VOICE ONSET TIME AS A PERCEPTUAL CUE TO VOICING CONTRAST

Register No. M 9308

A DISSERTATION SUBMITTED AS PART FULFILLMENT FOR THE
FINAL YEAR M.Sc. (SPEECH AND HEARING)
TO THE UNIVERSITY OF MYSORE

ALL INDIA INSTITUTE OF SPEECH AND HEARING
MYSORE - 570 006

MAY 1995

To,

My Dearest Ammayya,
An Human Dynamo, Whose
Zest and Zeal for Life
Never Cease to Surprise

CERTIFICATE

This is to certify that this dissertation entitled "VOICE ONSET TIME AS A PERCEPTUAL CUE TO VOICING CONTRAST" is the bonafide work, done in part fulfillment for the degree of "MASTER OF SCIENCE (SPEECH AND HEARING)", of the student with Register No. M 9308.

MYSORE

MAY, 1995

Dr. S. NIKAM
DIRECTOR
ALL INDIA INSTITUE OF
SPEECH AND HEARING
MYSORE -570 006

CERTIFICATE

This is to certify that this dissertation entitled "VOICE ONSET TIME AS A PERCEPTUAL CUE TO VOICING CONTRAST" has been prepared under my supervision and guidance.

MYSORE

MAY, 1995

Savithri S.R

Dr. S.R. SAVITHRI
LECTUTER
DEPARTMENT OF SPEECH SCIENCES
ALL INDIA INSTITUE OF
SPEECH AND HEARING
MYSORE - 570 006

DECLARATION

I hereby declare that this dissertation entitled ***"VOICE ONSET TIME AS A PERCEPTUAL CUE TO VOICING CONTRAST"*** *is the result of my own study under the guidance of* ***Dr. S.R. SAVITHRI,*** *Lecturer, Department of Speech Sciences, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier at any university for any other diploma or degree.*

MYSORE ***(REG NO. M 9308)***

MAY, 1995

ACKNOWLEDGEMENTS

*I extend my sincere gratitude to **Dr. (Mrs) S.R.. SAVITRI,** Dear Madam, I could not have spelt S   C C E S S without 'U'. Thank You.*

*My heartfelt thanks to **Dr. (Miss) S. NIKAM,** Director, AIISH, Mysore for allowing me to undertake this study.*

*My sincere gratitude to Dr. **N.P. NATRAJ,** Head of the Department of Speech Sciences, for allowing me access to Speech Science lab.*

***Mr. C.S. Venkatesh,** Thank You sir, for all the help that you provided.*

Thanks to all my subjects, who listened to the soporific litany of Kadi, Tada and Padi uncomplainingly.

*Dear **Amma and Appa,** Thanks for everything.*

*Dear **Anu**$_r$ **Mona, Ritu** & **Sashi,** you people have been a constant source of joy to me. Thankyou, for you .*

Thanks to all my classmates, for the moral support they provided.

***Rahul, Sara, Megha, Kanchan & Pradeep,** I would have missed out on a lot of things, if I had not known you. Thankyou.*

*Thanks to **Mr. Kiran, Sangeetha and Swarna** for their wonderful work.*

There is always room for improvement, it is the biggest room in the house.

(REG NO. M 9308)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

CHAPTER I

INTRODUCTION

He gave man speech and speech created thought,
which is the power of the universe.

Shelley

Speech is a form of oral communication in which transformation of information takes place by means of speech waves which are in the form of acoustic energy (Fant, 1960). Speech perception is the process where in speech is decoded and interpreted by the listener.

The ear seems to be custom built for the purpose of detecting and analysing sounds. The speech signals which are long spurts of a complex and constantly changing stream of sounds radiate from the speaker's lips, travel in air, impinge upon the ear drum of the listener and reach the higher cortical structures through middle and inner ears and the auditory pathway. The speech signal is analyzed temporally and spectrally at the lower centers to some extent (i.e, below the thalamus level) 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, Kallert, Korth, and Humes, 1983). The linguistic components are added only at the higher centers of the cortex to the already analyzed signal to reconstruct the percept intended by the speaker. When the listener has reconstructed this signal (i.e, decoded and interpreted) speech perception is said to have occurred.

Speech sounds are varied and have numerous acoustic cues. It seems that the auditory system depends on some of the acoustic cues of the speech sounds to identify and thus to perceive it.) From the speech production studies, it is known that speech sounds are bundles of different acoustic cues like the formants, their band widths and levels, Fundamental frequency (F0), energy, duration of closure, preceding vowel duration (PVD), burst energy and voice onset time (VOT).

The process of speech perception in human beings is of interest and extensive research has been conducted in the recent past to obtain a knowledge about the processing of speech signals in the auditory pathway. The results of these research have enhanced the knowledge about the process of speech perception and has provided information about the cues that could be used with the speech and language handicapped.

These studies involve spectral analysis or perceptual analysis. Spectrography has been extensively used to obtain production data, to gather the details of the parameters of speech sounds which are perceived. Parameters which characterize particular speech sounds, are considered to be the cues for the perception of those sounds.

In most of the speech perception studies, speech sounds are reconstructed from their known spectral and temporal Parameters and presented to the listeners for judgement. The 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 synthesis (Fant, 1960) and synthesis by rule (Flanagan et al, 1970). These techniques have been used to assess the role of temporal parameters and spectral parameters. Among the speech sounds, stop consonants have been worked on extensively. Dorman, Raphael and Liberman (1979a), Fischer Jorgensen(1979), Lisker and Price (1979), Bailey and Summer field (1980), Fitch, Halves, Erickson and Liberman (1980), Port (1980), Usha Rani (1989), Datta(1989) , Vinay Rakesh (1990) have worked on closure duration. Raphael (1972), Fruin and Bishoff (1976), Raphael (1980), Usha Rani (1989) and Vinay Rakesh (1990) have experimented with the preceding vowel duration as a cue to voicing of stop consonants. Keating and Blumstein (1978), Haskins group (1980), Raphael (1980) Usha Rani (1989) and Vinay Rakesh (1990) have worked with the transition duration of the preceding vowel which may cue the stop perception. Stevens and Klatt (1974), Lisker (1975), Summerfield and Haggard (1977), Ahmed and Gupta (1980), Bailey and Summer field (1980), Usha Rani (1989) and Vinay Rakesh (1990) have explored the possibility of the transition duration of the following vowel as a cue to voicing of stop consonant.

VOT is the duration of the interval between onset of burst resulting from stop release and the glottal signal (Lisker and Abramson, 1964). It has been claimed that short lag or lead VOT's cue voiced cognates and long lag VOT's cue voiceless cognates. Lisker and Abramson (1964), Lisker (1967, 1975), Stevens and Klatt (1974) Zlatin (1974), Lisker (1975), Darwin and Brady (1975), Moslin and John (1976), Williams (1976), Wood (1976), Carney (1977), Diehl et al, (1977), Summer field and Haggard (1977), Ohde (1978), Keating, Mikos and Ganong (1981), Elliott (1986), Usharani (1989), Vinay Rakesh (1990), Burnhem et al (1991), Volaitis and Miller (1992), Kluender et al, (1992), Nearey and Rochet (1992), Kim (1993) have worked on VOT as a cue to stop consonant voicing.

In spite of extensive research in this area, the way in which the cues combine to produce a percept is not yet known. Also, some of the temporal parameters seem to vary across languages. In this context, the present study was designed to find out the role of VOT as a perceptual cue to voicing of stop consonants in Kannada. It was also designed to evaluate the possible role of VOT as a cue to place of articulation.

REVIEW OF LITERATURE

CHAPTER II

REVIEW OF LITERATURE

I. Stop consonants:

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 examples for listener's use of the acoustic overlapping of phonemes in speech systems. Also, they have consistently produced evidence for phonetic level processing. They appear to be the 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 transient explosion (usually less than 20 m.sec) produced by shock excitation of the 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 a narrow constriction as in the homo organic fricatives.

4. A very brief period of aspiration (2-20 m.sec) with in which may be detected 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.

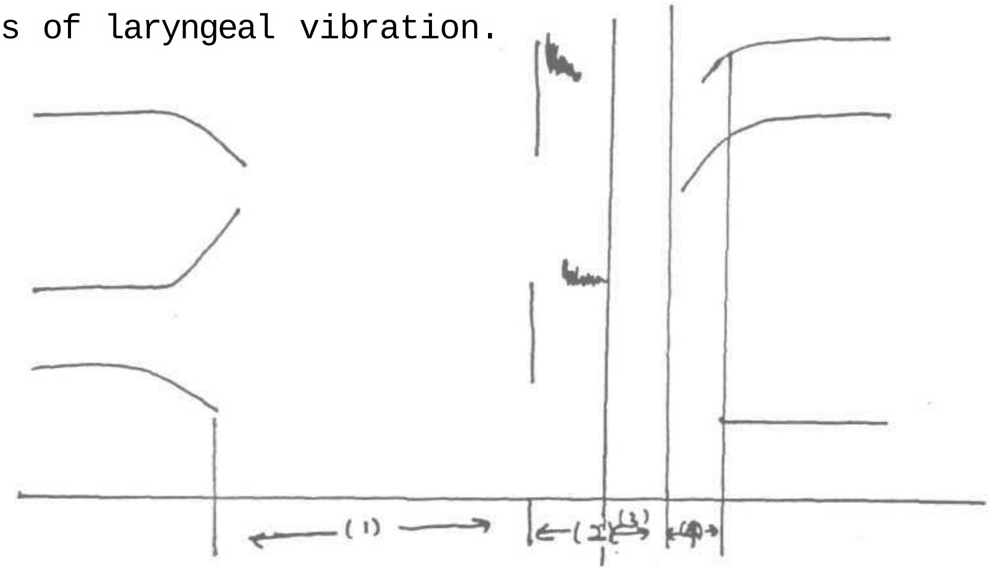

Figure 1 - Spectrogram depicting the salient features of /k/

1. Period of occlusion.
2. Period of frication.
3. Period of aspiration.
4. Voiced formant transition into the following vowel.

Several experiments have been conducted to investigate the perceptual cues of stop consonant in normal adults and children and in Speech handicapped population. The various parameters studied can be listed under the spectral and temporal characteristics.

The spectral parameters include :

1. Burst amplitude.
2. Burst frequency.
3. F0 change in the succeeding vowel.
4. Frequency of formants 1,2, and 3.
5. Band widths of formants 1,2, and 3.
6. Direction of second and third formant (F^ and F-J transitions.
7. Voicing during closure.

The temporal parameters include :

a) Preceding vowel duration (VD/PVD).
b) Closure duration (CD).
c) Voice onset time (VOT).
d) Voice offset time.
e) Stop consonant duration.
f) Off glide duration of the first formant, F1.
g) Off glide duration of the second formant, F2.
h) Burst duration (BD).

Voicing cues have been studied extensively. Lisker (1977) in his study has listed 16 parameters that cue voicing of stop consonants in medial position. They are :

1. Presence/absence of low frequency buzz during the closure interval.
2. Duration of closure.
3. F1 offset frequency before closure.

4. F1 offset transition duration.
5. F1 onset transition duration.
6. F1 onset frequency following closure.
7. Preceding vowel duration.
8. F1 cutback following closure.
9. F1 cutback before closure.
10. VOT delay after closure.
11. VOT cutback after closure.
12. F0 contour before closure.
13. F0 contour after closure.
14. Amplitude relative to vowel.
15. Decay time of glottal signal preceding closure.
16. Intensity of burst closure.

As the present study aims to study VOT as a perceptual cue to voicing of stop consonants in Kannada, the review here is restricted to studies of VOT as a perceptual cue.

II. Voice Onset Time:

(Voice Onset Time refers to the duration of the interval between onset of the burst resulting from stop release and the glottal signal. It has been found to be distinctive for voicing in many languages. Lisker and Abramson (1964),/ Lisker (1967) opine that differences in VOT provided a useful acoustic measure of the various phonemic categories such as 'voiced stop' and 'voiceless unaspirated stops'. Lisker and Abramson (1964) proposed that languages which have stop voicing contrasts have chosen among three VOT categories namely :

a) Voicing lead (negative VOT) or prevoicihg, in which voicing onset precedes the release burst.

b) Coincident or short lag VOT (with zero or low positive VOT values) in which voicing onset is simultaneous or briefly lags behind the release burst.

c) Long lag VOT (with high positive VOT values) in which the voicing onset lags behind the release burst.

Voiced plosives in English normally have a short lag VOT (less than 20-30 m.sec) and voiceless plosives on the other hand have relatively longer VOT values (more than 50 m.sec) (Lisker, 1975). After the initial work of Lisker and Abramson on VOT, several investigators have evaluated the role of VOT in various languages.

III. VOT as a perceptual cue to voicing:

Stevens and Klatt (1974) studied speech perception using analysis by synthesis method. Synthetic aspirated and unaspirated stop consonants were used. The perceptual task was to judge whether or not there was an interval of silence between the bursts of noise and onset of buzz. The VOT ranged from 0 to 40 m.sec and was varied in 5 m.sec steps. They found that the minimum VOT for 50% recognition of silent interval was 20 m.sec.

Zlatin (1974) conducted a study to examine the status of VOT in perception of word initial voiced and voiceless stops?

Synthetic CV syllables with VOT continuum ranging from -150 m.sec to + 150 m.sec and varied in 5 m.sec or 10 m.sec were used. Forced choice procedure was used for b Vs p, k Vs g, and d Vs t percept. Zlatin concluded that VOT is a primary cue for differentiation of homorganic stop consonants./

Lisker (1975) synthesized CV syllables. The stop consonants /k/ and /g/ were used with the vowel /a/. The temporal parameters were varied as follows - In the first condition, VOT and F1 onset were varied from 0 to 60 m.sec in 5 m.sec steps, the burst duration was 20 m.sec and the transition duration was 45 m.sec. In the second condition F1 was kept constant at 769 Hz for /a/. A forced choice task was used for /k/ and /g/ percept. The results showed that /g/ and /k/ were clearly divided at about 40 m.sec of VOT. Sharply rising F1 was not found to be a requirement for /g/ percept. VOT for /g/ was found to be less than 25 m.sec and /k/ had greater VOT values.

In an investigation on the role of VOT in distinguishing among Korean apical stop consonants, Moslin and John (1976) measured VOTs for word initial apical stops in the speech of four native Korean speakers. Words in citation form, in test sentences, in conversation among Korean adults and in mother's speech to children were used. Results indicated that although VOT is sufficient to distinguish the strong from the aspirated stops, it cannot effectively distinguish either of these from weak stops.

Williams (1976) used synthetically produced syllable initial stops with VOT ranging from - 40 m.sec to + 40 m.sec, these were given to Spanish listeners for judgement. Seven out of Eight listener's divided the series into voiced and voiceless portions with in the prevoiced region suggesting that prevoicing can be a sufficient voicing cue.

Mood (1976) employed Signal Detection methodology to study the phoneme boundary effect. Synthetic stimuli ranging from /ba/ to /pa/ (VOT's from - 50 m.sec to + 70 m.sec) were given to subjects in a same - different discrimination task. The results showed that there was a clear increase in discriminability and a marked shift in response bias from same to different occurred near the voiced-voiceless boundary. This was seen even when VOT were isolated from syllable context so that they were not categorized as phonemes. The results suggest that phoneme boundary effect for VOT is not due exclusively to phonetic categorization but may instead reflect acoustic and auditory properties which are distinct from phonetic processing.

Carney (1977) studied the discriminability of bilabial stop consonants differing in VOT (- 100 m.sec to + 100 m.sec) as measured in a same-different task, an oddity task and a dual response discrimination - identification task. After a moderate amount of training in a same-different task with a fixed standard and with feed back, subjects showed excellent within category discrimination in all three tasks , discrimination performance continuously improved and well-defined category boundaries fell at arbitrary values determined by the experimenters.

Darwin and Brady (1975) used a synthetic VOT continuum ranging from 5 to 55 m.sec in 5 m.sec steps, these were presented to the subjects in blocks A (15 -25 m.sec), B (15 - 35 m.sec), C (25 - 45 m.sec) and D (35 - 55 m.sec) and one block covering the whole range. These were given for perceptual analysis and the percent /d/ response calculated. Results show that the location of the voicing boundary in the perception of initial stop consonants is shown to vary according to the range of voice onset times used in a block of trials and according to the order in which blocks covering different ranges are presented. Subjects were more willing to perceive as unvoiced a sound to the long VOT end of a short VOT range than to perceive as voiced a sound to the short VOT end of the corresponding long - VOT range.

Diehl, Buchwald and Elman (1977) used synthetic CV syllables in an adaptation experiment. Each test syllable had a value of VOT which placed it near the English voiced-voiceless boundary. The investigators found that when the test syllables were preceded by a clear /b/ [VOT = - 100 m.sec] subjects tended to identify them as /p/, whereas when they were preceded by an unambiguous /p/ (VOT = + 100 m.sec), the syllables were labelled /b/. This contrast effect occurred even when the contextual stimuli were velar and the test stimuli were bilabial suggesting a featural rather than a phonemic basis for the effect.

Elman, Diehl and Buchwald (1977) studied the identification performance of three groups of subjects, monolingual English speakers, monolingual Spanish speakers and English - Spanish bilinguals. Naturally produced /ba/ and /pa/ syllables with VOT ranging from - 69 m.sec to + 66 m.sec were presented for identification. Results indicated that the two monolingual groups differed substantially in their identification performance with the English speakers tending to label most of the stimuli as /ba/ and Spanish speakers tending to label most of them as /pa/. The bilingual listener's placement of boundaries varied as a function of language set depending on whether they were strong, moderate or weak bilinguals.

Summerfield and Haggard (1977) used synthetic (g-kh) CV stimuli with VOT ranging from 0 to + 80 m.sec in a PEST method expeirment. It was found that with increased VOT values greater number of /kh/ percept was reported.

Ohde (1978) examined the effects of duration and number of repetitions of the adapting stimulus on the voicing feature scaling of stimuli varying in VOT before and after adaptation. The adapting stimulus was either 5, 25 or 55 m.sec VOT and the number of repetitions of adaptation trial was 5, 32 or 95. The findings were - the 55 m.sec adaptor was rated as p-like and the 5 and 25 m.sec adaptors were rated as b-like. Greater shifts were seen for longer VOT adaptors and for greater repetitions. THe researcher concludes that the results support a fatigue type model and effects of adaptor repetition support an auditory component of voicing analysis.

Keating, Mikos and Ganong (1981) studied the cross language differences between English and Polish. Three sets of synthesized apical stops followed by /a/ i.e, (ta.........da) were used in a forced choice task for labelling and discrimination. The VOT duration range were as follows :

1. VOT ranging from -100m.sec to +50m.sec.
2. VOT ranging from -100m.sec to +20m.sec.
3. VOT ranging from - 20m.sec to +80m.sec.

They found that Polish and English speakers use different VOT categories in their voicing distinctions and have corresponding different peaks in discrimination ; the English use higher boundaries. The poles were found sensitive to differences of VOT around 0 m.sec.

The researchers concluded that

1. Poles might not use VOT as temporal interval between the bursts and the voicing onset. The salience is more of Psychoacoustic short lag VOT.

2. Languages can differ in the range effects which could be due to internal composition of their phonemic categories.

Elliot (1986) used consonant vowel (CV) continuum in which VOT ranged from 0 to 35 m.sec, and values were varied in 10 m.sec and 20 m.sec steps, to study the age related differences in VOT discrimination. The two groups of children and one adult group were given these stimuli in a same-different identification task. Results showed that children displayed poorer discrimination than adults for CV pairs differing by both time intervals.

