed what is normally called a 'pneumotachograph unit'.

The outputs of both transducers were passed into a Gaeltec control unit, amplified and low passed filtered at a cut off frequency of 50 Hz in the aerodynamic speech analyser. The subjects were asked to breathe out once through their noses only and another time through their mouth only. In addition, the subjects were asked to read three runs of the test words in isolation and three more runs in a carrier sentence. Nasal airflow and oral airflow traces were recorded on two separate mingograph channels. The results of the preliminary checks showed that oral and nasal airflow were well separated in the speech of the two subjects and that no nasal airflow was observed in the production of non-nasal voiced and voiceless plosives, as well as emphatic and non-emphatic plosives in YSA.

On the basis of these results and the fact that no nasal sounds were investigated in this study, it was decided that an undivided single mesh (Rothenberg mask, 1977, Glottal Enterprises) would be more appropriate for our purpose than the separated mask for the following reason. Since it contained a large mesh area it gave a good acoustic output with distortion minimized. The total airflow from the mouth and nose combined was measured and the total flow for a certain speaker was equated with oral airflow.

The mask covered the speaker's mouth and nose. It was fitted with only one flowhead. The open end of it was connected to a Gaeltec transducer (8 T ± 2 cm H2 O). The signal went through an amplifier and was passed into an aerodynamic speech analyser (R. Caley) where it was low pass filtered at 50 Hz. It was then fed into the mingograph and displayed as channel two.

An unfiltered version of the oral airflow signal was transmitted to the FM tape recorder where a back up copy was recorded on the Ampex high precision magnetic tape.

5.1.4.2 Oral air pressure

Oral pressure traces were obtained by a small polyethylene tube open at one end and inserted into the mouth. The external diameter of the tube was about 3 mm and its internal diameter was about 2.5 mm. The choice of this pressure tube came after several trials of different tubes varied in their length and shapes. This tube fitted both subjects.

The curved open end of the tube was put behind the constriction pointing downwards across the airstream near the midline in the oral cavity. This position gave good air pressure results for the bilabials and denti-alveolar plosives: emphatics and nonemphatics, voiced and voiceless. This position, in addition, should "...minimize stagnation pressure effects" Scully (1990, p165).

From this position it was not possible to extract pressure values for the velar consonants /k, g/ and the uvular sound /q/. The constriction points for these consonants occurred further back in the mouth behind the pressure tube opening. Consequently no pressure values were obtained for these sounds.

The polyethylene tube was connected to a Gaeltec pressure transducer No.1 (type 3 CT ±10 cm $H2\ O$). The output signal was amplified by a Gaeltec pressure amplifier whose output gain was (70). The signal was then transmitted to an Aerodynamic Speech Analyser (R. Caley) where it was low-pass filtered at a cut off frequency of 50 Hz. It was then fed into the mingograph and displayed as channel three on the mingogram.

The reference tube had its open end inside the mask (in front of the speaker's face). In this way the difference (drop) of air pressure across an oral constriction (ΔP) was obtained.

5.1.4.3 Area

Traces for the minimum area of constriction (A) were obtained from the aerodynamic speech analyser. It was an analog-computer device that uses inputs of oral airflow (Uo) and pressure across a constriction (ΔP). It divides the signal for Uo by the square root of the signal for (ΔP) to give an output A. This A trace is an indication of the minimum cross-section of the area of the constriction in front of the open end of the pressure tube. The orifice equation:

$A = K\ Uo / \Delta P$ is assumed to have an empirical constant K of fixed value (Scully, 1990 p163).

The area signal was processed in the aerodynamic speech analyser and was low pass filtered at 50 Hz. It was then displayed by the mingograph on channel four. Since we are concerned with the plosives the minimum area of constriction is expected to be at its lowest point reflecting the complete closure of plosives in YSA. Some speakers are reported to have produced plosives with partial closure (Lisker and Abramson, 1967) or with no release at all (Miller and Daniloff, 1977). It would be interesting to see whether similar results could be found in our data for these YSA plosives.

5.1.4.4 The laryngograph signal (Lx):

Mingograph traces of the Lx signal were obtained by placing two electrodes (Laryngograph Ltd) horizontally on both sides of the external surface of the subject's larynx. The electrical resistance presented by the larynx was transmitted via the two electrodes to the laryngograph. The (Lx) signal was then inverted to match the signal from the microphone (See Figure 3.1.1). It was then displayed on channel eight of the mingograph.

5.1.4.5 The sound pressure signal (Sp):

Additional traces for the sound pressure signal (Sp), the fundamental frequency (F0) and the intensity level signals were all obtained from a B & K condenser microphone (type 4134) placed just outside the mask. The microphone was placed so as to avoid direct impingement of air from the speaker. At the start of the recording the B & K portable sound level meter (type 2231) the microphone amplifier showed a full scale deflection of 109.9 dB. The trace for the sound pressure signal (Sp) was obtained as follows:

The signal passed through a step attenuator (type A64 600 Ohms impedence). This later one was set at 26 dB. The signal went through the mingograph unit and was displayed on channel one. Two recorded versions of the sound pressure signal were obtained: one on the Revox tape recorder and the other on the FM tape recorder.

5.1.4.6 The Fundamental Frequency (F0):

The amplified microphone signal was passed through a fundamental frequency meter (F-J Electronics) where it was low passed at 120 Hz. The input level to the F0 meter was 1.75. The signal was then transmitted to the mingograph and displayed on channel seven.

5.1.4.7 The intensity level signals:

The amplified microphone signal was fed into two intensity meters (F-J Electronics). The signal was processed in an intensity level meter with a channel external filter (ref. no. 334) where it was high pass filtered at 3.9 kHz. The amplitude of this meter was 12 dB. The output signal was displayed on channel five.

At the same time the signal went through an intensity level meter with internal filter (ref. no 364) where it was high pass filtered at 500 Hz. The amplitude reading of this meter was also 12 dB. The signal was displayed on channel six. Both intensity level traces of this signal were produced with a log calibration scale.

Prior to the actual recording all eight mingographic channels were synchronised. The Mingograph (type Siemens Elema 803) was running at a speed of 10 cm/sec. Its output channels displayed on the mingograms were set to the following gains:

CH.1 at 7

CH.2 at 2.5

CH.3 at 6

CH.4 at .75

CH.5 at 4.5

CH.6 at 4

CH.7 at 5

CH.8 at 6

All materials were recorded with the same instrumental set-up for the two subjects. Calibrations for the oral airflow, oral pressure, intensity and fundametal frequency were made immediately after the recording was finished. The airflow recording system was calibrated against a rotameter and the flow values were expressed in cm^3/sec. The oral pressure recording system was calibrated with a U-tube manometer. The peak oral pressure measurements were expressed in centimeters of water (cm H2 O). This was done for both speakers. The calibration was recorded on the FM tape for later reference.

There were 28 test words in the emphatic/nonemphatic investigation, while in the voicing experiment there were 36 test words in the three vocalic contexts. They were read six times by each of the two subjects.

The material for both the emphatic and the voicing contrasts in the different contexts were randomised in six lists and were all recorded in one session. They were read by two subjects. All the words were embedded and uttered in the carrier sentence /ʔiqraʔaa _____sariiʕun/ 'read (you two) -------soon'. Each of the six runs began and ended with a dummy word. As part of the preparation for the aerodynamic recording, the material for emphasis and voicing was recorded on high quality open reel tape (type Ampex) in the studio of the Department of Linguistics & Phonetics. The speed of the tape for this initial recording was 7 ½ per second and the microphone used was AKGD 202 E1.

An auditory check was done by playing back the recording for all the material to make sure that the phonemes recorded were those intended. This sample was for use with the actual recording as a guide for the subjects. During the recording a gap was left after each sentence to give the subject a sufficient time to repeat it. By playing back this recording the subject was required to listen and repeat the sentence. It was hoped that this procedure would help to control the accentual patterns of the subjects as well as the speaking rate. Another aim of this recording was to get as much identical conditions as possible for the two subjects. In this way some of the extraneous variables would be controlled and the results obtained would allow valid comparisons.

5.1.5. Materials:

As far as emphasis is concerned, the sample consists of word-initial and word-medial plosives. The words in the latter position included consonants in geminated and single forms as shown in Tables 5.1.1 and 5.1.2

Word-initial		Word-medial	
emphatics	non-emphatics	emphatics	non-emphatics
Taab	taab	baTaaT	bataat
qaad	kaad	baqaaʔ	bakaaʔ
Tiir	tiil	faTiir	fatiil
qiis	kiis	faqiir	bakiil
Tuub	tuub	-	-
quut	kuut	-	-

Table 5.1.1 Word-initial and medial emphatics and non-emphatics.

Emphatics	Nonemphatics
baTTaat	fattaat
baqqaal	dakkaat

Table 5.1.2 Word-medial geminated emphatics and non-emphatics.

All the emphatics and the nonemphatics in this experiment were voiceless. They included /T/ versus /t/ and /q/ versus /k/. In word-initial position both emphatic pairs were followed by three long vocalic contexts: /aa/, /ii/ and /uu/, whereas in word-medial position they were only followed by the vocalic contexts /ii/ and /aa/ as can be seen from Tables 5.1.1 and 5.1.2. In the whole set there were a total of 20 words representing the emphatic contrasts in the two positions. The material also comprised four more geminated test words for the emphatics and the nonemphatics.

The voicing contrasts are investigated in word-initial and word-medial also in geminated consonants. The pairs investigated are /t/ versus /d/, /T/ versus /D/ and /k/ versus /g/. In word-initial these consonants were all followed by the three long vocalic contexts /ii/, /aa/ and /uu/ whereas in word-medial they were only followed by the vocalic contexts /ii/ and /aa/ (See Tables 5.1.3 and 5.1.4 below). In all there were 36 test words: 30 test words for the consonants in word-initial and word-medial positions and 6 more in geminated contexts.

Word-initial		Word-medial	
Voiceless	voiced	Voiceless	Voiced
taab	daar	fataa?	badaa?
Taar	Daar	xaTaa?	raDaa?
kaad	gaad	bakaa?	bagaa?
tiil	diir	fatiil	badiil
Tiir	Diiq	faTiir	faDiil
kiis	giis	bakiil	fagiir
tuub	duur	-	-
Tuur	Duur	-	-
kuut	guut	-	-

Table 5.1.3 Word-initial and medial voiced and voiceless plosives.

Voiceless	Voiced
fattaat	saddaad
baTTaal	waDDaah
dakkaat	baggaal

Table 5.1.4 The voicing categories in geminated forms.

All the words were meaningful and in minimal or near minimal pairs. In some cases the word-final consonants differed. Nasal sounds were avoided in our test words for experimental reasons. All the consonants in all the positions and contexts investigated occurred at the start of a stressed syllable. The same words used in our acoustic analyses were also used in this study except where the control conditions required otherwise (e.g. replacing nasal sounds by non-nasal ones). This implied a change not only in the words containing the nasal sounds but also in words containing their corresponding voicing cognates. The carrier sentence was included partly for naturalness and partly to keep as many variables as possible under control. It would help maintain normal stress and intonation patterns.

5.1.6 Subjects

Two native speakers of Yemeni Spoken Arabic served as the subjects of this study. Both subjects had a similar and widely used accent in Yemen. Subject two (WS) was twenty years old. He spent most of his life in Yemen. He came to Leeds in October, 1991 to continue his education.

As far as experimental phonetics is concerned, he is a complete novice. He had never been involved in any work similar to this before. For details about subject (AA) see section 3.1.3.

Both subjects are believed to have normal hearing and speaking habits and had not been treated for any known speaking or hearing illness.

5.1.7 Segmentation and measurement criteria

Segmentation for the purpose of aerodynamic and articulatory measurements was based on oral airflow and oral air pressure traces but other traces were used as indicators of the beginning and end of a certain speech segment e.g. the Lx , Sp and the area traces. Three measurements were related to the oral airflow trace. These were the peak oral airflow, closure duration and the duration of the articulatory release. The peak oral airflow point (Uo) was measured at the time of consonantal release. This was identified by the sudden rise of the airflow trace following a period of air closure. To locate the peak airflow point a baseline was drawn. The baseline and the highest point on the flow curve at the time of consonantal release were then connected. The peak oral airflow value was taken as the distance between these two points (a-b) on Figure 5.1.2.

For the sake of consistency in our measurements of articulatory durations certain criteria were adopted. In order to measure the duration of a certain articulatory segment on the flow trace two points were identified. They marked the beginning and end of this articulatory segment respectively. For example, the onset of the articulatory closure was characterised by a downward deflection. This was taken to indicate the cessation of the final vocoid in /?iqra?aa/ which preceeded the plosive consonant. In this way the closure onset was regarded as the point when airflow trace returns to its original baseline on the airflow trace point (a) in Figure 5.1.3. The closure termination point was regarded as that position where airflow initiated its rise from the baseline i. e. release point (b). The distance between (a) and (b) was measured as the articulatory closure duration in (msec) .

At the same time the termination point of the closure duration was regarded as the starting point of the airflow release. The onset of the articulatory release was defined as the point from which the airflow initiated its sudden rise while the termination point of the release was the one where the following vocoid began (point c on the airflow trace). The start of the vocoid itself was identified by the semi-regular movements of the sound wave on the Lx trace indicating the vocal folds vibration for the vocoid. It was also characterised by the small ripples on the flow trace. The sound pressure traces were also quite useful in differentiating between voicing as well as the emphatic contrasts. Thus the articulatory duration of the release was measured as the interval between points (b) and (c).

Underlying this method of segmentation and measurements was the assumption that the reduced oral airflow rate at the time of closure corresponded to the time at which the vocal tract constriction gradually got more closer till it reached complete closure. The airflow rate at this stage was at the zero point on the airflow trace. The peak oral pressure rise behind the oral constriction would also be expected to be near its maximum. It was also assumed that the sudden release of airflow would be a reflection of the very rapid separation of the articulators.

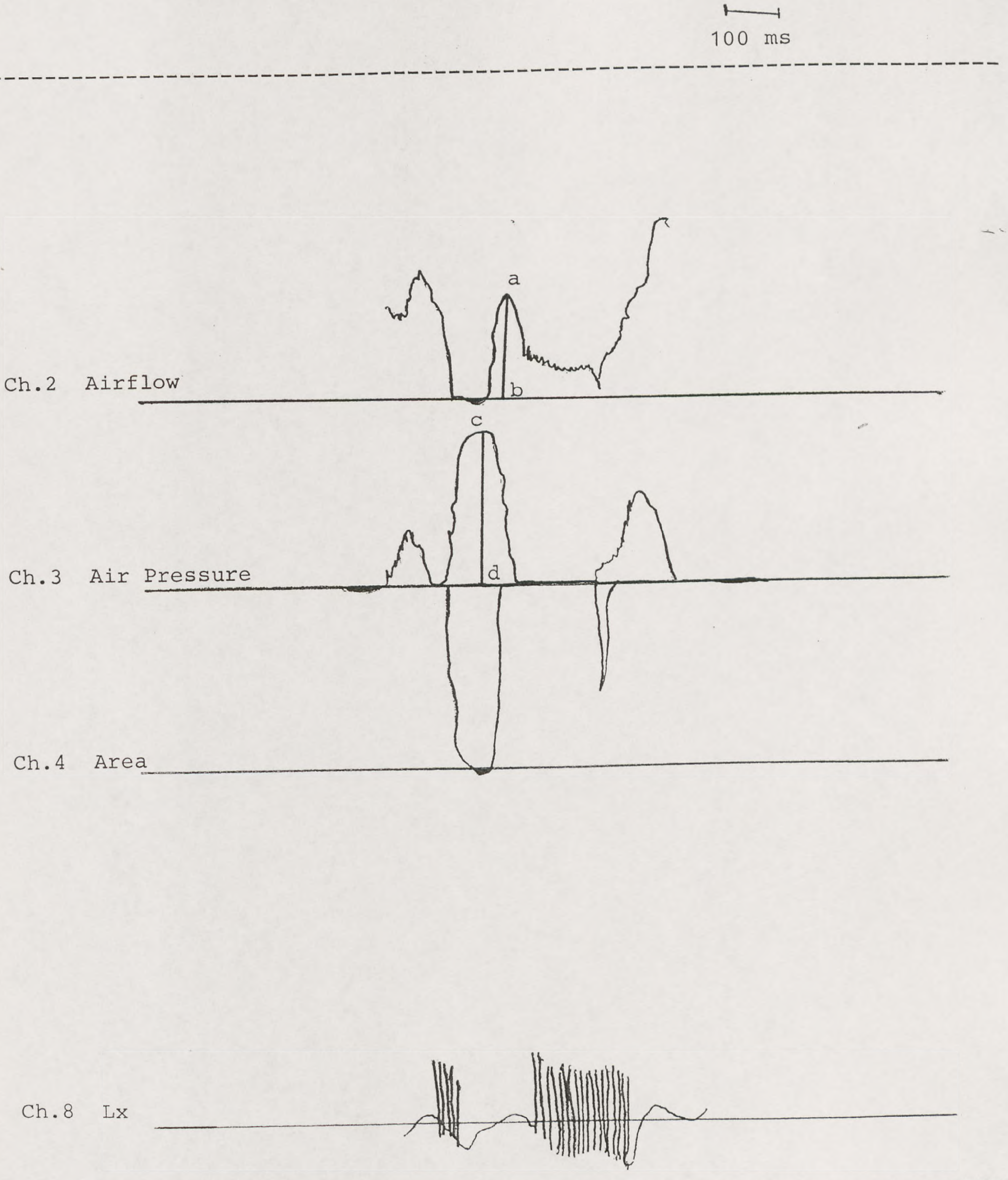


Figure 5.1.2 Mingogram traces of the word /fatiil/ showing the criteria of segmenting and measuring the peak oral airflow (a-b) and pressure (c-d) values in the aerodynamic study of emphasis and voicing.

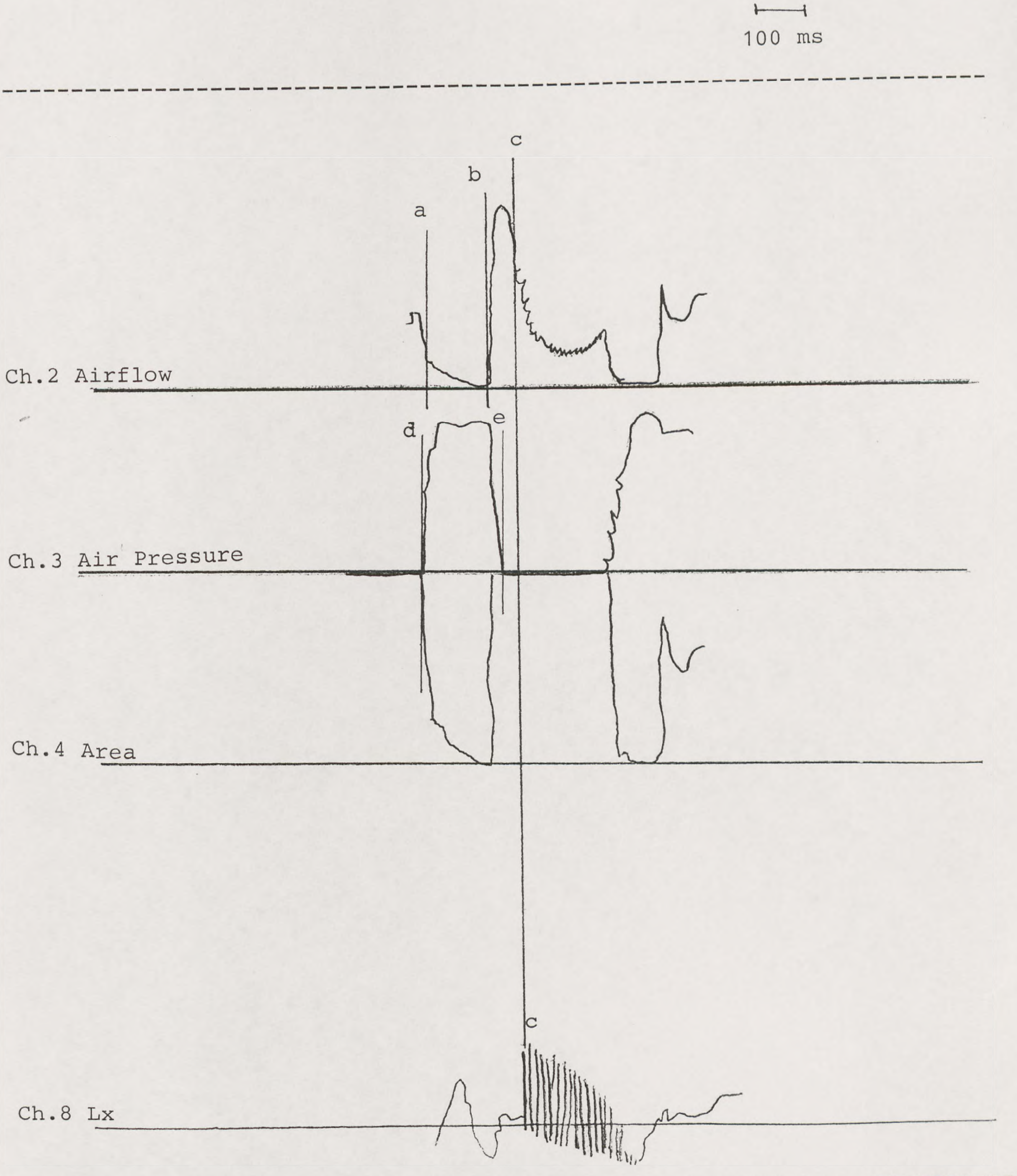


Figure 5.1.3 Mingogram traces of the word /taab/ illustrating the criteria of segmenting and measuring the articulatory closure (a-b), release (b-c) and the pressure durations (d-e) in the aerodynamic study of emphasis and voicing.

Two measurements were obtained from the pressure trace: the peak oral pressure (Po) at the time of closure and the total pressure duration (Pd). For the measurements of (Po) a baseline was established on the pressure trace. The peak oral pressure value was measured from the baseline to the highest point of the oral pressure trace at the time of closure (points c-d on Figure 5.1.2).

Since the articulatory closure for the linguo-velar and uvular consonants occurred posterior to the pressure tube placement in the mouth, oral pressure for these sounds could not be recorded on the pressure traces or measured.

Pressure duration was measured by locating the beginning and end points on the pressure trace. The initiation point (d) was identified as that point where pressure trace started to rise from the baseline (See Figure 5.1.3). The offset point was identified as that point where pressure returned to baseline again (e). This meant that the pressure duration as measured in this study included both the pressure rise and decay time. These two aspects were correlated with the articulatory closure and release respectively (Brown and Commerford, 1979). In this study no attempt was made to measure the duration of the pressure rise and decay separately. It was felt that measuring the total pressure duration would best serve our purpose in this study. It would allow valid comparisons between peak oral pressure and peak oral airflow as well as between the closure duration and the total pressure duration.

It was estimated that the accuracy of measurements for the three measures would be ± 4 msec for duration, ± 0.2 cm $H2\ O$ for peak oral pressure and ± 20 cm^3/sec for peak oral airflow. This needs to be kept in mind when it comes to the interpretation of the results and the patterns shown by the different contrasts.

These measurements of aerodynamic variables and articulatory durations were applied to all the contexts described in section 5.1.5.

5.2. Results for experiment 1: Aerodynamic patterns for the emphatic categories in YSA:

5.2.1 Peak oral pressure & emphasis:

The mean peak oral pressure values (Po) and the standard deviations for the emphatic and non-emphatic pairs in initial and medial positions as well as in geminated forms are displayed in Tables 5.2.1 and 5.2.2. The oral air pressure differences between the two cognates are quite small. Thus, the pair /Taab/ and /taab/ shows mean (Po) values of 7.1 cm H_2O and 7.3 cm H_2O respectively for subject (AA) and 10.3 and 9.4 cm H_2O for subject (WS). The speaker' variation is quite apparent for the same testwords. Subject (WS) produces higher (Po) values than subject (AA). The increase ranges between 2.5 cm H_2O and 3 cm H_2O.

		Word-initial		Word-medial			
Emphatic Contrast→	N	/T/	/t/	/T/	/t/	/TT/	/tt/
Vocalic Context↓							
/aa/	6	7.1	7.3	11.8	6.8	8.4	8.0
		(0.6)	(0.2)	(3.0)	(0.8)	(1.0)	(1.3)
/ii/	6	7.1	7.3	10.3	7.3	-	-
		(0.9)	(0.7)	(2.9)	(0.4)	-	-
/uu /	6	7.4	7.8	-	-	-	-
		(0.5)	(0.7)	-	-	-	-

Table 5.2.1 Means and standard deviations for the peak oral pressure (cm H_2O) in the word-initial and word-medial emphatics and non-emphatics (subject AA).

		Word-initial		Word-medial			
Emphatic Contrast→	N	/T/	/t/	/T/	/t/	/TT/	/tt/
Vocalic Context↓							
/aa/	6	10.3	9.4	9.2	9.4	11.0	10.4
		(1.3)	(0.9)	(1.6)	(1.4)	(1.5)	(0.7)
/ii/	6	10.9	10.5	10.2	10.4	-	-
		(1.8)	(0.9)	(0.6)	(0.8)	-	-
/uu/	6	10.5	10.1	-	-	-	-
		(1.1)	(0.6)	-	-	-	-

Table 5.2.2 Means and standard deviations for the peak oral pressure (cm/H_2O) in the word-initial and medial emphatics and non-emphatics (subject WS).

The emphatic /T/ seems to have higher peak oral air pressure values than /t/ in word-medial position for one speaker only, AA. Most pressure values are higher for WS than for AA. No other clear trends are apparent. The emphatics retain their increase over the nonemphatics in geminated forms. The values obtained for /baTTaat/ and /fattaat/ are of the order 11.0 and 10.4 cm H_2O. The increase is quite small and may be negligible. The geminated nonemphatics show higher mean (Po) values than the single nonemphatics. This is shown by both subjects. As far as the emphatics are concerned, gemination appears to have little effect on the peak oral pressure variations.

5.2.2.Pressure duration and emphasis:

The mean pressure durations (Pd) for the emphatic and the non-emphatic pairs as produced by the two subjects in the two positions and in geminated forms are shown in Tables 5.2.3 and 5.2.4. In initial position the mean (Pd) values show nonsystematic patterns and fewer variations between the contrasted pairs. Intersubject variations can be found. Subject (WS) consistently produced both the emphatics and the non-emphatics over a longer duration than subject (AA).

In medial position, the nonemphatics appear to show only slightly higher mean durational values than the emphatics with a few exceptions indicating that the differences may not be linguistically important. In spite of that, these differences are quite consistent for those pairs occurring in the presence of the vocalic context /ii/. Thus, the pair /faTiir/ and /fatiil/ shows mean (Pd) values of 74.2 and 81.7 msec for subject (AA).

The same pattern is found for the geminated emphatics and non-emphatics. The mean pressure durations for the geminated emphatics are almost double that of the single emphatics. Unlike (Po), the (Pd) measure shows marked differences between the geminated and single forms. This is true in both the emphatic and the nonemphatic environments.

The results show that the mean peak oral pressure values and the mean pressure durations are inversely related for the same test words and for the same subject.

An effect of position on the (Pd) is also seen. Mean (Pd) values of 110.8 and 85.8 msec are obtained for the initial and medial positions respectively in the non-emphatic class and 101.7 and 72.5 msec in the emphatic one.

		Word-initial		Word-medial			
Emphatic Contrast→	N	/T/	/t/	/T/	/t/	/TT/	/tt/
Vocalic Context↓							
/aa/	6	101.7	85.8	77.5	69.2	136.7	139.2
		(2.6)	(4.9)	(6.1)	(6.7)	(8.8)	(8.0)
/ii/	6	92.5	92.5	74.2	81.7	-	-
		6.9)	(7.6)	(3.8)	(4.1)	-	-
/uu/	6	90.8	90.8	-	-	-	-
		(5.9)	(6.7)	-	-	-	-

Table 5.2.3 Means and standard deviations for the pressure duration (msec) in the word-initial and medial emphatics and non-emphatics (subject AA).

		Word-initial		Word-medial			
Emphatic Contrast→ Vocalic Context↓	N	/T/	/t/	/T/	/t/	/TT/	/tt/
/aa/	6	101.7	110.8	72.5	85.8	138	148
		(4.1)	(5.9)	(9.9)	(8.0)	(8.2)	(9.9)
/ii/	6	108.3	119.2	85.0	104.2	-	-
		(9.3)				-	-
/uu /	6	106.7	110.8	-	-	-	-
		(7.5)	(7.4)	-	-	-	-

Table 5.2.4 Means and standard deviations for the pressure duration (msec) in the word-initial and medial emphatics and non-emphatics (subject WS).

5.2.3 Peak oral airflow and emphasis

The striking difference between the two cognates is shown by the aerdynamic variable of airflow. Figure 5.2.1 illustrates this distinction. It is observed that in a few cases the flow rate reached a negative point prior to the explosion of the emphatic sounds /T/ and /q/ especially in medial-position. Their occurrence is difficult to explain, but they are very rare. The downward movement of the jaw causing an inward movement of air may be a possible reason for their occurrence. It may have also been caused by the posterior movements of the tongue prior to the release. Negative flow, as Gilbert (1973) notes, does not always have to reflect a reversal of direction in the respiratory airstream.

