SPEECH PERCEPTION : PERCEPTUAL CUES OF KANNADA STOP CONSONANTS

FINAL REPORT

DST REF No : SP/ YS/L 57/87

Principaf Investigator

Dr. (Mrs). S. R. Savithri

Lecturer, Department of Speech Sciences

All India Institute of Speech and Hearing

Manasagangothri, Mysore—570 006

Funded By

Department of Science and Technology

Ministry of Science and Technology

New Delhi, India

14-6-1988 to 13-6-1990

The coiled structure of the inner ear
is responsible for the sense of hearing

Karika: vali

ACKNOWLEDGEMENTSNTS;

This project was made possible by the qrant.s obtained from the? Department of science & technology, Ministry of Science & *technolog*y to whom our foremost: thanks are. We acknowledge Dr . M. Ratna , Ex-Director , All India Institute of Speech & Hearing, Mysore & Dr.S Nikkam, Director, All India Institute of Speech & Hearing, Mysore for havinq provided us with all the facilities. Our thanks are to Dr.N.B.Nataraja , Professor & Head Department of Speech Sciences Who encouraged us to take up this; project & since then has provided us with all the -facilities of the department. Mr, C.S. Venkatesh, Lecturer, Department of Speech Sciences & Ms, R. Manjula , Lecturer , Department of Speech Patholoqy deserve special thanks for giving us moral suppor when required. We are thankful to the Director, National institute of Mental Health & Neurosciences, Bangalore for allowing us to utilize the Sound spectroqraph 7800. We also thank the staff, Department, of Speech Pathology & Audioloqy, National Institute of Mental Health & Neurosciences, Bangalore for providing necessary facilities for us during our stay at Banglore; Dr.T.V.Ananthapadmanabha & Mr.Jayasimha, Voice speech Systems, Bangalore who prepared the LPC softwares for analysis are deeply acknawledged. Our acknowledgments are to Dr, S.S.Agarwal, CEER1, Delhi, Dr.B.Yajnanarayana, IIT, Madras, Dr. A.K.Datta, ISI, Calcutta, Ms.E.Mani Rao, AVNIHH,Bombay, for providing us help at various stages of our project, & Dr. Studdert…Kennedy, Dr.B.Repp, Dr . A. Liberman. all front Haskins Laboratory, USA, Dr.G.Fani, Royal lnstitute *of* Technology, Sweden, Dr.V.Zue,, MIT, USA for providing us with the reprints of all their studies on Speech perception. Our sincere thanks are to all our subjects who voluntarily participated in the experiment & taught us about Speech Perception & to Ms.Indira Prakash, Itenerent Speech Therapist & Mr.T. Suresh, II M.Sc. student 1990-1991 who helped us. with their perceptual evaluation of the imitations. We also thank II B.Sc. students. 1989-1990 for the help they have rendered, Our acknawledgments *are* to Mr . G. V. Ganeshiah , Artist, All India Institute of Speech & Hearing, Mysore for the photographic assistance , Ms.Susheela Bai , Stenographer , All India Institute of Speech & Hearing, Mysore & Ms. Shubha, Stenographer, All India Institute of Speech & Hearing, Mysore for their secretarial assistance.

TABLE OF CONTENTS:

Acknowledgements	IV-V
Abstract	XIII-XV
Introduction	1-3
Method	4-16
Results and discussion	17-97
Appendix I: Review of literature.	98-135
Appendix II: List of words.	136-141
Appendix III: Articles published/sent *for* publication.	142-175
Appendix IV: Phonetic symbols used.	176
Biblioqraphy	

LIST OF TABLES;

1. Aqe of the subjects participated in this study, 4
2. Significance of difference between the temporal parameters of MI & 181. 18
3. Significance of difference between the scaled temporal parameters of MI & IS1. 26
4. Siqnificance of difference between the spectral parameters *of MI* & IS1. 31
5. Parameters imitated percent times by IS1 (of M1) 32
6. Significance of difference between the temporal parameters of MI & IS2. 40
7. Significance of difference between the scaled temporal parameters of MI & IS2. 41
8. Significance of difference between the spectral parameters of Ml S< IS2. 53
9. Parameters imitated percent times by IS2 (of Ml) 54
10. Significance of difference between the temporal parameters of M2 & I1. 61
11. Significance of difference between the scaled temporal parameters of M2 & I1. 62
12. Significance of difference between the spectral parameters of M2 & I1. 74

13. Parameters imitated percent times by I1 (of M2) 80

14. "temporal features used in imitation by all

the imitators. 85

15. Transition duration of imitator imitating two models. 89

16. Scaled temporal features of the three imitators. 90

17. FO contours of the models & the imitators. 93

18. Basis of perceptual judqements for the quality of imitation. 94

19. Types of functions attributed to the **hemispheres.** 95

LIST OF FIGURES

1. Closure durations of MI & ISI. 19

2. Voicing duration of MI & IS1. 19

3. Burst duration of MI & IS1. 20

4.. Affrication duration of MI & IS1. 20

5. Aspiration & murmer duration of M1 & IS1. 21

6. VOT of M1 & I S1. 21

7. Transition duration of the first formant of MI & I S 1. 22

8. Transition duration o+ the second •formant of MI & I S 1. 22

9. Transition duration of the third formant of Ml & IS1. 23

10. Speed of transition of the first formant of MI & IS1. 23

11. Speed of transition of the second formant. of MI & IS1. 24

12. Speed of transition of the third formant. of MI & IS1. 24

13. Word durations of Ml & IS1. 25

14. Scaled closure? duration of Ml & IS1. 27

15. Scaled voicinq duration of Ml & 1S1. 2.7

16. Scaled burst duration of Ml & IS1. 28

17. Scaled affrication duration of Ml & IS1. 28

18. Scaled aspiration duration of M1 & IS1. 29

19. Scaled VOT of Ml & IS1. 29

20. Formant -frequencies of M1 & IS1. 33

21. Burst, amplitudes of Ml & IS1. 34

22. Overall amplitude of M1 & IS1. 34

23. Amplitude of the following vowel of M1 & IS1. 35

24. Amplitude of the following consonant of M1 & IS1. 35

25. Terminal formant frequencies of < Preceding Vowel) Ml & IS1. 37

26. F0 contours for Vowels following voiced & voiceless consonants of Ml & IS1. 37

27. Closure durations of M1 & IS2. 42

28. Voicing duration of Ml & 1S2. 42

29. Durst duration of Ml & IS2. 43

30. Affricatian duration of Ml & 1S2. 43

31. Aspiration & murmer duration of Ml & 1S2. 44

32. VOT of MI & IS2. 44

33. Transition duration of the first formant of MI & 1S2. 45

34. Transition duration of the second farmant of MI & IS2. 45

35. Transition duration of the third formant of MI & IS2. 46

36. Speed of transition of the first formant of MI & IS2. 46

37. Speed of transition of the second formant of MI & IS2. 47

38. Speed of transition of the third formant of M1 & IS2. 47
39. Word durations of Ml & IS2. 48
40. Scaled closure duration of Ml & IS2. 48
41. Scaled voicing duration of Ml & IS2. 49
42. Scaled burst duration of M1 & IS2. 49
43. Scaled affrication duration of Ml & IS2. 5O
44. Scaled aspiration duration of Ml & IS2. 50
45. Scaled MOT of Ml & IS2. 51
46. Formant frequencies of Ml & IS2. 55
47. Burst amplitudes of M1 & IS2. 53
48. Overall amplitude of Ml &t **IS2.** 56
49. Amplitude of the following vowel of Ml & IS2. 56
50. Amplitude of the following consonant of Ml & IS2. 57
51. Terminal formant frequencies of (Preceding Vowel) Ml & IS2. 59
52. F0 contours for Vowels following voiced & voiceless consonants of Ml & IS2. 59
53. Closure durations of M2 & I1. 63
54. Voicing duration **of M2** & **I1.** 63
55. **Burst** duration of M2 & I1. 64
56. Affrication duration of M2 & I1. 64
57. Aspiration & murmer duration of M2 & I1. 65
58. VOT of M2 & I1. 65
59. Transition duration of the first, formant of M2 & I1. 66

60. Transition duration of the second forman£
at M2 & I 1. 66
6 1 . Transition duration of the third formant
of M2 & I1. 67
62. Speed of transition of the first formant:
of M2 & I 1, 67
63. Speed of transition of the second formant,
of M2 & I 1. 68
64. Speed of transition of the third formant
of M2 & I1. 68
65. Word durations of M2 & I1. 69
66. Scaled closure duration of M2 & I1. 69
67. Scaled voicing duration of M2 & I1. 70
68. Scaled burst duration of M2 & I1. 70
69. Scaled affrication duration of M2 & I1. 71
70. Scaled aspiration duration of M2 & I1. 71
71. Scaled VOT of M2 & Il. 72
72. Formant frequencies of M2 & I1. 75
73. Burst amplitudes of M2 & I1. 75
74. Overall amplitude of M2 & I1. 76
75. Amplitude of the following vowel of M2 & I1. 76
76. Amplitude of the following consonant of M2 & I1. 77
77. Terminal formant frequencies of (Preceding Vowel)
M2 & I1. 77
78. F0 contours for Vowels fallowing voiced & voiceless
consonants of M2 & I1. 79

79. Transition patterns of /a/ following stops/ affricates in various places of articulation. 83

ABSTRACT

The purpose of this project was to extract the perceptual cues of stops & affricates in an imitation paradigm. 551 Kannada words with stops/ affricates, as embedded in a carrier phrase were uttered by two models which were recorded on high fidelity audio-cassettes. These words were imitated by two imitators with special attention on stop consonants which were also recorded. These words were digitized using a 12 bit A/D & D/A converter at 8000 Hz. resolution. 46 parameters were extracted using both Spectrography & Computer analysis for all the key stops/ affricates in the words. The acoustic parameters of the model & the imitator were compared & Walsh test was performed to find out the significance of difference between the features used by the model & the imitator. Also, the words used by the model & the imitators were recorded on to an audio-cassette & two judges perceptually evaluated for the quality of imitation.

The results indicated that the transition durations, & the speed of transitions of the first two formants were imitated maximally (apart from the place cues). Also, some of the scaled temporal parameters (aspiration & affrication duration) were imitated maximum number of times. The results have lead to make several points. Firstly, the importance of transition durations & the speed of transitions calls

attention on coarticulation which is reportedly defective in speech handicapped , especially so in the hearing impaired. More emphasis on coarticulation in these children would enhance their speech.

Second, the study throws some light on the perceptual mechanism. The results suggest that the perception may rely on a scaling of the temporal parameters rather than the absolutist temporal values.

Third, voice identification deserves some comments. The results indicate that in spite of the perceived similarity in two voice/ speech samples, there exist wide differences in the acoustic features used & the imitators use greater effort, which *are* points of consideration in Forensic voice identification.

Fourth, as the perceptual evaluation reveals a voice identification rather than a speech identification & as the cues for voice similarities were mostly suprasegmental in nature, more participation of the right hemispheres in these tasks are likely.

Fifth, the fact that the imitators were aware of the articulatory maneuvers to be changed to suit the models vocal tract, supports the view of Speech perception as perception of gestures.

In this context, it is suggested that the future research be oriented on perception in children, perception in speech handicapped & perception of suprasegmentals to explore further the mechanism of speech perception which would help in serving the speech impaired better.

INTRODUCTION

Speech is a multidimensional signal that elicits *a* linguistic association (Flanagan, 19/2). When one says this is X , we do understand it. The process of thi» understanding is termed speech perception. Both speech production and perception are considered unique to human beings . Interest in human speech perception has continued since three decades and attempts have been made to understand it.

Several investigations using synthetic speech stimuli have been conducted in the past, in various languages. Of these, Dorman, Raphael and Librman (1979)a. Fischer Jorgnensen (1979), Lisker and Price (1979), Price and Lisker (1979), Bailey and Summerfield (1980), Fitch,Halwes, Erickson and Liberman (1980),Usha (1989),Datta (1989), Vinay (1990) have worked on **the** closure duration of stop consonants. Raphael (1972),Fruin and Bishoff (1976) , Raphael (1980), Usha <1989) and **Vinay** (1990) **have** experimented with the preceding vowel duration as a cue to voicing of stop consonants. Keating and Blumstein (1978), Hasking group (1980), Raphael (1980), Usha (1989) and Vinay (1990) have worke**d with the** transition duration of the preceding vowel which may cue the stop perception, stevens and Klatt (1974), Lisker (1975), Summerfield and Haqqard (1977), Ahmed and Gupta (1980), Bailey and Summerfield

(1980), Usha (1989) and Vinay (1990) have explored the possibility of the transition duration of tha following vowel as a cue to voicinq of stop consonant. Stevens and Klatt (1974). Lisker (1975), Moslin and John (1976), Darwin and Brady (1975), Diehl (1977), Summerfield and Haggard (1977), Ohde (1978), Keating, Mikos and Ganong (1981), Usha (1989), and Vinay (1990) have worked on VOT as a cue for stop consonant voicinq, Blumstein and Stevens (1976). Stevens (1977), Fisher-Jorqensen (1979), Dorman and Raphael (1977, 3.980), Erickson, Fitch, Halwes and Liberman (1977), Winitz, Schieb and Reeds (1972), Repp (1984) experimented with release bursts as a cue to stop consonants.

Dhala (1978), Gruenfelder (1979), Ohde (1982, 1984), Peterson (1983), have studied F0 contours of the following vowel Chasaide and Gobl (1987) studied the bandwidth of F1, Stevens and Klatt (1974) on the tradiang relationship between VOT and spectral cues. Summerfield and Haggard (19/7) on Fi, F2 transition, Reeds and Wang (1961), Walsh and Parker (1980) on aspiration as a cue separating voiced and voiceless in the initial position. Mann (1980), Mack and Blumstein (1983), Miller end Baer (1983), Walsh and Diehl (1987), Zakia and Kingston (1987) studied the cues differentiating stops from other consonants. (For detailed review see Appendix 1).

In spite of extensive research in this area, the way in which the cues combine to produce a percept is nat yet known. Also, some of the temporal parameters seem to vary across languages suggesting that the perception of native and non-native speakers may be different and that the selectivity of cues may vary across languages. In this context, the present study was designed to find out the perceptual cues of Kannada stop consonants and affricates in an imitation task. It is hoped that the outcome of this project would help develop new methods in teaching Speech handicapped.

METHODOLOGY

Matial Initially, a pilot study was conducted to select the environments for stop consonants. 551 meaningful kannada words were selected for the study with all possible environments. These words as embedded in sentences 'word anta heiltini' formed the test material (Appendix II).

Subjects: Two young adult Kannada speaking males served as the model subjects and two young adult Kannada speaking males served as imitators. The imitators were selected on the basis of their performance in mimicry. The age of all the speakers are in table 1.

Subject	Age in years
Model 1	17
Madel 2	29
Imitator 1	26
Imitator 2	*17*

Table 1: Age of the subjects participated in this study.

Method:

Recording: Kannada words were visually presented (as written on cards) one at a time and the model subject was instructed to utter them embedding in sentence in a natural manner into the microphone (cardiode-unidirectional) kept at a distance of 10 cms -from the mouth. All these were recorded on high •fidelity audio-tapes. These sentences were further transferred to audio--cassett.es and were given to the imitator for practice. Once the imitator felt confident about his proficiency, he was audio-presented with the sentences uttered by the model through earphones and was instructed to imitate the sentences with special attention to the stop consonants. These? were also recorded on high-fidelity audio-tapes. The key words containing the stop consonants, of all the subjects were extracted from the sentence and stored in an audio-cassette. The words were further digitally stored in magnetic diskettes using a PC/XT with 12 bit A/D and D/A converter at a sampling rate of 8000 Hz.

Acoustic analysis: Stop consonants in the words which were transformed digitally were computer analyzed using LPC autocorrelation method to extract F1 , F2 , F_3, Bl, E2, B3, L1, L2, L3. Further, using the sound spectrograph (VII 700), wideband bar-type and average amplitude type of spectrograms were obtained for all the words with stop consonants and 27 parameters were measured using these spectrograms, viz., Closure duration (C.D), VOT, Burst duration (B.D), Aspiration duration (A.D), Affrication duration (AFD) , Voicing duration (V.D) , Murmur duration (M.D), Stop-conanant duration, Word duration (W.D), Transition duration of Fl, F2, F3 of the preceding and the following vowel (TDF1, TDF2, TDF3), Speed of transition of Fl, F2, F3 of the preceding and the following vowel <STF1, STF2, STF3) , Burst amplitude (BA), Amplitude of the following vowel and consonant (AV, AC), and Overall amplitude (OA) . Totally 46 parameters were extracted for each consonant thus constituting 1,26,730 tokens. These parameters as used by the models were compared with those of the imitators.

Percentual analysis: The words as uttered by the models and the imitators were transferred in order on to an audio-cassette and a booklet was provided to two Kannada speakers (one male and one female) for perceptual evaluation- They were instructed to judge the imitation as very good/ fair/ poor/ bad and indicate the cues for their judgements. The perceptual judgements thus obtained were correlated with the acoustic parameters imitated to arrive at the perceptual cues of Kannada stops/ affricates.

Definition of the parameters used:

1. F1 , F2, F3 (Hz): Frequency of the first, second and third formants as obtained through LPC auto-correlation analysis. The formant frequencies were obtained at the instance of burst and at the voicing period when it was not available at the instance of burst.

4. Bl, B2, B3: Bandwidths of the Fl, F2 and F3 respectively as obtained through LPC auto-corrslation analysis.

7. L1, L2 L3: Levels of the Fl, F2 and F3 respectively as obtained through LPC auto-correlation analysis.

10. Steady F0 of the preceding & the following Vowel: F0 in the steady portion of the Vowel as displayed on the computer.

12. Terminal F0 of the preceding Vowel; F0 in the terminal portion of the Vowel as displayed on the computer.

13. Initial F0 of the following Vowel: F0 in the initial position of the Vowel as measured from the computer disp1ay.

The following parameters were mesured from the wide-band bar and average amplitude spectrograms:

1. Closure duration (msec): Duration between the offset of resonance for the preceding vowel and the onset, of burst for the stop consonant. in intervocalic condition and duration from the regular vertical striation on the baseline to the onset of burst, for the voiced stop an the-initial position.

2. Burst, duration (msec) : Duration of the vertical irregular striations depicting the articulatory release.

3. Aspiration duration (msec): Duration of the low frequency energy measured as the time between the onset of irregular vertical striation in the low frequency region to the end of the same between the burst, and the onset of resonances for the following vowel .

4. Voicing duration (msec): The time between the onset and offset of regular vertical Etriations within the closure duration.
5. Affrication duration (msec): The duration of high- low frequency energy depicted as irregular vertical striations starting from high frequency & gradually dipping to low frequency between the burst & the onset of resenance for the following vowel.

6. VOT (msec): The duration between the burst and the onset of tha resonances for the following vowel.

7. Murmur duration (msec): The duration of partial voicing depicted as regular vertical striations superimposed by irregular vertical striations on the spectrogram between the burst and the onset of resonances for the following. vowel.

8. Transition duration of Ft, F2 and F3 of the vowel (msec) s Duration between the onset of resonances changing to the onset of steady resonances for the following vowel and the duration between the onset/offset of resonances shifting for the preceding vowel.

14. Speed of transition (Hz/msec): The ratio of the frequency shift during transition to the duration of transition.

19. Stop duration <msec): The duration including the closure, burst and aspiration.

20. Word duration: The time between the onset/offset of regular vertical striations on the baseline for a word starting with voiced stop/vowel and as the time between the burst and the offset of regular vertical striations on the baseline for a word starting with voiceless stop.

21. Burst amplitude (R dB): Peak amplitude at the region of the burst as measured from the average amplitude spectrogram.

22. Overall amplitude (R dB) : Peak amplitude (except that of the burst) in this region of the atop aS measured from the average amplitude spectrogram.

23. Amplitude of the following vowel/consonant *(R* dB) : Peak amplitude within the duration of vowel/consonant following the Stop/ Affricate measured from the average amplitude spectrogram.

25. Terminal F1 ,F2 ,F3 , of the preceding Vowel: The terminal frequencies of the first three formants of the preceding Vowel as measured from the spectrogram.

28. Scaled temporal measures (CD, VD, AD, AFD, BD, VOT) : These parameters were scaled in terms of percent of the total stop duration by using the following formulae.

CD * 100

Scaled CD= -----------------------

Stop/ Affricate duration

Stop/ Affricate duration

VD * 100

£>C«1«U VU & ~*" .~**VD**—.* **100** – – ~– ~

Scaled VD = ---------------------------

Stop/ Affricate duration

Stop/ Affricate duration

AD * 100

AD * 100

Stop duration

AFD * **100**

Scaled AD = -----------------

Affricate duration

BD * **100**

3UB1SQ OU w **Stop duration** ~™~~™™·-“—·–----~

Stop/ Affricate duration

AFD * 100

VOT * 100

Stop/ Affricate duration

Scaled AFD = ----------------

Affricate duration

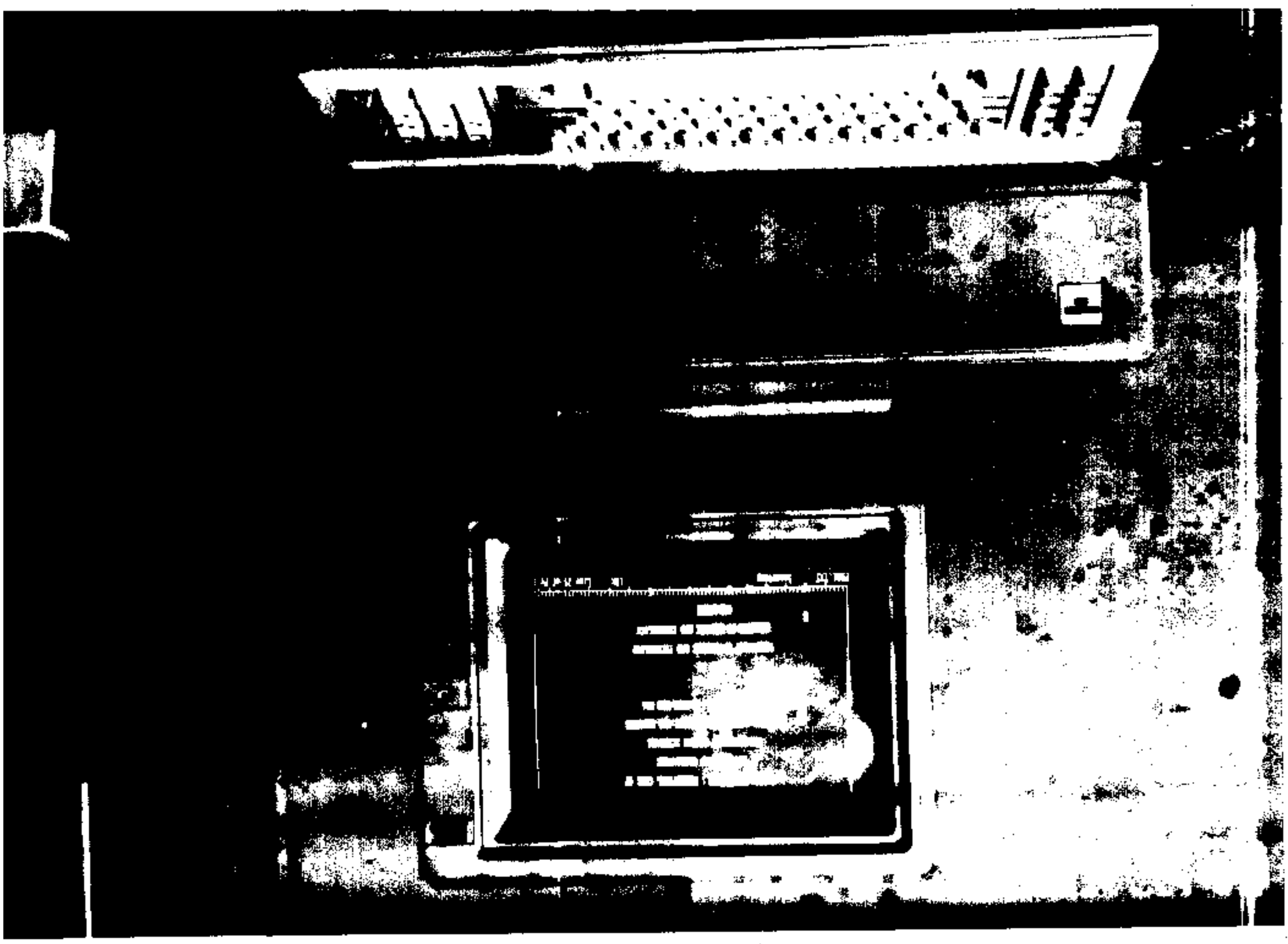
Scaled VOT = Stop/ Affricate duration

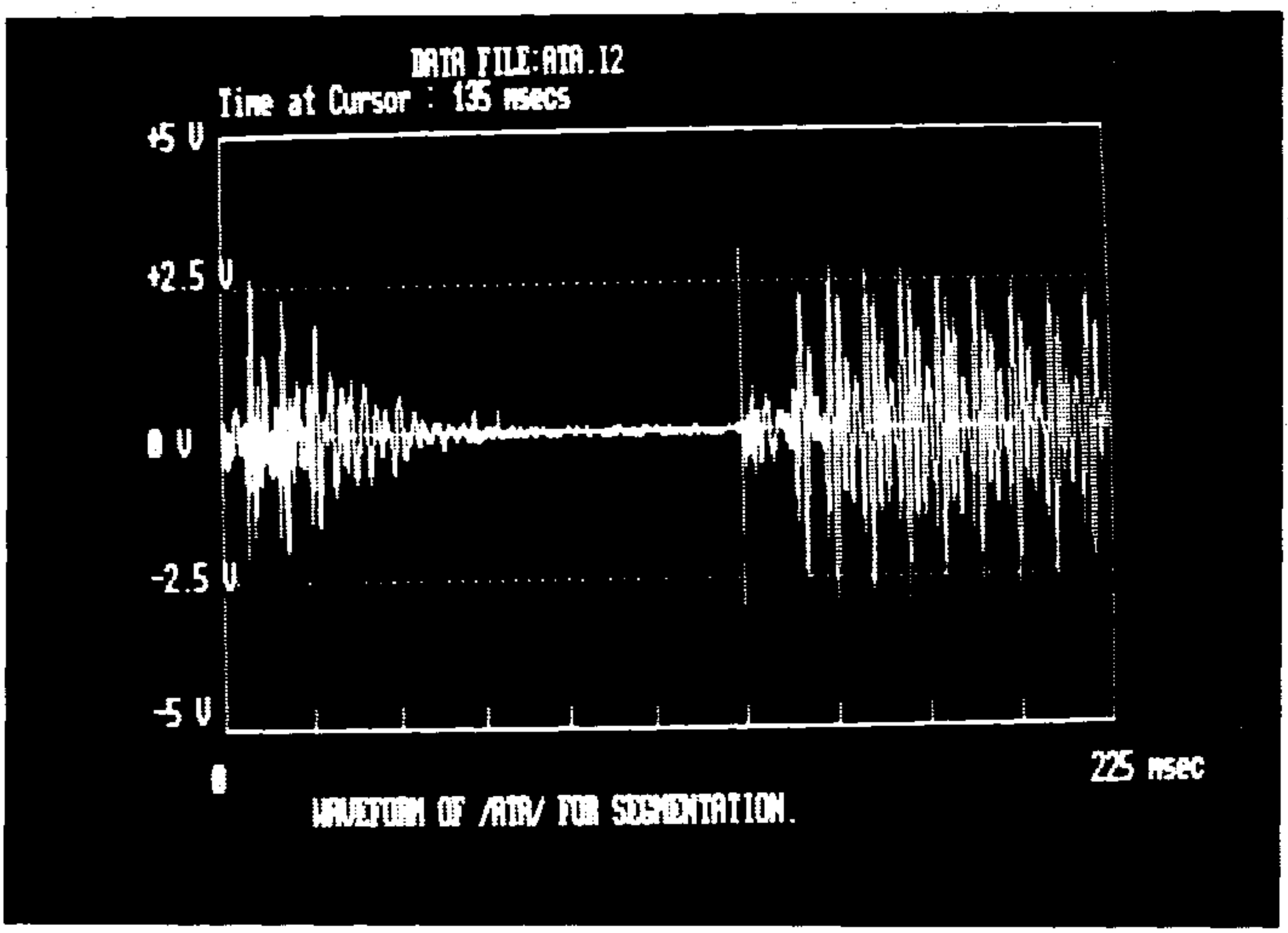
Scaled BD = Stop duration

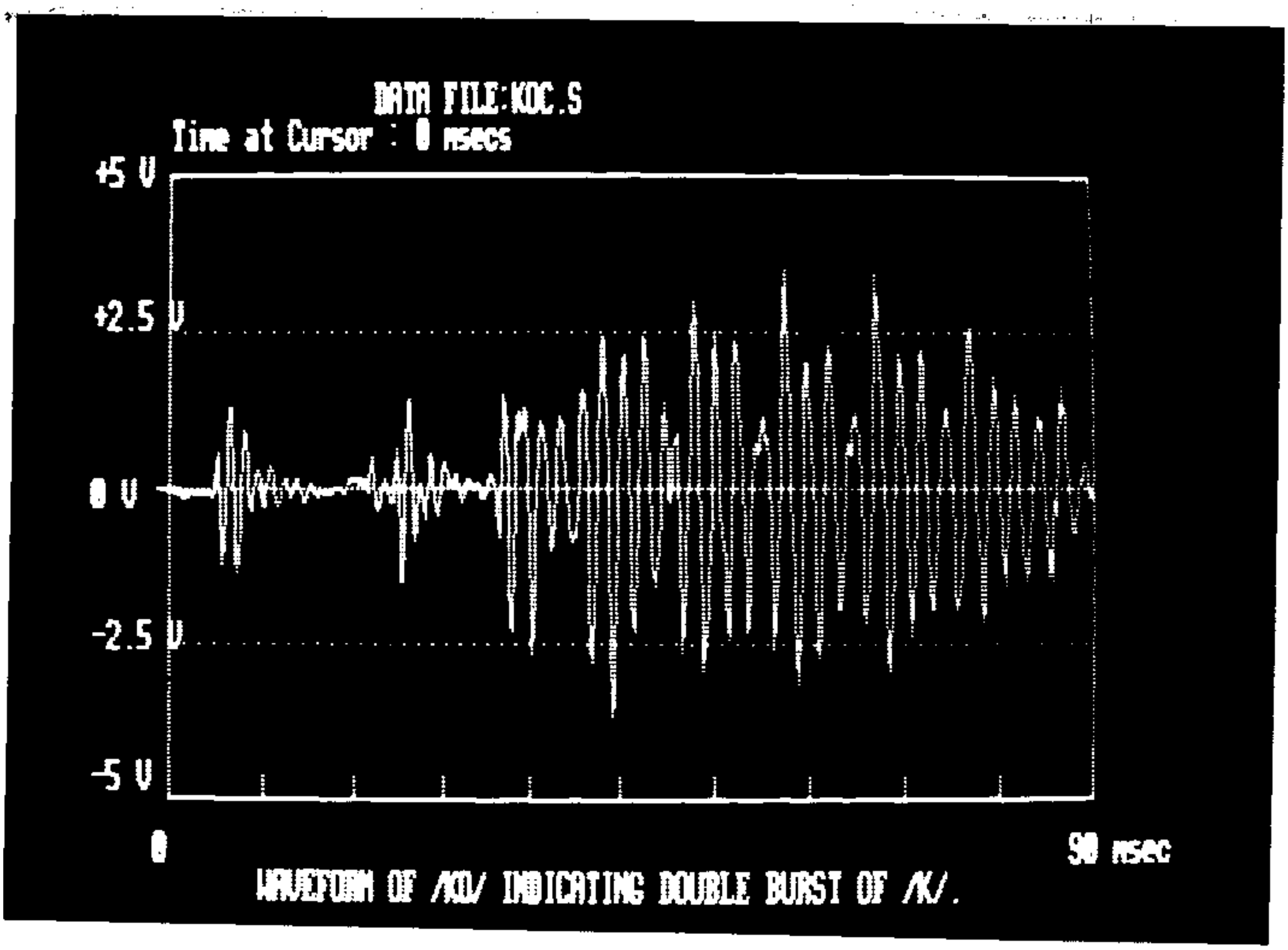
The temporal & the spectral measures of each phoneme for a given Vowel environment were measured & averaged. These measures of the model & the imitator were compared to find the significance of difference between the parameters used. Walsh test was adapted for the same. Percent times parameter imitated were calculated using the following formula.

$$\%\ \text{imitated} = \frac{\text{No. of times each parameter had significant difference}}{\text{Total number}}$$

All these measures are considered in discussing the results. Photographs following depict the computer used, wave of part of a word, double burst for the Velar */k/,* LPC analysis with a hamming window, forniants, bandwidths & formant levels extracted from LPC autocorrelation method, spectral display, waveform & LP spectral display, FO contour for voiced affricate, & waveform display comparing waveforms of the model & the imitator in order.




DATA FILE:ATA.12
Time at Cursor : 135 msecs
+5 V
+2.5 V
0 V
-2.5 V
-5 V
0
225 msec
WAVEFORM OF /ATA/ FOR SEGMENTATION.


DATA FILE:KOC.5
Time at Cursor : 0 msecs
+5 V
+2.5 V
0 V
-2.5 V
-5 V
0
90 msec
WAVEFORM OF /KO/ INDICATING DOUBLE BURST OF /K/.

RESULTS

The results are discussed with reference to each imitation and all the three imitations are compared and discussed with reference to the earlier findings.

MODEL (M1) Ys IMITATION (ISI)

The? temporal parameters TDF1 , STF1 , STF2 were maintained by the imitator for more than 75%. times. Significant differences were noticed in the closure durations of *gh*, j, d, bh, voicing duration of gh, d, bh, burst duration of k, kh, c, d, t, affrication duration of g, j, Aspiration duration of kh, th, murmer duration of bh, VOT of k, kh, c, t, d, p, bh, TDF1 of kh, g, j, TDF2 of kh, g, j, bh, TDF3 of k, kh, g, c, th, d, bh, STF1 of g, d, STF2 of j, STF3 of g, gh, c, d, p, bh, and word durations of all except d, Word durations in all the conditions except for /d/ were shorter in the imitator compared to the model. Figures 1 to 13 represent the temporal measures of the imitator and the model and the significance of difference between the temporal parameters of the model and the imitator *are* represented in table 2.