Usha Rani (1989) studied the effect of temporal parameters in cueing voicing in Kannada and Hindi. The stop consonants were studied in intervocalic context. The VCV syllable was stored, the CV segment was separated and the VOT truncated in steps of 10 m.sec and then the series of newly formed VCV syllables was given for perceptual analysis to adult speakers of Kannada and Hindi. Results indicated that as VOT was truncated no change in percept was reported except for /k/ by Hindi speakers. Usha Rani (1989) concluded that VOT did not cue the perception of voicing in intervocalic position but did cue place of articulation.

Vinay Rakesh (1990) studied the effect of five temporal parameters (closure duration, preceding vowel duration, transition duration of the preceding and following vowels and VOT) in cueing the cluster and voicing feature of unaspirated bilabial and velar stops in Malayalam and Telugu. The synthetic test stimuli (VCV) were presented to Malayalam and Telugu speakers for perceptual analysis. The results indicated that the closure duration and presence or absence of voicing are the major cues for perception of voicing and clustering while preceding vowel duration, preceding vowel transition duration, following vowel transition duration and VOT were found to be insufficient cues for voicing.

Burnham, Earnshaw and Clark (1991) used an infant speech identification procedure to test the identification of sounds on a native (voiced/voiceless bilabial stop) continuum and a non-native (Prevoiced/voiced bilabial stop) continuum. The subjects taken were infants, two and six-year old children and adults. Synthetic syllables (bilabial stop plus vowel /a/) with VOT of initial consonant ranging from - 70 m.sec to + 70 m.sec in 10 m.sec steps were used. The subjects were also tested for their identification of a continuum of harmonic tones varying in pitch. Results revealed that with age, the native contrasts were perceived more categorically where as Non-native contrast were perceived more noncategorically with age. The scores of the native contrast were not significantly greater than the scores of pitch identification. The scores of the pitch identification task were significantly greater than non-native contrasts.

Kim (1993) studied the production and perception of three classes of stop consonants in Korean : Voiceless unaspirated (t) Voiceless slightly aspirated (t) and voiceless heavily aspirated (t). For these classes of stops produced by male and female Korean speakers Voice onset time, amplitude of aspiration and F_0 of the first five glottal pulses were measured. VOT was longer for slightly aspirated than for unaspirated stops and longer still for heavily aspirated stops. Unaspirated and heavily aspirated stops were both produced with significantly higher F_0 than slightly aspirated stops. For the perception task six 12 step series of syllables differing in VOT (5 - 82 m.sec), F_0 (100, 125, 150 Hz) and amplitude of aspiration noise (- 37, - 21 dB) were synthesized. These were given to adult speakers for perceptual analysis. Results revealed that VOT and the other two stimulus variables were perceptually significant for the Korean listeners.

IV. VOT as a perceptual cue to place of articulation:

VOT is found to be dependant on the place of articulation. As the tongue moves back for the articulation of stops, VOT becomes longer. This is also true for the perception of voiceless stops (Delattre Liberman and Cooper, 1955) For labials the VOT is 25 m.sec for alveolars 35 m.sec and for velars 40 m.sec (Delattre, Liberman and Cooper, 1955).

This finding has been reported by other authors also. Lisker and Abramson (1964) also reported that VOT typically increases from labial to apical to velar points of articulation. }

Zlatin (1974) studied both perceptual and productive VOT characteristics of adults. Synthetic syllables with VOT ranging from - 150 m.sec to + 150 m.sec were given for perceptual analysis using a forced choice procedure. Results showed a consistent increase in cross over value as place of articulation moved back.

Miller (1977) assessed the location of voiced-voiceless boundary as a function of place of articulation. Synthetic stimuli, (the syllables /ba/, /pa/, /ga/, /ka/, /da/ and /ta/) with VOT ranging from 0 to + 50 m.sec were used in an identification task. The findings were similar to the findings of Lisker and Abramson's study ie, phonetic boundary systematically shifted towards the voiceless end of the series as the place of articulation varied from front to back. Miller concluded that atleast for stimuli near the phonetic boundary, the assignment to a voicing value is contingent on place value assigned.

Repp (1977) investigated the dependence of voicing boundaries on place cues by varying F_2 and F_3 transition onset frequencies of syllable initial stop consonants as well as their VOT. He reported evidence for changes in voicing boundary which was tied to the perceived place category. Also a dependence of the place boundary on VOT : labial - alveolar - velar boundaries converge as VOT increases resulting in a reduction of the size of the alveolar category was reported.

Usha Rani (1989) evaluated the role of VOT in cueing voicing contrast in Kannada and Hindi. Results indicated that VOT did not cue the perception of voicing but did cue place of articulation in Kannada.

Volaitis and Miller (1992) conducted a series of experiments to study affect of place of articulation on the perception of voicing. Synthetic /bi/ and /gi/ tokens were given for perceptual analysis. Results indicated that a change in place of articulation from labial to velar consonants resulted in a shift in the voiced-voiceless category boundary values towards longer VOT values. This effect was seen irrespective of the speaking rate.

Nearey and Rochet (1992) studied the effect of place of articulation on the perception of VOT continua. Twelve different continua (VOT ranging from - 80 to + 80 m.sec in 10 m.sec steps) were presented to English and French speakers for perceptual analysis. The consonants taken were /b, d, g, p, t.

k/. Results indicate that both French and English speakers shoved significant effects for place i.e, as place of articulation moved forward the voiced-voiceless cross over boundary values reduced.

V. VOT trade off with other cues:

There is voluminous evidence that homorganic stop consonants are distinguished on the bases of VOT relative to their supra glottal articulation. Following are a few studies which highlight the trade off of VOT with other cues.

Stevens and Klatt (1974) have emphasized the role of transition duration showing that with greater durations there is increase in the VOT value at the boundary between /da/ and /ta/ syllables. They have reported upto 15 m.sec change in the location of the perceived phoneme boundary as measured in terms of VOT.

Lisker, Liberman, Erickson, Dechovitz and Handler (1977) studied the relation between transition duration and VOT. These researchers found that the 50% crossover points along the VOT dimension moved to higher values with increasing transition duration. They also concluded that VOT is not alone sufficient cue but voiced transition duration is even less adequate by itself.

Summerfield and Haggard (1977) conducted a series of three experiments to assess ;

1. The role of F_1 onset as a cue to the voiced percept.

2. Whether spectral influences on the perception of voicing was a function only of the frequency of F_1 or the distribution of energy in both F_1 and higher formants.

3. Whether a rising F_1 transition was a positive cue to voicing independent of its onset frequency.

Synthetic /cv/ syllables were used and the variables were F_1 onset frequency, F_1 transition duration and VOT, each of these were manipulated independently. They reported that the major effect of F_1 in initial voicing contrast is determined by its perceived frequency at the onset of frequency and also that a periodically excited F_1 transition is not perse, a positive cue to voicing. They also concluded than in perception, the temporal separation component of VOT and the F_1 onset frequency component may be traded one for the other. The lower the frequency of F_1 at the onset of voicing, the longer the separation interval required to produce a voiceless percept.

Repp (1979) conducted two experiments to demonstrate that the amplitude of aspiration noise is a cue for the distinction between voiced and voiceless syllable initial stop consonants in English and that it can be traded for VOT. Results showed that the category boundary on the VOT continuum (/da/, /ta/) was a linear function of the amplitude ratio between the aspirated and vocalic portions. Burst and aspiration amplitudes were seen to effect voicing boundary but not independently. Thus, a synthetic stop sounded more voiceless, the more aspiration there is, the perceived amount of aspiration being a function of both noise duration and amplitude.

Kluender, Lotto and Jenison (1992) studied the effect of change in stimulus intensity and change in frequency difference between F_1 and F_2 on the perception of voicing. In a series of experiments several continua of synthesized CV's varying in VOT Mere played to listeners at levels ranging from 40 to 80 dBSPL. The frequency difference between F_1 and F_2 was also manipulated. Results showed that subjects labeled more CV's as voiceless as a function of increasing stimulus level and of decreasing $F_1 - F_2$ frequency difference. There was also an interaction between $F_1 - F_2$ frequency difference and stimulus intensity. Such that the effect of intensity was greater for smaller $F_1 - F_2$ differences. The results indicate presence of synchrony encoding for stop consonants.

The result of these studies indicate that VOT cues voicing and place of articulation. Also, it appears that VOT is language dependant. Though Usha Rani (1989) studied the perceptual cueing of VOT in Kannada, the stop consonants in her study were in the medial position of the words and it was to be expected that VOT did not cue voicing. In the present study the cueing of VOT is investigated with reference to stop consonants in the initial position of meaningful Kannada words.

METHODOLOGY

CHAPTER III

METHODOLOGY

Material:

Three plosives - voiced unaspirated velar /g/, voiced unaspirated alveolar /d/ and voiced unaspirated bilabial /b/ - were selected for the study. Three bisyllabic meaningful Kannada words with these plosives in the initial position were considered.These words formed a minimal pair with a change from voiced plosive to voiceless plosive. Table I shows the words.

Key	Phoneme	Bisyllabic Word Selected	
	:g:	gadi	(Kadi)
	:d:	dada	(Tada)
	:b:	badi	(Padi)

Table I : Material used for the present study

These three words were uttered into a microphone (cardiode) kept at a distance of 10 cms from the mouth, by a twenty one year old Kannada speaking normal female (experimenter). These were digitally recorded on a computer with a 12 bit ADC at a sampling frequency of 10 KHz. THe digitized waveform was displayed on the screen of the computer by using the program DWSSLC developed by

VSS (Voice and Speech Systems, Bangalore). Voice onset time was measured for each plosive from the waveform. Voice onset time was defined as the duration between the onset of the release of the stop consonant and the onset of the glottal pulse.

Using the waveform Editor, DWSSLC, voicing pulses were truncated in steps of three pulses (which was the closest approximation to a duration of 10 m.sec) until lead VOT was completely removed. Once pre-voicing was removed silence was added after the burst in 10 m.sec steps (the total duration of silence was equal to the duration of lag VOT for the same word as uttered by the same subject) thus approximating a voiceless plosive. A total of twelve tokens for /g/, eleven tokens for /d/ and ten tokens for /b/ were synthesized. Table II shows the original lead VOT and the values for the subsequent synthetic tokens. Fig 2 to Fig. 7 show the synthetic tokens.

TOKENS	GADI	DADA	BADI
1 (Original)	- 96	- 88.8	- 70
2	- 82	- 76.4	- 50
3	- 68	- 59.6	- 38.4
4	- 57	- 44.0	- 22.4
5	- 41	- 28.0	0
6	- 27	0	+ 10
7	0	+ 10.0	+ 20
8	+ 10	+ 20.0	+ 30
9	+ 20	+ 30.0	+ 40
10	+ 30	+ 40.0	+ 50
r-l r-l	+ 40	+ 50.0	
12	+ 50	-	

Table II VOT values for synthetic tokens (in m.sec)

Each word with its synthetic token was considered as a test and for each test the tokens were randomized and iterated ten times. These 330 tokens were audio-recorded on magnetic cassettes with an interstimulus interval of one second, which formed the material.

