It is the mark of an instructed mind
to rest satisfied
with that degree of precision
which the nature of the subject admits,
and not to seek exactness
where only an approximation
of truth is possible.

- ARISTOTLE

CROSS-LANGUAGE STUDY OF VELAR AND BILABIAL STOP PERCEPTION

VINAY RAKESH

REG.No.M8802

DISSERTATION SUBMITTED AS PARTIAL FULFILMENT FOR THE DEGREE OF

M.Sc. (SPEECH AND HEARING) TO UNIVERSITY OF MYSORE, MYSORE

ALL INDIA INTITUTE OF SPEECH AND HEARING: MYSORE

MAY, 1990

TO

MY GUIDE

AND

THE GEMINATE CLUSTER

WILD ROSES

MY

"EVERLASTING INSPIRATIONS"

CERTIFICATE

This is to certify that this Dissertation entitled: "CROSS-LANGUAGE STUDY OF VELAR AND BILABIAL STOP PERCEPTION" is a bonafide work done in part fulfilment for the degree of M.Sc.(Speech and Hearing) of the student with Reg.No. M.8802.

Mysore

May, 1990

Director
All India Institute of
Speech and Hearing
Mysore.

CERTIFICATE

This is to certify that this Dissertation entitled "Cross-Language Study Of Velar And Bilabial Stop Perception" has been prepared under my supervision and guidance.

Savithri S.R

Mysore
May, 1990

Dr.S.R.Savithri
Guide

DECLARATION

This dissertation entitled "Cross-Language Study of Velar and Bilabial Stop Perception" is the result of my own work undertaken under the guidance of Dr.Savithri S.R. Lecturer in Speech Sciences, All India Institute of Speech and Hearing, Mysore-6, and has not been submitted earlier at any University or Institution for any other Diploma or Degree.

Mysore.
May, 1990 Reg.No. M8802

ACKNOWLEDGEMENTS

I owe my sincere gratitude to:

Dr. Savithri, S.R., Lecturer in Speech Sciences, for her supervision and guidance throughout this study.

Dr.(Miss) S.Nikam, Director, A.I.I.S.H., for giving me an opportunity to take up this work.

Dr.Nataraja, N.P., Professor and Head of the Department of Speech Sciences, for permitting me to avail all the facilities of the department for the completion of this study.

Mr.Venkatesh, C.S., Lecturer in Speech Sciences, for his timely suggestions and encouragement.

Ms.Sridevi, H.S., Research Scholar, for the references and her kind help.

Ms.Usharani, P.N., for making this study simple and easier for me.

My friends, Srividya, Suresh, T, Mithilesh, Vinay Kumar, Anu, Sunil, Siddheshwar and all my classmates for their help and encouragement.

My subjects, for their kind co-operation and contribution to this study.

Ms.Rajalakshmi R Gopal for doing an excellent work.

C O N T E N T S

LIST OF TABLES

contd..

LIST OF FIGURES

1.INTRODUCTION

Speech is an acoustic stimulus which is a series of time related events and relationships of pressure values and energy distribution (Sanders, 1977). It is the conversion of language into sounds, it is a method of oral communication, the language written on breath. Communication through speech is transmission of thoughts or feelings from the mind of a speaker to the mind of a listener. According to Sanders (1977) it is a matter of evoking equivalent ideas rather than sending or exchanging them.

The speech signal which seems like long spurts of a complex and constantly changing stream of sounds radiates from the speaker's lips, travels in air, inpinges upon the eardrum of the listener and reaches the higher cortical centres through middle and inner ears and the auditory pathway. The speech signal is analyzed at lower centres (below the thalamus level) to some extent and processing of specific speech parameters and other complex acoustic features of natural stimuli begins only at the level of medial geniculate body (MGB), which is located in the thalamus (Kiedel, Kallett, Korth and Humes, 1983). The linguistic components are added only at the higher centres of cortex to the already analyzed signal to reconstruct the percept intended by the speaker. When the

listener has reconstructed this signal, it is said that he has perceived what the speaker has uttered. Thus the auditory perception of speech is a process of interpreting the instruction imprinted on the acoustic wave by the speaker over a time span or as Berry (1969), "Auditory states perception of speech per se deals mainly with the temporal management of information from the input" (Sanders, 1977).

The interest in perception and perceptual processing has grown since 1945 with increased neurological injuries because of the wars. The role of auditory perception in both perceptual and conceptual learning (particularly in the acquisition and use of language) has been recognized for many years. Moreover, research in this area has added and in future will add to the knowledge of human speech perception and understanding of the functions of neuro-anatomical structures. Also the information is of use for teaching speech to the non-verbally handicapped and for constructing speech synthesizers and text-to-speech conversion systems. This is also of help in automatic speech recognition. These utilities have triggered several experiments on human speech perception in the past few decades.

From the speech production studies it is known that speech sounds have different acoustic cues like the formants, band widths and level of formants, F_o , energy, duration of closure, burst energy, VOT etc.

In most of the perceptual studies speech sounds are reconstructed from their known spectral and temporal parameters and presented to the listeners for judgement. Various parameters of the acoustic signal can be altered individually or in combination to evaluate the effect of their cues on listener's perception. The different techniques used in the perceptual studies are analysis-by-synthesis (Halle and Stevens, 1959), articulatory studies (Pant, 1960) and synthesis-by-rule (Flanagan et al, 1970). These techniques have been used to assess the role of some of the temporal parameters like VOT (Lisker, 1975; Lahiri, 1980; Keating, Mikos and Ganong, 1981; Winitz et al, 1975), transition duration (Summerfield and Haggard, 1977); Dorman and Raphael, 1980) and spectral parameters like F_1 onset frequency, F_2 onset frequency (Lisker, 1977; Alwan, 1989) etc. within and across languages.

Linguistic expectation and articulatory knowledge are important in speech perception (Dorman, Raphael and Liberman, 1979; Lisker and Abramson, 1964). Bilingual subjects seem to divide stimulus (prefaced /ra/ and /la/ stimuli with questions spoken in English) according to phonemic contrasts of the particular language they were using immediately before each stimulus (Elman, Diehl, and Buchwald, 1977). Also, several studies reviewed by Strange and Jenkins (1978) offer evidence that the language experiences of adults can influence their perception. For

example, Spanish, French and Thai speakers use different VOT criteria for voicing contrast than do English speakers. In Arabic all syllables begin with a consonant. Hence, a steady-state stimulus, for example, which would be heard normally by English-speaking listeners as just a vowel will be heard by Arabs as (?V) where (?) is a glottal stop (Alwan, 1989). Thus the linguistic knowledge seems to influence the percept. This warrants research in and across various languages of the world. Till date, the process of human speech perception is not understood. The research has been mainly in languages other than Indian.

In this context, the present study is planned to investigate the effect of five temporal parameters (closure duration, Preceding vowel duration, Transition duration of the preceding and the following vowels, and VOT) in cueing the cluster and voicing feature of the geminate stop /kk/ /gg/, /pp/ and /bb/ as occuring in the words /akka/, /agga/, /appa/ and /abba/. The perceptual difference across native Telugu and Malayalam speakers, percept change for meaningful and nonsense stimuli by different language listeners and sex difference in the perception are also studied.

2. GENERAL REVIEW

This is a general review which deals with the theories of speech perception, cues for stop consonants and cross-linguistic studies in the perception of speech sounds. The review specific to the different parameters will be presented separately in the respective chapters.

2.1: Theories of Speech Perception:

There are two types of classification for the theories of speech perception.

1. Active and Passive
2. Variance and Invariance

2.1.1: Active theories - According to the active theories, speech perception depends on speech production. The sounds are sensed, analyzed for their phonetic properties by reference to how such speech sounds are produced and thus are recognized (Borden and Harris, 1980).

The two active theories are -

a) Motor theory of speech perception (Liberman et al, 1972);
b) Analysis-by-synthesis (Stevens and Halle, 1959).

2.1.1.1: Motor theory of speech perception (Liberman et al, 1972):- According to the Motor theory of speech perception, the production and perception are two different entities.

There is an inbuilt vocal tract in the brain, which assumes the articulatory posture similar to the vocal tract of the speaker, based on its articulatory knowledge in order to perceive the incoming auditory signal. Here the reference is mainly articulatory.

2.1.1.2: Analysis-by-Synthesis Theory (Stevens and Halle, 1959) :- It is similar to the motor theory in that the listener refers to production, but the reference is more acoustic, less articulatory and it relies on a system of matching. The listener receives the auditory pattern and analyzes it by eliciting an auditory model of his own production of the phoneme, and, if the patterns match, he accepts his perception as correct (Borden and Harris, 1980).

2.1.2: Passive theories - According to the passive theories the process of speech perception is primarily sensory. The message is sensed, filtered and mapped onto acoustic-phonetic features of language (Borden and Harris, 1980). Passive theory was advocated by Fant, (1960),

According to Fant, (1960), "the perceptual mechanism shares with the speech production mechanism a common pool of distinctive features, but the listeners need not refer to production for perception. Linguistic centres in the brain are common to both incoming and outgoing messages,

but the centres responsible for subphonemic, more peripheral aspects of production and perception are viewed as independent" (Border and Harris, 1980).

2.1.3: The Quantal Theory of Speech proposed by Stevens (1972) is not a theory of speech perception in the sense, but has some implications for theory of speech perception* It relates the articulatory changes to the acoustic results. The thesis of the quantal theory is that there is a discontinuity between changes in articulatory positioning and the resultant changes in acoustic output. There are vocal tract regions in which small continuous differences in articulatory position cause little or no differences in acoustic output, but there are other regions within which small articulatory differences result in large acoustic differences. In these critical regions, a slight adjustment of articulatory placement will cause a quantal leap in terms of sound change.

"Similar kinds of relations are observed between parameters that give a measure of auditory response to a speech like sound and parameters that specify some acoustic dimensions of sound. The tendency for quantal relations among these acoustic, auditory and articulatory parameters is a principal factor shaping the inventory of acoustic and articulatory attributes that are used to signal distinctions in language" (Stevens, 1989).

Of these, the analysis-by-synthesis theory is termed the theory of invariance and the others are called as the theories of variance.

2.2: The stop consonants:

The stop consonants are produced by occluding the oral cavity by an articulator. Air is held behind the articulator for sometime and is then released. The stops represent the non-linearity of the speech production system. They also demonstrate the redundancy of acoustic cues available to distinguish speech sounds. The nature of stop perception provides the best example of listener's use of the acoustic overlapping of phonemes in speech system. Also, they have consistently produced evidence for phonetic level processing. They appear to be most highly encoded speech sounds" (Day and Vigorito, 1973).

The salient features of stop consonants are -

1. A period of occlusion (silence/voiced)
2. A transition explosion (usually less than 20 m.secs) produced by shock excitation of vocal tract upon release of occlusion.
3. A very brief (0-10 m.sec) period of frication as articulators separate and air is blown through the narrow constriction, as in homorganic fricative.

4. A brief period of aspiration (2-20 m.secs) within which may be detected the noise excited formant transitions, reflecting shifts in vocal tract resonances as the main body of the tongue moves towards the position appropriate for the following vowel.
5. Voiced formant transitions, reflecting the final stages of articulatory movement into the vowel during the first few cycles of laryngeal vibration. (Fischer-Jorgensen, 1954; Halle, Hughes and Radley, 1957; Pant, 1969).

Several experiments have been conducted to gain insight into the perceptual cues of stop consonants. The various parameters studied can be listed under spectral and temporal categories.

The spectral parameters include :-

Frequency of formants, 1, 2 and 3; bandwidths of formants 1, 2 and 3; amplitude of formants, 1, 2, and 3; direction of second and third formant (F2 and F3) transitions, voice during closure, burst amplitude, burst frequency, double burst release, F_0 changes in the preceding and following vowel and amplitude of the following vowel etc. The temporal parameters include - preceding vowel duration, closure duration, voice onset time, voice offset time, stop consonant duration, off-glide

duration of the first formant (F2/2), off-glide duration of the second formant (F½), burst duration and transition duration and speed of transition of the formant of preceding vowel etc.

Various parameters cue the place, manner and voicing of stop consonants. The acoustic cues for manner of articulation of stops are, relative silence, burst and short transitions (Borden and Harris, 1980).

The cues for place of articulation of stops are -

(i) Frequency position of burst in relation to a vowel;- high frequency burst cues /t/, low frequency burst cues /p/, and bursts perceived as /k/ are slightly above the second formant of the particular two-formants vowel i.e. high burst /k/ percepts with front vowels and low bursts /k/ percepts with back vowels. Without bursts also, stops are perceived when accompanied by transitions.

(ii) Second formant transition of the following vowel;- All the rising F2/2 transitions are perceived as bilabials (/p/ and /b/),F½ slight fall for front vowels and sharp fall for back vowels are perceived as alveolars /t,d/, sharp F2/2- fall for front vowels and slight fall of Fl for back vowels are perceived as velars /k, g/ (Borden and Harris, 1980).

The various cues for voicing of stop consonants are:-

1. Presence/absence of low frequency buzz during the closure interval.
2. Duration of closure.
3. F1/1 offset frequency before closure.
4. F1/1 offset transition duration.
5. F1/1 onset frequency following closure.
6. F1/1 onset transition duration.
7. /ae/ duration (vowel duration).
8. F1/1 cut back before closure.
9. F1/1 cut back following closure.
10. VOT cut back before closure.
11. VOT delay before closure.
12. F_o contour before closure.
13. F_o contour after closure.
14. Amplitude of /i/ relative to /ae/
15. Delay time of glottal signal preceding closure.
16. Intensity of burst following the closure.(Lisker, 1977).

It has been found that these parameters in combination (Raphael, Dorman and Liberman, 1980; Bailey and Summerfield, 1980) or in trading relation with one another (Liberman and Studdert-Kennedy, 1977; Bailey and Summerfield, 1978; Fitch, Hallwes, Erickson and Liberman, 1980), cue for the perception of stop consonants.

2.3 : Cross linguistic studies:

The influence of linguistic knowledge is one of the important aspects in speech perception. Lotz, Abramson,

Gerstman, Ingemann and Nemser (1960) studied the perception of English stops by speakers of English, Spanish, Hugarian and Thai using tape-cutting method. They concluded that there is a hierarchy among the cues in the acoustic stimulus for the perception of stops in various languages. For American English, the lack of aspiration is a dominant cue for forcing the evaluation of the stops in the direction of /b,d,g/, whereas in the languages where the distinctions exist, the evaluation is different. This examplify the reflection of the linguistic categories of the listener's native language in his perception of speech sounds.

Singh and Black (1966) studied the production and perception of twentysix intervocalic consonants in the speakers of four languages - Hindi, English, Arabic and Japanese. They reported that listening groups differed and that consonants were unequal in their intelligibility. On the whole, all speakers spoke better and all listeners listened better when saying and hearing sounds of their native language.

Elman, Diehl and Buchwald (1977) found the language 'set' that listeners have when making decisions about speech sound identity to be capable of changing the boundary between categories. Bilingual subjects divide such stimuli according to the phonetic contrast of the particular language they are using immediately before each stimulus.

Strange and Jenkins (1978) reviewed the studies of monolingual and bilingual speakers. There are evidences that the language experience of admits cam influence their perception. For example Spanish, French and Thai speakers use different VOT criteria for voicing contrasts than do English speakers.

However, recently Fox and Lehiste (1989) studied discrimination of duration ratios in bisyllabic token by native English and Estonian listeners and concluded that neither the linguistic background of the listener nor the linguistic status of the stimulus token (i.e. noise burst or bisyllables) seem to have significant effect upon the ability to make precise discrimination.

In Indian languages ,(Usha Rani,(l989) studied the effect of five temporal parameters (closure duration, preceding vowel duration, transition duration of preceding and following vowels and the voice onset time) on the perception of bilabial and velar unaspirated geminate stop consonants in Kannada and Hindi speakers. She found that there was no significant difference in the percepts by the listeners of Kannada and Hindi languages.

Though, extensive research has been carried out, stop perception by human beings is still not understood and

further the production data (Slis and Damste, 1967; Lindquist, 1972; Frokjaer-Jensen et al, 1973; Benguerel, et al 1978; Keating, 1984; Brownman and Goldstein,1986)' indicates language differences, especially for VOT. This warrants further cross-language studies.

The present study attempts at evaluating the temporal perceptual cues of velar and bilabial geminate, unaspirated, stop consonants in Telugu and Malayalam speakers.