Pho neme	CD	VD	BD	AFD	AD	MD	VOT	TD Fl	TD F2	TD F3	ST Fl	ST F2	ST F3	WD
k			+				+	+						+
kh			+		+		+	+	+	+				+
9													+	+
gh	+	+											+	+
c			+							+				+
j	+			+				+	+					+
th					+									+
d	+	+	+							+	+			+
t			+				+			+				
d							+							+
P							+						+	+
b														+
bh	+ +					+	+		+	+			+	+

Table 2: Significance of differences between the temporal parameters of M1 and IS1.

parameters of Ml and ISI.

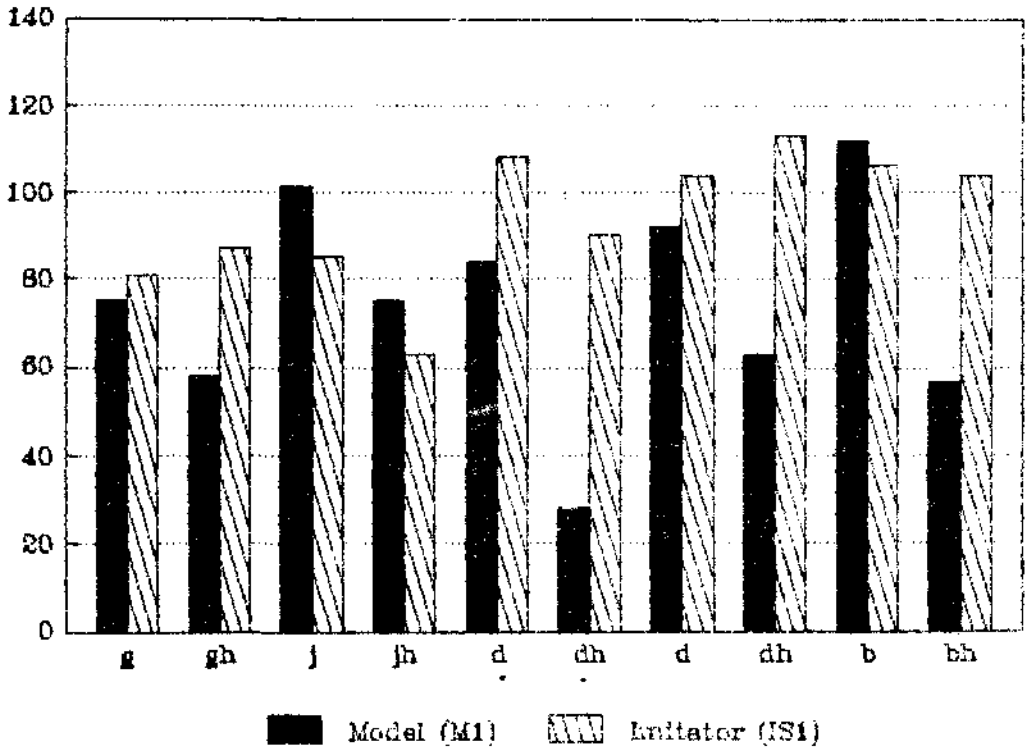


Fig.1 Closure durations of MI & IS1 (M.sec)

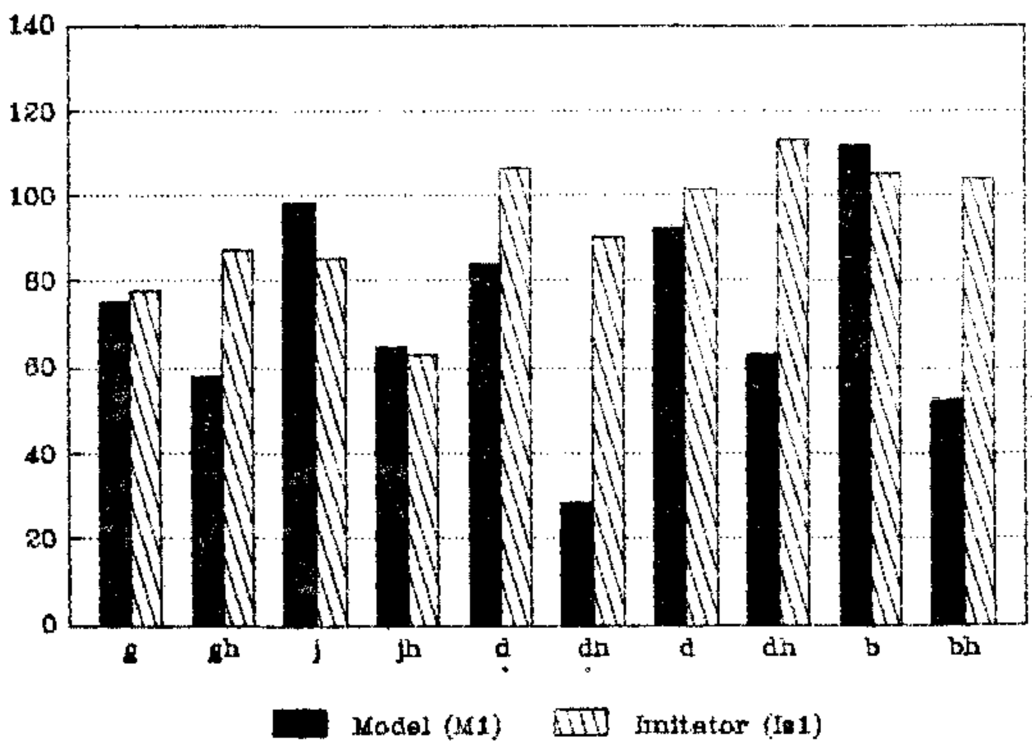


Fig.2 Voicing duration of M1 & IS1 (M.sec)

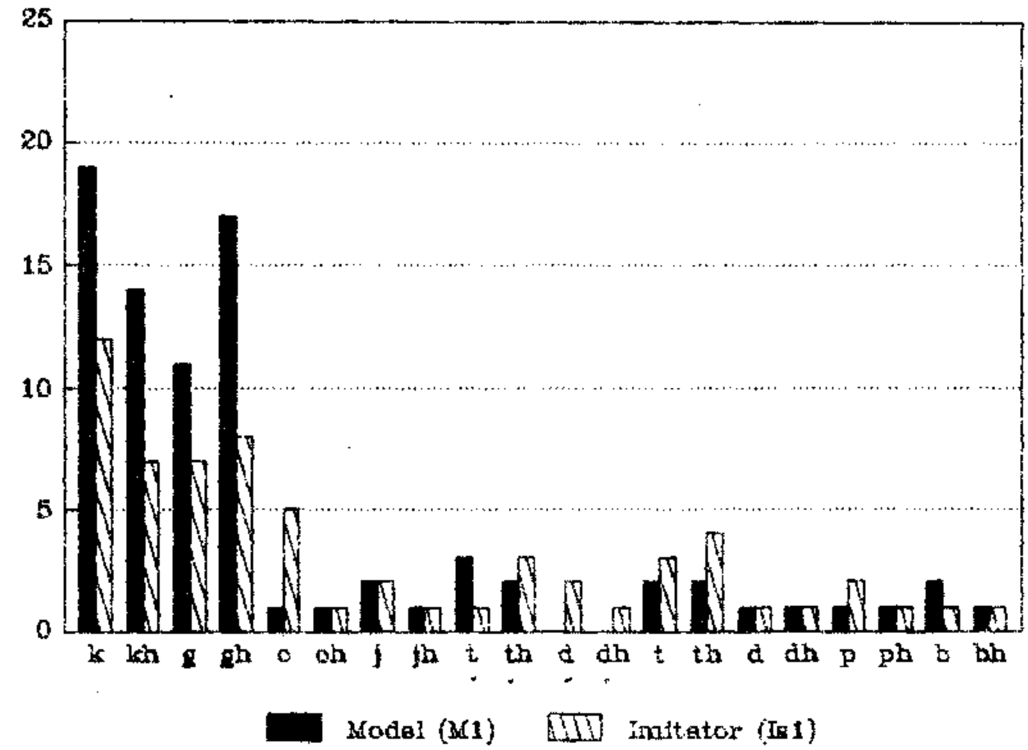


Fig.3 Burst duration of M1 & IS1 (M.sec)

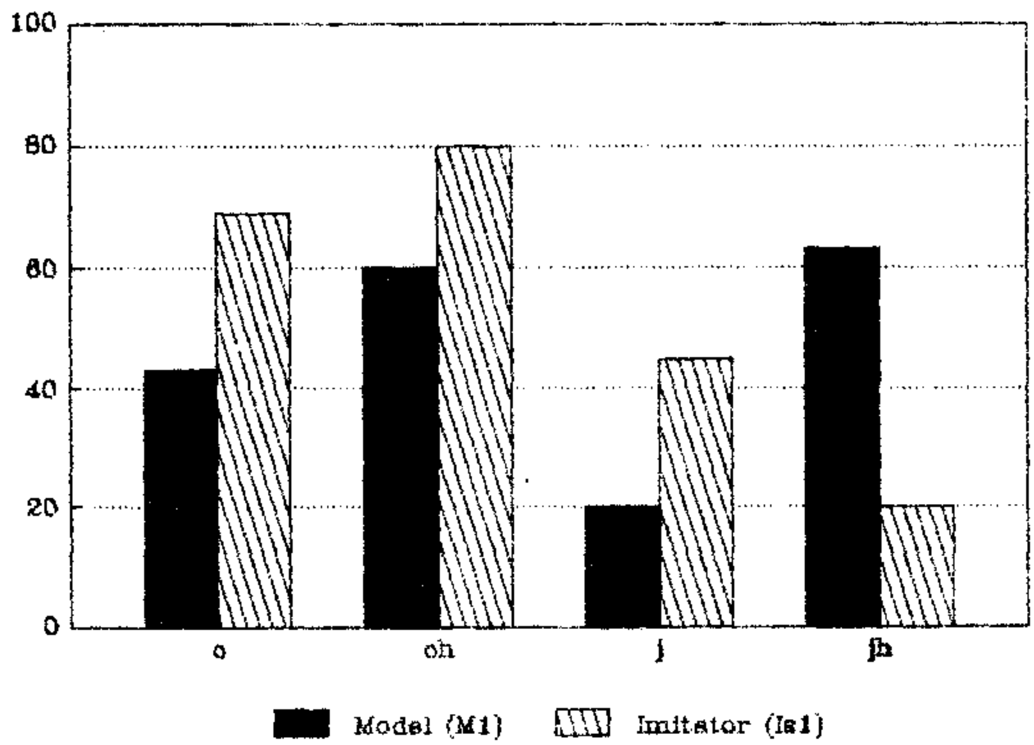


Fig.4 Affrication duration of M1 & IS1 (M.sec)

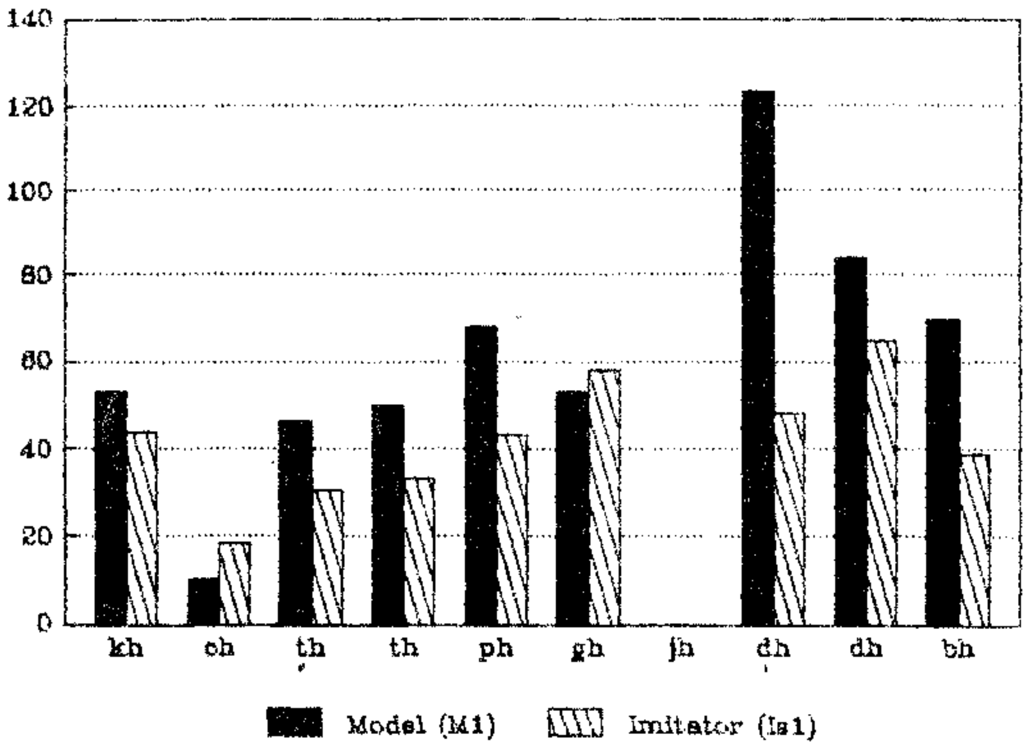


Fig.5 Aspiration & murmur duration of M1 & IS1 (M.sec)

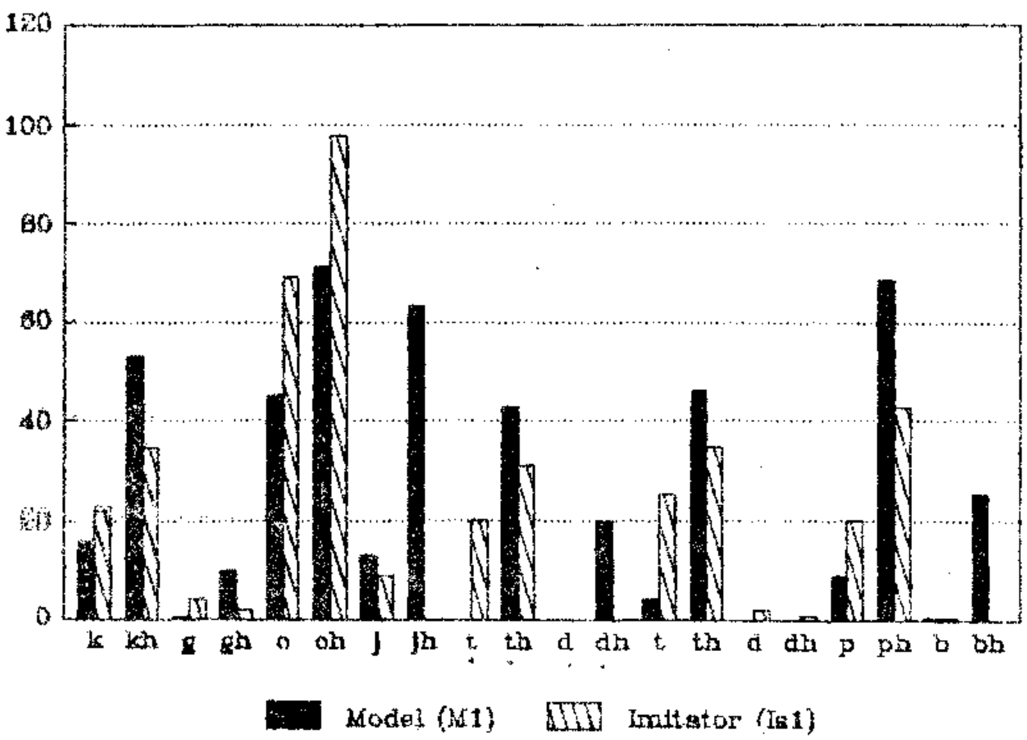


Fig.6 VOT of M1 & IS1 (M.sec)

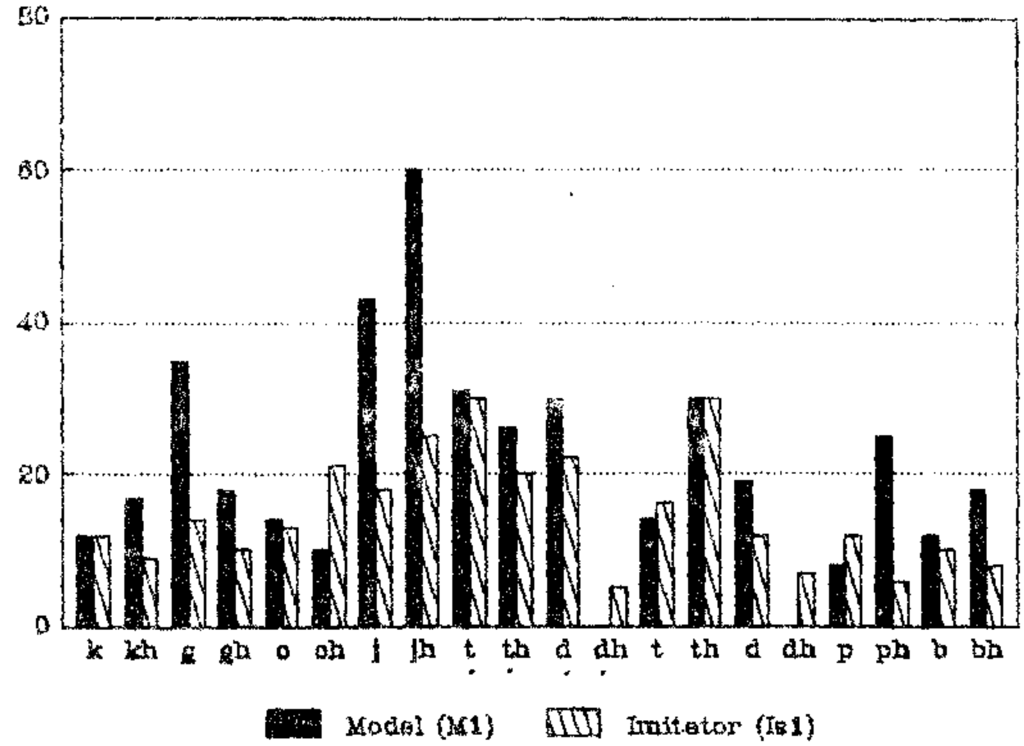


Fig.7 Transition duration of the first formant of M1 & IS1. (M.sec)

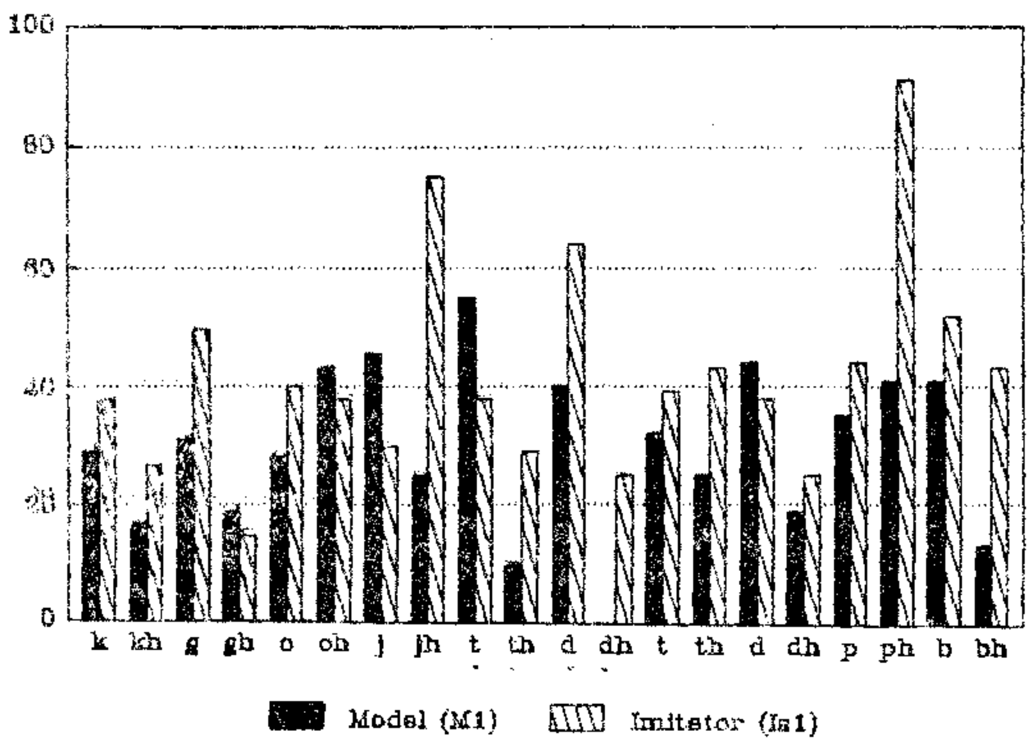


Fig.8 Transition duration of the second formant of M1 & IS1. (M.sec)

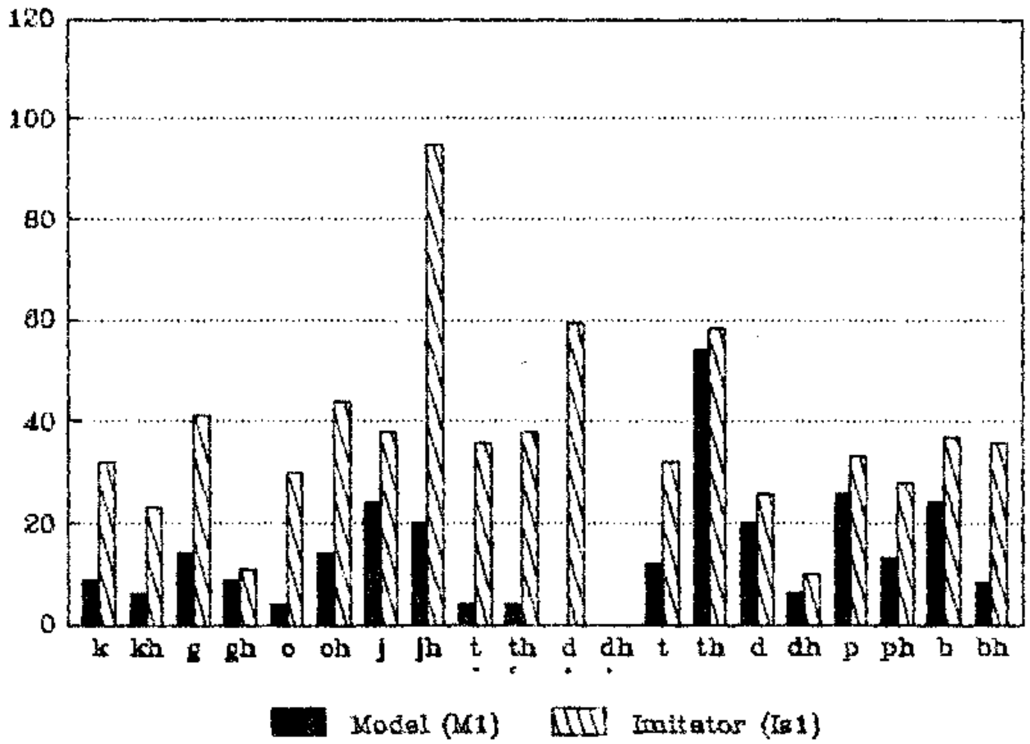


Fig.9 Transition duration of the third formant of M1 & IS1 (M.sec)

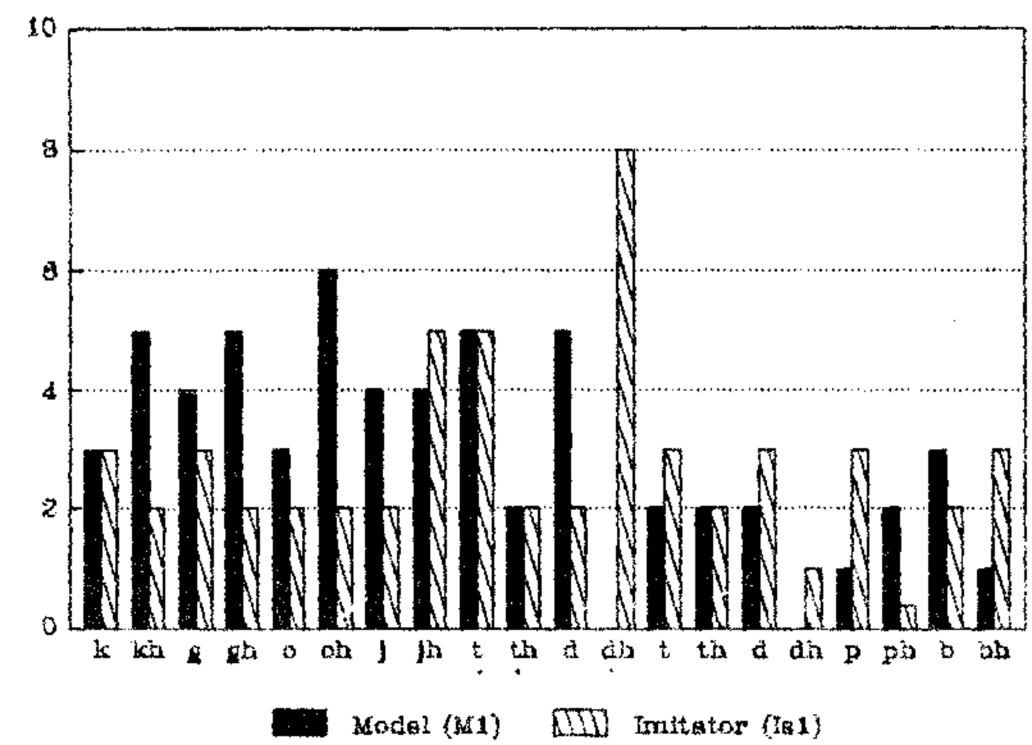


Fig.10 Speed of transition of the first formant of M1 & IS1 (Hz/M.sec)

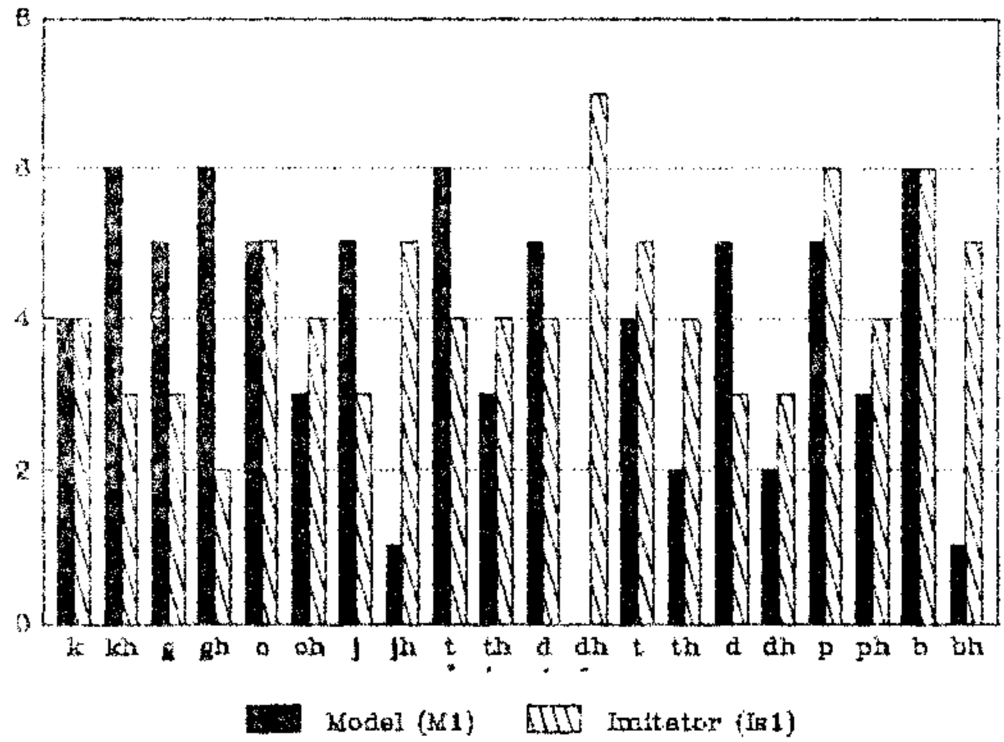


Fig.11 Speed of transition of the second formant of M1 & IS1 (Hz/M.sec)

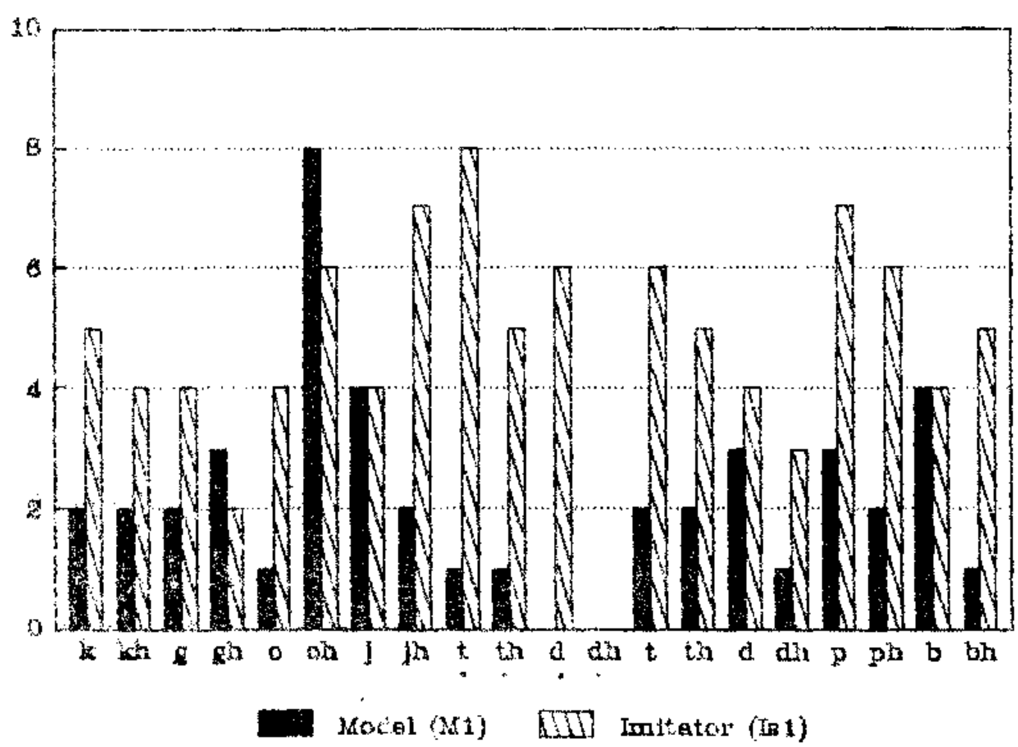


Fig.12 Speed of transition of the third formant of M1 & IS1. (Hz/M.sec)

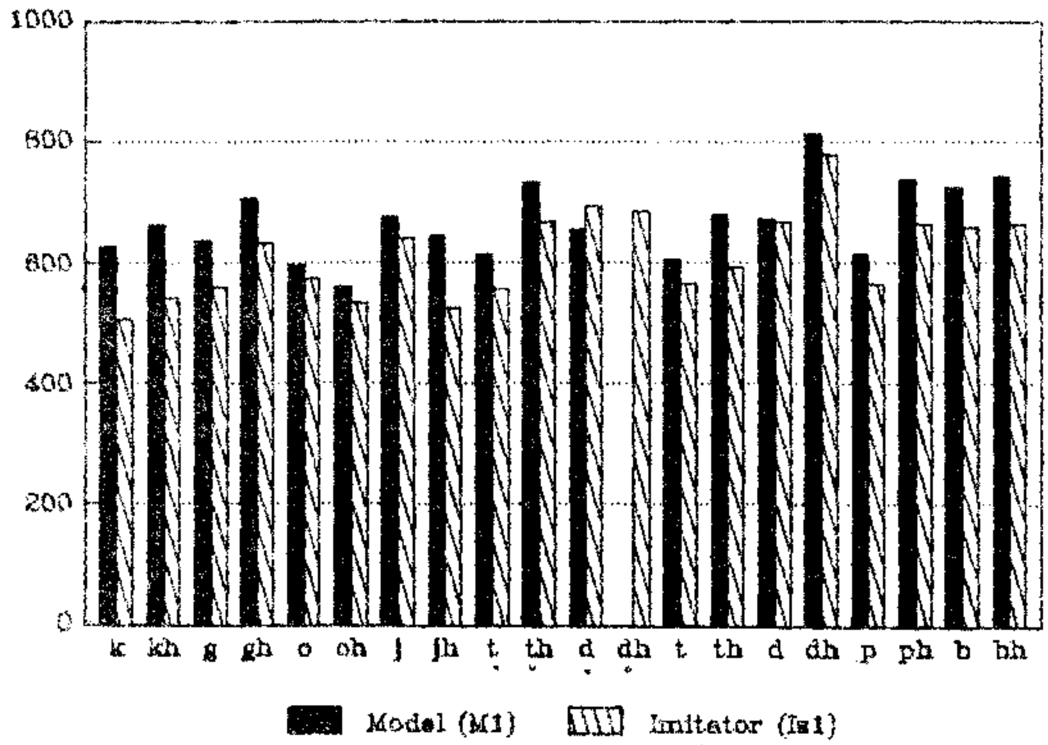


Fig.13 Word durations of M1 & IS1 (M.sec)

On scaling the temporal measures, Walsh test indicated no significant differences between scaled affrication and aspiration durations af Ml and IS1 (Table 3). Figs 14 to 19 represent the various scaled temporal measures for Ml and Is1.

Phoneme	CD	VD	BD	AFD	AD	VOT
k			+			
kh			+			
g						
gh		+	+			
c						
j						
th						
d	+	+	+			
t						+
d						
p						
b						
bh						

Table 3: Significance of differences between the scaled temporal parameters of Ml and Is1.