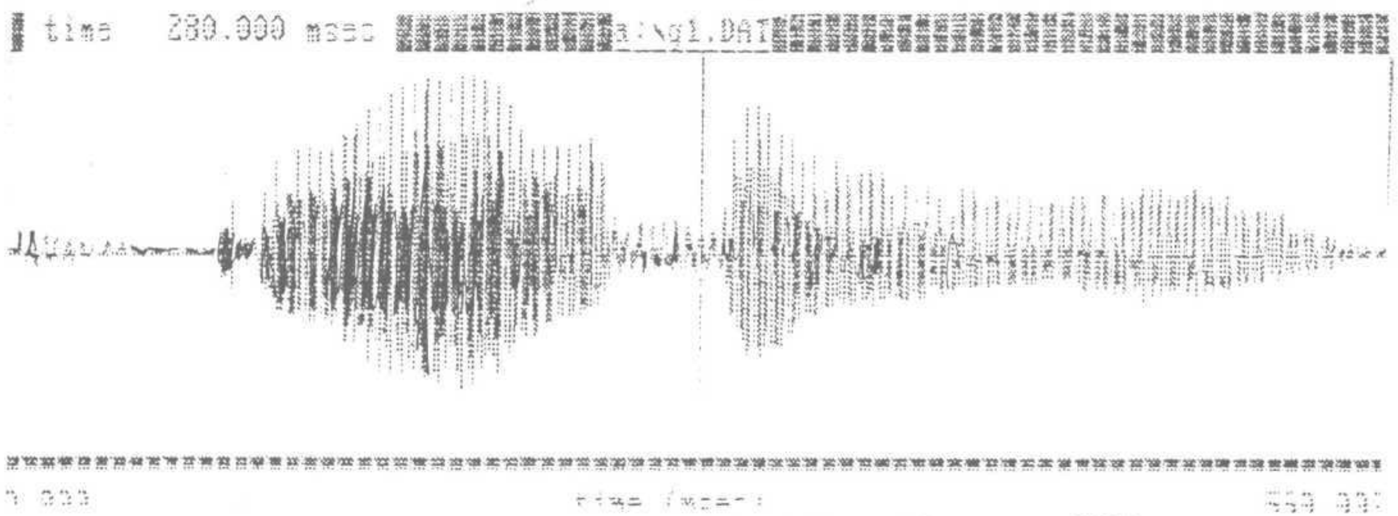

Fig 2. Synthesized /Gadi/ with - 82 m.sec VOT

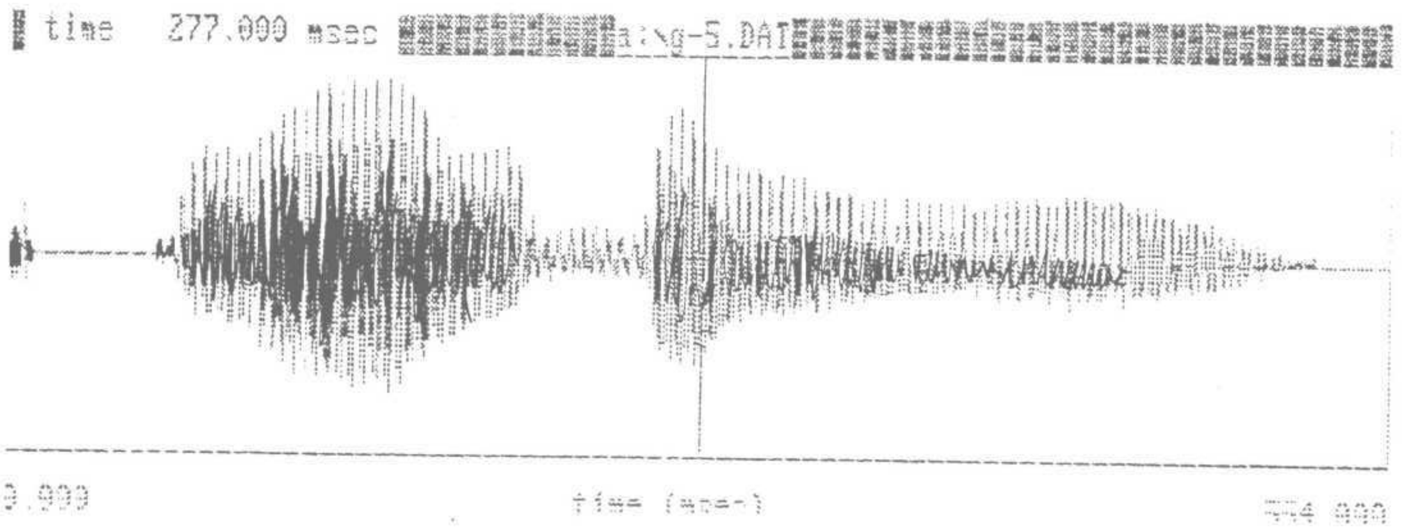

Fig 3. Synthesized /Gadi/ with + 50 m.sec VOT

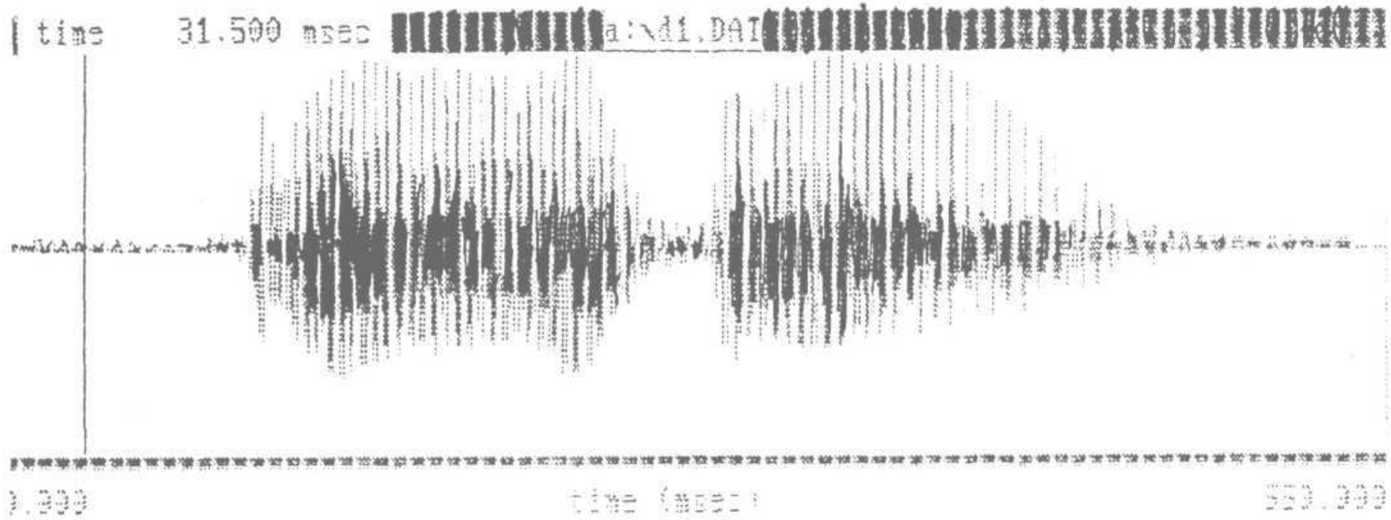

Fig 4. Synthesized /Dada/ with - 76.4 m.sec VOT

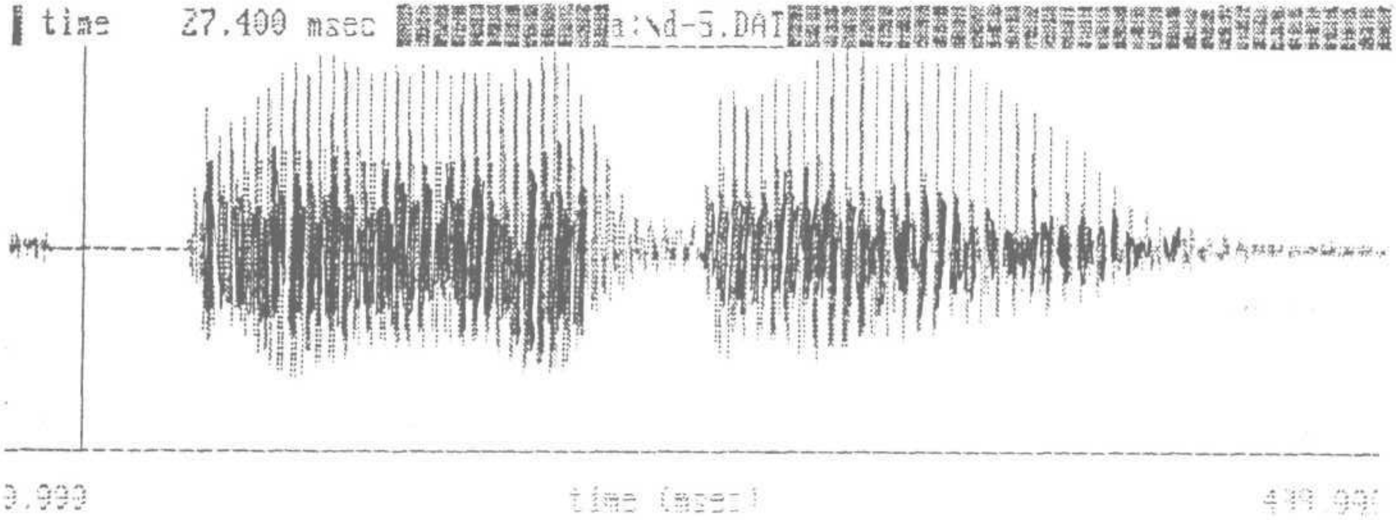

Fig 5. Synthesized /Dada/ with + 50 m.sec VOT

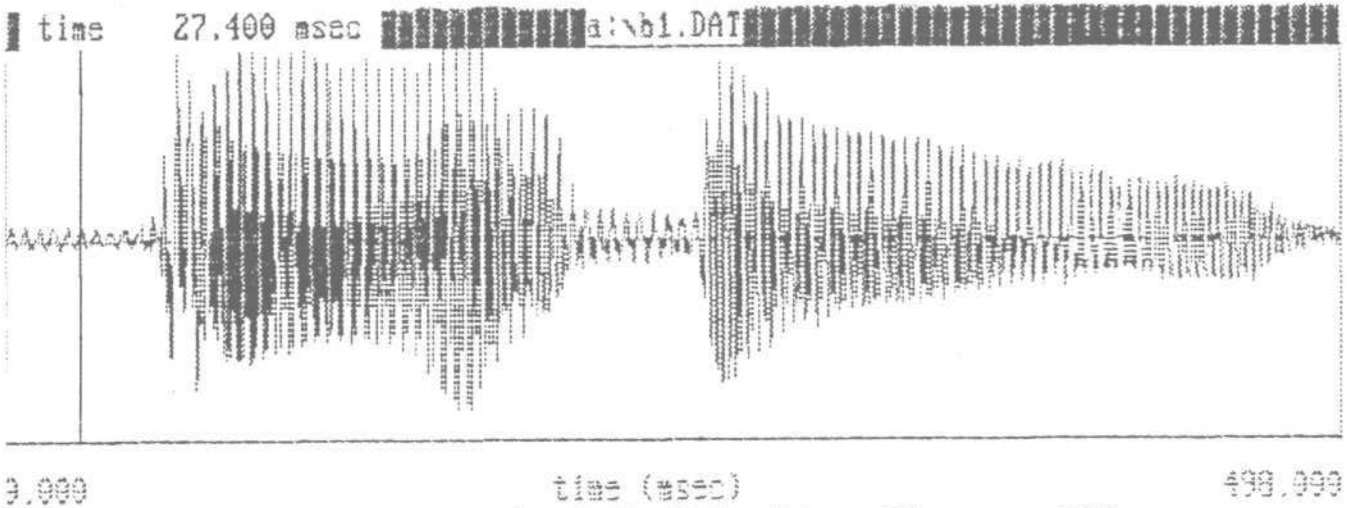

Fig 6. Synthesized /Badi/ with - 50 m.sec VOT

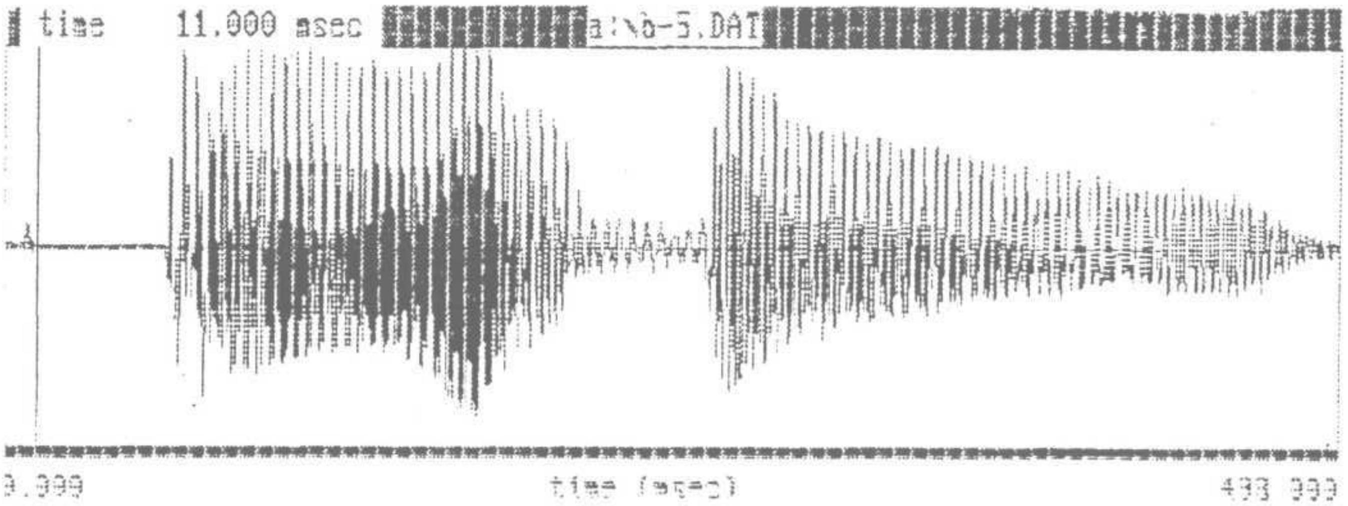

Fig 7. Synthesized /Badi/ with + 50 m.sec VOT

Subjects:

Twenty Kannada speaking normal adults (ten males and ten females) in the age range of 18 - 35 years served as subjects, all of them had normal speech and hearing. Males ranged from 20 to 35 years with a mean age of 24.4 years and females ranged from 18 to 35 years with a mean age of 27.6 years.

Method:

The subjects were tested individually in a quiet room and the stimuli were audio-presented through earphones at a comfortable listening level. An alternate forced-choice identification task was used. The subjects were instructed to listen to the tokens carefully and mark (*) for the identified phoneme on a response sheet provided.

Analysis:

The data thus obtained was tabulated and percentage response for the stimulus was calculated by the following formula.

$$\text{Percent response} = \frac{\text{Obtained number of responses}}{\text{Expected number of responses}} \times 100$$

For example, if the total or expected number of response for a stimulus was 10 and the obtained number of response was 4 then the percentage response was

$$\frac{4}{10} \times 100 = 40\%$$

The percent response for voiced and voiceless plosives were tabulated for each of the test stimulus on the basis of which the identification and discrimination functions for each plosive were plotted.

Four measurements were obtained from the identification function (modified from Lisker and Abramson (1964), originally given for VOT and Doughty, (1949).

1. Lower limit of the phoneme boundary width - was that point along the VOT continuum where an individual identified voiced stop 75% of the time. For eg: in Figure 8 value of the lower limit Vis - 34 m.sec of voice onset time i.e, point A on the x-axis.

2. 50% crossover - It was that point on the graph which was the actual or interpolated point about the VOT continuum for which 50% of the subjects response corresponded to the voicing category, eg. in figure 50% crossover will be - 10 m.sec of VOT i.e, point B' on the X-axis.

3. Upper limit of phoneme boundary width - was defined as the corresponding point for the identification of voiceless cognate 75% of the time. For eg: in figure + 7.5 m.sec, the point C' on the x-axis is the upper limit.

4. Phoneme boundary width (in m.sec) between voicing category was defined as the arc boundary cross point along the VOT continuum and was determined by subtracting the lower limit

from the upper limit, eg : In figure 8, the phoneme boundary width is 41.5 m.sec.

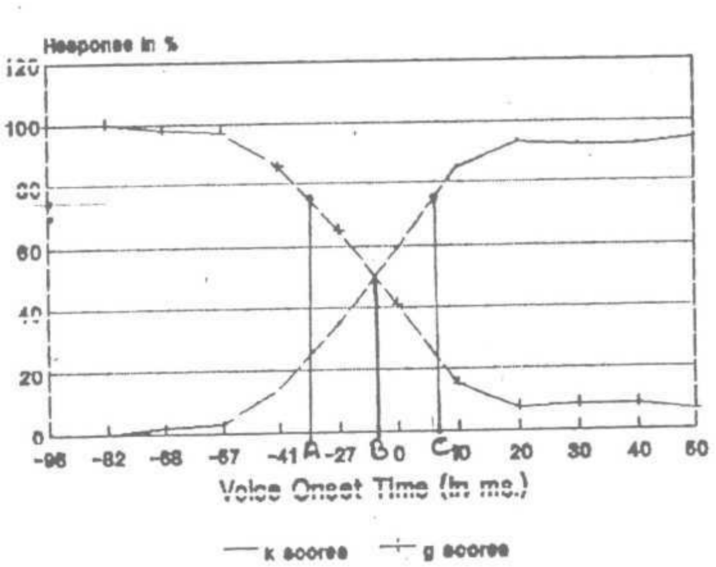

Figure 8 depicts the showing lower limit, upper limit, 50% crossover and phoneme boundary width.

These measurements were used to analyse the effect of VOT on the perception of stop consonants in Kannada.

RESULTS & DISCUSSION

CHAPTER IV

RESULTS AND DISCUSSION

RESULTS:

I. Lower limit of the phoneme boundary:

In general, the lower limit of the phoneme boundary was in the lead VOT range and the lower limit was longest for bilabial followed by alveolar and velar. When males and females were compared, females had longer limits for /k-g/ and /t-d/ percepts. The lower limit decreased as place of articulation moved back in the oral cavity. Table III depicts the lower limit of the phoneme boundary.

PERCEPT	FEMALES	MALES	AVERAGE
k - g	– 30.5	– 34	- 32.25
t - d	– 30.0	– 34	- 32.00
p - b	– 24.4	– 20.4	- 22.40

Table III : Lower limit of the phoneme boundary (in m.sec)

II. 50% cross over:

The 50% cross over from voiced stop to voiceless stop occurred in the lead VOT range. Shift occurred earlier for /t-d/ percept when compared to the others and for /k-g/ percept the shift occurred at a later lead VOT (- 6.75 m.sec). The shift from voiced to voiceless percept occurred at an earlier VOT for /p-b/ in males and for /k-g/ percept in females. Table IV depicts the 50% crossover figures 9 to 17 indicate the 50% cross over for each percept.

PERCEPT	FEMALES	MALES	AVERAGE
k - g :	- 6 75	- 10.2	- 6 75
t - d :	- 14.0	- 10.5	- 14.00
P - b :	- 11.2	- 5.6	- 8.40

Table IV : 50% crossover points (in m.sec)

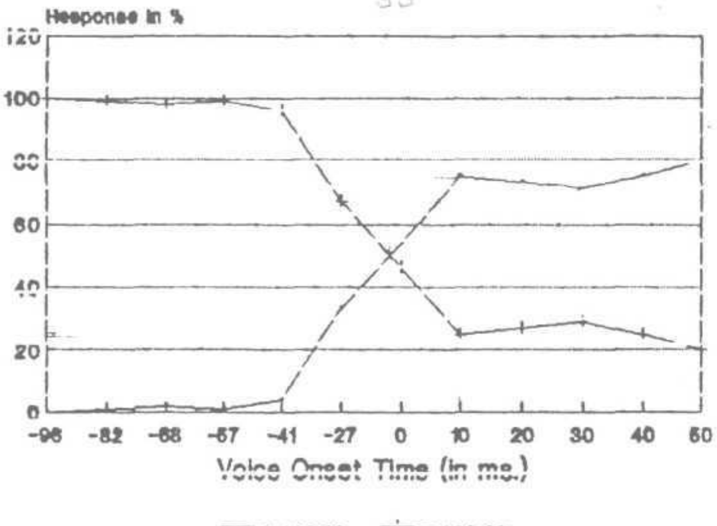

Fig 9 :- Identification discrimination function of females for /k-g/ continuum.

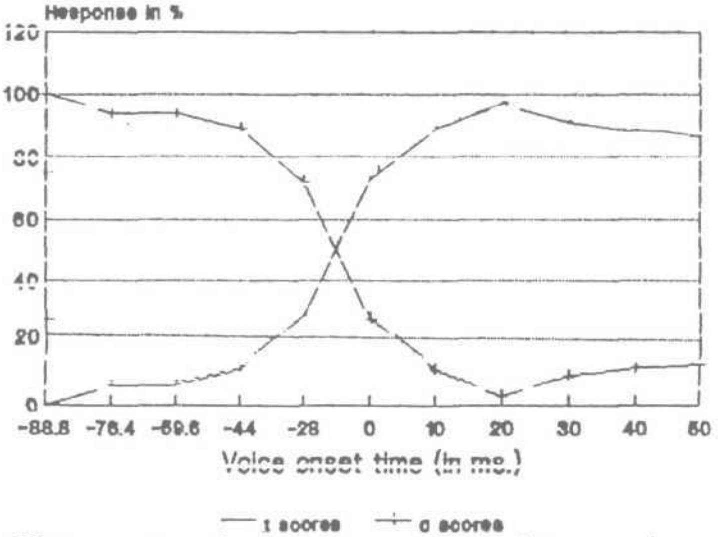

Fig 10 :- Identification discrimination function of females for /t-d/ continuum.

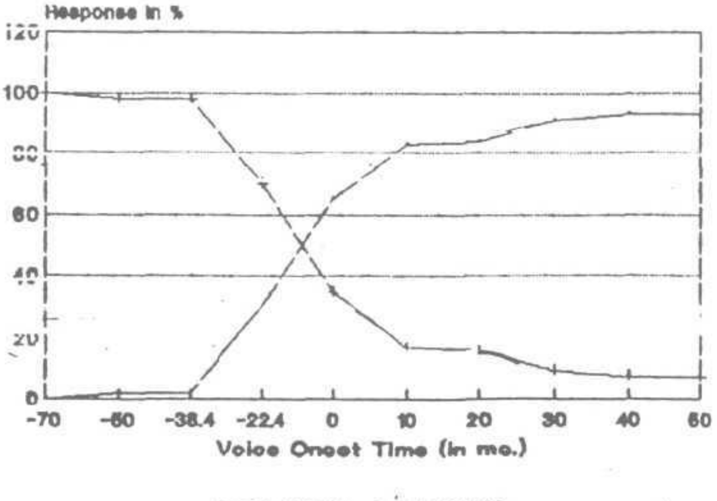

Fig 11 :- Identification discrimination function of females for /p-b/ continuum.

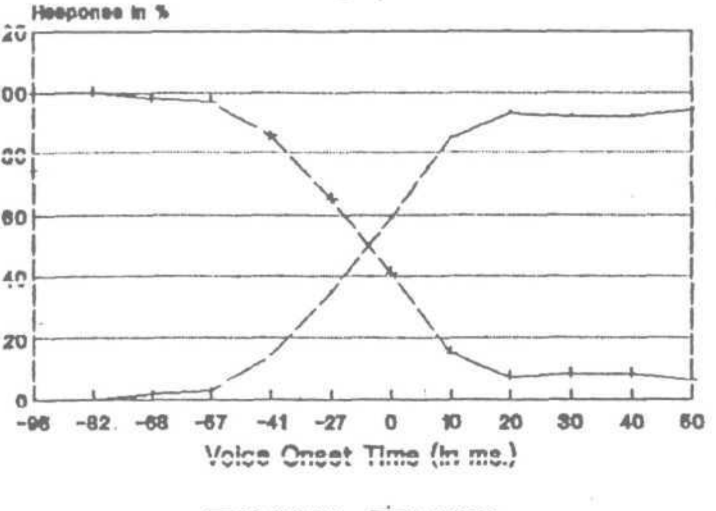

Fig 12 :- Identification-discrimination function of males for /k-g/ continuum.

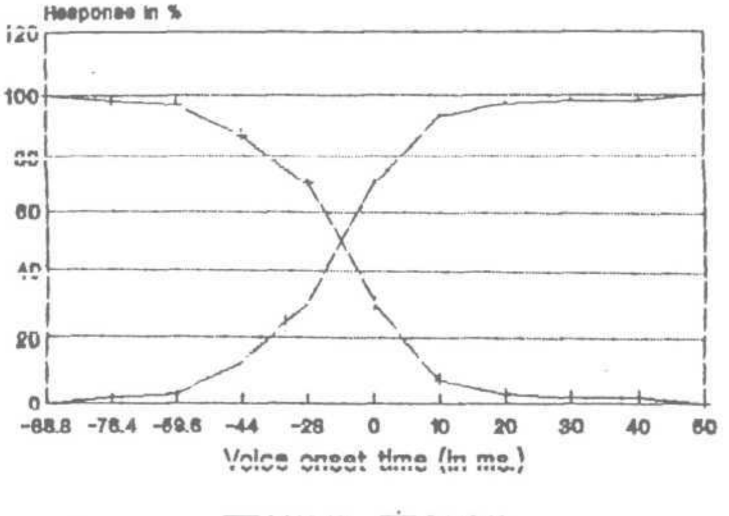

Fig 13 :- Identification-discrimination function of males for /t-d/ continuum.

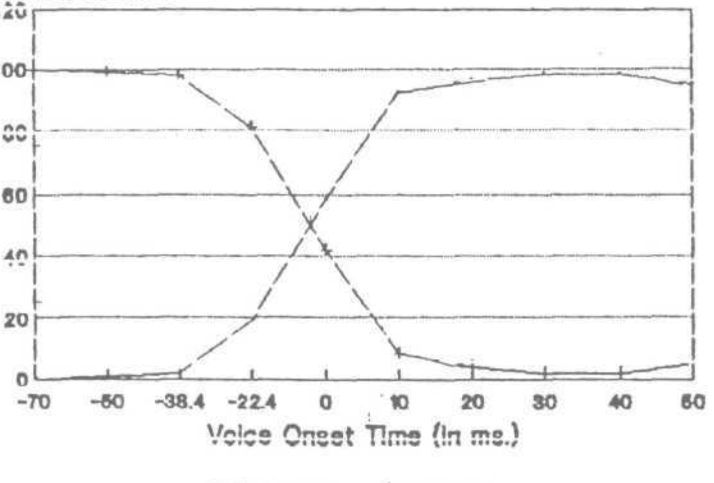

Fig 14 :- Identification - discrimination function of males for /p-b/ continuum.

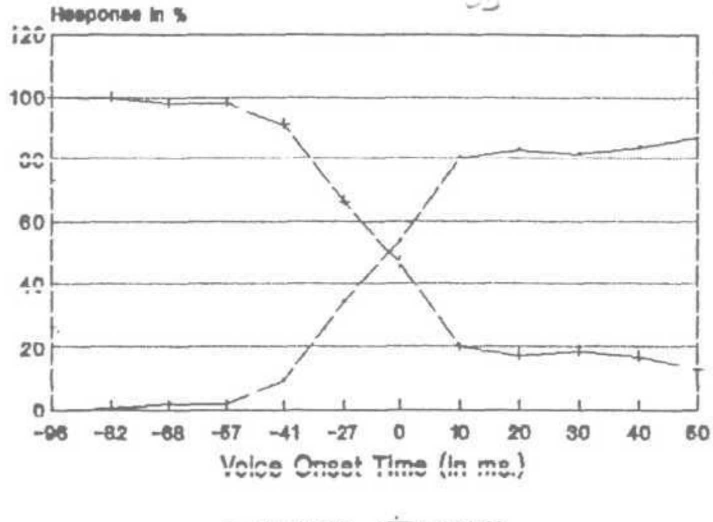

Fig 15 :- Identification-discrimination function of adults for /k-g/ continuum.