3. GENERAL METHODOLOGY

In the present study five experiments were conducted (effect of (1) closure duration (2) preceding vowel duration (3) Transition duration of the preceding vowel (4) Transition duration of the following vowel and (5) VOT on the perception of geminate velar and bilabial stop consonants). The test environment, subjects and procedure of data collection are common to all the five experiments. The speech materials are different and they will be described in the respective chapters.

3.1: Material:

The test materialagenerated by Usha Rani (1989) were used for this study. It comprised of four synthetic $V_1CC_1V_2$ words (/akka, agga, appa and abba/) varying in one of the temporal parameters viz. closure duration, VOT of the consonant (C_1), duration of the preceding vowel (V_1), transition duration of the preceding vowel (V_1) and transition duration of the following vowel (V_2). These were generated from a PC/XT with 12 bit A/D and D/A converter at a sampling rate of 8 KHz using analysis-by-synthesis technique. In total there were one hundred and fifty seven stimuli with three token of each stimulus.

(kk - voiceless unaspirated geminated velar stop consonant, gg - voiced unaspirated geminated velar stop consonant, pp - voiceless unaspirated geminated bilabial stop consonant, bb - voiced unaspirated geminated bilabial stop consonant).

3.2: Subjects:

Ten native adult Malayalam speakers (five males and five females) and ten native adult Telugu speakers (seven males and three females) served as subjects in the perceptual judgement tasks for all the experiments. The subjects were not exposed to Kannada, so the words /akka/ and /appa/ were meaningful and /agga/ and /abba/ were nonsense for them. All of them had normal hearing and had no reported history of any ear diseases. None of them were exposed to synthetic speech stimulus previously. The age range of these subjects are in Table-1.

	Sex	No.	Age range	Mean age
Malayalam speaker	Males	5	17 - 18	17.2
	Females	5	17 - 18	17.8
Telugu speakers	Males	7	17 - 19	18
	Females	3	17 - 19	18

Table-1: Age range of subjects studied.

3.3: Test environment:

The stimuli were audio presented to the subjects at

comfortable listening levels in a quiet room. Only one subject participated at a time.

3.4: Procedure:

The subjects were instructed to listen to the audio-stimulus and indicate their responses by marking ( ) under the appropriate percepts on the response sheet provided. Closed choice of four percepts were indicated on the response sheets (Appendix-I). The subjects were allowed to listen to the stimuli three times before making perceptual judgement.

The perceptual judgement of these subjects were tabulated to evaluate the effect of temporal parameters on the perception of velar and bilabial unaspirated geminate clusters in words among the Malayalam and Telugu speakers and across languages for meaningful and nonsense stimuli. Also, the results are compared with those obtained from Kannada and Hindi speakers.

4. EXPERIMENT-I

CLOSURE DURATION AND STOP PERCEPTION

Introduction:

Closure duration is the time for which the articulator is held in position for a stop consonant.

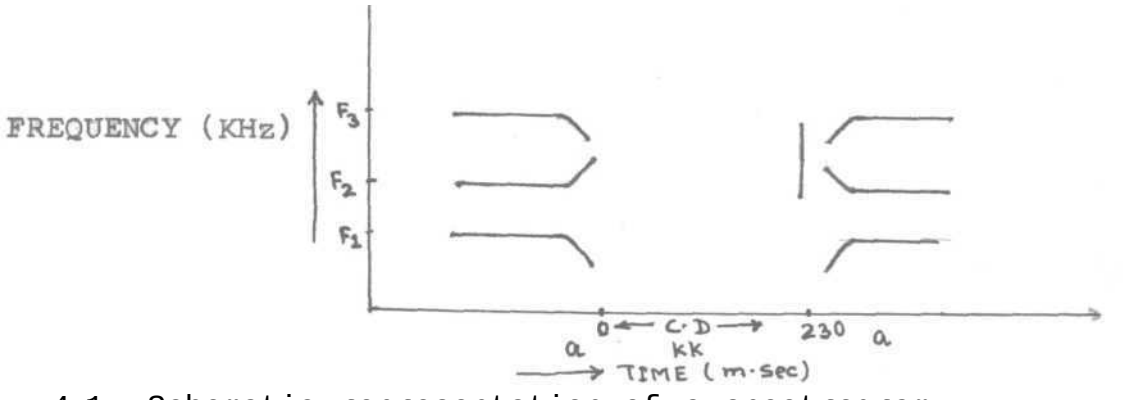

Fig. 4.1 Schematic representation of a spectrogram depicting closure duration.

It has been established by several studies (Lisker, 1957; Gupta, Agrawal and Ahmad, 1973; Savithi, 1989; Vanden Berg, 1988 and Datta, 1989) that silence plays an important role in perception of the manner,voicing and place features of a stop consonant. Depending on the duration of the silence between the syllables, the inter-vocalic stops in trochees were perceived as voiced or voiceless (eg. rapid or rabid), (Lisker, 1957). Port (1976) gives an account of 'rapid' being perceived as 'ratted', when the duration of silence between the syllables was reduced.

FILE:a:apa.230 STARTING DURATION: 0 ENDING DURATION: 290

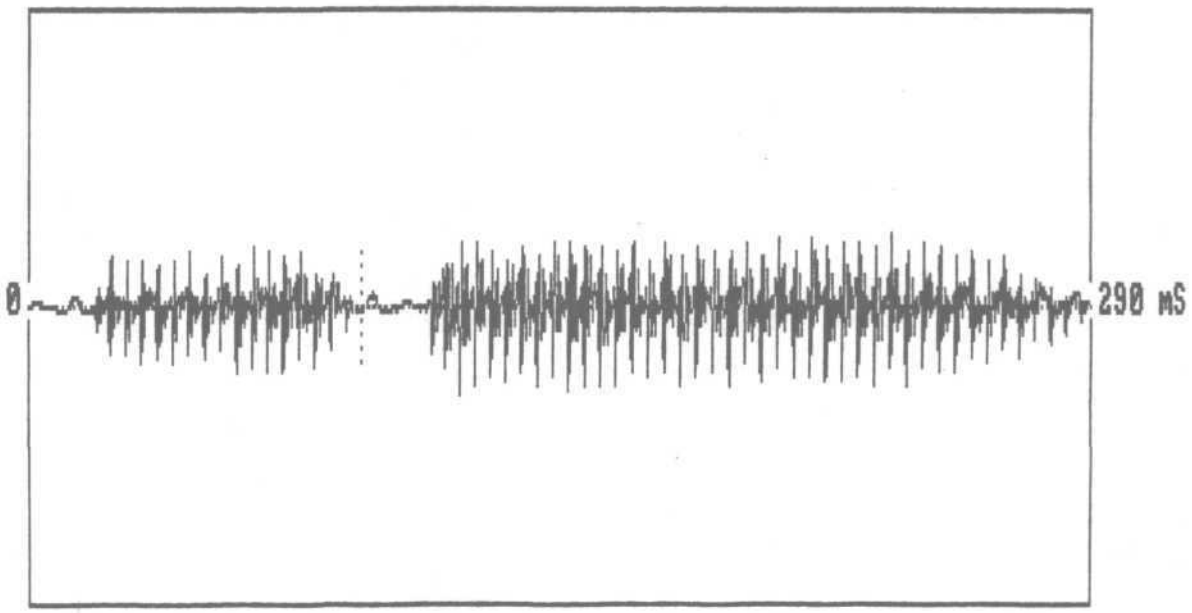

Fig.4.2a: : SYNTHESIZED /APPA/ WITH -230 M.Secs CLOSURE DURATION.

28

FILE:a:apa.10 STARTING DURATION: 0 ENDING DURATION: 560

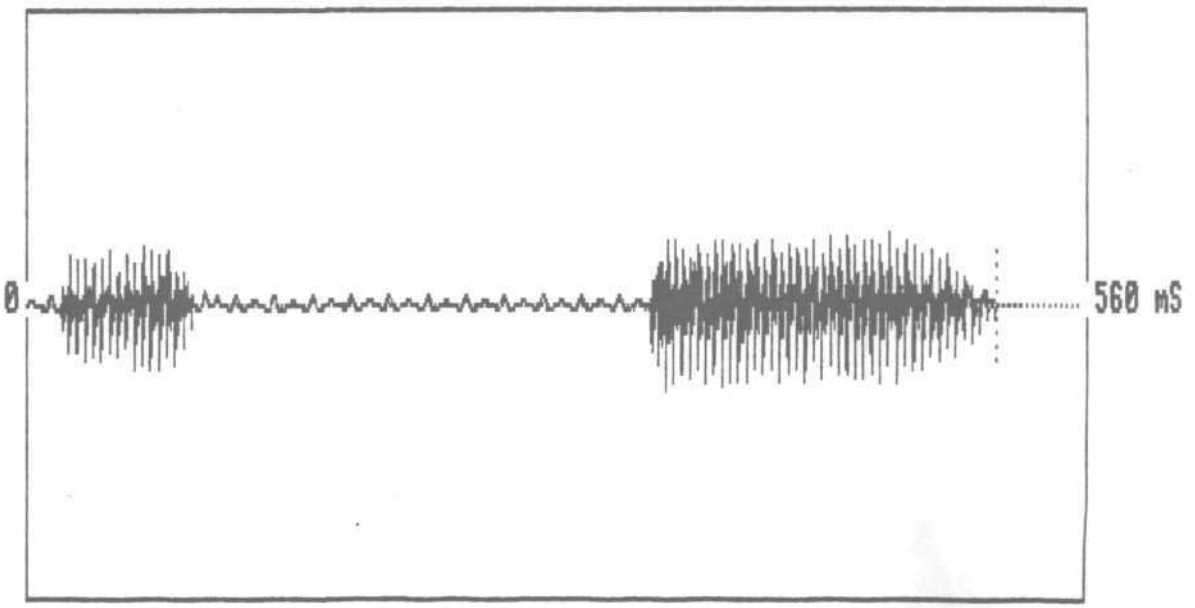

Fig.4.2b: SYNTHESIZED /APPA/ WITH -10 M.Secs CLOSURE DURATION.

Fig.4.2(a,b) waveforms of some synthetic stimuli.

27

Spectrograms of English medial voiceless stops before unstressed syllables showed longer closure durations than do the voiced stops (Lisker, 1957). However, Lisker (1978) contradicted this view.

Usha Rani (1989) studied role of closure duration on the perception of medial geminate stop consonants (kk, gg, pp, bb) in meaningful Kannada words to assess the perception of:

a) Cluster vs non-cluster and

b) Voicing feature (voiced vs voiceless).

Differences in perception between sex and a cross-languages (Hindi and Kannada) were assessed.

It was reported that closure duration provides the cue for manner and place of articulation, though to some extent. The present experiment aims to assess the perception of the above features in native Telugu and Malayalam speakers. Since the words /agga/ and /abba/ are meaningless for both Telugu and Malayalam speakers a fifth variable is added to the above mentioned four i.e. difference in perception between nonsense and meaningful words and cross-language differences between Telugu, Malayalam speakers.

Review:

The various experiments on closure duration as a cue to the perception of voicing of stop consonants are enlisted in Table 4.1

Table: 4.1

Review on closure duration as a cue to stop perception.

Author 1.	Material used 2.	Method employed 3.	Language 4.	Nature of task 5.	Durational parameters. 6.	Results 7.
Dorman Raphael & Liberman (1979-a)	Synthetic syllables / PE/ & / KE/	Analysis-by-synthesis	English	Forced choice Identification g/ PE/, / KE/ and / E/	Silent intervals between the —noise and /bE/ and /gE/ with their burst removed ranging from 20-100 ms in 20 ms steps.	l.When the silent interval was less than 20 ms. listeners reported no stop consonant and thus only / E/ response. 2.Stop heard when the silent interval exceeded about 40 ms.
Dorman Raphael & Liberman (1979-b)	Synthetic 'split'	Analysis-by-gynthesis	English	Forced choice percept of 'slit' 'split' or 's' followed by 'lit'	Appended 'S' noise and 'lit', with varying intervals of silence separating the two segments, ranging from 0-100 ms. in 15 ms steps and from 100 to 650 ms in steps of 50ms.	1.Silent intervals of less than 60 ms perceived as 'slit'. 2.At longer intervals out to about 450 ms,'split' perceived. 3.At the longest intervals of silence, stop not heard. Perceived as 's' silent-lit'.
Fischer-Jorgonsen (1979)	Synthetic CV syllables /p,t,k/ and /b,d,g/	Synthesis-by-rule	Danish	Forced choice percept of the syllables.	—	Closure duration increased from /g/ to /b/ percepts.

1	2	3	4	5	6	7
Lisker & Price (1979)	Synthetic 'sender' and 'center'	Analysis-by-synthesis	English	Forced choice percept of /t/ and /d/	/t/-nasalised flap /d/-sequence of nasal+stop closure duration shortened	Shortening the oral voiced closure of 'sender' yields a voiceless percept of 'center' when closure duration is reduced to 40 ms or less.
Price & Lisker (1979)		Analysis-by-synthesis	English	Forced choice percept of /P,b/	a. Shortened closure duration of /p/ b. Lengthened closure duration of /b/.	a. Relatively little effect b. Produced distinctive shift in labelling.
Bailey & Summerfield (1980)	Six synthetic consonant-vowel syllables of the type /s/ vowel and /s/ stop-vowel	Analysis-by-Synthesis	English	Forced choice percept of /s/ vowel, or /s/ stop vowel	Total duration of cv syllables was 300 m.sec. Duration of silence between /s/ and the vowel ranged from 0-90 ms in 10ms steps.	The closure durations corresponding to cross-overs between /s/ vowel and /s/ stop-vowel were:
Fitch Halves, Erickson & Liberman (1980)	Synthetic 'slit' and 'split'	Synthesis-by-rule	English	Forced choice identification of 'slit' and 'split'	Silence interval varying from 8-160ms in 8 ms steps.	1. Silent duration effective in contrasting 'slit' from 'split' 2. As silent duration increased, judgements shift from 'slit' to 'split'. 3. Smallest shift was 8 ms and the highest being 40 ms.

1.	2.	3.	4.	5.	6.	7.
Port (1980)	Synthetic /dipper/ and /dibber/		English	Forced choice by 16 listeners	Five durations of dib(140-260 ms) and 9 medial stop closure durations(20-140ms)	No significant effect.
Raphael (1981)	Natural speech /peg/ excised from two different contexts		English	Forced choice identification of peg or peck		1.Generally sufficient cue-closure voicing. 2.Idiosyncratically sufficient cue-vowel duration.
Datta (1989)	Natural speech /stri/ and /ri/	Electronic gating.	Bengali		1.Silence between /a/ & /tri/ removed. 2.Silence of about 30 ms introduced between / / and /ri/	1.Perceived as /Sri/ 2.Perceived as / kri/
Usharani (1999)	Synthetic /akka/,/agga/ /appa/ and /abba/	Analysis-by-synthetis	Kannada & Hindi	Forced choiced	Closure duration was truncated from 0-120 ms.	As the closure duration decreased percept changed from cluster to non-cluster and voiceless to voiced.

Method:

Stimulus:

The stimuli prepared by Usha Rani (1989) were used in this experiment. The temporal parameters for each word are in Table 4.2.

	akka	agga	appa	abba
Total duration	615	432	535	520
Closure duration	240	215	238	214.
Preceding vowel duration	127	143	89	97

Table 4.2 Temporal parameters of the stimuli synthesized (in m.secs).

(The syllable /akka/, /agga/, /appa/, and /abba/ were recorded as uttered by a female speaker, digitized and stored in a computer memory. The initial vowels and CCV of these words were extracted and stored separately. In the CCV syllable /kka/, /gga/, /ppa/ and /bba/, the closure duration was truncated in 10 m.secs steps to produce stimuli with reduced closure durations. The truncation was performed until the closure duration was almost removed (less than 10 m.secs). The resultant CCV stimuli were then recombined with the initial vowels of the corresponding syllables. Each word had the following number of stimuli: akka-23, appa-24, agga-22, end abba-23. These stimuli were randomized.

converted into analog signal and audiorecorded. Each stimulus was recorded thrice with an inter stimulus interval of 2 seconds)

The concatenated VCCV waveforms are in Fig. 4.1.

Results and Discussion:

The results are presented in Table 4.3.