The area trace also touches the baseline during the plosive closure phase reflecting the complete closure of the articulators. The emphatics show some influence on the neighbouring vowels. It is observed that the area gets lower for vowels following emphatics as compared with the same vowel following nonemphatics.

The mean peak oral airflow values (Uo) for the emphatic/non-emphatic contrasts in the two positions and in geminated forms are also presented in Tables 5.2.5-6. In initial position there is a considerable difference between all pairs in the mean (Uo) values regardless of variations in vocalic contexts, subject or phonetic position. The nonemphatics are observed to have mean (Uo) values which are four times as high as the emphatics. For example, subject (AA) produces mean values of 222 and 800 cm³/sec for the pair /qaad / and /kaad/ respectively. Similarly, subject (WS) shows mean values of the order 293 and 1270 cm³/sec for the same pair.

The mean (Uo) values in medial position and in geminated forms for the emphatic/non-emphatic pairs show a similar pattern to that found in initial position. However, both the emphatic and the nonemphatic components are found to have lower mean (Uo) values in medial than in initial position.

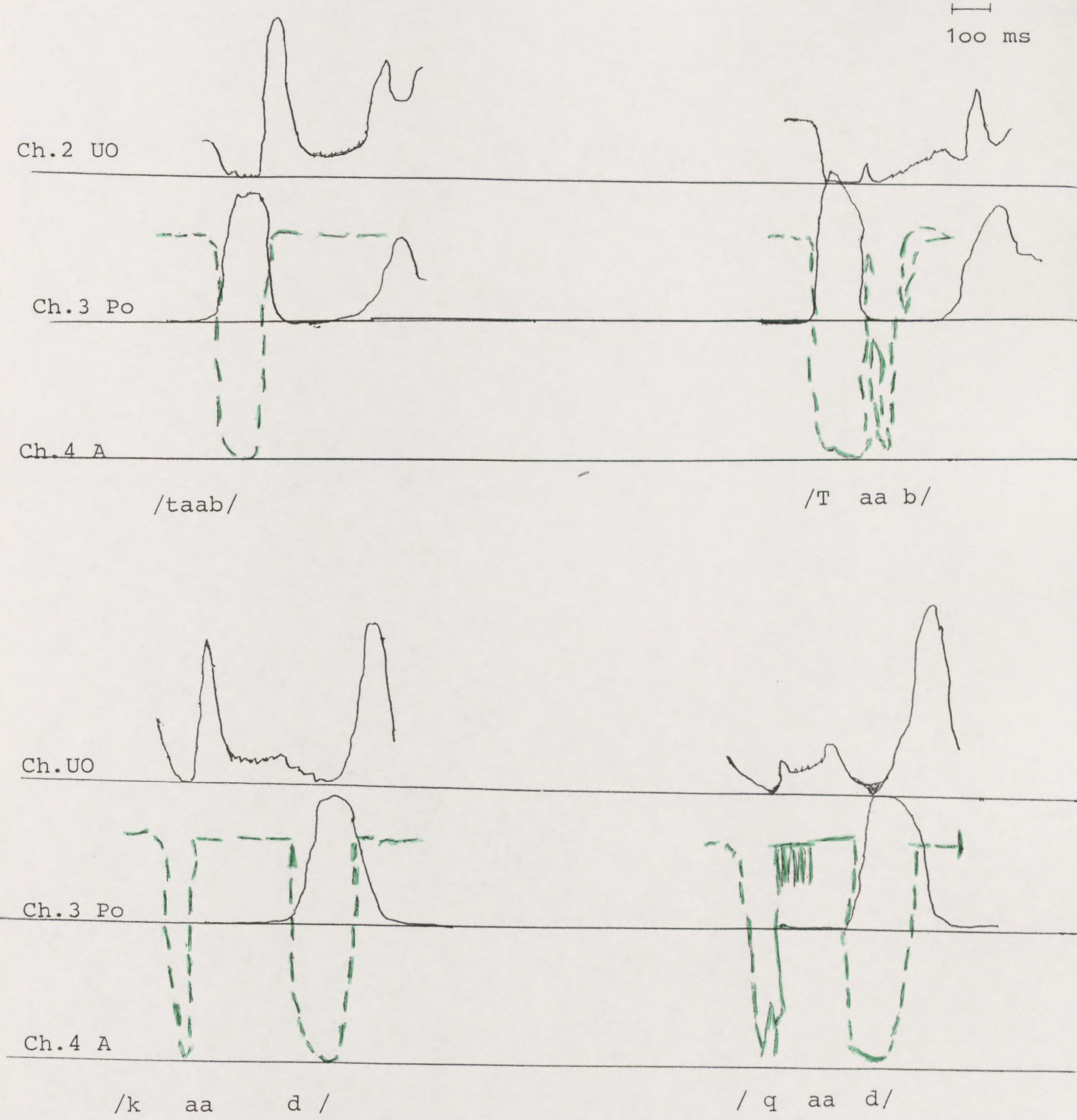


Figure 5.2.1 Mingogram traces of some test words illustrating the oral air pressure and airflow patterns for the emphatic/non-emphatic distinctions in YSA.

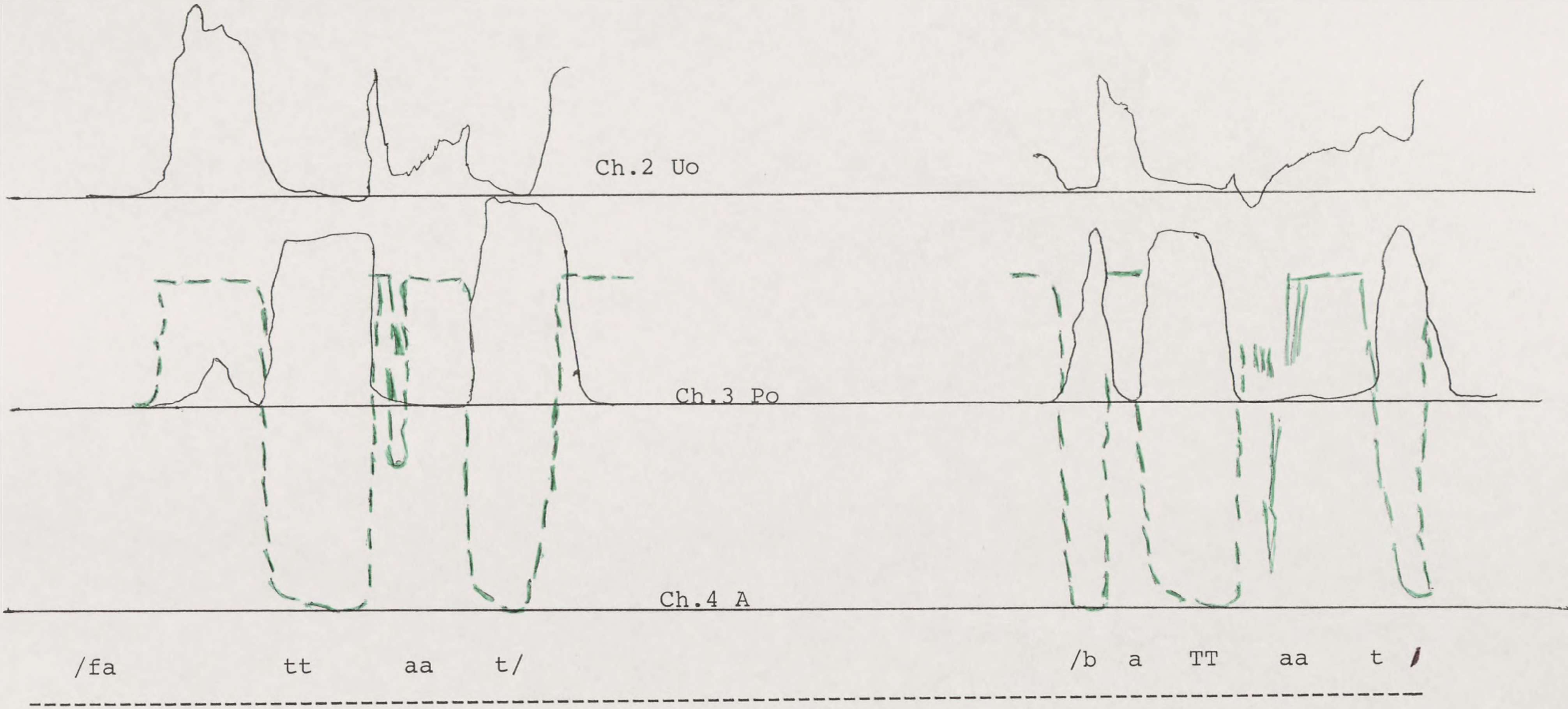


Figure 5.2.1 (Continued).

There is a slight increase in the mean (Uo) values for the geminated emphatics over the single ones in word-medial position for AA only. The mean (Uo) values obtained for the pair /baTTaat/ and /baTaaT/ are 208.2 and 120 cm³/sec respectively. A similar increase is found for the geminated and single nonemphatics.

		Word-initial position				Word-medial position							
Vocalic context↓	N	/T/	/t/	/q/	/k/	/T/	/t/	/q/	/k/	/TT/	/tt/	/qq/	/kk/
/aa/	6	263	1040	222	800	120	742	75	965	208	820	167	1063
		(43)	(83)	(40)	(155)	(105)	(103)	(77)	(301)	(173)	(159)	(37)	(116)
/ii/	6	270	867	74	677	250	849	65	728	-	-	-	-
		(32)	(134)	(109)	(121)	(49)	(139)	(170)	(106)	-	-	-	-
/uu/	6	272	927	178	493	-	-	-	-	-	-	-	-
		(74)	(266)	(42)	(73)	-	-	-	-	-	-	-	-

Table 5.2.5 Means and standard deviations for the peak oral airflow (cm³/sec) in the word-initial and word-medial emphatics and non-emphatics (subject AA).

		Word-initial position				Word-medial position							
Vocalic context↓	N	/T/	/t/	/q/	/k/	/T/	/t/	/q/	/k/	/TT/	/tt/	/qq/	/kk/
/aa/	6	143	1173	293	1270	250	657	172	965	128	695	237	1130
		(67)	(203)	(94)	(164)	(47)	(114)	(48)	(185)	(66)	(110)	(54)	(92)
/ii/	6	269	757	348	622	270	617	373	790	-	-	-	-
		(76)	(105)	(77)	(183)	(42)	(94)	(74)	(73)	-	-	-	-
/uu /	6	103	1127	175	683	-	-	-	-	-	-	-	-
		(110)	(113)	(64)	(164)	-	-	-	-	-	-	-	-

Table 5.2.6 Means and standard deviations for the peak oral airflow (cm³/sec) in the word-initial and medial emphatics and non-emphatics (subject WS).

5.2.4 Articulatory closure duration and emphasis

The closure duration results for the emphatics and the non-emphatics in the two positions and in geminated forms are displayed in Tables 5.2.7and 5.2.8. The mean closure durations (Uc) show a clear-cut distinction between the two pairs. The emphatics are consistently produced with higher (Uc) values than the non-emphatics. It is more likely that the lower mean peak airflow values associated with the emphatics may have been caused by this increase in the closure duration.

Although similar closure duration patterns are found, the values for all pairs in medial position appear to be lower than those observed in initial position.

The emphatic/non-emphatic distinction is also retained in geminated forms. It is not as marked as that for single forms. The greatest effect on the closure duration for both emphatics and nonemphatics is caused by gemination. The mean closure duration values for the geminated emphatics are almost double that of the single emphatics. For example, the mean values shown for /baTTaat/ and /baTaaT/ are 117 and 63 respectively. The same observations can be said about geminated and single nonemphatics. These differences reached a highly significant level ($p \leq 0.01$). On the other hand, the effect of gemination on the aerodynamic variables of (Po) and (Uo) is not as strong as on the durational measures of (Pd) and (Uc).

The two vocalic contexts appear to have inconsistent and nonsystematic effect on the (Uo) and (Uc) values in the two classes: emphatics and non-emphatics. Thus the nonemphatics show higher mean (Uo) values and shorter mean (Uc) values in the environment of /aa/ than in the environment of /ii/. In the emphatics it is just the opposite (i.e. lower mean (Uo) values and longer (Uc) values in the environment of /aa/ than in the environment of /ii/). These durational differences indicate that the emphatic/non-emphatic pairs are very different in their constriction time regardless of position, vocalic context, subject and gemination. This observation is in marked contrasts with Ali and Daniloff's (1972) who claim that these contrasted pairs are found to have similar durations of articulatory closure.

		Word-initial position				Word-medial position							
Vocalic context↓	N	/T/	/t/	/q/	/k/	/T/	/t/	/q/	/k/	/TT/	/tt/	/qq/	/kk/
/aa/	6	77	53	45	32	63	41	43	34	117	111	89	99
		(8)	(13)	(5)	(4)	(3)	(7)	(7)	(9)	(9)	(10)	(11)	(14)
/ii/	6	69	53	27	30	57	47	46	44	-	-	-	-
		(7)	(4)	(14)	(7)	(5)	(6)	(6)	(6)	-	-	-	-
/uu/	6	69	48	52	38	-	-	-	-	-	-	-	-
		(6)	(6)	(7)	(8)	-	-	-	-	-	-	-	-

Table 5.2.7 Means and standard deviations for the closure duration (msec) in the word-initial and medial emphatics and non-emphatics (subject AA).

		Word-initial position				Word-medial position							
Vocalic context↓	N	/T/	/t/	/q/	/k/	/T/	/t/	/q/	/k/	/TT/	/tt/	/qq/	/kk/
/aa/	6	87	81	79	57	48	41	79	50	123	121	114	101
		(3)	(5)	(8)	(3)	(14)	(12)	(7)	(8)	(9)	(12)	(6)	(8)
/ii/	6	87	77	86	55	56	47	64	62	-	-	-	-
		(10)	(9)	(12)	(5)	(5)	(6)	(13)	(7)	-	-	-	-
/uu/	6	97	78	83	63	-	-	-	-	-	-	-	-
		(18)	(14)	(11)	(4)	-	-	-	-	-	-	-	-

Table 5.2.8 Means and standard deviations for the closure duration (msec) in the word-initial and medial emphatics and non-emphatics (subject WS).

5.2.5 Articulatory release duration and emphasis

The mean release duration values (Ur) for the emphatic and the nonemphatic pairs in the two positions and in geminated forms are displayed in Tables 5.2.9 & 5.2.10. The emphatics are observed to have lower mean (Ur) values than the non-emphatics. This is just the opposite of what is observed for the closure duration. The pattern is consistent and systematic in the two positions, the two vocalic contexts, geminated forms and for the two subjects. The results also show that the mean release duration for any test word varies directly with the peak oral airflow for the same word. In other words, there is a one to one positive correlation between (Uo) and (Ur). The higher the peak oral airflow value, the longer the release duration is found to be for the same test word as produced by either of the two subjects.

These results suggest that (Ur) is not less effective than the (Uo) or the (Uc) measures. The emphatics and the nonemphatics are not only differentiated by the duration of their articulatory closure but also by their articulatory release time.

		Word-initial position				Word-medial position							
Vocalic context↓	N	/T/	/t/	/q/	/k/	/T/	/t/	/q/	/k/	/TT/	/tt/	/qq/	/kk/
/aa/	6	18	29	15	33	17	26	16	37	19	27	19	40
		(3)	(2)	(3)	(4)	(4)	(4)	(4)	(5)	(4)	(4)	(4)	(0)
/ii/	6	21	43	18	40	23	38	24	43	-	-	-	-
		(2)	(6)	(5)	(3)	(4)	(9)	(12)	(3)	-	-	-	-
/uu/	6	18	37	16	38	-	-	-	-	-	-	-	-
		(3)	(3)	(2)	(7)	-	-	-	-	-	-	-	-

Table 5.2.9 Means and standard deviations for the release duration (msec) in the word-initial and medial emphatics and non-emphatics (subject AA).

		Word-initial position				Word-medial position							
Vocalic context↓	N	/T/	/t/	/q/	/k/	/T/	/t/	/q/	/k/	/TT/	/tt/	/qq/	/kk/
/aa/	6	18	59	24	68	27	43	21	23	18	33	21	48
		(3)	(3)	(2)	(12)	(8)	(9)	(2)	(4)	(3)	(5)	(4)	(7)
/ii/	6	30	71	26	53	27	55	51	53	-	-	-	-
		(6)	(9)	(5)	(19)	(8)	(5)	(6)	(17)	-	-	-	-
/uu/	6	20	63	19	57	-	-	-	-	-	-	-	-
		(5)	(8)	(4)	(12)	-	-	-	-	-	-	-	-

Table 5.2.10 Means and standard deviations for the release duration (msec) in the word-initial and medial emphatics and non-emphatics (subject WS).

Although the t-test is useful for comparing two group means its disadvantage is that it cannot handle data with several independent and dependent variables interacting simultaneously. In particular our data have five dependent variables (Po, Pd, Uo, Uc and Ur) and at least four independent variables: emphasis, position, vocalic context and subjects each consisting of at least two levels. To account for the variations in the dependent variables it is considered necessary to use a more powerful statistical test such as the analysis of variance (ANOVA). It will enable us to find the relative effect caused by each of the independent variables on the dependent ones.

The ANOVA test has similar assumptions to the t-test with some extra ones. These are random assignment and equal variance within the sample sets, normality of population values within sets and additivity of the different sources of variability which contribute to the main effect (Meddis, 1973; Miller, 1984; Butler, 1985; Hatch and Farhady, 1982).

The analysis of variance is done on the SPSS package for PC installed on the Novell Network of the University of Leeds Computing Service. Two data files were created: one for the emphasis data and the other for the voicing data (see section 5.3). All measurements obtained from the two subjects participated in these experiments were first entered into the computer using Data Entry Two.

The emphasis data are arranged into eight separate columns. The independent variables are indicated by the symbols PS, V, E and S for position , vocalic contexts, emphasis and subject in that order (See Appendix 8 for definition of these symbols). Each of these has two levels identified by the Figures (1) and (2) when entering the data. Thus, the figure (1) stands for initial position in C1, vocalic context /aa/ in C 2, emphatics in C3, subject AA in C5, whereas the figure (2) stands for medial position in C1, vocalic context /ii/ in C2, nonemphatics in C3 and subject WS in C5. The other columns are occupied by actual measurements of the dependent variables. There are 192 cases for the variables (Uo), (Uc), (Ur) and 96 cases for the variables (Po) and (Pd). All token within each level are of similar conditions. The number of cases is not only identical within each level but also in all comparable columns as appears from Appendix (9).

Tables 5.2.11 -14 show the main effect of the different factors and their interactions on the dependent variables. For each contrast the F-ratio and the P-value are presented. The former is meant to show the variations attributed to each factor. Since there are equal cases and levels within each factor the degrees of freedom are not included in the results. The results in Table 5.2.11 show that most factors have a highly significant main effect on the dependent variables at ($p \leq 0.01$) level. The peak oral airflow variations reach a significant level with all factors except position. The extent of the differences caused by each factor varies greatly as can be seen from the F-ratio. The F-ratio for the emphatic distinction greatly exceeds that of the other factors (F=776.727). This result shows that emphasis is the factor which carries the primary effect on the (Uo) variations. As far as the other variables are concerned the only insignificant main effect observed is that by the vocalic context on the peak oral pressure (p= 0.249) and on the closure duration (P= 0.983). With the pressure duration, two factors, position and subjects prove to be responsible for a large part of the variations. With the closure duration, emphasis is shown to be the second influential factor after subjects whereas in release duration emphasis is the main source of variation. Thus from these results of the ANOVA main effect, it can be said that emphasis is a major primary source of variation affecting all dependent variables.

Table 5.2.12 shows the results of two-way interaction effect on the dependent variables. A vertical inspection of the table reveals fewer (Po) variations as a result of the interaction between emphasis and vocalic context (P= 0.255) and between position and vocalic context (p= 0.735). All other two-way interactions show significant effect at ($P \leq 0.05$) on the peak oral pressure. The most effect is found for the interaction between position and subject followed by that of emphasis and subject. The F-values observed for the two contrasts are (16.629) and (9.058) respectively.

With the pressure durations all contrasts involving two-way interactions reach a highly significant effect. In particular the interaction of emphasis and subject is found to have the most effect (F= 40.511). The interactions of position and vocalic context as well as position and subjects show only significant effect rather than a highly significant one, indicating that whenever position or vocalic context interact with any other factor, they reduce the extent of variation and thereby the significance level of the interaction.

Four of the six two-way interaction contrasts show non-significant effect on the peak oral airflow at ($P \leq 0.05$) suggesting that the significant main effect of emphasis on (Uo) found in Table 5.2.11 remains the most influential one. With regard to the closure duration only three contrasts result in a significant effect ($p \leq 0.05$). These are position and subject, vocalic context and subject and emphasis and vocalic context.

The interaction effect on (Ur) is also significant with three contrasts: emphasis and position ($p \leq 0.01$) emphasis and subject ($p \leq 0.01$) and position and subject ($P \leq 0.01$). All other interactions do not seem to have significant effect on Ur.

Table 5.2.13 shows the results of the three-way interactions. A horizontal inspection of this table shows that the interaction of the factors emphasis, position and subject to have a higly significant effect on (Po, Uo and Ur). Their interaction effect on (Pd) and (Uc) are, however, not found to be significant at the 5% level. Non-significant F-value is also observed for the three-way interaction of emphasis, position and vocalic context.

The interaction effect of the four factors on the dependent variables is shown on Table 5.2.14. This interaction effect is more apparent on the aerodynamic variables (PO) and (UO) ($p \leq 0.01$). It is less on the release duration ($p \leq 0.05$). No significant effect is observed with (Uc) and (Pd). This is a clear indication that the variations observed in the dependent variables (Po), (Uo) and (Ur) cannot only be attributed to one factor. All factors have been found to contribute significantly to this variation. However, some caused more variations than others. On the other hand, the insignificant results obtained for the closure duration and the pressure duration suggest that the significant two-way interaction found previously remains the main contributor to the variations in the (Uc) variable.

			Dependent Variables				
ANOVA main effect	N	Value	Po	Pd	Uo	Uc	Ur
Position	192	F	5.916	232.585	3.539	23.213	5.985
		P	0.017 *	0.000**	0.062	0.000**	0.015*
Emphasis		F	14.139	10.429	776.727	46.55	472.64
		P	0.000**	0.002**	0.000**	0.000**	0.000**
Vocalic Context		F	1.350	27.56	14.419	0.000	23.489
		P	0.249	0.000**	0.000**	0.983	0.000**
Subject		F	43.99	112.188	3.920	82.365	132.09
		P	0.000**	0.000**	0.000**	0.000**	0.000**

Table 5.2.11 Analysis of variance showing the main effect of the independent variables on Po, Pd, Uo, Uc and Ur.

			Dependent Variables				
2-Way interaction effect	N	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
Emphasis &	192	F	8.714	6.95	0.281	1.093	10.74
Position		P	0.004**	0.010**	0.597	0.297	0.001
Emphasis &		F	1.317	11.013	41.654	6.353	0.769
Vocalic Context		P	0.255	0.001**	0.000**	0.013**	0.382
Emphasis &		F	9.058	40.511	0.120	0.000	43.567
Subject		P	0.003**	0.000**	0.729	0.983	0.000**
Position &		F	0.115	4.988	1.093	2.545	0.479
Vocalic Context		P	0.735	0.028*	0.297	0.112	0.490
Position &		F	16.629	4.712	3.177	18.40	10.144
Subject		P	0.000**	0.033*	0.076	0.000**	0.002**
Vocalic Context		F	3.939	15.766	8.307	4.06	3.241
Subject		P	0.051*	0.000**	0.004**	0.045*	0.074

Table 5.2.12 ANOVA showing the two-way interaction effects of the independent variables on Po, Pd, Uo, Uc and Ur.

			Dependent Variables				
3-Way interaction effect	N	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
Emphasis, Position &	192	F	0.470	0.521	0.357	2.343	0.616
Vocalic Context		P	0.495	0.473	0.551	0.128	0.434
Emphasis, Position &		F	19.574	0.008	8.586	0.707	4.706
Subject		P	0.000**	0.930	0.004**	0.402	0.031*
Emphasis, Vocalic Context		F	0.359	6.669	23.837	1.33	1.792
Subject		P	0.551	0.012**	0.000**	0.250	0.182
Position, Vocalic Context &		F	0.470	0.003	3.57	1.005	0.002
Subject		P	0.495	0.957	0.060	0.318	0.963

Table 5.2.13 Analysis of variance showing 3-way interaction effect of the independent variables on Po, Pd, Uo, Uc and Ur.

			Dependent Variables				
4-Way interaction effect	N	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
Emphasis, Position, Vocalic Context &	192	F	8.48	0.525	9.198	2.72	5.115
Subject		P	0.000**	0.471	0.003**	0.101	0.025*

Table 5.2.14 Analysis of variance showing 4-way interaction effect of the independent variables on Po, Pd, Uo, Uc and Ur.

5.2.6.1 Discussion

Our data show that the peak oral pressure differences for the emphatics and non-emphatics are quite small. Statistical analyses show significant differences in (Po) between the two groups but only in some positions and in some contexts. When all measurements are pooled for the two subjects and subjected to an analysis of variance, the F-ratio reaches a significant level. Similarly, the peak oral airflow values for the emphatics and nonemphatics are also significantly different in pooled data.

Since the aerodynamic parameters help to suggest articulatory behaviour during speech, it is important to reflect on the significance of such variations and the implications of these findings. Previous research shows that subglottal pressure below the glottis does not seem to vary significantly for different plosive categories (Netsell 1969, Ohala 1980, Lofqvist 1976, McGlone and Shipp, 1971). If that is the case, one may assume that any changes found in the peak oral airflow or peak oral air pressure values for the emphatic/nonemphatic contrast may be attributed to glottal and supraglottal actions.

Although the glottal action is crucial for the (Po) variations in some linguistic contrasts (e.g. voicing) it may not be so for the emphatic/non-emphatic distinction. As is pointed out earlier both the emphatic and non-emphatic consonants in this investigation are voiceless. The intention behind such a choice is to neutralize the voiced voiceless effect so that any observed differences can only be attributed to emphasis and its associated articulatory characteristics. Consequently, the vocal folds for these voiceless sounds are believed to be abducted and the glottal resistance to the pulmonary airflow to be minimal (Van den Berg, 1957; Rothenberg, 1968; Warren, 1976). This makes us contemplate the possibility that the dynamic changes in the pharyngeal and the oral tracts may be the factors which result in a significant obstruction to airflow.

The peak oral airflow is shown to be a strong and reliable measure of emphasis. There are several possible explanations as to why these two phonological categories display significant differences in (Uo). As is indicated the emphatics are believed to have a secondary constriction in the pharynx, in addition to their primary one at the denti-alveolar region. The non-emphatics are assumed to have the same place of articulation as that for the emphatics. The airflow is obstructed not only by the raising of the tongue against the roof of the mouth at the denti-alveolar region, as is the case with the non-emphatics, but also by the pharyngeal obstruction, for the secondary constriction. The mandible and lips' movements (Elgendy, 1990) and the retraction of the tongue (Ali and Daniloff, 1972) are also other factors contributing to that obstruction. Our main prediction in this study is that the more total obstruction there is, the lower the airflow out of the mouth is expected to be. It is, therefore, reasonable to assume that the main differentiating articulatory factor for the emphatics and the non-emphatics is the secondary articulation at the pharynx. These findings confirm our prediction that the secondary constriction during the production of emphatics may be associated with a greater airflow resistance as compared with the non-emphatics.

Another explanation for the (Uo) variations between the emphatics and the non-emphatics may be related to differences in the cavity size at the point of vocal tract constriction. Support for this finding has come from Isshiki and Ringel (1964). In this study it is stated that the relationship between the cavity size and the oral airflow rate is an inverse one. As the cavity size becomes smaller, one would expect the air pressure to rise and consequently the oral airflow rate at a plosive release is expected to rise as well.

Another finding of this study is related to the pressure-flow relationship for the emphatics and the non-emphatics. The two parameters have not been found to vary linearly as is suggested by many studies. The magnitude of pressure build-up during the closed phase of the emphatic and the nonemphatic consonants does not determine the rate of oral airflow during the release phase. It is observed that the peak oral airflow values greatly increase for the non-emphatics even when the peak oral pressure is equal to or lower than that for the emphatics. The implication of such a remark is that air pressure is not a major determinant of peak oral airflow. This finding is in contrast with the claim of some researchers that the two factors vary together. Thus, Subtelny et. al.(1966) conclude that "...the air pressure posterior to a given articulatory stricture is one of the major determinants of the rate of oral airflow (volume per unit)". Lubker, too, (1973) reports that the oral airflow and air pressure "...were generally, perhaps linearly related".