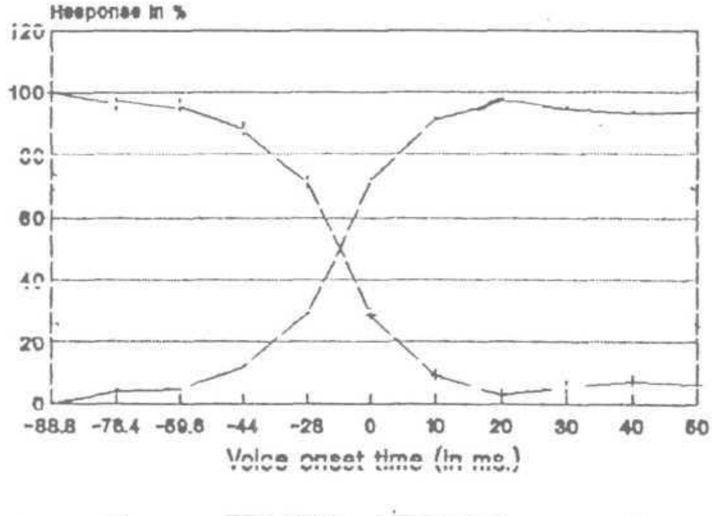

Fig 16 :- Identification - discrimination function of adults for /t-d/ continuum.

2

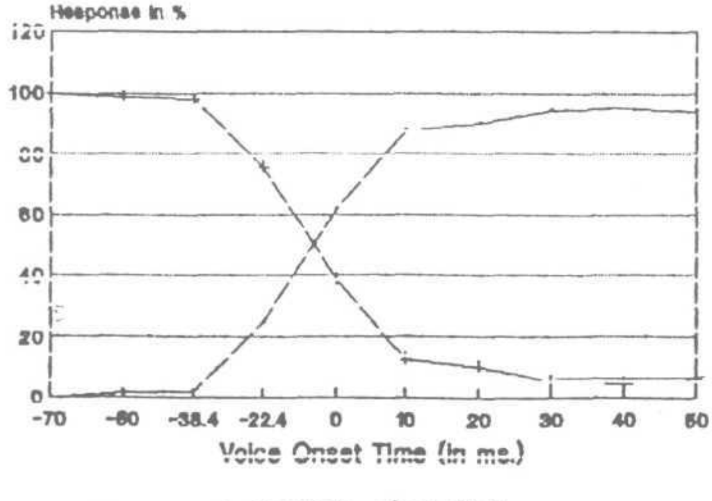

Fig 17:- Identification - discrimination function of adults for /p-b/ continuum.

III. The upper limit of the phoneme boundary:

The upper limit of the phoneme boundary was in the lag VOT range for all the phonemes.The boundary was low for /t-d/ percept and high for /k-g/ percept and the upper limit value for /p-b/ was intermediate between the two. The upper limit occurred at a longer VOT in females for /k-g/ percept when compared to males and upper limit of the male group for /t-d/ percept was longer. (Table V)

PERCEPT	FEMALES	MALES	AVERAGE
k - g	+ 10.0	+ 7.5	+ 7.50
t - d	+ 0.6	+ 2.5	+ 1.25
p - b	+ 6.25	+ 6.25	+ 6.25

Table V : Upper limit of the boundary (in m.sec)

IV. Phoneme boundary width:

The boundary was largest for velars followed by alveolars and bilabials. In males the boundary width was larger than in females for velar and alveolar percepts. (Table VI)

PERCEPT	FEMALES	MALES	AVERAGE
k - g	40.5	41.5	39.75
t - d	30.6	36.5	33.25
p - b	30.65	26.65	28.65

Table VI : Phoneme boundary width (in m.sec)

The raw scores of male and female subjects were subjected to statistical analysis. The results indicated that there was significant difference between the responses of males and females for /k-g/ percept (T = 3.81, P = 0.0002) and no significant differences were found for /t-d/ percept (T = 0.64, P = 0.52)and /p-b/ percept (T = 0.95, P = 0.34).

V. Individual Responses:

Individual variations were seen in the perceptual judgements of the various subjects, figures 18 to 37 represent the individual responses for /k-g/ percept, figures 38 to 57 represent the individual responses for /t-d/ percept and figures 58 to 77 are the responses for /p-b/ percept. For /k-g/ percept fourteen subjects showed good perpetual judgement (single crossover) and six showed poor perceptual judgement (multiple crossovers). For the /t-d/ percept nineteen subjects showed good perceptual judgement and one showed poor perceptual judgement. For /p-b/ percept none of the subjects showed multiple cross overs. These findings may reflect practice and task familiarity effect as the /k-g/ tokens were presented first arid the /p-b/ tokens were presented later. For /k-g/ percept seven subjects located the crossover in the lag region, one had crossover at coincident VOT and twelve subjects had the crossover in the lead VOT region. For /t-d/ percept five subjects had

Identification discrimination functions of individual subjects [S] for /k-g/ continuum