1. In general, three percepts were identified for the voiceless consonants and two percepts were identified for the voiced consonants both by Malayalam and Telugu listeners.
2. At reduced closure durations the stimulus was perceived as voiced and at lengthened closure duration it was perceived as voiceless. As the closure duration increased the percept changed from non-cluster voiced to non-cluster voiceless and finally to cluster voiceless for the voiceless consonants, and for the voiced consonant it changed from non-cluster to cluster.
3. In Malayalam listeners the change in percept was not continuous. For example, inthe percept /kk/ the scores were scattered at 110 and 120 m.secs of closure duration and hence were not included into any category. In case of the percepts /gg/ and /bb/ break was observed at closure duration of 90-100 m.secs and 100-120 m.secs respectively. In case of /pp/ there was an overlap at 20 m.secs of closure duration. In Telugu listeners, the scatter of score (break) was noticed at a closure duration of 120 m.secs for the percept /pp/.

Percept identified	Closure duration (ms)		% identification				Percept identified	Closure duration (ms)		% identification			
VL Velar /kk/	Malayalam	Telugu	M.M.	M.F.	T.M.	T.F.	VL Bilabial /PP/	Malayalam	Telugu	M.M.	M.F.	T.M.	T.F.
Aga	0-50	0-30	63.33	70.0	53.56	91.66	Aba	0-20	0-30	60-00	53.33	60-71	66.66
Aka	60-100	40-100	72.0	68.0	69.38	80.94	Apa	20-120	40-110 120	56.36	81.81	64.28	87.49
Akka	^ 110,120 130-210	110-210	62.22	88.88	81.81	78.78	Appa	130-230	130-230	61.81	92.72	79.21	93.93
VD Velar /gg/							VD Bilabial /bb/						
Aga	0-80	0-90	77-77	93.33	71.42	96.66	Aba	0-90	0-90	72.0	82.0	67.13	100.0
Agga	H 90,000 110-210	100-210	92.72	98.18	83.33	97.22	Abba	H 100-120 130-210	100-210	88.88	80.0	66.66	88.88

Table 4.3: Percentage of different percepts identified by Native Telugu and Malayalam speakers for closure duration (No. indicate the duration of closure of the stop in m.secs.)

MM - Malayalam Males H - break
MF - Malayalam Females S - overlap
TM - Telugu Males
TF - Telugu Females.

4. There were no clearcut differences in perception between sex and across languages (Malayalam and Telugu).
5. Meaningfulness of the stimulus did not affect the percept.
6. There was an overlap in the perception of voicing between the voiceless and the voiced percept,eg. when the percept For was /kk/, the percept changed from voiced to voiceless at a closure duration of 60 and 40 m.secs respectively for Malayalam and Telugu listeners. However, when the percept was /gg/, Malayalam and Telugu listeners perceived it as voiced singleton till a closure duration of 80 and 90 m.secs respectively. Similarly in case of /pp/, the percept changed from voiced to voiceless at a closure duration of 20 and 40 m.secs respectively for Malayalam and Telugu listeners. Again, when the stimulus was /bb/, both Malayalam and Telugu listeners perceived it as voiced singleton /b/ until a closure duration of 90 m.secs. This might be because of the presence of vocal fold vibration cueing voicing.

The results of this study are in consonance with the findings (production data) of Lisker (1957) in English, Nagamma Reddy (1985) in Telugu and Savithri (1989) in Kannada. These studies report that the closure duration for voiceless sounds is longer than that of voiced. The perceptual data of the present study supports this. Also, the results support

the findings of the earlier studies (Dorman, Raphael, Liberman, 1979, a, b; Datta, 1989 etc) in that a voiceless stop in perceived at increased closure durations.

However, they do not support the findings of closure duration as a cue for place of articulation (Fischer-Jorgensen, 1979). This is expected as a forced choice percept was provided.

The responses of Malayalam and Telugu listeners differed from those of Kannada and Hindi who participated in the experiment conducted by Usha Rani (1989). In Kannada and Hindi listeners, the percept changed from non-cluster to cluster when the closure duration was greater than 60-70 m.secs whereas in Malayalam and Telugu listeners this change was observed at around 100 m.secs of closure duration. Kannada and Hindi listeners reported voiced percepts for closure durations of less than 60 m.secs. Whereas Malayalam and Telugu listeners reported voiced percept for closure durations of less than 30 m.secs (except Malayalam males who reported voicing till 50 m.secs). There was no difference for the percepts /gg/ and /bb/ across languages. Scatter of scores or breaks were also reported in Kannada and Hindi listeners around 100-110 m.secs of closure duration which was treated as "perceptual confusion*. The production data and the perceptual responses of the four different languages are in table 4.4.

Language		Closure duration in production	Closure duration in perception
Telugu	/g/	105	Less than 30
	/b/	100	Less than 30
Malayalam	/g/	NIL	Less than 50
	/b/	NIL	Less than 20
Kaanada	/g/	72	Less than 60
	/b/	106	Less than 60
Hindi	/g/	100	Less than 60
	/b/	90	Less than 60

Table 4.4 Closure duration (m.secs) in production and perception.

Of the four languages the production and perceptual data seem to correlate better in Kannada language. This might be perhaps because (1) the speaker had Kannada as her first language and (2) all the words used were meaningful to the listeners of Kannada language.

The three phonetic dimensions voicing, aspiration and articulatory force have been widely cited as bases for separating the stop categories of English and many other languages (Lisker and Abramson, 1967). In the present study all the four sounds are unaspirated and still it was

observed that the perception changes from Voiceless to voiced as the closure duration was decreased. As the articulatory force was essentially constant for all the synthetic words the sound is perhaps perceived as voiceless if the articulators are closed for a longer time. Thus, it seems that the closure duration becomes more important than the articulatory resistence.

5. EXPERIMENT-II

PRECEDING VOWEL DURATION AND STOP PERCEPTION

Introduction:

The duration of the preceding vowel is often cited as an important cue to the voicing feature of final stop consonants in English. Preceding vowel duration has been called under certain conditions as a primary (Klatt, 1976) and even necessary cue (Raphael, 1972) to voicing distinctions.

"In English and some other languages vowels preceding final voiced consonants are longer than those preceding final voiceless consonants and this difference is perceptually relevant" (Denes, 1955; Raphael, 1972). Irrespective of the consonants following a stressed vowel or preceding an unstressed vowel, the length of vowel cues voicing of the consonant. The vowel is shortened preceding a voiceless consonant (Denes, 1955; House, 1961).

However, the phonetic categories that contrast voiced vs voiceless are not the same across languages (Goldstein and Brownman, 1986). Vowels in both French and English are longer before the phonologically voiced stops than before the phonologically unvoiced stops (Mack, 1982). Vowel length

difference for voicing contrasts in Dutch (Slis and Cohen, 1969; Vanden Berg, 1988), Korean (Chen, 1970) and Swedish (Lindblom and Repp, 1973) are also noted.

In Indian languages, preceding vowel duration was not found to cue for the voiced-voiceless distinction and place of articulation in Kannada and Hindi (Usha Rani, 1989). Though, the production data in Kannada (Savithri, 1989) and Telugu (Nagamma Reddy, 1985) indicate lengthening of the vowel preceding voiced stop, it was not significant.

This experiment aims to evaluate the effect of preceding vowel duration in cueing the perception of /kk, gg, pp and bb) for voicing, Difference between sex and across listeners of Malayalam and Telugu languages are also studied.

Review:

Most of the experiments on the cueing value of preceding vowel duration, employing the patterns of synthetic speech have used only the steady-state portion as the variable. However, the syllable-initial transitions in CV syllables do contribute to the 'vowel' duration, to almost same degree as an equal duration of steady state (Raphael, Dorman and Liberman, 1980). Thus, the effective duration of a vowel extends over all parts of the acoustic signal that may be said to have influenced by it, including especially, the transition that

Table:5.1

Review on preceding vowel duration as a cue to stop perception.

Author 1.	Material used 2.	Method employed 3.	Language 4.	Nature of task 5.	Durational parameters. 6.	Results 7.
Raphael (1972)	Synthetic word final clusters of stop+stop. stop+fricative & fricative+ stop types i.e. /p-b/-gape-gabe. /t-d/-bet-bed. /k-g/-Burke-Burg /pt-bd/-bopped-bobbed,/st-zd/-cost-caused.	Pattern playback	English	Forced choice format	Steady state vowel varying from 150-350 ms for voiced.	Longer preceding vowel duration gave a voiced percept & shorter durations gave voiceless percepts.
Wardrip-fruin and Bischoff (1976)	Naturally produced speech parts of vowel. removed from initial,medial & final portions of the vowel.	Naturally produced speech	English	Forced choice percept of consonant voicing.	Vowel duration reduced in 20 ms steps.	No r elation seem between the preceding vowel duration and consonant voicing.
Raphael (1980)	/b d/-/ d/ /b t/-/ t/	Synthesis	English	-	-	d-t phonemic boundaries plotted as a function of vocalic duration. fell within 3 ms of each other, indicating the effect of vowel duration.
Wardrip-Fruin (1982)	Natural speech	Analysis-by-Synthesis	American English	Forced choice	Preceding vowel duration	In natural speech preceding vowel duration is not a cue for voicing of final stops. Voicing during closure may be required to cue for voiced final stops.
Usharani (1989)	Synthetic /apga//abba/	Analysis-by-synthesis	Kannada and Hindi	Forced choice	Preceding vowel duration was truncated from 0 to 50 ms in 10 ms steps.	No change in percept was observed.

reflect the consequences of the coarticulations of vowel and consonant (Raphael, Dorman and Liberman, 1980). The studies of relevance to the preceding vowel duration are in Table 5.1.

Method:

Stimulus;

The stimuli prepared by Usha Rani (1989) were used in this experiment. (Two stimulus words /agga/ and /abba/ varying in the duration of the initial vowel were synthesized as follows. A female speaker recorded the words /agga/ and /abba/ (VCCV), which were stored in a computer memory. The duration of the initial vowel in the steady-state was truncated in 10 m.secs steps until the vowel was reduced to half its duration and these stimuli were stored in the computer memory to generate VCCV stimuli which differed in duration of the preceding vowel (Fig. 5+1 a and 5.1 b).

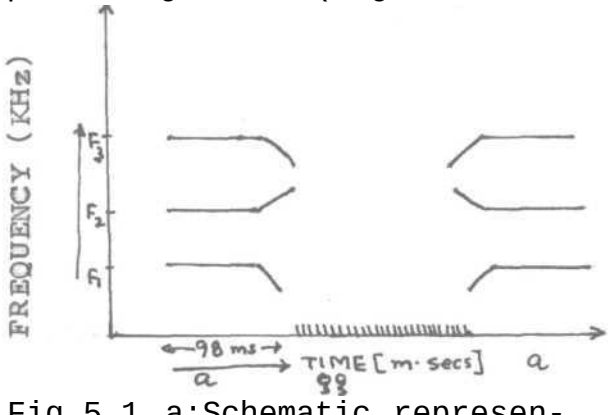

Fig.5.1 a:Schematic representation of spectrogram depicting /agga/ before truncating the duration of preceding vowel.

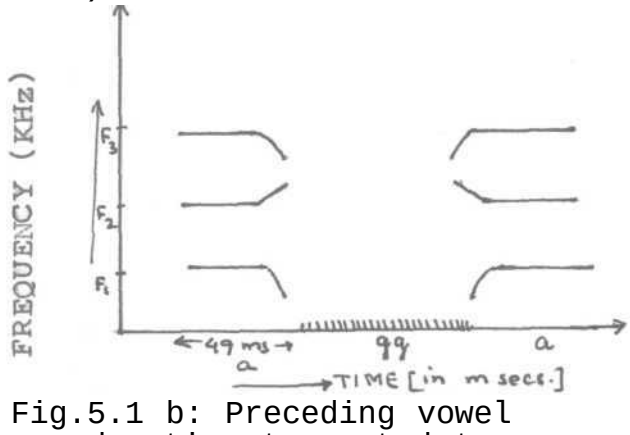

Fig.5.1 b: Preceding vowel duration trancated to half its value of /agga/.

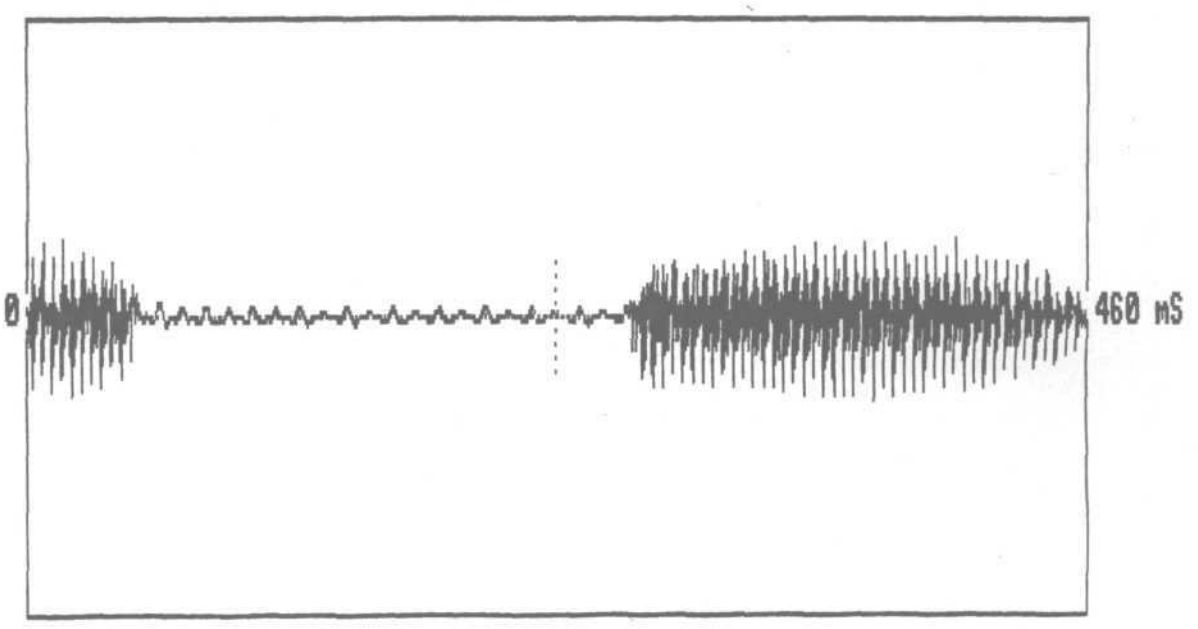

Fig.5.2a:= SYNTHESIZED /ABBA/ WITH -50 M.Secs OF PRECEDING VOWEL DURATION.

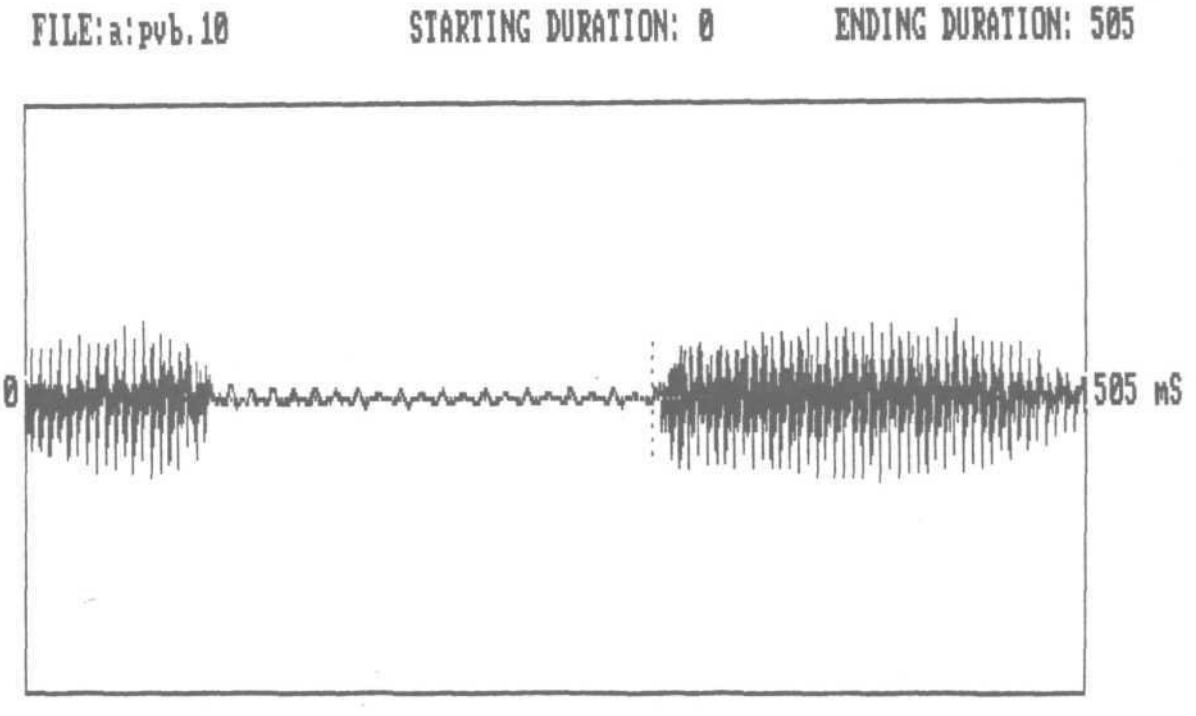

Fig.5.2b:— SYNTHESIZED /ABBA/ WITH -10 M.Secs OF PRECEDING VOWEL DURATION.