One possible explanation for the contradiction between the results of those studies and ours may be related to the nature of the linguistic feature under investigation and the material used. Their data are based on material for the voiced and voiceless sounds while ours are based on material from emphatics and non-emphatics. The two phonetic features vary greatly with regard to the articulatory behaviour and the glottal and supraglottal articulation. It seems that the pressure build-up behind the articulatory closure can only affect the airflow rate in the presence of voicing. Further evidence for this statement is presented by Haag (1977). Using aerodynamic data from German voiceless plosives he argues that no relationship between the oral pressure and the oral airflow can be found in his data. In addition, Hixon, Minifie and Tait (1967) conduct a study on the air pressure-flow relationships in the English voiceless consonants [s] and [sh]. Among other things, they find that the flow rate for [sh] as compared with [s] increases even when the pressure remains unchanged. As a result Hixon et. al. (1967) attribute these unrelated changes in airflow rate and oral pressure to the mode of constriction in the vocal tract. They indicate that [sh] is produced with a larger minimum area of constriction than [s].

On the other hand, significantly longer closure duration values for the emphatics are obtained as compared with the nonemphatics. This demonstrates that there may be more than one factor contributing to the (Uo) variations for the emphatic/non-emphatic distinction. It is not only the degree of constriction that may be resposible for the (Uo) variations but also the timing of the articulatory closure. This last parameter could be a decisive one particularly if one considers the point that both the emphatics /T/ and /q/ and their nonemphatic counterparts /t/and /k/ involve complete closure and that the minimum area of constriction may not be easily compared. Thus, the variations in the closure as well as in the release durations may reflect the speed of articulators' movements and the time they remain in a particular articulatory position until the appropriate amount of frication noise has been generated and achieved.

One more finding of this study is that the peak oral airflow, closure and release durations covary together from one underlying articulatory contrast. The inverse relationship between the peak oral airflow and the duration of articulatory closure appears to be a relevant one for the emphatic and the nonemphatic contrasts. It is found that the longer the plosive closure in the same test words, the lower the peak oral airflow value will be for the same subject. On the other hand, a one to one positive correlation is found for the peak oral airflow and the release duration. The peak oral airflow value becomes greater as the time required for the release of a certain sound is found to be longer.

To sum up: of the variables investigated in this study the peak oral airflow, the durations of articulatory closure and release are the ones that contribute significantly to the emphatic/non-emphatic distinction in YSA. The systematic variations in these parameters may reflect the rapid dynamic changes in the vocal tract. As far as emphasis is concerned, the pressure build-up behind the constriction seems to have nothing to do with the (Uo) variations for the emphatics and non-emphatics. It seems from the results that the (Uo) variations may be explained by the following factors: (1) the obstruction of airflow at the pharyngeal and the oral tracts (2) the durations of the constriction and release and the expiratory activities involved in this distinction (3) the jaw and lips movements which are believed to determine not only the location of the pharyngeal constriction but also the degree of the tongue's retraction (Elgendy, 1990).

5.2.6.2 Phonetic position

The air pressure results are found to vary as function of phonetic position. The two positions are found to vary systematically with the medial plosives having consistently higher peak oral pressure values than the initial ones. This finding is in agreement with Arkebauer, Hixon and Hardy (1967) who report significantly greater oral air pressure values in the intervocalic position than either the pre- or post-vocalic positions. It is also consistent with that of Malècot (1955) in which the word-medial are reported to have higher peak oral pressure than either the word-initial or word-final plosives. However, Lisker (1970) points out that the word-initial plosives are associated with higher oral air pressure values than the word-medial plosives, an observation which stands in contrast to Malècot (op. cit.).

One possible explanation for this inconsistency in results for phonetic position may be related to the degree of tongue resistance to the respiratory airstream. Since the same consonants show different peak oral air pressure values in different positions it is tempting to assume that a speaker's articulatory behaviour during a particular consonant depends partly on the phonetic position.

The results in Tables 5.2.1 and 5.2.2 show that there are a few differences in (Po) values for the same sounds in initial and medial positions. These, however, are few in number and limited to certain variables. They are sensitive to subject variations.

Another probable explanation of a more speculative nature is that the subjects' control over the articulatory timing and/or the vocal tract shapes as well as the expiratory effort may have been similar in the two positions; thus resulting in correspondingly similar (Po) values for the same plosive sounds.

A more likely reason for the lack of (Po) variations in the two positions may be caused by the use of the carrier sentence. All our word-initial and word-medial plosives are embedded in a carrier sentence. As such all the consonants of the sample may be regarded as word-medial since they all occur within the carrier sentence. Nevertheless, the instructions given to the subjects are to read each test word in a clear and natural manner so as to avoid any ambiguity in their pronunciation. A subsequent inspection of the recording confirms that the recorded words within the carrier sentence are pronounced in a consistent speech pattern with a clear and sufficient distinction.

On the other hand, several studies report contradictory results to those shown in this study (Black, 1960; Brown et al, 1970; Dixit and Brown, 1978; Subtelny et. al., 1966). For example, Black finds decreasing oral pressure values for the word-initial, word-medial and word-final plosives in that order. Brown et al. (1970) presented higher oral pressure values for the phonemes /p/, /b/, /t/ and /d/ in initial position than in medial position. Dixit and Brown (op cit.) show larger peak pressure values for Hindi plosives in medial positions than those in initial position. Although such a finding is consistent with ours, its validity is overshadowed by the fact that in their material, position and stress are interacting in such a way that it is difficult to state confidently whether the effect reported is related to one or the other. In recognition of this close connection between position and stress Malècot (1968) notes that the two factors "operate in an independent and additive fashion" (p. 101).

There is also a strong indication that the consonant duration greatly influence the phonetic position. The emphatic and the non-emphatic categories are observed to have different pressure duration in initial and medial positions. Similar differences are also present in the duration of articulatory closure and release as function of position. With a few exceptions all durational differences suggest that the articulators require a longer time to create a word initial plosive than they take to produce the same sound in medial position. It is not clear why such an increase in duration is a characteristic of the initial plosives but not of the medial ones. One may speculate that the articulators for the word-initial plosives take a longer time because they start from the rest position and take sometime to reach the point of contact, whereas in the word-medial the active articulators (e.g. the tongue) is continuing the movement from the point of contact for the preceding vowel and not from the rest position. In other words the extent of the tongue's displacement and other relevant articulators may explain the duration differences in the two positions.

Another related reason may be that the acoustic requirements of the word-medial plosives do not necessitate the articulators to remain constricted for a longer time as compared with the word-initials.

In addition the emphatics and non-emphatics show significantly greater peak oral airflow values in initial position than in medial position. Such variations may reflect the aerodynamic interaction between the obstruction to the airflow and the pressure duration. Longer pressure duration may result in higher peak oral airflow values. This finding is consistent with that of previous studies (e.g. Emmanuel and Counihan, 1970; Gilbert, 1973; Dixit and Brown; 1986). Emmanuel et. al. report higher airflow rates for the word-initials as compared with the word-medial plosives in CV and VCV syllables. Gilbert (1973) indicates that flow rates are higher for the pre- than either the intervocalic or the post-vocalic positions. He adds that the pre-and intervocalic contexts are associated with greater flow rates than the post-vocalic contexts.

Contrary to our results Dixit and Brown (1986) show word-medial plosives to be associated with higher airflow rates than word-initial ones. The contradiction between our airflow results and those of others may be understood with reference to the type of material used and the task required to be achieved by the subjects. In some studies the stress placement on the test words is not indicated whether it is in initial or medial position (e .g. Isshiki and Ringel 1964). In other studies the target plosives are unstressed in word-initial position , and stressed in word-medial position (e.g. Dixit and Brown, 1986). In our material the plosives investigated in both initial and medial positions occur as the start of a stressed syllable. Hence any variations in the peak oral airflow values can reasonably be attributed to position. Such observations caution against direct comparisons of results without taking into account the source of the material and its aims.

5.3 Results for experiment 2: Aerodynamic patterns of voicing in YSA.

5.3.1 The peak oral pressure and the voicing distinction:

The means and the standard deviations for the peak oral air pressure in the voicing contrasts are displayed in Tables 5.3.1 and 5.3.2. In word-initial position the peak oral pressure values (Po) for the voiceless plosives are considerably higher than the voiced ones in all vocalic contexts and for the two subjects. Unlike the emphatics and the nonemphatics, the voicing categories are well separated by this measure. For example, for subject AA the values obtained for the sounds /t/ and /d/ in the pairs /tuub/ and /dʊʊr/ is 7.8 and 3.8 cm H_2O respectively. Subject WS shows similar pattern but with higher values than those for subject AA. The values obtained for this subject are 10.1 and 8 cm H_2O for the same pair. These values obtained by the two subjects are typical of the peak oral pressure values in the voicing contrasts in YSA .

In word-medial position the two voicing categories show systematic and similar properties to that found in initial position. The voiceless sounds are observed to have higher (Po) values than those in initial position. This remark does not apply to the voiced sounds. The pattern found for the voicing distinction is also retained with the geminated forms. Thus the pair /fattaat / and /saddaad/ obtains values of 8 and 5.5 cm H_2O respectively for subject AA. The values obtained for subject WS were 10.4 and 5.8 cm H_2O for the same pair. These results clearly indicate that gemination has little effect on the (Po) variations for the voicing distinction. When the peak oral pressure values for the geminated and single forms are compared for the same voicing category (e.g. voiceless plosives) no real differences are found.

The mingograms in Figure 5.3.1 show that the voiced sounds are characterised by rapid fluctuations in flow and pressure traces. The shape and size of the oral pressure waveforms for the voiced and the voiceless plosives are quite different. The pressure rise is a very rapid one for the voiceless plosives and less rapid for the voiced ones. The voiceless plosives are associated with a smoothed waveform, whereas the voiced ones are observed to have more or less a linear appearance with ripples indicating the presence of voicing and the vocal folds vibrations.

To test the significance of these differences all (Po) measurements for the voiced and voiceless sounds are pooled into two groups for each subject separately and statistically treated by the t-test for groups. The t-test results show the differences between the voicing cognates to be highly significant at ($p \leq 0.01$) and for the two subjects. This finding confirms our earlier remark that the oral pressure measure is a reliable indicator of the voicing distinction. Its variation appears to be much influenced by the glottal obstruction to airflow. It is also consistent with a number of studies (Black, 1950, Subtelny et. al. (1966), Miller and Daniloff, 1977, among others).

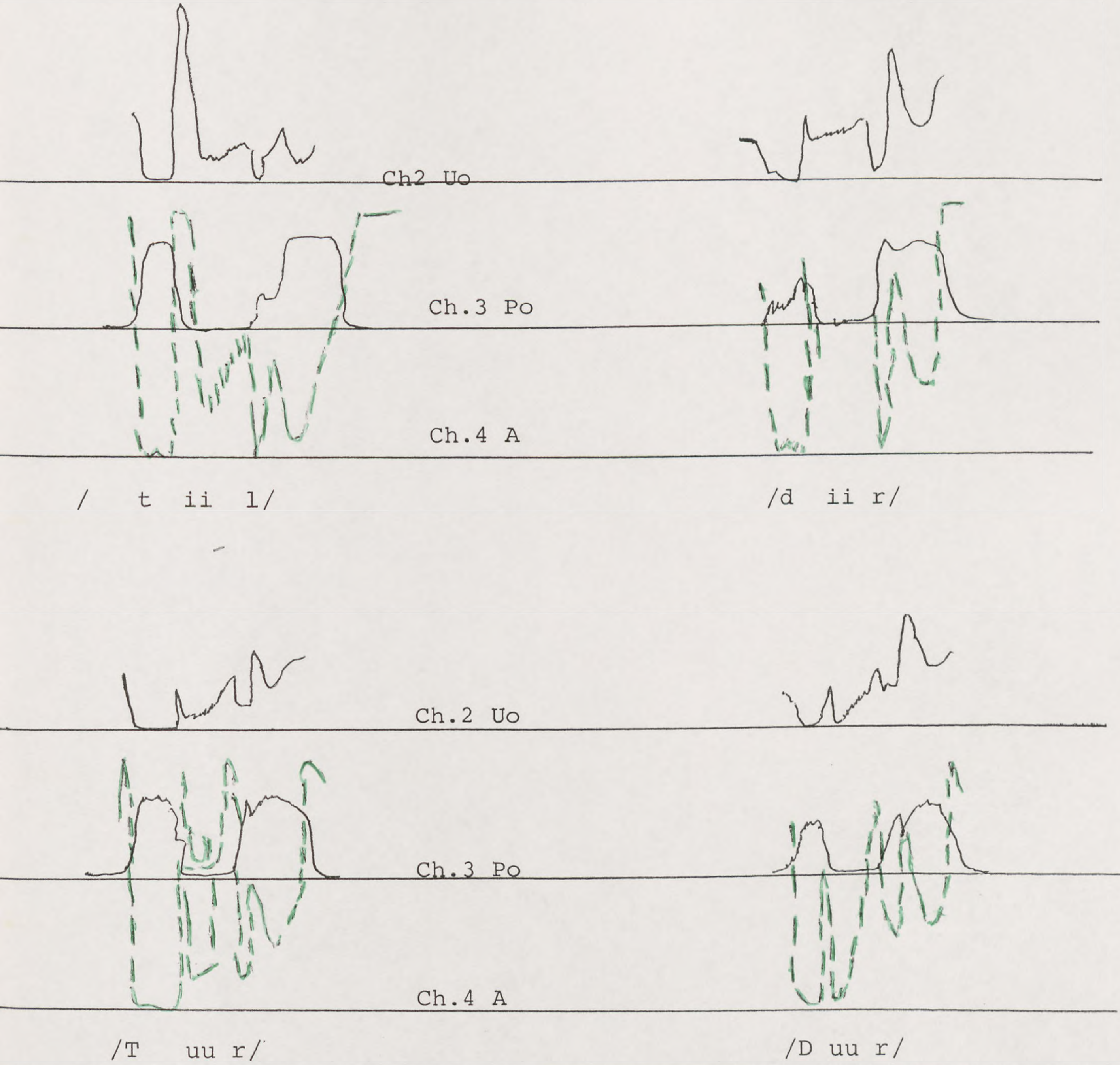


Figure 5.3.1 Mingogram traces of some test words illustrating the oral air pressure and airflow patterns for the voicing distinction.

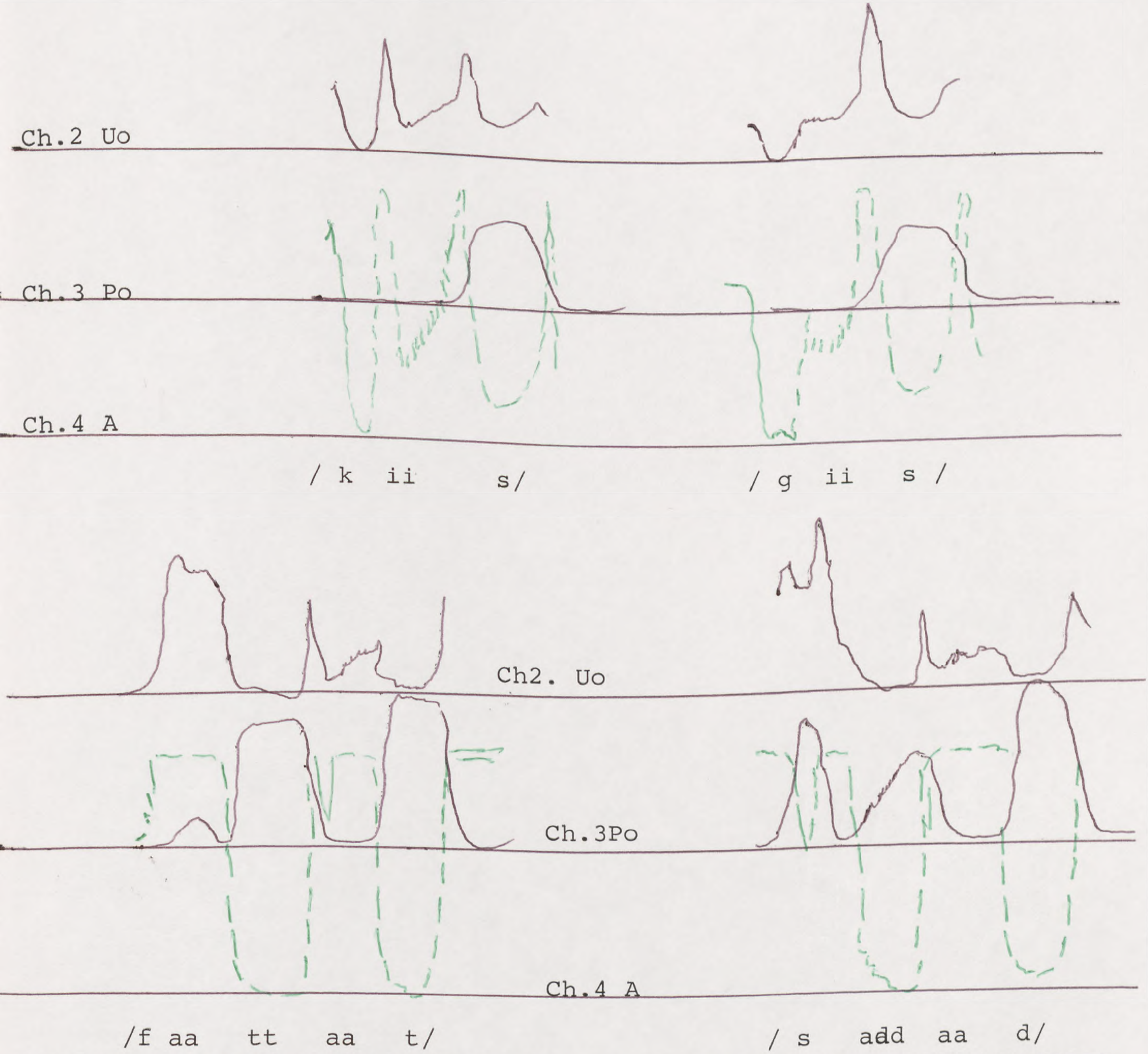


Figure 5.3.1 (Continued).

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/uu/	/aa/	/ii/
/t/	6	7.3	7.4	7.8	6.8	7.3
		(0.2)	(0.7)	(0.7)	(0.8)	(0.4)
/d/	6	3.3	3.4	3.8	3.1	3
		(1.1)	(0.4)	(0.4)	(0.3)	(1.1)
/T/	6	7.3	7.1	7.5	11.8	10.3
		(0.7)	(0.9)	(0.4)	(3.0)	(2.9)
/D/	6	2.4	4	3.3	2.9	2.5
		(0.9)	(1.2)	(0.9)	(0.6)	(0.6)
/tt/	6	-	-	-	8	-
		-	-	-	(1.3)	-
/dd/	6	-	-	-	5.5	-
		-	-	-	(0.8)	-
/TT/	6	-	-	-	8.4	-
		-	-	-	(1.0)	-
/DD/	6	-	-	-	5.9	-
		-	-	-	(1.2)	-

Table 5.3.1 Mean peak oral pressure values (cm H2 O) and standard deviations for the word-initial and medial voicing contrasts (subject AA).

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/uu/	/aa/	/ii/
/t/	6	9.4	10.5	10.1	9.4	10.4
		(0.9)	(0.9)	(0.6)	(1.4)	(0.8)
/d/	6	6.1	8.2	8.0	3.8	6.9
		(0.9)	(0.8)	(0.8)	(1.2)	(0.6)
/T/	6	10.0	10.9	10.5	9.2	10.2
		(0.8)	(1.8)	(0.9)	(1.6)	(0.6)
/D/	6	4.5	4.1	5.2	3.7	3.4
		(0.8)	(0.5)	(1.3)	(0.9)	(0.6)
/tt/	6	-	-	-	10.4	-
		-	-	-	(0.7)	-
/dd/	6	-	-	-	5.8	-
		-	-	-	(1.2)	-
/TT/	6	-	-	-	11.0	-
		-	-	-	(1.5)	-
/DD/	6	-	-	-	5.7	-
		-	-	-	(1.0)	-

Table 5.3.2 Mean peak oral pressure values (cm H2 O) and standard deviations for the word-initial and medial voicing contrasts (subject WS).

5.3.2 Pressure duration and voicing

The mean pressure duration values (Pd) and the standard deviations for the voicing cognates are shown in Tables 5.3.3 and 5.3.4. The voiceless sounds in word-initial position are consistently produced over a longer pressure duration than their voiced counterparts. The (Po) and (Pd) covary in such a way that the voiceless plosives are not only characterised by higher oral pressure values but also by longer durations as compared with the voiced ones. The pressure duration is as effective in distinguishing the voicing pairs in YSA as the peak oral pressure. The pair /Tʊʊr/ and /Dʊʊr/ show (Pd) values of 90.8 and 62.5 msec respectively for subject AA. For subject WS the values are 106.7 and 82.5 msec for the same pair. These results show a difference of about 20 msec. This remark applies to all pairs regardless of vocalic context and position. It is also true for both subjects.

In medial position the (Pd) values reinforce the remarks made about the voicing distinction in initial position. Similar patterns are observed with the voiced plosives having consistently lower pressure durations than the voiceless ones. The (Pd) values found in medial position for both voiced and voiceless sounds are noticeably lower than those in initial position.

Results concerning the relationship between gemination and the voicing distinction show no obvious change from those in initial and medial positions. Thus the pair /fattaat / and /saddaad/ displays (Pd) values of the order 139.2 and 123.3 msec for subject AA. Subject WS shows the values 148.3 and 125.0 msec for the same pair. When comparing geminated and single forms in the voicing categories the differences are quite large. The geminated forms show values which are twice as long as those of the single forms. This large difference is unexpected especially since the peak oral pressure values for the geminated and single forms are not much different. This finding indicates that our earlier observation regarding covariation of (Po) and (Pd) is not a straightforward one. The (Pd) value increase dramatically in geminated forms as opposed to single ones even though no corresponding (Po) increase is found. One more finding of this study is that the pressure duration is a very strong and reliable indicator of gemination in YSA.

There is, in addition, some indication that the (Pd) values undergo some variations in initial position as a function of vocalic context. The results show that in both the voiceless and the voiced sounds the (Pd) values are significantly greater in the environment of /ii/ than in the environment of /aa/. The differences in some environments (e. g. /aa/ and / uu/ or /ii/and /ʊʊ/) are to some extent similar. The consistent and systematic variations in the (Pd) values for the voicing cognates in the environments of /aa/ and /ii/ suggest that the vowel height may be responsible for some of this variation.

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/ʊʊ/	/aa/	/ii/
/t/	6	85.8	92.5	90.8	69.2	81.7
		(4.9)	(7.6)	(6.7)	(6.7)	(4.1)
/d/	6	77.5	77.5	74.2	60	64.2
		(4.2)	(6.9)	(7.4)	(4.5)	(5.9)
/T/	6	95.8	92.5	90.8	77.5	74.2
		(5.9)	(6.9)	(9.2)	(6.1)	(3.8)
/D/	6	66.7	79.2	62.5	55.8	60
		(7.5)	(2.0)	(2.7)	(5.9)	(3.2)
/tt/	6	-	-	-	139.2	-
		-	-	-	(8.0)	-
/dd/	6	-	-	-	123.3	-
		-	-	-	(10.3)	-
/TT/	6	-	-	-	136.7	-
		-	-	-	(8.8)	-
/DD/	6	-	-	-	128.3	-
		-	-	-	(11.3)	-

Table 5.3.3 Mean pressure duration values (msec) and standard deviations for the word-initial and medial voicing contrasts (subject AA).

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/ʊʊ/	/aa/	/ii/
/t/	6	110.8	119.1	110.8	85.8	104.2
		(5.9)	(11)	(7.4)	(8.0)	(4.9)
/d/	6	82.5	99.2	88.3	66.7	81.7
		(8.2)	(3.8)	(6.1)	(7.5)	(2.6)
/T/	6	85.8	108.3	106.7	72.5	85.0
		(3.8)	(9.3)	(7.5)	(9.9)	(5.8)
/D/	6	75	76.7	82.5	67.5	59
		(7.8)	(10.3)	(6.9)	(6.9)	(7.4)
/tt/	6	-	-	-	148.3	-
		-	-	-	(8.8)	-
/dd/	6	-	-	-	125	-
		-	-	-	(11.0)	-
/TT/	6	-	-	-	138.0	-
		-	-	-	(8.2)	-
/DD/	6	-	-	-	115.8	-
		-	-	-	(12.4)	-

Table 5.3.4 Mean pressure duration values (msec) and standard deviations for the word-initial and medial voicing contrasts (subject WS).

5.3.3 Peak oral airflow and the voicing distinction

The mean peak oral airflow values (Uo) and the standard deviations for the voicing contrasts are displayed on Tables 5.3.5 and 5.3.6. In word-initial position the voiceless plosives show higher (Uo) values than the voiced ones for both subjects. For example, the plosive contrast /k/ and /g/ in the pair /kaad/ and /gaad/ obtained values of the order 800 and 362 cm^3/sec for subject AA. Subject WS shows values of 1270 and 410 cm^3/sec for the same pair respectively. This result is not unexpected. It gives further support to the claim that the peak oral airflow is largely determined by the pressure build-up behind the articulatory closure (Isshiki and Ringel, 1964). It is likely that the higher (Po) values found for the voiceless plosives may explain the peak oral airflow variations for the voicing categories. Such pressure-flow relationship is not found in the study of emphasis. This remark cautions against generalistion from one phonetic feature to another.

The results for both voiced and voiceless plosives in the vicinity of the three vocalic contexts reveal higher (Uo) values for these sounds in the context of /aa/ than in the context of /ʊʊ/. The lowest values are displayed in the environment of /ii/. This (Uo) variation may have to do with the vowel height and the size of constriction of the following vowel. It is likely that the narrow constriction of /ii/ is a factor contributing to the (Uo) reduction for the preceding voiced category. Similarly, the higher (Uo) values in the environment of /aa/ may be a reflection of the more open articulation in this vowel.

Another factor which is of interest to us in this study is the place of articulation and its possible effect on the (Uo) pattern found for the voicing distinction in YSA. Three voiceless plosives representing three places of articulation are investigated: the post-dental emphatic /T/, the post-dental nonemphatic /t/, and the velar /k/. In addition, the voiced sounds have three places of articulation. These are the post-dental non-emphatic /d/, the post-dental emphatic /D/ and the velar sound /g/. The results show that the post-dental nonemphatics are the ones produced with the highest (Uo) values, next come the velars and the lowest values are seen with the post-dental emphatics. The three places show progressive (Uo) values for the voiceless sounds /T/, /k/, and /t/ in the words /Taar/, /kaad/ and /taab/. Similarly, with the voiced sounds the order effect is /D/< /g/ < /d/ in the words /Daar/, /gaad/ and /daar/. The same order effect is observed for these sounds in the context of /ii/ and /ʊʊ/.

In word-medial position similar (Uo) patterns for the voicing distinction to that in initial position are seen. However, the (Uo) results reveal that a slight reduction is associated with those in medial position as compared with those in initial position.

There are some differences caused by the vocalic context in this position. The two voicing categories show higher (Uo) values in the environment of /aa/ than in the environments of /ʊʊ/ or /ii/. In addition, they do not differ much in the context of /ʊʊ/ and /ii/. This result gives further support for the remark made earlier regarding the influence of vowel height and the size of constriction on the neighbouring consonants.

The order effect for the place of articulation in this position undergoes some changes. The (Uo) values diminish for velars, post-dental non-emphatics and post-dental emphatics in that order. This suggests that the place effect on the peak oral airflow is not consistent in the two positions. The same airflow patterns for the voicing distinction are retained in the geminated context. The voiceless geminated plosives show higher (Uo) values than the voiced geminated. The geminated forms, in the voiced and the voiceless classes, are observed to have only a slight increase over their single counterparts indicating that the peak oral airflow measure is not as effective as the pressure duration in distinguishing geminated from single forms in YSA.

		Word-initial position			Word-medial position	
Vocalic Context →	N	/aa/	/ii/	/ʊʊ/	/aa/	/ii/
Voicing Contrasts↓						
/t/	6	1020	867	927	742	849
		(83)	(134)	(266)	(103)	(139)
/d/	6	450	407	440	432	370
		(43)	(71)	(83)	(76)	(63)
/T/	6	278	255	242	120	250
		(50)	(32)	(33)	(105)	(49)
/D/	6	202	288	243	188	125
		(18)	(26)	(43)	(37)	(56)
/k/	6	800	677	493	965	728
		(155)	(121)	(73)	(301)	(106)
/g/	6	362	248	255	415	151
		(54)	(26)	(34)	(58)	(89)
/tt/	6	-	-	-	820	-
		-	-	-	(159)	-
/dd/	6	-	-	-	508	-
		-	-	-	(85)	-
/TT/	6	-	-	-	208	-
		-	-	-	(173)	-
/DD/	6	-	-	-	497	-
		-	-	-	(123)	-
/kk/	6	-	-	-	1063	-
		-	-	-	(116)	-
/gg/	6	-	-	-	453	-
		-	-	-	(63)	-

Table 5.3.5 Mean peak oral airflow (cm^3/sec) and standard deviations for the word-initial and medial voicing contrasts (subject AA).