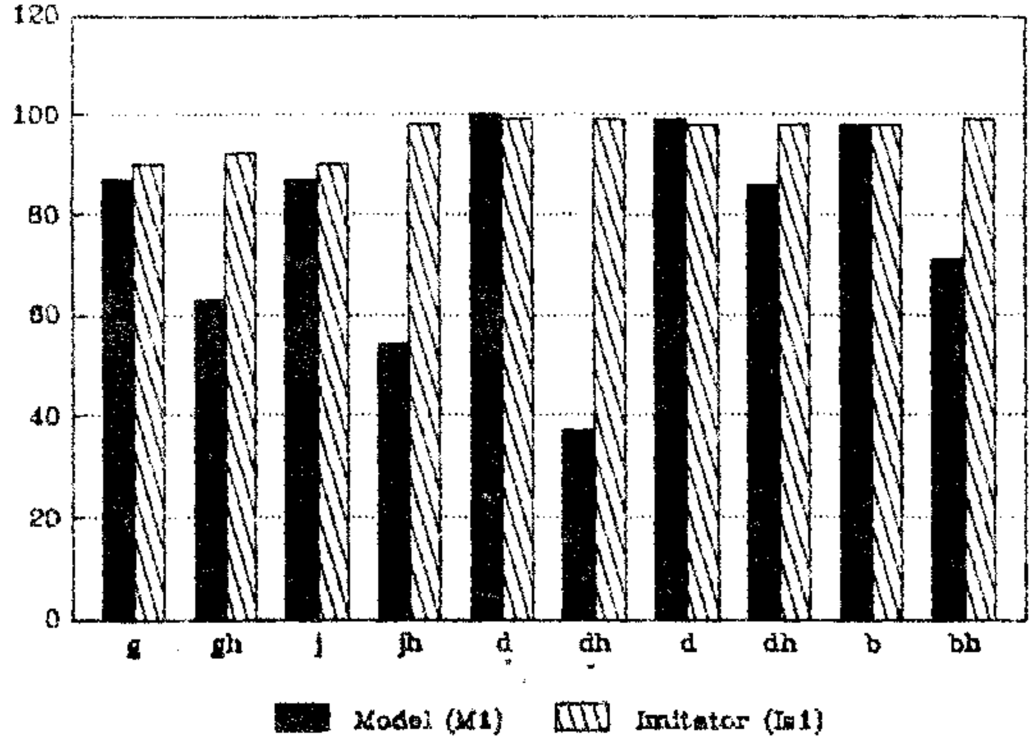


Fig.14 Scaled closure duration of M1 & IS1.

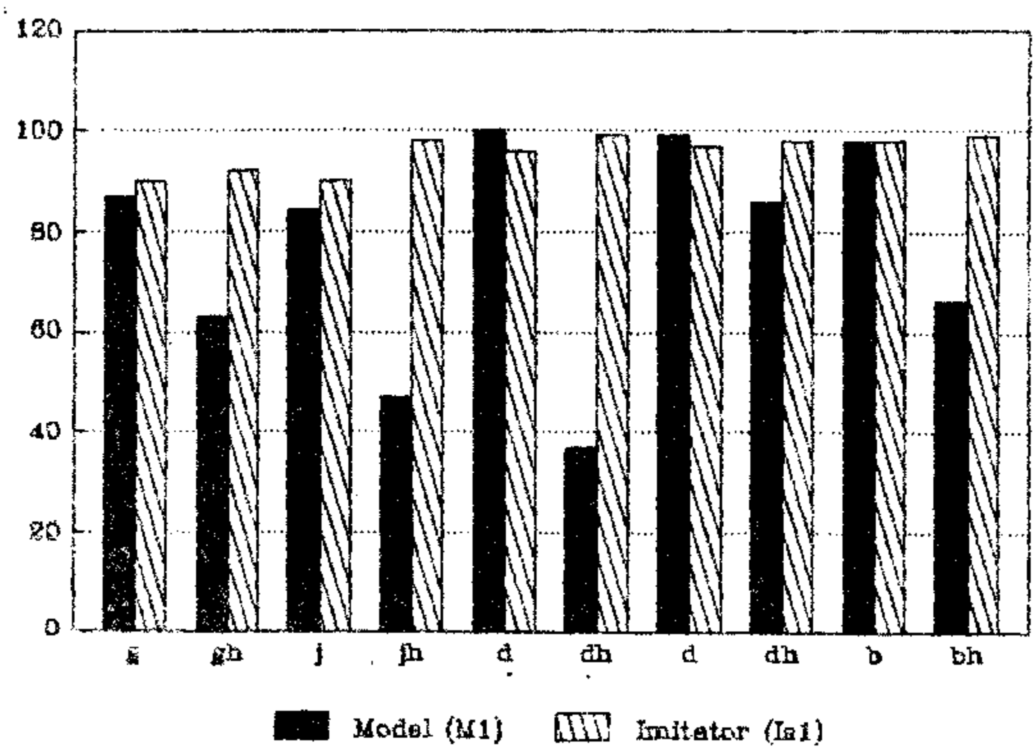


Fig.15 Scaled voicing duration of M1 & IS1.

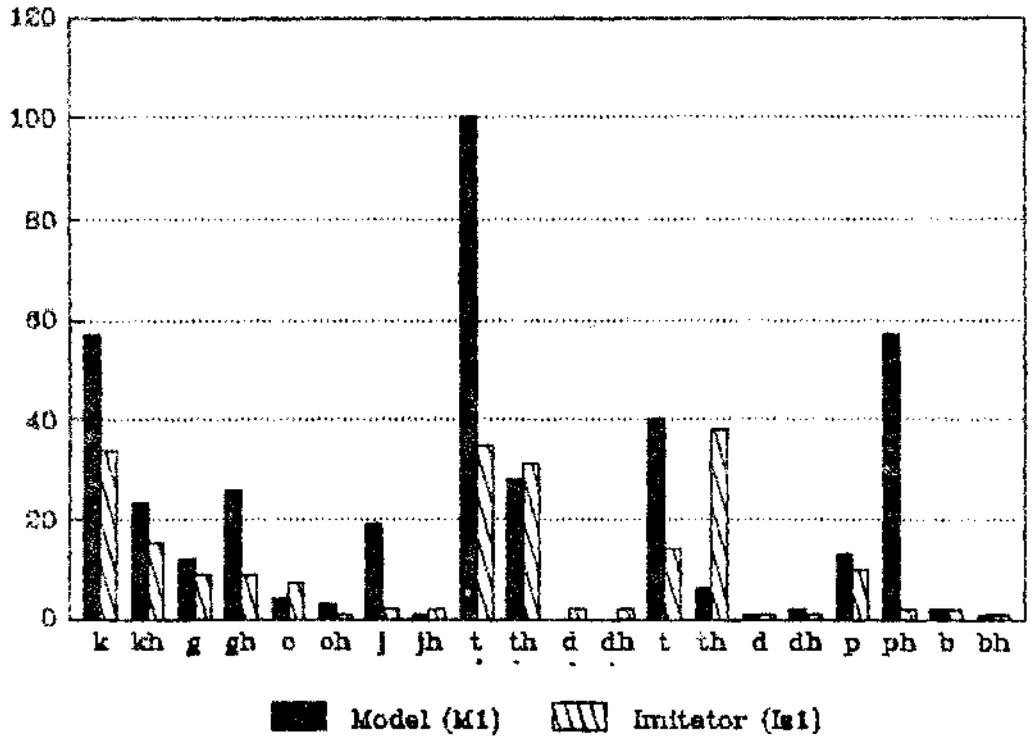


Fig.16 Scaled burst duration of M1 & IS1.

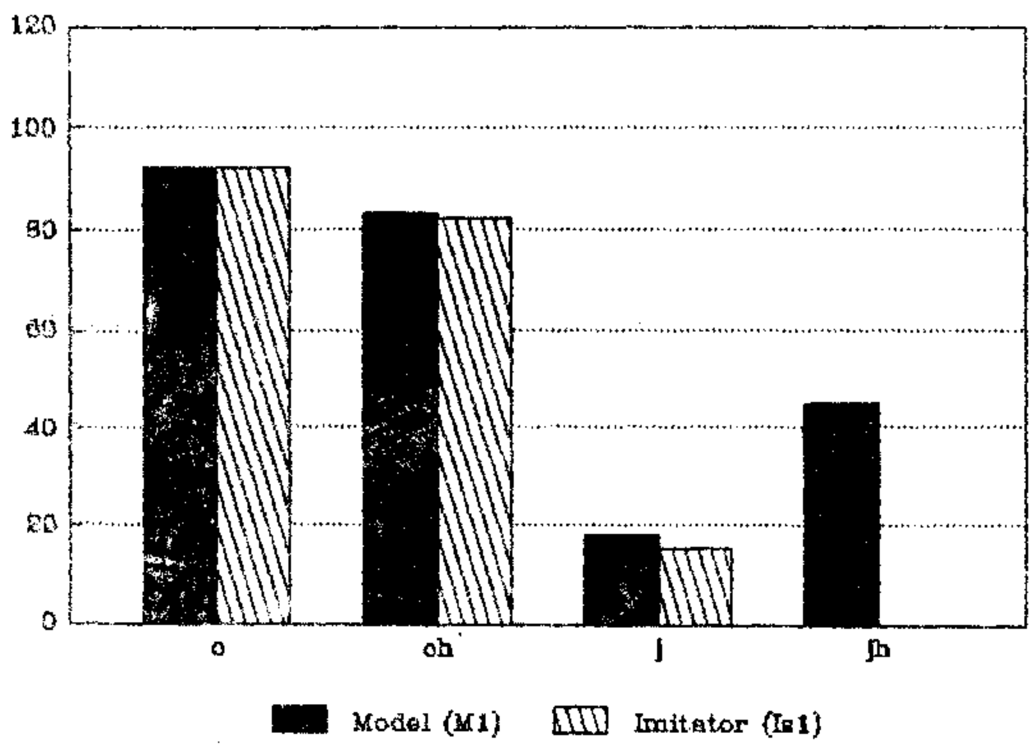


Fig.17 Scaled affrication duration of M1 & IS1.

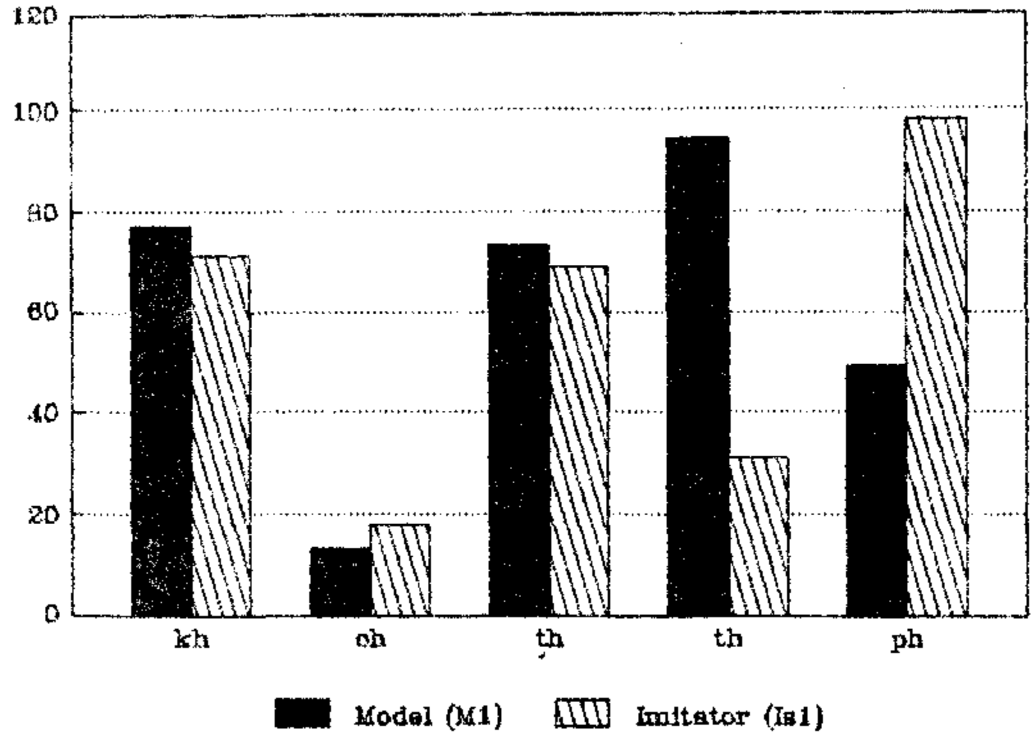


Fig.18 Scaled aspiration duration of M1 & IS1.

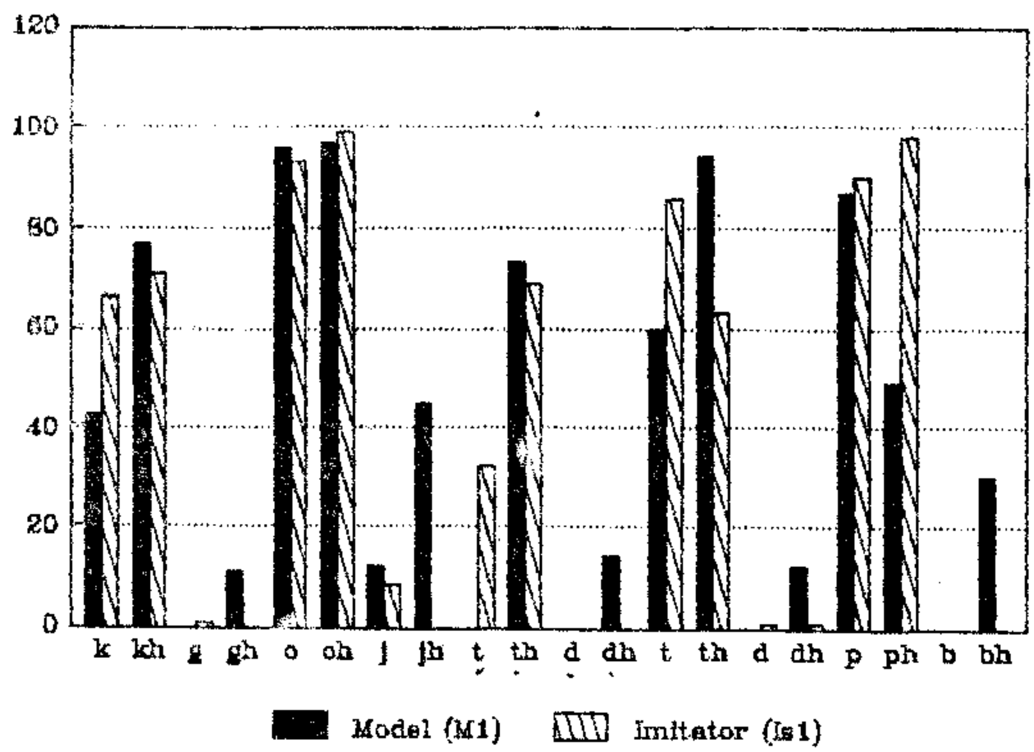


Fig.19 Scaled VOT of M1 & IS1.

Of the spectral parameters, except the burst amplitude and the amplitude of the following consonant none of the parameters were maintained for more than 75% of times by the imitator. F1 was significantly different in k ,g ,d ,p ,b ,bh , F2 was significantly different in kh, g, d, p, b, t, d, dh & F3 was significantly different in all consonants except dh, ph and affricates. Table 4 shows the significant differences between the various parameters of Ml and Isl. Figures 20 to 24 show the spectral paramters of Ml & IS1.

Graded parameters exhibited that scaled aspiration and affrication durations, STF1, STF2, amplitude of the following consonant, TDF1 and burst amplitude were similar in the model and the imitator for more than 75% of times.

Pho neme	F1	F2	F3	B1	B2	B3	L1	L2	L3	BA	DA	AV	AC
k	+		+	+	+			+			+	+	
kh			+							+		+	
g	+	+	+								+	+	
gh			+	+	+						+		
c						+	+	+				+	
j								+		+	+		
th											+		
d	+	+	+	+	+	+	+	+	+	+			
t		+	+		+								
d				+	+		+		+		+		+
dh		+							+				
P	+						+	+	+		+		
ph													
b		+	+	+	+		+		+			+	
bh	+								+			+	

Table 4s Significance of difference between the spectral parameters of Ml & IS1.

Parameters	%times imitated
Scaled aspiration duration	100
Scaled affrication duration	100
STF2	92
Amplitude of following consonant, STF1	85
TDF1 and Burst amplitude	77
B3	71
TDF2, Scaled VOT	69
L1	64
Burst duration	61
Scaled closure duration, scaled voicing duration and L3	57
STF3, Scaled burst duration	54
FO movemant of that vowel following voiceless stops	52
murmur duration, F1, Bl, 82, L2	50
Amplitude of the following vowel	46
Closure duration, Voicinq duration, VOT	43
F2	42
TDF3, Overall amplitude	38
Fo movement of the vowel following voiced stops	36
F3	28
Word duration	a
Aspiration duration	0
Affrication duration	0

Table 5 : Parameters imitated percent times by IS1 (of Ml)

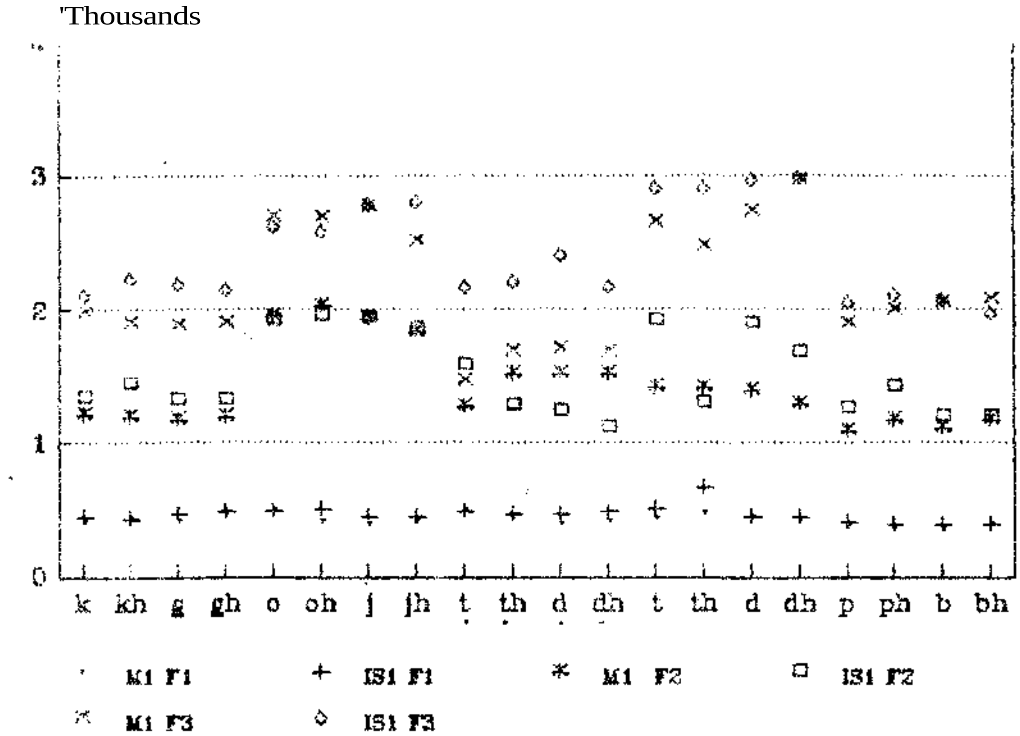


fig.20 Formant frequencies of M1 & IS1 (KHz)

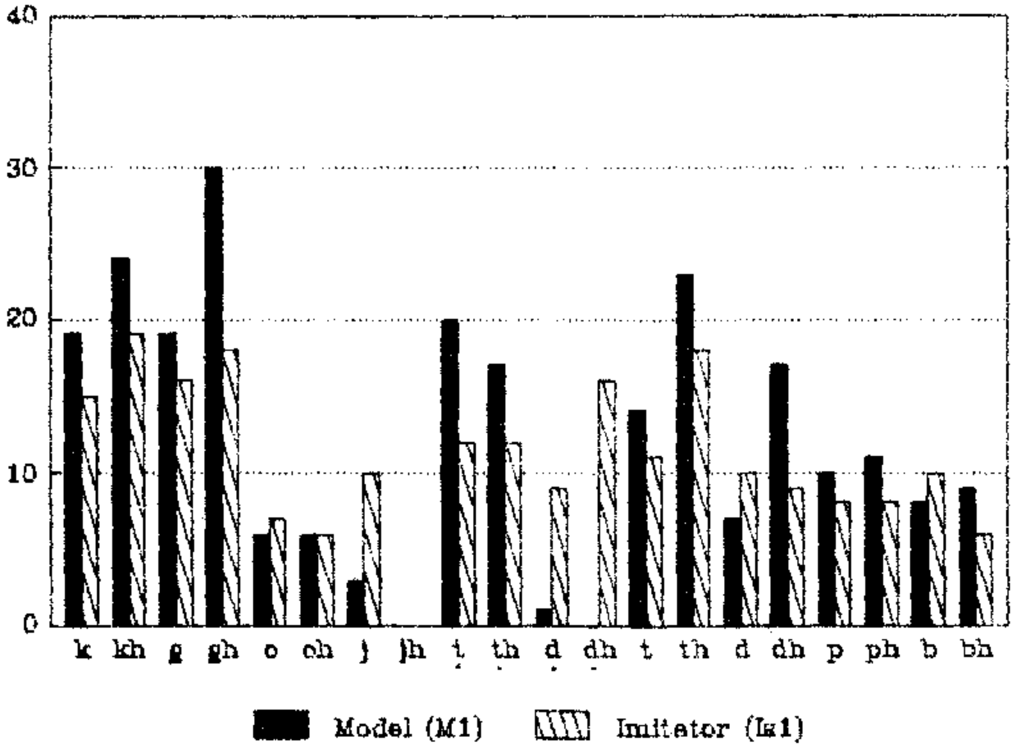


Fig.21 Burst amplitudes of M1 & IS1 (RdB)

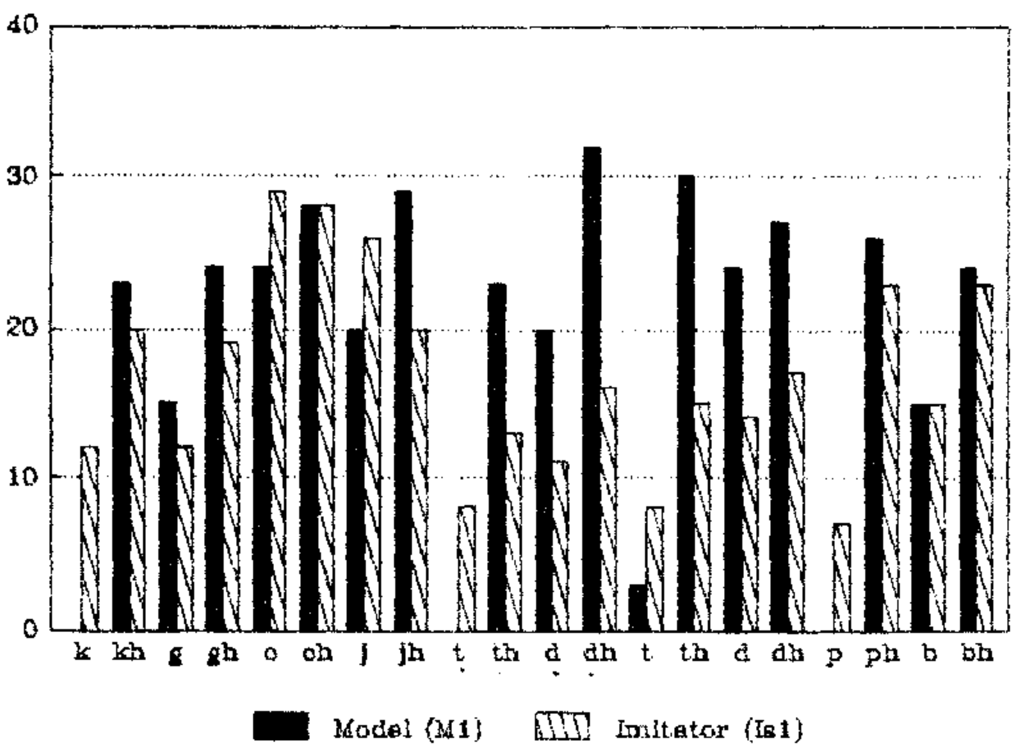


Fig.22 Overall amplitude of M1 & IS1 (RdB)

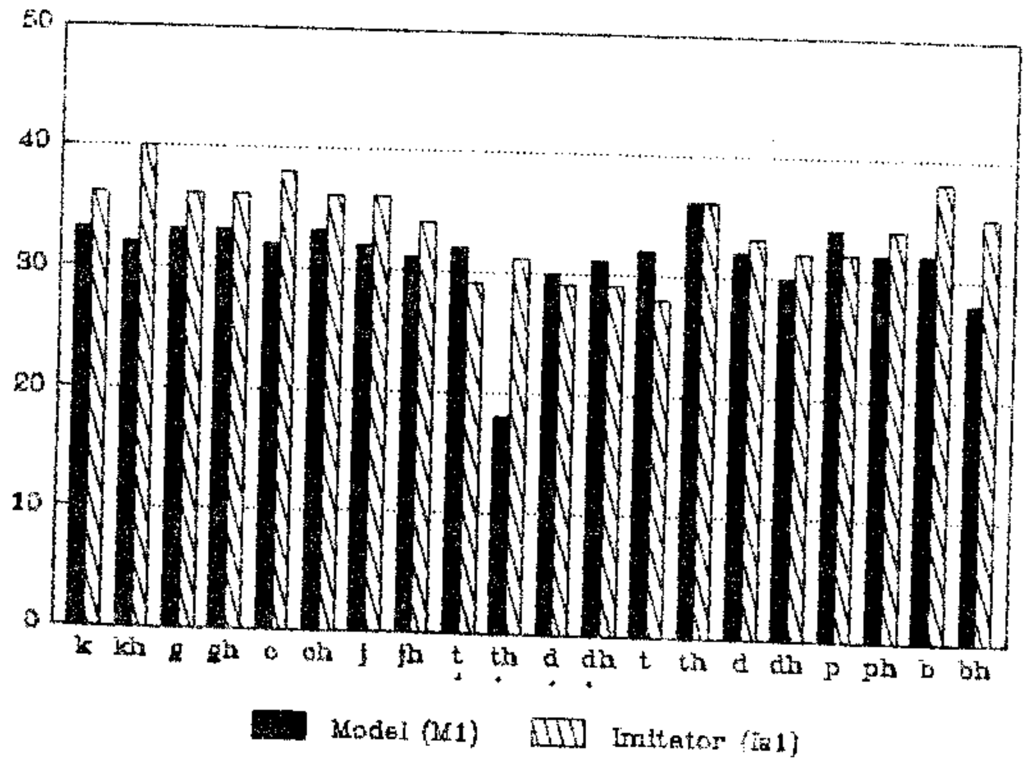


Fig.23 Amplitude of the following vowel of M1 & IS1 (RdB)

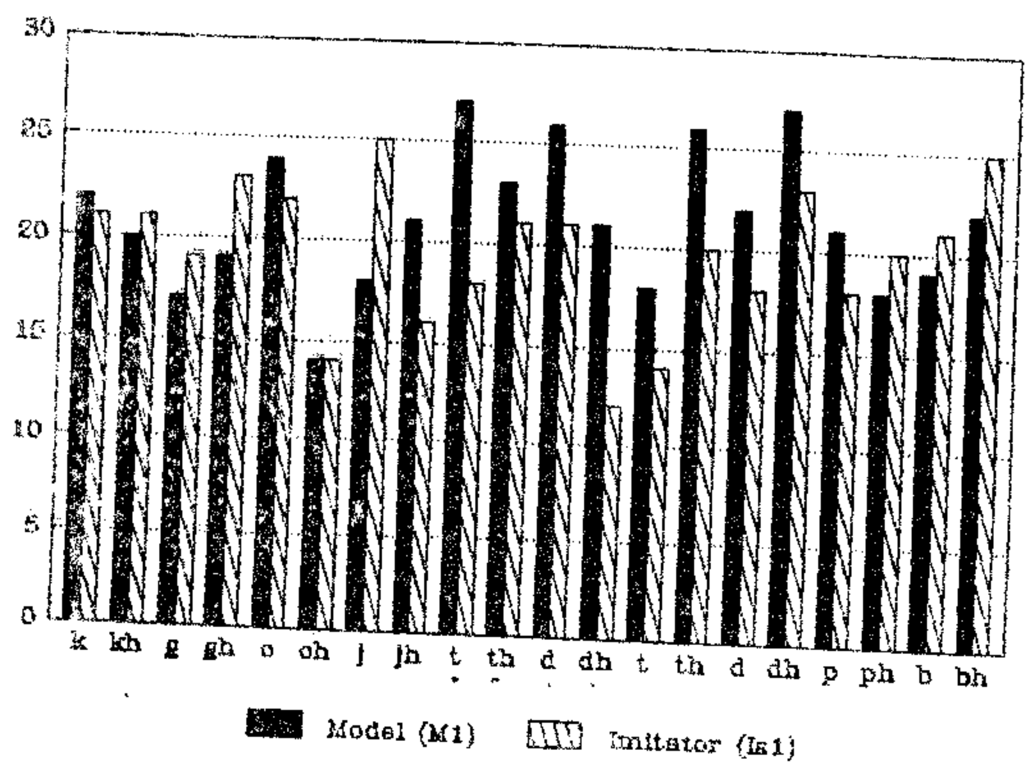


Fig.24 Amplitude of the following consonant of M1 & IS1 (RdB)

In VCV condition, the terminal *F1.* F2, & F3.5 of the imitator did not seem to match these of the model. Especially the terminal F2.s of the velars in the imitator were lower than that of the model. Fig 25 shows the terminal F1, *F2, &* F3 of the model and the imitator.

IS1 exhibited 52% fall patterns of Fo in the vowel following of the voiceless stops/affricates, 43% fall pattern, 57. rise? patterns. In vowels following voiced stops/affricates, 187. had falling Fo, 467. had flat, Fo and 35% had rising Fo patterns. In the model 69%, 34%, 21%, 25% & 10%, 41% had falling, flat and rising patterns in voiceless & voiced conditions respectively. Fig 26 depicts the percent of rise (R) , fall (F) & flat (Ft.) patterns of Fo of the vowel following voiced and voiceless stops/affricates.

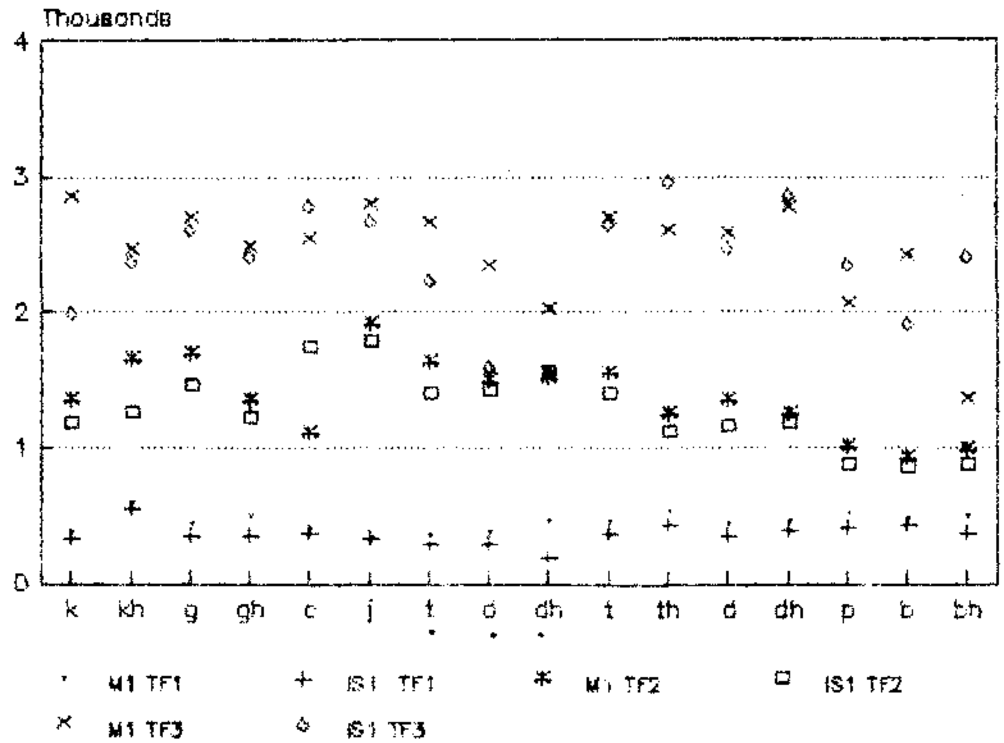
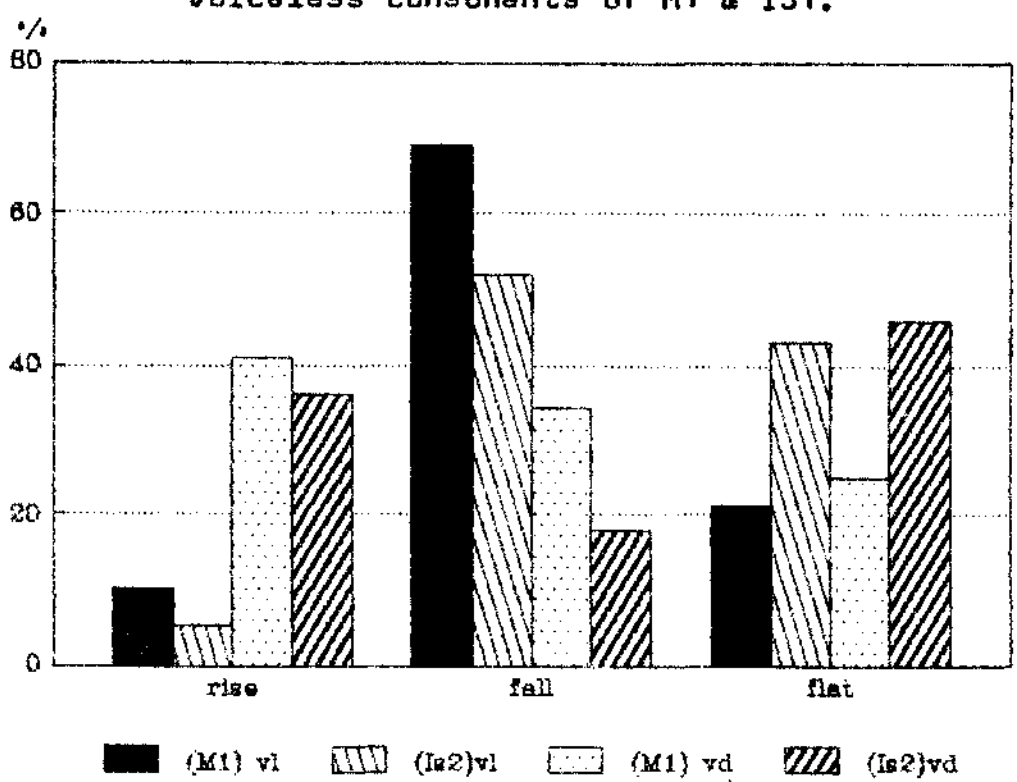


Fig.25 Terminal formant frequencies of (Preceding Vowel) M1 & IS1 (KHz)

Fig.26 F0 contours for Vowels following voiced & voiceless consonants of M1 & IS1.