Fig.5.2(a,b) Waveforms of some synthetic stimuli.

The stimuli was randomized, converted to analog signals and audio recorded. Each stimulus was audio recorded thrice with an interstimulus interval of 2 seconds). The durational parameters and the stimuli generated are in the table 5.2.and the concatenated wave forms are in Fig 5.2(a) and

Words	abba	agga
Duration of initial vowel (in ms)	97.4	143
Ending of truncation (in ms).	48.7	70
Number of stimuli obtained	5	7

Table 5.2 Durational parameters of the words /agga/ and /abba/

Results and Discussion:

Results are presented in Table 5.3.

In general, there was no change in the percept with decrease in preceding vowel duration. Also, there were no apparent differences in percept for meaningful/meaningless stimuli. However for the stimulus /bb/, Telugu male listeners reported a voiceless percept as the preceding vowel duration was decreased. This agrees with the findings

of Smith and Muse (1988), Kluender, Dichl and Wright (1988), Vanden Berg (1988), Raphael (1972) and Denes (1955).

Identified Percept	Malayalam		Telugu	
	Male	Female	Male	Female
agga:				
PVTD	0 - 70	0 - 70	0 - 70	0 - 70
Percentage Identification.	85.71	94.23	73.76	100
abba:				
PVTD	10 - 50	10 - 50	10 - 20	10 - 50
Percentage Identification	68	68	71.43	93.33
appa:				
PVTD			30 - 50	
Percentage Identification.			61.9%	

Table 5.3: Perceptual judgement of Malayalam and Telugu listeners as a function of preceding vowel duration.

PVTD indicates the preceding vowel truncated (steady-state) in m.secs.

It seems that preceding vowel duration may not cue for voicing in Malayalam for velar and bilabial stops. However, it seems to cue voicing in Telugu bilabial stops for males. Thus, there was a sex difference among Telugu listeners and

a language difference for /bb/ percept. The data of this study do not support the findings of the other studies with English listeners, but they are similar to the production data of Telugu speakers (Nagamma Reddy, 1985). Nagamma Reddy (1985) states that in Telugu, there exists no apparent difference between the duration of the vowels preceding voiced and voiceless consonants and that "the maximum difference found in vowels before voiced and voiceless plosive is about 20 m.secs. In Kannada (Savithri, 1986) the average duration of vowels preceding voiced and voiceless are 86,86 and 66,72 respectively for velars and bilabials with no significant difference. Usha Rani (1989) reported no change in percept as a function of preceding vowel duration in Kannada and Hindi. Kluender et al.(1988) used the principle of durational contrast to explain the compensatory action for syllables - a long vowel should make a short closure interval seem even shorter and hence more voiced and a short vowel should make a long closure interval seem even longer and hence more voiceless. This compensatory action for syllables is not supported by then perceptual data in Malayalam, Kannada or Hindi.

6. EXPERIMENT-III

TRANSITION DURATION OF THE PRECEDING VOWEL AND STOP PERCEPTION

Introduction:

Transitions are defined as the movements of the formants which are due to the movement of articulators from one phoneme to another. Delattre, Liberman and Cooper (1955); Liberman (1957); LaRiviere, Winitz and Harriman (1975) and Wolf (1976) in their experiments, altered the transition of vowel formants and the results proved the effective role of transitions in cueing the place of articulation. Harris (1958) and Dorman and Raphael (1980) also reported the importance of transition durations in the perception of various consonantal sounds. Walsh, Parker and Miller (1987) reported that for any steady-state duration, the more rapid the F1/1 decline state, the more likely the listener is to perceive it as voiced. The rate of F1/1 decline systematically effects the perception of (± voice) in a following stop. Any change in the Fl configuration has cueing value, but the cueing strength of any rate of Fl decline or steady-state duration is dependent on the value of other cues also. Alwan (1989) found that the onset value of F1/1 (Fl_o) is an important cue in discriminating between the two sounds when F2/2 onset value is held constant at a value appropriate for either pharyngeal or uvular.

There have been some studies in Indian languages. Gupta, Agrawal and Ahmad (1973) suggested that the transition of initial and final vowels terminating into the consonant are of maximal importance for the recognition of intervocalic consonants. Tyagi, Agrawal and Pavate (1987) reported that the rate of the formants 1 and 2 follows in the decreasing order - bilabials, dentals, retroflexes and velars.

Usha Rani (1989) reported no change in the percept of bilabial and velar consonants as a function of preceding vowel transition duration in Kannada and Hindi listeners.

This experiment aims at determining the role of the transition duration of the preceding vowel on the perception on intervocalic consonants /kk, gg, pp, bb/ for voicing percept differences between sex and language (Malayalam and Telugu) and meaningful/nonsense stimuli.

Review:

The studies on transition duration of the preceding vowel as a cue to the perception of consonants are in Table 6.1.

Method:

Stimulus:

The stimuli prepared by Usha Rani (1989) were used in this experiment. (The words /akka/, /agga/, /appa/ and

Table: 6.1

Review on transition duration of preceding vowel as a cue to stop perception.

Author 1.	Material used. 2.	Method employed	Language 4.	Nature of task 5.	Duration parameters. 6.	Results. 7.	Discussion 8
Keating and Blumstein (1978)	Synthetic /da-ga/	-	-	Forced choice percept by 20 untrained listeners.	Lengthened transitions (45,95,145 ms)	1.Lengthened transitions minimally affect the perception of stop consonants. 2.The identification and discrimination scores did not vary systematically with increase in the length of transition duration.	
Haskins group (1980)	Synthetic and/g /		English	Forced choice percept.	Second formant transitions varied systematically.	1.Short duration F2 transitions perceived as stops /b/ and /b/. 2.Longer transition durations (40-50ms) perceived as / / and /j/. 3.Transitions of 150-200ms as /u / and /i /.	Formant transition durations cue the manner of articulation.
Raphael (1980)	Spoken syllable /s d/ edited into 3 continua: /s d/-/s t/ /d d/-/d t/ / t/-/ t/	Deletion of pitch pulses, noise & initial transitions in that order for the 3 contlnua.	English	Identification of final consonants as either voiced or voiceless.	-	/d/-/t/ phoneme boundary,plotted as a function of vocalic duration, fell within 3 ms of each other.	The initial CV transitions contribute equally with steady state formants to cue the final consonant voicing
Walsh et.al (1987)	Synthetic 'bad' & 'bat'	Analysis-by-synthesis		Forced choice envircling bad or bat.	1.Steady state duration(100 150,200,250 ms). 2.F1 decline rate(0,3.7, 6.7 & 9.7 Hz/ms)	l.For any steady-state duration rapid Fl decline rate was perceived as voiced. 2.For any rate of Fl decline,greater steady state duration was perceived as voiced. 3.Cueing strength of Fl decline rate and steady-state duration is dependent on other cues.	
Usharani (1989)	Synthetic /akka/ /agga//appa/ and /abba/	Analysis-by-synthesis	Kannada and Hindi	Forced choice	Preceeding vowel transition duration was truncated in 10ms. steps 0-30, FC for /appa & akka/ 0-50,FC for /abba/& /agga/	No change in percept was seen with decrease in preceding vowel duration.	

/abba/ were synthesized with varying preceding vowel transition duration. The four words were uttered by a female speaker, recorded, digitized and stored in a computer memory. The CCV syllables and the initial vowels in the words were extracted and stored separately. The initial vowels were analyzed to obtain their spectral and temporal characteristics. A parametric synthesis was performed by referring to the values obtained from analysis of the same and the vowel /a/ was generated. The parameters used were:

/a/ of /appa/:

Duration (m.secs)	0-50	50-80
F_1(Hz)	925	700
F_2(Hz)	1400	1367
F_3(Hz)	2646	2625

Duration (m.secs)	0-10	20-70	-80
F_oHz)	183	- 235	230
Energy	267	400	317

/a/ of /abba/

Duration (m.secs)	0-40	50-100
F_1 (Hz)	967	667
F_2(Hz)	1383	1283
F_3(Hz)	2688	2625

Duration (m.secs)	0-20	20-100
F_o(Hz)	223	233

Duration (m.secs)	0-20	30-80	50-100
Energy	367	483	358

/a/ of /akka/:

Duration (m.secs)	0-80	80-120	
F_1(Hz)	1038	812	
F_2(Hz)	1433	1900	
F_3(Hz)	3313	2438	
Duration (m.secs)	0-50	60-120	
F_o(Hz)	183	235	
Duration (m.secs)	0-40	50-110	110-120
Energy	158	383	300

/a/ of /agga/:

Duration (m.secs)	0-80	80-140
F_1(Hz)	1000	642
F_2 (Hz)	1550	2050
F_3 (Hz)	3167	2750
Duration (m.secs)	0-10	20-140
F_o (Hz)	183	235

Duration (m.secs)	0-40	40-50	50-110	110-140
Energy	300	350	400	206

The synthetic vowels were used to create. /Stimulus /a/ varying in their transition durations, by truncating, the transiting portion of F1/1+ F2/2 and F3/3 simultaneously, in 10 m.secs steps, keeping the terminal F_1, F_2 and F_3 constant. Stimuli were also synthesized with F_1 cutback (Fig. 6.1(a) and 6.1 (b)

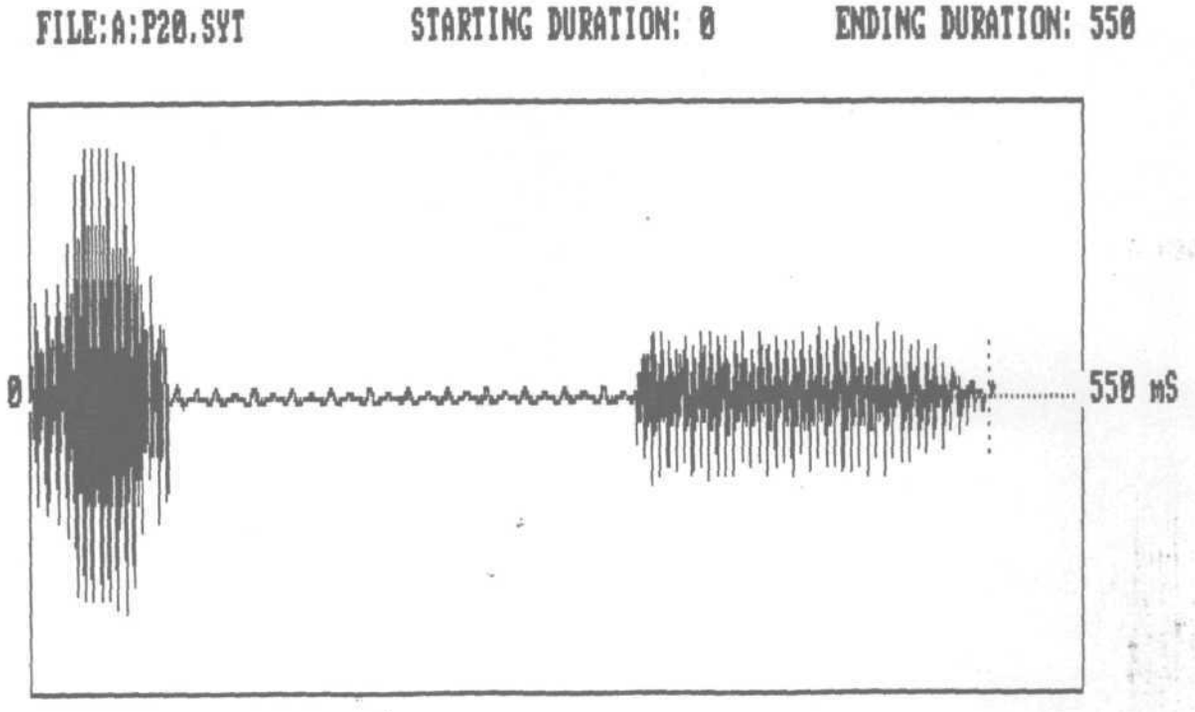

Fig.6.2a: : SYNTHESIZED /APPA/ WITH -20 M·Secs OF TRANSITION DURATION (PRECEDING VOWEL).

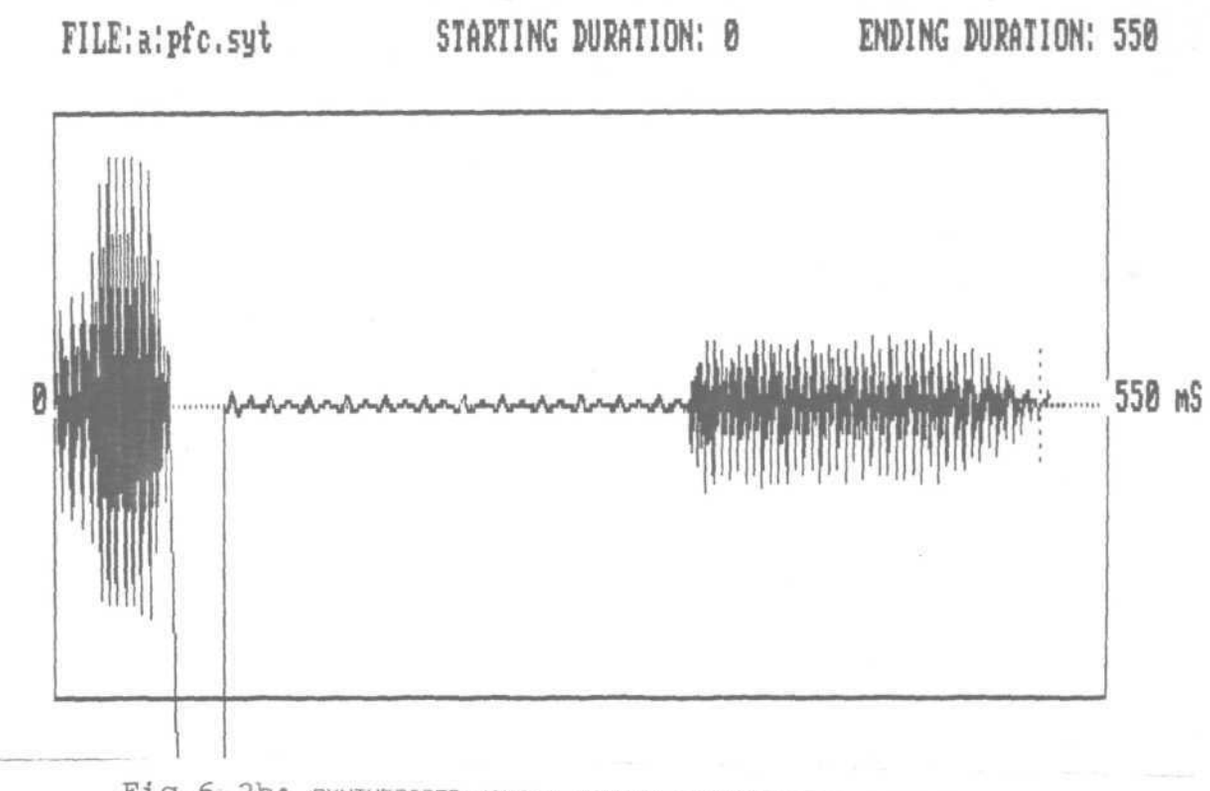

Fig.6.2b: SYNTHESIZED /APPA/ WITH F1 CUTBACK OF THE PRECEDING VOWEL.

Fig.6.2 (a,b) waveform of some synthetic stimuli.