		Word-initial position			Word-medial position	
Vocalic Context →	N	/aa/	/ii/	/ʊʊ/	/aa/	/ii/
Voicing Contrasts↓						
/t/	6	1173	757	1127	657	617
		(203)	(105)	(113)	(114)	(94)
/d/	6	533	458	632	262	383
		(40)	(67)	(83)	(96)	(56)
/T/	6	182	269	221	250	270
		(37)	(76)	(34)	(49)	(42)
/D/	6	251	295	208	230	305
		(43)	(37)	(54)	(76)	(61)
/k/	6	1270	622	683	965	790
		(164)	(181)	(164)	(185)	(73)
/g/	6	410	622	273	398	387
		(164)	(181)	(59)	(76)	(88)
/tt/	6	-	-	-	695	-
		-	-	-	(110)	-
/dd/	6	-	-	-	477	-
		-	-	-	(99)	-
/TT/	6	-	-	-	128	-
		-	-	-	(66)	-
/DD/	6	-	-	-	135	-
		-	-	-	(60)	-
/kk/	6	-	-	-	1130	-
		-	-	-	(92)	-
/gg/	6	-	-	-	452	-
		-	-	-	(52)	-

Table 5.3.6 Mean peak oral airflow (cm^3/sec) and standard deviations for the word-initial and medial voicing contrasts (subject WS).

.3.4 Articulatory closure duration and the voicing distinction

The mean closure durations and the standard deviations for the voicing distinction as produced by the two subjects are displayed in Tables 5.3.7 and 5.3.8. In word-initial position the voiceless plosives show a slight increase in the closure duration (Uc) over the voiced ones. In some occurrences, the results show that some contrasted pairs have approximately similar mean (Uc) values for the voiced and the voiceless consonants. For example, the pairs /taab/ and /daar/ obtain values of the order 53 and 53 msec for subject AA. Similarly, the pair /kʊʊt/and /gʊʊt/ as produced by subject WS shows values of 62.5 and 62.5 msec respectively. This result is clearly in agreement with Kent and Moll (1968) who find similar articulatory closure durations for the voiced and the voiceless plosives. In spite of that, the voicing contrasts in other contexts show clear differences as can be seen in the pair /Taar/ and /Daar/ where the mean (Uc) values obtained are 83 and 41 msec respectively.

Concerning the effects of other factors, the voiceless plosives show quite similar (Uc) values in the three vocalic contexts. The voiced plosives show some variations in their closure duration values in the context of /aa/ and /ii/ and in the context of /aa/ and /ʊʊ/. In the environment of /ii/ and /uu/ they are quite similar.

Similarly, the place of articulation results show different order effect in the two voicing categories. Thus, for voiceless plosives the order shown is as follows : /T/ > /t/ > /k/. This pattern is consistent in the vocalic contexts /ii/ and /aa/. In the voiced ones the order is /d/ > /g/ > /D/. There is a lack of consistency. It is not clear why the articulators would require a longer time for /T/ than for /t, k / and for /d/ than for/g , D /. It may be that the tip of the tongue involved in the production of /T, t, d, D/ is faster than the back of the tongue which is involved in the production of /k, g/. If that is the case, then it is difficult to justify the longer closure duration obtained for /g/ over / D/. The relationship appears to be more complicated and there may be other interacting factors underlying these variations.

As far as the voicing distinction is concerned the results in medial position reveal similar pattern to that in initial position. Both the voiced and voiceless plosives, however, are observed to have lower mean (Uc) values in this position as compared with the initial one.

The closure durations for the geminated voiced and voiceless plosives follow basically the same patterns as in the previous two positions, but there is a marked increase for both voicing categories in geminated forms as compared with single forms. An example of this is shown by the pair /fattaat/ and /fataat/ which shows values of 111 and 41 msec respectively. The articulatory closure duration is, therefore, a reliable measure for distinguishing geminated and single forms in YSA.

To test the significance of the differences observed in (Uc) values the results are pooled and arranged in a similar manner to that of (Po), (Pd) and (Uo). The differences beween the two voicing categorieas are highly significant at (p≤0.01) for each subject separately.

5.3.5 Articulatory release duration and the voicing distinction

Tables 5.3.9-10 display the mean release duration values (Ur) and the standard deviations for the voicing pairs. An investigation of the results in word-initial plosives reveals that the voiceless sounds are released over a longer duration than their voiced counterparts. This is true regardless of vocalic context or subject variations. There appears to be a positive relationship between the peak oral airflow and the release duration. The higher the (Uo) value for a certain consonant the longer the (Ur) value will be. This may be another explanation for the higher peak oral airflow values associated with the voiceless plosives.

The effect of the vocalic context on the release duration for the voiced and the voiceless sounds is also considered. No difference in articulatory release duration for a particular voicing category is found in the environment of /aa/ and /υυ/. In the environment of /ii/ and /υυ/ or /ii/ and /aa/ quite clear differences between these sounds are found. The voiceless plosives show highly significant differences in the environment of /aa/ and /ii/.

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/υυ/	/aa/	/ii/
/t/	6	53	53	48	41	47
		(13)	(4)	(6)	(7)	(6)
/d/	6	53	57	52	41	42
		(6)	(7)	(7)	(2)	(4)
/T/	6	74	69	67	63	57
		(8)	(7)	(8)	(3)	(5)
/D/	6	38	35	34	26	35
		(7)	(5)	(6)	(10)	(6)
/k/	6	33	30	38	34	44
		(4)	(7)	(8)	(9)	(6)
/g/	6	39	43	39	44	53
		(12)	(6)	(10)	(2)	(10)
/tt/	6	-	-	-	111	-
		-	-	-	(10)	-
/dd/	6	-	-	-	103	-
		-	-	-	(11)	-
/TT/	6	-	-	-	117	-
		-	-	-	(9)	-
/DD/	6	-	-	-	104	-
		-	-	-	(9)	-
/kk/	6	-	-	-	99	-
		-	-	-	(14)	-
/gg/	6	-	-	-	95	-
		-	-	-	(15)	-

Table 5.3.7 Mean closure durations (msec) and standard deviations for the word-initial and medial voicing contrasts (subject AA).

On the other hand, the release duration may be affected by the place of articulation. For this measure the order effect is /k/ > /t/ > /T/ in the voiceless plosives and /g/ > /D/ > /d/. These orders do not remain constant in different vocalic contexts.

		Word-initial position			Word-medial position	
Vocalic Context →	N	/aa/	/ii/	/ʊʊ/	/aa/	/ii/
Voicing Contrasts↓						
/t/	6	80.8	77	78	41	47
		(5)	(9)	(14)	(12)	(6)
/d/	6	59	69	65	42	47
		(11)	(7)	(5)	(5)	(6)
/T/	6	83	87	86.7	48	56
		(4)	(10)	(8)	(14)	(5)
/D/	6	41	49	50	40	25
		(2)	(13)	(8)	(3)	(0)
/k/	6	57	55	62.5	50	62
		(3)	(5)	(4)	(8)	(7)
/g/	6	49	57	62.5	41	55
		(7)	(4)	(8)	(5)	(11)
/tt/	6	-	-	-	121	-
		-	-	-	(12)	-
/dd/	6	-	-	-	113	-
		-	-	-	(15)	-
/TT/	6	-	-	-	123	-
		-	-	-	(9)	-
/DD/	6	-	-	-	79	-
		-	-	-	(21)	-
/kk/	6	-	-	-	101	-
		-	-	-	(8)	-
/gg/	6	-	-	-	98	-
		-	-	-	(11)	-

Table 5.3.8 Mean closure durations (msec) and standard deviations for the word-initial and medial voicing contrasts (subject WS).

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/υυ/	/aa/	/ii/
/t/	6	29	43	37	26	38
		(2)	(6)	(3)	(4)	(9)
/d/	6	21	32	23	22	24
		(4)	(3)	(3)	(3)	(2)
/T/	6	18	21	18	17	23
		(3)	(2)	(3)	(4)	(4)
/D/	6	28	29	23	28	34
		(7)	(5)	(3)	(7)	(12)
/k/	6	33	40	38	37	43
		(4)	(3)	(7)	(5)	(3)
/g/	6	32	27	28	27	30
		(13)	(4)	(12)	(5)	(8)
/tt/	6	-	-	-	27	-
		-	-	-	(4)	-
/dd/	6	-	-	-	28	-
		-	-	-	(3)	-
/TT/	6	-	-	-	19	-
		-	-	-	(4)	-
/DD/	6	-	-	-	28	-
		-	-	-	(4)	-
/kk/	6	-	-	-	40	-
		-	-	-	(0)	-
/gg/	6	-	-	-	30	-
		-	-	-	(6)	-

Table 5.3.9 Mean release durations (msec) and standard deviations for the word-initial and medial voicing contrasts (subject AA).

		Word-initial position			Word-medial position	
Vocalic Context → Voicing Contrasts↓	N	/aa/	/ii/	/υυ/	/aa/	/ii/
/t/	6	59	71	63	43	55
		(3)	(9)	(8)	(9)	(5)
/d/	6	25	30	28	36	34
		(5)	(5)	(4)	(9)	(7)
/T/	6	18	30	18	27	27
		(3)	(6)	(4)	(8)	(5)
/D/	6	33	33	18	34	43
		(8)	(9)	(5)	(7)	(8)
/k/	6	68	53	57	51	53
		(12)	(19)	(12)	(6)	(17)
/g/	6	28	38	30	28	31
		(4)	(15)	(11)	(9)	(7)
/tt/	6	-	-	-	33	-
		-	-	-	(5)	-
/dd/	6	-	-	-	26	-
		-	-	-	(2)	-
/TT/	6	-	-	-	18	-
		-	-	-	(3)	-
/DD/	6	-	-	-	38	-
		-	-	-	(12)	-
/kk/	6	-	-	-	48	-
		-	-	-	(7)	-
/gg/	6	-	-	-	33	-
		-	-	-	(3)	-

Table 5.3.10 Mean release durations (msec) and standard deviations for the word-initial and medial voicing contrasts (subject WS).

For the same reasons mentioned in the study of emphasis the data will further be analysed using the ANOVA test. So far the t-test has enabled us to compare two group means at a time for some factors only. It is indicated that there are more than two independent variables interacting with the dependent ones. In order to see the extent of the variations caused by each of these independent variables it is considered necessary to use the ANOVA test. Measurements for the dependent variables of the voicing categories are entered in the same way as the emphatics and the non-emphatics. The independent variables differ with regards to the levels investigated within each factor. There are four factors identified by the symbols P, V1, V and S for the place of articulation, voicing, vocalic contexts and subjects (See Appendix 10).

The levels within each factor are represented by the figures 1, 2 and 3. Thus, the figure (1) stands for denti-alveolar emphatics in C1, voiced in C2, vocalic context /aa/ in C3 and subject AA in C4. The figure (2) stands for denti-alveolar non-emphatics in C1, voiceless in C2, vocalic context /ii/ in C3 and subject WS in C4. The figure (3) stands for velars in C1, vocalic context /υυ/ in C3, see Appendix (11).

In this experiment there are 216 cases for the variables (Uo), (Uc) and (Ur) whereas for (Po) and (Pd) there are only 144 cases. For practical reasons air pressure measurements are not obtained for back consonants. The units for each dependent variable are as given above (See sections 5.1.3 and 5.2.).

The method used for data presentation is the same as that used in emphasis (See section 5.2). The main effect of the independent variables on the dependent ones is shown in Table 5.3.11. A highly significant effect on almost all variables is observed at ($p \leq 0.01$). The results, however, show that the most influential factor on the peak oral air pressure is voicing. The F-value observed for this factor reaches 806.3. The other factors influencing (Po) are subject variations, place of articulation and vocalic context as is seen from their F-values. The same order effect is found for the pressure duration and peak oral airflow with voicing as the primary factor. In fact the (Uc) variation as function of the vocalic context is found to be insignificant at ($p \leq 0.05$). The conclusion which may be drawn from these results is that the voicing effect on all dependent variables is a primary one. The other factors play only a secondary role in their variations.

Table 5.3.12 shows the two-way interaction effects on the dependent variables. The interaction between place of articulation and voicing is found to be highly significant at ($p \leq 0.01$). Thus, the F-values found for all variables, except pressure durations, suggest that this particular combination is the most influential one. The voicing and subject interaction are found to have a nonsignificant effect on the peak oral pressure, the peak oral airflow and the release duration suggesting that the main effect exerted by voicing is a strong one and remains unaffected by these interactions. Other interactions involving the vocalic context appear to result in insignificant p-values.

The three-way interaction results are displayed on Table 5.3.13. A horizontal inspection of this table shows that almost all interactions resulted in a highly significant effect on the peak oral airflow. The interactions among the place, voicing and subject are shown to be significant at ($p \leq 0.05$). The same result is found for the interaction between the place, vocalic context and subject. All other three-way comparisons particularly those involving vocalic contexts and place are not significant at the 5% level. The four-way interactions are shown on Table 5.3.14. These interactions exert a highly significant effect on all variables at ($p \leq 0.01$) except Uc.

		Dependent Variables				
ANOVA main effect	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
Place of Articulation	F-ratio	24.54	26.08	428.52	78.27	78.43
	P-value	0.000**	0.000**	0.000**	0.000**	0.000**
Voicing	F-ratio	806.28	352.10	544.44	160.82	123.99
	P-value	0.000**	0.000**	0.000**	0.000**	0.000**
Vocalic Context	F-ratio	10.68	10.42	32.92	1.59	6.502
	P-value	0.000**	0.000**	0.000**	0.207	0.002**
Subj.	F-ratio	356.895	155.50	22.42	284.99	113.49
	P-value	0.000**	0.000**	0.000**	0.000**	0.000**

Table 5.3.11 Analysis of variance showing the main effect of place of articulation, voicing, vocalic context and subject on the dependent variables.

		Dependent variables				
ANOVA	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
2-Way Interactions						
place &	F-ratio	35.42	8.815	140.15	130.73	94.14
voicing	P-value	0.000**	0.004**	0.000**	0.000**	0.000**
place &	F-ratio	0.408	0.733	26.88	2.159	2.408
Voc. Context	P-value	0.666	0.482	0.000 **	0.075	0.051*
place &	F-ratio	12.344	12.694	8.96	4.394	12.73
subj.	P-value	0.001**	0.001**	0.000**	0.014**	0.000**
voicing &	F-ratio	0.780	0.691	14.16	3.27	0.839
Voc. Context	P-value	0.461	0.503	0.000**	0.040*	0.434
voicing &	F-ratio	0.249	8.815	1.49	22.499	43.122
subj.	P-value	0.619	0.004**	0.224	0.000**	0.000**
Voc. Context &	F-ratio	1.404	2.835	9.93	4.26	0.667
subj.	P-value	0.250	0.063	0.000**	0.016*	0.515

Table 5.3.12 Analysis of variance showing the two-way interaction effects of place of articulation, voicing, vocalic contexts and subject on the dependent variables.

		Dependent Variables				
ANOVA	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
3-Way Interactions						
place, voicing &	F-ratio	0.152	0.268	4.704	0.605	2.011
voc. context	P-value	0.859	0.765	0.001**	0.660	0.095
place, voicing &	F-ratio	27.47	2.031	6.680	3.116	12.26
subj.	P-value	0.000**	0.157	0.002**	0.047*	0.000**
place, voc. con. &	F-ratio	3.65	5.120	5.514	0.677	0.290
subj.	P-value	0.029*	0.007**	0.000**	0.609	0.884
voicing, voc. con.	F-ratio	2.390	1.88	3.996	0.897	2.499
subj.	P-value	0.096	0.158	0.020*	0.410	0.085

Table 5.3.13 Analysis of variance showing the three-way interaction effects of place of articulation, voicing, vocalic contexts and subject on the dependent variables.

		Dependent Variables				
ANOVA	Value	**Po**	**Pd**	**Uo**	**Uc**	**Ur**
4-Way Interactions						
place, voicing, voc. context	F-ratio	4.462	5.25	3.71	0.556	3.797
subj.	P-value	0.014**	0.007**	0.006**	0.695	0.005**

Table 5.3.14 Analysis of variance showing the four-way interaction effects of place of articulation, voicing, vocalic contexts and subject on the dependent variables.

5.3.6.1 Discussion

The results of this study show significantly higher peak oral air pressure values for the voiceless plosives than for the voiced ones in YSA. The finding is consistent with previous studies (e.g. Black, 1950; Arkebauer, Hixon and Hardy (1967) Dixit and Brown, 1978, Miller and Daniloff 1977, Subtelny, Worth and Sakuda (1966). All of them report higher (Po) values for the voiceless plosives as compared with the voiced ones. There are some minor differences which need to be mentioned. For example, the (Po) values obtained in the early study by Black are very high. If converted to cm H_2O the values would be in the range from 105.3 cm H_2O to 185.2 cm H_2O. The discrepancy may perhaps be attributed to the transducer system in use at that time.

The results are not in agreement with that of Lisker (1970). Lisker finds significant amount of overlap in the oral pressure values for the two voicing categories. The apparent incompatibility in these findings could be explained if the glottal and the supraglottal strategies used by the subjects of the two studies were known. Lisker's speech material is based on one American subject. In this language the voicing contrast /t/ vs /d/ in medial position is believed to be neutralized with some American subjects. This in part may explain the lack of contrast of oral air pressure in his study. Since there is only one subject the effect may be a reflection of idiosyncratic factors. Contrary to our results, Nihalani (1974) does not find any differences in the (Po) values between the voiced and the unvoiced both in the aspirated and the unaspirated class of Sindhi plosives. Similarly, Dixit and Brown (1978) from data on Hindi plosives, another four category language, conclude that differences in peak oral pressure values for the voiced and voiceless plosives are only significant in the aspirated class but not in the unaspirated one. In YSA both the voiced and the voiceless plosives are unaspirated and yet highly significant differences between them are found in all positions and contexts examined.

Similar conclusions for English and some other languages that have aspiration and voicing in their phonetic system, are also reached. Dixit and Brown (1978) explain the contradiction between their results and those of others by suggesting that the variety of the unvoiced plosives used in English studies, for example, are in fact of the aspirated class, thus implying that the differences in peak oral pressure values in the voicing categories may have been caused by aspiration rather than voicing. Assuming that such explanation may be true for English which contains aspiration and voicing in its phonological system, this generalization cannot be taken for granted in YSA because this language does not make use of aspiration noise in the distinction between its categories. In our view the lack of (Po) variations between the voiced and the unvoiced unaspirated plosives in Hindi may be a special case of that language and perhaps other four categories languages like Sindhi. As the results of this study show there are subject variations as well as language variations which cannot be ignored. Thus any generalization regarding the superiority of a specific phonetic feature over the other should take into consideration these subject and language-specific differences.

Several explanations have been advocated as possible reasons for the pressure variations in the voicing distinction. The most likely one may be the glottal actions at the time of closure and release. Results from previous research seem to point towards this conclusion (Lisker et. al.; 1969 Warren 1976; Van den Berg 1957; Van den Berg et. al. 1958; Rothenberg, 1968). Lisker et. al. observe that as the size of the oral port increases after the release of the voiceless plosives, the oral resistance decreases. In correspondence with that the oral airflow increases since the resistance at the glottis continues to remain low as a result of the vocal folds being abducted. According to them the vocal folds are adducted and the glottis is closed and remains closed at and after the release of the oral closure for voiced plosives in English. Although, the oral resistance to peak oral airflow has greatly decreased after the release and the articulators' separation, the glottal resistance is still present. It is, therefore, reasonable to assume that this reduction in the peak airflow and air pressure values for the voiced sounds is caused by this glottal impedence. It is suggested that a decrease or increase in the peak oral air pressure is correlated with the presence or absence of the vocal folds vibrations and glottal impedence.

Rothenberg (1968) states that resistance at the glottis varies from less than 1 cm $H2$ O/Liter per second during quiet breathing to as much as 100 cm $H2$ O/Liter per second during voicing. For example, the difference between subglottal and oral pressure during vowel articulation is believed to be approximately between 4-8 cm $H2$ O and between 1-4 cm $H2$ O during vowel/consonant articulation. He adds that during phonation the oral air passage produces resistance of less than 1 cm $H2$ O /Liter per second for certain vowels to an infinite resistance for plosives. He observes a larger transglottal pressure drop with voiced plosives as compared with the voiceless ones. Similarly, Van den Berg (1957) finds that glottal tissue resistance increases as the area of constriction decreases. From these observations one may conclude that the variation in peak oral air pressure values for the voicing cognates in YSA is to a large extent determined by the glottal action. A more constricted glottis may result in a higher transglottal pressure drop. This in its turn may correspond with lower peak oral pressure values and vice versa.

On the other hand, the vocal folds action is not enough by itself. It has to be maintained by the aerodynamic forces of air pressure and airflow as well as the relevant muscular contraction. Without these aerodynamic forces, the vocal folds may not be able to perform their function. For this reason voicing may be regarded as an aerodynamic phenomenon which requires both vocal folds action and pressure variations.

Another factor which may have contributed to the depressed (Po) values of voiced plosives is the expansion of the pharyngeal cavity. This action is believed to delay the moment when air pressure below and above the glottis is equalized in the vocal tract. Some researchers, however, appear to undermine the role of the glottal action in the pharyngeal expansion. Perkell (1969) suggests that the wall of the pharynx is lax during voiced plosives and that this laxing is responsible for the pharyngeal enlargement during the constriction phase of their production when oral air pressure rises, applying a force at the pharynx walls.

Other researchers believe that the pharyngeal expansion associated with the voiced plosives is more likely to be caused by an active articulation than by a passive response to overpressure (Kent and Moll, 1968 , Rothenberg, 1968, Lubker 1973, Bell Berti, 1980). The results from Lubker (1973) also give some evidence of an inverse relationship between the peak oral pressure value and the size of the pharyngeal cavity for the voicing cognates at the time of closure. The voiced plosives are produced with lower peak oral pressure values than voiceless ones.

Another possible explanation for the reduced (Po) values in the voiced plosives is reported by Bell-Berti (1980). It is observed that the voiced plosives are associated with "greater levator palatini" activity and higher velar elevation than their voiceless counterparts. Such an action could result in an increase in the pharyngeal volume during the closed phase of voiced plosives. This again may explain the lower (Po) values found for these sounds.

Articulatory duration may be another reason for the (Po) variations in the voicing cognates. This study shows that the (Po) and the (Uc) values vary together for the same test word of YSA. Higher (Po) values are associated with longer pressure duration and longer closure duration values. Such a finding is consistent with Miller and Daniloff (1977) Malècot (1966) and Subtelny et. al. (1966). Miller and Daniloff (1977) report a positive correlation between consonantal closure duration and oral air pressure for the voiced plosives.

5.3.6.2 Place of articulation

The order effect for the different places of articulation with regard to the peak oral airflow is as follows: /T/ < /k/ < /t/ in initial position and /T/ < /t/ < /k/ in medial position. For the voiced plosives the order effect is /D/ < /g/ < /d/ in initial position and /D/ < /d/ < /g/ in medial position. These order effects are consistent regardless of the change in vocalic environments /aa/, /ii/ and /ʊʊ/. These results are partly consistent with those of Gilbert (1973). He reports /t/ to have significantly higher (Uo) values than either /k/ or /p/. Similarly, his results show the voiced plosive /d/ to exihibit higher (Uo) values than either /b/ or /g/ in the initial and the medial positions.

Contrary to the results of this study Dixit and Brown (1986) find nonsignificant differences in the peak oral airflow values as function of place of articulation for Hindi plosives.

There appears to be no systematic relationship between the size of the oral cavity behind the point of constriction and the peak oral airflow values for the different places of articulation. Although the place differences are quite apparent in our study they may have been caused by phonetic position and other related factors. As the results have shown the order effect is quite different in the two positions. The denti-alveolar emphatics /T/ and the uvular /q/ places of articulation are shown to have non-significant (Uo) differences in the voiceless category. Similarly, the denti-alveolar nonemphatics /t/ and the velar /k/ places of articulation are shown to have non-significant (Uo) differences. This confirms that the source of variability in these sounds is caused by emphasis rather than by the place of articulation. When the feature emphasis is neutralised (e.g . /T/ vs /q/ and / t/ vs / k/) in these voiceless plosives non-significant differences are obtained. However, when /T/ is compared with /t/ and /q/ with /k/ (see section 5.2) significant differences are found. Therefore, these differences are caused by emphasis rather than by the place of articulation.

As far as the closure duration is concerned the denti-alveolar sound /T/ is observed to have longer closure duration values than /t/ in initial and medial positions. The order effect is /k/ < /t/ < /T/ in initial position and /t/ < /k/ < /T/ in medial position. The voiced sounds have shown that the closure durations decrease in this order: /d/< /g/ < /D/ in initial position and /g/ < /d/ < /D/ in medial position. These results again confirm the remarks made above that emphasis is more likely to be responsible for these variations than the place of articulation. These place-related-differences may be attributed to the speed of articulators. The pharyngeal secondary articulation during the production of the emphatics /T/ and /q/ may have caused the delay time in the closure duration (see section 5.2.6 for more details on this point).

In brief the place of articulation effect in our study is not systematic. It is inconsistent in the two positions and varies from one vocalic environment to the other. The sources of variability in (Uo), (Uc) and (Ur) appear to be the outcome of more than one factor (e.g. emphasis, position, vocalic context and subject).

5.3.6.3 Vocalic Contexts

Results from the previous two sections (5.2.) and (5.3) suggest that differences in vocalic context can affect the peak oral air pressure values for the plosive consonants in YSA. All consonants show higher peak oral air pressure when in /ii/ and /uu/ context than in /aa/ context. However, when all measurements are pooled and subjected to ANOVA test the differences are shown to be insignificant (P= 0.249).

The two voicing categories are significantly higher in the presence of /uu/ and /ii/ than in /aa/. This last finding is in good agreement with Karnell and Willis (1982) who indicate that both unvoiced and voiced bilabial plosives have higher mean peak oral air pressure when in the context of /u/ than when in the context of /a/. It is also consistent with Brown et. al. (1973) who investigate the lingual pressure and the oral air pressure. They find air pressure values for /t/, /d/ and /n/ in /u/ and /a/ contexts vary across three intensity levels. Lingual pressure shows similar trends. Ashley et. al. (1985) conclude from a study on vocalic contexts that consonants in the environment of /i/ have significantly greater oral pressure values than the same consonants with /a/.

A possible explanation for these (Po) variations as function of vocalic context may be understood with reference to velar elevation associated with the different vowels. Previous studies on velo-pharyngeal function have repeatedly shown that the velar height varies depending on whether the vocalic context is /aa/, /uu/,or /ii/ (Moll 1962; Fritzell;

1979; Bell Berti 1980). Moll concludes that velar port closure and hence velar-elevation is greater for high vowels than for low vowels. Similarly Fritzell in a review of the velopharyngeal function reached the conclusion that the velar height increases through the vowel series [a] [e] [o] [u] and [i]. He also adds that such variations in velar height as function of vowel quality has a direct influence on the 'levator activity' for the neighbouring sounds. Bell Berti (1980) based on data obtained by a fiberoptic endoscope concludes that the velar elevation is greater for [i] than for [a] and more importantly that vowel environment has a significant influence on velar elevation for consonants. These findings may explain why the plosives in the proximity of /ii/ and /uu/ are produced with higher (Po) values than those in the proximity of /aa/. As the vowel height and position move from low back /aa/ to high front /ii/, peak oral air pressure as well as duration for the preceding consonants are expected to increase.

On the other hand, the closure duration for the plosives in the voiced-voiceless and the emphatic-nonemphatic categories are quite similar in the two vocalic contexts. Consonants in the presence of /aa/ are shown to have a shorter release duration than those of /ii/.

Chapter 6: Summary and Conclusions

The foregoing chapters of this study are an attempt to elucidate certain phonetic and phonological aspects of the voiced/voiceless and emphatic/nonemphatic plosives in Yemeni Spoken Arabic. Most conclusions to be drawn from them are already set out in the relevant places in the preceding chapters. A few general points will be reproduced in the following sections for convenience together with suggestions for further research which have been stimulated by the various results of the present study.

6.1 Acoustic analysis:

The study begins with an acoustic investigation of voicing and emphasis in three related experiments. In the first one a detailed investigation of the voice onset time in citation forms and connected speech is conducted. The results show the voicing categories to be associated with characteristically distinct acoustic properties reflected by the varied VOT values for the two sets of consonants on the VOT dimension. They occupy two ranges: the voicing lead and the short voicing lag. Some voiceless plosives (e. g. /k/ and /t/ in the environment of /ii/ extend their range to the long lag. Although the differences are found to be statistically significant indicating the robustness of this measure as a reliable correlate of voicing in YSA, instances of failure have also been found as shown by the overlap of some occurrences for the voiced and voiceless emphatics. This finding indicates that the two features (i.e. emphasis and voicing) overlap. Consequently, it is predicted that other acoustic cues may be operating together with VOT. Some of these are the formant transitions, the acoustic closure, the durational relationship between consonants and vowels. In addition, other phonetic phenomena such as the place of articulation, vocalic context and gemination are expected to contribute to that interaction. Although the findings of that analysis are clear-cut and conclusive, they also point to the necessity of carrying out more investigations on another phonological contrast and on another dimension. Accordingly, the first three formant transitions from the emphatic/nonemphatic consonants to the following vowel are explored. The consonantal loci and the transitional targets at the following vowels are measured. It is assumed that the transitional movements vary for different consonant-vowel combinations and consequently display characteristically varied transition information. The results of this experiment demonstrate that the major acoustic feature operating on this contrast is the onset frequency of formants particularly that of the second formant. With regard to the loci position, the extent of the frequency shifts are very marked. The magnitude of the displacement appears to relate to the emphatic/nonemphatic elements and the extent to which they are linguistically utilized as a contrastive device. The F3 transitions are the same in both consonant categories and so are not likely to be distinctive. Nevertheless a combination of the second and the third formant transitions is found to provide more additional acoustic information for the characterization of these consonants than is shown by each one of them separately.