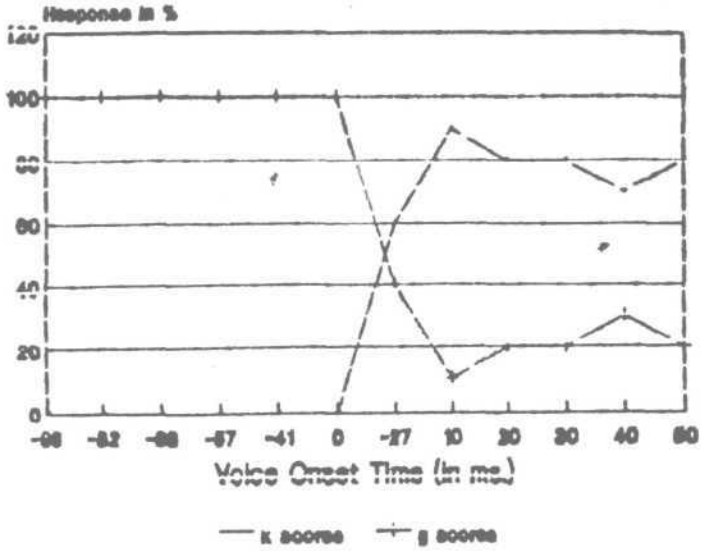

Fig 18. S_1's response

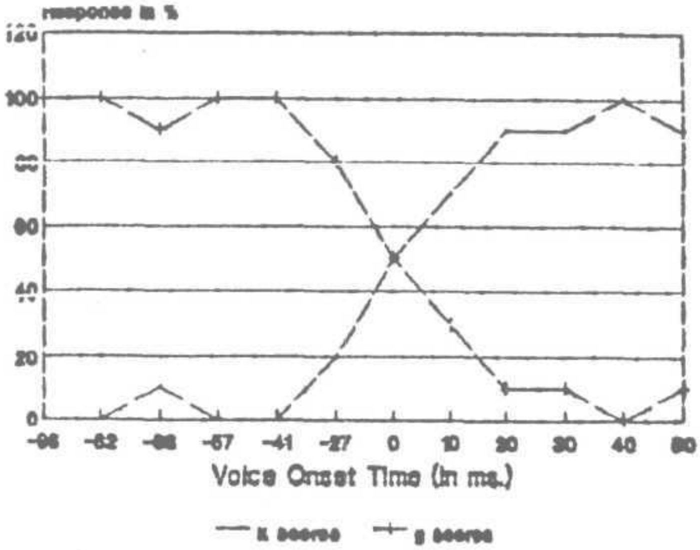

Fig 19. S_2's response

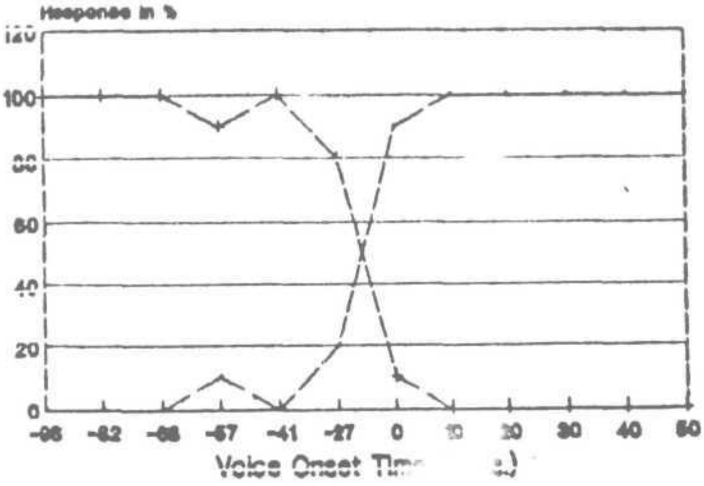

Fig 20. S_3's response

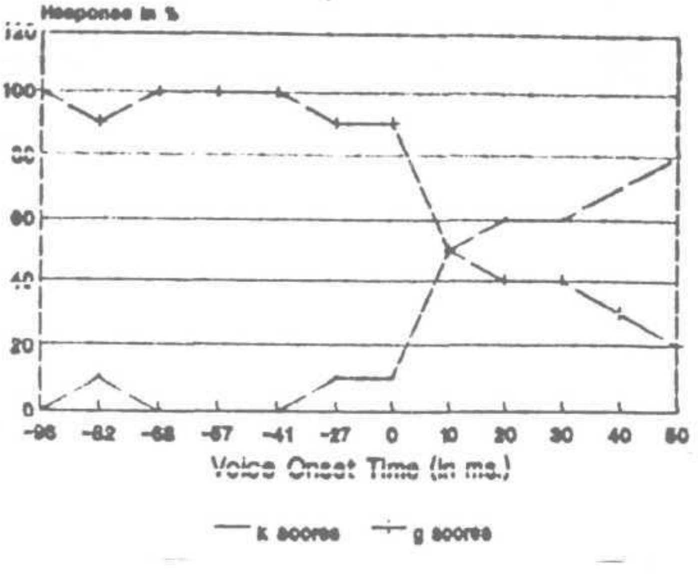

Fig 21. S_4's response

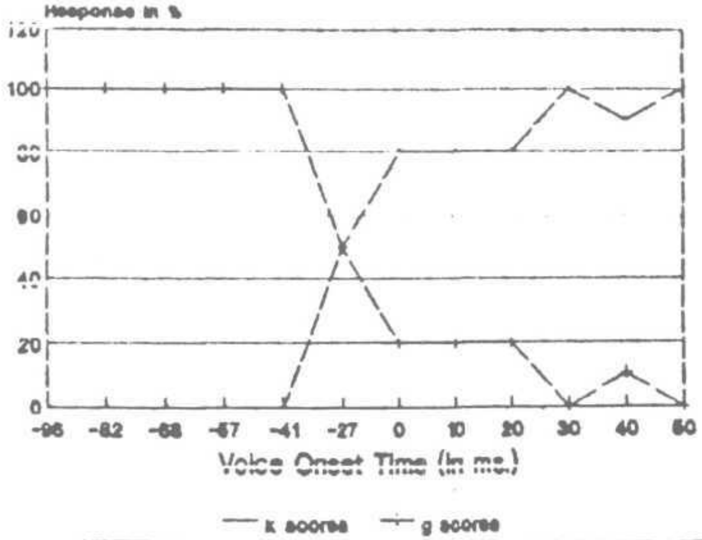

Fig 22. S_5's response

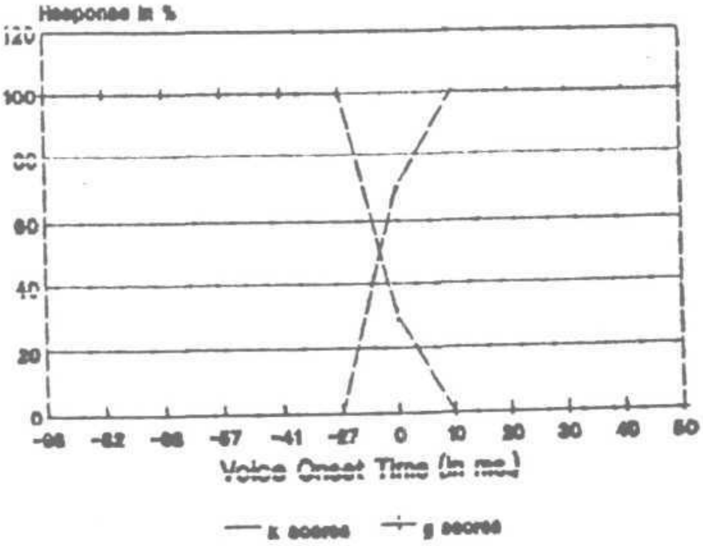

Fig 23. S_6's response

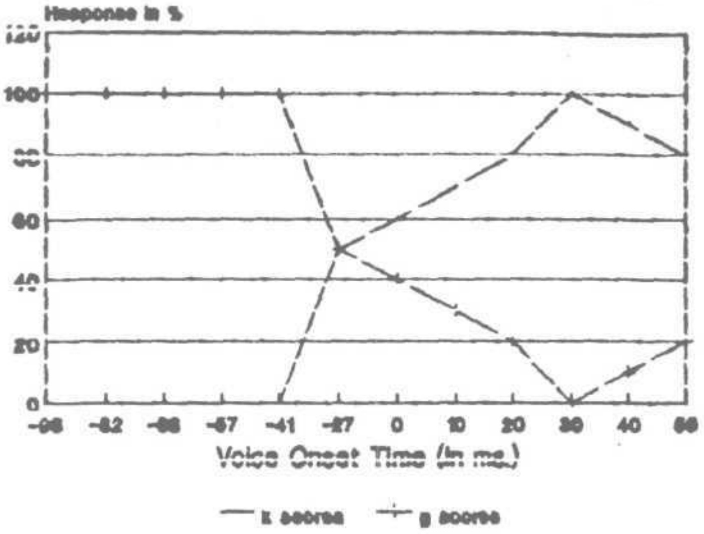

Fig 24. S_7's response

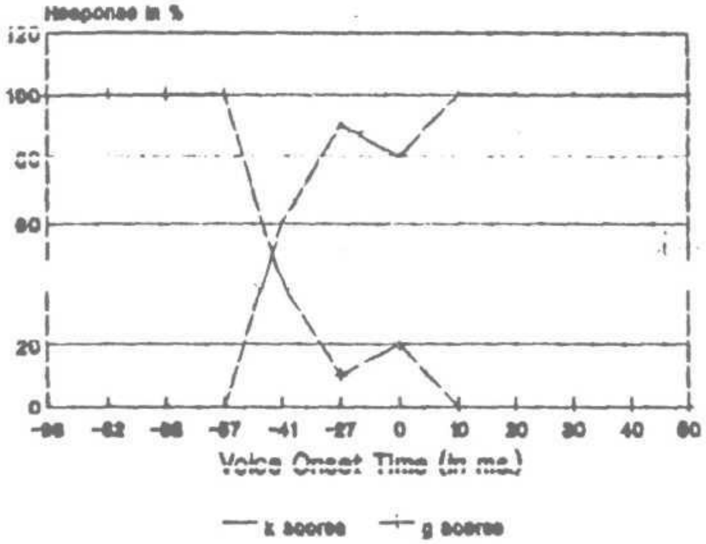

Fig 25. S_8's response

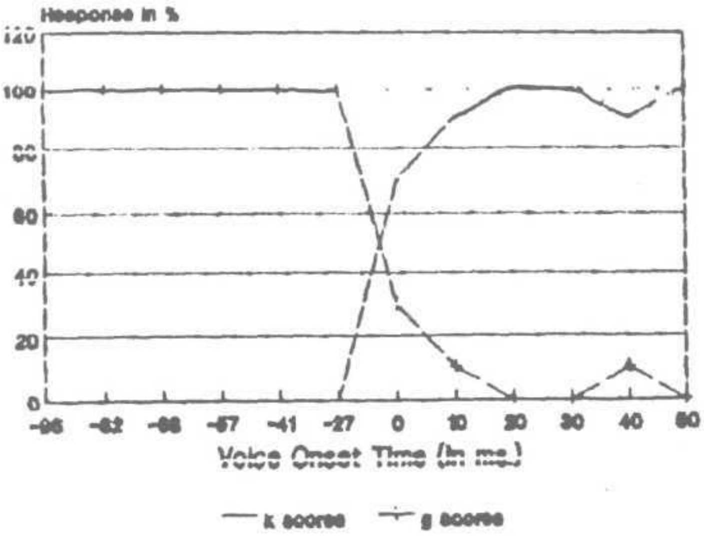

Fig 26. S_9's response

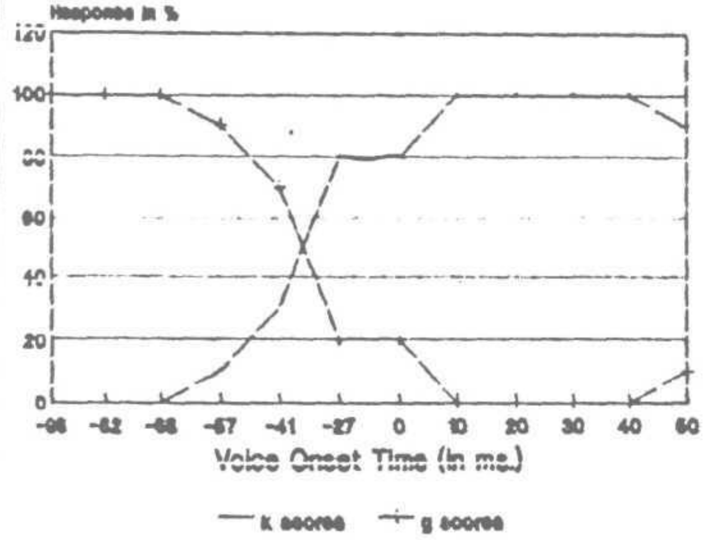

Fig 27. S_{10}'s response

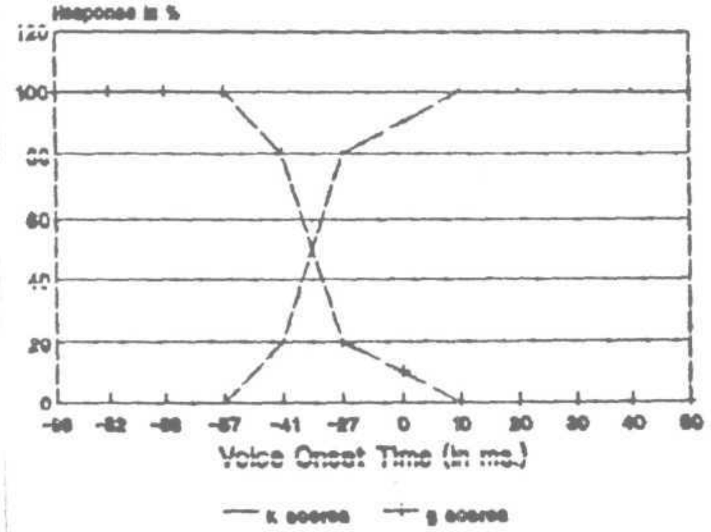

Fig 28. S_{11}'s response

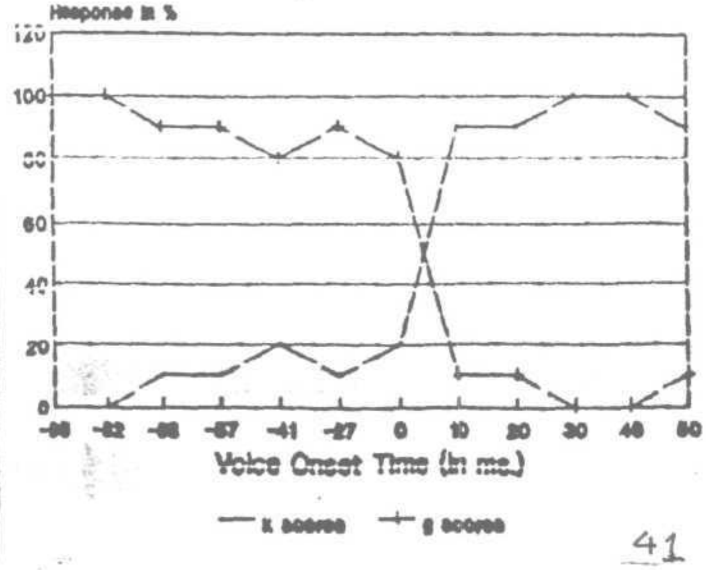

Fig 29. S_{12}'s response

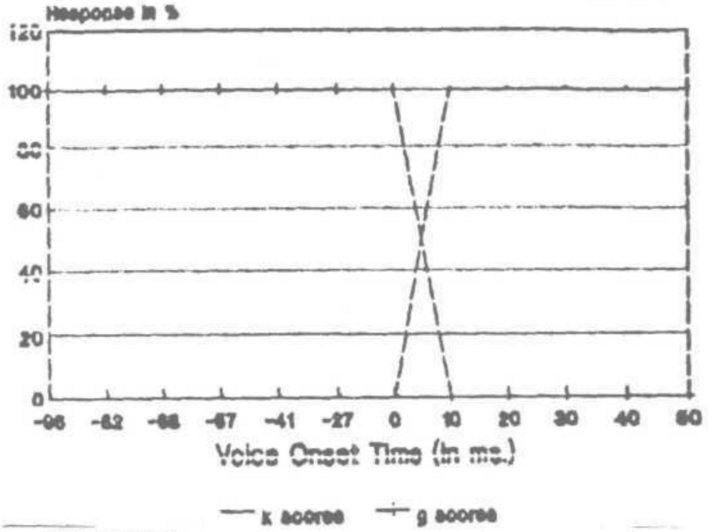

Fig 30. S_{13}'s response

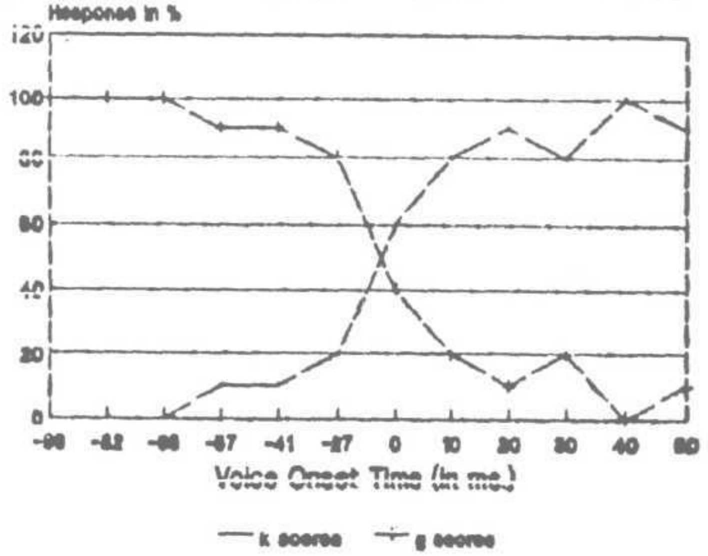

Fig 31. S_{14}'s response

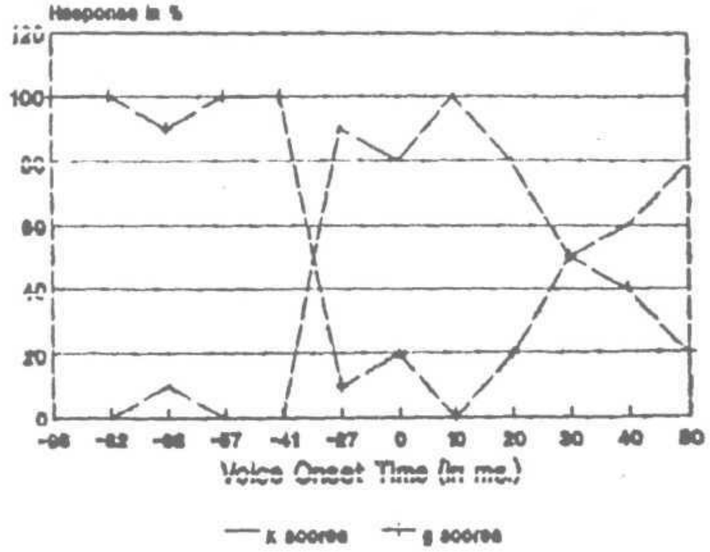

Fig 32. S_{15}'s response

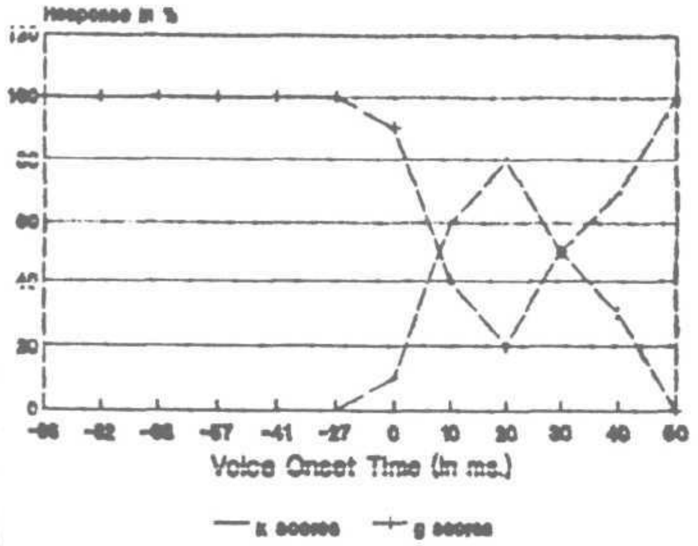

Fig 33. S_{16}'s response

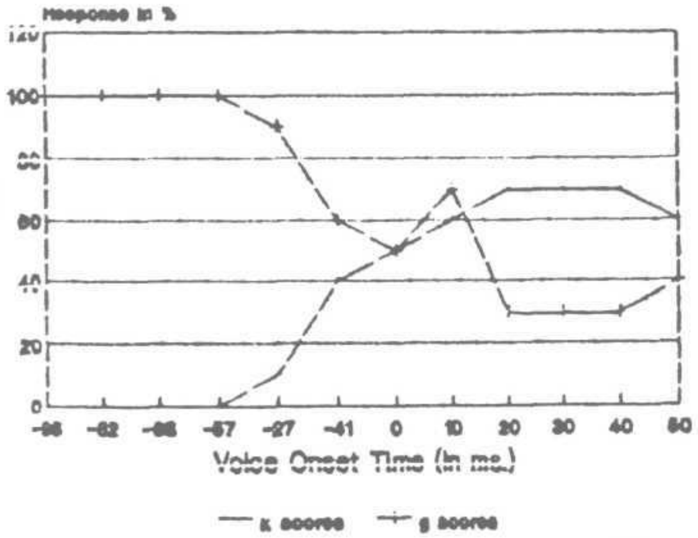

Fig 34. S_{17}'s response

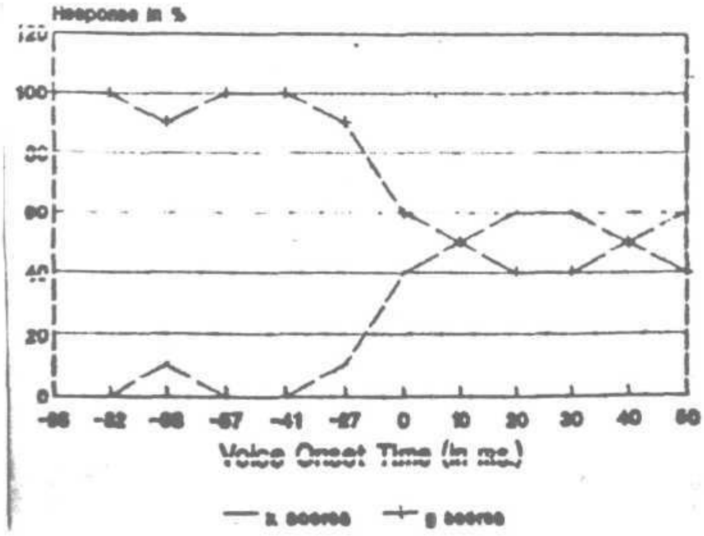

Fig 35. S_{18}'s response

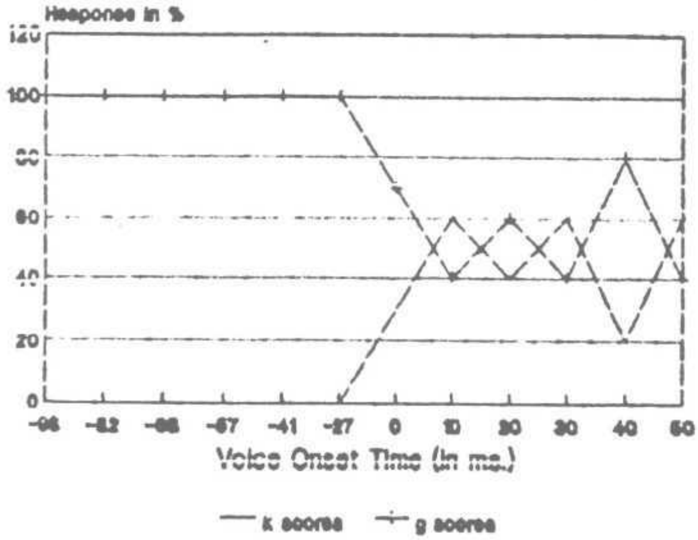

Fig 36. S_{19}'s response

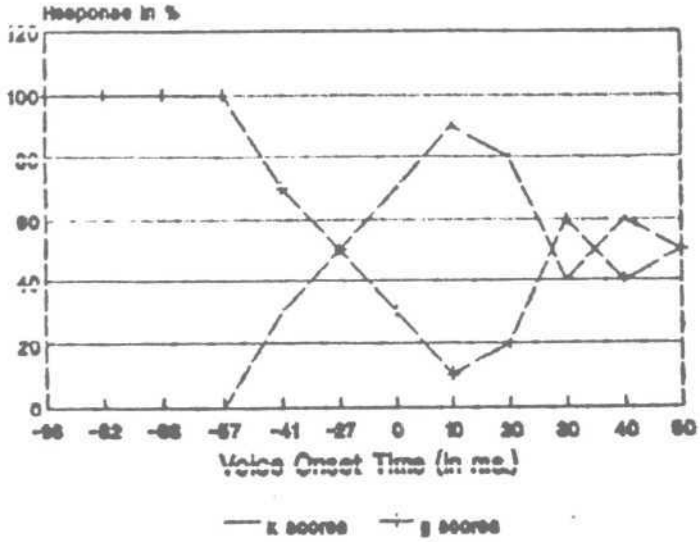

Fig 37. S_{20}'s response

Identification discrimination functions of Individual subjects [S] for /t-d/ continuum