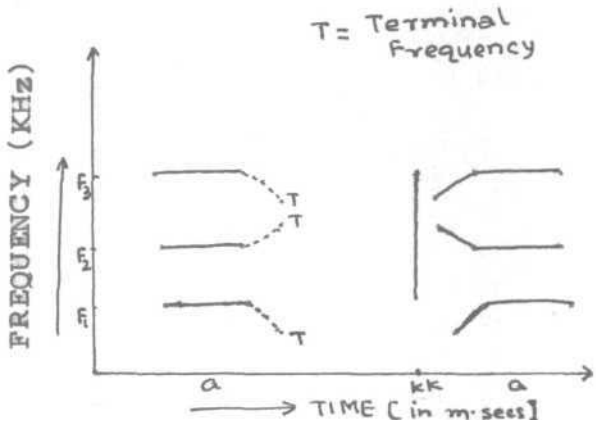

Fig.6.1(a):Schematic diagram of truncation of F1, F2, and F3 transition of preceding vowel in /akka/.

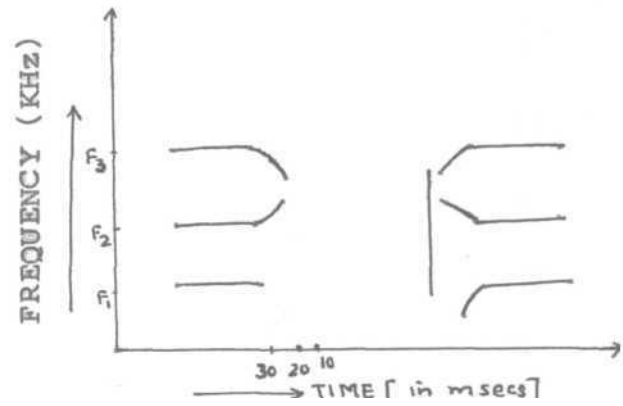

Fig.6.1(b): Schematic diagram of Fl cutback /akka/.

The synthesized vowels were then concatenated with their respective CCV syllables of /kka/, /gga/, /ppa/, and /bba/ to produce the complete stimulus word. The number of stimuli synthesized were: /akka/-5; /agga/-7; /appa/4; /abba/-7. Thesedigitized stimuli were randomized,converted to analog stimuli and were audiorecorded. Each stimulus was recorded thrice with an interval of 2 sees, between stimuli). Fig.6.2 shows the waveforms of a few stimuli.

Results and Discussion:

There was no change in the percept of /gg/, /kk/ and /pp/ as the duration of the preceding vowel transition decreased. However, stimuli /bb/ was perceived as /pp/ by Malayalam male listeners when the transition of the preceding vowel was truncated by 50 m.secs and it was perceived

Percept identified	Preceding vowel transition duration truncated (in ms)				% Identification			
	Malayalam				Telugu	M.M	M.F.	T.M.	T.F.
	Male (MM)	Female (MF)	Male (TM)	Female (TF)					
Abba	F_1C,0-30	F_1C.00	F_1C,00	F_1C,0-50	84	60	57.14		66
Appa	H 40 50	H 10-50	10-30 40	-	100	72	85.71		
Agga	F_1C,0-50	F_1C,0-50	F_1C,0-40	F_1C,0-50	77.14	68.57	61.99		85.71
Akka	F_1C,0-30	F_1C,0-30	F_1C,0-30	F_1C,0-30	76	80	71.42		100
Appa	0-20	F_1C,0-20	0-20	F_1C,0-20	93.33	95	71.42		91.66

Table: 6.2 - Percentage of different percept identified by native Telugu and Malayalam speakers for preceding vowel transition duration: F_1C - F_1 cutback: 'H' - break

as /pp/ at truncation duration of 10-50 m.secs by Malayalam females and 40 m.secs by Telugu males. Telugu females did not report any change in percept. The change in percept from voiced to voiceless by some of these listeners was not continuous. F_1 cutback had no effect on perception. Thus the preceding vowel transition duration did not seem to consistently cue voicing and sex and language differences were not apparent. Also, the meaningfulness of the stimuli did not seem to play any role.

The results of this experiment are similar to the findings of Usha Rani (1989) and Keating and Blumstein (1978) in that shortening the transition duration had no effect on the percept. Cooper, Delattre, Liberman, Borst and Gerstman (1952) pointed to a rising first formant as a cue to voicing and Liberman, Delattre and Cooper (1958) singled out first formant 'cut back' as a formidable cue to the English voiceless stops, but this was not supported by the results of the present experiment. However, the preceding vowel transition duration might cue for voicing in combination with other cues like preceding vowel duration (Walsh et al. 1987), VOT (Stevens and Klatt, 1974), release burst and closure duration (Gupta, Agrawal and Ahmad, 1973). This needs to be experimented further.

7. EXPERIMENT-IV

TRANSITION DURATION OF THE FOLLOWING VOWEL AND STOP PERCEPTION

Introduction:

Transition for a following vowel occurs, when a stop consonant occlusion is released, and the vocal tract shape associated with the particular formant resonance, changes towards the following vowel (Delattre, Liberman, and Cooper, 1955). Research with synthetic speech has revealed that both bursts and voiced formant transitions may serve as separate cues to place of articulation of initial /b,d,g/. Many studies have shown that transitions of the second and third formants are sufficient cues for the place distinction (Delattre, Liberman and Cooper, 1955; Liberman, Delattre, Cooper and Gerstman, 1954). Dorman, Studdert-Kennedy and Raphael (1977) reported that bursts and transitions are equivalent and complementary. Cooper, Delattre, Liberman, Borst and Gerstman (1952) had also suggested that bursts and transitions complement each other in the sense that when one cue is weak, the other is usually strong. The two are equivalent but not alterative (Dorman, Studdert-Kennedy and Raphael, 1977).

This experiment aims to determine the role of the transition duration of the following vowel on the perception

of atop consonants for voicing and place. Difference across sex and languages (Malayalam and Telugu) and meaningful, meaningless stimuli were also tested.

Review:

Table 7.1 briefly presents the review in various languages.

Method:

Stimulus:

The stimuli prepared by Usha Rani (1989) were used ia this experiment. (Four stimuli varying in the transition duration of the vowel following the stop consonants for the four words /akka/, /agga/, Zappa/, and /abba/ were synthesized. The words /akka/, /agga/, /appa/ and /abba/ were uttered by a female speaker, recorded, digitized and stored in computer memory. VOC and V following the consonant were extracted from the four words and stored separately. These final vowels were acoustically analysed to obtain their spectral and temporal measures. A parametric synthesis of the vowel /a/ was performed on the basis of the values obtained by analysis. The spectral and temporal characteristics for synthesis of /a/ are below;

/a/ of /akka/:

Duration (m.secs)	0-60	60-170
F_1(Hz)	775-900	900-975
F_2(Hz)	1600-1583	1583-1550
F_3(Hz)	2542-2646	2646-2521

Table:7.1
Review on the transition duration of the following vowel

Author	Material used	Method employed	Language	Nature of work	Durational parameters.	Results	Discussion
1.	2.	3.	4.	5.	6.	7.	8.
Stevens & Klatt (1974)	Synthetic aspirated vs. unaspirated stop consonants. Synthetic consonant vowel stimuli.		English	Forced choice percept of /da/vs/ta/	4 formant trajectories from moderate to rapid consonant-vowel transition times producing 16 stimuli. VOT ranged simultaneously from 0-40 ms in 5ms steps.	l. VOT less than 20 ms + rapid transition lead to voiced consonant percept. 2. VOT greater than 20 ms + rapid transition with onset before voicing lead to voiceless percept. 3. VOT greater than 20 ms + rapid transition onset after voicing lead to voiced percept; and 4. VOT greater than 20 ms + moderate transitions lead to voiceless percept.	
Lisker et al (1975)	Synthetic /ba-p^ha/ /da-t^ha/ & /ga-k^ha/	—	English	Forced choice percept	Identifical Fl transitions for 3 continua to test the phonetic boundary for the place of articulation.	Boundaries on the 3 continua coincided perfectly, showing a clear phonetic class.	
Dorman Studdert-Kennedy & Raphael (1977)	Natural speech (CVC) syllable	Masking laboratories pulse code modulation system.	American English	Forced choice.	Release burst & following vowel transition duration.	Burst and transition are acoustically and perceptually equivalent.	
Summerfield & Haggard (1977)	Exemplars from /p-k^h/ (CV) from a VOT continue of 0 to +80ms in 1ms steps.	Closed loop algorithm controlled stimulus.	English	Identification of initial consonant of each stimulus as /g/ or /k^h/	Fl transitions varying from 0-36 ms in 6 ms steps at 5 Hz/ms, after voicing onset, Fl onset constant at 250Hz.	Variation in F1 trasition duration has small effect on the perception of voicing	

1.	2.	3.	4.	5.	6.	7.	8.
Summerfield & Haggard (1977)	Synthetic CV syllable /k^h-g/ F1 onsets of 208,311 & 412 Hz combined with F1 transitions of 200.100 and OHz.	Closed loop algorithm controlled stimulus.	English	Forced choice percept of /k^h-g)	Decreasing Fl to zero and F2 excited by noise,paired with varying VOT.	1.Increased Fl transition from 0-100 or 200Hz decreased /g/ responses. 2.Increasing Fl transition from 100-200HZ had no systematic effect.	
Ahmad and Gupta (1930)	/p,p^h,t,t, t,t^h,d,g, g^h,k,k^h,b, b^h,& d,d^h/ with /a/.	Electronic gating.	Hindi	Perceptual judgement of an unrecognizable sound	Transition duration of syllable durations were varied in 10 ms steps.	1.Place identification of velars lost when 10 ms of the initial part of the vowel transition was removed. 2.Place feature identification of /b/ lost when approximately ¼ of the initial part of transition with the voice bar and burst were removed. 3./k/ gets affected & /p/ remain s unaffected when 20ms of terminating transition deleted. 4.20ms of terminating transition to be added for /b/ percept. 5.Addition of terminating transition enhance: identification of /p,t,k/.	Vowel transition and the steady state of vowel is enough to ldentify/p,t,k/
Bailey & Summerfield (1980)	Six CV syllables	Synthesis-by-rule.	English	Forced choice percept of 5-vowel and s-stop-vowel.	Transition duration were 35 ms. 3.values of Fl transition extent were OHz, 156Hz & 309Hz closure duration varied from 0-90 ms in 10 ms steps.	l.With increased Fl transition extent. crossovers between s-vowel & s-stop vowel occur at shorter intervals of silence.	l.For a stop percept, trading relationship between Fl transmission and closure duration seen.

1.	2.	3.	4.	5.	6.	7.	8.
						2.. Reduced magnitude of transition with a lowered Fl onset frequency. increased the percentage of /sk/ percepts. 3.. Longer, slower Fl transitions for velar percept in comparison with alveolars.	2. The duration of the stop closure is inversely related to the rate at which the oral constriction is released i.e. longer closures characterize bilabials than alveolars & velars. 3. For a given place. shorter closures precede open vowels /a/ than the more closed vowels /i/ and /u/
Alwan (1989)	Synthetic CV stimuli	Analysis-by-synthesis.	Arabic	Forced choice between voiced pharyngeal & voiced uvular consonants.	1. F1 and F2 were varied independently in 50Hz and 60Hz steps respectively. 2. Bandwidth of F2(pharyngeal) and Fl(uvular) was changed.	1. Onset value of Fl (Fl.) is a cue for place of articulation and F2 is not. 2. Increase in bandwidth of Fl increases the percentage of uvular responses and enhances naturalness of uvular stimuli. Increasing F2 band width had no effect	
Usharani (1989)	Synthetic /akka//agga/, /appa//abba/	Analysis-by-synthesis	Kannada and Hindi.	Forced choice	Transition duration of the follow-up vowel was truancated from 0 to 50 ms to 10 ms steps and Fl-cutback.	No change in percept was seen.	Transition duration of following vowel is not a cue for voicing or clustering.

Duration (m.secs)	0-10	10-50	60-170	150-170
F_0(Hz)	250	243	235	243

Duration (m.secs)	0-10	20-80	90-170
Energy	417	417-367	367-300

/a/ of /agga/:

Duration (m.secs)	0-120
F_1(Hz)	725-900
F_2(Hz)	1683-1533
F_3(Hz)	2542-2604

F_0 constant at 235 Hz.

Duration (m.secs)	0-40	50-1000	110-120
Energy	383	-350	-292

/a/ of /appa/:

Duration (m.secs)	0	-80	-170
F_1(Hz)	783	-900	-958
F_2(Hz)	1317	-1467	-1733
F_3(Hz)	2604	-2646	-2646

Duration (m.secs)	0-10	-40	50-170
F_0(Hz)	243	235	228

Duration (m.secs)	0-10	-50	-100	-120	-170
Energy	600	-500	-4000	-500	-333

/a/ of /abba/:

Duration (m.secs)	0-80	90-200
F_1(Hz)	717-883	880-958
F_2(Hz)	1340-1400	-1525
F_3(Hz)	2583-2688	-2688

Duration (m.secs)	0-10	20-150	160-180	190-200
F_o Hz)	235	228	220	230

Duration (m.secs)	0-100	110-140	150-160	170-200
Energy	500	492	-442	-292

The transition durations of these vowels were truncated simultaneously for F1/1, F2/2 and F3/3 in 10 m.secs steps and various vowel stimuli varying in transition duration were generated. However, the terminal frequencies and the steady state frequencies of F1/1, F2/2 and F3/3, were kept constant (Fig.7ll(a). Vowel /a/ with F1/l cutback was also synthesized (Fig.7.1(b)).

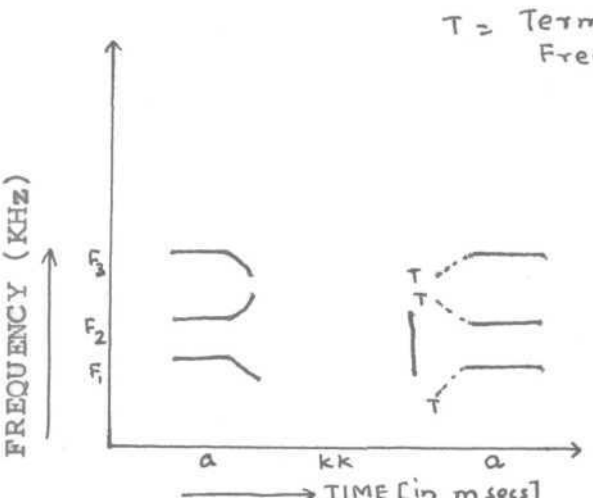

Fig.7.l(a): (Schematic representation of truncation of F1, F2 and F3 transition durations for the following vowel in /akka/.

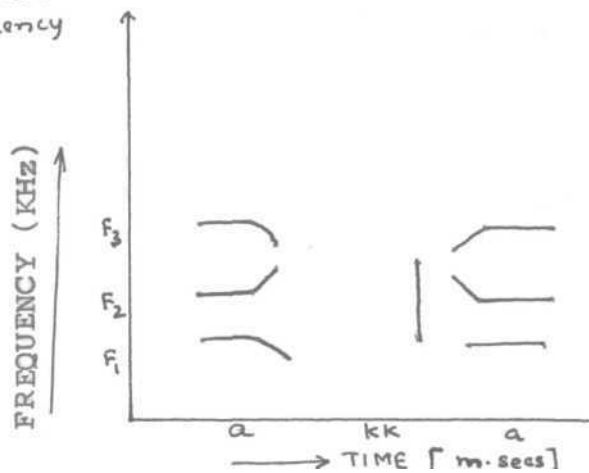

Fig.7.1(b): Schematic diagram of F_1 cutback in /akka/

The synthesized vowels were then concatenated with their respective initial VCC syllables to form complete stimuli. The number of stimuli varying in the transition duration of the following vowel obtained for each word were - /akka/-6; /agga/-6; /appa/-7; and/abba/-7. The stimuli were randomized, converted to analog stimuli and audio recorded. Each stimulus was recorded thrice with an interstimulus interval of 2 seconds). Fig. 7.2 shows the waveforms of some synthetic stimuli.