The data also show spectrographic evidence of progressive and regressive coarticulation. The formant transition effect of the emphatic consonants spreads into the middle of the following vowels. It also extends to the preceding vowel even if it thereby crosses a word-boundary. This latter phenomenon is likely to be the result of articulator positions being modified in the direction of the preceding phoneme representation. That is, the production of the vowels involves concomitant articulatory adjustments more like that of the emphatic consonants. If the secondary articulation is to produce an acoustic effect (for emphasis) then it must be in place either as the plosive closure is made or as the plosive releases. This means that formant evidence of the secondary constriction is still there in the following vocoid. However, this inevitable small extension in time of the secondary articulation and hence acoustic formant effects may well be extended further than it absolutely has to be. Possibly the concomitant effect is allowed to become bigger than it needs, to enhance and strengthen and make more reliable the acoustic and perceptual contrasts between emphatic and non-emphatic plosives and their vocalic context. An other related observation for that phenomenon is that the formant targets, particularly that of the second formant, occupy quite distinct frequency locations for the emphatics and the nonemphatics. These vowel targets are always lower in the former environment than in the latter. All these remarks point to the conclusion that the minimum domain of emphasis is in fact the (CV-) or even (-VCV-) syllables and that this feature should be treated accordingly.

The other variable found to be of influence on these categories is the acoustic closure duration. In this experiment the temporal characteristics of the YSA plosives are examined to determine the extent to which the phonological features are found to control timing and to see if emphasis or voicing may have as one of their acoustic manifestation a difference in any of these three measures: (1) acoustic closure duration (2) the contoid duration and the (3) vocoid duration.

The findings of this investigation have clearly shown that the three durational intervals vary not only with change of cognates within each feature, but also with change in the phonological features that serve to distinguish words. The acoustic closure duration is less than satisfactory in separating the voiced/voiceless cognates but it is of some importance for the emphatic/nonemphatic pairs. The contoid duration is, however, found to be a good indicator of the voicing contrast. The inclusion of the release phase as part of that interval appears to enhance the distinction. The contoid duration is also useful for separating consonants produced with different manner of articulation. The single and geminate fricatives examined in this study are characterised by longer duration than that of the plosives. This may have been caused by the gradual controlled movement of the articulators during the production of the fricative consonants as compared with the "abrupt ballistic movements during the plosive production".

In contrast to Port's (1981) findings for English, the results of this study clearly show that the contoid/ vocoid ratio is an unstable and perhaps redundant measure of voicing. It changes from one context to another. It differs for the two subjects.

Similarly, in distinguishing a voiced from a voiceless consonant, the duration of the following vocoid is found to be a negligible factor. Vowels following voiced consonants are not significantly different from those following voiceless consonants. The data do not provide evidence in support of the concept of temporal compensation (MacNeilage, 1968). For example, no compensatory relation between the lengthening of the closure time for an intervocalic consonant and a following vowel is found. The increase in the closure duration for the geminate consonants is accompanied by a shortening in the duration of the following vowels. On the surface this may look like a compensatory process. There are certain points that need to be clarified. First, the compensation is not complete in the sense that the closure duration for the geminate consonant and the following vowel durations do not increase and decrease respectively by the same duration. The increase associated with the former is much greater than the decrease in the latter. Moreover, the compensatory temporal adjustment requires the syllable units to have equal duration and if one segment is lengthened or shortened then that is expected to be compensated by a neighbouring one (MacNeilage, 1968). This requirement is not met in our data. Some investigators (e.g. Homma, 1981) maintain that the compensatory process is not only confined to the syllable units but also operates over whole words. Since our material is not particularly suitable for checking that possibility, this last claim cannot be accepted or rejected until it is tested by other appropriate material.

6.2 Perceptual analysis:

Perceptual tests using synthetically generated speech are also used with the aim of assessing the auditory effect of variation in certain acoustic cues on the listeners judgements regarding the phoneme boundary. Two variables are manipulated in two separate experiments. These are VOT and the second formant transitions.

These cues are not only shown to be acoustically effective in separating their respective contrasts but also but also they maintain the same discriminatory power in our perceptual tests presented to the YSA listeners.

Results show that the shift from the voiced category to the voiceless one is quite sharp. The listeners can distinguish the voiced from the voiceless consonants when the 50% crossover point occurs within the VOT range -17 ms and 14 ms. The voiced plosives are not heard without a voicing lead. This contrasts with results for English in which the /d/ sound is perceived at the short lag region and /t/ at the long lag region (Abramson and Lisker, 1965, 1968).

The findings confirm our predictions that the relative onset time between the plosive release and the voicing onset is a very reliable indicator of voicing not only acoustically but also perceptually.

In our perception experiment two cues are varied in two separate tests (1) The F2 onset frequency and (2) The F2 steady state frequency to determine the phoneme boundary for the emphatic/nonemphatic categories. The findings of this investigation show that the average onset frequency at the 50 % response crossover point ranges between 1274 Hz and 1620 Hz. Thus one would limit the region of emphasis between these two values. Similarly, the average F2 steady state response at the 50 % crossover point occurs at 1193 Hz. The shift from the emphatic response to the non-emphatic one is quite sudden. The emphatics are heard at lower F2 onset frequency and lower F2 steady state frequency than their nonemphatic counterparts. One of the articulatory correlates of this phenomenon is the narrowing of the pharyngeal cavity. This may correspond to the down shift of the F2. Both tests point to the conclusion that the F2 transition and steady state work in the same range and produce similar results when manipulated separately. The role of context as determined by this experiment is as effective in cueing the emphatic/nonemphatic categories as the F2 transitions. The implication of such a finding is that emphaticness is not only a property of the consonant itself but rather it extends to the neighbouring vowels.

6.3 Articulatory and aerodynamic analysis:

Our conclusions for the four phonetic categories from the particular contexts investigated on the articulatory and aerodynamic levels can be summarized as follows:

1. Voiceless consonants are accompanied by higher peak oral air pressure values than their voiced parteners regardless of whether they occur word-initially, medially in geminated forms or in various vocalic contexts and places of articulation.

2. Peak oral airflow is considerably greater for the nonemphatics as compared with the emphatics. However, little difference in peak oral air pressure values or pressure durations is found between the two categories.

3. The differences in peak oral airflow for the emphatics/nonemphatics and air pressure for the voiced/voiceless categories appear to reflect the degree of constriction in the supraglottal tract. But there may be other factors operating , too (e.g. glottal constriction). It may be the most influential factor on the air pressure variations for the voiced and the voiceless categories. As the glottal constriction becomes more severe the airflow resistance increases. Similarly, the secondary articulation associated with the emphatics appears to correspond to the airflow reduction in the emphatics as compared with the nonemphatics. Partial evidence for this explanation is seen when the emphatics are produced with significantly longer articulatory closure duration than their nonemphatic cognates.

4.The voiceless plosives are produced with longer articulatory release than their voiced counterparts. There is also a positive one to one correlation between the peak oral airflow values and the release duration. The higher the peak oral airflow value for a certain category the longer the release duration is found to be and vice versa. This covariation for (Uo) and (Ur) is believed to reflect the same underlying articulatory strategy.

5. From the articulatory point of view there is no evidence that the voiced/voiceless or emphatic/nonemphatic geminates involves a reiteration of the same contoid. It seems to be articulated as one whole long unit with no indication of double articulation.

6.The results also show that duration plays a primary role in distinguishing single from geminate consonants.

7. There is no systematic air pressure patterns that are shown to correspond to the cavity size behind the place of articulation.

8. Position in the word, vocalic context as well as between subject variations are found to contribute in different ways to the aerodynamic changes associated with the phonetic categories investigated in this study.

The perception of the voicing contrast are found to require +VOT vs -VOT and the acoustic analyses show this for /t/ vs /d/ and /T/ vs /D/. *But VOT values are not the same when emphatic /T/ and non-emphatic /t/ are compared. It is indicated that the main acoustic cue for emphasis is the second formant transitions. These are produced (it seems) from a specific change to articulation i.e. addition of a secondary articulation. As a side effect (concomitant effect) of this extra constriction, airflow is greatly reduced for emphatics, during the release phase (i. e. from plosive release to the start of voicing). The data show that airflow and long release phase go together. They are positively correlated. Higher airflow for /t/ than for /T/

goes with longer VOT for /t/ than for /T/. Therefore, the main essential articulatory contrast for emphasis creates (inevitably perhaps) an additional acoustic contrast - different VOT values. This seems to be acceptable because the pairs /t/ vs /d/ and /T/ vs /D/ are still separated by VOT. Thus the small difference in VOT is used by listeners to judge whether they hear /T/ or /t/. This suggests that VOT may be a secondary acoustic cue of emphasis.

On the other hand, there is no evidence that the emphatics are associated with stronger air pressure than the nonemphatics. It seems, therefore, inappropriate to describe the emphatics as phonemes with 'strong pressure' or as 'heavy' consonants (Jakobson, 1957). The use of emphasis as a cue to the identification of stressed syllables is not supported by this study since both the emphatics and non-emphatics show almost similar peak oral air pressure values.

6.4 Some implications for foreign language teaching:

The major objective of the present study is to find out the acoustic, perceptual and aerodynamic characteristics of emphasis and voicing as they are realized in the plosives of Yemeni Spoken Arabic. Although the findings contribute directly to the description of the phonemic characteristics that are relevant for speech communication and technology, they also have some implications for the teaching and testing of foreign languages. For example, the study highlights some phonetic and phonological differences between Yemeni Spoken Arabic and some other languages (e.g. English). We believe that such differences in the systems of the two languages should be brought to the attention of the foreign learner willing to learn either of these two languages. This will be of great help in his/her pronunciation learning. Some of the aspects which this study shows, for example, is the absence of aspiration in the voiceless plosives of YSA. The presence of a big voicing lead with voiced plosives indicates that these sounds are rarely devoiced at least in the contexts investigated. Additional places of articulation are also found in YSA as exemplified by the back articulation (i.e. the uvular and pharyngeal articulations). One voicing cognate is missing in the phonemic system of this language, that is the voiceless bilabial plosive /p/. The length feature whether it is associated with the vowel or the consonant (particularly geminates) also functions differently in YSA from English. Furthermore, a whole group of consonants (namely the emphatics) which constitute about a third of the phonemic system in YSA are not found in English or any other European language. Previous research on second language acquisition shows that Arab students learning English frequently confuse the English segment /p/ with /b/. They also misplace the stress. Such problems are found to result in serious communicative problems (Gimson,1989).

English learners of Arabic also find difficulty in pronouncing emphatics and pharyngeal sounds. They also stress syllables that should not be stressed (Mitchell, 1990; Gairdner, 1925). It is our belief that such learning difficulties faced by learners of the two languages can be reduced if the learners are made aware of these differences in the phonetic and phonological systems of the two languages.

These communicative problems arise when the learners employ their own language rules in the organisation of the target language situations. This strategy (transfer) is frequently used by learners of foreign languages. It is defined as the use of the native language rules, items and phrases in the organisation of target language situations (Littlewood, 1984). In some situations some rules appear to be similar in both the target language and the native language. Thus, if the learners find themselves in a difficult communicative situation, they resort to borrowing their own native language rules. Transfer comes in different forms one of which is 'anglification' in which the learners foreignize a word by making appropriate modifications in pronunciation and inflectional endings of nouns, verbs and adjectives in a word. Transfer is a common strategy in the speech of people whose languages come from the same language family but also it is used by learners whose native languages come from a different language family. The relevant point which must be made is that transfer can be negative or positive. In the former the learner tends to generalize and uses it much in his conversations without learning. Because of the learner's frequent transfer to his native language and whether he is in real communicative difficulty or not, the native language rules appear to interfere with his learning. In the latter, on the other hand, transfer is only used in really difficult communicative situations and it leads to learning. In this respect it is the latter that should be encouraged. It leads him to discover the differences and similarities between the first and the second languages. At the same time it is the duty of the teacher to facilitate learning. He/she should help them discover the source of difficuly and accordingly develop suitable materials and use good techniques that will help them get rid of the negative transfer. Yemeni students learning English have difficulty distinguishing certain contrasts (e.g. /p/ vs /b/ /f/ vs /v/ and /ŋ, n/) because YSA lacks one of those components in its phonological system. Such a gap in the learner's knowledge could be filled by presenting him with materials that cover this area and that enable him to make use of the positive transfer. For example, they may be asked to listen in the language laboratory to minimal pair words containing these contrasts. They may also be asked to find or comment on the differences to make sure that they have acquired them. If it is found from their comments that they are still unable to find these distinctions, then it may be necessary for the teacher to explain these contrasts explicitly and ask them to listen again paying attention to the notes he/she told them about.

However, transfer as a strategy of learning may not be suitable to all situations. As this study has shown the realization of certain phonetic and phonological rules is quite different in YSA and English. The relationship between vowel length (syllable duration) and stress in YSA is quite strong to the extent that syllables with long vowels are almost always stressed. It is a primary correlate of stress. In English, on the other hand, duration is only one correlate of stress and in certain contexts it may be overriden by other factors such as the fundamental frequency of voicing or intensity. Consequently, syllables with long vowels should be given their due weight and failure to take this point into consideration may result in misunderstanding.

The second differentiating factor between English and YSA is related to the concept of stress and emphasis. Although in English this concept is often associated with stress, there is little evidence of a positive relationship between emphasis and stress in YSA. As indicated above no real differences in peak oral air pressure exist between emphatics and non-emphatics. This is an important point for English learners of Arabic who in their attempt to pronounce emphatic consonants may stress every syllable in which they are found. The risk of this generalization rule is that they may stress the emphatics even if they are not in syllables with long vowels.

The third area of difference between YSA and English is related to the presence of the back articulation in YSA represented by the pharyngeals and glottals. This difference again cannot be learned by a transfer strategy. The learners need to listen to the pronunciation of native speakers and try to imitate them. It is only by practice, experience and observation that these differences may be learned. It is particularly important to seize any opportunity for training their ears to Arabic utterances.

In addition, our perception experiments show that the listeners are sensitive to the language specific characteristics in that the phoneme boundaries are assigned differently in YSA from those shown by English listeners (Abramson and Lisker, 1965, 1968). The language experience of the listeners seems to determine the perception of the category boundary. The theory of categorical perception states that categorization can be improved by training. If that is the case, then an appropriate training scheme of how to categorize certain contrasts may be developed and be incorporated as part of a syllabus for foreign language teaching and learning.

6.5 Limitations and suggestions for future research:

Although this study gives a detailed account of what is known at present about the emphatic/nonemphatic articulation, it has also raised some questions concerning the nature of the active and passive articulators contributing to their realizations. It is not possible, for example, from the data of this study to decide whether the secondary constriction results from the interaction between the pharyngeal wall, as a passive articulator, and any of these three possible active articulators: (1) the backward movements of the epiglottis (2) the retraction of the back of the tongue (3) the raising of the hyoid bone as a consequence of the jaw movements or the relevant muscle forces or even the raising of the larynx. Another unanswered question is the exact size and location of this secondary constriction (i.e. whether it is at the lower, middle or upper pharynx).

Evidence on these aspects is only possible with simultaneous physiological and acoustic observation of the articulators and the relevant muscles involved in the production of these consonants. One way such questions may be answered is by positioning a fiberscope at the pharynx so that those articulators may be directly observed or by obtaining cinefluourographic motion film pictures that allow the direct observation of the articulators as they move across time. Other recently developed techniques, such as magnetic resonance imaging (MRI) and ultrasonic imaging can be used as well for the same purpose. Electromyographic data will be needed to register the muscle activities as these consonants are produced.

To get a more complete picture, an additional experiment using electropalatography could be helpful to examine the tongue-palate contact during their primary articulations. This could help to resolve the question of whether the two categories share the same primary place of articulation and constriction shapes during the closure and release phases or whether they involve some variations.

It will also be of interest to us to register the timing of glottal opening and closing gestures in relation to the supraglottal constriction while producing voiced and voiceless consonants. To achieve this goal one will need to use electroglottography to measure the relative amount of light transmitted through the glottis. It allows one to obtain information on the area of glottal opening as an indirect measure of vocal folds adjustment. It gives no information on the shape of the glottis but only on the extent of the opening (Borden and Harris, 1984). The shape of the glottis can also be observed directly by a fiberscope.

On the perceptual level, it would also be interesting if a more extensive experiment using synthetic speech are to be conducted with the aim of finding out whether the voiced/voiceless or the emphatic/nonemphatic contrast reflects a genuine case of the phenomenon of categorical perception. In this respect one would not only need to do an identification test but also an ABX discrimination test to judge from the listeners judgements whether they can discriminate more than they can identify and whether the discrimination curve is associated with peaks at the phoneme boundary and "troughs" near the chance level on either side of the phoneme boundary. Similar experiments would also need to be carried out to demonstrate non-categorical perception.

Appendix 1
Some examples of consonant clusters in MSA and YSA

1.Clusters with /-d-/

Consonant Clusters /d-/	Examples	Equivalent translation	Consonant Clusters /-d/	Examples	Equivalent translation
db	tadbiir	planning	bd	mubdiʕ	creative
dt	ʕudtum	you returned	td	-	-
dT	-	-	Td	-	-
dd	ħaddaad	blacksmith	dd	ħaddaad	blacksmith
dD	-	-	Dd	ʔaDdaad	antithesis
dz	-	-	zd	ʔaZdaad	grandfathers
df	madfaʕ	canon	fd	wafd	delegation
dθ	-	-	θd	-	-
dð	-	-	ðd	-	-
dδ	-	-	δd	-	-
dS	-	-	Sd	ʔaSdaaʔ	echoes
ds	madsuus	entered illegally	sd	masduud	closed
dz	-	-	zd	ʔizdihaar	progress
dʃ	-	-	ʃd	maʃduud	tightened
dm	radmaan	a name	md	ɣamdaan	a name
dn	ʕadnaan	a name	nd	sindaan	
dr	madruus	well planned	rd	marduud	rejected
dl	badlah	suit	ld	-	-
dk	madkuuk	flattened	kd	-	-
dq	madquuq	struck	qd	ʔaqdaam	feet
dx	madxal	entrance	xd	ʔaxdaam	servants
dɣ	ʔadɣaal	forests	ɣd	maɣduur	was deceived
dħ	madħuur	was defeated	ħd	maħduud	limited
dʕ	madʕuum	was supported	ʕd	maʕduum	unavailable
dh	madhʊʊn	painted	hd	mahd	cradle
dw	ʕidwaan	attack	wd	mawduuʕ	was saved
dy	ʔadyaan	religions	yd	maydaan	stadium
dʔ			ʔd		

Appendix 1 (Cont.)

Some examples of consonant clusters in MSA and YSA

2.Clusters with /-t-/

Consonant Clusters /t-/	Examples	Equivalent translation	Consonant Clusters /-t/	Examples	Equivalent translation
tb	ʔatbaaʕ	followers	bt	mubtaʕiθ	a student
tt	ħatta	until	tt	ħatta	until
tT	-	-	Tt	-	-
td	-	-	dt	-	-
tD	-	-	Dt	-	-
tdʒ	ʔitdʒaar	trading	dʒt	ʔidʒtinaab	avoidance
tf	ħatfuh	his destiny	ft	miftaaħ	a key
tθ	-	-	θt	-	-
tð	-	-	ðt	-	-
tδ	-	-	δt	-	-
tS	ʔintSaar	victory	St	-	-
ts	natsaabiq	go racing	st	ʔistilaam	reception
tz	natzaawar	mutual visits	zt	-	-
tʃ	-	-	ʃt	muʃtaaq	looking forward
tm	ʔitmaam	completion	mt	ʔimtiħaan	Exam
tn	natnaaSaf	equal shares	nt	ʔintiħaar	suicide
tr	traakum	compilation	rt	murtaaħ	comfortable
tl	jattlaazim	association	lt	multaam	was to blame
tk	natkalam	speaking	kt	maktuub	written
tq	ʔitqaan	accuracy	qt	maqtuul	was killed
tx	mutxaaridz	graduate	xt	maxtuum	signed up
tɤ	burtɤaal	Burtugese	ɤt	muɤtar	arrogant
tħ	fatħ	a name	ħt	maħtuum	unavoidable
tʕ	mutʕalim	learner	ʕt	taʕtiim	blackening
th	muthaawin	irresponsible	ht	buhtaan	lies
tw	mutwalii	to be in charge of	wt	mawt	death
ty	mutyaqiD	awake	yt	bayt	home
tʔ			ʔt	muʊʔtamn	honest

Appendix 1 (Cont.)

Some examples of consonant clusters in MSA and YSA

3.Clusters with /-D-/

Consonant Clusters /D-/	Examples	Equivalent translation	Consonant Clusters /-D/	Examples	Equivalent translation
Db	maDbuuT	exact	bD	-	-
Dt	-	-	tD	-	-
DT	ʔiDTiraab	confusion	TD	-	-
Dd	ʔaDdaad	antithesis	dD	-	-
DD	XuDDaab		DD	XuDDaab	
Dz	maDzaʕ	bedroom	ZD	-	-
Df	maDfuur		fD	ʔafDaʕ	greater
Dθ	-	-	θD	-	-
Dð	-	-	ðD	-	-
Dδ	-	-	δD	-	--
DS	-	-	SD	-	-
Ds	-	-	sD	-	-
Dz	-	-	zD	-	-
Dʃ	-	-	ʃD	-	-
Dm	maDmuun	guaranteed	mD	-	-
Dn	miDmaar	intention	nD	minDaar	telescope
Dr	maDrab		rD	ʔarD	earth
Dl	ʔiDlaal		lD	-	-
Dk	-	-	kD	makDuum	covered up
Dq	-	-	qD	ʔaqDaab	
Dx	ʔaDxam	greater	xD	maxDʊʊb	mixed
Dɤ	maDɤuuT	compressed	ɤD	maɤDuub	hated
Dħ	-		ħD	maħDuur	prohibited
Dʕ	ʔaDʕaaf	multiplication	ʕD	muʕDil	problematic
Dh	maDhar	appearance	hD	mahDuum	digested
Dw	ʔaDwaʔ	lights	wD	mawDuuʕ	subject; topic
Dy			yD	bayDaat	eggs
Dʔ			ʔD		

Appendix 1 (Cont.)
Some examples of consonant clusters in MSA and YSA

4.Clusters with /-T-/

Consonant Clusters /T-/	Examples	Equivalent translation	Consonant Clusters /-T/	Examples	Equivalent translation
Tb	maTbuuʕ	printed	bT	mabTuuħ	lying on the ground
Tt	-	-	tT	-	-
TT	xaTTaaT	writer	TT	xaTTaaT	writer
Td	-	-	dT	-	-
TD	-	-	DT	ʔiDTiraar	obliged to
TZ	-	-	ZT	-	-
Tf	ʔaTfaal	children	fT	mafTuum	was prevented from
Tθ	-	-	θT	-	-
Tð	-	-	ðT	-	-
Tδ	-	-	δT	-	-
TS	-	-	ST	ʔiSTilaaħ	standard, criterion
Ts	-	-	sT	-	-
Tz	-	-	zT	-	-
Tʃ	-	-	ʃT	maʃTuub	omitted
Tm	ʔaTmaaʕ	ambitions	mT	ʔamTaar	rains
Tn	ʔaTnaan	tons	nT	qinTaar	
Tr	muTrib	singer	rT	ʔarTaal	pounds
Tl	raTl	a pound	lT	ʔilTaaf	a name
Tk	-	-	kT	-	-
Tq	ʔaTqum	sets	qT	ʔaqTaab	
Tx	-	-	xT	maxTuuf	hostage
Tɤ	-	-	ɤT	-	-
Tħ	maTħuun	grounded	ħT		
Tʕ	maTʕam	restaurant	ʕT	maʕTuuf	curved
Th	taThiir	purification	hT	-	-
Tw	ʔaTwaar	Subsequent periods	wT	ʔawTaan	One's lands
Ty	ʔaTyaan	lands	yT	-	-
Tʔ			ʔT	-	-

Appendix: 2 Sample material for the word-initial plosives in connected speech

1. *taab* ðaalik ʔalʔinsaan , laaʔinahu laa *dub* ʃadiid ulbaladah, walaa *dim* xafiif ʔalħiilih, walakinah ʔimaa *daab* zaaħif, waʔimaa *duuq* ʃaaja, ðʊ tuuq saabiq. wa *qad* qiil lahʊ ʔilaa *diin* Saaħib ʔalhaq *tʊb* Saraħatan, wa fi mʊsaʕadat ʔal ʔama *tim* fiʕalik. wa laa tadʊb kama *dab* abʊ lahab ul-laði tab, θum *tub* fi ʔarDin laa hia ðuu θamar wala hia ðuu *tiin.*

2. ʔal θalidʒ ʔalmutsaaqiT min ʔasamaaʔ jabduu ðuu *Dil* ʃaahiq, walaakinahuu jatawi alaa *Til* saaqiʕa, jahTul alaa jabal møl ðaalik ʔallaði fi *Tuur* siniin. miθl haða ʔaljaw *Daar* , fahwa yuʔadi ʔilaa *Diiq* ʃaddid fi ʔilaiʃ, wa adam ʔilqudrah ala *Tiiq* ʃuʔaun ʔilħiat, wala Ala *Dur* sakan ʔilʃitaaʔ ʔilbaarid. wamaa *Dar* sʊkaan bilaadin maa ʔadaa ʔilaa *Duur*ahum, kamaa juʔaadi ʔilaa *Tar*ahum. øumaa ʔinahu qad jusabbib hijrathum miøl ʔilTujuur ʔillati ʔauDTirt aw bimaʔnaa *Taar*t ʃitaaʔn ʔilaa ħaiø ʔildifaʔ baadama *Taar* ʃuhuur ʔilaa makaan tatawafir fiih muqawimaat ʔilhiaT ʔilDaruriah.

3. *kaad* θalaaθah junuud ʔan jahlaku, lawla ʔan ʔillaði *gaad* ðaalik ʔiljaiʃ ʔangaðahʊm, ħiinama ʔaʕaTaa *kuub* ʃaaj likuln minhum liaʃrabuh. wabaad fatrah ʃaʕaruu binaqS fi *kad* øalaaθathum, θuma talaaʃaa ʔaxiirun. ama ʔilgaaʔid fagaal lajundi *gum* θaabir, *guud* ʃujaʕaan ʔiljaiʃ . θumaa gaal likul jundi kif ʃar ʔalʔaðaa, laaʔin ʔilʔaðaa *gad* sabbab halaak ʔunaas ʔabriaʔ. Wabaʃarahum jamiiʔan bi ʔilnaSir, mimaa ʔadaa ʔilaa *kiif* ʃaʕaruu baadʊh biθigah bil nafs. wa ʔaxiiran, gaala ʔilgaaʔid la jundi *giis* ʔilmasaafah ħalamaa traa *gis* fii ʔilTariiq minhunaak ʔilaa ʔilmadinah. wa gaala lahu laa tansaa *kub* kul ʔilxyuul ʔind ʔadatik.

4. jaaʔa ʔilqaaʔid ʔillaði *qaad* ʃuja'aan kaθiiruun. waqaala lajundi *qum* sariiaʕun ʔilaa ʔiljaiʃ, *quud* θuwaar ʔilghaDib ʔilaa ʔilnaSir. waqaala muDiifun laa tansaa *qiis* ʔilmasafah min hunaak hataa traa qis ʔilmadinah, wanaSaħahu biaʕadam ʔilʔaðyia lalabria,laʔan ʔilaðaa *qad* sabbab fi halaak ʔaumum sabaqat.

5. *baat* ðaalik ʔilmusaafir fi *biir* ʃadiidat ʔilaʔaumq, walakinahu ʔaufzia ħiinama ʔirtafaʕaa *buuq* Saaxib, ðuu *buq* θaaqib, yamsakahu ħaras *bir* saabiq, limaukib ʔilħaakim ʔillaðibat fii ʔamuur kaθiirah.