On perceptual analysis 63% was evaluated as very good imitation, 14% as fair, 6% as poor and *17%* as bad by the first judge & 9% as very good, 52% as fair, 36% as poor & 3% as bad by the second judge. The evaluators felt that of the three imitations this was the best and whenever it was rated very good, it was the best of the two imitations. Whenever the imitation resembled the model completely in terms of articulation, stress, aspiration, syllable duration and pause?, it was judged very good. Deviations in nasality were judged fair (The model exhibited nasal tinge in almost all the words). Distortion of the consonants were judged poor and mi sarticulation (substitution and distortion) especially on retroflexs, improper aspiration and durations in stress, syllable deviations were judged bad.

MODEL (Ml) Vs IMITATION (IS25

On Walsh test it was observed that the temporal parameters TDF1, TDF2, TDF3, STF1, STF2, STF3 were imitated by IS2 most of the times. However, individual differences were present. While the closure durations of */g/* and /b/ were significantly **different**, voicing durations of */g/ &* /bh/ were different significantly. Significant differences were present in the burst durations of k , g, gh, d, and t, aspiration duration of kh, murmer duration of gh, VOT of k , kh, g, th, t. & bh, TDF1 of g, TDF2 of *d* & b, STF2 of k, STF3 of gh. Word durations in all the 551 tokens of the Imitator were longer than the model. Figures 27 to 39 represent the averaged temporal parameters used by MI & 1S2. Table 6 represents the significant difference between the various temporal parameters of Ml & IS2.

The relative timings of the various temporal parameters **indicated** that IS2 used closure duration, aspiration **duration,** affrication duration and voicing duration, in a scaled manner. Figures 40 to 45 represent the scaled temporal parameters of M1 & IS2 and Table *7* depicts the significance of difference between the various scaled temporal parameters of the model and the imitator.

Pho neme	CD	VD	BD	AFD	AD	MD	VOT	TD	TD F1	TD F2	ST F3	ST F1	ST F2	WD F3	
k				+				+					+		+
kh						+		+							+
g		+	+	+				+	+						+
gh				+			+							+	+
c															+
j					+			+				+			+
th								+							+
ḋ															+
t́				+				+							+
d										+					+
p															+
b		+								+					+
bh			+					+							+

Table 6: Significant difference in the temporal parameters of M1 and IS2.

Phonemes	CD	VD	BD	AFD	AD	VOT
k			+			+
kh			+			
g			+			+
gh						
c						
j						
ṭ			+			+
ṭh						+
ḍ						
ḍh			+			+
t			+			+
th			+		+	+
d			+			+
dh						+
p			+			+
ph						
b						+
bh						+

Table 7 : Significant differences between scaled temporal parameters of M1 & IS2.

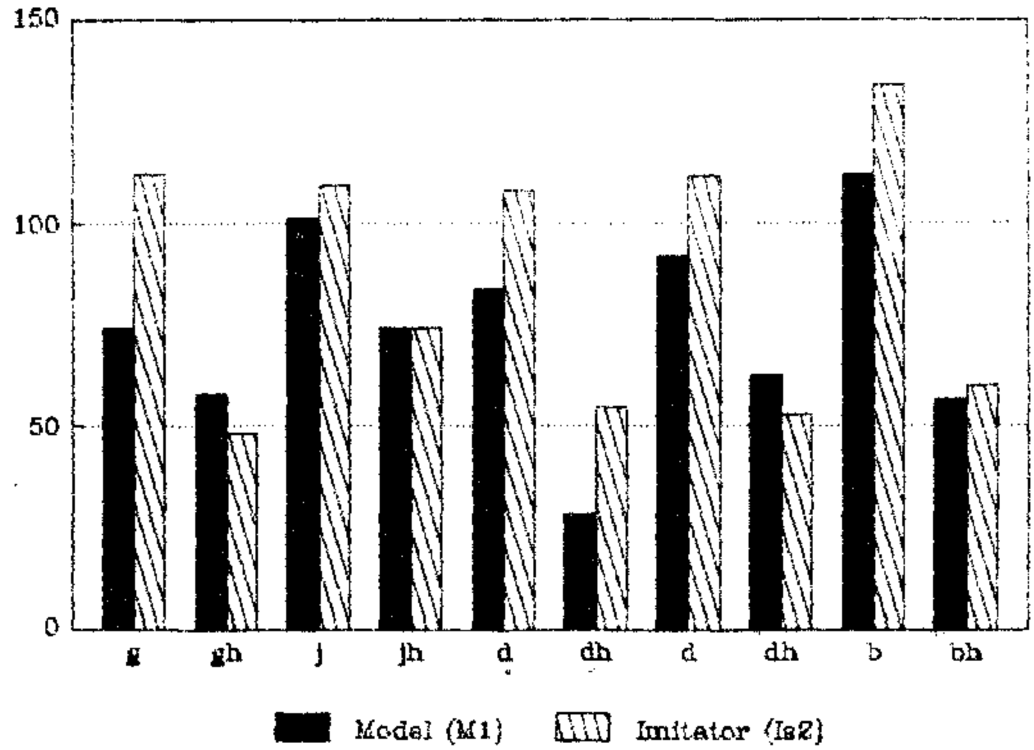
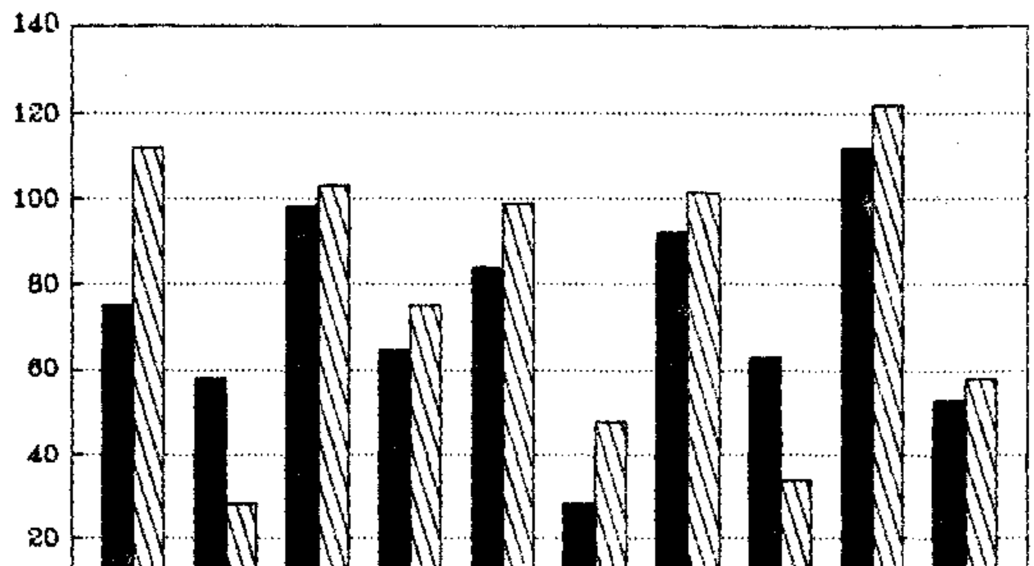


Fig.27 Closure durations of M1 & IS2 (M.Sec)

Fig.28 Voicing duration of M1 & IS2 (M.sec)



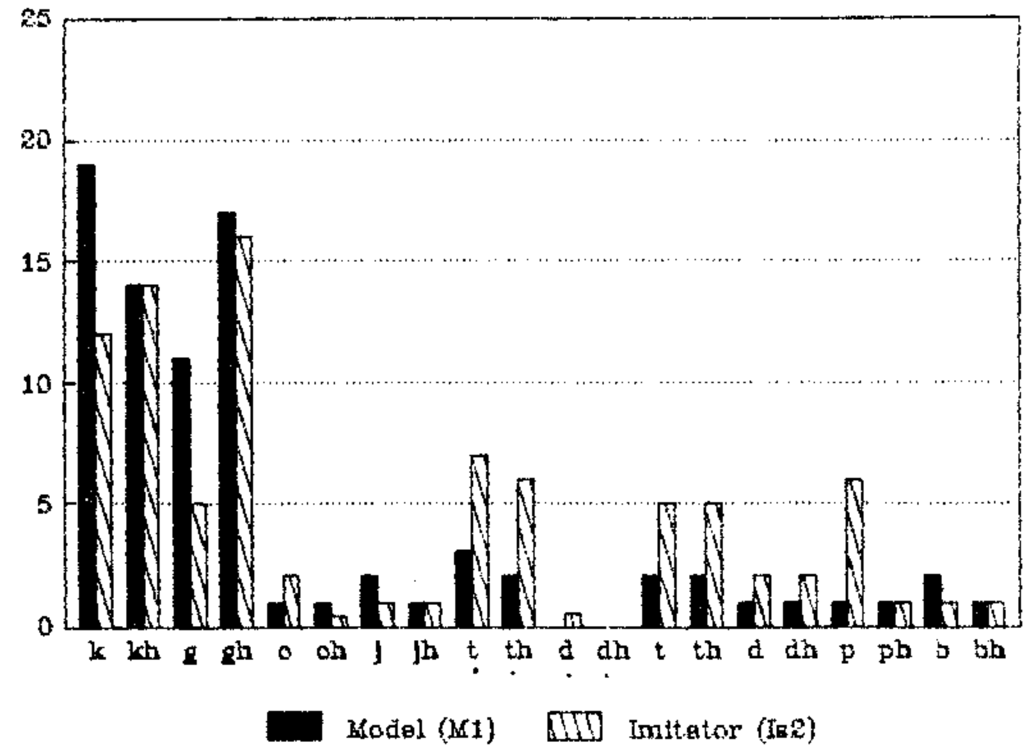
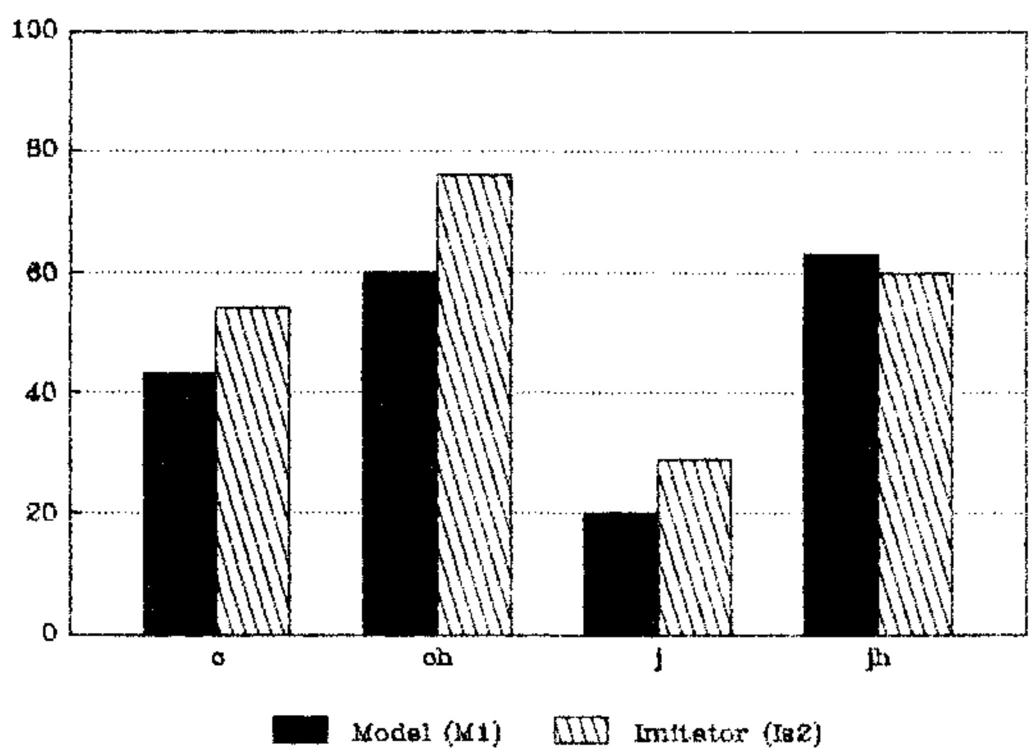


Fig.29 Burst duration of M1 & IS2 (M.sec)

Fig.30 Affrication duration of M1 & IS2 (M.sec)



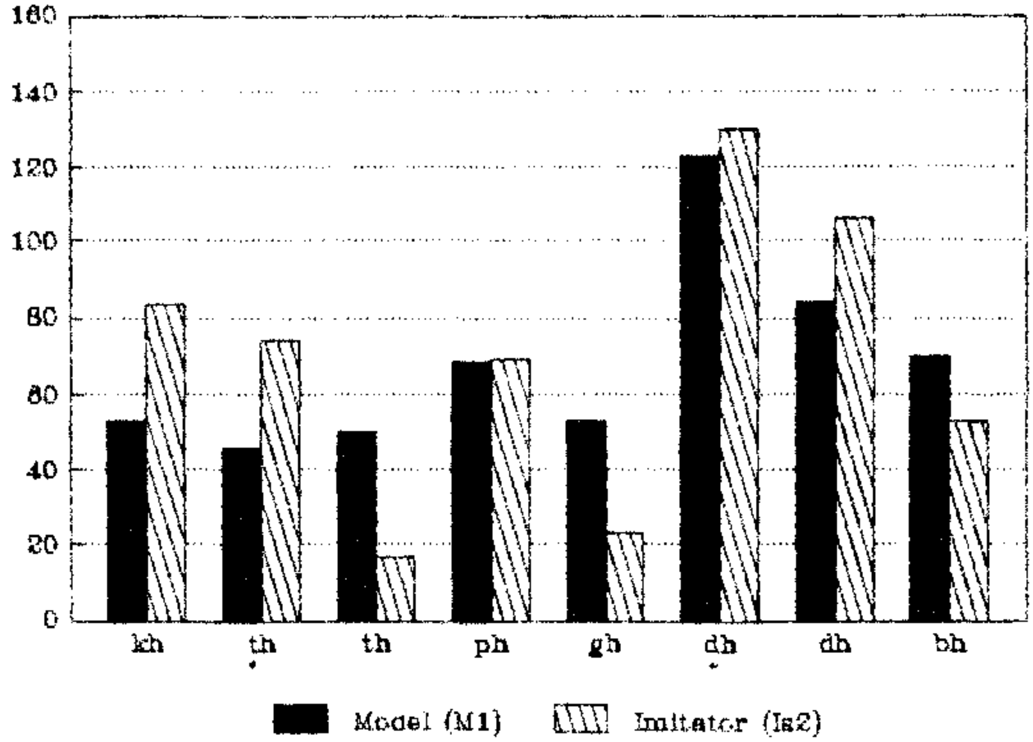
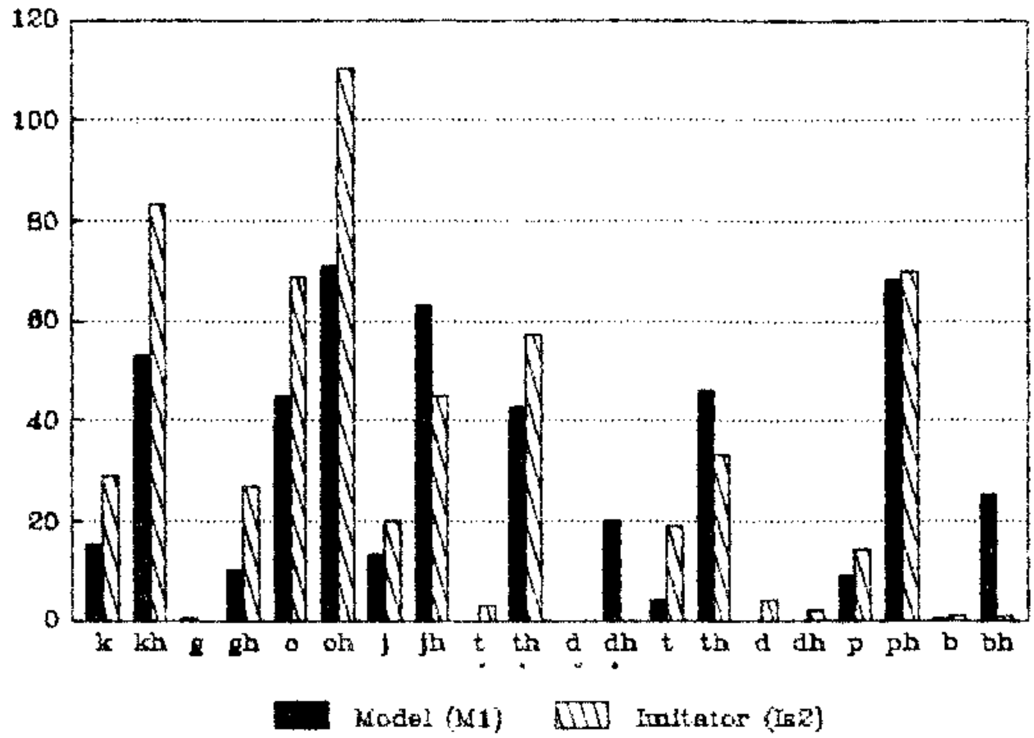


Fig.31 Aspiration & murmur duration of M1 & IS2 (M.Sec)

Fig.32 VOT of M1 & IS2 (M.sec)



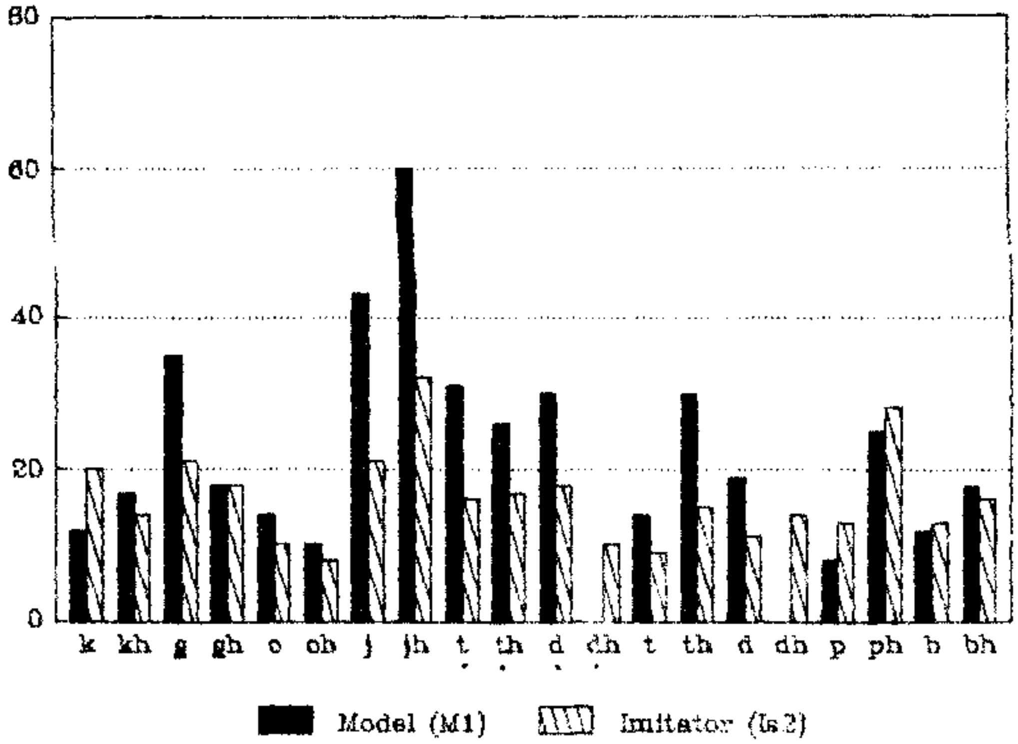
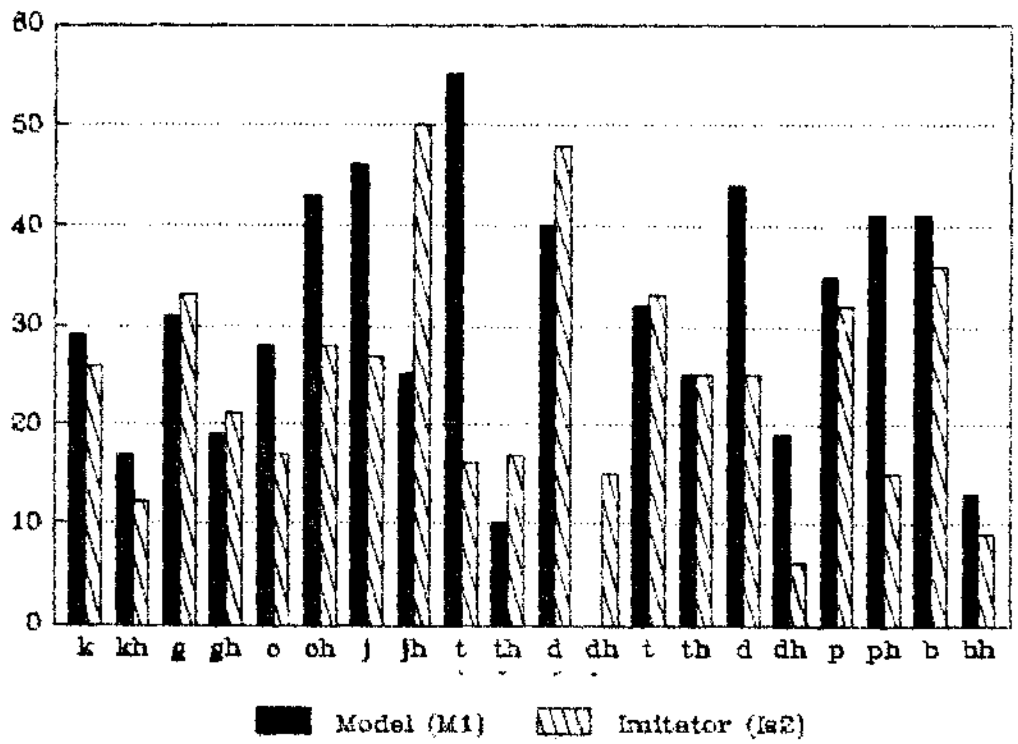


Fig.33 Transition duration of the first formant of M1 & IS2. (M.sec)

Fig.34 Transition duration of the second formant of M1 & IS2 (M.sec)



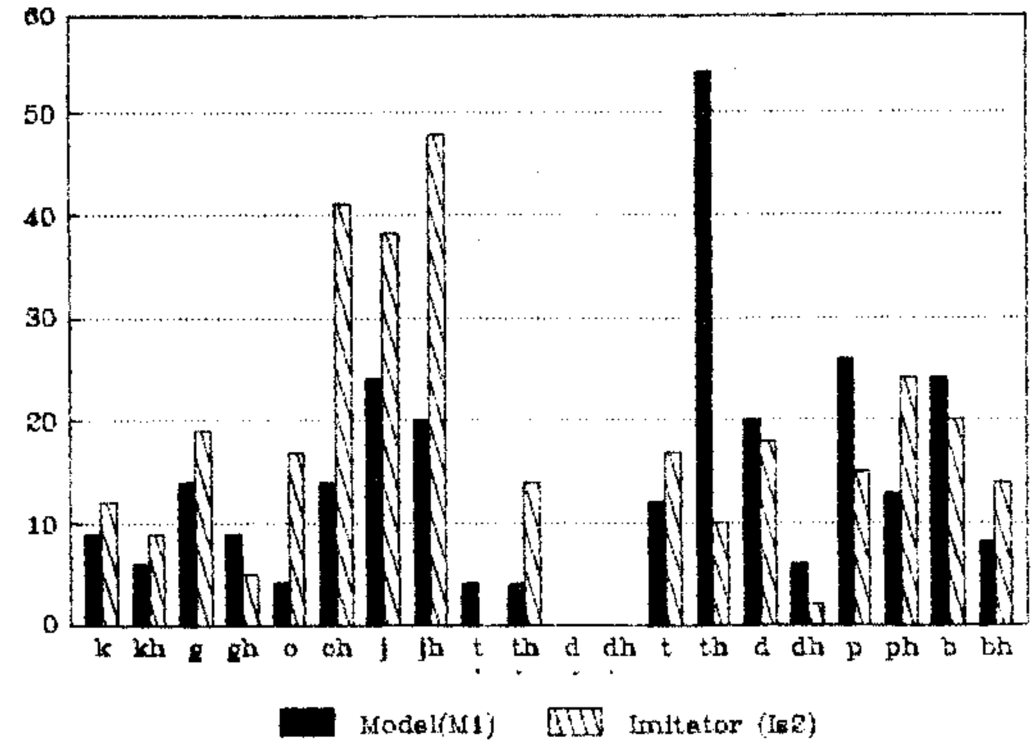
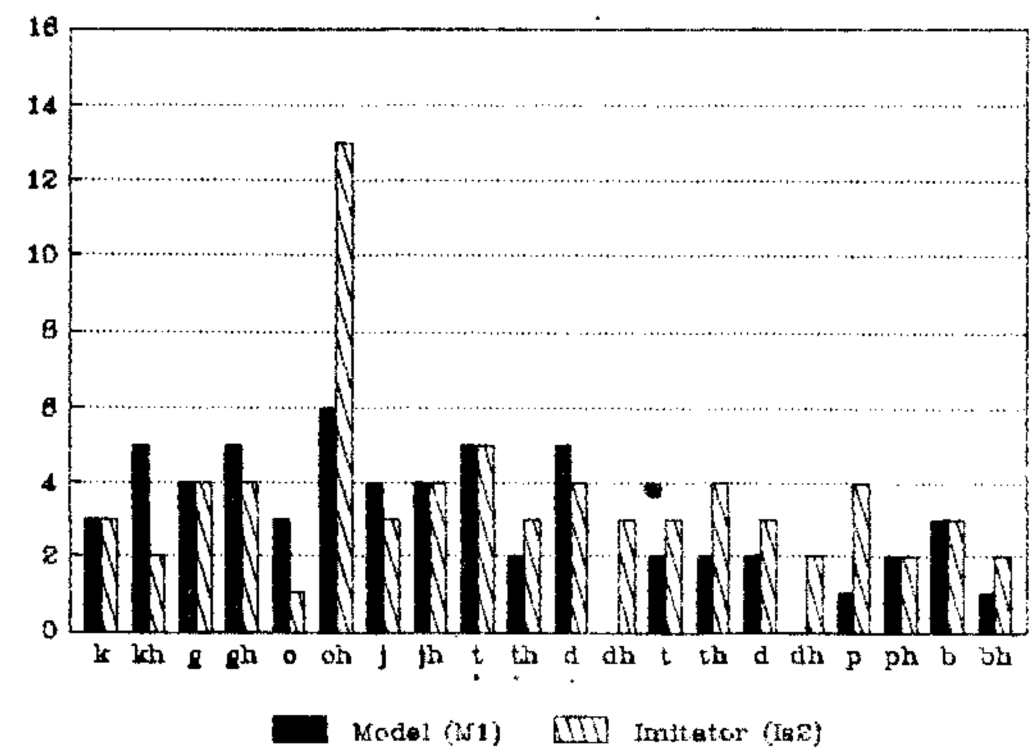


Fig.35 Transition duration of the third formant of M1 & IS2 (M.sec)

Fig.36 Speed of transition of the first formant of M1 & IS2 (Hz/M.sec)



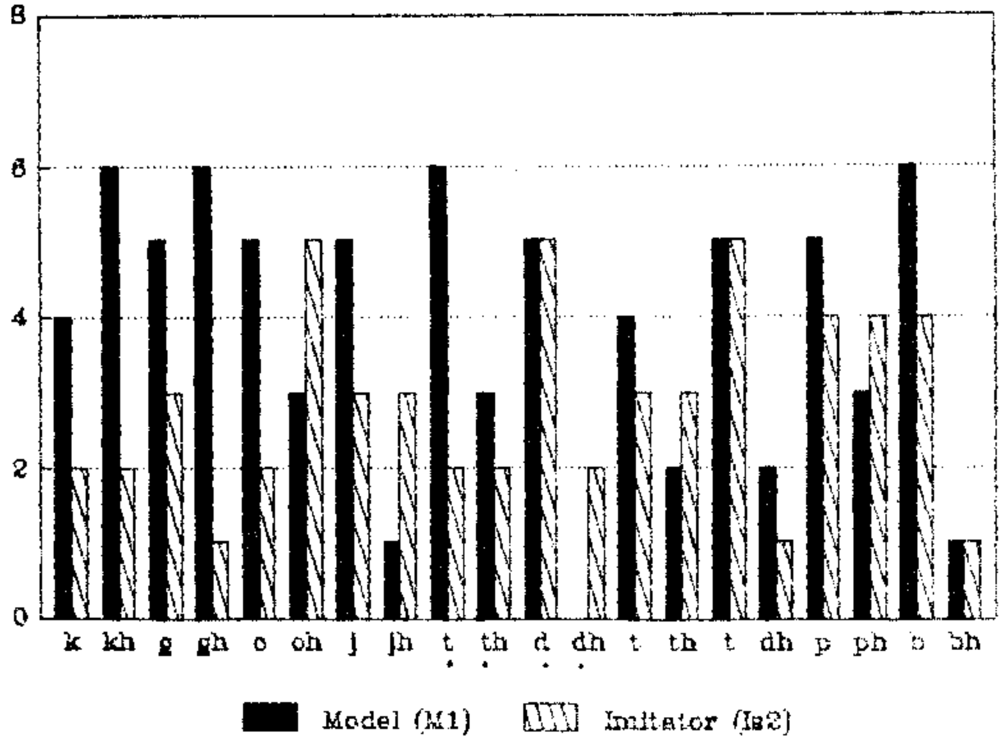
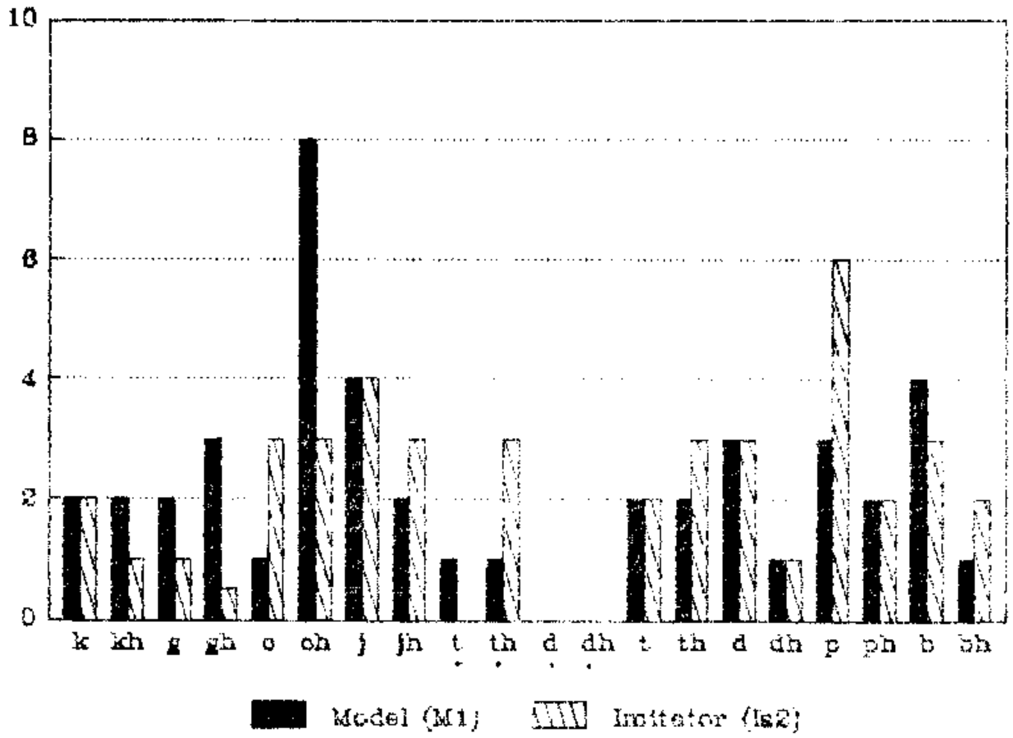


Fig.37 Speed of transition of the second formant of M1 & IS2 (Hz/M.sec)

Fig.38 Speed of transition of the third formant of M1 & IS2 (Hz/M.sec)



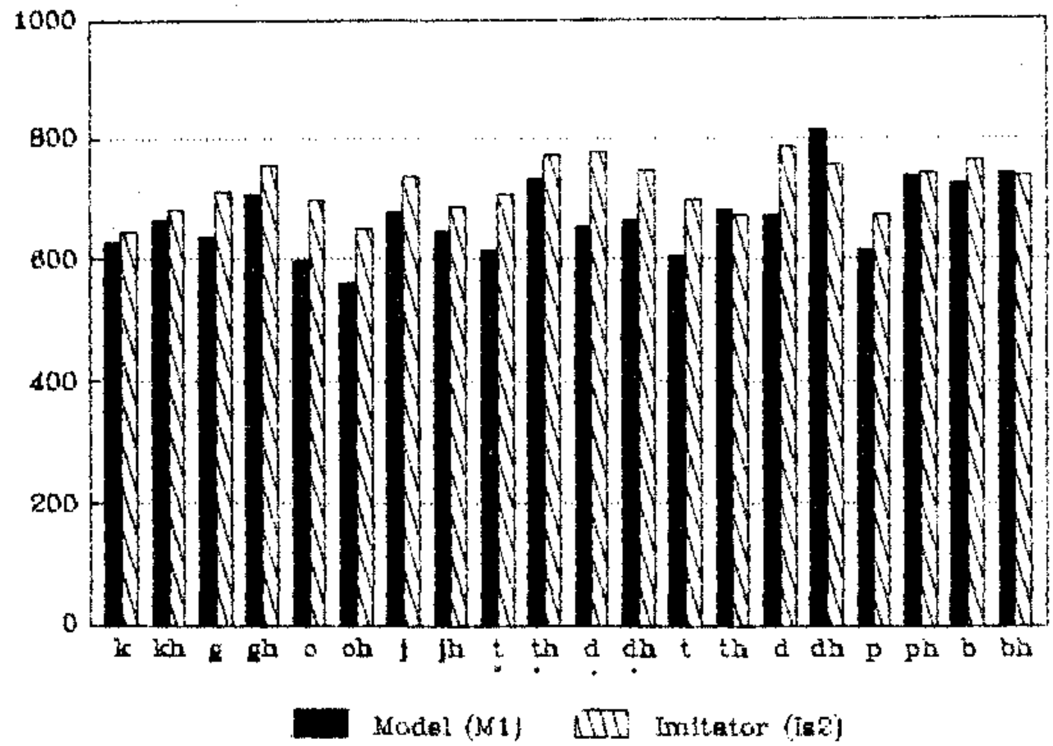
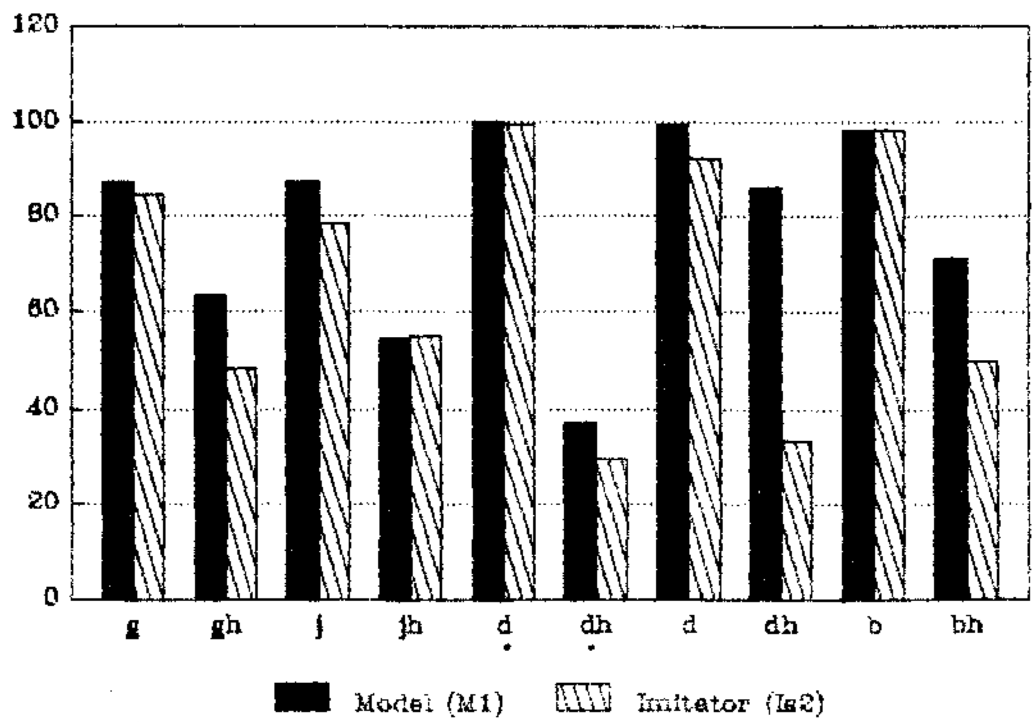


Fig.39 Word durations of M1 & IS2 (M.sec)

Fig.40 Scaled closure duration of M1 & IS2.