Results and Discussion:

There was no change in perception accompanied with reduced transition duration of the following vowels. The percepts of /kk/ and /pp/ were not continuous. At some transition durations the scores were scattered. Table 7.2 shows the percentage percept with respect to the following vowel transition duration. The percept /pp/was perceived as /bb/ at the truncation duration of 30 m.aecs by Malayalam males, 20 m.secs by Malayalam females, 0 m.sec by Telugu males and 20 to 40 m.secs by Telugu females. F_1 cutback had no affect on perception. There was no major differences between sex andacross languages of Malayalam and Telugu. The results were similar to the findings of Usha Rani (1989) in Kannada and Hindi listeners.

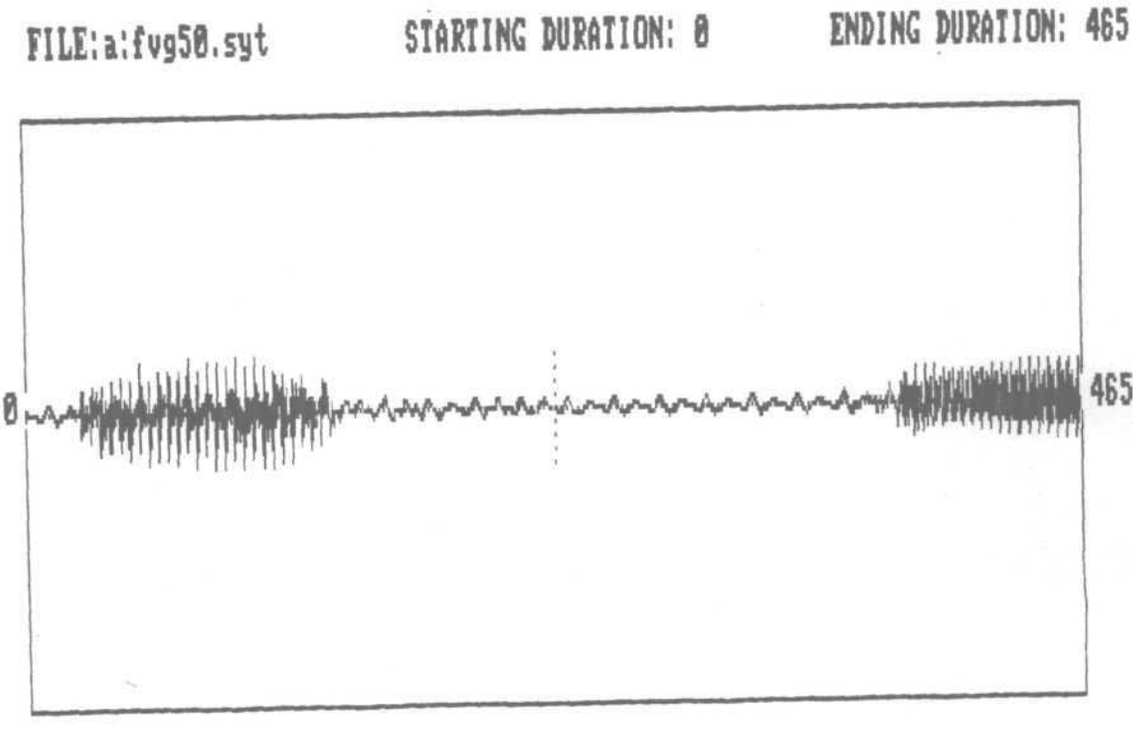

Fig. 7.2a: SYNTHESIZED /AGGA/ WITH -50 M.Secs OF TRANSITION DURATION (FOLLOWING VOWEL).

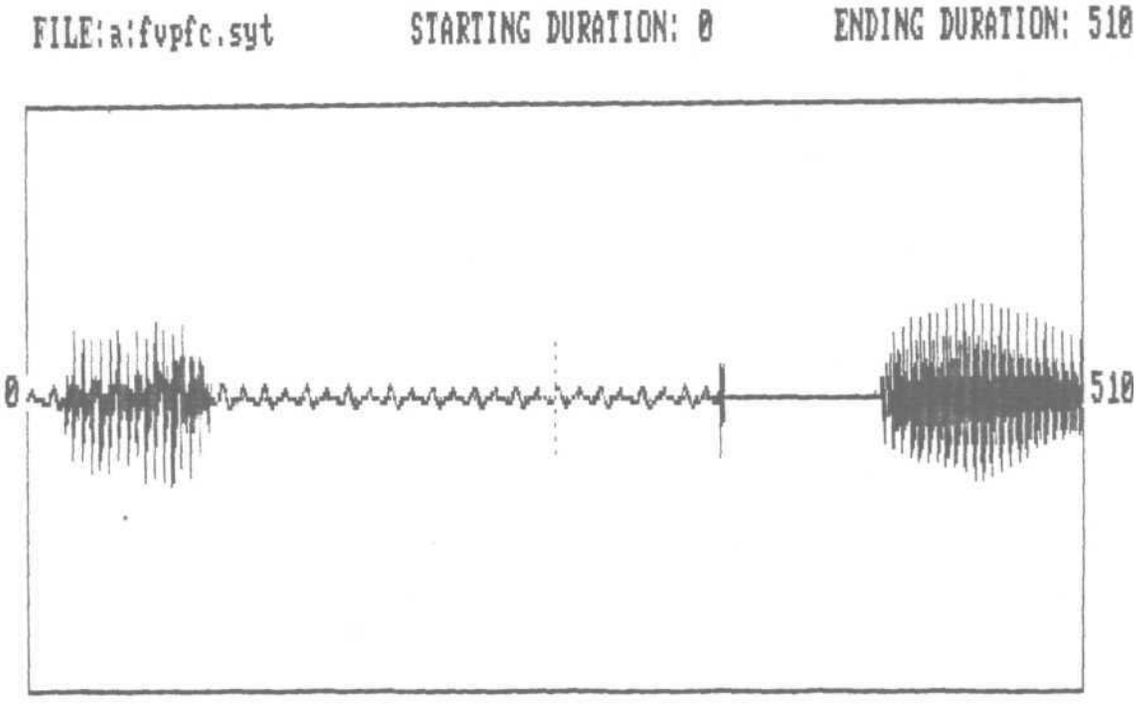

Fig.7.2b: SYNTHESIZED /APPA/ WITH Fl CUTBACK OF THE FOLLOWING VOWEL.

Fir.7.2 (a,b): waveforms of some synthetic stimuli.

Percept Identified	Following vowel transition duration truncated (in ms)				% identification			
	Malayalam		Telugu		M.M.	M.F.	T.M.	T.F.
	Male(MM)	Female(MF)	Male(TM)	Female(TP)				
Abba	F_1C.0-30	0-50	F_1 C 0-50	F_1C.0-50	64	80	63.22	95.24
Agga	0-50	0-50	0-50	0-50	86.66	100	69.04	72.21
Akka	0-30 & 50	0-30 & 50	0.20,50	0-50	76	84	71.42	88.88
Appa	F_1C, 0-10 & 40-50	F_1C,0-10 30-50	F_1 C,20 & 40-50	F_1C,0-10 & 50	68	86.66	74.99	66.66
Abba	30	20	0	20 40	60	60	57.14	66.66

Table:7.2 - Percentage of different percepts identified by native Telugu and Malayalam speakers for following vowel transition duration. F_1C- cutback.

In Kannada (Savithri, 1989) the transition durations for velars and bilabials are as in Table 7.3. There seems

	TDF1/1	TDF2/2	TDF3/3
k	13	30	8
g	32	31	13
p	15	33	23
b 14		40	24

Table 7.3: Transition durations (following vowel) of first three formants for velars and bilabials (m.secs).

to be no consistence increase of F1/1 transition form voiced to voiceless. However, F2/2 and F3/3 transitions were longer in voiced than in voiceless. These inconsistencies in production may be reflected in the perceptual responses. Thus, the transition duration of the following vowel did not seem to be a sufficient cue for voicing or place.

8. EXPERIMENT-V

VOICE ONSET TIME AND STOP PERCEPTION

Introduction:

VOT is defined as the time difference between the articulatory release and onset of voicing for the following vowel.

The VOT measure provides rather good separation for labial, dental, alveolar, retroflex and velar stops across the variety of language that have two or three distinct classes at each place of articulation (Lisker and Abramson, 1964, 1967) VOT may be said to distinguish the voiced aspirated (murmured) stops of Hindi and Marathi from voiceless stops but certainly not from unaspirated stops (Abramson, 1977). Pant (1960) and Lisker and Abramson (1964) reported that homorganic stops are distinguishable on the basis of VOT relative to their supraglottal articulation.

VOT has been found to be distinctive in many languages like English (Lisker, 1957), Hindi, Punjabi and Bengali (Lahiri, 1980), Polish (Keating, Mikos and Ganong, 1981). However, cross language differences are also found in initial stops with VOT as reported in English (Lisker and Abramson, 1964) and Danish (Fischer-Jorgensen, 1972).

In four-category languages like Hindi, Marathi, Punjabi and Bengali. VOT is insufficient to distinguish all stop categories. In these languages, the voiced aspirates are distinguished from voiceless aspirates and voiced and voiceless unaspirated by a pattern of prevoicing followed by approximately 100 m.secs of silence before resumed phonation. The lead, coincident and lag VOTs become important in these languages (Ahmad and Gupta, 1980).

VOT is found to be dependent on the place of articulation. As the tongue moves back for the articulation of stop, VOT becomes longer. This is also true for the perception of voiceless stop (Delattre, Liberman and Cooper, 1955). For labials the VOT is 25 m.secs, for alveolars 35 m.secs. and that for velars is 40 m.secs (Delattre, Liberman and Cooper, 1955) voiced plosives in English normally have a short VOT (less than 20-30 m.secs) and voiceless plosives have relatively longer VOTs (greater than about 50 m.secs), (Stevens and Klatt, 1974). In Telugu speakers, VOT values for/K/ is 35 m.secs,/g/ is 25 m.secs, /p/ and /b/ is 20 m.secs and it is insufficient to cue the place of articulation or voicing even in case of clusters (Nagamma Reddy, 1985). Usha Rani (1989) reported that VOT not a cue for the perception of voicing or clustering for the Kannada and Hindi listeners.

The aim of this experiment is to assess the role of VOT on the stop consonant perception for voicing. Difference among sex and across languages (Malayalam and Telugu) and with meaningful and nonsense stimuli were also evaluated.

Review:

The studies related to VOT are in Table 8.1.

Method:

Stimulus:

The stimuli prepared by Usha Rani (1989) were used in this experiment. (Stimuli with velar and bilabial geminate stop consonants varying in VOT were synthesized for the word /akka/ and /appa/. The words /akka/ and /appa/ were uttered by a female speaker, recorded, digitized and stored in a computer memory. The VOTs for /akka/ and /appa/ were 26.7 m.secs and 23.8 m.secs respectively. The initial vowels along with the closure duration and burst (V+CD+B) of the consonants /kk/ and /pp/ and consonant with the following vowel (VOT+V) were separated from the word /akka/ and /appa/ and stored. Stimuli were generated from the (VOT+V) syllables by truncating the initial duration (before the voicing for the vowel started) in 10 m.secs steps, till the VOT was reduced to less than 10 m.secs. Thus four stimuli, two for each stop were obtained.

Table: 8.1

Review on VOT as a cue to stop perception

Author 1.	Material used 2.	Method employed 3.	Language 4.	Nature of work 5.	Durational parameters 6.	Results 7.	Discussion 7.
Stevens & Klatt (1974)	Synthetic aspirated vs. unaspirated stop consonants.	Analysis-by-synthesis	-	To judge whether or not there was an interval of silence between the bursts of noise and onset of buzz.	VOT ranging from 0-40 ms in 5ms steps.	Minimum VOT for 50% recognition of silent interval was 20ms.	
Darwin & Brady (1975)	Synthetic /de-t_he/ & /dri-$t_h$ri/			Forced choice percept.		Boundaries of /dri-thri/continuum had shorter VOTs than /de-t^he/ continuum.	
Lisker (1975)	Synthetic CV syllable /k/ and /g/ with vowel /a/		English	Forced choice paradigm for /k/, /g/ percept.	1. VOT & Fl onset varied from 0-60ms in 5ms steps. Burst duration was 20ms & transition duration was 45 ms. 2. No Fl transition Fl constant at 769Hz for /a/.	1. /g/ and /k/ clearly divided at about 40 ms of VOT. 2. Sharply rising Fl not a requirement for /g/ percept. VOT for /g/ less than 25 ms & for /k/ greater VOT values.	
Moslin & John (1976)	Natural speech	VOT in word initial, apical stops. in test sentences, in conversations.	Korean	-	-	VOT cannot effectively distinguished strong stops from weak stops.	
Diehl (1977)	Synthetic CV syllables	—	-	-	-	Subjects identified /b/ when preceded by unambiguous /p/ when preceded by a single clear /b/ (VOT-100ms).	

1.	2.	3.	4.	5.	6.	7.	8.
Liaker etal (1977)	Synthetic speech patterns.		English	Forced choice labelling /da/ or /ta/	VOT varied from 5 to 65 ma in 19 ms steps. l.six transition duration from 20 to 85 ms. 2. 40 to 115 ms. transition duration.	50% crossover points along VOT dimension increase with incre-sing transition duration 25 ms shift in VOT was seen by a change of 95 ms in transition dura-tion.	VOT boundary is not fixed they vary directly with transition duration.
Summerfield & Haggard (1977)	Exemplars from /g-k^h/ (CV)	Closed loop algorithm controlled stimulus.	English	Forced choice iden-tification of /k^h-g/	VOT of 300 ms stimuli varying from+ 15 to +50 ms in 5 ms steps.	Increased /k^h/ per-cent with increased VOT.	
Ohde (1978)	Synthetic speech			Forced choice percept		55 ms of VOT rated as /p/ like and 5-25ms VOTs rated as /b/ like.	
Keatinh Mikos & Ganong (1981)	3 sets of synthesized apical stops followed by /a/ i.e. /ta-da/		English and Polish	Forced choice for labelling & discri-mination.	1.V0T ranging from -100-+50ms in 10ms steps 2.VOT ranging from -100 - +20 ms. 3.VOT ranging from -20 - + 80 ms.	1.Polish and English speakers use di-fferent VOT cate-gories in their voicing distinc-tion & have corres-ponding different peaks in discri-mination, the English using higher boundaries. 2.Poles are sensi-tive to diffe-rences in VOT around 0ms.	I.Poles might not use VOT as a temporal inter-val between the bursts & the voicing onset. The salience is more of psycho-acoustic short lag VOT. 2.Languages can differ in the range effects which could be due tothe internal compo-sision of their phonemic catego-, ries.
Usharani (1989)	Synthetic speech/akka/ /agga//appa/ /abba/	Analysis-by-synthesis	Kannada and Hindi	Forced choice	VOT of 0 and 10 ms.	No change in percept was observed.	

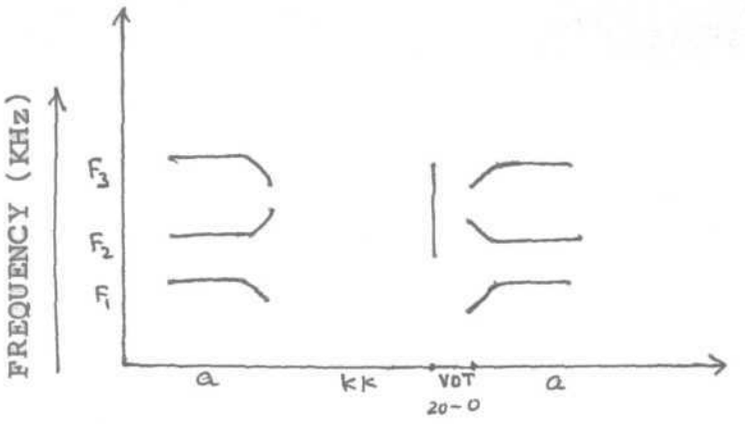

Fig. 8.1: Schematic diagram of VOT in /akka/.

These stimuli were concatenated with the (V+CD+B) stimuli appropriately to create the word (Fig.8.1). These stimuli were randomized, converted to analog and audio-recorded. Each stimulus was recorded thrice with 2 m.secs interstimulus interval). Fig.8.2 shows waveforms of stimulus with different VOTs.