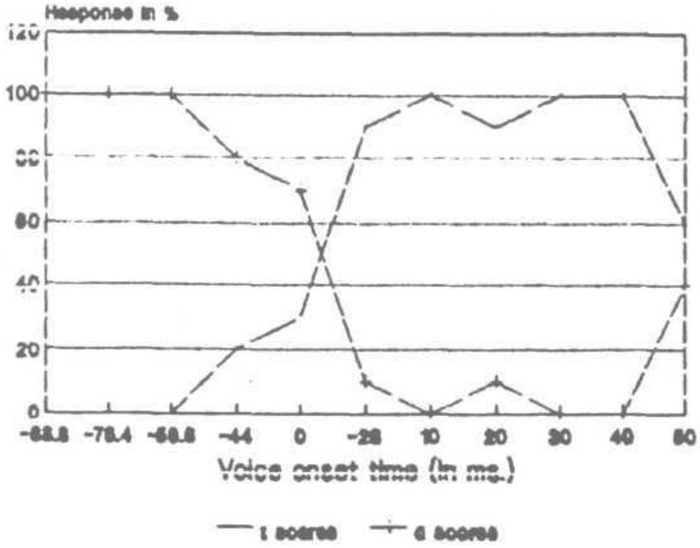

Fig 38. S_1's response

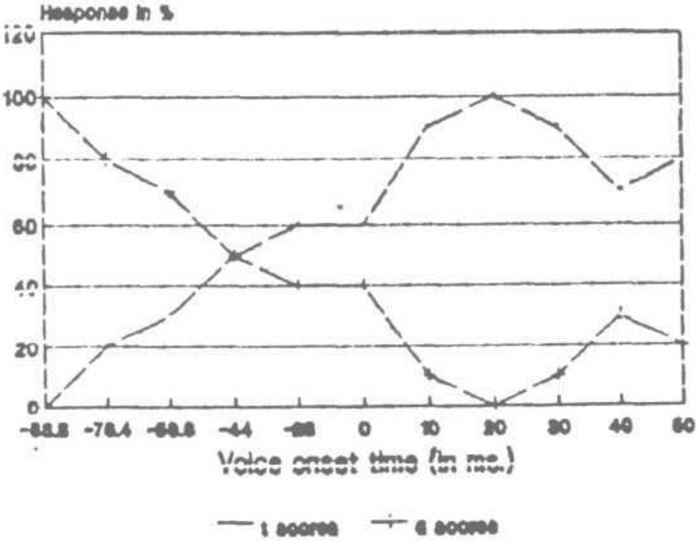

Fig 39. S_2's response

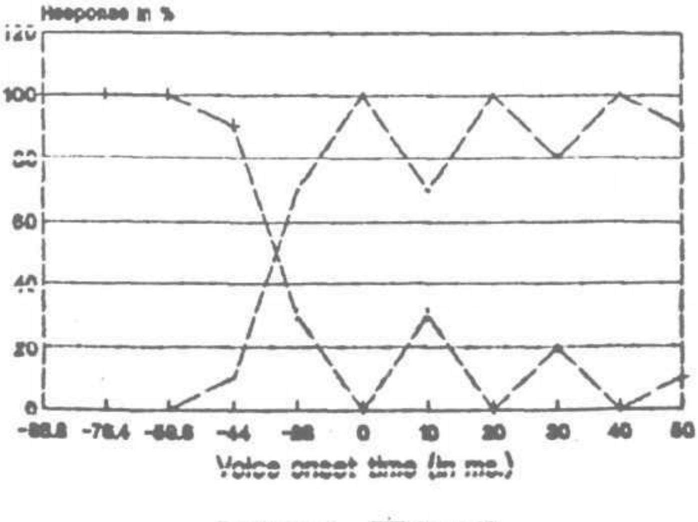

Fig 40. S_3's response

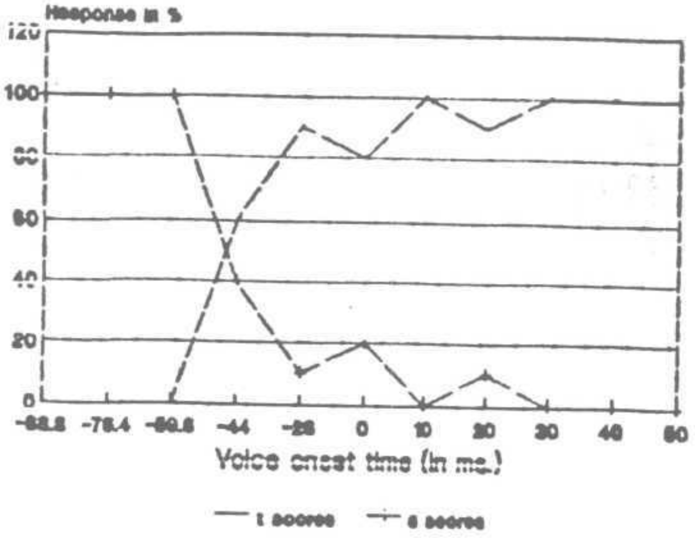

Fig 41. S_4's response

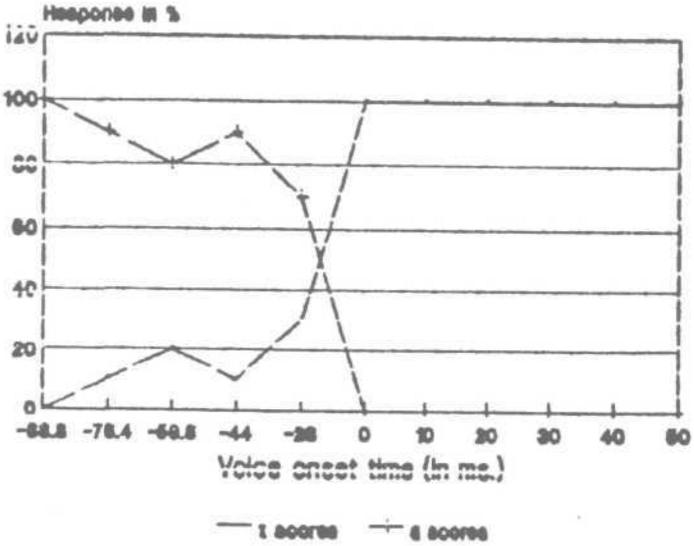

Fig 42. S_5's response

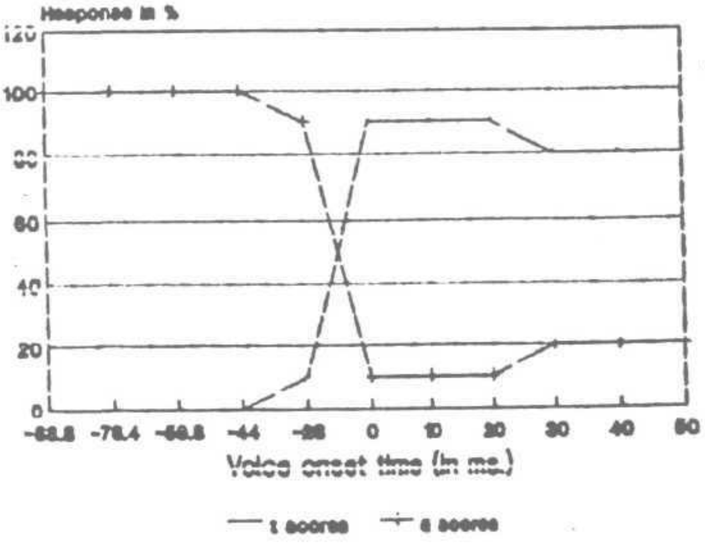

Fig 43. S_6's response

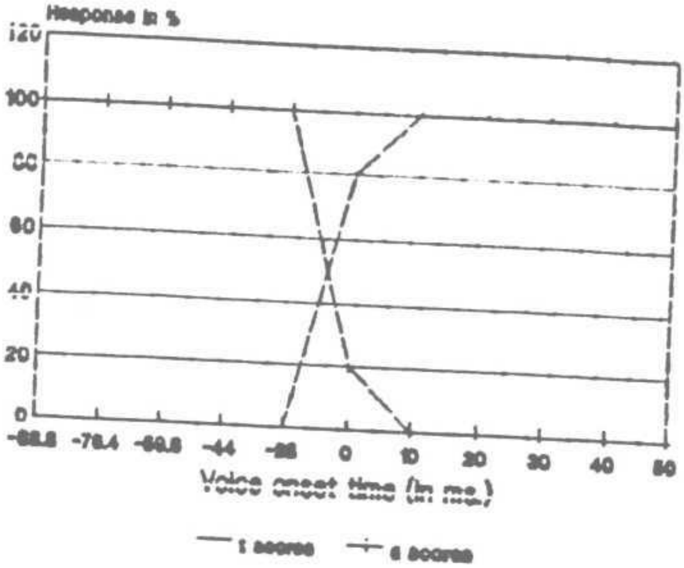

Fig 44. S_7's response

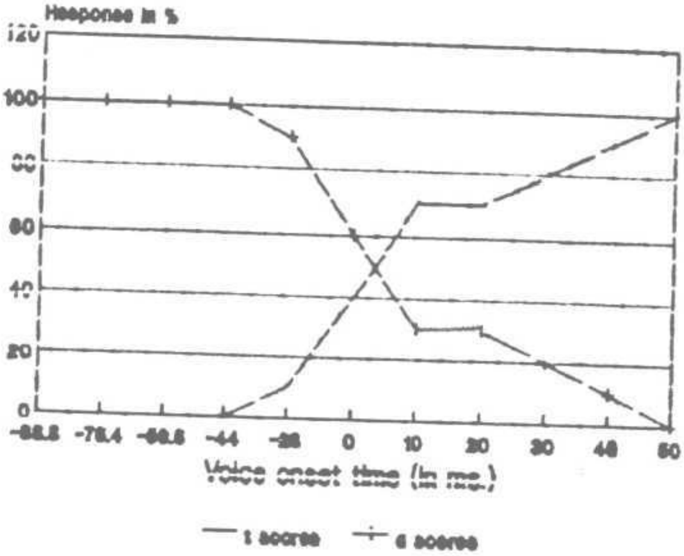

Fig 45. S_8's response

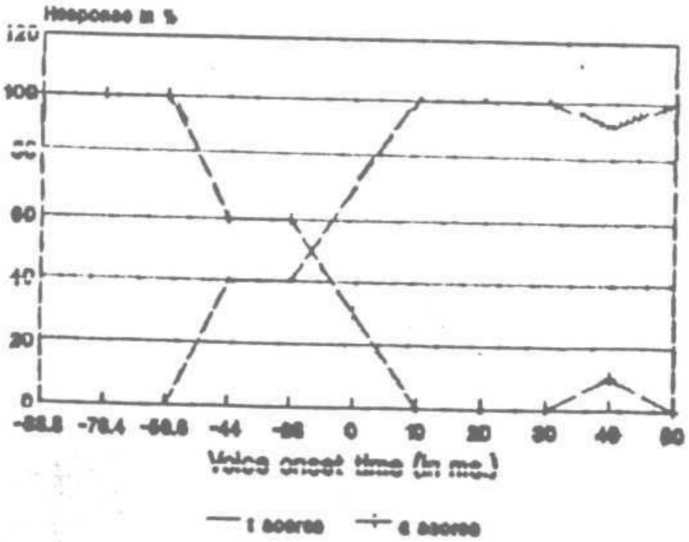

Fig 46. S_9's response

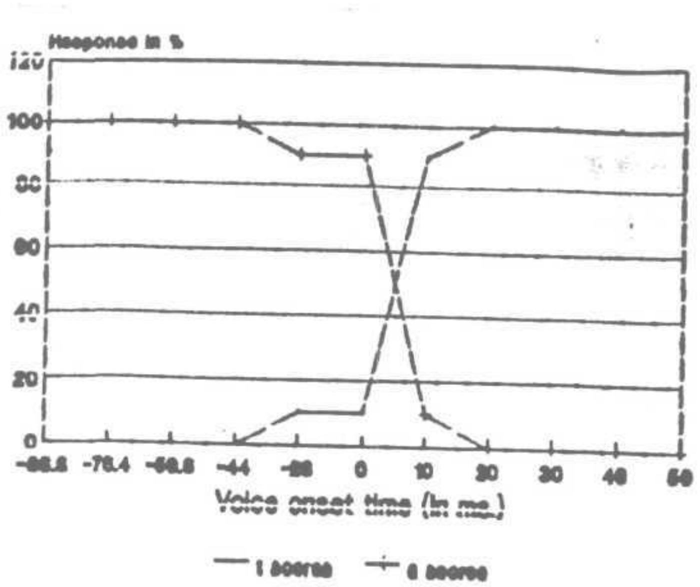

Fig 47. S_{10}'s response

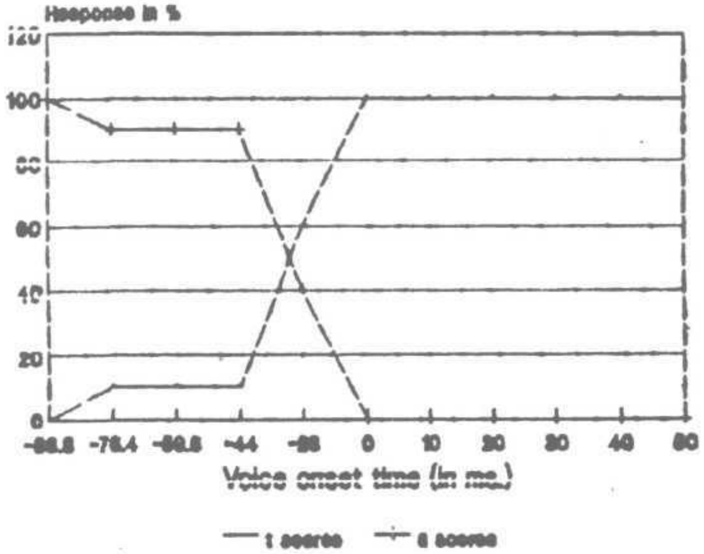

Fig 48. S_{11}'s response

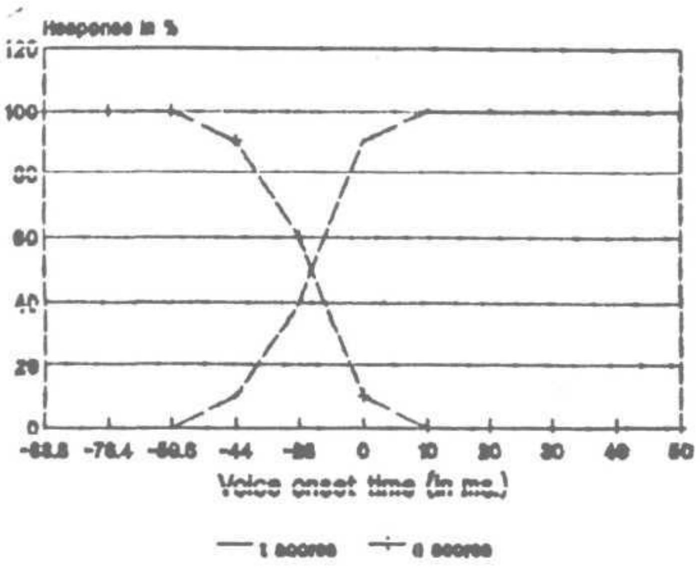

Fig 49. S_{12}'s response

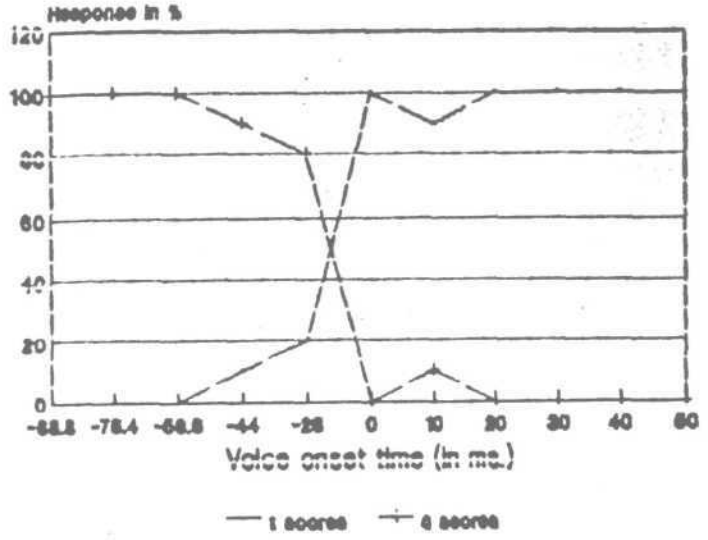

Fig 50. S_{13}'s response

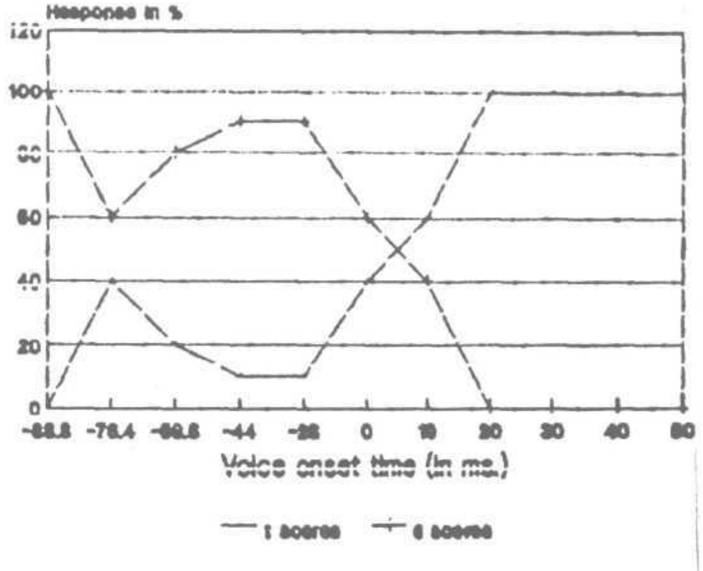

Fig 51. S_{14}'s response

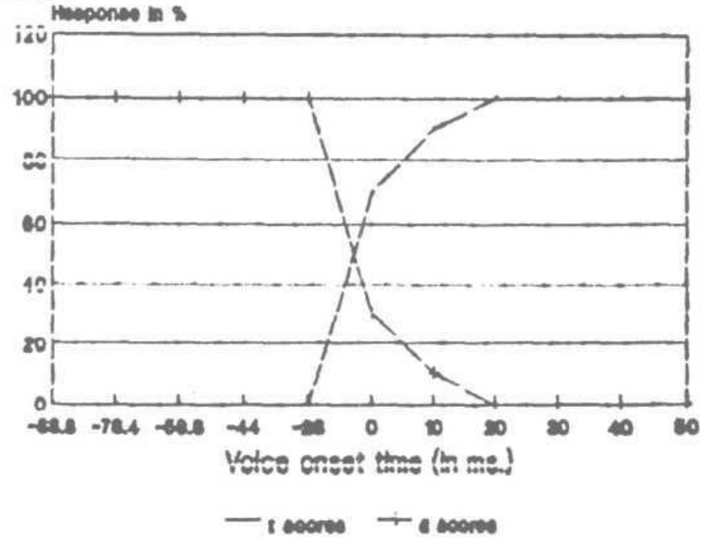

Fig 52. S_{15}'s response

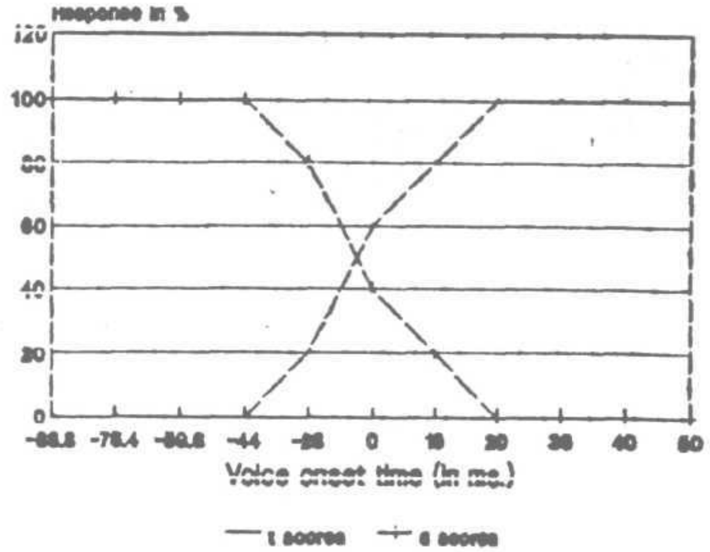

Fig 53. S_{16}'s response

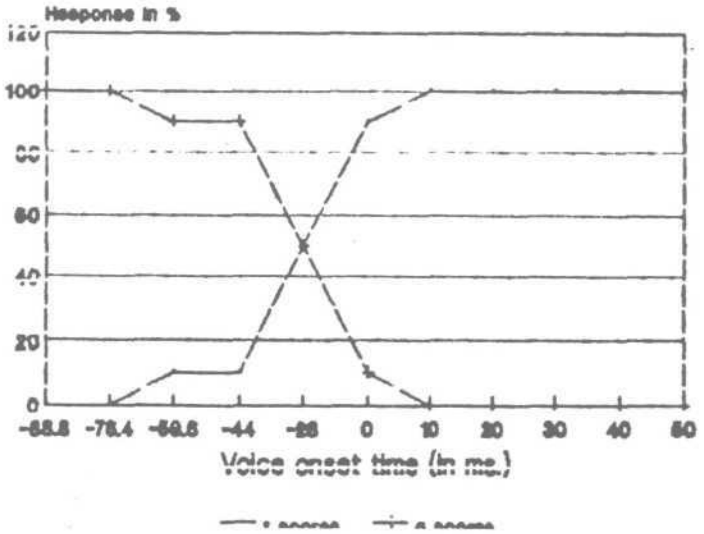

Fig 54. S_{17}'s response

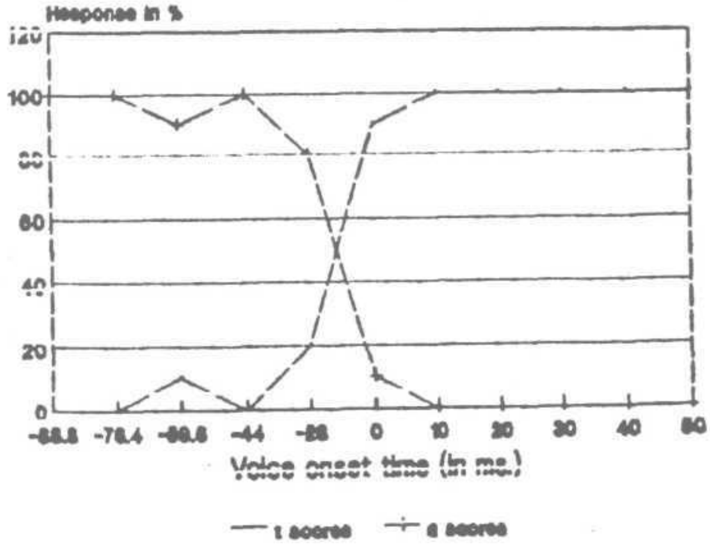

Fig 55. S_{18}'s response

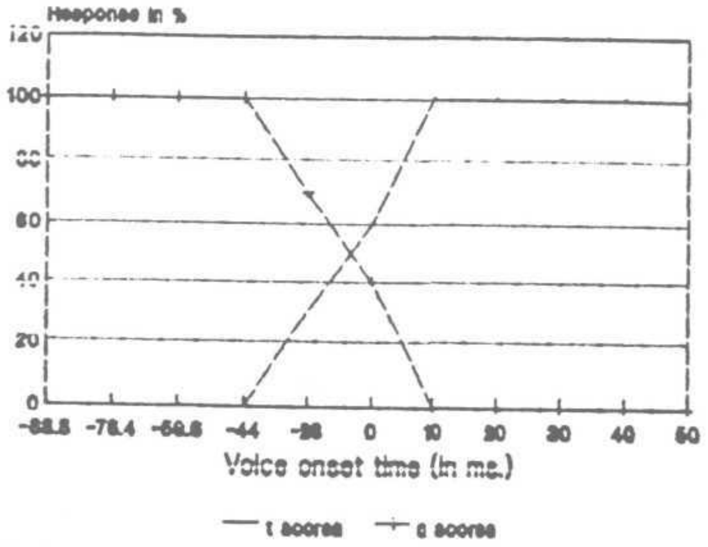

Fig 56. S_{19}'s response

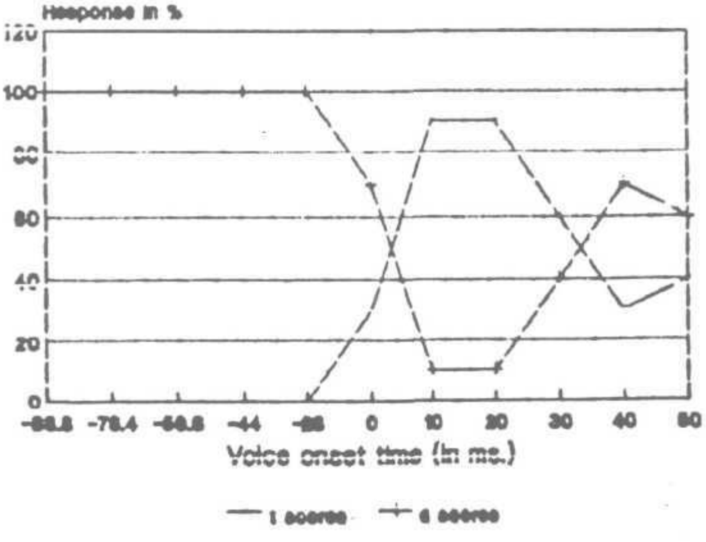

Fig 57. S_{20}'s response

Identification discrimination functioning of individual subjects [S] for /p-b/ continuum