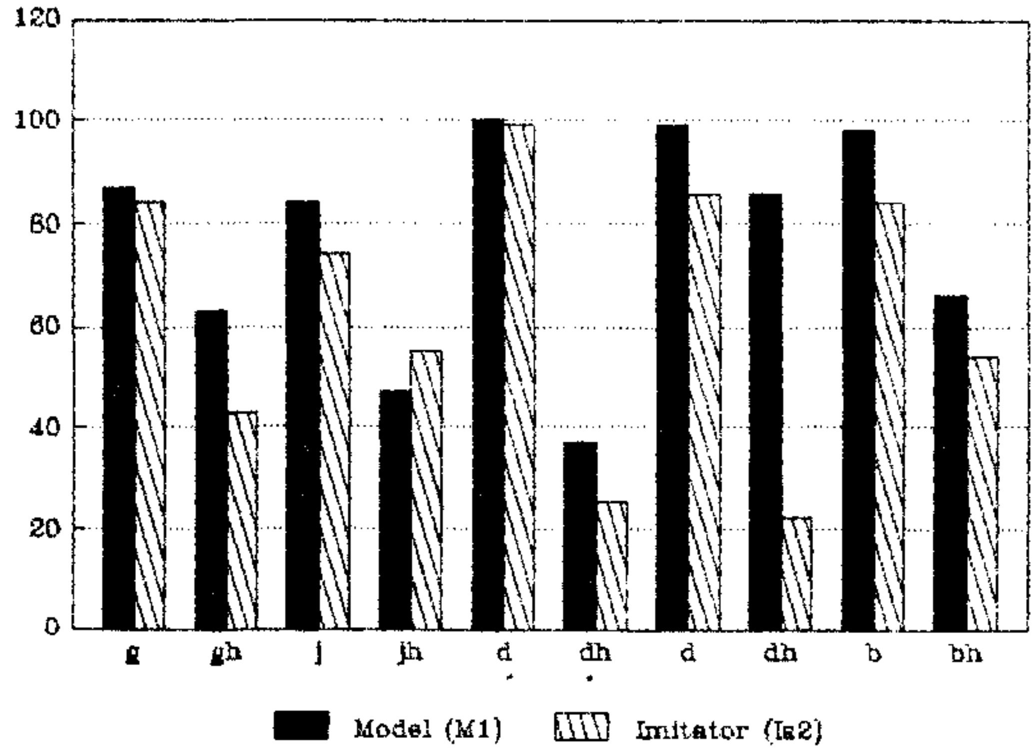
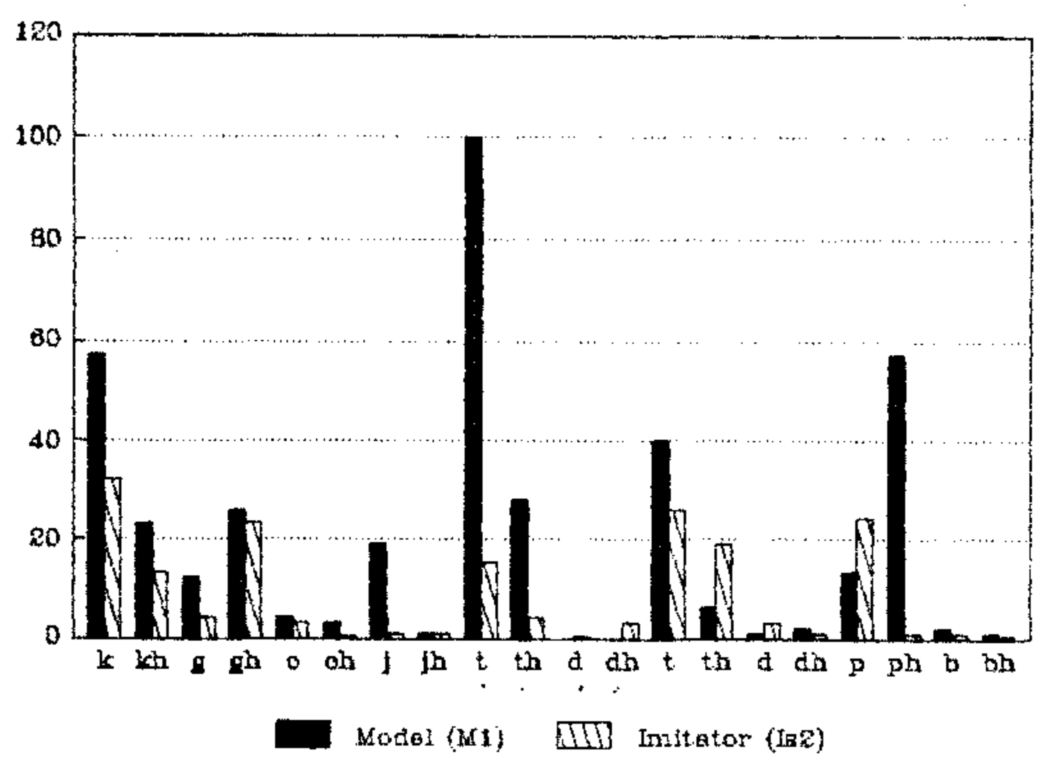


Fig.41 Scaled voicing duration of M1 & IS2.

Fig.42 Scaled burst duration of M1 & IS2.



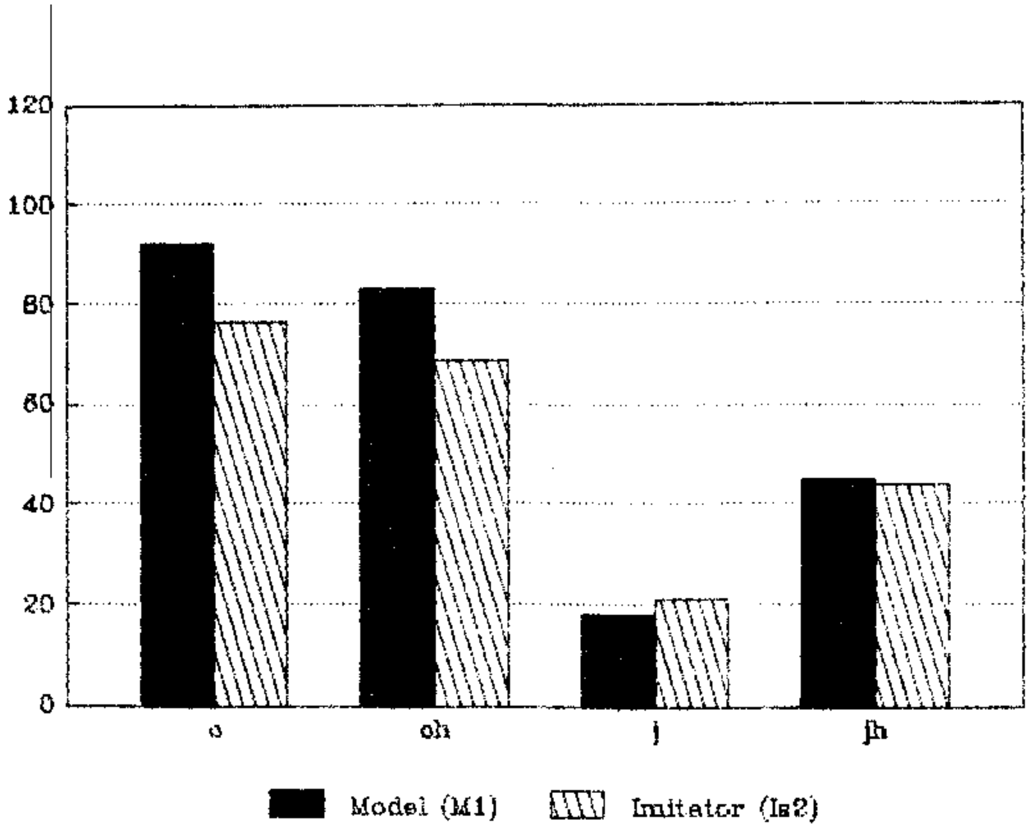
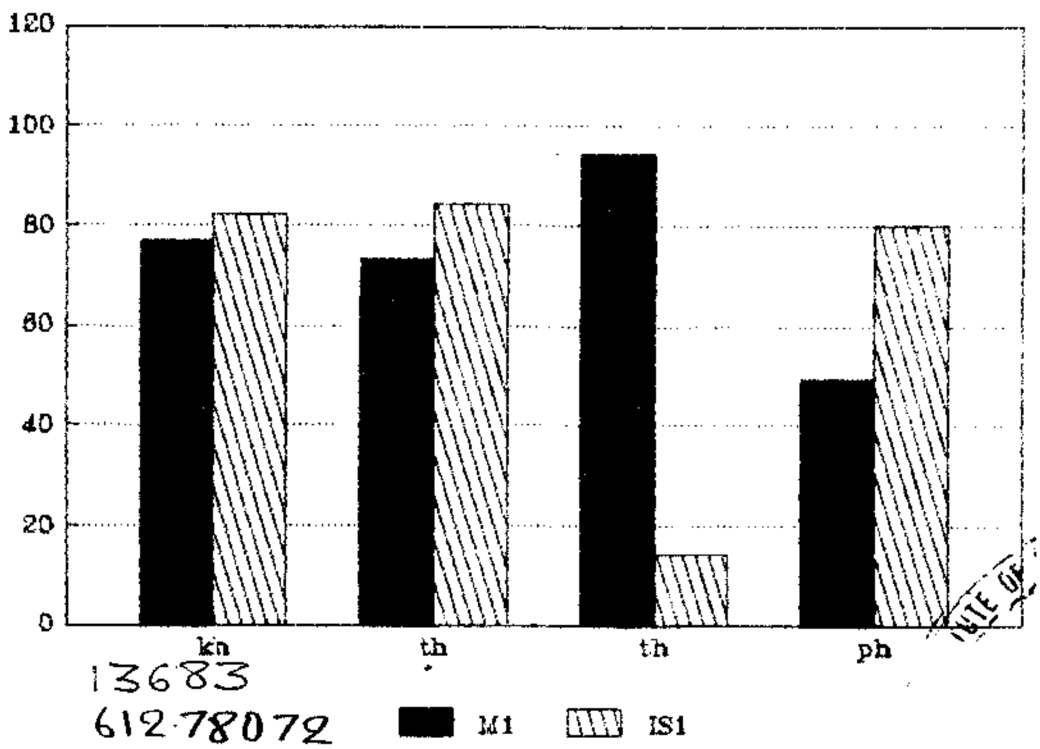


Fig.43 Scaled affrication duration of M1 & IS2.

Fig.44 Scaled aspiration duration of M1 & IS2.



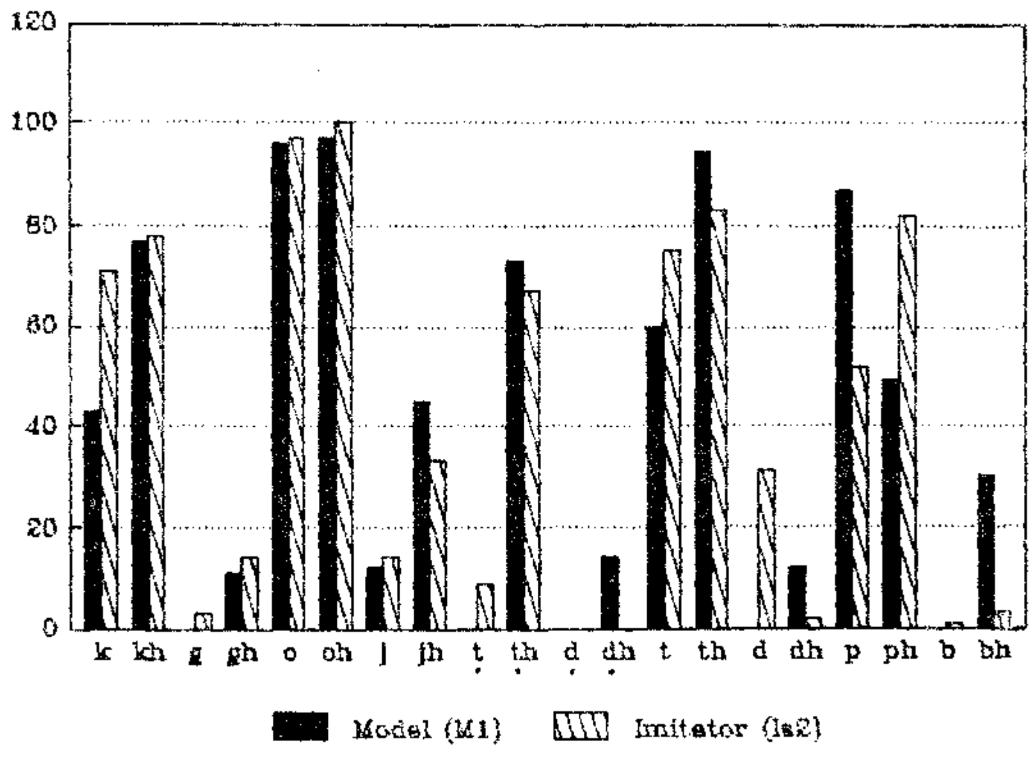


Fig.45 Scaled VOT of M1 & IS2.

Of the spectral parameters B2, B3, L2, & burst amplitude were the parameters imitated. The formant patterns of the model and the imitator did not agree inspite of the perceptual judgements being good imitation. Though the formant frequencies were imitated, there were significant differences in the F1 of g, d, p, ph, b, bh, F2 of g, gh, dh, p, ph, b, bh, & F3 of k, kh, g, gh, t & p. (As the LPC measurements provided erroneous bandwidths and levels, they were considered only for the purpose of comparison. However, they were not considered as values in production). Figures 46 to 50 depict the various spectral parameters of the model and the imitator and table 8 shows the significant differences in the spectral parameters of Ml & IS2.

On grading the temporal, scaled temporal and spectral paramters, it was noticed that scaled voicing duration, scaled closure duration, TDF3, TDF1, STF1, STF2, STF3, L2, B2, burst amplitude, TDF2 and B3 were maintained by the imitated for more than 75% of times. Table 9 depicts the parameter imitated by IS2 in percent.

Phone names	F_1	F_2	F_3	B1	B2	B3	L1	L2	L3	BA	DA	AV	AC
k			+								+	+	+
kh			+	+					+			+	+
g	+	+	+		+	+					+	+	
gh		+	+				+		+	+	+	+	
c													
ch							+				+		
j				+									
jh													
ṭh													+
ḍ	+				+						+		
ṭ			+									+	
d						+				+	+		+
dh		+				+	+		+		+		+
p	+	+	+	+			+	+				+	
ph	+											+	
b	+	+		+			+		+			+	
bh	+	+		+			+		+		+		

Table 8: significant difference between the spectral parameters of M1 & IS2.

Parameter	% times imitated
Scaled closure duration	100
Scaled voicing duration	100
TDF3	100
L2	94
TDF1, STF1, STF2, STF3	92
B2, Burst amplitude	88
TDF2	85
B3	82
Scaled A.D	
Closure duration, Voocing duration, B1, L3, Amplitude of the following consonant	71
Burst duration	70
Fo movement of the vowel Following voiceless stops	69
F_1, F_2, F_3, L1	65
Overall amplitude, Amplitude of the following vowel	53
Affrication, Aspiration, murmer duration, Scaled BD	50
VOT	46
Fo movement of the vowel following vd stops	35
Scaled VOT	33
Word duration	0

Table 9: Parameters imitated percent times by IS2.

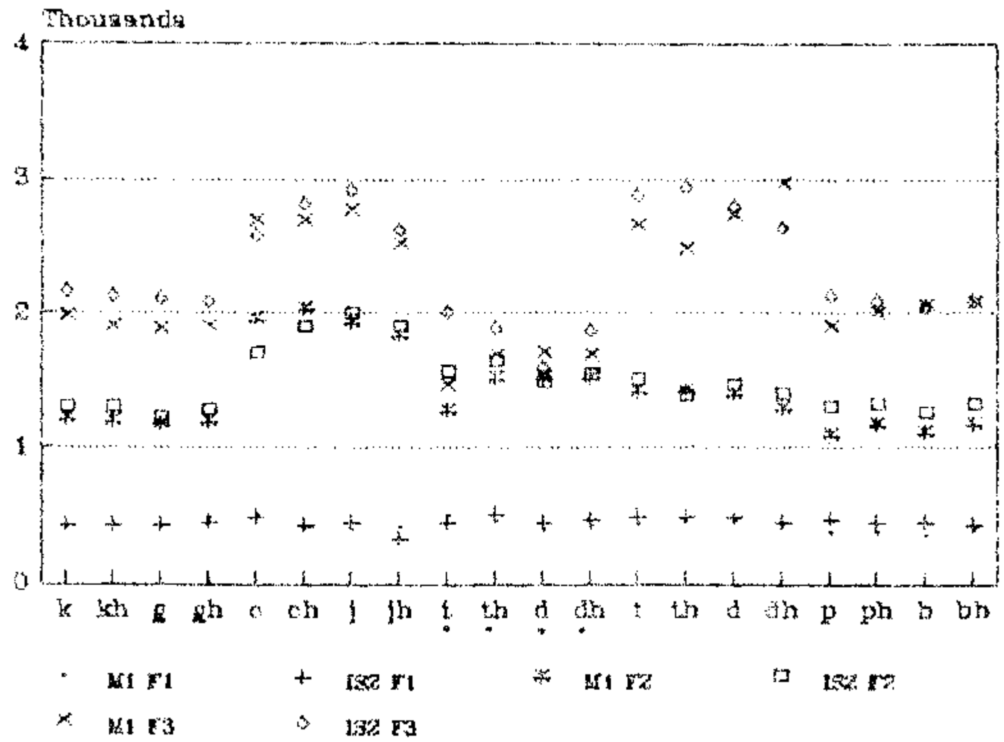
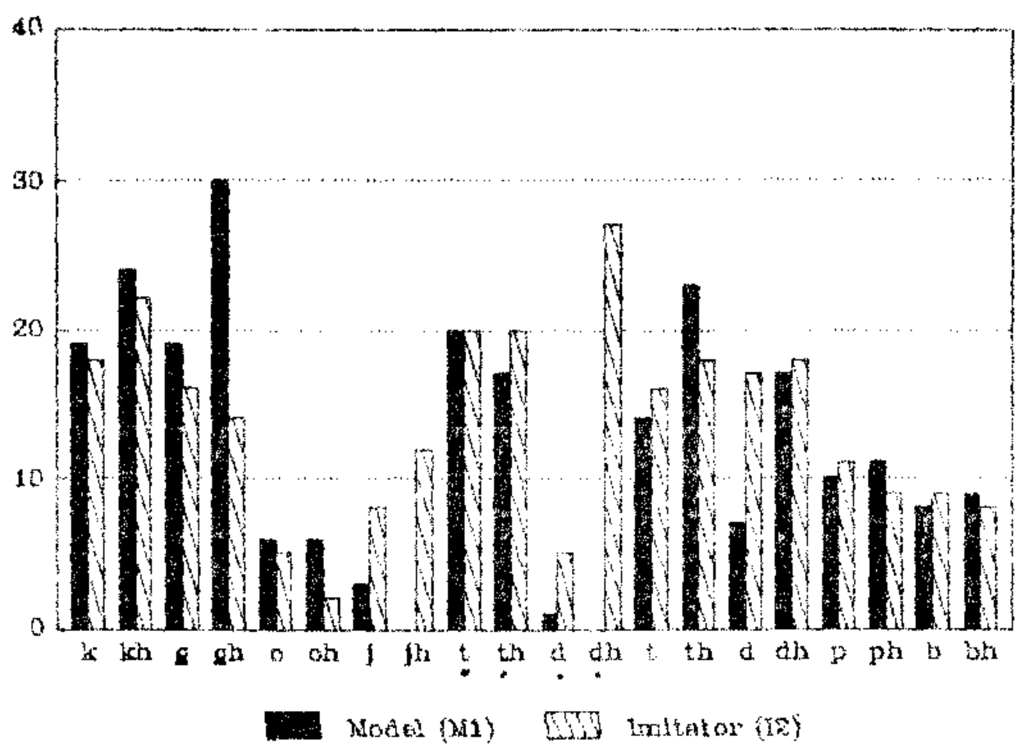


Fig.46 Formant frequencies of M1 & IS2, (KHz)

Fig.47 Burst amplitude of M1 & IS2 (RdB)



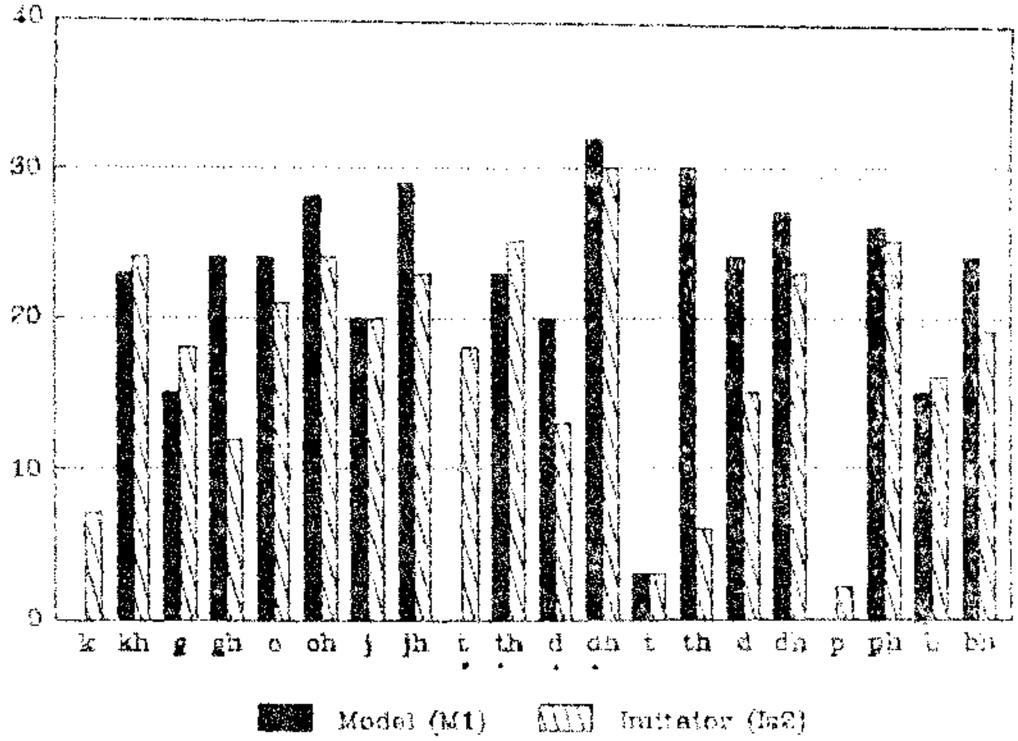
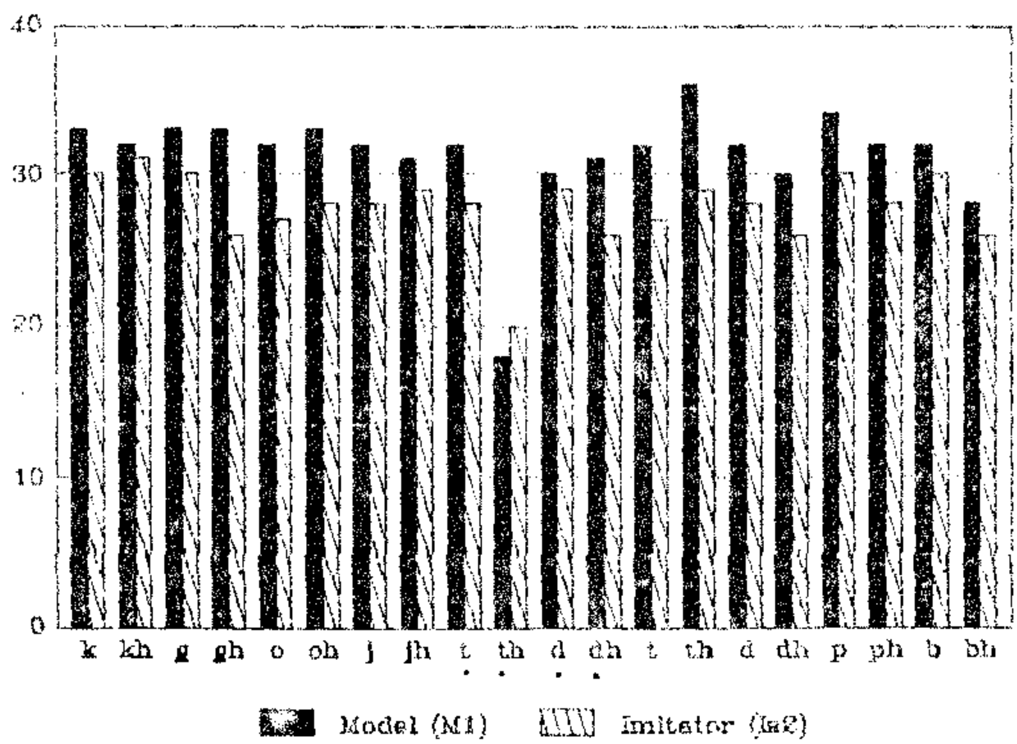


Fig.48 Overall amplitude of M1 & IS2 (RdB)

Fig.49 Amplitude of the following vowel of M1 & IS2 (RdB



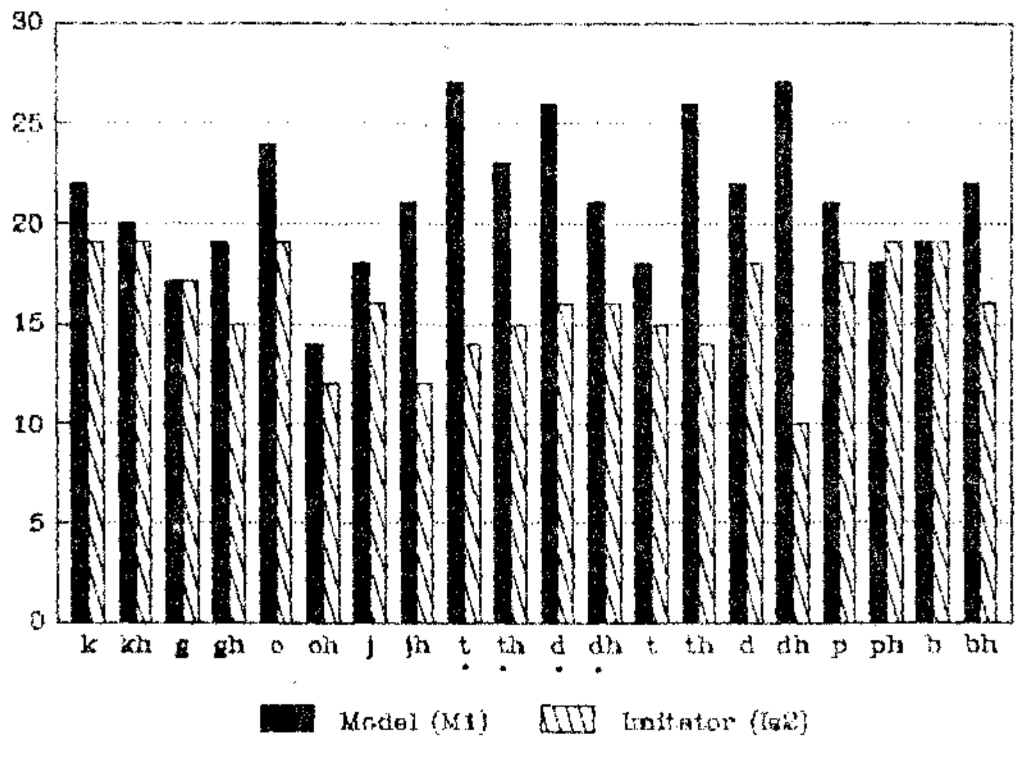


Fig.50 Amplitude of the following consonant of M1 & IS2 (RdB)

In VCV condition the terminal frequencies of the first two formants of the imitator matched that of the model . However, the third formant frequency of IS2 did not seem to match the model. Figure 51 shows the terminal frequencies of Ml & IS2.

Fc movement of the following vowel in voiceless atop/affricate condition had 71% fall pattern (F) , 21% flat pattern (Fl> & 87. rise pattern (R) in the imitator and 69% fall, 21% flat and 10% rise in the model. Following the voiced stop/affricates the FO of vowel had 287. and 34% fall pattern, 37"/. and 25% flat pattern and 35%., 41% rise pattern in the imitator and the model respectively. Fig 52 represent the different Fo movements in M1 & IS2.

Perceptual analysis revealed judgement of 30% as very good imitation, 39% ass fair, 23% as poor and 8% as Bad by the first imitator 22 as very good, 587. as fair, 19% as poor & 1% as bad by the second imitator. When the imitation resembled the model to the extent of articulation, stress, aspiration, pause and syllable duration it was rated very good. Any deviation in nasality (nasal tinge) was judged fair and consonant, distortions were rated poor and deviations in articulation, stress, aspiration & syllable duration were judged poor.

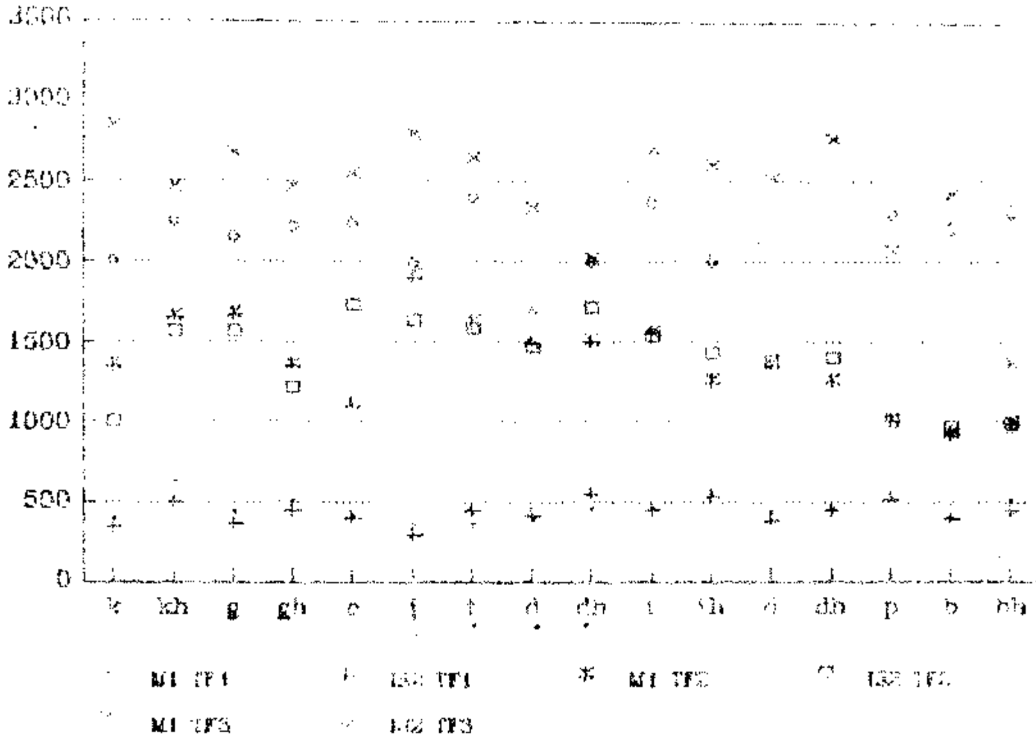
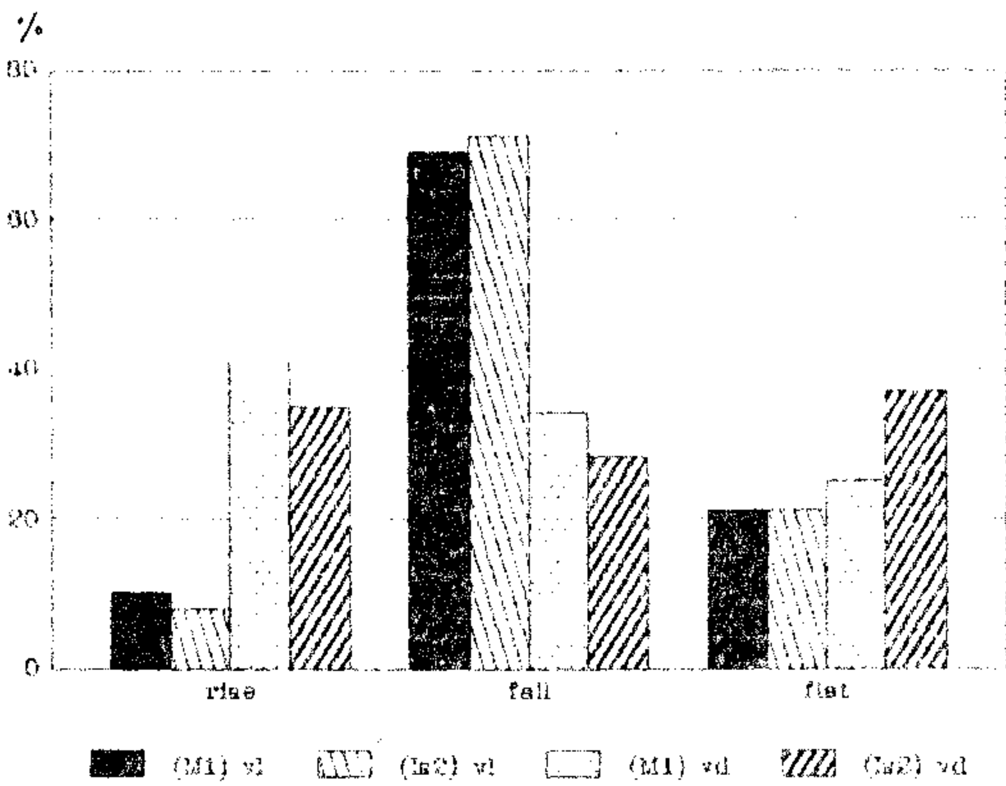


Fig.51 Terminal formant frequencies of (Preceding Vowel) M1 & IS2 (Hz)

Fig.52 F0 contours for vowels following voiced & voiceless consonants of M1 & IS2.