Results and Discussion:

As the VOT decreased the percept changed from voice-less to voiced for velar /kk/. However, it was not so far bilabials, in case of Malayalam listeners. In case of Telugu males the velar /kk/ was perceived as /gg/ at VOT of 10 m.secs and the scores were scattered at VOT of 0 m.sec. The bilabial /pp/ was perceived at VOT of 0 m.secs and /bb/ was perceived of VOT of 10 m.secs. No apparent differences between sex, and meaningfulness of the stimulus were noticed. (Table 8.2)

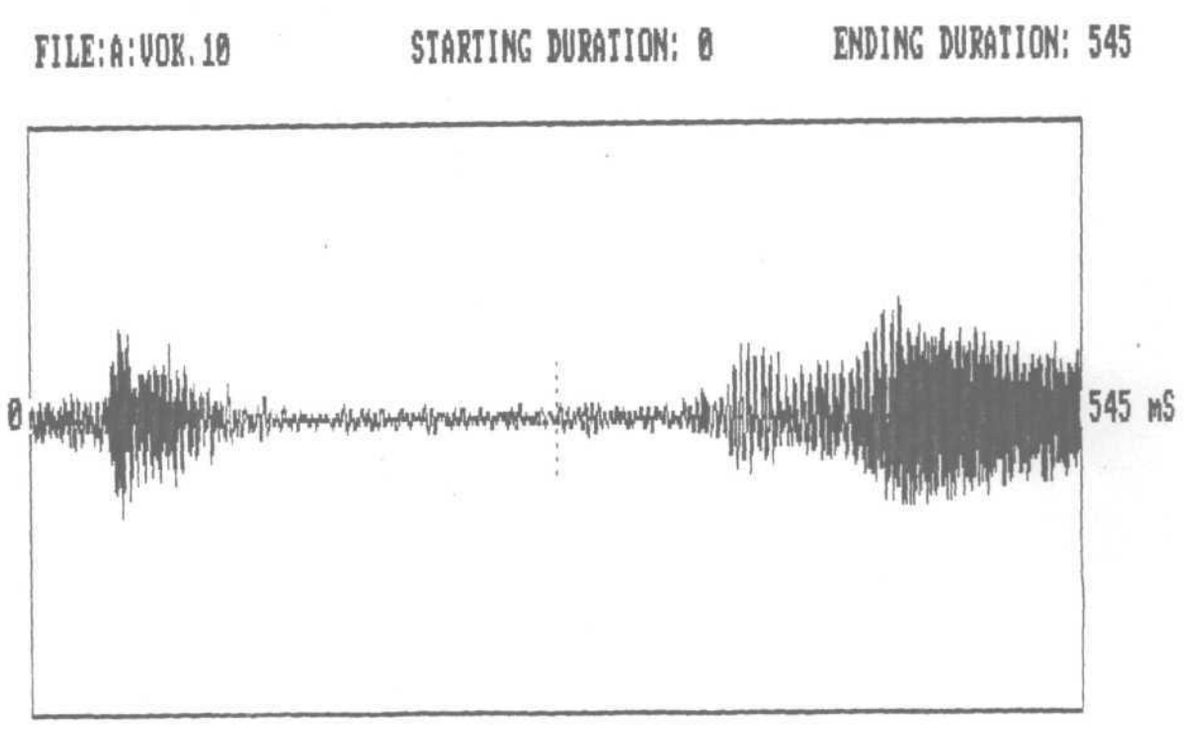

Fig.8.2a: SYNTHESIZED /AKKA/ WITH -10 M.Secs OF VOT.

99

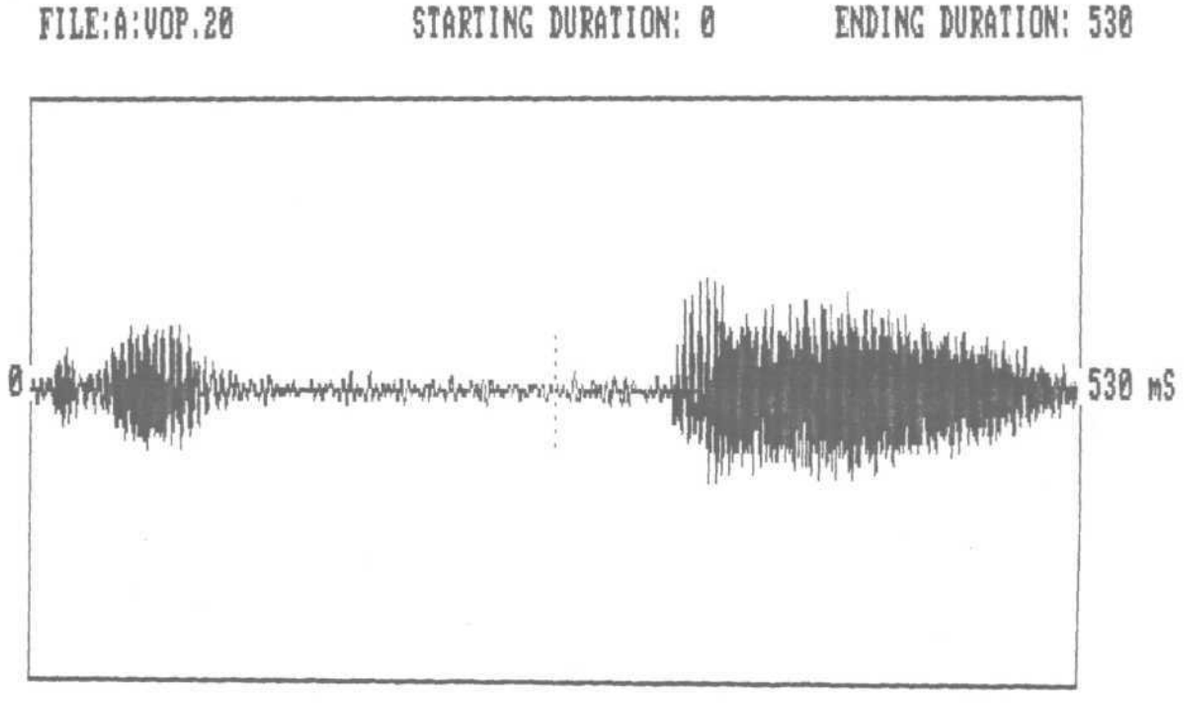

Fig.8.2b: SYNTHESIZED /APPA/ WITH -20 M.Secs OF VOT.

Fig.8.2 (a) (b) waveforms of stimulus with different VOTs.

100

Percept identified	Voice onset time (VOT) (in ms)				% identification			
	Malayalam		Telugu		M..M	M.F.	T.M.	T.F.
	Male (MM)	Female (MF)	Male (TM)	Female (TF)				
Akka	10	10	-	0,10	60	80	-	83.33
Agga	0	0	10	-	40	80	57.14	-
Appa	0,10	0,10	0	0,10	70	70	57.14	83.33
Abba	-	-	10	-	-	-	57.14	

Table: 9.2 - Percentage of different percepts identified by native Telugu and Malayalam speakers for VOT.

Part of the findings in both the languages are in accordance with the studies of Lisker and Abramson (1964, 1967), Abramson (1977) with English subjects and Vanden Berg (1988) with Dutch subjects, and part of them are not.

On the studies in Indian languages Abramson (1977) reports that VOT cannot separate the voiced aspirated stops from unaspirated stops. Nagamma Reddy (1985) reporte equal VOTs of 20 m.secs for /kk/, /pp/ and /gg/ in Telugu, Lisker and Abramson (1964) reported a VOT of 18 m.secs and above for /k/ in different languages including Tamil and Hindi. The average VOT for /p/ was found to be 0 to 13 in various languages.

The perception of velars by Malayalam and Telugu listeners is similar to that of other languages but the perception of bilabial in Telugu males is in contrast to the earlier studies in both Indian and foreign languages in that stimulus with greater VOTs (10 m.secs) was perceived as voiced. The results of present study are in accordance, with that of Usha Rani (1989),where she reported that VOT was not a cue for the perception of voicing in Kanada and Hindi.

VOT has been cited to cue for voicing and place of articulation (Lisker and Abramson, 1964, 1967). In this study the percept did not change probably because of the forced choice and because VOT of 0 and 10 m.secs were used. VOT of /k/ is more than 18 m.secs in different languages (Lisker and Abramson, 1964). Thus in these two languages VOT was not sufficient to cue voicing and place and a major like the closure duration might be operating as a cue for voicing.

9. GENERAL DISCUSSION

The effect of five temporal parameters, closure duration, preceding vowel duration, transition durations of the preceding and following vowels and the voice onset time, on the perception of bilabial and velar unaspirated medial geminate stop consonants was studied in Malayalam and Telugu listeners and the results were compared with the findings of Usha Rani (1989) for Kannada and Hindi listener, production data in Telugu (Nagamma Reddy, 1985) and studies in other languages (Lisker and Abramson, 1964; Abramson, 1977).

Of the five parameters closure-duration was found to be an important cue for voicing and clustering. The other parameters were not found to be sufficient to bring a change in the percept on their own. Thus the results of this study did not support the earlier findings that VOT, preceding vowel duration cue for voicing. On the basis the findings of this study the different parameters can be divided into two groups. The closure duration being the major cue while the other parameters, preceding vowel duration transition durations of the preceding and following vowels and VOT being the minor cues.

It might be possible that the minor cues interact among themselves or combine with the major cue to signal a particular percept. Kluender,Diehl and Wright (1988) suggested that the language communities intentionally vary vowel length in order to enhance auditorily the closure-duration cue for voicing distinctions. By the principle of duration contrast, a long vowel should make a short closure interval seem even shorter and hence more voiced, whereas a short vowel should make a long closure interval appear langer and hence more voiceless. Dorman, Studdert-Kennedy and Raphael (1977) suggested that an increase in time taken for the consonantal release (closure duration) leads to an increase in the time taken for development of a transglottal pressure drop sufficient to initiate voicing, and so to an increase in VOT. If VOT is increased, transitions into the following vowel may be largely complete at the voicing onset, so that the duration of devoiced transitions is increased. Stevens and Klatt (1974) reported a significant trading relation-ship between VOT and transition duration of the following vowel.

The functional equivalence of release bursts and transitions have been also reported by Raphael (1977). Walsh, Parker and Miller (1987) suggested that the cueing

strength of any rate of F_1 decline or steady-state duration of preceding vowel is dependent on the value of the other cues. Wardrip-Furin (1982) suggested that in addition to the preceding vowel duration the voicing during the closure may be required to disambiguate the find voiced stops.

In the light of above studies, it is clear that in natural speech, change in one parameter brings a change in the other as all these parameters are related to each other. In the present study change in one of the minor cues was not sufficient to bring a change in the percept perhaps because all the other four parameters were kept constant. For closure duration of 30 m.secs and less the percept changed from voiceless to voiced even in the absence of voicing during closure.

Only the percepts of closure-duration and VOT were in consonance with the product-data given by Nagamma Reddy (1985) and Savithri (1989) out of the five temporal parameters selected for this study. The four minor parameters were not sufficient to cue for voicing and this goes along with the findings of Nagamma Reddy (1985) in Telugu speakers. Production data in Malayalam was not available for comparision with the perceptual data. The results of this study also supported the findings of Usha Rani (1989)

in that only closure-duration was found to be a definite cue for voicing and clustering. Thus, it seems that the findings of this study partially support the motor theory of speech production, because of the similarities in perception and production seen in Telugu language.

Some differences across the languages and sex were observed but these differences were not apparent. The word /appa/ was meaningful for some of the Malayalam listeners and /akka/ was meaningful for Telugu listeners. Rest of the words were meaningless for both Malayalam and Telugu listeners. However, apparent differences emerged no on the semantical basis in this study. These findings further need to be confirmed by embeding the synthetic words in natural sentences in different languages and evaluating it for perceptual judgement by native speakers. If the perception is at the level of words and semantics does not play significant role, as seen in this study, then, no change in percepti for the synthetic words embedded in natural sentences in different language context can be found. On the other hand if semantics, does play a role or if the listener, on hearing the word of his language, switches to his language code, major differences across language could be observed. Also, multiple cues could be changed simultaneously as in the production dat* to evaluate the effectiveness of each feature by its own or as an interactive cue for a given speech sound.

10. SUMMARY AND CONCLUSION

Human speech perception is a topic of research since four decades and still the process of speech perception is for from being understood. The knowledge from the speech production studies that the acoustic cues of the speech sounds differ from language to language has triggered many studies in speech perception across languages of the world*

This study was aimed at evaluating the cueing strength of closure duration, preceding vowel duration, transition duration of preceding and following vowel and VOT for the perception of medial geminate unaspirated bilabial and velar stop consonants in Malayalam and Telugu slisteners and compare the results with the findings of Usha Rani (1989) among the Kannada and Hindi listeners.

Four Kannada words with geminate bilabial and velar stop consonants (/akka/, /agga/, /appa/ and /abba/) where taken for the experiments. Out of these four words /akka/ was meaningful in Telugu /appa/ was meaningful in Malayalam and /abba/ and /agga/ were meaningless inboth the languages.

The test stimuli synthesized by (analysis-by-synthesis/ synthesis-by-rule method) Usha Rani (1989) were used. Five different experiments were conducted for each of the para-meters mentioned earlier. The first experiment dealt with

closure duration (92 stimuli), second with preceding vowel duration (12 stimuli), third with preceding vowel transition duration (23 stimuli), fourth with following vowel transition duration (26 stimuli) and the last with VOT (4 stimuli). In total there were 157 stimuli, each audio-recorded three times with an interstimulus interval of two seconds and were givenfor the perceptual judgement of 20 subjects. Ten native Nalayalam (5 males and 5 females) speakers in the age range of 17-18 years and tea native Telugu (7 males and 3 females) speakers in the age range of 17-19 years participated in the perceptual judgement. All of them had arrived in Mysore within one month and had no exposure to Kannada. All the subjects were tested individually and were instructed to select one of the four forced-choice percept for each stimuli.

The effect of temporal parameters were evaluated in terms of voicing and clustering features of stop consonants. The other variables in the study were sex, language variation andhrole of semanticity.

The following were the results:

1. The percept changed from cluster to non-cluster as the closure duration was reduced.
2. The percept changed from voiceless to voiced as the closure duration was reduced.

3. Presence or absence of voicing during closure played an important role in perception, although, it was not a parameter under study.
4. Preceding vowel duration, preceding vowel transition duration, following vowel transition duration and VOT were found to be insufficient to cue for voicing.
5. Semanticity (meaningful vs nonsense words) did not seem to play any role.
6. No apparent differences were found between males and females and across the two languages under study.
7. The results of this study were similar to those of Usha Rani (1989) in Kannada and Hindi speakers*

It can be concluded that the closure duration and presence or absence of voicing are the major cues for the perception of voicing and clustering, while the other parameters turned out to be the minor cues. The minor cues probably interacting together or along with the major cues may signal the distinction between two phonemes.

The participation of semantic processing involving the higher cortical areas (Wernicke's, Supramarginal and angular gyrus) was not indicated by the results. It is suggested that the interaction of the different cues should

be studied by embedding the synthetic words in natural sentences and also among the patient groups with lesion in specific areas of the auditory system to gain a better understanding of the process of speech perception.

BIBLIOGRAPHY

Abramson, A.S. (1977), "Laryngeal timing in consonant distinctions", Phonetica, 34:295-303.

Ahmad, A. and Gupta, S. (1980), "Cues for the perception of (Hindi) stops", J.Acoust.Soc. India, VIII(3), 32-38.

Alwan, A. (1989), "Perceptual cues for place of articulation for the voiced pharyngeal and uvular consonants", J.Acoustic.Soc.Am., Vol.86, No.2, 549-556.

/Bailey, P.J. and Summerfield, A.O. (1978), "Some observations on the perception of /s/ + stop cluster", (Haskins Laboratories status report on Speech Research),New Haven, Haskins, Laboratories.

Bailey, P.J. and Summerfield, A.O. (1980), "Information in Speech: Observations on the perception of /s/ - stop clusters", J.of Exp.Psy: Human Perception and Performance, Vol.6(3), 536-563.

Benguerel, A.P, Hirose, H, Sawashima, M., and Ushijima, T. (1978), cited in"Representation of Voicing contrasts using articulatory gestures", by Goldstein, L. and Brownman, C.P. (1986), J.of Phonetics, 14(2), 339-342.

Berry, M.F. (1969), cited in "Auditory perception of speech: An Introduction to principles and problems" by Sanders, D.A, Prentice-Hall, INC, Englewood Cliffs, New Jersey, 98.

Borden, G.J. and Harris, K.S. (1980), "Speech science primer Physiology, Acoustics and Perception of Speech", Williams and Wilkins, Baltimore, USA, 171-214.