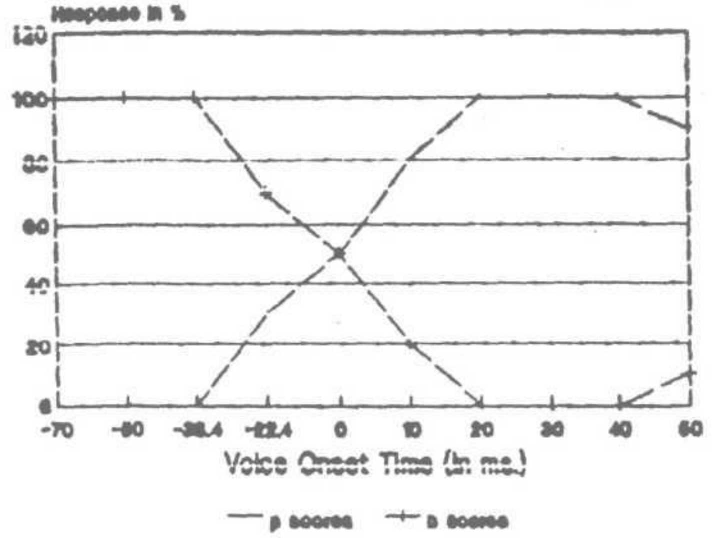

Fig 58. S_1's response

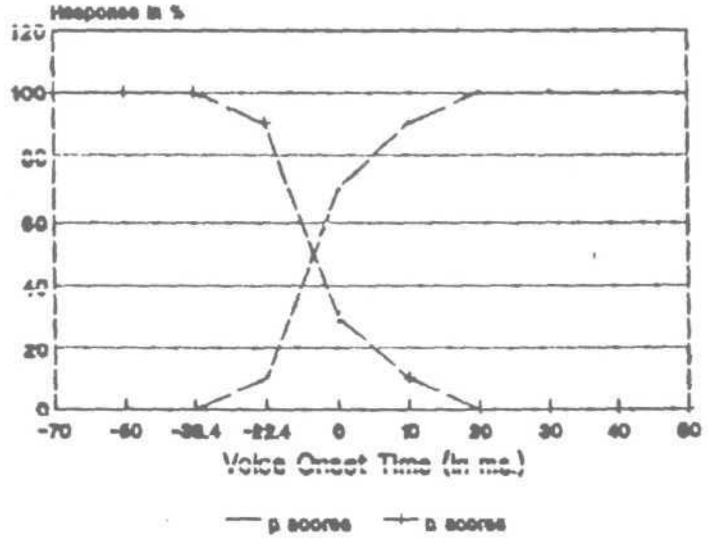

Fig 59. S_2's response

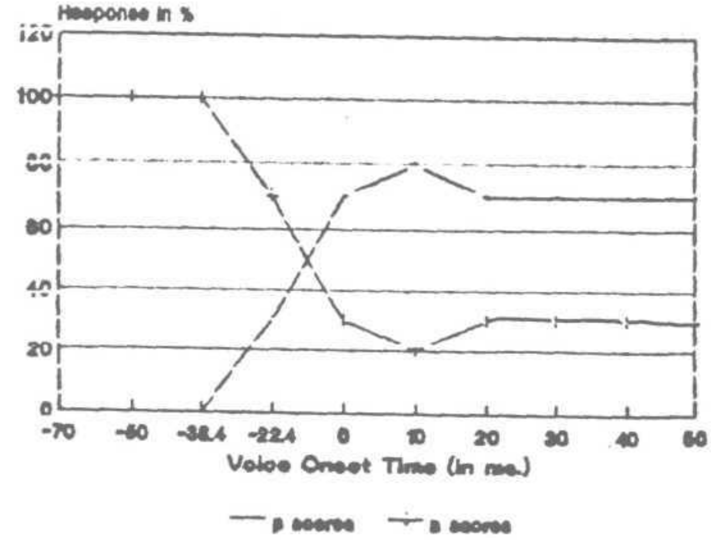

Fig 60. S_3's response

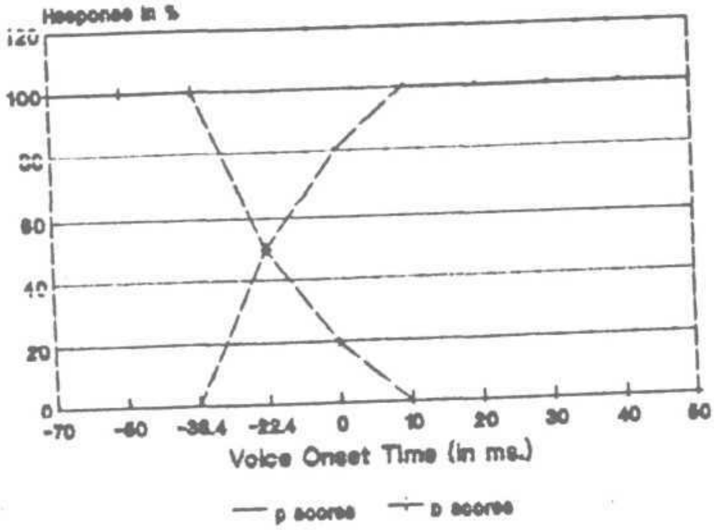

Fig 61. S_4's response

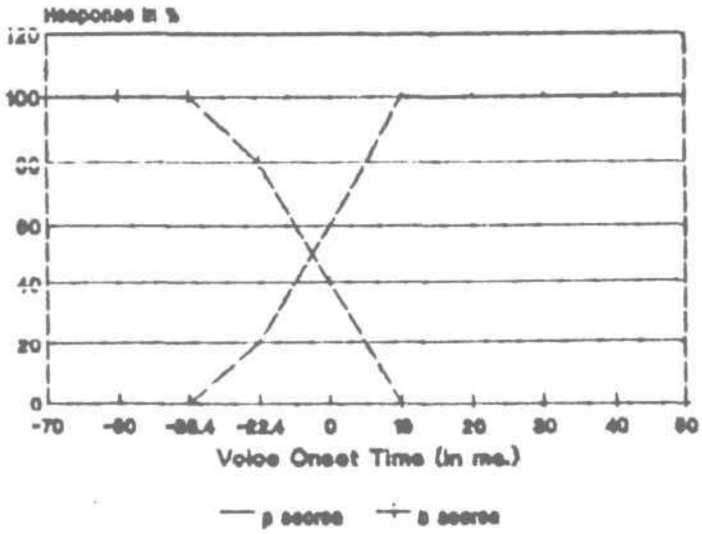

Fig 62. S_5's response

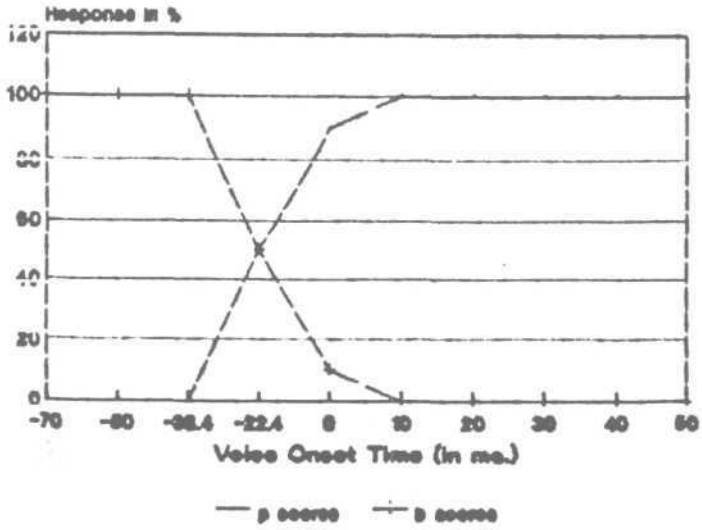

Fig 63. S_6's response

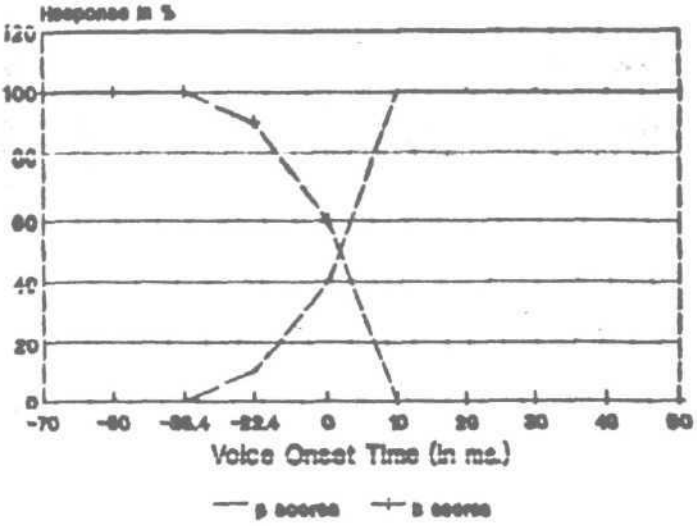

Fig 64. S_7's response

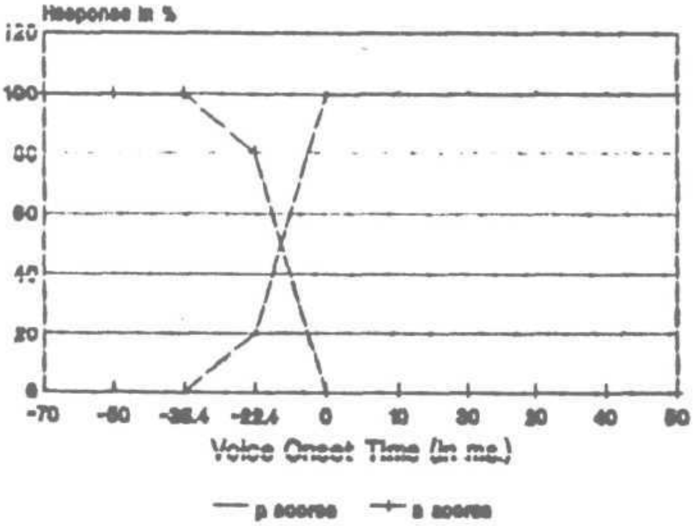

Fig 65. S_8's response

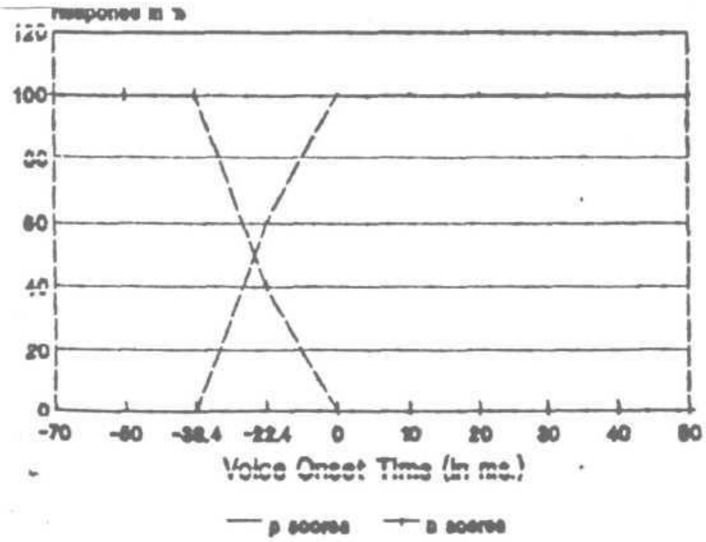

Fig 66. S_9's response

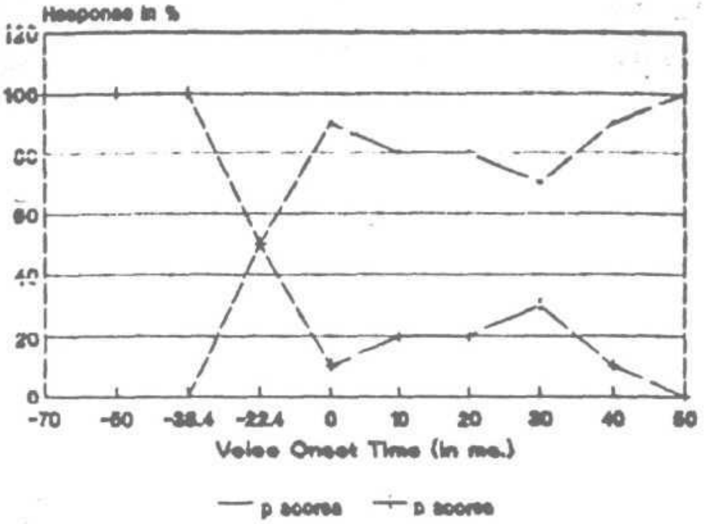

Fig 67. S_{10}'s response

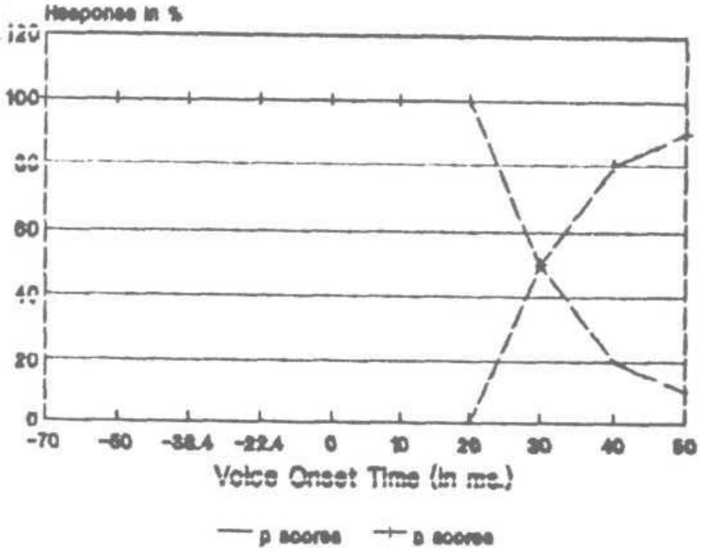

Fig 68. S_{11}'s response

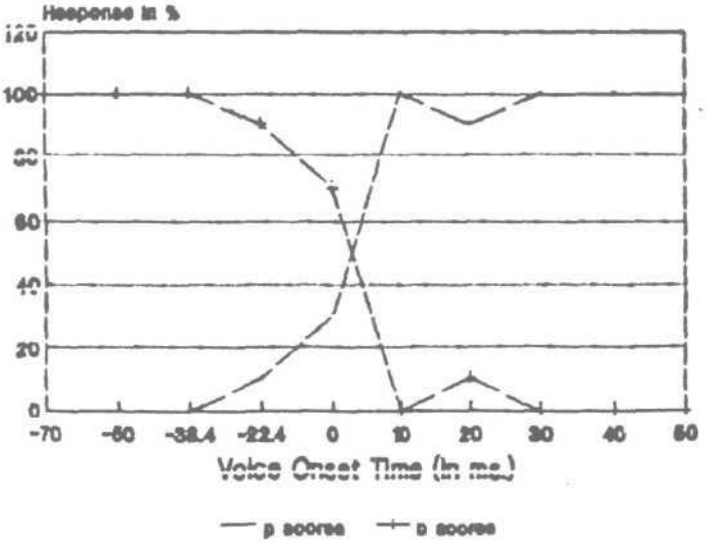

Fig 69. S_{12}'s response

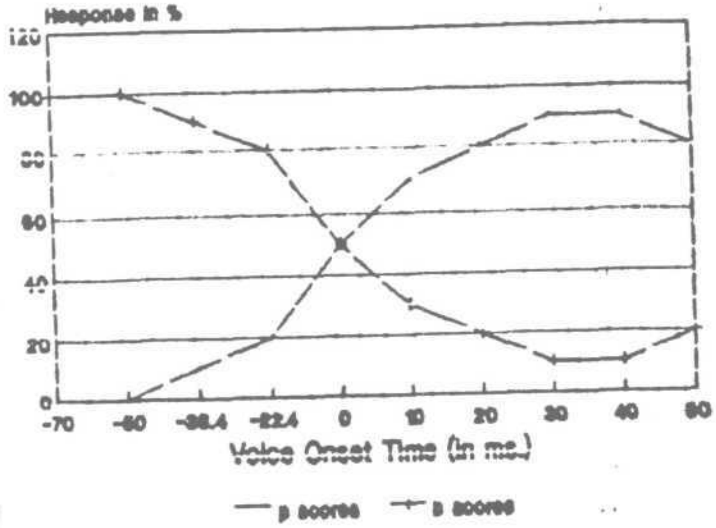

Fig 70. S_{13}'s response

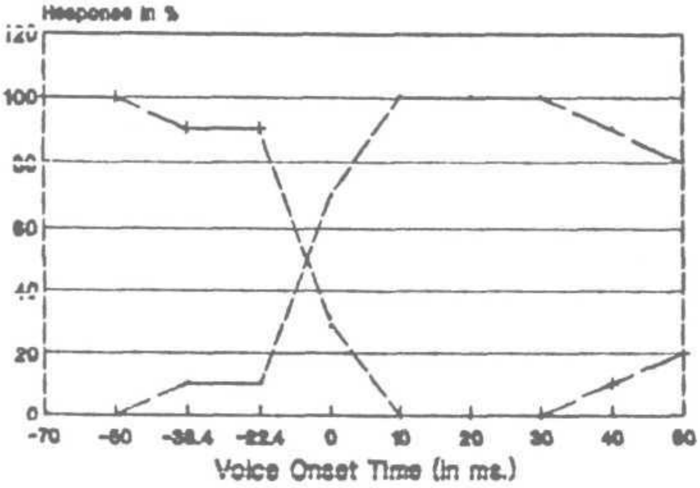

Fig 71. S_{14}'s response

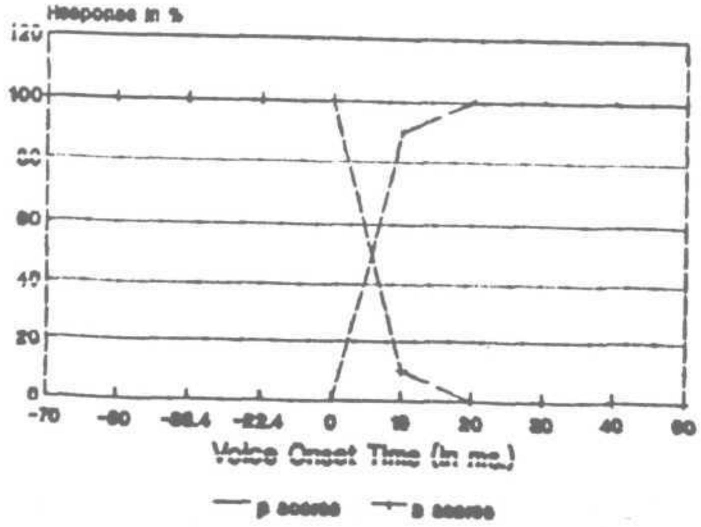

Fig 72. S_{15}'s response

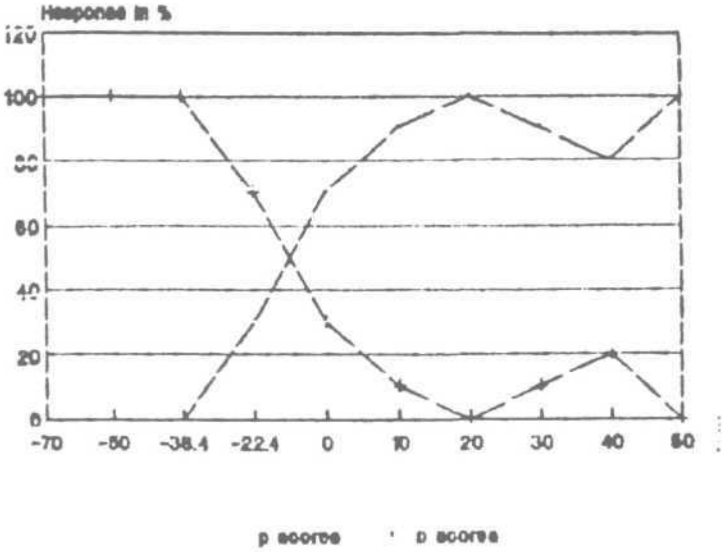

Fig 73. S_{16}'s response

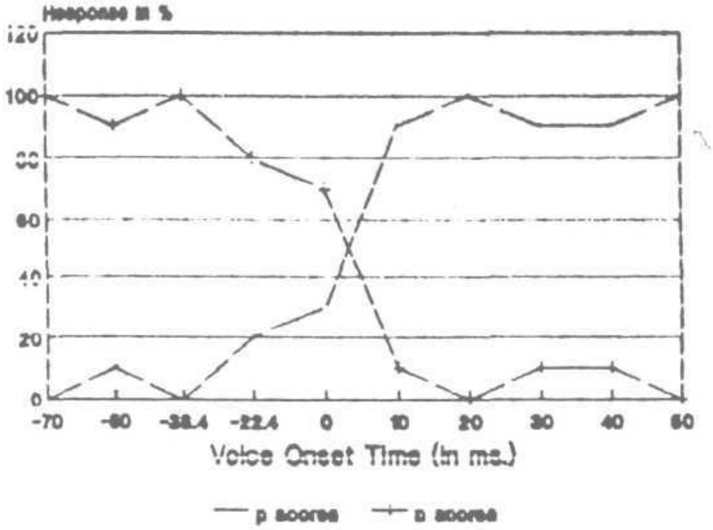

Fig 74. S_{17}'s response

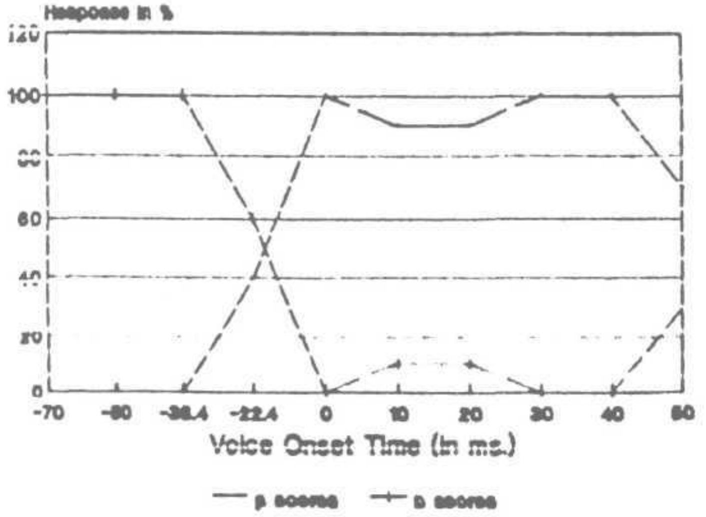

Fig 75. S_{18}'s response

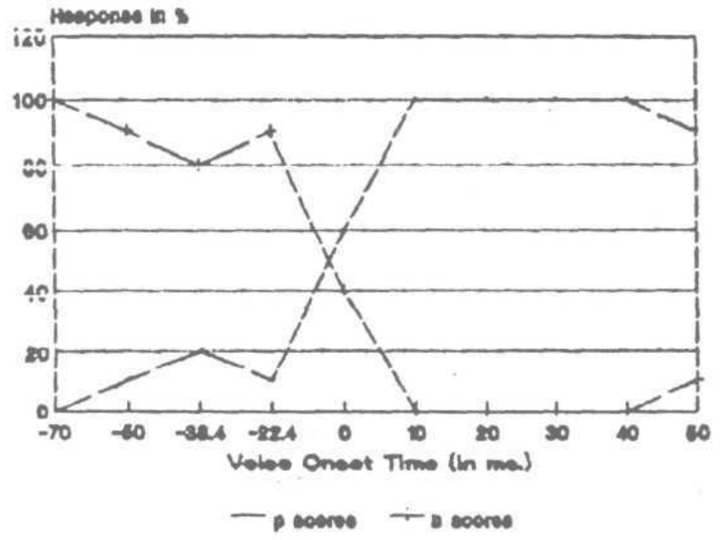

Fig 76. S_{19}'s response

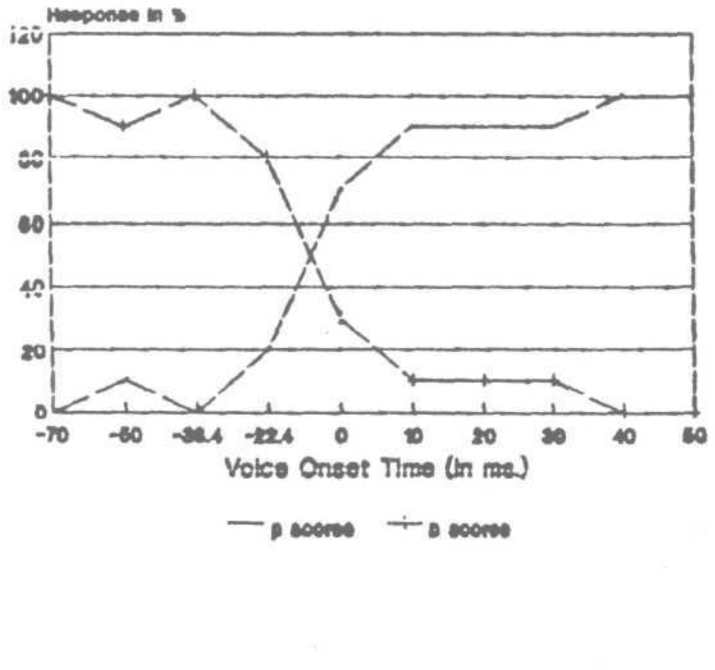

Fig 77. S_{20}'s response

crossover in the lag VOT region and fifteen subjects shoved crossover in lead VOT region. For the /p-b/ percept, only five subjects showed crossover in the lag VOT region and one subject had crossover at coincident VOT and remaining fourteen exhibited crossover in lead VOT region.

To summarize , the results indicated the following ;

1. As the VOT values changed from lead to lag there was a change in percept from voiced stop to voiceless stop.
2. The crossover from voiced to voiceless percept occurred in the lead portion of the VOT continuum for all of the three stops used in the study, 50% crossover occurred earlier for alveolar followed by bilabial and velar.
3. The boundary width decreased from velar to alveolar to bilabial.
4. Significant differences between the perceptual response of males and females for /k-g/ percept were seen.
5. Individual variation were observed among the subjects.

Discussions

Results revealed that VOT cues voicing in Kannada. At long lead VOTs subjects perceived voiced stop and at short lead and lag VOTs voiceless percept was reported. The crossover from voiced to voiceless percept occurred in the lead portion of the VOT continuum for all of the three stops.

The finding of the present study that VOT cues voicing is in agreement with the findings of Lisker and Abramson (1964), Zlatin (1974), Lisker (1975), Williams (1976), Diehl (1977), Ohde (1978), Keating, Mikos and Ganong (1981), Elliot et al (1986), Burnham et al (1991), who found VOT to cue the voicing contrast. The findings of the present study are contradictory to the findings of Usha Rani (1989) and Vinay Rakesh (1990), who did not find VOT to cue voicing contrast in Hindi, Kannada, Malyalam and Telugu respectively. The difference may be attributed to the place of the stop consonants in the word. Both these studies considered the effect of VOT on perception of voicing in intervocalic stop consonants where a myriad of other cues also come into play. Whereas in the present study, the stop consonant was in the word initial position.

Lisker and Abramson (1964) and Zlatin (1974) in their studies located the 50% cross over in the lag VOT range. Lisker and Abramson reported crossover values of + 25 m.sec, + 35 m.sec and + 45 m.sec for labial, apical and velar cognates. Williams (1976) in his study in Spanish found the crossover to be in lead VOT range. In the present study the 50% crossover was observed in the lead VOT range. In this regard Kannada, as a language using VOT to cue voicing, appears to be similar to Spanish and not to English.

Another reason for this discrepancy might be that while in English, by rule, the initial voiceless stop consonants are aspirated it is not so in Kannada. This rule based aspiration induces longer lag VOT's for voiceless stop consonants in English.

The findings of the present study are compared with the speech production studies in Kannada. Shukla (1989) reports an increase in both lag and lead VOT as the place of articulation moves backwards in the oral cavity (Table VII). Sridevi's study (1990) and Savithri et al's study (1995) contradict this in that while lag VOT increases as the place of articulation moves back in the oral cavity, lead VOT does not (Sridevi, 1990). The lead VOT decreases as the place of articulation moves backward in the oral cavity (Sridevi, 1990). In contrast Savithri et al's study (1995) reveals that the mean lead VOT's increase as the place of articulation moves backwards in the oral cavity. The differences between these findings may be attributed to the differences in the material used.

INVESTIGATORS	MEAN VOT (in m.sec)					
	P	b	t	d	k	g
Shukla (1989)	1	- 67	12	- 76	30	- 79
Sridevi (1990)	14	- 134	18	- 107	31	- 48
Savithri et al (1995)	33	- 80	15	- 85	19	- 85

Table VII : Mean VOT value as reported by various Investigators

As the speech production data of Savithri et al's (1995) study uses the same subjects as used in this study, their data will be compared with the speech perception data of the present study (Table VIII). On comparing the lower limits of the perception data and the voiced stop production several points of interest are revealed. VOT has higher values in speech production and the place effect evident in the perception data is not evident in production. The lower limit of speech production data is maximum for /d/ and minimum for /g/ whereas in speech perception an increase in VOT value as place of articulation moves back is evident in lower limit values. Conversely, the upper limit of VOT is maximum for /k/ and is minimum for /p/ whereas in perception the upper limit is maximum for /k - g/ and minimum for /t - d/. In speech production data the voiced and voiceless categories are clearly seperated by lead and lag VOT. In perception data it is not so. In production the stops showing greatest seperation are /t - d/ and least seperated are /k - g/. In perception data velars show maximum seperation and labials show minimum separation.

MEASUREMENTS	/k-	-g/	/t-	d/	/P-	b/
	SPEECH PRODUCTION	SPEECH PERCEPTION	SPEECH PRODUCTION	SPEECH PERCEPTICN	SPEECH PRODUCTION	SPEECH PERCEPTICN
Lower limit	-60	- 32	- 71	- 32	- 68	-22
Upper limit	+ 45	+ 8	+ 20	+ 1	+ 19	+ 6
S-Ratio/ Boundary Width	+ 47	+ 40	+ 70	+ 33	+ 66	+ 29

Table VIII : A comparison between the various measurements of speech production and speech perception in the data of same subjects.

This comparison reveals that no one to one comparison between speech perception and speech production exists.

Finally, it can be concluded that VOT acts as a cue for voicing in Kannada, though, the distinctions are not similar to English.

SUMMARY AND CONCLUSIONS

CHAPTER V

SUMMARY AND CONCLUSIONS

The present study was designed to investigate the effect of VOT on perception of stop consonant voicing. The stop consonants chosen for the study were voiced unaspirated bilabial stop, voiced unaspirated alveolar stop and voiced unaspirated velar stop. Three bisyllabic meaningful Kannada words,* with these voiced plosives in the initial position, as uttered by a twenty one year old Kannada speaking female were digitally recorded on a computer. Synthetic tokens were created by truncating voice pulses in steps of three pulses and adding silence after the burst in 10 ms steps once the pre-voicing was removed. The tokens thus synthesized were iterated ten times and audio-recorded with an inter-stimulus interval of one m.sec. A total of 330 tokens were audio presented to twenty adult-subjects (10 males and 10 females) for perceptual judgement. An alternate-forced choice task was chosen for the study and the subjects were required to indicate the phoneme they heard on the response sheet. Percent response for the voiced and voiceless plosives were tabulated for each of the test stimulus and the identification and discrimination functions were plotted for each individual. From these identification functions four measurements i.e, lower limit of the phoneme boundary, 50% crossover point, upper limit of the phoneme boundary and the phoneme boundary width were calculated. Lower limit of the phoneme boundary was the point on the identification-discrimination function where the subject identified the token as being voiced 75% of the time. The 50% cross over point was considered to be the point where the subject had equal percent of responses for the voiced and voiceless percepts, the upper limit was the 75 % response value for the voiceless percept. The difference between the lower limit and upper limit was considered as the phoneme boundary width. Paired 't' test was also done to know the significance of the difference in the responses of male and female subjects.