MODEL M2 VS IMITATION II

Of the temporal parameters, STFl, STF2, TDFl, affrication duration & aspiration duration were similar in the model and the imitator for more than 75% of times. However, there were significannt differences between the voicing durations of g, gh, j, d, b, bh, closure durations of d, b, bh, burst durations of c, j, t, d, p, b, bh, affrication duration of c, murmtsr duration of gh, VOT of g, c, j, t, d, TDF1, of g, c, j, TDF2 of c, j, TDF3 of kh, *c,* j, d, p, STF1 of g, t, STF2 of c & STF3 of kh, c, t, t, d, Word durations in all the 551 conditions were longer in the imitator than in the model. The temporal measures of the model and the imitator *are* represented in figs 53 to 65. Table 10 shows the significance of difference between the various parameters of the M2 & I1.

On comparing the scaled temporal measures of M2 *And* II, it was observed that Aspiration duration was used by the imitator as a cue. Of the six scaled temporal parameters, voicing duration was a parameter which was used least in imitation. Figures 66 to 71 show the various scaled temporal parameters of M2 & II. Table 11 represents the significance of difference between M2 & 11, on various scaled temporal parameters.

Pho neme	CD	VD	BD	AFD	AD	MD	VOT	TD F1	TD F2	TD F3	ST F1	ST F2	ST F3	WD
k			+						+	+			+	+
kh										+			+	+
g		+					+	+	+	+	+		+	+
gh		+				+								+
c							+	+	+	+		+		+
j		+	+				+	+						+
ṭh														+
ḍ	+	+	+				+			+			+	+
t													+	+
d	+	+	+				+			+			+	+
p			+							+				+
b	+	+	+											+
bh	+	+	+											+

Table 10: Significance of difference between various temporal parameters of M2 & I1.

Phonemes	CD	VD	BD	AFD	AD	VOT
k			+			+
kh						
g		+	+			+
gh		+				
c			+	+		
j	+	+	+	+		+
ṭh						
ḍ						
th						
d	+	+				+
p			+			
b		+	+			
bh	+	+	+			

Table 11: Significance of difference between various scaled temporal parameters of M2 & I1.

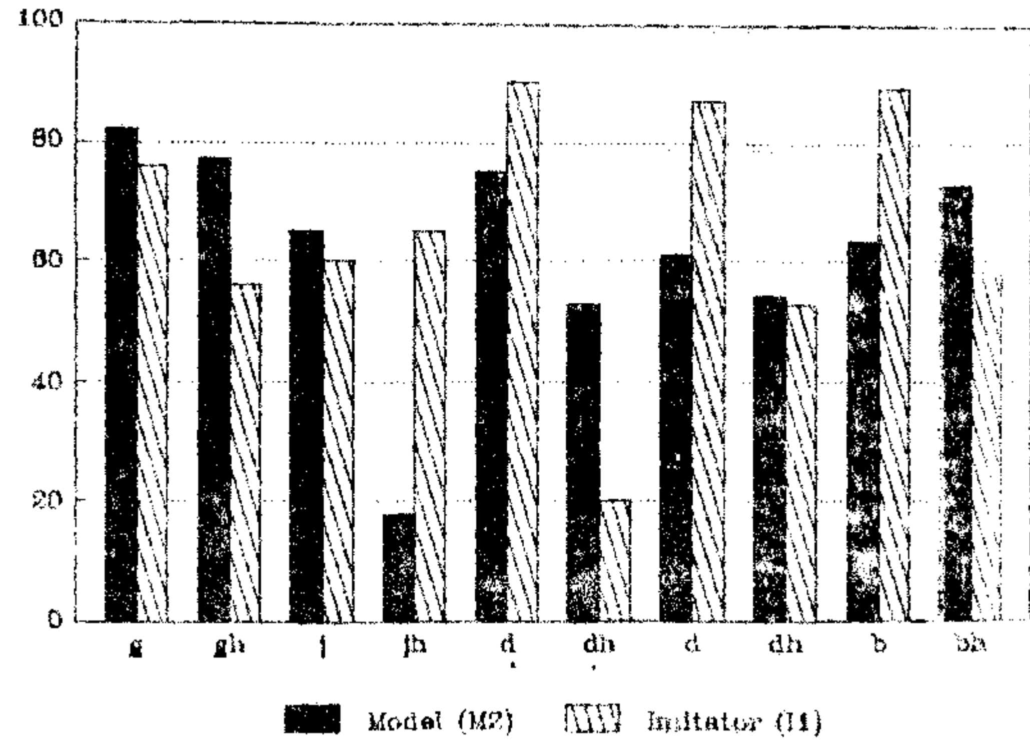
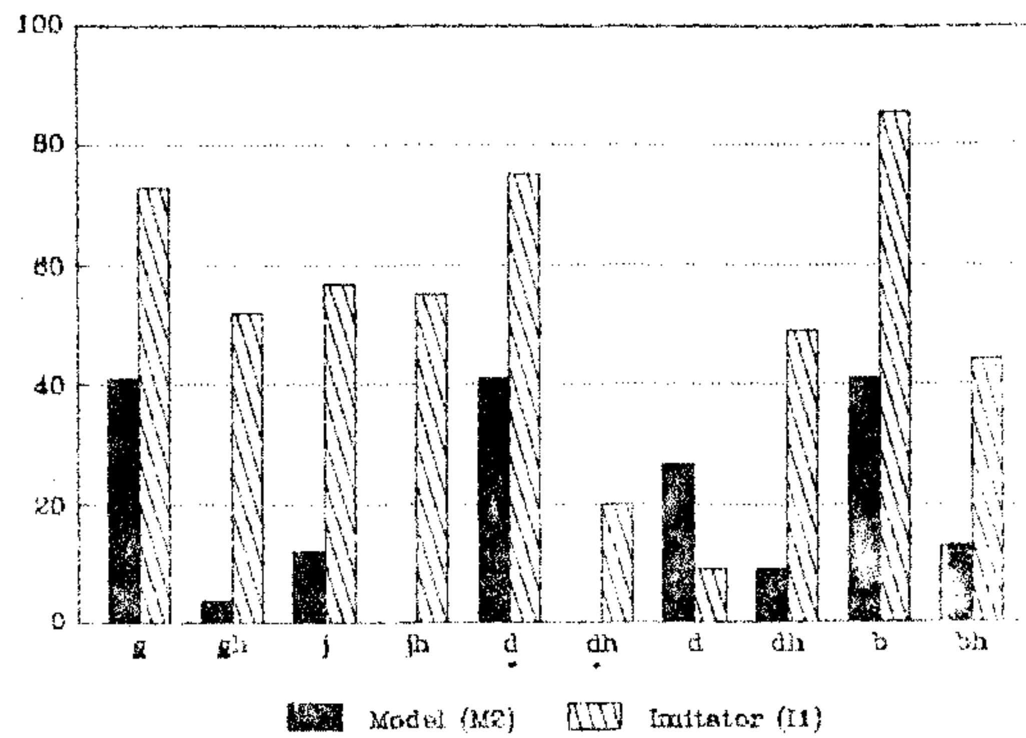


Fig.53 Closure durations of M2 & I1 (M.sec)

Fig.54 Voicing duration of M2 & I1 (M.sec)



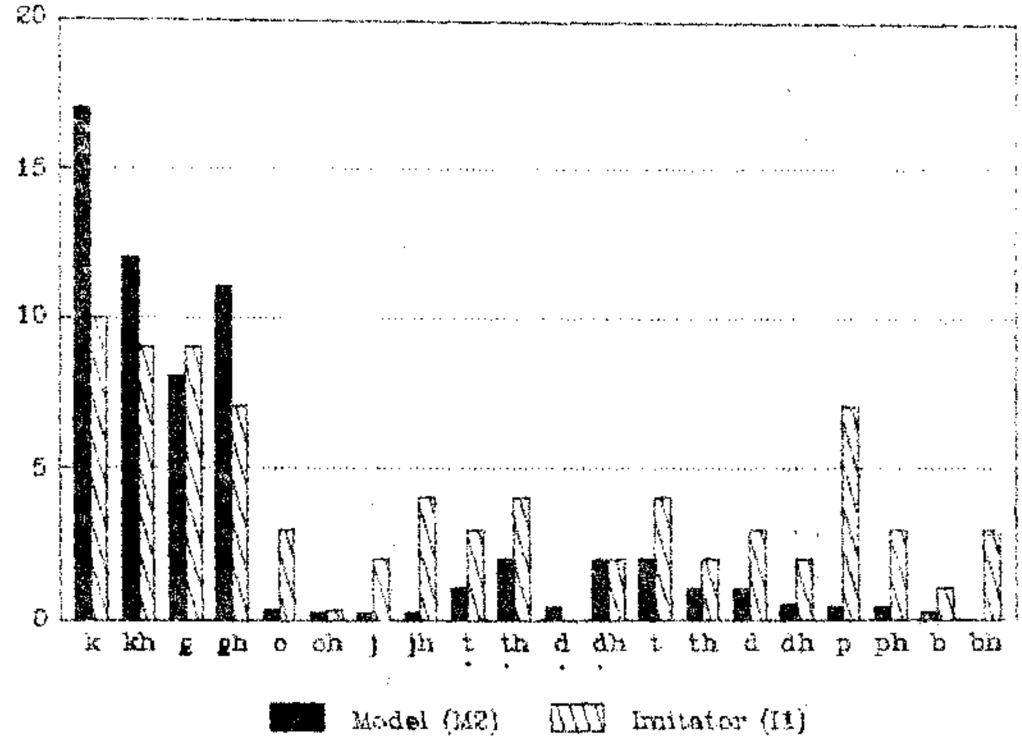
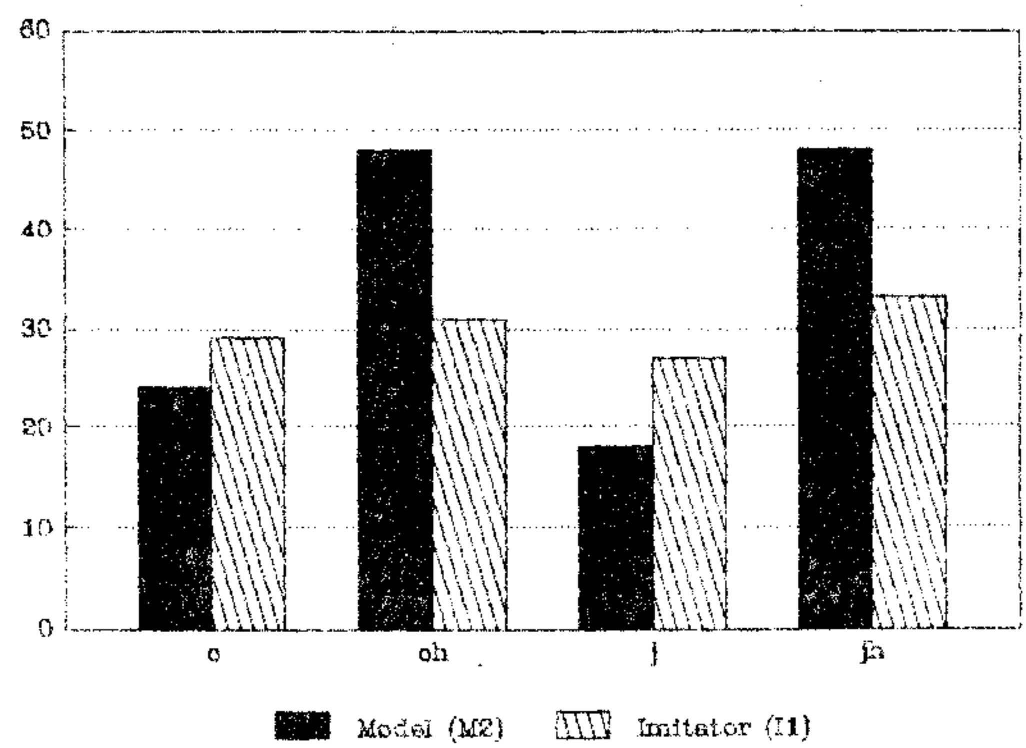


Fig.55 Burst duration of M2 & I1 (M.sec)

Fig.56 Affrication duration of M2 & I1 (M.sec)



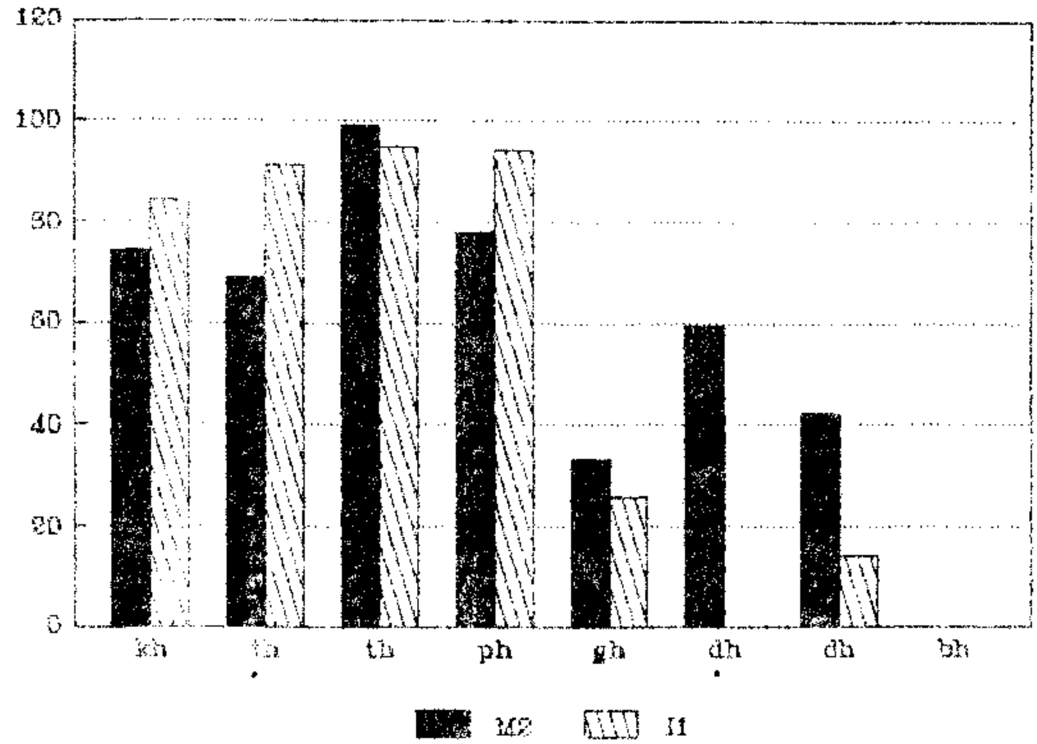
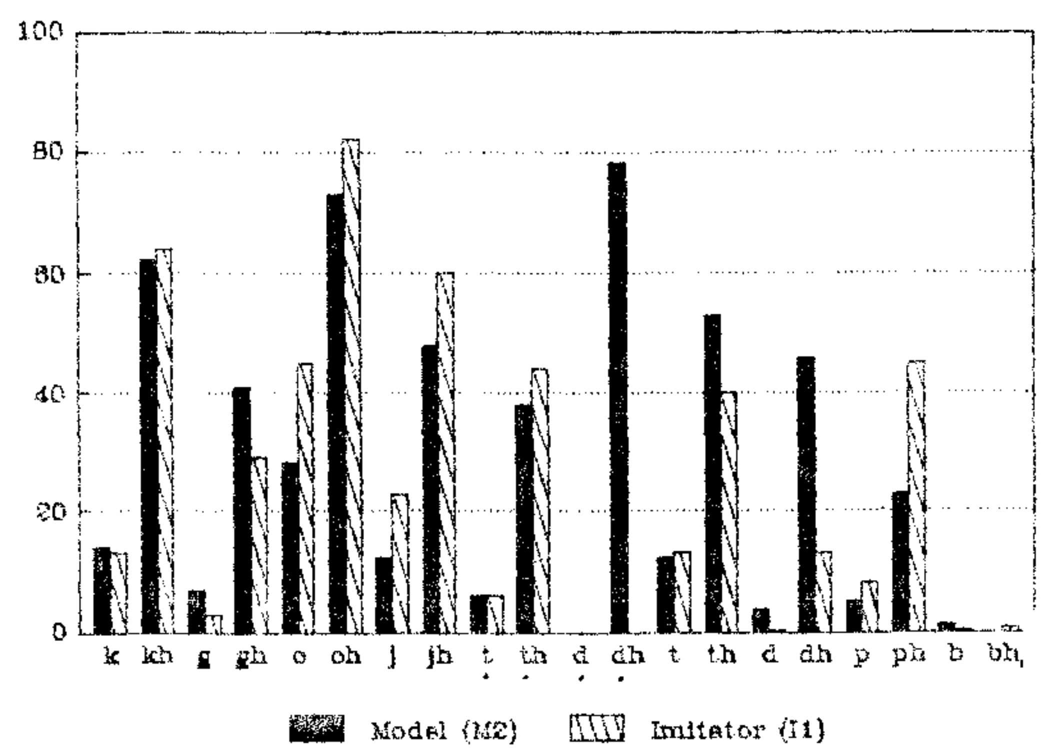


Fig.57 Aspiration & murmur duration of M2 & I1 (M.sec)

Fig.58 VOT of M2 & I1 (M.sec)



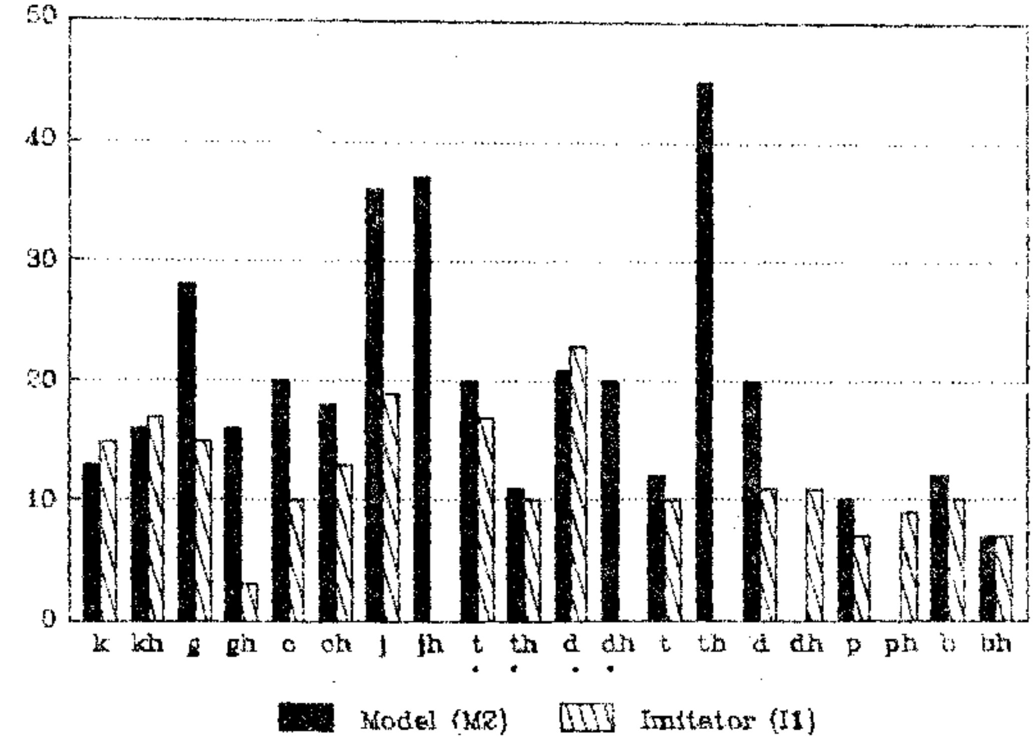
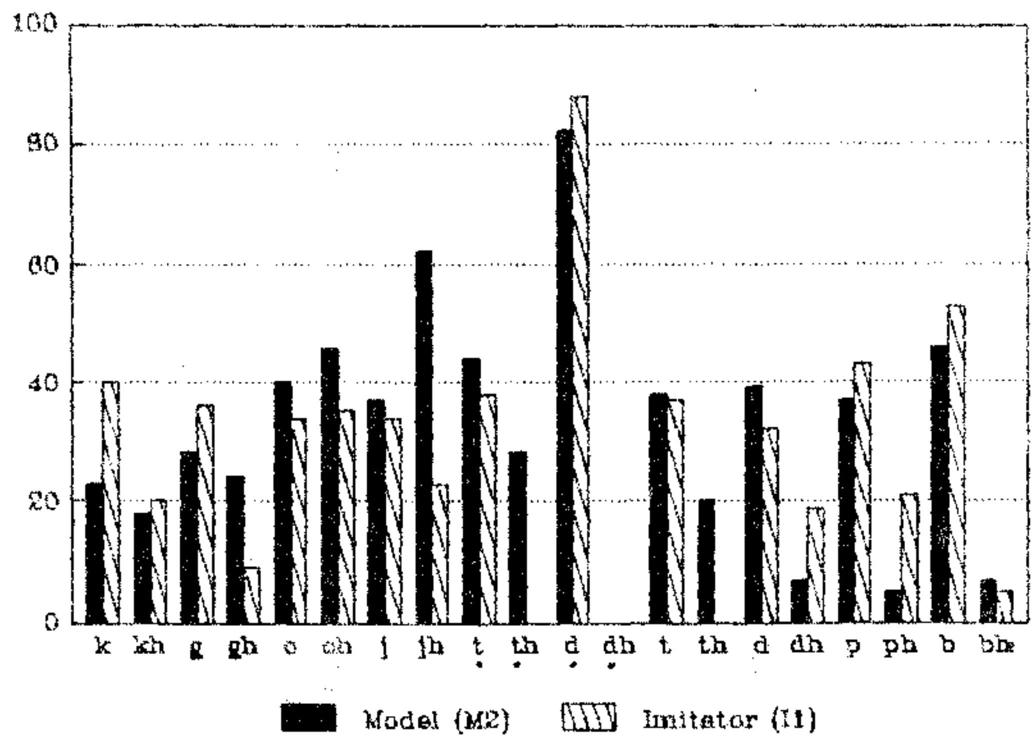


Fig.59 Transition duration of the first formant of M2 & I1 (M.sec)

Fig.60 Transition duration of the second formant of M2 & I1 (M.sec)



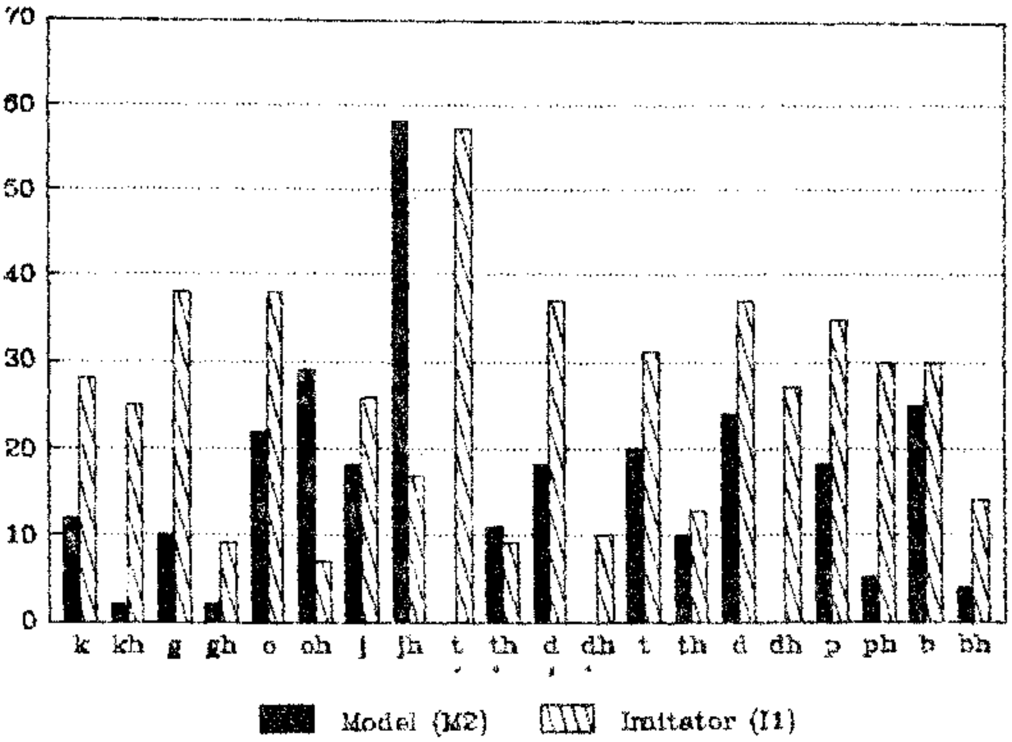
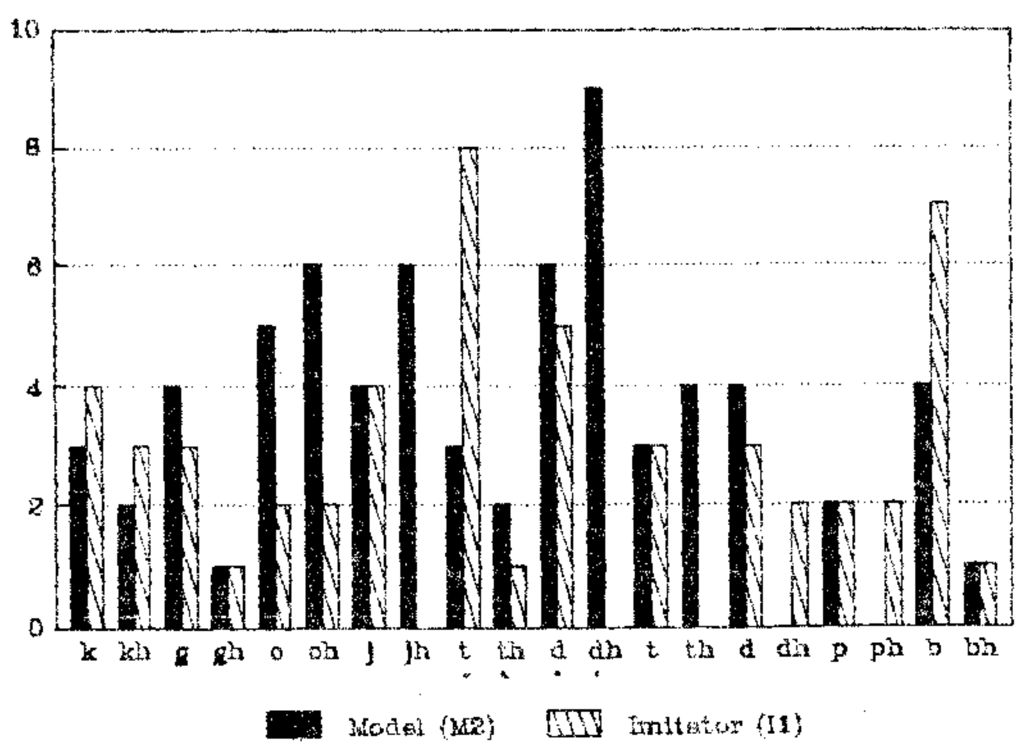


Fig.61 Transition duration of the third formant of M2 & I1 (M.sec)

Fig.62 Speed of transition of the first formant of M2 & I1 (Hz/M.sec)



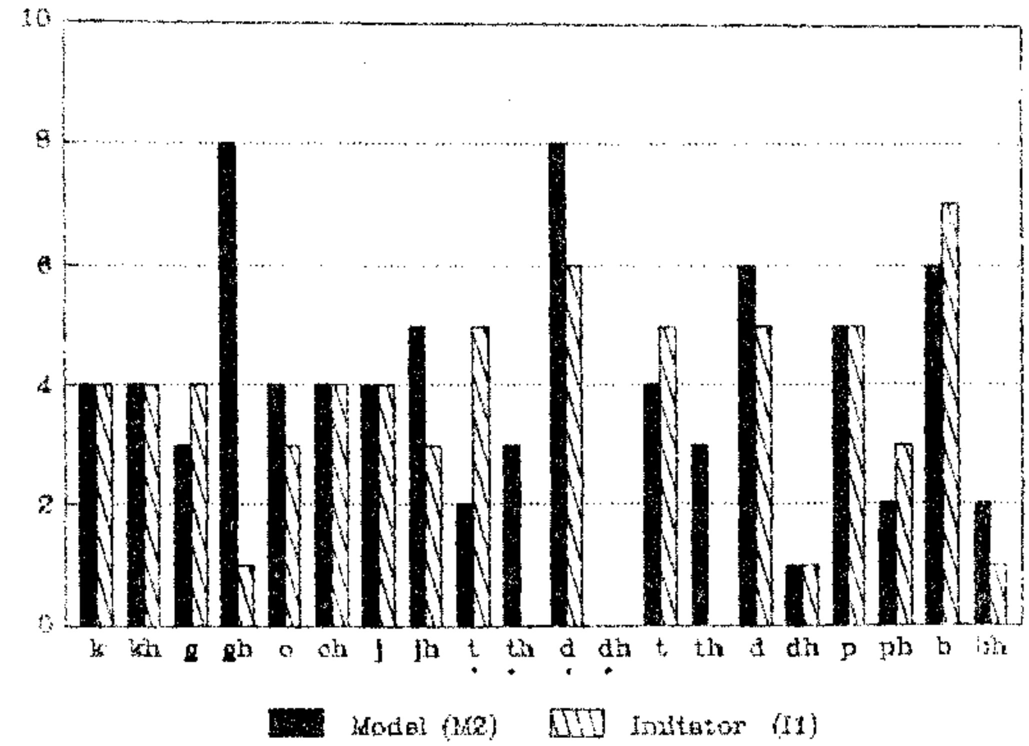
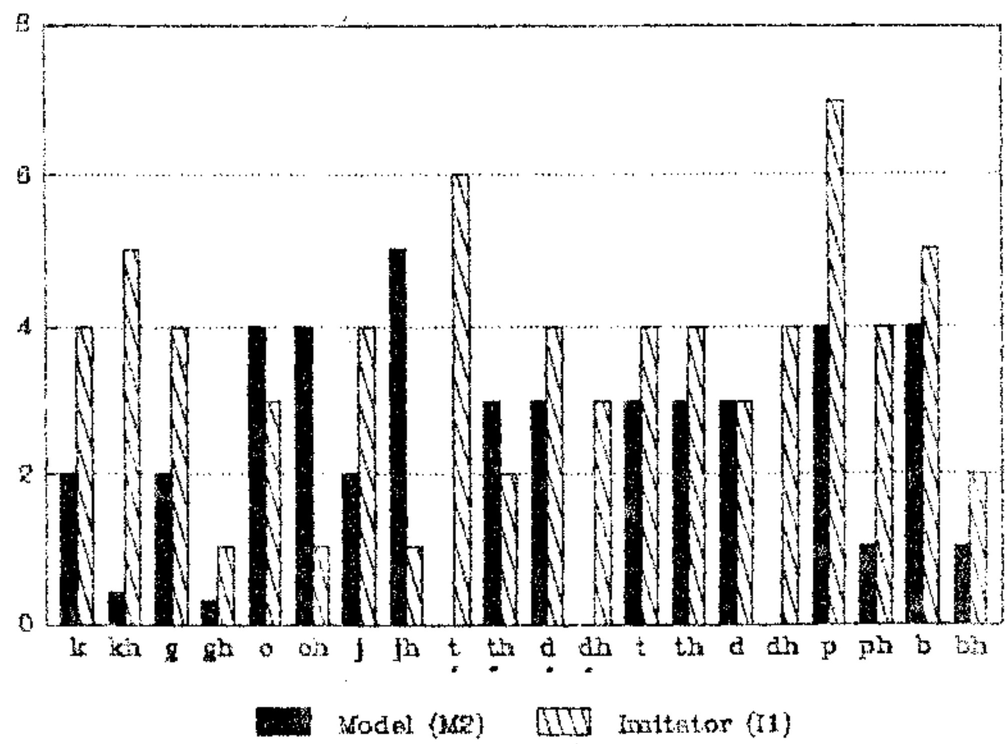


Fig.63 Speed of transition of the second formant of M2 & I1 (Hz/M.sec)