Appedix 3 Synthesis specification for the voicing contrast /daa/ versus /taa/.

```
   Max output signal (overload if greater than 0.0 dB) is  -4.2 dB
   Total number of waveform samples = 4050

CURRENT CONFIGURATION:
  47 parameters

    SYM V/C  MIN   VAL   MAX          SYM V/C  MIN   VAL   MAX
    ------------------------          ------------------------
    sr   C  5000 10000 20000          nf   C      2     5     8
    du   C    30   405  5000          ss   C      1     1     2
    ui   C     1     5    20          rs   C      1     1 99999
    f0   V     0  1000  5000          av   V      0    60    80
    F1   V   180   500  1300          b1   v     30    60  1000
    F2   V   550  1500  3000          b2   v     40    90  1000
    F3   V  1200  2500  4800          b3   v     60   150  1000
    F4   V  2400  3250  4990          b4   v    100   200  1000
    F5   v  3000  3700  4990          b5   v    100   200  1500
    f6   v  3000  4990  4990          b6   v    100   500  4000
    fz   v   180   280   800          bz   v     40    90  1000
    fp   v   180   280   500          bp   v     40    90  1000
    ah   V     0     0    80          no   v     10    30    65
    at   v     0     0    80          tl   V      0     0    28
    af   V     0     0    80          sk   v      0     0   100
    a1   V     0     0    80          p1   V     30    80  1000
    a2   V     0     0    80          p2   V     40   200  1000
    a3   V     0     0    80          p3   V     60   350  1000
    a4   V     0     0    80          p4   V    100   500  1000
    a5   V     0     0    80          p5   V    100   600  1500
    a6   v     0     0    80          p6   v    100   800  4000
    an   v     0     0    80          ab   V      0     0    80
    ap   v     0     0    80          os   C      0     0    20
    g0   v     0    60    80

Varied Parameters:
time   f0   av   F1   F2   F3   F4    ah   tl   af   a1   p1   a2   p2   a3   p3   a4   p4   a5   p5   ab
   0 1650    0  400 1800 3100 3250    50    0   60   60   80   50  200   40  350   30  500   30  600   60
   5 1650    0  400 1800 3090 3250    45    0   60   60   80   50  200   40  350   30  500   30  600   60
  10 1650    0  400 1800 3081 3250    40    0   60   60   80   50  200   40  350   30  500   30  600   60
  15 1650    0  400 1800 3072 3250     5    0   60   60   80   50  200   40  350   30  500   30  600   60
  20 1650    0  400 1800 3062 3250     0    0   60   60   80   50  200   40  350   30  500   30  600   60
  25 1650    0  500 1800 3053 3250    30    0   60   60   80   50  200   40  350   30  500   30  600   60
  30 1650    0  500 1800 3044 3250     0    0   60   60   80   50  200   40  350   30  500   30  600   60
  35 1650    0  500 1800 3035 3250     0    0   60   60   80   50  200   40  350   30  500   30  600   60
  40 1650    0  500 1800 3025 3250     0    0   60   60   80   50  200   40  350   30  500   30  600   60
  45 1650    0  500 1800 3016 3250     0    0   60   60   80   50  200   40  350   30  500   30  600   60
  50 1650    0  500 1800 3007 3250     0    0   60   60   80   50  200   40  350   30  500   30  600   60
  55 1650    0  550 1800 2998 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  60 1650    0  600 1800 2988 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  65 1650    0  650 1700 2979 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  70 1650    0  650 1696 2970 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  75 1670    0  650 1693 2961 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  80 1690    0  650 1690 2951 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  85 1690    0  650 1687 2942 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  90 1690    0  650 1684 2933 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
  95 1670    0  650 1681 2924 3250     0   10   60   60   80   50  200   40  350   30  500   30  600   60
 100 1650    0  650 1678 2914 3250     0   10    0   60   80   50  200   40  350   30  500   30  600    0
 105 1650   60  650 1675 2905 3250     0   10    0   60   80   50  200   40  350   30  500   30  600    0
 110 1650   60  650 1672 2896 3250     0   10    0   60   80   50  200   40  350   30  500   30  600    0
```

Appendix 3 (Continued).

115	1650	60	650	1669	2887	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
120	1650	60	650	1666	2877	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
125	1635	60	650	1663	2868	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
130	1620	60	650	1660	2859	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
135	1608	60	650	1657	2850	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
140	1597	60	650	1654	2840	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
145	1585	60	650	1651	2831	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
150	1574	60	650	1648	2822	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
155	1562	60	650	1645	2812	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
160	1551	60	650	1650	2803	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
165	1539	60	650	1650	2794	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
170	1528	60	650	1650	2785	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
175	1516	60	650	1650	2775	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
180	1505	60	650	1650	2766	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
185	1493	60	650	1650	2757	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
190	1482	60	650	1650	2748	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
195	1470	60	650	1650	2738	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
200	1459	60	650	1650	2729	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
205	1447	60	650	1650	2720	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
210	1436	60	650	1650	2711	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
215	1424	60	650	1650	2701	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
220	1413	60	650	1650	2692	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
225	1401	60	650	1650	2683	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
230	1390	60	650	1650	2674	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
235	1378	60	650	1650	2664	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
240	1367	60	650	1650	2655	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
245	1355	60	650	1650	2646	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
250	1344	60	650	1650	2637	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
255	1332	60	650	1650	2627	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
260	1321	60	650	1650	2618	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
265	1310	60	650	1650	2609	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
270	1298	60	650	1650	2600	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
275	1287	60	650	1650	2512	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
280	1275	60	650	1650	2525	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
285	1264	60	650	1650	2538	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
290	1252	60	650	1650	2551	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
295	1241	60	650	1650	2564	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
300	1229	60	650	1650	2576	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
305	1218	49	650	1650	2589	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
310	1206	48	650	1650	2602	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
315	1195	47	650	1650	2615	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
320	1183	46	650	1650	2628	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
325	1172	45	650	1650	2641	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
330	1160	44	650	1650	2653	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
335	1149	43	650	1650	2666	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
340	1137	42	650	1650	2679	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
345	1126	41	650	1650	2692	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
350	1114	40	650	1650	2705	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
355	1103	35	650	1650	2717	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
360	1091	30	650	1650	2730	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
365	1080	25	650	1650	2743	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
370	1068	20	650	1650	2756	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
375	1057	12	650	1650	2769	3250	0	10	0	60	80	50	200	40	350	30	500	30	600	0
380	1045	5	650	1650	2782	3250	0	0	0	60	80	50	200	40	350	30	500	30	600	0
385	1034	2	650	1650	2794	3250	0	0	0	60	80	50	200	40	350	30	500	30	600	0
390	1022	0	650	1650	2807	3250	0	0	0	60	80	50	200	40	350	30	500	30	600	0
395	1011	0	650	1650	2820	3250	0	0	0	60	80	50	200	40	350	30	500	30	600	0
400	1000	0	650	1650	2833	3250	0	0	0	60	80	50	200	40	350	30	500	30	600	0

Appendix 4 Synthesis specification for the voicing contrast /Daa/ versus /Taa/.

```
     Max output signal (overload if greater than 0.0 dB) is  -4.5 dB
     Total number of waveform samples = 4060


  CURRENT CONFIGURATION:
    47 parameters

       SYM V/C  MIN   VAL   MAX          SYM V/C  MIN   VAL   MAX
       ------------------------          ------------ -----------
       sr   C  5000 10000 20000          nf   C     2     5     8
       du   C    30   405  5000          ss   C     1     1     2
       ui   C     1     2    20          rs   C     1     1 99999
       f0   V     0  1000  5000          av   V     0    60    80
       F1   V   180   500  1300          b1   v    30    60  1000
       F2   V   550  1500  3000          b2   v    40    90  1000
       F3   V  1200  2500  4800          b3   v    60   150  1000
       F4   v  2400  3250  4990          b4   v   100   200  1000
       F5   v  3000  3700  4990          b5   v   100   200  1500
       f6   v  3000  4990  4990          b6   v   100   500  4000
       fz   v   180   280   800          bz   v    40    90  1000
       fp   v   180   280   500          bp   v    40    90  1000
       ah   v     0     0    80          no   v    10    30    65
       at   v     0     0    80          tl   V     0     0    28
       af   V     0     0    80          sk   v     0     0   100
       a1   V     0     0    80          p1   V    30    80  1000
       a2   V     0     0    80          p2   V    40   200  1000
       a3   V     0     0    80          p3   V    60   350  1000
       a4   V     0     0    80          p4   V   100   500  1000
       a5   V     0     0    80          p5   V   100   600  1500
       a6   v     0     0    80          p6   v   100   800  4000
       an   v     0     0    80          ab   V     0     0    80
       ap   v     0     0    80          os   C     0     0    20
       g0   v     0    60    80


Varied Parameters:
time   f0   av   F1   F2   F3    tl    af   a1   p1   a2   p2   a3   p3   a4   p4   a5   p5   ab
   0 1000   60  200 1600 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
   2 1010   60  200 1580 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
   4 1020   60  200 1560 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
   6 1030   60  200 1540 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
   8 1040   60  200 1520 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  10 1050   60  200 1500 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  12 1285   60  200 1480 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  14 1500   60  200 1460 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  16 1501   60  200 1440 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  18 1502   60  200 1420 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  20 1503   60  200 1400 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  22 1504   60  200 1380 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  24 1505   60  200 1360 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  26 1506   60  200 1340 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  28 1507   60  200 1320 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  30 1508   60  200 1300 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  32 1509   60  200 1280 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  34 1510   60  200 1260 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  36 1511   60  200 1240 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  38 1512   60  200 1220 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  40 1513   60  200 1200 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  42 1514   60  200 1180 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
  44 1515   60  200 1160 2500    28     0   70   80   55  200   50  350   40  500   30  600    0
```

Appendix 4 (Continued).

46	1516	60	200	1140	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
48	1517	60	200	1120	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
50	1518	60	200	1100	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
52	1520	60	200	1080	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
54	1521	60	200	1060	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
56	1522	60	200	1040	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
58	1523	60	200	1020	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
60	1524	60	200	1000	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
62	1525	60	210	980	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
64	1526	60	220	960	2500	28	0	70	80	55	200	50	350	40	500	30	600	0
66	1527	60	230	940	2500	28	60	70	80	55	200	50	350	40	500	30	600	60
68	1528	60	240	920	2500	28	60	70	80	55	200	50	350	40	500	30	600	60
70	1530	60	250	900	2500	28	60	70	80	55	200	50	350	40	500	30	600	60
72	1554	60	270	900	2500	19	60	70	80	55	200	50	350	40	500	30	600	60
74	1578	60	290	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
76	1602	60	310	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
78	1626	60	330	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
80	1650	60	350	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
82	1670	60	384	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
84	1690	60	418	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
86	1692	60	452	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
88	1695	60	486	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
90	1697	60	520	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
92	1700	60	546	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
94	1700	60	572	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
96	1700	60	598	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
98	1700	60	624	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
100	1700	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
102	1700	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
104	1700	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
106	1700	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
108	1700	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
110	1700	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
112	1694	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
114	1688	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
116	1682	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
118	1676	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
120	1670	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
122	1670	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
124	1670	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
126	1670	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
128	1670	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
130	1670	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
132	1670	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
134	1670	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
136	1670	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
138	1670	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
140	1670	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
142	1665	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
144	1660	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
146	1656	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
148	1651	60	650	900	2500	0	0	70	80	55	200	50	350	40	500	30	600	0
150	1647	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
152	1642	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
154	1637	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
156	1633	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
158	1628	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
160	1624	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
162	1619	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
164	1614	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
166	1610	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
168	1605	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
170	1601	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
172	1596	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
174	1591	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
176	1587	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0

178	1582	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
180	1578	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
182	1573	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
184	1568	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
186	1564	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
188	1559	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
190	1555	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
192	1550	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
194	1545	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
196	1541	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
198	1536	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
200	1532	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
202	1527	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
204	1522	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
206	1518	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
208	1513	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
210	1509	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
212	1504	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
214	1500	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
216	1493	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
218	1486	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
220	1480	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
222	1473	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
224	1467	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
226	1460	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
228	1453	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
230	1447	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
232	1440	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
234	1434	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
236	1427	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
238	1421	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
240	1414	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
242	1407	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
244	1401	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
246	1394	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
248	1388	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
250	1500	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
252	1493	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
254	1487	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
256	1480	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
258	1474	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
260	1467	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
262	1461	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
264	1454	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
266	1448	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
268	1441	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
270	1435	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
272	1428	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
274	1422	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
276	1415	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
278	1409	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
280	1402	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
282	1396	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
284	1389	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
286	1383	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
288	1376	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
290	1370	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
292	1363	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
294	1357	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
296	1350	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
298	1344	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
300	1337	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
302	1331	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
304	1324	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
306	1318	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0
308	1311	60	650	900	2500	10	0	70	80	55	200	50	350	40	500	30	600	0

Appendix 4 (Continued).

Appendix 4 (Continued).

```
310 1305   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
312 1298   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
314 1292   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
316 1285   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
318 1279   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
320 1272   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
322 1266   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
324 1259   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
326 1253   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
328 1246   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
330 1240   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
332 1233   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
334 1227   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
336 1220   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
338 1214   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
340 1207   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
342 1201   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
344 1194   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
346 1188   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
348 1181   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
350 1175   60  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
352 1168   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
354 1162   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
356 1155   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
358 1149   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
360 1142   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
362 1136   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
364 1129   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
366 1123   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
368 1116   50  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
370 1110   40  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
372 1103   37  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
374 1097   35  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
376 1090   30  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
378 1084   25  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
380 1077   20  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
382 1071   17  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
384 1064   14  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
386 1058   11  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
388 1051    8  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
390 1045    5  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
392 1038    2  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
394 1032    0  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
396 1025    0  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
398 1019    0  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
400 1012    0  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
402 1006    0  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
404 1000    0  650  900 2500   10    0   70   80   55  200   50  350   40  500   30  600    0
```

Appendix 5 Response sheet.

Name ________________________________

RESPONSE SHEET File HA.ans Created Tue Mar 17 22:26:47 1992

1. ______
2. ______
3. ______
4. ______
5. ______
6. ______
7. ______
8. ______
9. ______
10. ______

11. ______
12. ______
13. ______
14. ______
15. ______
16. ______
17. ______
18. ______
19. ______
20. ______

21. ______
22. ______
23. ______
24. ______
25. ______
26. ______
27. ______
28. ______
29. ______
30. ______

31. ______
32. ______
33. ______
34. ______
35. ______
36. ______
37. ______
38. ______
39. ______
40. ______

41. ______
42. ______
43. ______
44. ______
45. ______
46. ______
47. ______
48. ______
49. ______
50. ______

51. ______
52. ______
53. ______
54. ______
55. ______
56. ______
57. ______
58. ______
59. ______
60. ______

61. ______
62. ______
63. ______
64. ______
65. ______
66. ______
67. ______
68. ______
69. ______
70. ______

71. ______
72. ______
73. ______
74. ______
75. ______
76. ______
77. ______
78. ______
79. ______
80. ______

81. ______
82. ______
83. ______
84. ______
85. ______
86. ______
87. ______
88. ______
89. ______
90. ______

91. ______
92. ______
93. ______
94. ______
95. ______
96. ______
97. ______
98. ______
99. ______
100. ______

101. ______
102. ______
103. ______
104. ______
105. ______
106. ______
107. ______
108. ______
109. ______
110. ______

111. ______
112. ______
113. ______
114. ______
115. ______
116. ______
117. ______
118. ______
119. ______
120. ______

121. ______
122. ______
123. ______
124. ______
125. ______
126. ______
127. ______
128. ______
129. ______
130. ______

131. ______
132. ______
133. ______
134. ______
135. ______
136. ______
137. ______
138. ______
139. ______
140. ______

141. ______
142. ______
143. ______
144. ______
145. ______
146. ______
147. ______
148. ______
149. ______
150. ______

151. ______
152. ______
153. ______
154. ______
155. ______
156. ______
157. ______
158. ______
159. ______
160. ______

Appendix 5 Answer sheet.

```
Page 1

File HA.ans   Created Tue Mar 17 22:26:47 1992

  1.  11      51.   8     101.  12     151.  11
  2.   3      52.  11     102.  14     152.   3
  3.  13      53.   4     103.   2     153.  13
  4.   5      54.  12     104.   4     154.   5
  5.   7      55.   2     105.  15     155.   7
  6.  12      56.  14     106.  13     156.  12
  7.  14      57.  10     107.   8     157.  14
  8.  15      58.   7     108.   3     158.  15
  9.   9      59.   6     109.  10     159.   9
 10.  10      60.   9     110.  15     160.  10

 11.   2      61.  15     111.   1
 12.   8      62.   9     112.   2
 13.   1      63.   6     113.  11
 14.   4      64.   8     114.   4
 15.   6      65.   2     115.   6
 16.   4      66.  12     116.   5
 17.   8      67.   5     117.   9
 18.  12      68.   1     118.  12
 19.   1      69.  13     119.   7
 20.   7      70.  10     120.  14

 21.   3      71.   4     121.  15
 22.   2      72.  11     122.   2
 23.  15      73.   3     123.   6
 24.  14      74.   7     124.  10
 25.  10      75.  14     125.   4
 26.   6      76.   8     126.  14
 27.  11      77.   4     127.  11
 28.   9      78.  13     128.  12
 29.   5      79.  15     129.   3
 30.  13      80.   1     130.   7

 31.   9      81.  14     131.   9
 32.  10      82.   5     132.   1
 33.   5      83.  12     133.   8
 34.   2      84.   3     134.  13
 35.  14      85.   7     135.   5
 36.  12      86.  11     136.   1
 37.  15      87.   9     137.   5
 38.   3      88.   6     138.  13
 39.   6      89.  10     139.  15
 40.   1      90.   2     140.   2

 41.   8      91.   9     141.   8
 42.  13      92.  10     142.   7
 43.  11      93.   7     143.   9
 44.   4      94.   3     144.  12
 45.   7      95.   5     145.  11
 46.   3      96.  11     146.   6
 47.   1      97.  13     147.   4
 48.  15      98.   8     148.  10
 49.   5      99.   6     149.  14
 50.  13     100.   1     150.   3
```

```
Max output signal (overload if greater than 0.0 dB) is  -13.2 dB
Total number of waveform samples = 3050


CURRENT CONFIGURATION:
  47 parameters

   SYM V/C  MIN   VAL   MAX           SYM V/C  MIN   VAL   MAX
   -------------------------          -------------------------
   sr   C  5000 10000 20000           nf   C     2     4     8
   du   C    30   305  5000           ss   C     1     1     2
   ui   C     1     5    20           rs   C     1     1 99999
   f0   V     0  1000  5000           av   V     0    60    80
   F1   V   180   500  1300           b1   V    30    60  1000
   F2   V     0  1500  3000           b2   V    40    90  1000
   F3   V  1200  2500  4800           b3   V    60   150  1000
   F4   V  2400  3250  4990           b4   V   100   200  1000
   F5   V  3000  3700  4990           b5   V   100   200  1500
   f6   V  3000  4990  4990           b6   V   100   500  4000
   fz   V   180   280   800           bz   V    40    90  1000
   fp   V   180   280   500           bp   V    40    90  1000
   ah   V     0     0    80           no   V    10    30    65
   af   V     0     0    80           tl   V     0     0    24
   ai   V     0     0    80           sk   V     0     0   100
   a1   V     0     0    80           p1   V    30    80  1000
   a2   V     0     0    80           p2   V    40   200  1000
   a3   V     0     0    80           p3   V    60   350  1000
   a4   V     0     0    80           p4   V   100   500  1000
   a5   V     0     0    80           p5   V   100   600  1500
   a6   V     0     0    80           p6   V   100   800  4000
   an   V     0     0    80           ab   V     0     0    80
   ap   V     0     0    80           os   C     0     0    20
   g0   V     0    60    80


Varied Parameters:
time   f0   av   F1   F2   F3   F4    af    a1    p1   a2   p2   a3   p3   a4   p4   a5   p5   p6   ab
   0    0    0  180    0 2800 3900    60     0    80   40  200   50  350   80  500   80  600  800   60
   5    0    0  205    0 2800 3900    57     0    80   40  200   50  350   80  500   80  600  800   60
  10    0    0  230    0 2800 3900    55     0    80   40  200   50  350   80  500   80  600  800   60
  15    0    0  255    0 2800 3900    52     0    80   40  200   50  350   80  500   80  600  800   60
  20    0    0  280    0 2800 3900    50     0    80   40  200   50  350   80  500   80  600  800   60
  25    0    0  305    0 2800 3900    45     0    80   40  200   50  350   80  500   80  600  800   60
  30    0    0  330    0 2800 3900    40     0    80   40  200   50  350   80  500   80  600  800   60
  35    0    0  330    0 2800 3900    35     0    80   40  200   50  350   80  500   80  600  800   60
  40    0    0  330    0 2800 3900    30     0    80   40  200   50  350   80  500   80  600  800   60
  45    0    0  330  700 3000 3600    25     0    80   40  200   50  350   80  500   80  600  800   60
  50 1231    0  330  818 3000 3600    20     0    80   40  200   50  350   80  500   80  600  800    0
  55 1222   60  330  936 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  60 1213   60  330 1054 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  65 1203   60  330 1172 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  70 1194   60  330 1290 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  75 1185   60  330 1409 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  80 1176   60  330 1527 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  85 1167   60  330 1645 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  90 1157   60  330 1763 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
  95 1148   60  330 1881 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
 100 1139   60  330 2000 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
 105 1130   60  330 2200 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800    0
 110 1121   60  330 2200 3000 3600     0     0    80   40  200   50  350   80  500   80  600  800
```

Appendix 6 Synthetic specification for the emphatic-nonemphatic contrast in test one.

Appendix 6 (Continued).

```
115 1111    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
120 1102    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
125 1093    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
130 1084    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
135 1075    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
140 1065    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
145 1056    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
150 1047    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
155 1038    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
160 1028    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
165 1019    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
170 1010    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
175 1001    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
180  992    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
185  982    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
190  973    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
195  964    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
200  955    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
205  946    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
210  936    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
215  927    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
220  918    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
225  909    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
230  900    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
235  880    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
240  861    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
245  842    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
250  823    60  330 2200 3000 3600    0    0   80   40  200   50  350   80  500   80  600  800    0
255  803    53  330 2200 3000 3600    0    0   80   40  200   50  350   71  500   71  600  800    0
260  784    46  330 2200 3000 3600    0    0   80   40  200   50  350   62  500   62  600  800    0
265  765    40  330 2200 3000 3600    0    0   80   40  200   50  350   53  500   53  600  800    0
270  746    33  330 2200 3000 3600    0    0   80   40  200   50  350   44  500   44  600  800    0
275  726    26  330 2200 3000 3600    0    0   80   40  200   50  350   35  500   35  600  800    0
280  707    20  330 2200 3000 3600    0    0   80   40  200   50  350   26  500   26  600  800    0
285  688    13  330 2200 3000 3600    0    0   80   40  200   50  350   17  500   17  600  800    0
290  669     6  330 2200 3000 3600    0    0   80   40  200   50  350    8  500    8  600  800    0
295  650     0  330 2200 3000 3600    0    0   80   40  200   50  350    0  500    0  600  800    0
300    0     0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
```

Appendix 7 Synthetic specification for the emphatic-nonemphatic contrast in test two.

```
    Max output signal (overload if greater than 0.0 dB) is  -2.1 dB
    Total number of waveform samples = 4050

CURRENT CONFIGURATION:
  47 parameters

   SYM V/C  MIN    VAL   MAX           SYM V/C  MIN    VAL   MAX
   -------------------------           -------------------------
   sr   C  5000 10000 20000            nf   C     2      3     8
   du   C    30   405  5000            ss   C     1      1     2
   ui   C     1     5    20            rs   C     1      1 99999
   f0   V     0  1000  5000            av   V     0     60    80
   F1   V   180   500  1300            b1   v    30     60  1000
   F2   V     0  1500  3000            b2   v    40     90  1000
   F3   V  1200  2500  4800            b3   v    60    150  1000
   F4   v  2400  3250  4990            b4   v   100    200  1000
   F5   v  3000  3700  4990            b5   v   100    200  1500
   f6   v  3000  4990  4990            b6   v   100    500  4000
   fz   v   180   280   800            bz   v    40     90  1000
   fp   v   180   280   500            bp   v    40     90  1000
   ah   v     0     0    80            no   v    10     30    65
   at   v     0     0    80            tl   v     0      0    24
   af   V     0     0    80            sk   v     0      0   100
   a1   V     0     0    80            p1   V    30     80  1000
   a2   V     0     0    80            p2   V    40    200  1000
   a3   V     0     0    80            p3   V    60    350  1000
   a4   V     0     0    80            p4   V   100    500  1000
   a5   V     0     0    80            p5   V   100    600  1500
   a6   V     0     0    80            p6   V   100    800  4000
   an   v     0     0    80            ab   V     0      0    80
   ap   v     0     0    80            os   C     0      0    20
   g0   v     0    60    80

Varied Parameters:
time   f0   av   F1   F2   F3   af   a1   p1   a2   p2   a3   p3   a4   p4   a5   p5   a6   p6   ab
   0 1500    0  490 2000 2700   60   65   80   55  200   54  350   55  500   50  600   60  800   60
   5 1493    0  500 1870 2697   60   65   80   55  200   54  350   55  500   50  600   60  800   60
  10 1487    0  510 1740 2695   60   65   80   55  200   54  350   55  500   50  600   60  800   60
  15 1481    0  520 1610 2692   60   65   80   55  200   54  350   55  500   50  600   60  800   60
  20 1475    0  530 1480 2690    0   65   80   55  200   54  350   55  500   50  600   60  800    0
  25 1468    0  540 1350 2687    0   65   80   55  200   54  350   55  500   50  600   60  800    3
  30 1462    0  550 1220 2685    0   65   80   55  200   54  350   55  500   50  600   60  800    7
  35 1456   60  600 1090 2682    0   65   80   55  200   54  350   55  500   50  600   60  800   11
  40 1450   60  600  960 2680    0   65   80   55  200   54  350   55  500   50  600   60  800   15
  45 1443   60  600  830 2677    0   65   80   55  200   54  350   55  500   50  600   60  800   18
  50 1437   60  600  700 2675    0   65   80   55  200   54  350   55  500   50  600   60  800   22
  55 1431   60  600  700 2672    0   65   80   55  200   54  350   55  500   50  600   60  800   26
  60 1425   60  600  700 2670    0   65   80   55  200   54  350   55  500   50  600   60  800   30
  65 1418   60  600  700 2667    0   65   80   55  200   54  350   55  500   50  600   60  800   33
  70 1412   60  600  700 2665    0   65   80   55  200   54  350   55  500   50  600   60  800   37
  75 1406   60  600  700 2662    0   65   80   55  200   54  350   55  500   50  600   60  800   41
  80 1400   60  600  700 2400    0   65   80   55  200   54  350   55  500   50  600   60  800   45
  85 1393   60  600  700 2433    0   65   80   55  200   54  350   55  500   50  600   60  800   48
  90 1387   60  600  700 2466    0   65   80   55  200   54  350   55  500   50  600   60  800   52
  95 1381   60  600  700 2500    0   65   80   55  200   54  350   55  500   50  600   60  800   56
 100 1375   60  600  700 2533    0   65   80   55  200   54  350   55  500   50  600   60  800   60
 105 1368   60  600  700 2566    0   65   80   55  200   54  350   55  500   50  600   60  800    0
 110 1362   60  600  700 2600    0   65   80   55  200   54  350   55  500   50  600   60  800    0
```

Appendix 7 (Continued).