The results of the study revealed that long lead VOT cues voiced stops and short lead VOT and lag VOT cue voiceless stops in Kannada, and the change in the percept from voiced to voiceless stop occurs in the lead VOT range. Thus, it appears that VOT cues voicing in Kannada. Significant differences were found between the performance of males and females for /k-g/ percept. Also, the boundary width decreased from velar to bilabial place of articulation. Individual variations were observed in that some subjects showed multiple crossovers.

A comparison of the speech perception data with that of speech production data (of the same subjects) revealed no one to one correlation between the two. It appears that depending on the VOT structure in a given language, the 50% crossover changes while in English, voiceless stops are aspirated by rule in the initial position, in Kannada it is not so. In English the 50% crossover occurs in the lag VOT range. Williams (1976) in his study has reported that the crossover occurs in the lead VOT range in Spanish. In this, Kannada appears to be similar to Spanish. In conclusion, VOT appears to serve as a cue for voicing in Kannada though not in a similar way to English.

BIBLIOGRAPHY

BIBLIOGRAPHY

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

Bailey, P.J, and Summerfield, A.Q. (1980) "Information in speech : Observation on the perception of /s/- stop clusters" J. of Exp. Psy : Human Perception and performance. 6(3), 536-563.

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

Burnham, D. K. and Earn Shaw, L.J. (1991) "Development of categorical identification of native and non-native bilabial stops : infants, children and adults " J. Child. Lang, 18, 231-260.

Carney, A. E. (1977) "Non categorical perception of stop consonants differing in VOT", Acoustic Soc. America. 62 (2).

Datta, A. K. (1989) "Machine emulation of audition", J. of Acoust. Soc. India, XVII(3 & 4) , 1-8.

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

Day and Vigorito (1973), Cit. in "Auditory and phonetic levels of processing in speech perception", by Wood, C.C. Thesis.

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.

Diehl, R. L. (1977) "Phonetic context effects on the identification of stop consonants", J. Acoust. Soc. Am. 62, Suppl. No. 1, HH 11.

Dorman, M. F., Studdert-Kennedy, H. and Raphael, L.J. (1977)."Stop consonant recognition release bursts and formant transitions as functionally equivalent context dependent cues" Perception and psychophsics, 22(2), 109-122.

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. 65(6), 1518-1532.

Doughty, J. (1949), Cit. in Zlatin (1974) "Voicing contrast : Perceptual and Productive VOT characteristics of adults". J. Acoust. Soc. Am., 56(1), 981-989.

Elliot, L. (1986) "Discrimination and Response bias for CV syllables differing in VOT among children and adults", J. Acoust Soc. Am, 80(4), 1250-1255.

Elman, J. L., Diehl, R.L., Buchwald, S.E. (1977), "Perceptual Switching in bilinguals", J. Acoust. Soc. Am., 62(4), 971-975.

Fant, G.M. (1960), "Acoustic theory of speech production", Morton, The Hague.

Fischer-Jorgensen, E. (1979) "Temporal relations in consonant-vowel syllables with stop consonants based on Danish material", in B. Lindblom and S. Ohman (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, 27(4), 343-350.

Flanagan, J.L., 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.

Fruin, C.W., and Bischoff, D.M. (1976), "Vowel duration as a cue for consonant voicing ?", J. Acoust. Soc. Am. 60, Suppl.No.l, S91.

Halle and Stevens (1959), cited in Borden, G.J. and Harris, K.S. (1980) "Speech science primer-physiology, acoustics and perception of speech", Williams and Wilkins, Baltimore, U.S.A., 171-214.

Haskins group (1980), Cit. in "Speech Science Primer, Physiology, Acoustics and Perception of speech" by Borden, G.J. and Harris, K.S. (1980), Williams and Wilkins, Baltimore.

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.

Kiedel, W., Kallert, S., Korth, M. and Humes, L (1983), Cited in "Tutorial : Neuroanatomy, Neurophysiology, and Central auditory assessment, Part-1 : Brainstem", Museik, F.E. and Barah, J.A., Ear and Hearing, Vol. 7(4), 214, 1986.

Kim, M. (1993) "Production and perception of word initial stops by Korean adults" J. Acoust. Soc. Am, 94, Suppl.3, S9. Kluender, K.R., Lotto, A.R. and Jenison, L.R. (1992) "Perception of voicing in syllable initial stops at different intensities : Does neural synchrony encode voice-onset ?" J. Acoust. Soc. America, 92(4), S2.

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

Lisker, L. (1977), "Rapid 7s Rabid : A catalogue of acoustic features that may cue the distinction". J. Acoust. Soc. Am. 62, Suppl. No.l, 577.

Lisker, L. and Abramson, A. (1964), "A cross-language study of voicing in initial stops; Acoustic measurements", Word, 20, 384-422.

Lisker, L., Liberman, A.M., Erickson. D.M. and Dechovitz.D. (1975). "On pushing the VOT boundary about", Haskins lab status report on speech research, SR-42/43, 257-264.

Lisker, L., and Price, V. (1979), "Context determined effects of varying closure duration", J. Acoust. Soc. Am. 65, Suppl. No.l, 57.

Moslin, B.J. and John, K.S., (1976), "Korean apical stop production : a VOT analysis", J. Acoust. Soc. Am. 60, Suppl.No.l, S-45.

Miller, J.L. (1977) "Properties of feature detectors for VOT : the voiceless channel of Analysis", J. Acoust. Soc. Am, 62(3), 636-642.

Miller, J.L. (1977) Non independence of feature processing in initial consonants" JSHR, 20, 519-528.

Kearey, M.T. and Rochet, B.L. (1992) "Effects of place of articulation and vocalic context on the perception of VOT continua in French and English", J. Acoust. Soc. Am, 91, Suppl.4, S9.

Ohde, R.N. (1978) "Effects of VOT duration and number of adaptor repetitions on the scaling of stop consonant voicing", J. Acoust. Soc. Am. 64, Suppl. No. 1, S-19.

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

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

Raphel, L.J. (1980), Cited in, Borden, G.J. and Harris, K.S. (1980), "Speech science primer"-Physiology, acoustic and perception of speech". Williams and Wilkins, Baltimore, USA, 171-214.

Repp, B.H. (1977) , "Cross talk between voicing and place cues in initial stops", J. Acoust. Soc. Am., 62, Suppl. No. 1, S79.

Repp, B.H. (1979) "Perceptual trading relation between aspiration amplitude and VOT", J.Acoust. Soc. Am., 65, Suppl. No.l, S8

Savithri, S.R., Pushpa ; Sujatha. S. (1995), VOT in Speech production, paper presented at the Indian Speech and Hearing Association, Madras, 1995.

Shukla, R.S. (1989) "Voice onset times in the Kannada Language". J. of AIISH, XX, 13-18.

Sridevi, H.S. (1990) "VOT as a voicing cue to stop consonants in Kannada" J. Acoust. Soc. India, 16-18.

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

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

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

Vinay, R. (1990), "Cross language study of stop perception" Unpublished Master's Dissertation submitted to the University of Mysore.

Volaitis, L. and Miller, J.L. (1992) "Phonetic prototypes : Influence of place of articulation and speaking rate on interval structure of voicing categories" J. Acoust. Soc. Am., 92(2) , 723-735.

Williams, L. (1976) "Prevoicing as a perceptual cue to voicing in Spanish "., J. Acoustic. Soc. AM. 59, Suppl. 1, S-41.

Wood, C.C. (1976) "Discriminability, response bias, and phoneme categories in discrimination of VOT". J. Acoust. Soc. Ame. 60(2) , 1381-1386.

Zlatin, M.A. (1974) "Voicing contrast : Perceptual and Productive VOT characteristics of adults". J. Acoust. Soc. Am., 56(1), 981-989.

APPENDIX

APPENDIX

RESPONSE SHEET

Subject's Name

Age / Sex

SI.No.	Kadi	Gadi	Tada	Dada	Padi	Badi
1.						
2.						
3.						
4.						
5.						
6.						
7.						
8.						
9.						
10.						
11.						
12.						