Fig.64 Speed of transition of the third formant of M2 & I1 (Hz/M.sec)



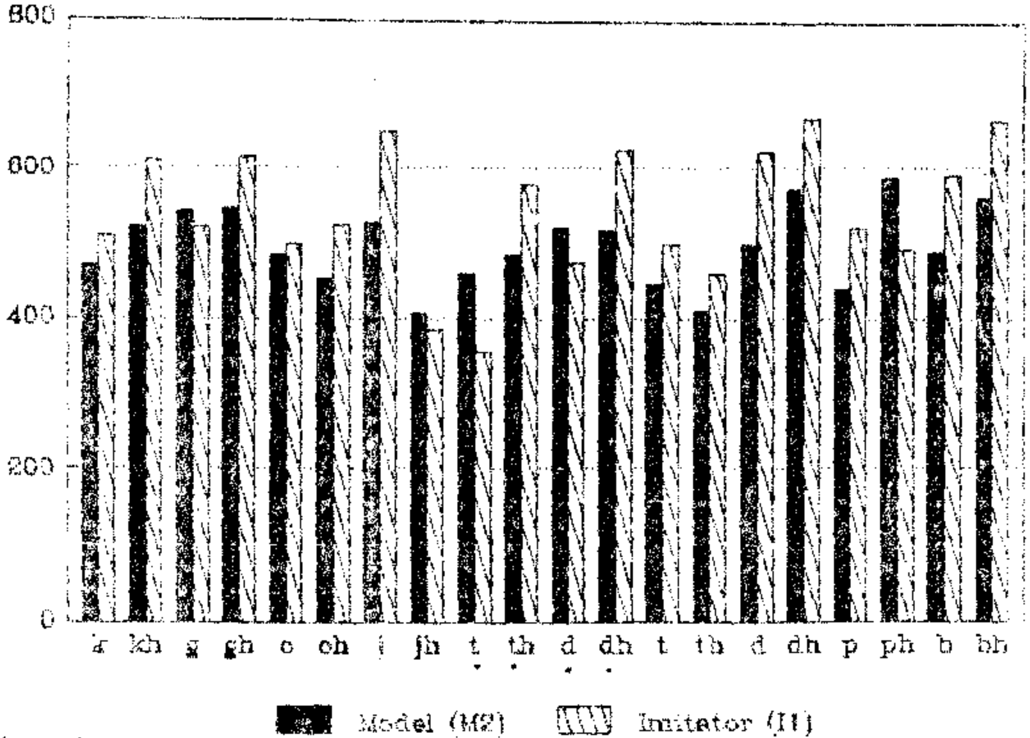
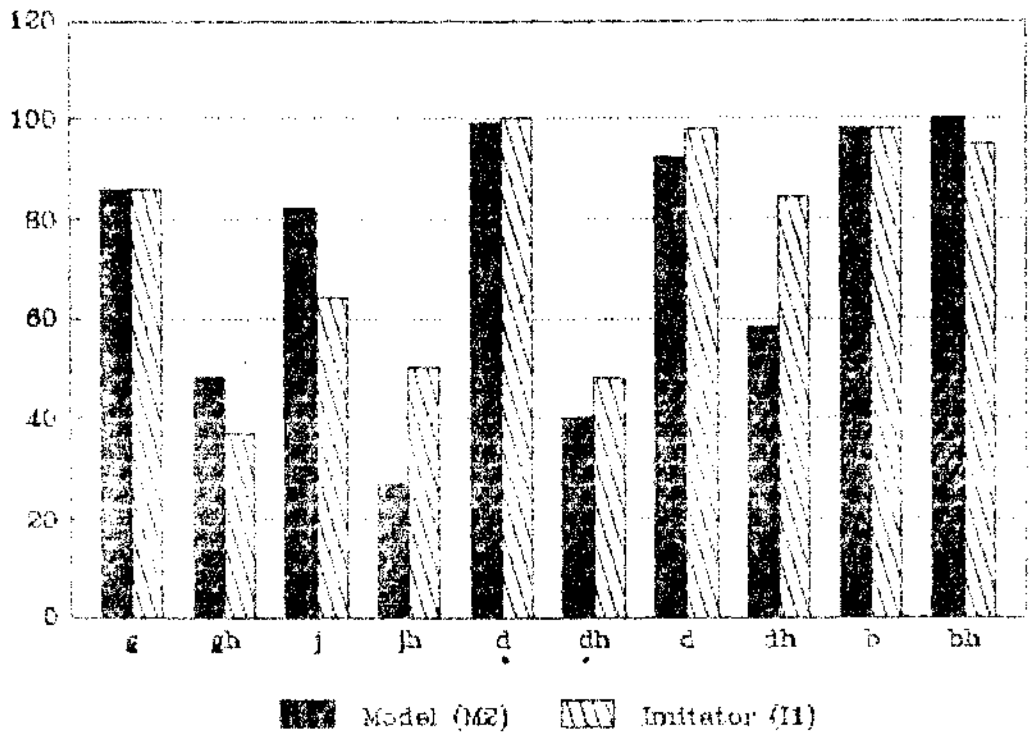


Fig.65 Word durations of M2 & I1 (M.sec)

Fig.66 Scaled closure duration of M2 & I1



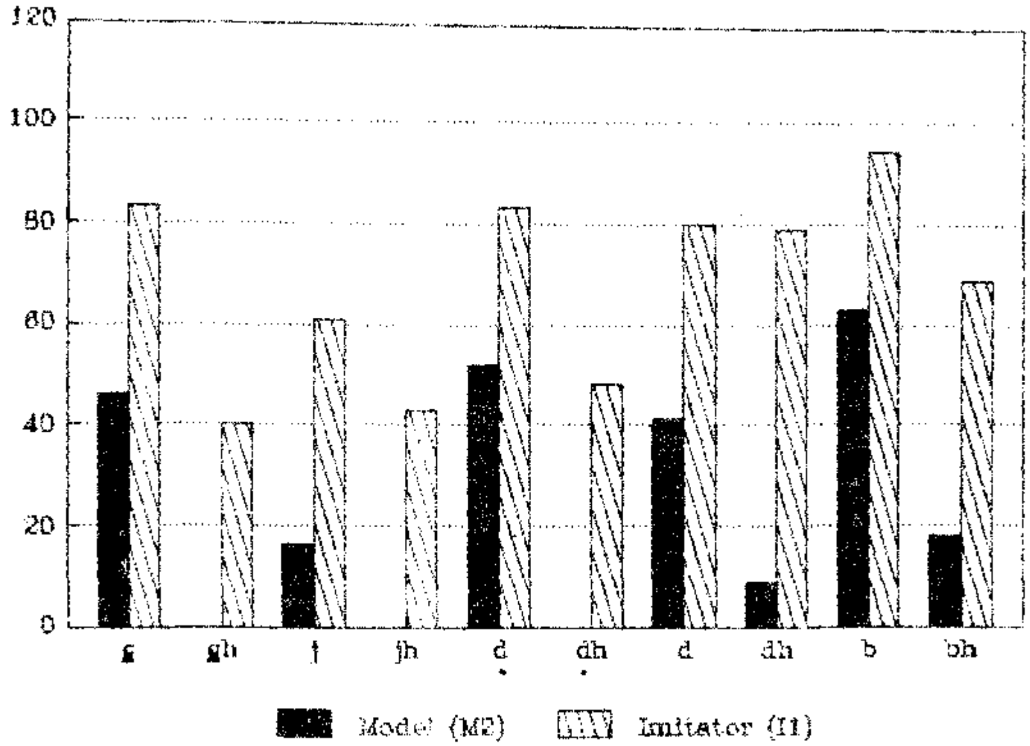
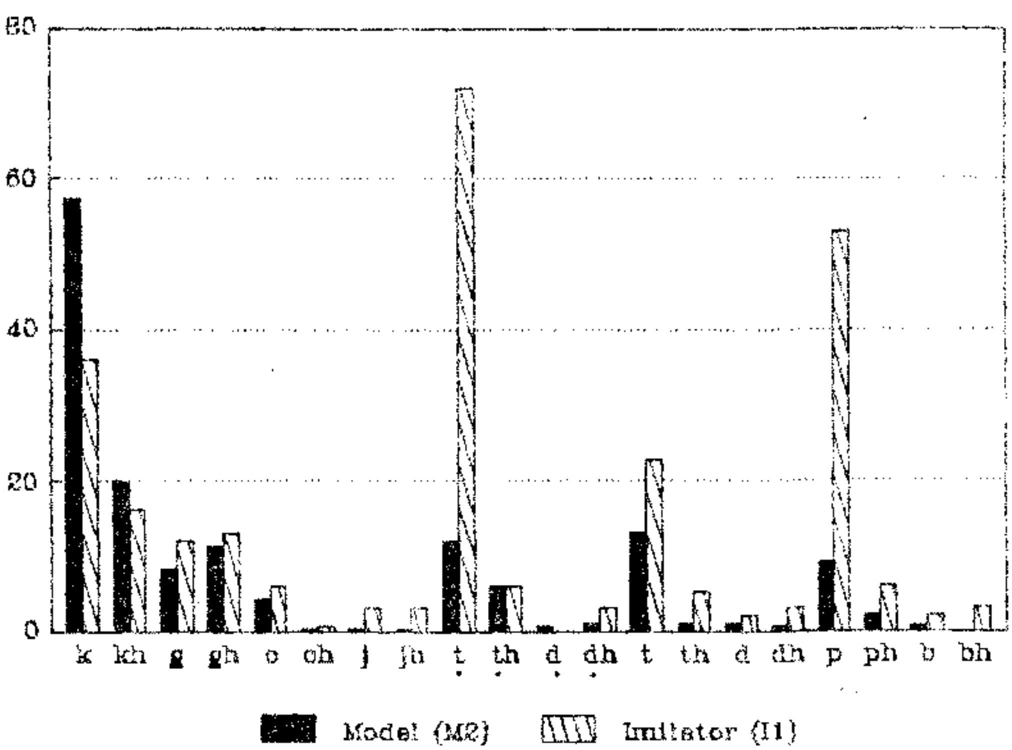


Fig.67 Scaled voicing duration of M2 & I1.

Fig.68 Scaled burst duration of M2 & I1.



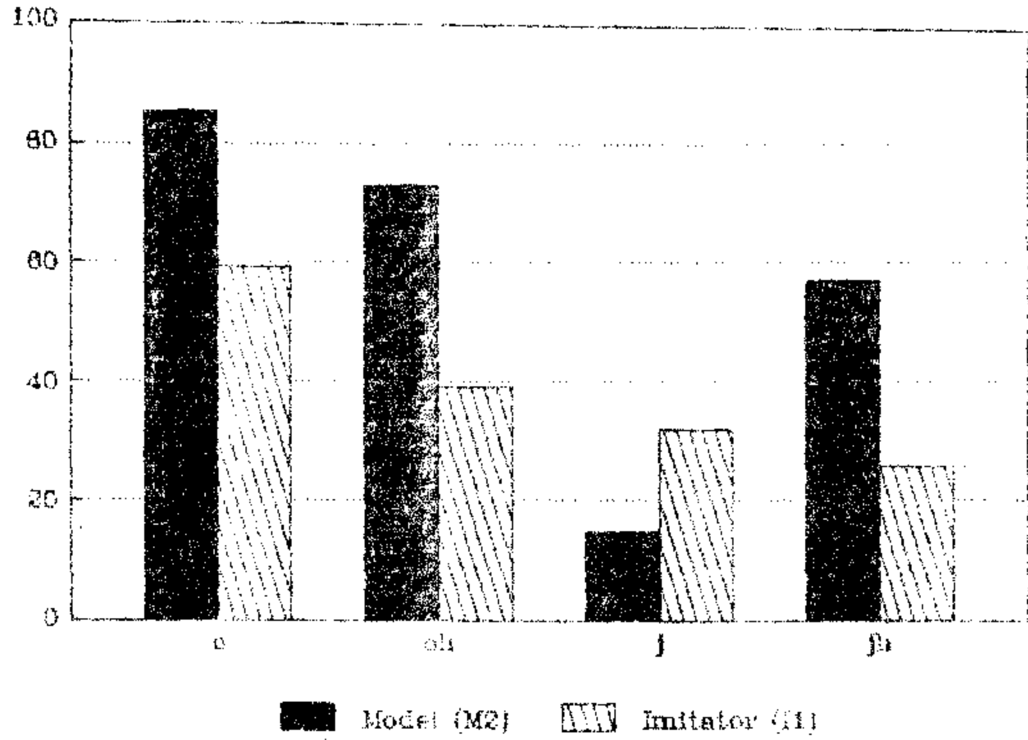
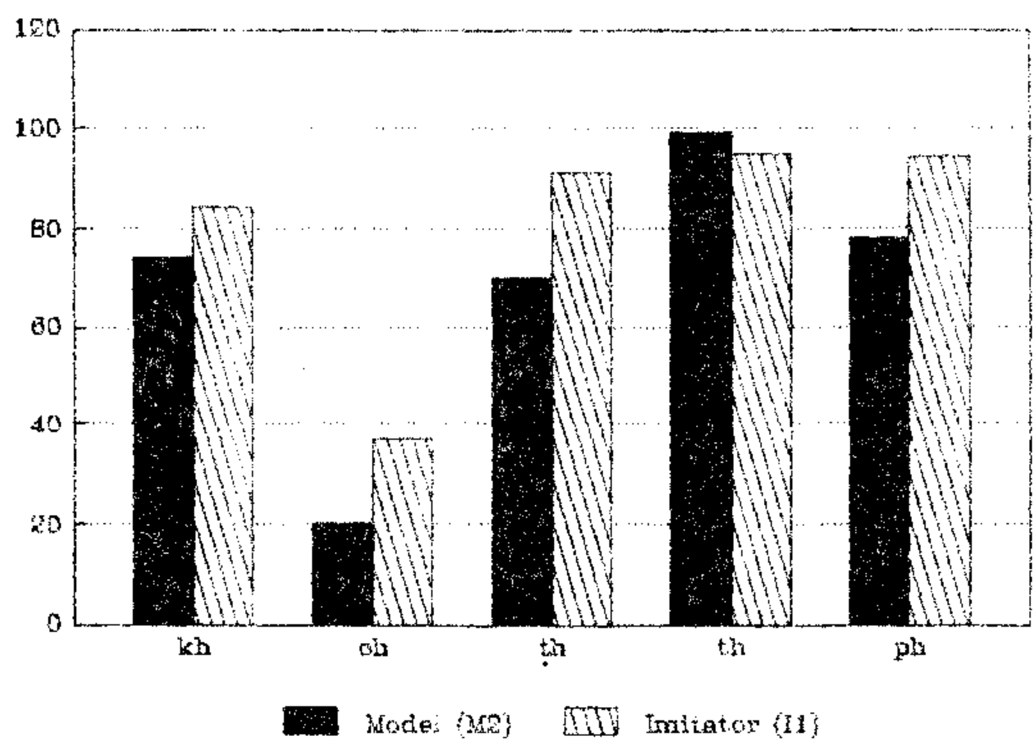


Fig.69 Scaled affrication duration of M2 & I1.

Fig.70 Scaled aspiration duration of M1 & I1.



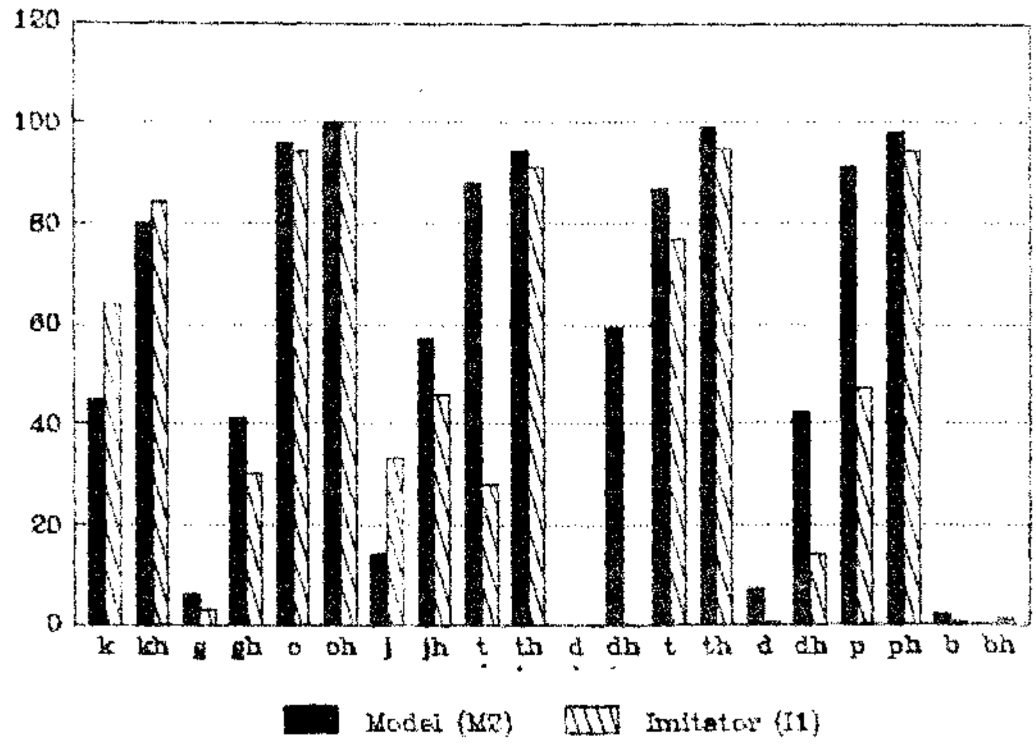


Fig.71 Scaled VOT of M2 & I1.

Of the spectral parameters the imitator used L1 & F3 for more than 75%. at times as the features for imitation. The for**man**t structure in the speech of M2 & I1 in general, ,and not seem to *agree,* F1, f2 & F3, of M2 &. *I*1 for various,stops and affricates are shown in Fig 72 and table 12 shows. the significance of difference between the spectra1 parameters s of M2 & I1. It could be observed that though F3. was used *more* often in the imitation paradigm, significant differences were present for F1, , of /kh/, /th/, /d/, /dh/, /b/ F2,of k, kh , gh, th, d, b, ph F3 of q, d, & b. (As the band widths and levels of the first three formants were very high and as the calculations were erroneous, they have not been considered. However, the average values have been provided) . Fig 73 to 76 depicts the amplitude values of M2 & I1.

In VCV condition the terminal F1 , F2 & F3, **the** preceding vowel was also measured. Fig 77 represents the terminal F1, F2,F3 of the precedinq vowel for both M2 & *I1.* It was observed that, especially for retroflexs, the? terminal F2 & F3. of the imitator *never* matched that of the model but however, was still perceived as a retroflex (on perceptual judgement). Wide variations in the terminal frequencies of M2 & I1 were noticed.

Phonemes	F1	F2	F3	B1	B2	B3	L1	L2	L3	BA	OA	AV	AC
k		+			+			+	+			+	
kh	+	+			+	+	+	+	+		+	+	
g			+						+			+	
gh		+							+				
c									+		+	+	
ch			+	+	+	+			+		+	+	+
j				+					+				
ṭ										+		+	
ṭh		+		+	+				+	+	+	+	
ḍ		+	+	+	+				+		+		+
t	+				+		+		+	+			
d	+			+				+	+				
dh	+				+		+	+	+	+	+		
p					+	+			+		+		
ph		+				+					+	+	+
b	+	+	+			+		+			+		+
bh					+				+	+	+		+

Table 12s Significant differences between the spectral parameters of M2 & I1.

1

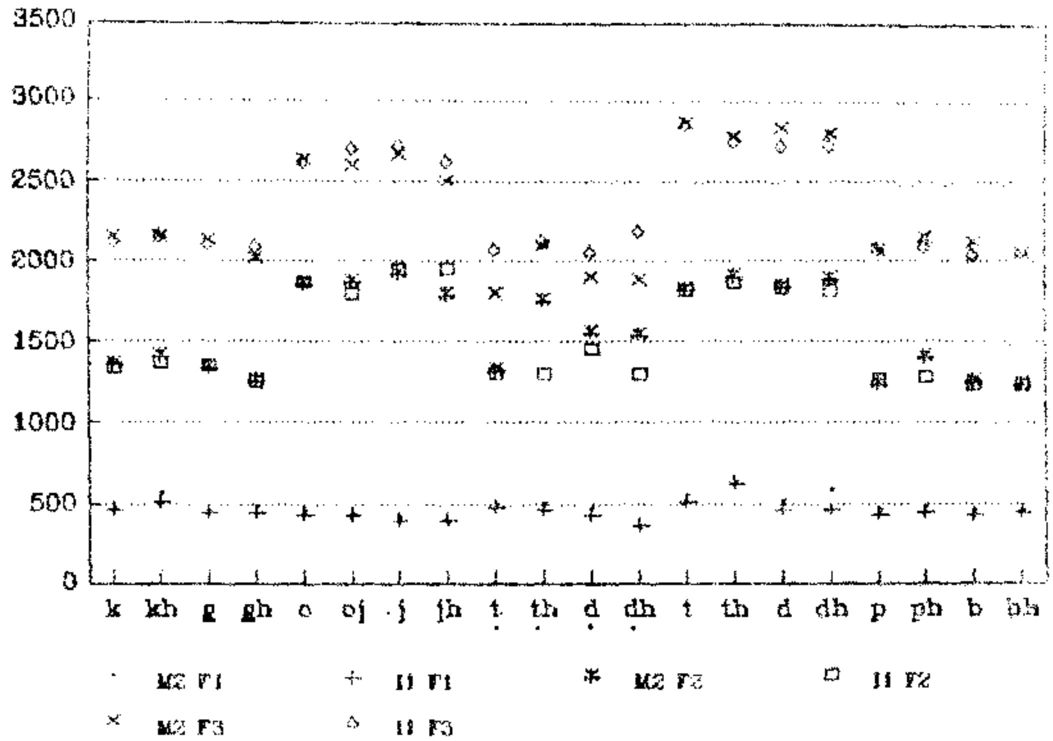
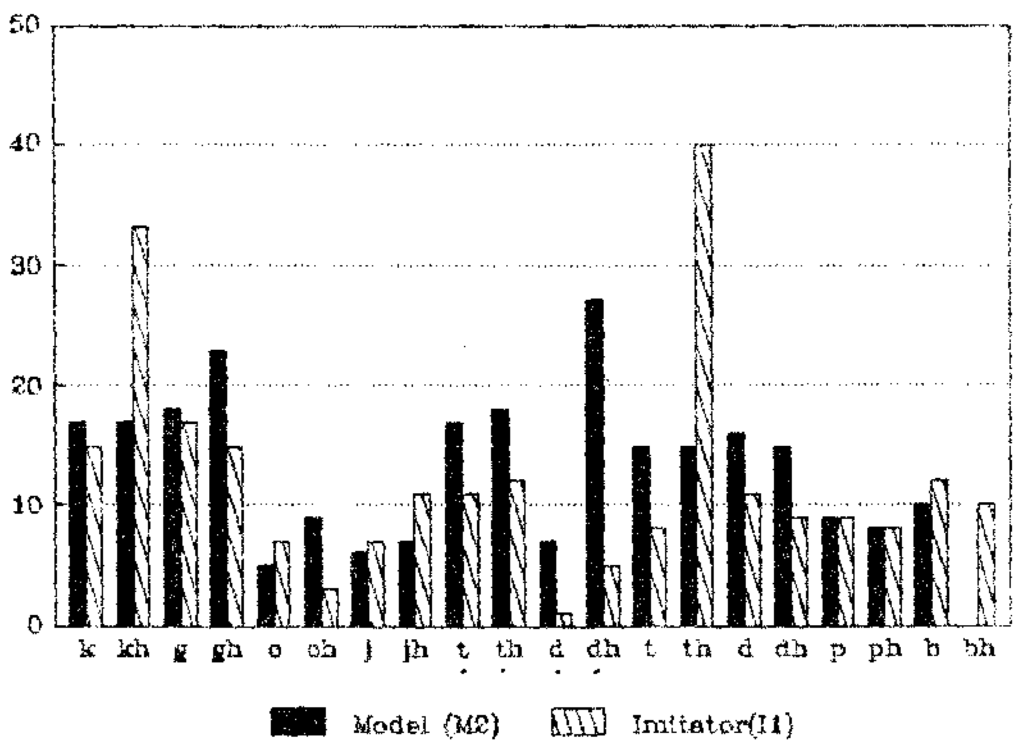


Fig.72 Formant frequencies of *P12* & 11 *(Hz)*

Fig,73 Burst amplitudes of M2 & 11 (RdS)



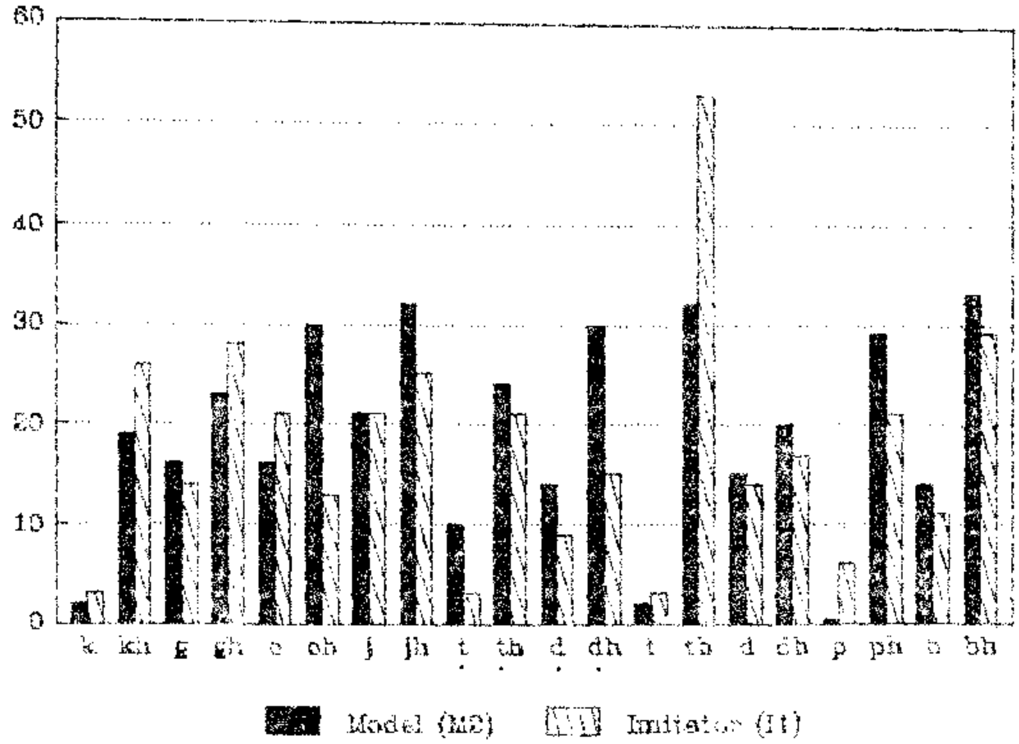
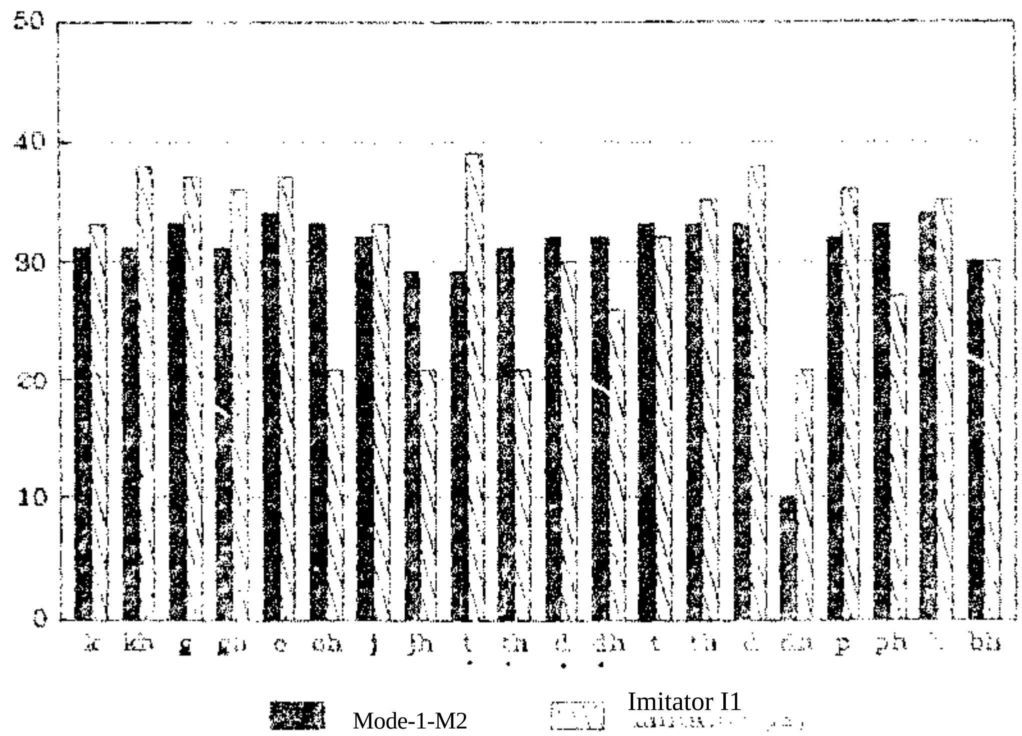


Fig.74 Overall amplitude of PI2 & I1 (RdB)

Fig.75 Amplitude of the following vowel of M2 & II (RdB)



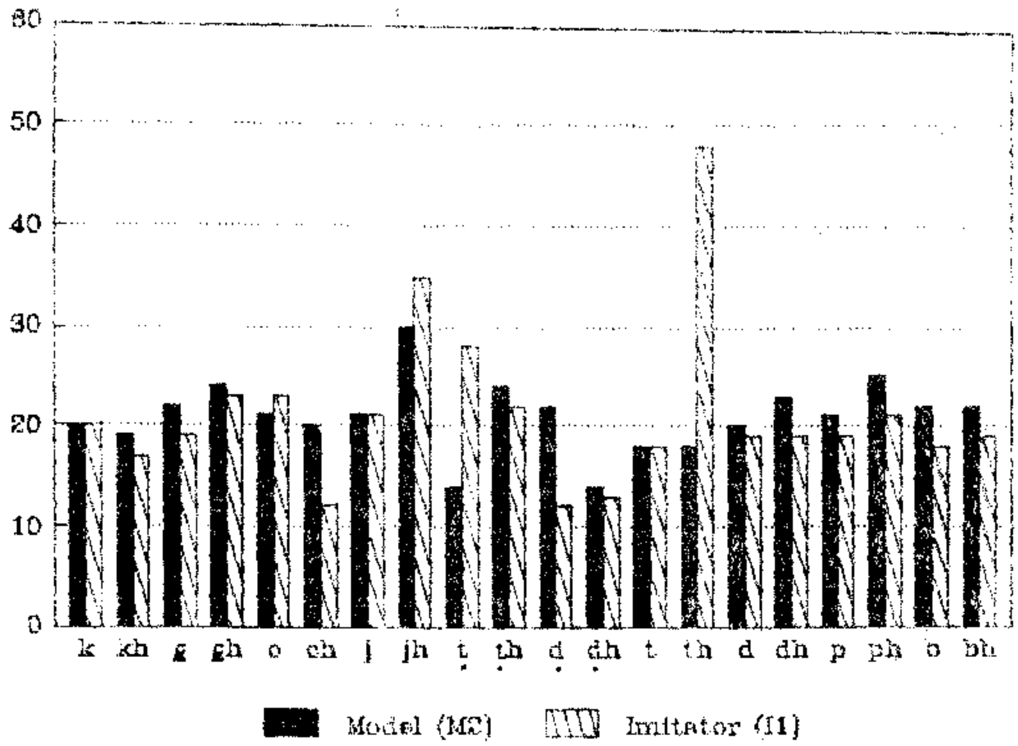
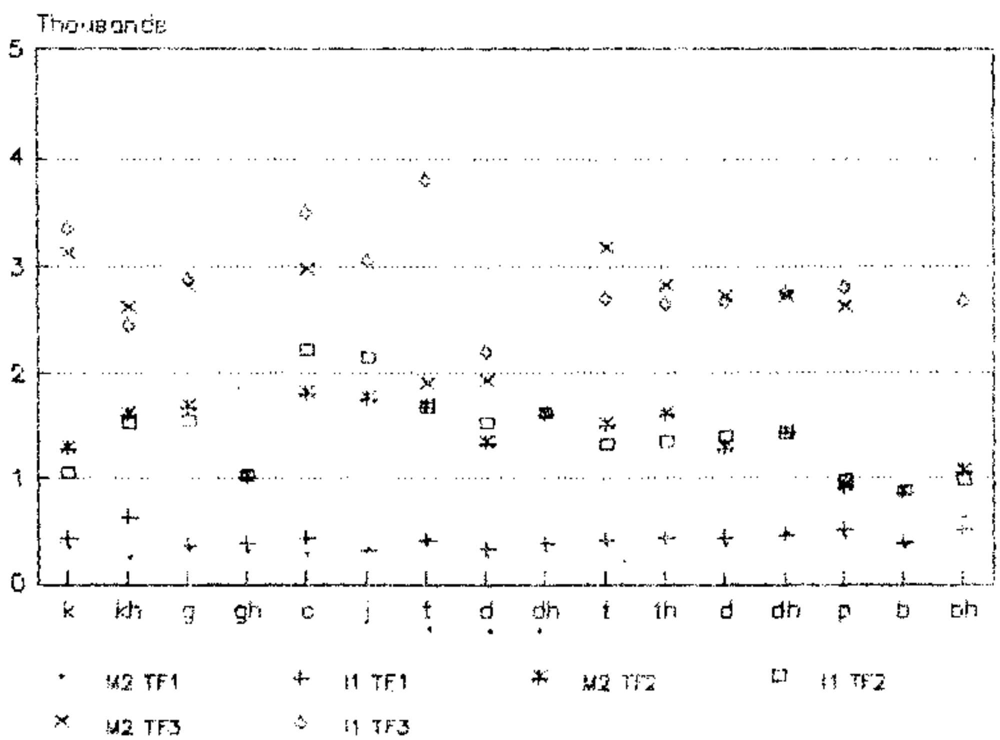


Fig.76 Amplitude of the following Consonant of M2 & I1 (Rdb)

Fig,77 Terminal farmant frequencies of (Preceding *Vawel)* M2 & I1 (KHz)



F0 movement. the following vowel was also computed. Vowels follownig voiced stops/affricates had 70%. rise pattern, 30%. fall pattern & 0%. flat pattern in the model and 0%. rise pattern, 80%, fall pattern and 207, flat pattern in the imitator. In Vowels following voiceless stops/affricates 80%. and 20% were rise patterns, 10 & 20%. were fall patterns and 10 & 60% were flat patterns in the model and the imitator respectively. The rise patterns in the vowels following

K voiceless stops/affricates were not imitated by I1. Fig 78 shows the percent of rise (R) fall (F) & Flat (F1) patterns of F0 of Vowel. following voiced arid voiceless stop/ affricates.