115	1356	60	600	700	2600	0	65	80	55	200	54	350	55	500	50	600	60	800	0
120	1350	60	600	700	2600	0	65	80	55	200	54	350	55	500	50	600	60	800	0
125	1343	60	600	700	2600	0	65	80	55	200	54	350	55	500	50	600	60	800	0
130	1337	60	600	700	2600	0	65	80	55	200	54	350	55	500	50	600	60	800	0
135	1331	60	600	700	2550	0	65	80	55	200	54	350	55	500	50	600	60	800	0
140	1325	60	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
145	1318	64	600	700	2450	0	65	80	55	200	54	350	55	500	50	600	60	800	0
150	1312	63	600	700	2400	0	65	80	55	200	54	350	55	500	50	600	60	800	0
155	1306	63	600	700	2412	0	65	80	55	200	54	350	55	500	50	600	60	800	0
160	1300	63	600	700	2425	0	65	80	55	200	54	350	55	500	50	600	60	800	0
165	1293	63	600	700	2437	0	65	80	55	200	54	350	55	500	50	600	60	800	0
170	1287	62	600	700	2450	0	65	80	55	200	54	350	55	500	50	600	60	800	0
175	1281	62	600	700	2462	0	65	80	55	200	54	350	55	500	50	600	60	800	0
180	1275	62	600	700	2475	0	65	80	55	200	54	350	55	500	50	600	60	800	0
185	1268	62	600	700	2487	0	65	80	55	200	54	350	55	500	50	600	60	800	0
190	1262	62	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
195	1256	61	600	700	2300	0	65	80	55	200	54	350	55	500	50	600	60	800	0
200	1250	61	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
205	1243	61	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
210	1237	61	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
215	1231	60	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
220	1225	60	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
225	1218	60	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
230	1212	60	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
235	1206	60	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
240	1200	59	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
245	1193	59	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
250	1187	59	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
255	1181	59	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
260	1175	58	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
265	1168	58	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
270	1162	58	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
275	1156	58	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
280	1150	58	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
285	1143	57	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
290	1137	57	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
295	1131	57	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
300	1125	57	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
305	1118	56	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
310	1112	56	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
315	1106	56	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
320	1100	56	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
325	1093	56	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
330	1087	55	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
335	1081	55	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
340	1075	55	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
345	1068	55	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
350	1062	55	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
355	1056	49	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
360	1050	44	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
365	1043	38	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
370	1037	33	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
375	1031	27	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
380	1025	22	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
385	1018	16	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
390	1012	11	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
395	1006	5	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0
400	1000	0	600	700	2500	0	65	80	55	200	54	350	55	500	50	600	60	800	0

Appendix (8): Definition of terms and symbols used in the Analysis of Variance for the emphasis data in section 5.2

Symbol	Stands For	Levels	Variable Type
PS	Position	1= initial 2= medial	Independent
V	Vocalic Context	1= Vowel /aa/ 2= Vowel /ii/	Independent
E	Emphasis	1= Emphatic 2= Nonemphatic	Independent
S	Subject	1= Subject (AA) 2= Subject (WS)	Independent
Uo	Peak Oral Airflow Measurements		Dependent
Uc	Articulatory Closure Duration Measurements		Dependent
Ur	Articulatory release duration Measurements		Dependent
Po	Peak oral air pressure		Dependent
Pd	Pressure duration		Dependent

App. (9) (ANOVA) DATA FOR THE STUDY OF EMPHASIS

Cases		PS	V	E	S	UO	UC	UR	PO	PD
1	>	1	1	1	1	270	80	20	6.50	100
2	>	1	1	1	1	310	65	20	6.40	100
3	>	1	1	1	1	200	80	20	7.60	100
4	>	1	1	1	1	240	85	15	7.30	105
5	>	1	1	1	1	250	80	15	6.80	100
6	>	1	1	1	1	310	70	20	7.80	105
7	>	1	1	1	2	210	90	20	10.00	105
8	>	1	1	1	2	150	85	20	8.80	100
9	>	1	1	1	2	170	85	15	10.20	100
10	>	1	1	1	2	190	90	20	10.30	105
11	>	1	1	1	2	25	85	15	10.00	95
12	>	1	1	1	2	110	85	20	12.70	105
13	>	1	2	1	1	240	75	20	8.00	100
14	>	1	2	1	1	290	60	25	6.70	85
15	>	1	2	1	1	320	75	20	6.50	100
16	>	1	2	1	1	250	65	20	6.00	85
17	>	1	2	1	1	240	65	20	7.60	90
18	>	1	2	1	1	280	75	20	8.00	95
19	>	1	2	1	2	380	90	30	14.30	110
20	>	1	2	1	2	260	80	35	9.00	100
21	>	1	2	1	2	200	95	20	10.40	115
22	>	1	2	1	2	345	100	35	10.50	120
23	>	1	2	1	2	210	80	30	10.30	110
24	>	1	2	1	2	220	75	30	10.70	95
25	>	1	2	1	1	-110	50	15	7.00	95
26	>	1	2	1	1	150	85	20	7.40	85
27	>	1	2	1	1	50	40	10	7.00	85
28	>	1	2	1	1	25	40	15	7.60	80
29	>	1	2	1	1	150	45	15	7.20	85
30	>	1	2	1	1	180	45	15	7.30	85
31	>	1	2	1	2	400	50	25	9.00	105
32	>	1	2	1	2	340	85	25	8.00	110
33	>	1	2	1	2	390	85	25	9.70	110
34	>	1	2	1	2	360	80	25	9.60	105
35	>	1	2	1	2	200	65	20	9.40	120
36	>	1	2	1	2	400	75	25	10.80	115
37	>	1	1	1	1	200	20	25	7.60	80
38	>	1	1	1	1	250	55	20	6.40	95
39	>	1	1	1	1	290	25	20	7.60	100
40	>	1	1	1	1	190	15	15	7.00	90
41	>	1	1	1	1	200	20	10	7.00	100
42	>	1	1	1	1	20	20	15	8.20	90
43	>	1	1	1	2	380	85	25	11.00	110
44	>	1	1	1	2	200	100	30	8.80	125
45	>	1	1	1	2	400	65	20	10.90	125
46	>	1	1	1	2	180	90	30	10.50	105
47	>	1	1	1	2	340	90	20	11.00	115
48	>	1	1	1	2	260	85	30	10.80	135
49	>	1	1	2	1	110	70	30	10.40	70
50	>	1	1	2	1	1050	40	30	9.00	85
51	>	1	1	2	1	110	60	30	11.80	70
52	>	1	1	2	1	900	45	30	15.70	80
53	>	1	1	2	1	1030	60	25	15.40	80
54	>	1	1	2	1	940	40	30	9.00	80
55	>	1	1	2	2	1090	85	50	6.20	55
56	>	1	1	2	2	1150	80	75	9.40	70
57	>	1	1	2	2	1050	85	50	11.00	80
58	>	1	1	2	2	1560	75	65	9.70	80
59	>	1	1	2	2	990	85	55	9.00	80

```
 60  >  1  1  2  2  1200   75   60  10.00   70
 61  >  1  2  2  1   780   55   35   8.00   70
 62  >  1  2  2  1   980   50   50  12.00   80
 63  >  1  2  2  1   950   45   50  11.00   75
 64  >  1  2  2  1   650   55   40   9.00   75
 65  >  1  2  2  1   850   55   40   7.00   70
 66  >  1  2  2  1   990   55   45  15.00   75
 67  >  1  2  2  2   620   85   60   9.00   85
 68  >  1  2  2  2   800   75   85  10.60   85
 69  >  1  2  2  2   890   80   75  10.30   80
 70  >  1  2  2  2   800   60   65  10.20   80
 71  >  1  2  2  2   640   80   75  10.40   85
 72  >  1  2  2  2   790   80   65  10.60   95
 73  >  1  1  2  1   950   35   35   5.30   60
 74  >  1  1  2  1   890   25   30   6.70   65
 75  >  1  1  2  1   770   35   30   6.80   80
 76  >  1  1  2  1   950   35   40   7.40   70
 77  >  1  1  2  1   590   35   30   7.50   70
 78  >  1  1  2  1   650   35   30   7.00   70
 79  >  1  1  2  2  1220   30   55   9.50   85
 80  >  1  1  2  2  1300   55   80  10.20   80
 81  >  1  1  2  2  1500   60   75   6.90   80
 82  >  1  1  2  2  1140   55   55  10.00   90
 83  >  1  1  2  2  1400   60   80   9.00   80
 84  >  1  1  2  2  1060   55   60  10.70  100
 85  >  1  2  2  1   600   20   40   6.80   80
 86  >  1  2  2  1   700   30   45   7.00   80
 87  >  1  2  2  1   540   25   40   7.40   90
 88  >  1  2  2  1   700   30   40   7.00   80
 89  >  1  2  2  1   890   35   40   8.00   80
 90  >  1  2  2  1   630   40   35   7.40   80
 91  >  1  2  2  2   800   60   60  11.90  100
 92  >  1  2  2  2   600   55   60  10.60  105
 93  >  1  2  2  2   600   50   60  10.00  100
 94  >  1  2  2  2   350   55   15  10.00  110
 95  >  1  2  2  2   530   50   60  10.40  110
 96  >  1  2  2  2   850   60   65   9.70  100
 97  >  2  1  1  1   -25   60   20
 98  >  2  1  1  1   200   60   15
 99  >  2  1  1  1   100   60   10
100  >  2  1  1  1    25   65   15
101  >  2  1  1  1   240   65   20
102  >  2  1  1  1   180   65   20
103  >  2  1  1  2   200   25   25
104  >  2  1  1  2   170   60   20
105  >  2  1  1  2   305   60   20
106  >  2  1  1  2   200   50   25
107  >  2  1  1  2   250   40   40
108  >  2  1  1  2   250   55   30
109  >  2  2  1  1   180   50   20
110  >  2  2  1  1   200   65   20
111  >  2  2  1  1   300   60   25
112  >  2  2  1  1   260   55   20
113  >  2  2  1  1   270   55   30
114  >  2  2  1  1   290   55   25
115  >  2  2  1  2   190   50   20
116  >  2  2  1  2   300   55   25
117  >  2  2  1  2   280   60   25
118  >  2  2  1  2   290   50   25
119  >  2  2  1  2   300   60   30
120  >  2  2  1  2   260   60   35
121  >  2  1  1  1   230   60   20
122  >  2  1  1  1    25   50   20
123  >  2  1  1  1    40   40   15
```

```
124  >  2  1  1  1    50   55   10
125  >  2  1  1  1    80   55   15
126  >  2  1  1  1    25   50   15
127  >  2  1  1  2   150   75   20
128  >  2  1  1  2   170   70   20
129  >  2  1  1  2   210   80   25
130  >  2  1  1  2    90   80   20
131  >  2  1  1  2   190   80   20
132  >  2  1  1  2   210   90   20
133  >  2  2  1  1  -230   50   30
134  >  2  2  1  1   170   40   20
135  >  2  2  1  1   250   55   15
136  >  2  2  1  1   150   45   45
137  >  2  2  1  1    50   46   15
138  >  2  2  1  1     0   40   20
139  >  2  2  1  2   450   70   25
140  >  2  2  1  2   400   60   20
141  >  2  2  1  2   270   45   30
142  >  2  2  1  2   310   60   25
143  >  2  2  1  2   360   85   20
144  >  2  2  1  2   450   65   20
145  >  2  2  2  1   580   30   20
146  >  2  2  2  1   750   40   30
147  >  2  2  2  1   700   50   25
148  >  2  2  2  1   730   45   25
149  >  2  2  2  1   890   40   30
150  >  2  2  2  1   800   40   25
151  >  2  2  2  2   610   40   35
152  >  2  2  2  2   650   45   45
153  >  2  2  2  2   650   20   60
154  >  2  2  2  2   680   55   40
155  >  2  2  2  2   850   45   35
156  >  2  2  2  2   500   40   40
157  >  2  2  2  1   900   50   35
158  >  2  2  2  1   830   55   30
159  >  2  2  2  1   700   50   30
160  >  2  2  2  1  1065   40   55
161  >  2  2  2  1   900   45   40
162  >  2  2  2  1   700   40   40
163  >  2  2  2  2   750   50   55
164  >  2  2  2  2   580   50   55
165  >  2  2  2  2   670   55   60
166  >  2  2  2  2   630   65   50
167  >  2  2  2  2   600   60   60
168  >  2  2  2  2   470   60   50
169  >  2  1  2  1   850   40   35
170  >  2  1  2  1   540   40   30
171  >  2  1  2  1  1130   45   35
172  >  2  1  2  1   780   25   35
173  >  2  1  2  1  1390   30   45
174  >  2  1  2  1  1100   25   40
175  >  2  1  2  2  1220   55   55
176  >  2  1  2  2  1130   40   60
177  >  2  1  2  2   800   55   45
178  >  2  1  2  2  1020   60   50
179  >  2  1  2  2   820   40   50
180  >  2  1  2  2   800   50   45
181  >  2  2  2  1   600   45   45
182  >  2  2  2  1   690   45   45
183  >  2  2  2  1   900   50   40
184  >  2  2  2  1   700   50   40
185  >  2  2  2  1   680   40   40
186  >  2  2  2  1   800   35   45
187  >  2  2  2  2   760   70   50
```

188	>	2	2	2	2	900	60	60
189	>	2	2	2	2	850	50	55
190	>	2	2	2	2	750	65	65
191	>	2	2	2	2	780	60	20
192	>	2	2	2	2	700	65	65

Appendix (10): Definition of terms and symbols used in the Analysis of Variance for the voicing data in section 5.3

Symbol	Stands For	Levels	Variable Typ
PS	Place of articulation	1 = Denti-alveolar emphatic 2 = Denti-alveolar non-emphatic 3 = Velars	Independent
V	Vocalic Context	1= Vowel /aa/ 2= Vowel /ii/ 3 = Vowel /uu/	Independent
V1	Voicing	1 = Voiced 2 = Voiceless	Independent
S	Subject	1= Subject (AA) 2= Subject (WS)	Independent
Uo	Peak Oral Airflow Measurements		Dependent
Uc	Articulatory Closure Duration Measurements		Dependent
Ur	Articulatory release duration Measurements		Dependent
Po	Peak oral air pressure		Dependent
Pd	Pressure duration		Dependent

Appendix 11:(ANOVA) DATA FOR THE STUDY OF VOICING

Cases		P	V1	V	S	UO	UC	UR	PO	PD
1	>	1	1	1	1	220	40	25	9.00	65
2	>	1	1	1	1	220	37	40	2.30	65
3	>	1	1	1	1	200	37	25	3.00	60
4	>	1	1	1	1	170	30	20	3.00	70
5	>	1	1	1	1	200	30	30	3.00	60
6	>	1	1	1	1	200	50	25	2.00	80
7	>	1	1	1	2	200	41	25	4.30	80
8	>	1	1	1	2	200	45	40	3.80	80
9	>	1	1	1	2	300	40	45	5.60	80
10	>	1	1	1	2	251	41	33	3.70	65
11	>	1	1	1	2	290	40	25	4.30	65
12	>	1	1	1	2	265	40	30	5.40	80
13	>	1	1	2	1	290	35	30	4.00	80
14	>	1	1	2	1	290	30	29	2.90	80
15	>	1	1	2	1	240	35	20	4.00	80
16	>	1	1	2	1	300	30	35	2.80	80
17	>	1	1	2	1	320	40	30	6.00	80
18	>	1	1	2	1	290	40	30	4.00	75
19	>	1	1	2	2	290	49	25	4.30	80
20	>	1	1	2	2	350	60	30	3.70	80
21	>	1	1	2	2	300	55	30	4.00	80
22	>	1	1	2	2	310	45	30	4.60	70
23	>	1	1	2	2	270	60	30	4.40	90
24	>	1	1	2	2	250	25	50	3.40	60
25	>	1	1	3	1	230	30	20	3.70	65
26	>	1	1	3	1	250	35	25	3.00	65
27	>	1	1	3	1	240	25	25	2.30	60
28	>	1	1	3	1	190	35	25	3.90	60
29	>	1	1	3	1	320	40	25	2.30	60
30	>	1	1	3	1	230	40	20	4.40	65
31	>	1	1	3	2	170	50	20	5.70	75
32	>	1	1	3	2	150	50	30	3.80	80
33	>	1	1	3	2	240	55	25	6.90	80
34	>	1	1	3	2	190	50	35	5.70	85
35	>	1	1	3	2	200	60	25	3.50	95
36	>	1	1	3	2	300	35	30	5.50	80
37	>	1	2	1	1	370	85	15	7.40	105
38	>	1	2	1	1	250	65	15	6.00	90
39	>	1	2	1	1	250	65	20	7.50	90
40	>	1	2	1	1	290	75	20	7.00	95
41	>	1	2	1	1	230	80	15	8.00	100
42	>	1	2	1	1	280	75	20	8.00	95
43	>	1	2	1	2	190	85	20	11.20	100
44	>	1	2	1	2	170	85	20	9.30	105
45	>	1	2	1	2	150	75	15	10.00	105
46	>	1	2	1	2	250	85	20	8.80	100
47	>	1	2	1	2	150	85	15	10.30	100
48	>	1	2	1	2	180	80	20	9.10	95
49	>	1	2	2	1	240	75	20	8.00	100
50	>	1	2	2	1	290	60	25	6.70	85
51	>	1	2	2	1	320	75	20	6.50	100
52	>	1	2	2	1	250	65	20	6.00	85
53	>	1	2	2	1	240	65	20	7.60	90
54	>	1	2	2	1	280	75	20	8.00	95
55	>	1	2	2	2	380	90	30	14.30	110
56	>	1	2	2	2	260	80	35	9.00	100
57	>	1	2	2	2	200	95	20	10.40	115
58	>	1	2	2	2	345	100	35	10.50	120
59	>	1	2	2	2	210	80	30	10.30	110

```
 60  >  1   2   2   2    220    75    30   10.70    95
 61  >  1   2   3   1    230    65    20    7.40    90
 62  >  1   2   3   1    230    75    20    7.90   100
 63  >  1   2   3   1    300    55    15    7.70    80
 64  >  1   2   3   1    200    70    20    7.00    95
 65  >  1   2   3   1    250    60    20    8.00    80
 66  >  1   2   3   1    240    75    15    7.30   100
 67  >  1   2   3   2    200    90    20   11.00   120
 68  >  1   2   3   2    275   100    20    9.20   110
 69  >  1   2   3   2    210    80    20   10.00   105
 70  >  1   2   3   2    200    80    15   11.00   100
 71  >  1   2   3   2    190    90    15   10.30   105
 72  >  1   2   3   2    250    80    25   11.60   100
 73  >  2   1   1   1    490    55    20    2.20    75
 74  >  2   1   1   1    460    45    25    2.40    75
 75  >  2   1   1   1    420    45    25    2.50    75
 76  >  2   1   1   1    460    60    20    5.00    85
 77  >  2   1   1   1    490    55    20    4.00    80
 78  >  2   1   1   1    380    55    15    3.90    75
 79  >  2   1   1   2    575    65    30    6.30    90
 80  >  2   1   1   2    510    65    20    4.50    80
 81  >  2   1   1   2    500    55    20    5.70    75
 82  >  2   1   1   2    500    60    25    7.20    80
 83  >  2   1   1   2    590    40    30    6.80    75
 84  >  2   1   1   2    520    70    25    6.30    95
 85  >  2   1   2   1    500    65    25    4.00    85
 86  >  2   1   2   1    400    60    25    3.20    80
 87  >  2   1   2   1    470    60    25    3.70    80
 88  >  2   1   2   1    300    45    20    3.30    65
 89  >  2   1   2   1    380    55    20    2.80    75
 90  >  2   1   2   1    390    55    25    3.50    80
 91  >  2   1   2   2    480    75    30    9.60   100
 92  >  2   1   2   2    580    65    30    8.00   100
 93  >  2   1   2   2    390    70    35    8.40   100
 94  >  2   1   2   2    440    65    35    8.00    95
 95  >  2   1   2   2    420    80    25    7.60   105
 96  >  2   1   2   2    440    60    25    7.50    95
 97  >  2   1   3   1    540    40    25    3.30    65
 98  >  2   1   3   1    400    60    20    3.30    85
 99  >  2   1   3   1    470    50    25    4.00    70
100  >  2   1   3   1    440    55    25    4.00    80
101  >  2   1   3   1    490    50    20    4.00    70
102  >  2   1   3   1    300    55    25    4.00    75
103  >  2   1   3   2    600    65    30    9.00    90
104  >  2   1   3   2    800    65    25    8.50    85
105  >  2   1   3   2    600    70    25    7.20   100
106  >  2   1   3   2    600    60    35    8.50    85
107  >  2   1   3   2    590    60    30    7.80    85
108  >  2   1   3   2    600    70    25    6.90    85
109  >  2   2   1   1   1100    70    30    7.00    95
110  >  2   2   1   1   1050    40    30    7.40    85
111  >  2   2   1   1   1100    60    30    7.00    85
112  >  2   2   1   1    900    45    30    7.60    80
113  >  2   2   1   1   1030    60    25    7.20    85
114  >  2   2   1   1    940    40    30    7.30    85
115  >  2   2   1   2   1090    85    50    9.00   105
116  >  2   2   1   2   1150    80    75    8.00   110
117  >  2   2   1   2   1050    85    50    9.70   110
118  >  2   2   1   2   1560    75    65    9.60   105
119  >  2   2   1   2    990    85    55    9.40   120
120  >  2   2   1   2   1200    75    60   10.80   115
121  >  2   2   2   1    780    55    35    7.60    80
122  >  2   2   2   1    980    50    50    6.40    95
123  >  2   2   2   1    950    45    50    7.60   100
```

124	>	2	2	2	1	650	55	40	7.00	90
125	>	2	2	2	1	850	55	40	7.00	100
126	>	2	2	2	1	990	55	45	8.20	90
127	>	2	2	2	2	620	85	60	11.00	110
128	>	2	2	2	2	800	75	85	8.80	125
129	>	2	2	2	2	890	80	75	10.90	125
130	>	2	2	2	2	800	60	65	10.50	105
131	>	2	2	2	2	640	80	75	11.00	115
132	>	2	2	2	2	790	80	65	10.80	135
133	>	2	2	3	1	1080	55	35	7.00	95
134	>	2	2	3	1	1100	40	35	7.00	80
135	>	2	2	3	1	1100	45	40	8.00	95
136	>	2	2	3	1	470	45	35	7.80	85
137	>	2	2	3	1	1080	55	35	9.00	95
138	>	2	2	3	1	730	45	40	7.80	95
139	>	2	2	3	2	1030	70	55	10.00	100
140	>	2	2	3	2	1150	100	65	9.90	120
141	>	2	2	3	2	1110	75	60	10.30	110
142	>	2	2	3	2	1290	85	75	10.70	115
143	>	2	2	3	2	980	60	55	9.10	105
144	>	2	2	3	2	1200	75	65	10.40	115
145	>	3	1	1	1	430	30	25		
146	>	3	1	1	1	390	30	25		
147	>	3	1	1	1	270	40	25		
148	>	3	1	1	1	380	60	50		
149	>	3	1	1	1	350	45	20		
150	>	3	1	1	1	350	30	45		
151	>	3	1	1	2	290	45	25		
152	>	3	1	1	2	450	45	25		
153	>	3	1	1	2	480	60	25		
154	>	3	1	1	2	400	55	30		
155	>	3	1	1	2	440	45	25		
156	>	3	1	1	2	400	45	35		
157	>	3	1	2	1	240	40	20		
158	>	3	1	2	1	300	40	25		
159	>	3	1	2	1	240	40	30		
160	>	3	1	2	1	230	55	30		
161	>	3	1	2	1	230	40	30		
162	>	3	1	2	1	250	40	27		
163	>	3	1	2	2	100	60	40		
164	>	3	1	2	2	300	50	60		
165	>	3	1	2	2	200	55	20		
166	>	3	1	2	2	200	60	50		
167	>	3	1	2	2	350	55	30		
168	>	3	1	2	2	250	60	30		
169	>	3	1	3	1	240	40	25		
170	>	3	1	3	1	225	35	28		
171	>	3	1	3	1	250	25	20		
172	>	3	1	3	1	210	40	50		
173	>	3	1	3	1	250	40	15		
174	>	3	1	3	1	160	55	30		
175	>	3	1	3	2	220	75	25		
176	>	3	1	3	2	380	65	50		
177	>	3	1	3	2	220	60	20		
178	>	3	1	3	2	260	60	20		
179	>	3	1	3	2	270	65	30		
180	>	3	1	3	2	290	50	35		
181	>	3	2	1	1	950	35	35		
182	>	3	2	1	1	890	25	30		
183	>	3	2	1	1	770	35	30		
184	>	3	2	1	1	950	30	40		
185	>	3	2	1	1	590	35	30		
186	>	3	2	1	1	650	30	30		
187	>	3	2	1	2	1220	55	55		

188	>	3	2	1	2	1300	60	80
189	>	3	2	1	2	1500	60	75
190	>	3	2	1	2	1140	55	55
191	>	3	2	1	2	1400	55	80
192	>	3	2	1	2	1060	55	60
193	>	3	2	2	1	600	20	40
194	>	3	2	2	1	700	30	45
195	>	3	2	2	1	540	25	40
196	>	3	2	2	1	700	35	40
197	>	3	2	2	1	890	30	40
198	>	3	2	2	1	630	40	35
199	>	3	2	2	2	800	60	60
200	>	3	2	2	2	600	55	60
201	>	3	2	2	2	600	50	60
202	>	3	2	2	2	350	55	15
203	>	3	2	2	2	530	50	60
204	>	3	2	2	2	850	60	65
205	>	3	2	3	1	550	35	40
206	>	3	2	3	1	550	40	50
207	>	3	2	3	1	530	25	35
208	>	3	2	3	1	530	45	40
209	>	3	2	3	1	400	40	35
210	>	3	2	3	1	400	45	30
211	>	3	2	3	2	880	60	70
212	>	3	2	3	2	800	60	65
213	>	3	2	3	2	700	60	45
214	>	3	2	3	2	650	65	65
215	>	3	2	3	2	670	60	55
216	>	3	2	3	2	400	70	40

References

Abercrombie, D. (1967) *Elements of General Phonetics*. Edinburgh University Press. Edinburgh.

Abramson, A. and Lisker, L. (1965) Voice Onset Time in stop consonants: Acoustic Analysis and Synthesis. Proc. 5th Inter. Cong. of Acoustics. Liegè Belgium.

Abramson, A. and Lisker, L. (1968) Voice timing: Cross-language experiments in identification and discrimintion. Haskins Laboratories: Status Report on Speech Research, SR-13/14, 49-63.

Abramson, A and Lisker, L (1972) Voice timing perception in Spanish word-initial stops. *J. of Phon.*, 1: 1-8.

Al-ani, S. H (1970) *Arabic Phonology:* An acoustical and physiological investigation. Janua 61: The Hague: Mouton.

Al-ani, S.H. (1976) The developments and distribution of the Arabic sound 'Qaaf' in Iraq. In *Readings in Arabic Linguistics* by S. Al-ani (ed.). Indiana University Club, Bloomington: 103-112.

Al-ani, S. H. (ed.) (1979) *Readings in Arabic Linguistics*. Indiana University Club, Bloomington.

Al-ani, S.H. and May, D.R. (1979) The phonological structure of the syllable in Arabic. In *Readings in Arabic Linguistics*, S. Al-ani (ed.). Indiana University Club, Bloomington: 113-126.

Al-AnTaakii, M. (n. d.) *Al-muHiiT* "The comprehensive book". Vol, 1. Damascus.

Ali, L. H. and Daniloff, R. G. (1972) A contrastive cinefluorographic investigation of the articulation of emphatic cognate consonants. *Studia Linguistica*, 26: 81-105.

Ali, L. H. and Daniloff, R. G. (1974) The perception of coarticulated emphaticness. *Phonetica*, 29: 225-231.

Al-Makhzuumii, M. (1960) Al-Khaliil Ibn Ahmad al-Farahiidii. ?aζmaaluh wa manhajuh. *'A biography of Al-Khaliil Ibn Ahmad al-Farahiidii'*. Baghdaad, Az-zahraa? Press. A revised version of an M.A. thesis, University of Cairo.

Aniis, I. (1952) Fial-lahajaat ?al Arabiiyah. *'Arabic Dialects'*. Cairo.

Aniis, I. (1971) ?al-?aSwaat ?al-luɣawijja. *'The Linguistic Sounds'*. The Anglo-Egyptian bookshop Cairo.

Arkebauer, H. J.; Hixon, T. J. and Hardy, J. (1967) Peak intraoral air pressures during speech. *JSHR* 16: 141-150.

Ashley, M. J. Brackett, I. P. and Lehr, R. B. (1985) A descriptive study of the relationship of intraoral air pressure to EMG activity of the levator palatini during CV syllables. Folia *Phoniatrica*, 37: 66-74.

Baran, J. A. Laufer, M. Z. and Daniloff, R. (1977) Phonological contrastivity in conversation: a comparative study of voice onset time. *J. of Phon.*, 5: 339-350.

Bell-Berti, F (1980) Velopharyngeal function: A Spatial-temporal Model. In N.J. Lass (ed.) *Speech and Language: Advances in Basic Research and Practice*. Vol.4 : 291-316.

Birkeland, H. (1954) *Stress Patterns in Arabic*. Oslo. Jacob Dybwad.

Black, J.W. (1950) The pressure component in the production of consonants. *J. Speech and Hearing Disorders*. 15: 207-210.

Blanc, H. (1953) *Studies in North Palestinian Arabic*. Jerusalem, Central Press.

Blanc, H. (1967) The 'sonorous' vs 'muffled' distinction in old Arabic phonology. In *To Honor Roman*

Jakobson, 1: 295-390. Mouton, The Hague.

Blumstein, S. E. and Stevens, K. N. (1980) Perceptual invariance and onset spectra for stop consonants in different vowel environments. *JASA* 67: 648 - 662.

Bohas, G; Guillaume, J. P. and Koulloughli, D. A. (1990) *The Arabic Linguistic Tradition*. Routledge, USA.

Borden , G.J and Harris ,K.S. (1984) *Speech Science Primer: Physiology, Acoustics and Perception of Speech.* (2nd edition): Library of Congress.

Bravmann, Meïr. M. (1942) Notes on the study of the Arabic dialects of Yemen. *Tarbiz*, 13: 165-177.

Brown,W.S. and McGlone, R.E .(1969) Relation of intraoral air pressure to oral cavity size. *Folia Phoniatrica* 21: 321-331

Brown, W. S., McGlone, R.E and Proffit (1973) Relationship of lingual and intraoral air pressures during syllable production. *JSHR* 16: 141-150.

Brown , W.S. and McGlone, R. E. Tarlow, A. and Shipp, T.H. (1970) Intraoral air pressures associated with specific phonetic positions. *Phonetica* 22: 202-212.

Brown, W. S. and Commerford, M.A. (1979) Supraglottal air pressure and lingual occlusion in midvocalic stop consonant production. *J. of Phon.*: 269-277.

Butler, C. (1985) *Statistics in Linguistics*. Basil Blackwell. New York.

Caramazza, A. Yeni-Komshian, G., Zurif, E. and Carbone, E. (1973) The acquistion of a new phonological contrast: the case of stop consonants in French-English bilinguals. *JASA*, 54: 421-8.

Caramazza, A. Yeni-Komshian, G. (1974) Voice onset time in two French dialects. *J. of Phon.*, 2:239-245.