Brownman, C.P. and Goldstein, L. (1986), Cited in "Representation of voicing contrasts using articulatory gestures", by Goldstein, L. and Brownman, C.P. (1986), J.of Phonetics, 14, 339-342.

Chen, M. (1970), "Vowel length variation as a function of the voicing of consonants environment", Phonetica, 22; 129-159.

Cooper, F.S. Delattre, P.C. Liberman, A.M, Borst, J.M, and Gerstman, C.J. (1952), cited in Dorman, M.F, Studdert-Kennedy, M, and Raphael, L.J. (1977), "Stop-consonant recognition: Release bursts and transition as functionally equivalent, context dependent cues", Perception and Psychophysics, 22(2): 109-122.

. Datta, A.K. (1989), "Machine emulation of audition", J. of Acoustic Soc.Ind., Vol.XVII (3 & 4): 1-6.

Darwin, C.J, and Brady, S.A. (1975), cited in Brady, S.A, and Darwin, C.J. (1978) "Range effect in perception of voicing", J.Acoust.Soc.Am. Vol.63, 1556-1558.

Day, R.S, and Vigorito, J.M, (1973),cited in"Speech science primer - Physiology, Acoustics and Perception of Speech" - by Borden, G.J. and Harris, K.S. (1980), Williams and Wilkins, Baltimore, USA, 202-210.

Delattre, P.C, Liberman, A.M, and Cooper, F.s. (1955), "Acoustic loci and transitional cues for consonants", J.Acoust.Soc.Am, 27:769-773.

Denes, P. (1955), "Effect of duration on the perception of voicing", J.Acoustic.Soc.Am. Vol.27:761-769.

Diehl,R.L. (1977), "Phonetic context effects on the identification of stop consonants", J.Acoust.Soc. Am. 62, Suppl.No.l, HH 11.

Dorman, M.F. and Raphael, L.J. (1980), "Distribution of acoustic cues for stop consonant place of articulation in VCV syllables", J.Acoust. Soc.Am. Vol.67(4): 1333-1335.

Dorman, M.F, Raphael, L.J, and Liberman, A.M. (1979), "Some experiments on the sound of silence in phonetic perception", J.Acoust.Soc.Am. Vol.65(6), 1518-1532.

Dorman, M.F, Studdert-Kennedy, M, Raphael, L.J. (1977), "stop consonant recognition: Release bursts and formant transition as functionally equivalent context-dependent cues", Perception and Psycho physics. Vol.22(2): 109-122.

Elman, J.L, Diehl, R.L, and Buchwald, S.E. (1977), cited in "Speech science primer - Physiology, Acoustics and Perception of Speech", by Borden, G.J. and Harris, K.S. (1960), Williams and Wilkins, Baltimore, USA, 194.

Pant, G.M. (1960), "Acoustic theory of speech production", Monton, The Hague.

Fant, G.M. (1969), Cited in "Stop-consonant recognition: Release burst and formant transitions as functionally equivalent, context-dependent cues" by Dorman, M.F, Studdert-Kennedy, M, and Raphael, L.J., Perception and Psycho-physics, (1977), 22(2)1109-122.

Fischer-Jorgensen, E. (1954), cited in "Stop-consonant recognition: Release burst and formant transitions as functionally equivalent, context-dependent cues" by Dorman, M.F. Studdert-Kennedy, M, and Raphael, a.J., Perception and Psychophysics, (1977), 22(2):1O9-122.

Fischer-Jorgensen, E. (1972), cited in "Speech Perception", Studdert-Kennedy, M. (1980), Language and Speech, 23(1), 45-65.

Fischer-Jorgensen, E. (1979), "Temporal relations in consonant-vowel syllables with stop conso-nants based on Danish material", in B. Lindblom and Sohman (Eds) "Frontiers of Speech Communication Research", Academic Press, London, 1979.

Fitch, H.L, Hallwes, T, Erickson, D.M, and Liberman, A.M. (1980), "Perceptual equivalence of two acoustic cues for stop consonant manner". Perception and Psychophysics, Vol.27(4): 343-350.

Flanagan, J.H, Coker, C.H, Rabiner, L.R, Schafer, R.W, Umeda, N. (1970), cited in "speech analysis synthesis and perception", 2nd edition, by Flanagan, J.L., Springer-Verlag, Berlin, 1972, 268.

Fox, R.A, and Lehiste,I. (1989), "Discrimination of duration ratio in biayllabic tokens by native English and Estonian Listeners", J.of Phonetics,17: 167-174.

Frojkaer-Jensen, B, Ludvissen, C. and Rishel, J. (1973): cited in "Representation of voicing contrasts using articulatory gestures", by Goldstein, L. and Brownman, C.P, J.of Phonetic (1986), 14: 339-342.

Goldstein, L, and Brownman, C.P, (1986), "Representation of voicing contrasts using articulatory gestures", J.of Phonetics, 14: 339-342.

Gupta, S, Agrawal, S.S, and Ahmad, A. (1973); "Perceptual study of segments of intervocalic stops and affricates, J.of Acoust.Soc.Am. Vol.21(2): 108-114.

Halle, M, Hughes, G.W, and Radley, J.P.A. (1957),Cited in "Stop-consonant recognition: Release burst and formant transitions as functionally equivalent, context dependent cues", Dorman, M.F, Studdert-Kennedy, M, and Raphael,L.J. (1977), Perception and Psychophysics,22(2): 109-122.

Halle, M, and Stevens, K.N. (1959), cited in Borden, G.J. and Harris,K.S. (1980), "Speech science primer - Physiology, Acoustics and Perception of speech", Williams and Wilkins, Baltimore, USA, 171-214.

Harris, K.s, (1958), cited by Michael Studdert-Kennedy "Speech perception" in N.J.Lass (Ed) (1976), A contemporary issues in experimental phonetics". Academic Press, New York, 243-293.

Haskins group (1980), Cited in BOrden, G.J. and Harris K.S. "Speech science primer - Physiology, Acoustics and perception of speech," Williams and wilkins, Baltimore, USA.

House, A.A. (1961), "On vowel duration in English", J.Acoust. Soc.Am. Vol.33: 1764-1771.

Keating, P.A, and Blumstein, S.E. (1978), "Effects of transition length on the perception of stop consonants", J.Acoust.Soc.Am. 64(1): 57-64.

Keating, P.A, Mikos, M.J. and Ganong, W.F. (1981), "A cross language study of range of voice onset time in the perception of initial stop voicing", J.Acoust.Soc.Am. 70(5):1261-1271.

Keating, P.A. (1984), "Phonetic and phonological representation of stop consonant voicing". Language, 60, 286-319.

Kiedel, W, Kallett, S, Korth, M, and Humes, L. (1983), cited in "Tutorial: Neuroanatomy, neurophysiology, and Central auditory assessment, Part-1; Brain stem", Musseik, F.E. and Barah, J.A. Ear and Hearing, Vol.7(4):214rl986.

Klatt, D.H. (1936), "Linguistic cues of segmental duration in English: Acoustic and Perceptual evidence", J.Acoust.Soc.Am, Vol.59(5), 1208-1221.

Kluender, K.R, Diehl, R.L, Wright, B.A. (1988), "Vowel length differences before voiced and voiceless consonants: an auditory explanation, J. of Phonetics (1988), 16:153-169.

Lahiri, A. (1980), "Voicing contrasts in four category stop consonants", J.Acoust.Soc.Am. 67, Suppl.No.l, S.51.

LaRiviere, C, Winitz, H, and Harriman,E. (1975), "Vocalic transition in the perception of voiceless initial stops", J.Acoust.Soc.Am. Vol.57(2): 470-475.

Liberman, A.M. (1957), "Some results of research on speech perception", J.Acoust.Soc.Am. 29:117-123.

Liberman et al. (1972), cited in "Auditory perception of speech: An introduction to principles and problems", by Sanders, D.A., Prentice Hall INC, Englewood Cliffs, New Jersy.

Liberman, A.M, Delattre, P.C, and Cooper,F.S. (1958), "Some cues for the distinction between voiced and voiceless stops in initial positions", Language and speech, 1:153-167.

Liberman, A.M, Delattre, P.C, Cooper, F.s, Gerstman, L.J. (1954),"The role of consonant-vowel transitions in the perception of the stop and the nasal consonants", Psychological Monographs, 68:1-13.

Liberman, A.M, and Studdert-Kennedy, M (1977), cited in "Perceptual equivalence of two acoustic cues for stop consonant manner", by Fitch, H.L, Hallwes, T, Erickson, D.M. and Liberman, A.M, Perception and Psychophysics (1980), Vol.27(4): 343-350.

Lisker, L, (1957), "Closure duration and the voiced-voiceless distinction in English", Language,33:42-49.

Lisker, L. (1975), "Is it VOT or a first formant transition detector"? J.Acoust.Soc.Am 57(6); 1547-1551.

Lisker, L. (1977), "Rapid vs rabid: A catalogue of acoustic features that may cue the distinction", J. acoustic.Soc.Am. 62: Suppl.No.1,5.77.

Lisker, L, (1978), cited in "Durations and contexts as cues to word final cognate opposition in English", Raphael,L.J, Phonetica: 38: 126-147(1981).

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 VOT in English stops", Language and Speech, Vol.10, Part-1, 1-28.

Lisker, L, Liberman, A.M, Erickson,D.M, Dechritz, D, and Handler, R. (1977), "On pushing the voice onset time boundary about",J. of Language and Speech, Vol.20; 209-216.

Lisker, L, and Price, V. (1979), "Context determined effects of varying closure duration", J.Acoust.Soc. Am.65, Suppl. No.l, S-7.

Lindblom, B. and Repp, R. (1973), cited in "Representation of voicing contrasts using articulatory gestures", Goldstein, L. and Brownman, C.P. J.of Phonetics, Vol.14, 1986, 339-342.

Lindquist, J. (1972), cited in "Representation of voicing contrasts using articulatory gestures", Goldstein L, and Brownman, C.P. (1986), J.of Phonetics, 14, 339-342.

Lotz, J, Abramson, A.S., Gerstman, L.J, Imgemann, F. and Nemser, W.J. (1960),"The perception of English stops by speakers of English, Spanish, Hungarian and Thai: A tape-cutting experiment, "Lang.Speech, Vol.3:part-2:71-77.

Mack, M. (1982),"Voicing dependent vowel duration in English and French monolingual and bilingual production", J.Acoust.Soc.Am. 71:173-78.

Moslin, B.J. and John, K.S. (1976), "Korean apical stop production: A VOT analysis", J.Acoust.Soc. Am, 60, Suppl.No.1,S.95.

Nagamma Reddy, (1985),"The duration of Telugu speech sounds: An acoustic study. J.INSTN. Electronics and Telecom. Enhrs, Vol.34, No.l, 1985:57-63.

Ohde, R.N. (1978), "Effects of VOT duration and number of adaptor repetitions on the scaling of stop consonants Voicing", J.Acoust.Soc.Am. 64, Suppl.No.l, 5-19.

Port, R.F. (1976), cited in Dorman, M.F, Raphael, L.J, and Liberman, A.M. "Some experiments on the sound of silence in phonetic perception", J.Acoust.Soc.Am 65(6), 1578-1532.

Port, R.F. (1980), "V/C ratio as a post vocalic voicing cue", J.Acoust.Soc.Am. 67, Suppl.No.l, S51.

Price, P., and Lisker, L. (1979), "(/b/-/p/) but F (/p/-/b/)", J.Acoust.Soc.Am. 65,suppl.No.l, S7-8.

Raphael,L.J. (1960),"Contribution of C.V. transition duration to the perception of final consonant voicing in natural speech", J.Acoust.Soc. Am.67, Suppl.No.l, 5-51.

Raphael, L.J. (1972), "Preceding vowel duration are cue to the perception of the voicing characteristics of word-final consonants in American English",J.Acoust.Sos.Am. 51(4): 1297-1303.

Raphael, L.J. (1981), "Duration and contexts as cues to word-final cognate opposition in English", Phonetic a, 38: 126-147.

Raphael, L.J, Dorman M.F. and Liberman, A. (1980), cited in "Durations and contexts as cues to word-final cognate opposition in English", Phonetica,38: 126-147.

Sanders, D.A.(1977), "Auditory perception of speech: An Introduction to principles and problems", Prentice-Hall INC, Englewood Cliffs, New Jersey. 98.

Savithri, S.R. (1986), "Durational analysis of Kannada Vowels", J.Acoust.Soc.Ind. 24(2): 34-41.

Savithri, S.R. (1989), "A databank on the duration of Kannada speech sounds", J.Acoust.Soc.Ind. Vol.17: 1&2, 308-312.

Singh, S, and Black, J.W. (1966), "Study of Twenty-six intervocalic consonants as spoken and recognized by four language groups", J.Acoust.Soc. Am. Vol.39(2), 372-386.

Slis, I.H, and Damste, P.H. (1967), cited in"Representation of voicing contrasts using articulatory gestures", Goldstein, L. and Brownman, C.P. J.of Phonetics, 14: 339-342.

Slis, I.H. and Cohen, A.(1969), "On the complex regulating the voiced-voiceless distinction II, Language and Speech, 12:137-55.

Smith, B.L. and Muse, A.M. (1987),"Kinematic characteristics of post vocalic labial stop consonants produced by children and adults", Phonetica 44: 227-237.

Stevens,K.N. (1972), cited in "speech science primer-physiology, acoustics and perception of speech" by Borden, G.J. and Harris, K.S. (1980), Williams and Wilkins, Baltimore, USA, 203-210.

Stevens, K.N. (1989), "On Quantal nature of speech", J.of Phonetics, 17:205-212.

Stevens, K.N. and Halle,M (1959), cited in "Auditory perception of speech: An introduction to principles and problems" by Sanders, D.A., Prentice-Hall INC, Englewood Cliffs, New Jersey, 98.

Stevens, K.N. and Klatt, D.H. (1974), "Role of formant transition in the voiced voiceless distinction for stops"., J.Acoust.Soc.Am. 55(3), 653-659.

Strange, W. and Jenkins,J.J. (1978), cited in "Speech science primer-Physiology, Acoustics and Perception of Speech", by Borden, G.J. and Harris, K.S. (1980), Williams and Wilkins, Baltimore, USA, 194.

Summerfield, Q and Haggard, M. (1977), "On the disociation of spectral and temporal cues to the voicing distinction in initial stop consonants", J. Acoust.Soc. Am. Vol.62(2), 435-448.

Tyagi, S.K, Agrawal, S.s. and Pavate, K.D. (1987), "Acoustic analysis of Hindi stop consonants", J.of Acoustic Soc.Ind, Proceedings of National Symposium on Acoustics, 1987.

Usha Rani, P.N. (1989),"Temporal perceptual cues of velar and bilabial stop consonants", Dissertation submitted to University of Mysore.

Van den Berg, R.J.H. (1988), "The perception of voicing in Dutch two obstruent sequence, a comparision of synthetic and natural speech", J. of Phonetics, 16.

Walsh, T, Parker, F. and Miller, C.J. (1987), "The contribution of rate of F. decline to the perception of (± voice)", J.of Phonetics, 15:101-103.

Wardsip-Fruim, C, and Bischoff, D.M. (1976), "Vowel duration as a cue for consonant voicing?", J.Acoustic Soc. Am. 60, Suppl.l, S91.

Wardrip-Fruim, C, (1982),"On the status of temporal and phonetic categories, PVD as a cue to voicing in final stop categories", J.Acoust.Soc.Am. 71, 187-195.

Winitz, H. et al (1975), "Variations in VOT for English initial stops",J.Phonetics, 3:41-52.

Wolf, C.G. (1976),"Perception of voicing in final stop", J. Acoust.Soc.Am. Vol.60,Suppl.No.l, S-91.

APPENDIX - RESPONSE-SHEET

Name: Rasitha Rabindran Age: 18 Mother tongue: Malayalam

Qualification: I[st] BSc

Sample: 1

Sets	akka	agga	aka	aga
1.	✓			
2.	✓			
3.	✓			
4.	✓			
5.	✓			
6.	✓			
7.	✓			
8.			✓	
9.	✓			
10.			✓	
11.			✓	
12.				✓
13.			✓	
14.			✓	
15.				✓
16.	✓			
17.	✓			
18.	✓			
19.		✓		
20.		✓		
21.			✓	
22.			✓	