Chen, M. (1970) Vowel length variation as function of the voicing of the consonant environment. *Phonetica*, 22: 129-159.

Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*. New York. Harper and Row.

Cole, R. A. and Scott, B. (1974a) The phantom in the phoneme: invariant cues for stop consonants. *Per. and Psychphys.*, 15: 101-107.

Cole, R. A. and Scott, B. (1974b) Toward a theory of speech perception. *Psych. Rev.* 81: 348-374.

Cooper, F. S. Delattre, P. C. Liberman, A. M. Borst, J.M. and Gerstman, L. J. (1952) Some experiments on the perception of synthetic speech sounds. *JASA*, 24: 597- 606.

Couper-Kuhlen, E. (1986) *An Introduction to English Prosody*. Edward Arnold, Baltimore, USA.

Crystal, T. H. and House, A.S. (1982) Segmental durations in connected speech signals: Preliminary results. *JASA*, 72: 705-716.

Crystal, T. H. and House, A.S. (1988) The duration of American English stop consonants: an overview. *J. of Phon.*, 16: 285-294.

Dart, S. (1987) Aerodynamic study of Korean stop consonants: measurements and modeling. *JASA* 81: 138-147.

Delattre, P. C. Liberman, A. M. Cooper, F. S. (1955) Acoustic loci and transitional cues for consonants. *JASA* 27, 769-773.

Delattre, P. (1962) Some factors of vowel duration and their cross-linguistic validity. *JASA*, 34: 1141-1143.

Delattre, P. (1971) Pharyngeal features in the consonants of Arabic, German, Spanish, French and American English. *Phonetica*, 23: 129-155.

Denes, P. (1955) Effect of duration on the perception of voicing. *JASA*, 25: 105-113.

Dixit, R.P. and Brown , W. S. (1978) Peak magnitudes of supraglottal air pressure associated with affricated and non-affricated stop consonant productions in Hindi . *J. Phon*. 6 :355-365.

Dixit, R. P. (1975) *Neuromuscular Aspects of Laryngeal Control: With Special Reference to Hindi*. Doctoral Dissertation, University of Texas, Austin.

Dixit, R. P and Brown, W. S. (1986) Peak magnitudes of oral airflow during Hindi stops. *J. Phon.* 13: 219-234.

Dorman, M. F. Studdert-Kennedy, M. and Raphael L .J. (1977) Stop-consonant recognition: release burst and formant transitions as functionally equivalent, context-dependent cues. *Perc. and Psychphys.* 22: 109- 122.

Elgendy, A.E. (1992) The role of the jaw in constriction adjustments during pharyngeal and pharyngealized articulation. PERILUS, 14: 59-83.

El-HaleeeS (1985) The role of F1 in the place-of-articulation distinction in Arabic. *J. of Phon*., 13: 287-298.

Emanuel, F. and Counihan, D.T. (1970) Some characteristics of oral and nasal airflow during plosive consonant production. *Cleft Palate Journal* 7: 249-260.

Fant, G. (1960) *Acoustic Theory of Speech Production*. Mouton: The Hague, Paris.

Fant, G. (1970) Analysis and synthesis of speech processes. In B. Malmberg (ed.), *Manual of Phonetics:* 173-277. Amsterdam: North Holland.

Fant, G.(1973) 'Stops in CV-syllables'. In S*peech Sounds and Features*, edited by Gunnar Fant (MIT, Cambridge MA): 110-139.

Fant, G. Kruckenberg, A. and Nord, L. (1991) Durational correlates of stress in Swedish, French and English. In Proc. of the Leeds Seminar on Speech Production. (J. Phon. 19: 351-365).

Ferguson, C. A. (1956b) Review of stress patterns in Arabic by Birkeland. *Language*, 32: 384-387.

Ferguson, C. A. (1956a) The emphatic /L/ in Arabic. *Language*, 32: 446-452.

Fischer-Jorgensen. E (1954) Acoustic Analysis of Stop Consonants. *Miscellanea Phonetica* 2: 42-59.

Fischer-Jorgensen. E and Tybjaerg Hansen (1959) An electrical manometer and its use in phonetic research. Phonetica, 4: 43-53.

Fischer-Jorgensen. E (1964) Sound duration and place of articulation. In: *Zeitschrift fur Sprachwissenschaft und kommunikations forschung*, 17: 175-207.

Fischer-Jorgensen. E (1977) Temporal relations in consonant-vowel syllables with stop consonants based on Danish material. In: *Frontiers in Speech Communication Research*. Lindblom, B. and Öhman, S. (eds.). Academic, New York: 109-123.

Flege, J. E. and Port, R. (1981) Cross-language phonetic interference: Arabic to English. *Language and Speech,* 24: 125-146.

FrayHa, A. (1981) Naδaryaat fil-lughah. *'Theories of Language'*. The University Press, Beirut.

Fritzell, B. (1979) Electromyography in the study of the velopharyngeal function: A review. *Folia*

Phoniatrica, 31: 93-102.

Fry, D. B. (1958) Experiments in the perception of stress. *Language and Speech*, 1: 126-152.

Fry, D. Abramson, A. Eimas, P. and Liberman, A. (1962) The identification and discrimination of synthetic vowels. *Language and Speech* 5: 171-189.

Fujimura, O. (1971) Remarks on stop consonants - synthesis experiments and acoustic cues. In Hammerich, Jakobson and Zwirner (eds.) *Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer Jorgensen.* Copenhagen: 221-232.

Gairdner, W. H. T. (1925) *The Phonetics of Arabic.* Oxford University Press.

Gairdner, W. H. T. (1935) The Arab phoneticians on the consonants and vowels. *Moslem World*, 25: 242-257.

Garbell, I. (1958) Remarks on the historical phonology of an East-Mediterranean Arabic dialect: *Word*, 14: 303- 337.

Gay, T. (1979) Coarticulation in some consonant-vowel and consonant cluster -vowel syllables. In *Frontiers in Speech Communication Research* edited by Lindblom, B and Öhman, S. Academic Press.

Ghazeli, C. (1977) *Backing feature and backing coarticulation in Arabic.* Ph.D Dissertation, University of Texas at Austin

Giannini, A. and Pettorino, M. (1982) The emphatic consonants in Arabic. Speech lab. SLR No. 4: (Instituto Universitario Orientale, Napoli).

Gilbert, H.R. (1973) Oral Airflow during stop consonant production. *Folia Phoniatrica*, 25: 288-301.

Gimson, A. C. (1989) An introduction to the pronunciation of English. Edward Arnold London. (4th edition).

Greenberg, J.H. (1950) The patterning of root morphemes in Semitic languages. *Word*, 6: 162-181.

Grimm, W. A. (1966) Perception of segments of English spoken consonant-vowel syllables. *JASA*, 40: 1454-1461.

Haag , W. K. (1977) Oral air pressure and airflow in German stop consonants. *Phonetica*, 34: 31-39.

Haggard, M. B. Ambler, S. and Callow, M. (1970) Pitch as a voicing cue. *JASA*, 47: 613-617.

Haggard, M. Summerfield, Q. and Roberts, M. (1981) Psychoacoustical and cultural determinants of phoneme boundaries: evidence from trading F0 cues in the voiced-voiceless distinction. *J.of Phon.* 9: 49-62.

Halle, M.; Hughes, G. W. and Radley, A. (1957) Acoustic properties of stop consonants. *JASA*, 29: 107-116.

Hardcastle, W. J. (1973) Some observations on the tense-lax distinction in initial stops in Korean. *J. of Phon.,* 1: 263-272.

Hardcastle, W.J. (1976) *Physiology of Speech Production.* Academic Press.

Harrell, R. (1957) *The Phonology of Colloqial Egyptian Arabic.* American Council of Learned Societies. New York.

Harrell, R. (1962) Consonant, vowel and syllable in Moroccan Arabic. Proc. of the Fourth Inter. Cong. of Phonetic Sciences: 643-647. Mouton, The Hague.

Harris, Z.S. (1942) The phonemes of Moroccan Arabic. *J. of the American Oriental Society* 62: 309 - 318.

Hatch, E. and Farhady, H. (1982) *Research Design and Statistics for Applied Linguistics.* New York: Harper and Row.

Hazan,V. and Rosen, S. (1991) Individual variability in the perception of cues to place contrasts in initial stops. *Perc. and Psychphys.* 49: 187-200.

Hirose, H .Lisker, L. and Abramson, A.S. (1972) Physiological aspects of certain laryngeal features in stop production. Haskins Laboratories: Status Report on Speech Research. SR No. 31/32: 181-191.

Hixon, T.J. Minifie, F.D. and Tait, A.T. (1967) Correlates of turbulence noise production for speech. *JSHR* 10: 133-140.

Homma, Y. (1981) Durational relationships between Japanese stops and vowels. *J. of Phon.* 9: 273- 281.

House, A.S. and Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels. *JASA*, 25: 105-112.

Hudgins, C.V. and Stetson, R.H. (1937) Voicing of consonants by depression of the larynx. *Arch. Neerland. de Pho. Exper,* 13. 85- 94.

Huggins, A. (1972) Just-noticeable differences for segment duration in natural speech. *JASA*, 51: 1270- 1278.

Huizinga, M. (1984) *Spoken Yemeni Arabic*. Amsterdam.

Ibn Siinaa, Abu Ali Al-Husayn (c. circa. 1353) Asbaab Huduθ ?al-Huruf *'The Causes of Sound Production'*. Cairo: Al-Maktaba Al-Salafiyah.

Isshiki, N. and Ringel, R. (1964) Airflow during the production of selected consonants. *JSHR* 7: 233-244.

Jakobson, R. and Halle, M. (1956) *Fundamentals of Language*. 's-Gravenhage.

Jakobson, R. (1957) Muffaxxama, the 'emphatic' phonemes in Arabic. *Studies Presented to Joshua Whatmough. In Jakobson, R. Selected Writings*. (1962). Mouton, the Hague. Paris..

Johnstone, T. M. (1963) The affrication of 'Kaf' and 'Gaf' in Arabic dialects of the Arabian Peninsula. *J. of Semitic Studies,* 8: 212-226.

Johnstone, T. M. (1967) *Eastern Arabian Dialect Studies*. London. OUP.

Jones, D. (1976) *An Outline of English Phonetics*. Cambridge University Press (9th Edition).

Karnell, M.P. and Willis, C. R. (1982) The effect of vowel context on consonantal intraoral air pressure. *Folia Phoniatrica*. 34: 1-8.

Kent, R.D. and Moll, K.L. (1969) Vocal-tract characteristics of the stop cognates. *JASA* 46: 1549-1555.

Kewley-Port, D. Pisoni, D. B. and Studdert-Kennedy. M. (1983) Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *JASA*, 73: 1779-1793.

Kewley-Port, D (1982) Measurements of formant transitions in naturally produced stop consonant-vowel syllables. *JASA* 72: 379-389.

Kim, C. W. (1965) On the autonomy of the tensity feature in stop classification with special reference to Korean stops. *Word*, 21: 339-359.

Kim, C. W. (1970) A theory of aspiration. *Phonetica*, 21: 107-116.

Klatt, D. H. and Stevens, K.N (1969) Pharyngeal consonants. *QPR-RLE.* (MIT), Cambridge, 93: 207- 216.

Klatt, D. H. (1975) Voice onset time, frication and aspiration in word-initial consonant clusters. *JSHR*, 18: 686-706.

Klatt, D. H. (1976) Linguistic uses of segmental duration in English: acoustical and perceptual evidence.

JASA, 59: 1208-1221.

Klatt, D. H. (1977) Synthesis by rule of segmental duration in English sentences. In: *Fronties in Speech Communication Research*. Lindblom, B. and Öhman, S. (eds.). Academic, New York: 109-123.

Klatt D. H. (1980) Software for a cascade/parallel formant synthesizer. *JASA*, 67: 971-995.

Kunze, L.H. (1964) Evaluation of methods of estimating subglottal air pressure. JSHR 7: 151-164.

Ladefoged, P. (1963) Some physiological parameters in speech. *Language and Speech* 6: 109-119.

Ladefoged, P. (1982) *A Course in Phonetics*. Harcourt-Brace.

Lahiri, A. and Hankamer, J. (1988) The timing of geminate consonants. *J. of Phon*., 16: 327-338

Lane, H.L. (1965) The Motor Theory of speech perception: A critical review. *Psych. Rev*. 72: 275-309.

LaRiviere, C. Winitz., H. and Herriman, E. (1975) Vocalic transitions in the perception of voiceless initial stops. *JASA* 57: 470-475.

Laufer, A. and Baer, T. (1988) The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech* 31: 181-205.

Lee, C.Y. and Smith, T.S. (1972) Oral and direct subglottal pressure in Korean stops. *JASA* 51: 102

Lehiste, I. and Peterson, G. E. (1961) Transitions, glides and diphthongs. *JASA*, 33: 268-277.

Lehiste, I. (1970) *Suprasegmentals.* M. I. T. Cambridge, MA.

Lehn, W. (1963) Emphasis in Cairo Arabic. *Language*, 39: 29-39.

Lehn, W. (1967) Vowel contrasts in Najdi Arabic. In *Readings in Arabic Linguistics* by S. Al-ani (ed.). Indiana University Club, Bloomington: 321-328.

Liberman, A. M. Delattre, P.C. Cooper, F. S. and Gerstman, L .J. (1954) The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs* 68 (whole No. 379).

Liberman, A. M. Harris, K. Hoffman, H.S. and Griffith, B.C. (1957) The discrimination of speech sounds within and across phoneme boundaries. *J. of Exper. Psych*., 54: 358- 368.

Liberman, A. A. Delattre, P. and Cooper, F.S. (1958) Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech* 1: 153-167.

Liberman, A. M. Cooper, F. S. Shankweiler, D. and Studdert-Kennedy, M. (1967) Perception of the speech code. *Psyc. Rev*., 74: 431- 460.

Lisker, L. (1957) Closure duration and the intervocalic voiced-voiceless distinction in English. Language, 33: 42-49.

Lisker, L. (1958) The Tamil occlusives: short vs long or voiced vs voiceless? Indiana Linguistics, Vol. 1: 294-301.

Lisker, L. and Abramson, A. (1964) A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20: 384-422.

Lisker, L. and Abramson, A. (1967) Some effects of context on voice onset time in English stops. *Language and Speech,* 10: 1-28.

Lisker, L. Abramson, A.S. Cooper, F.S. & Schevy, H. (1969) Transillumination of the larynx in running speech. *JASA* 45: 1544-1546.

Lisker, L (1970) Supraglottal air pressure in the production of English stops. *Language and Speech*, 13: 215-229.

Lisker, L. and Abramson, A. (1970) The voicing dimension: some experiments in comparative phonetics. In Proceedings of the sixth Inter. Cong. of Phonetic Science, Prague, (1967): 563-567.

Lisker, L. and Abramson, A. (1971) Distinctive features and laryngeal control. *Language*, 47: 776-785.

Lisker, L. (1972) Stop duration and voicing in English. In: *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre* (A.Valdman, ed.), Paris, Mouton: 339-343.

Lisker, L. (1975) Is it VOT or a first formant detector? *JASA*, 57: 1547-1551.

Lisker, L. Liberman, A. M. Erickson, D. and Dechvitz, D. (1977) On pushing the voice onset time (VOT) boundary about. *Language and Speech*. 20: 209-216.

Littlewood, W. (1981) Communicative language teaching. Cambridge University Press.

Lofqvist, A (1975) A study of subglottal pressure during the production of Swedish stops. *J. Phon.* 3: 175-189.

Lorge, B. (1967) A study of the relationship between production and perception of initial and intervocalic /t/ and /d/ in individual English speaking adults. Haskins Laboratories: Status Report on Speech Research, SR-9: 3. 1-3.18.

Lubker, J. F. (1968) An Electromyographic-Cineradiographic investigation of velar function during normal speech production. *Cleft Palate Journal*, 5: 1-18.

Lubker, J.F (1973) Transglottal airflow during stop consonant production. *JASA* 53: 212-215.

Luce, P.A. and Charles-Luce, J. (1985) Contextual effects on vowel duration, closure duration and the consonant/vowel ratio in speech production. *JASA*, 78: 1949-1957.

Mack, M. (1982) Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *JASA*, 71: 173-178.

MacNeilage, P. F. (1968) The serial ordering of speech sounds. Project on linguistic analysis. Phonology Laboratory, University of California, Berkeley 2.8: 1-52.

MacNeilage, P. F. and DeClerk, J. L. (1969) On the motor control of coarticulation in CVC monosyllables. *JASA*, 45: 1217-1233.

Malècot, A. (1955) An experimental study of force of articulation. Studia Linguistica, 9: 35-44.

Malècot, A. (1966a) The effectiveness of intraoral air-pressure pulse parameters in distinguishing between stop cognates. *Phonetica*, 14: 65-81.

Malècot, A. (1966b) Mechanical pressure as an index of 'Force of Articulation'. *Phonetica* 14: 169-180.

Malècot, A. (1968) The force of articulation of American stops and fricatives as a function of position. *Phonetica*, 18: 95-102.

Malècot, A. (1970) The lenis-fortis opposition: its physiological parameters. *JASA*, 47: 1588-1592.

MarÇais, P. (1948) L'articulation de l'emphase dans un parler arabe maghre'bin. In Annales de l'Institut d'etudes orientales, vol. 7, Alger.

Massaro, D.W. and Cohen, M.M. (1983) Consonant/vowel ratio: An improbable cue in speech. *Per. and Psychphys*. 33: 501-505.

Mattingly, I. Liberman, A. M. Syrdal, A. K. and Halwes, T. (1968) Discrimination of F2 transitions in speech context and in isolation. Haskins Laboratories, SR 15/16: 67- 79.

McGlone, R. E. and Shipp, T. (1971) Comparison of subglottal air pressures associated with /p/ and /b/. JASA, 51: 664-665.

Meddis, R (1973) *Elementary analysis of variance for the behavioural sciences.* London: McGraw-Hill Book Company, UK.

Menon, K. M. N.; Rao, P.V.S.; and Thosar, R.B (1974) Formant transitions and stop consonant perception in syllables. *Language and Speech,*17: 27-46.

Mermelstein, P. (1978) Difference limens for formant frequencies of steady-state and consonant-bound vowels. *JASA* 63: 572-580.

Miller, C.J. and Daniloff, R. (1977) Aerodynamics of stops in continuous speech. *J. of Phon.* 5: 351-360.

Miller, S. (1984) *Experimental Design and Statistics.* (2nd edition).

Miller, J. D. Pastore, R. E. Weir, C. C. Kelly, W. J. and Dooling , R. J. (1974) Discrimination and labelling of noise-buzz sequences with varying noise lead times. *JASA*, 55: P390.

Mitchell, T. F. (1960) Prominence and syllabication in Arabic. In *Readings in Arabic Linguistics*, S. Al-ani (ed.). Indiana University Club, Bloomington: 329-355.

Mitchell, T. F. (1990) *Pronouncing Arabic.* Clarendon Press, Oxford.

Moll, K. L. (1962) Velopharyngeal closure in vowels. *JSHR* 5: 30-37.

Muller, E. M. and Brown, W.S. (1980) Variations in the supraglottal air pressure waveform and their articulatory interpretations in N.J. Lass (ed.) *Speech and Language: Advances in Basic Research and Practice.* Vol.4. New York Academic Press: 317-389.

NaSr, R.T. (1959a) Velarization in Lebanese Arabic. *Phonetica*, 3: 203-209.

NaSr, R.T. (1959b) The predictibility of stress in Lebanese Arabic. *Phonetica*, 4: 89-94.

Netsell, R. (1969) Subglottal and intraoral air pressures during intervocalic contrasts of /t/ and /d/. *Phonetica*, 20: 68-73.

Nihalani, P. (1974) An aerodynamic study of stops in Sindhi. *Phonetica*, 29: 193-224.

Nihalani, P. (1975a) Velopharyngeal opening in the formation of voiced stops in Sindhi. *Phonetica*, 32: 89-102.

Nihalani, P. (1975b) Airflow rate in the production of stops in Sindhi. *Phonetica* 31: 198-205.

Obrecht, D.H. (1965) Three experiments in the perception of geminate consonants in Arabic. *Language and Speech*, 8: 31-41.

Obrecht, D.H. (1968) *Effects of the second formant on the perception of velarization consonants in Arabic.* The Hague: Mouton.

O'Connor, J.D. (1973) *Phonetics.* Harmondsworth, Penguin Books.

Odisho, E. Y. (1973) *The role of the rear section of the vocal tract in Baghdadi Arabic.* An Unpublished M.Phil. Thesis, Department of phonetics, the University of Leeds.

Ohala, J (1990) Respiratory activity in speech. In Hardcastle, J.W. and Marchal, A. (eds.) *Speech Production and Speech Modeling.* Kluwer Academic Press: 23-53.

Ohala, M (1980) Phonological features of Hindi stops. Report of the Phonology Laboratory, 5: 96-105. University of California Berkeley.

Öhman, S. E. G. (1966) Coarticulation in VCV utterances: spectrographic measurements. *JASA*, 39:151-168.

O'Shaughnessy, D. (1981) A study of French vowel and consonant durations. *J. of Phon.*, 9: 385-406.

O'Shaughnessy, D. (1984) A multi-speaker analysis of durations in read French paragraphs. *JASA*, 76: 1664-1672.

Painter, C. (1979) *An Introduction to Instrumental Phonetics.* University Park Press, Baltimore.

Panconcelli-Calzia, G. (1924) Die experimentelle phonetik in inhrer Anwendung auf die Sprachwissenschaft, Berlin.

Pastore, R. E. (1976) Categorical perception: A critical re-evaluation. In *Hearing and Davis: Essays Honounring Hallwell Davis.* edited by S.K. Hirsh and D.H. Eldredge, I. J.Hirsh S.R. Silverman (Washington University) St. Louis: 253-264.

Perkell, J.S. (1969) *Physiology of Speech Production: Results and Implications of Quantitative Cineradiographic Study.* (MIT Press, Cambridge).

Peterson, G. and Lehiste, I. (1960) Duration of syllable nuclei in English. *JASA*, 32: 693-703.

Pickett, J.M. and Decker, L.R. (1960) Time factors in perception of a double consonant. *Language and Speech,* 3: 11-17.

Pickett, J. M. (1980) *The Sounds of Speech Communication.* University Park Press, Baltimore.

Pisoni, D.B. and Lazarus, J. H. (1974) Categorical and non-categorical modes of speech perception along the voicing continuum. *JASA*, 55: 328-333.

Port, R.F. (1979) Influence of tempo on stop closure duration as a cue for voicing and place. *J. of Phon.*, 7: 45-56.

Port, R.F. (1981) Linguistic timing factors in combination. *JASA*, 69: 262-274.

Port, R.F. and Dalby, J. (1982) Consonant/vowel Ratio as a cue for voicing in English. *Per. and Psychphys.*, 32: 141-152.

Potter, R. Kopp, G. A. and Green, H.G. (1947) *Visible Speech.* (Van Nostrand Comp.).

Qafisheh, H. A. (1990) *Yemeni Arabic I.* Librairie Du Liban, Beirut.

Rabin (1971) Semitic languages. In *Encyclopedia Judaica.* 14: 1149-1157. Kater, Jerusalem.

RaHiim, A. J. (1980) *The Phonology of Spoken Iraqi Arabic from the Functional Point of View.* Unpublished Ph.D. Thesis, Department of Linguistics and Phonetics, University of Leeds.

Raphael, L.J. (1981) Duration and contexts as cues to word-final cognate oppositions in English. *Phonetica*, 38: 126-147.

Repp, B.H. (1979) Relative amplitude of aspiration noise as a voicing cue to syllable initial stop consonants. *Language and Speech* 22: 173-189.

Roach, P. J. (1983) *English Phonetics and Phonology*. Cambridge University Press. London.

Rothenberg, M. (1968) Breath-stream dynamics of simple-released-plosive production. Bibliographica, *Phonetica*, 6: 6-22.

Rothenberg, M. (1977) Measurement of airflow in speech. *JSHR* 20: 155-176.

Rubin, J. (1975) What the good language teacher can teach. TESL: 41-51.

Schatz, C. D. (1954) The role of context in the perception of stops. *Language* 30: 47-56.

Scully, C. (1969) Problems in the interpretation of pressure and airflow data in speech. In Phonetics Department Reports, University of Leeds, No. 2: 53-92.

Scully, C. (1974) A synthesizer study of aerodynamic factors in speech segment durations. In Proceedings of Speech Communication Seminar, Stockholm (Fant, G. ed.). Vol. 2: 227-237.

Scully, C. (1976) Some interrelated factors in the release stage of plosives. Occasional. papers University of Essex, Dep. Lang. Ling. 17: 143-149.

Scully, C. (1983) Voice onset time: acoustic contrasts of homorganic plosives. Unpublished paper. The Department of Linguistics and Phonetics; University of Leeds.

Scully, C. (1990) Articulatory synthesis. In *Speech Production and Modeling.* NATO Series, 151: Hardcastle and Marchal (eds.).

Sharf, D.J. (1962) Duration of poststressed intervocalic stops and preceding vowels. *Language and Speech* 5: 26-30.

Sharf, D.J. and Hemeyer, T. (1972) Identification of place of consonant articulation from vowel formant transitions. *JASA*, 51: 652-658.

Sibawayhi, Abu Bishr ζamr (1966) *Al-kitaab, 'The Book'* . Vol.1. Edited with an Introduction by Abd al-Salaam Haruun. Daar al-Qalam. Cairo.

Simon, C. and Fourcin, A. J. (1978) Cross-language study of speech learning. *JASA*, 63: 925-935.

Smith, B.L. (1978) Temporal aspects of English speech production. *J. of Phon.*, 6: 37-68.

Stathopoulos, E.T. and Weismer, G. (1983) Closure duration of stop consonants. *J. of Phon.*, 11: 395- 400.

Stetson, R. H. (1951) *Motor Phonetics.* Amsterdam (2nd Editition.).

Stevens, K. N. and House, A. S. (1956) Studies of formant transitions using a vocal tract analog. *JASA*, 28: 578-585.

Stevens, K. N. (1966) On the relation between speech movements and speech perception. In *Seminar on Speech Production and Perception*. Cistovic and Fant (eds). Leningrad, USSR: 102-106.

Stevens, K. N. and Klatt, D.H. (1974) Role of formant transitions in the voiced -voiceless distinction for stops. *JASA*, 55: 653-659.

Stevens, K. N. and Blumstein, S. E . (1978) Invariant cues for place of articulation in stop consonants. *JASA* 64: 1358-1368.

Studdert-Kennedy, M. Liberman, A. M. Harris, K. and Cooper, F. S. (1970) The Motor Theory of speech perception: A Reply to Lane's Critical Review. *Psych. Rev.* 77: 234-249.

Suomi, K. (1985) The vowel-dependence of gross spectral cues to place of articulation of stop consonants in CV-syllables. *J. of Phon.*, 3: 267-284.

Subtelny, J. D. Worth , J.H. & Sakuda, M. (1966) Intraoral pressure and rate of flow during speech. *JSHR* 9: 498-518.

Trubetzkoy, N. S. (1935) *Introduction to the Principles of Phonological Description*. The Hague: Martinus

Nijhuff. (English translation by L. A. Murray (1968).

Umeda, N. (1977) Consonant duration in American English. *JASA*, 61: 846-858.

Van den Berg, J.W. Zantema, J.T. and Doornenbal, P. (1957) On the air resistance and the Bernoulli Effect of the human larynx. *JASA* 29: 626-631.

Van den Berg, J.W. (1958) Myoelastic-aerodynamic theory of voice production. *JSHR* 1: 227-244.

Wajskop, M. (1977) Segmental durations of French intervocalic plosives. In: *Frontiers in Speech Communication Research.* Lindblom, B and Öhman, S (eds). Academic Press, New York: 109- 123.

Warren, D.W. and Wood, M.T. (1969) Respiratory volumes in normal speech: a possible reason for intraoral pressure differences among voiced and voiceless consonants. *JASA* 45: 466-469.

Warren, D. (1976) Aerodynamics of speech production. In *Current Issues in Experimental Phonetics.* (Lass, N. ed.): 105-135. New York Academic Press.

Watson, I. (1992) Cross-linguistic and individual variability in the perception of voicing distinction. Oxford Phonetics Laboratory: Progress Report No. 5: 73-83.

Williams, L. (1977) The voicing contrast in Spanish. *J. of Phon.*, 5: 169-184.

Winitz, H. Scheib, M. E. and Reeds, J. A. (1972) Identification of stops and vowels from the burst portion of /p, t, k / isolated from conversational speech. *JASA* 51: 1309 -1317.

Yeni-Komshian, G. H. Caramazza, A. and Preston, M.S. (1977) A study of voicing in Lebanese Arabic. *J. of Phon.* 5: 35-48.

Zimmermann, S.A. and Sapon, S.M. (1958) Note on vowel duration seen cross-linguistically. *JASA*, 30: 152-153.

Zlatin, M. A. (1974) Voicing contrast: perceptual and productive voice onset time characteristics of adults. *JASA,* 56: 981-994.

Zue, V.W. (1976) *Acoustic characteristics of stop consonants: A controlled study.* Doctoral Dissertation, M.I.T., Cambridge, MA.