A grading of the temporal, scaled temporal and spectral parameters revealed that aspiration duration, scaled aspiration duration, STF1 , STF2, L1, F2, TDF1 & TDF2 were maintained by the imitator for more than 757. of the times. Table 13 shows the parameters imitated in percent.

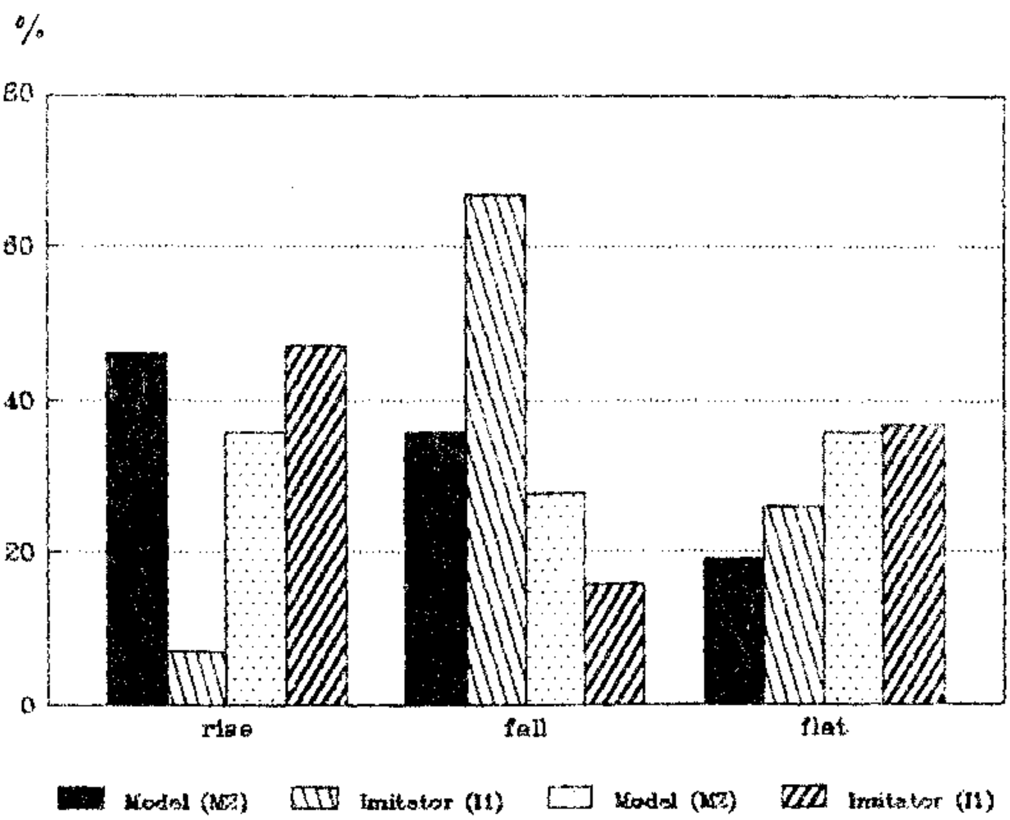


Fig.78 FO contours for Vowels following voiced & voiceless consonants of M2 & I1.

Parameter	% times imitated
Aspiration duration	100
Africatian duration	100
Scaled aspiration duration	100
STF1, STF2	92
L1	84
F2	78
TDF1, TDF2	77
L2, F2, Amplitude of the following consonant, Bl, B3, Burst, amplitude	72
VOT	61
F_3	60
Scaled closure duration	57
Amplitude of the following vowel	54
STF3, Murmer duration	50
B2	48
TDF3, Burst duration, Scaled Burst duration	43
overall amp1itude	42
Scaled VOT	30
Closure duration	29
Scaled voicing duration	14
L3	18
Fo movement for voiced stops/affricates	10
Voicing duration	0
word duration	0
Fo movement for voiceless® stops/affricates	

Table 13: Parameters imitated percent times by 11 (of M2).

Perceptual analysis by kannada speaker judging imitation revealed that, of the 551 imitations, 47% were judged 'very good', 45% as fair, 8% as poor and 0% as bad. The perceptual -features used to judge the imitation as 'very good' were similarity in pitch, syllable duration, inflection and rate of utterance. Whenever, the imitation changed in any one of these parameters it was judged 'fair'. In addition, any nasalization or substitution was judged to be poorimitation.

SPEAKER SPECIFIC FEATURES

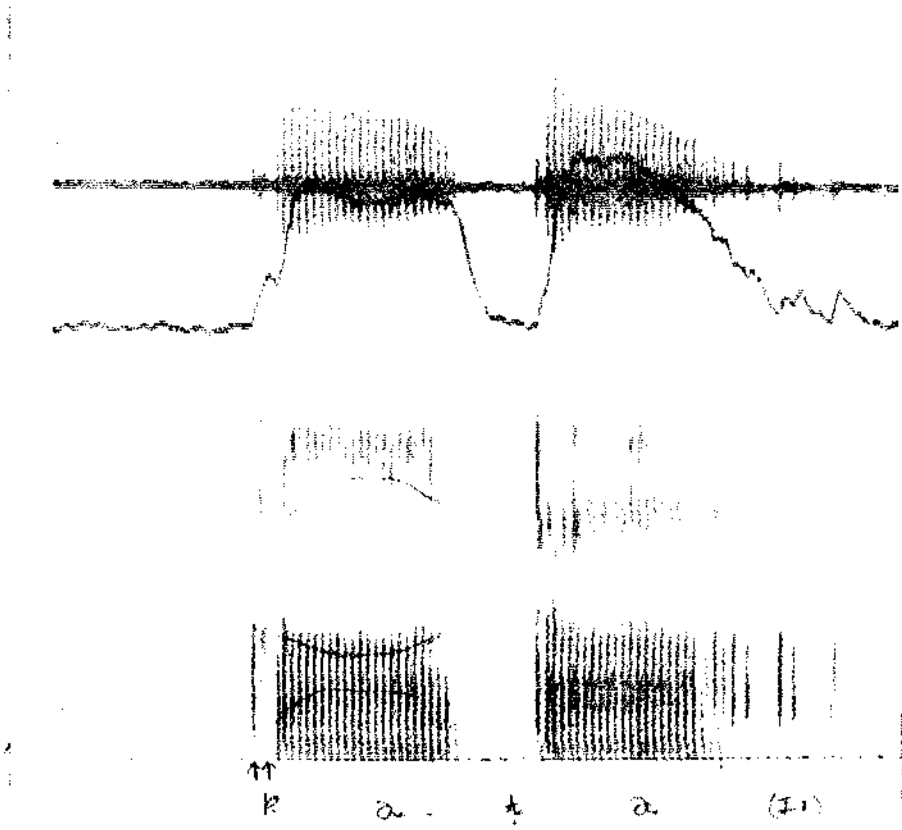
Some speaker specific characteristics were also observed. In speaker IS1 imitating Mi, multiple bursts extending upto 5 KHz was noticed. Formant structure® were observed during the voiced period. Whenever aspirated stops were the key phonemes to be uttered, they were unaspi rated, with the extension at aspiration to the following consonant if it was a plosive. No cessation of voicing or fading of F3 was observed in II.

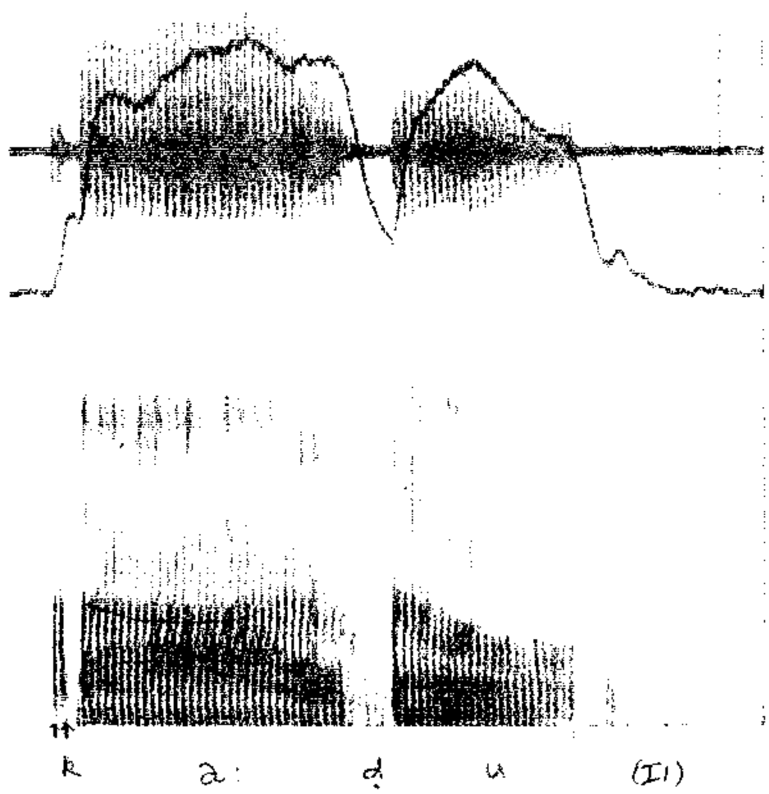
Model (M2) exhibited very clear bursts and few instances of multiple bursts in velars and bilabials were noticed. Third formant was absent often, especially following /u/. Words ended often with glottal stops, and at the end of the words, only voicing bars were observed with ceased formants.

Often /g/ and /b/ exhibited formants which indicated lax articulation. No cessation of voicing in the closure period was observed. The imitator (IS2), in contrast, showed glottal stops at the beginning of the word. Velars had multiple bursts and sometimes were aspirated. Word endings had murmer instead of voicing. The bursts of the model were extended upto 5 KHz. However, bursts of IS2 were restricted to 3.5 KHz. F3 was not visible when associated with high back vowels. The average amplitude curves showed multiple peaks during voicing of stops indicating different aerodynamic mechanisms. No cessation of voicing was observed.

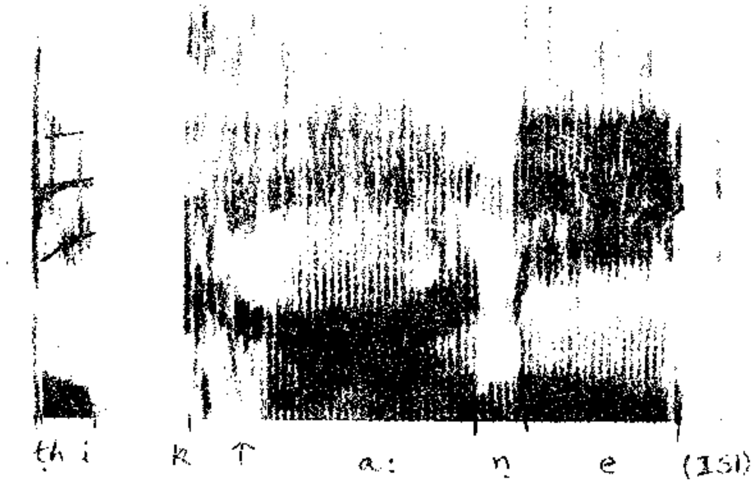
M2 exhibited glottal stops before the onset of the word. Multiple bursts (velars) were common and F3 was faint or absent especially when a labial constriction occured. Cessation of voicing during closure was another feature. In the spectrograms of the imitation, glottal stops were more frequent at the end of the words and multiple bursts (velars) mainly at the region of *F1 &* F2 appeared. The voiced stops often showed formant patterns as in Vowels/ Semivowels. Aspiration of unaspirates was also a feature and cessation of voicing during closure was noticed. A few Spectrograms showing speaker specific features follows.

 (a)

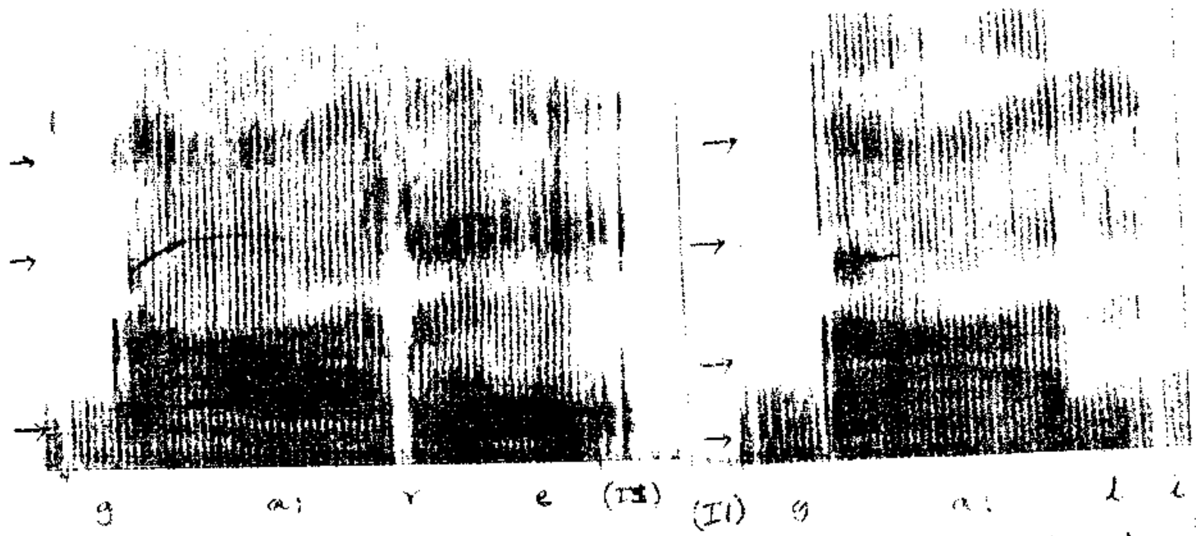




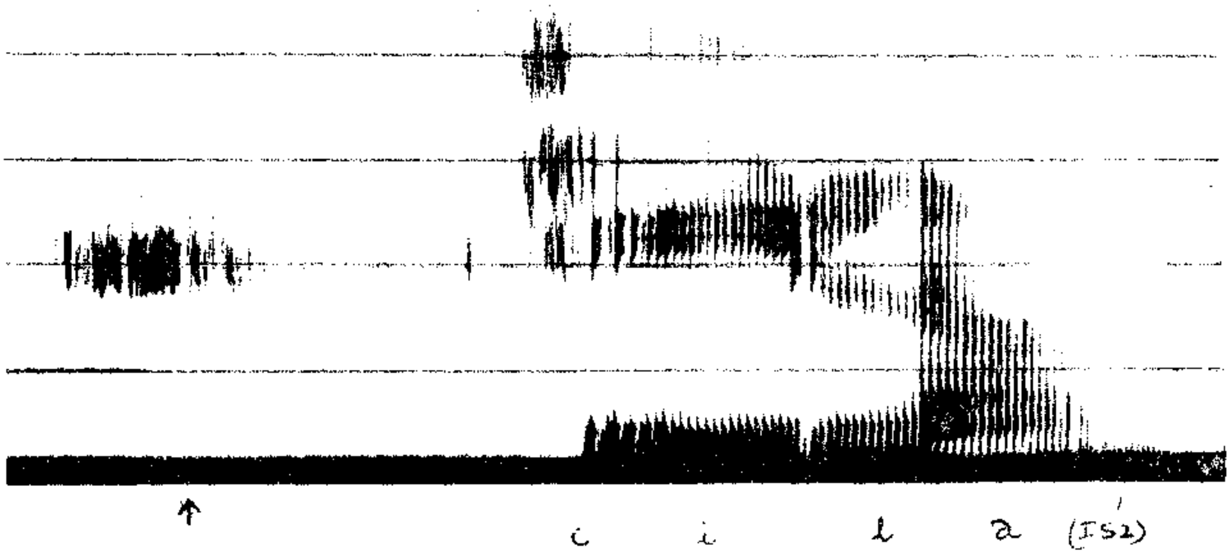
SPECTROGRAMS DEPICTING MULTIPLE BURSTS (MARKED ↑)



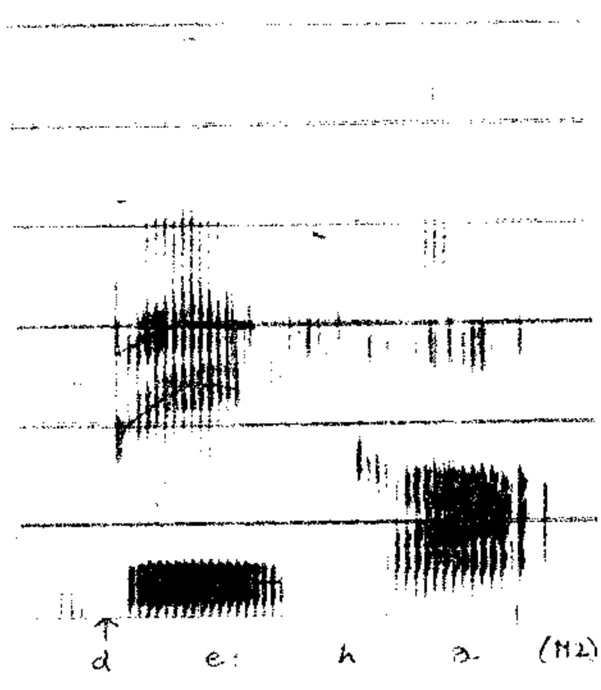
SPECTROGRAM DEPICTING SPREAD OF ASPIRATION TO THE FOLLOWING CONSONANT (MARKED ↑).



SPECTROGRAMS DEPICTING FORMANT STRUCTURE FOR VOICED-STOP (MARKED ↑)



SPECTROGRAM DEPICTING EXTRA EFFORT IN IMITATION (MARKED ↑)



SPECTROGRAM DEPICTING CESSATION OF VOICING (MARKED ↑).

FEATURES IN PRODUCTION

The production data deserves some comments. The formant patterns of the five subjects exhibited similar structure (except retroflex of IS2) as cited in the literature. Palatal had high F2, dentals had high F3, bilabials had low F1 & F2 and retroflexs had a pattern where F2 & F3 almost met. The transitions into the following vowels from different places of articulation were as in Fig 79.

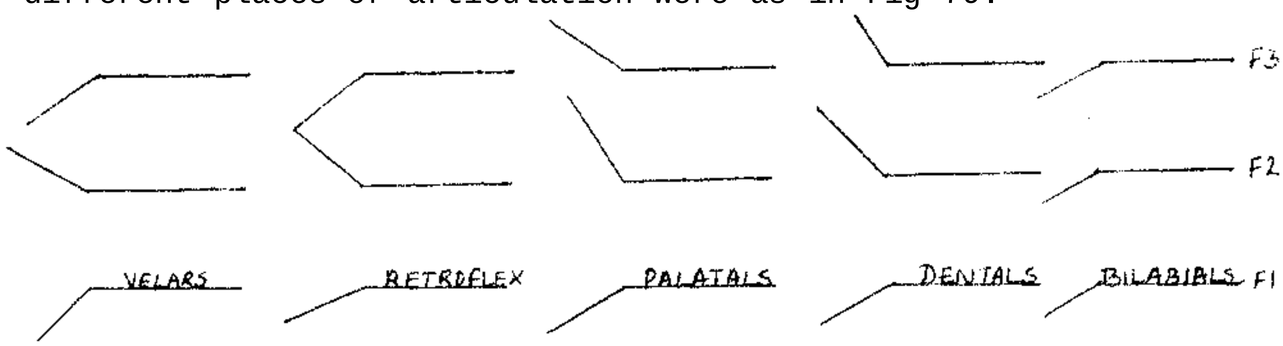


Fig 79 s Transitions of stops/affricates with different place of articulation to the vowel /a/.

In general, F2 & F3 seem to raise when followed by front high vowels & bilabilals. Of all the places of articulation, velars exhibited higher burst amplitudes. Aspirates showed higher burst amplitudes and the overall amplitude in voiced stops was greater than the voiceless. Dentals and bilabilals exhibited longer closure and voiced durations and retroflexs had shortest, closure durations. Velars exhibited longest burst & aspiration duration and VOT and tailabials had longest. murmur duration. Transition durations of the first and the second formants were longer in voiced than the voiceless.

Speed of transition of the first two formants were longer in voiced than the voiceless. The temporal parameters of the consonant (CVC) seemed to lengthen when the following consonants (CVC) had retroflex and bilabial places of articulation and nasal, lateral and tap manner. Some of these are in consonance with earlier studies, the details of which are not within the perview of this report.

DISCUSSION

From the results of this study several points can be made. First of all, among the temporal parameters, all the three imitators used transition durations and i-oeed of transition of the first three formants, more so of the first two formants, as cues for imitation. Word duations, durations pertaining to several gestures of the stops/affricates were the cues which were imitated 1east Table 14 shows the features used by all the three imitators.

	Ml vs IS1	MI vs IS2	M2 vs I1
Features imitated Maximum No. of times	TDF1 STF1 STF2	TDF1 TDF2 TDF3 STF1 STF2 STF3	TDF1 TDF2 STF1 STF2 Aspiration duration
Features imitated minimum no of times	Durations of word, voicing closure Affrication Aspiration VOT TDF3	Durations of word, V0T Affrication Aspiration	Durations of word, voicing closure burst TDF3

Table 14: Temporal features used in imitation by all the imitators

Several studies have been conducted in the past on the importance of transition duration and speed of transition in cueing for place and voicing . Of these some claim that they are cues and some deny it. Summerfield and Haggard (1972) have noted the presence of an Fl transition after voicing onset as a cue for voicing. Stevens and Klatt (1974) found that the absence of rapid spectrum change at the onset of voicing indicated a voiceless consonant. Both these studies emphasize the importance of Ft transition as a cue to manner. However, Keating and Blumstein (1978) explored the effect of lengthened transitions of F2 and F3 (45, 95 and 145 msec) on the perception of /da-ga/ continue and the results suggested that the lengthened transitions minimally affect the perception of place of articulation of stops.

Walsh and Parker (1987) comment that it is meaningless to speak of the absolute cueing value of either the rate of Fl decline or steady-state duration, since steady-state duration is dependent on the value of the other cue. Walsh and Parker (1987) using Fl dec line during the offset transition in /bat/ and /bad/ forming different rates of Fl decline (0, 3.7, 6.7, 9.7 Hz/msec) found that rapid Fl declines were perceived as /d/ (voiced). Porter et al (1987) noticed that listeners use perceived rate change than perceived duration to discriminate speech like sounds.

These call for comment on the categorical percepticn of stop consonants. It is commonly believed that consonants are perceived mores categorically than are vowels (Libersan etal,1957; Fry *et* al 1962; Stevens et al 1969). "he hypothesis originally offered to explain this dichotomy in consonant and vowel perception was that there are two separate mechanisms used to process speech, one which uses acoustic attributes of the stimulus at all the stages and one specially suited to converting acoustic information from eg. a stop consonant into a phonetic label. The specialised speech processor was thought to literally strip sway the acoustic: information from a stimulus in making the identification, so that such information was no longer available (Liberirian, Hattinqly and Turvcy, 1972). An alternative? hypothesis offered tav Fujisaki and Kawashima (1969) is that the difference attributed to consonant and vowel perception are due to intrinsic; nature of the stimuli and not to the mechanisms which presumably process these stimuli. Fujisaki and Kawashima hypothesised that auditory short-term memory does not retain transient cues such as consonantal formant transitions, and that therefore identification is stored in phonetic memory. EtecauBt? the low level acoustic information extracted from stop consonants deteriorates rapidly, it is generally not available for discrimination judgements. These judgements must not be based on a comparison of the stored phonemic labels. In contrast, since vowel stimuli *are* not transient they remain accessible in auditory short term memory. Discrimination judgements can be based on direct comparisons of the stimuli (Keating and Blumstein, 1978). The imitation data of the present study does not seam to support this hypothesis as the transition durations and the speed of transitions, which are considered inaccessible by Fujisaki and Kawashima, top the features perceived and hence imitated.

Liberman etal (1956) experimenting an synthetic speech sounds found that short transition durations signalled stop manner and long transition durations signalled Semivowel. Daniloff (1980) is of the opinion that as the distance between coarticulated targets shrinks, not only is the distance to be moved minimized but lesser force and a lower velocity of movement is needed to traverse the lesser distance. Accordingly, it could he interpreted that the srticuiatar while transiting from Vowel to Semivowel or vice-versa will have maximum speed i.e. the speed of transition depends on the distance to be moved and thus the speeds of transitions for different manners of speech sounds has to be different and specific. Considering these, it could be argued that all the three imitators either have perceived cues for stop and voicing manner only or have used appropriate transition durations and speed of transitions for stops in which case STF1, STF2, STF3 and TDF1 , TDF2 and TDF3 cannot really be considered as the major cues.

TDF1 IS2	TDFl I1	TDF2 IS2	TDF2 I1
20	15	26	40
14	17	12	20
21	15	33	36
18	3	21	9
10	10	17	34
8	13	28	35
21	19	27	34
32	0	50	23
16	17	16	38
17	10	17	0
18	23	48	88
0	0	0	0
9	11	33	37
0	0	0	0
11	11	25	32
0	11	0	19
13	7	32	43
28	9	15	21
13	10	36	52
16	7	9	5

TABLE 15: Transition durations *of* imitator imitating two models.

However, the data (table 15) on these measures of one imitator imitating two models indicate individual variations in transition durations and speeds of transitions. Thus, it could also he argued that transition durations and speeds of transitions are efficient cues in CVC conditions. This calls for comment on coarticulation, which is reportedly defective in speech handicapped children. More emphasis on transition durations and speeds of transitions could therefore, enhance the target productions in these children. Also, another point to be noted is that the absolute durations of several gestures involved in the production of stop consonants are not important cues for a stop. A scaling of these parameters may be of significance rather than the absolute values.

Among the scaled temporal parameters, aspiration duration was a parameter commonly used by all the three imitator and VOT and burst durations are parameters least used (Table 16). It is well known that neither two speakers can have the same temporal features nor a single speaker can maintain exactly same temporal features for a phoneme uttered twice. This and the fact that, the imitator have used some temporal parameters in a scaled manner suggest that the perceptual mechanism may depend on the relative durations of a speech sound rather than on absolute values. The perception by children of adult forms also supports this.

	M1 vs IS1	Ml vs IS2	M2 vs I2
Features imitated maximum no. of times	AD AFD	AD CD VD	AD
Features irnitated minimum no. of times	BD CD VD	VOT	VOT BD VD AFD

Table 16: Scaled temporal features used by all the three imitators.

The third point to be noted is that among the spectral parameters, burst amplitude? seems to be one of major parameters imitated (excluding the levels and bandwidths of the formants which were erroneous on LPC measurements). However, the formant structure in the speech of the imitators did not agree with that of the models. Earlier research on imitation by Kersta (1962b) who disclosed spectrograms of a speech made by President Kennedy and copied by his imitator Elliot Reed, revealed obvious differences in the spectrograms, despite the fact that, according to Kesta, the voices sounded extremely similar. Endress, Bombach and Flosser (197) studied the impersonations of five public figures by two well known German imitators and suggested that they were not able to adopt the parameters to match *or* even be similar to those of the imitated persons. Shafer (1974) in an experiment on adults repeating abnormal child forms found that the adult renditions failed to preserve all the acoustic cues present in the children's forms. Hall and Tosi (1975) experimenting on mimicked voices reported interspeaker differences. Nolan (1983)(pp 59) comments that "the dimensions of a person's vocal tract, or the length and mass of his vocal folds, will in some sense determine his formant frequencies and FO respectively and may even define optimum values for him in these parameters; but the plasticity of the vocal tract is such that his scope for variation in these parameters is considerable". In this context, it is worth noting that one of the imitator, a linguist and a speech pathologist, reported that he had to lengthen his vocal tract, open the velopharyngeal port a little, lower the larynx and had to have a firm contact of the articulator in imitating stops/affricates of the first model and had to puff his cheeks and had to have a firm contact of the articulator in imitating the second model. Also, while imitating the second model,it was extremely difficult for him tp modulate the FO movements in the CV syllables, which, according to him, were very fine. Also, he reported that he had to make a very smooth contact of the vocal folds which was not a characteristic in his natural speech. Inspite *of* the imitator changing his plastic vocal apparatus, he failed to bring the formants closer to that of the imitator. Thus, inspite of the imitation being judged as good, the characteristics of the vocal apparatus are reflected in the spectral measurements.

With respect to the absolute FOs, IS1 while imitating Ml imitated the FO of Ml. However, he did not imitate the FO of M2. Also, the second imitator did not imitate the FO of M2. The FO contours of the model (Ml) were imitated for vowels following voiceless stops/affricates. However, M2 showed rise- patterns for both voiceless and voiced which is in contrast with the earlier findings of Ohde (1984) that vowels fallowing voiceless stops have fall pattern (Table 17).

Vowels following	Ml vs Is1	H1 vs Is2	M2 VS I1
voiceless	rise—flat	fall-fall	fall-fall
voiced	rise-flat	rise-flat	rise-flat

Table 17: FO contours of the models and the imitators.

On perceptual evaluation, the cues (table 18) used to judge imitation as 'very good' were mostly suprasegmentai in nature. Articulatory precisions were also considered. Thus, it seems that the judges performed a voice identification rather than speech identification task. It should hr noted that inspite of the absence of similarities in the aroustic features imitated, the voices sounded extremely similar as evidenced by the percent (92, 69, 77 & 81, 61) judged as very good and fair. Also, inspite of the word durations being significantly different in the imitators, they were perceived as similar to the models. Apart from significant differences in word durations, the imitators used deep inspirations before uttering each sentence suggesting greater efforts, which would be a point of consideration in voice? print Forensic applications.

Ratings	Ml vs I SI	M1 vs IS2	M2 vs I1
Very good	Siff«iIaritvin stress articulation aspiration syllable du-ration pause	Simi1arity in strees articulation aspiration syl1able du-ration pause	Similarity is pitch syl1able du-ration inflection rate
Fair	deviation innasality	deviation in nasality	chanqa in any of the above
Poor	distartion	distartion	change in any of the above + nasalization + substation
Bad	deviation in more than one parametre	cleviation in more than one paramertef	

Table 18: Basis of perceptual judgements for the quality of imitation.

The fact that the judges participating in perceptual evaluation reported articulation as one of the cues for rating and the reported changes in the vocal tract made by the imitator reflects the capacity to perceive changes in articulatory gestures and supraseqmentals. Also, the fact that the imitator was aware of the changes to be made in his tract in order to sound similar to the model suggests that he perceived the articulatory gestures *in* the model's speech which supports the notion that speech perception is perception of gestures.

From the data, it appears that in deciding two speech samples as the same, it is not necessary that the listener extracts acoustic features pertaining to each speech sound. Suprasegmentale are sufficient in these decisions and this calls for a greater participation of the right hemisphere. Several functions have been attributed to both the hemispheres which are in table 19.

Left Hemishpere	Right Hemisphere
Speech/Language	Music
Sequencing (temporal ordering)	Spatial /artistic
Detailed	General
Analytic	Figure and iacia1 recognition
Reading and writing	Emotional
Concrete	Abstract.
Active	Receptive

Table 19: Types of functions attributed to the hemispheres.

The functions general, abstract of the right hemisphere suggests greater participation of right hemisphere in this decision making process.

CONCLUSIONS & DIRECTIONS FOR FUTURE RESEARCH

The results indicated that the transition durations, & the speed of transitions of the first two formants were imitated maximally (apart from the place cues). Also, some of the scaled temporal parameters (aspiration & affrication duration) were imitated maximum number of times. The results have lead to make several points. Firstly, the importance of transition durations & the speed of transitions calls attention on coarticulation which is reportedly defective in speech handicapped , especially so in the hearing impaired. More emphasis on coarticulation in these children would enhance their speech.

Second, the study throws same light on the perceptual mechanism. The results suggest that the perception may rely on a scaling of the temporal parameters rather than the absolute? temporal values.

Third, voice identification deserves some comments. The results indicate that in spite of the perceived similarity in two voice/ speach samples, there exist wide differences in the acoustic features used & the imitators use greater effort, which are points of consideration in Forensic voice identification.

Fourth, as the perceptual evaluation reveal *a* voice identification rather than a speech identification & as the cues for voice similarities were mostly suprasegmental in nature, more participation of the right hemispheres in these tasks *are* likely.

Fifth, the fact that the imitators were aware of the articulatory maneuvers to be changed to suit the models vocal tract, supports the view of Speech perception as perception of gestures.

In this context, it, is suggested that the future research be oriented on perception in children, perception in speech handicapped & perception of suprasegmentals to explore further the mechanism of speech perception which would help in serving the speech impaired better.

APPENDIX 1

REVIEW OF LITERATURE

THE STOP CONSONANTS: THEIR SPECIAL NATURE

The stop consonants are produced by occluding the oral cavity by an articulator. Air is held behind the articulator for sometime and is released. The stops are special in that they represent the nonlinearity 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 use of the acoustic overlapping of phonemes in the speech stream. Also, they have consistently produced evidence of phonetic level processing. They appear to be the most highly encoded speech sounds (Day and Vigorito, 1973) and they are the information bearing elements of speech.

" Five qualitatively distinct segments can be identified for stop consonants:-

1. A period of occlusion (silent / voiced),
2. a transient explosion (usually 20 IDS) produced by shock excitation of the vocal tract upon release of occlusion,
3. a very brief (0-10 ms) period of frication, as articulators seperate and air is blown through a narrow constriction, as in the homorganic fricative,

4. a brief period of aspiration (2-20 ms) within which may be detected noise excited formant transitions, reflecting shift in vocal tract: responses as the main body of tongue moves towards a position appropriate for the following vowel and
5. voiced formant. transitions, reflecting the final stages of tongue movement in to the vowel during the first few cycles of laryngeal vibration.

Out of these five segments, bursts and voiced formant transitions may serve as seperate cues to place of articulation of initial / b,d,g /. Many studies have been shown that transitions of second and third formants are sufficient cues for the place distinction (for example Delattre, Liherman and Cooper, 1955; Liberman,Delattre,Cooper and Gerstman, 1954) and these are infact the standard cues used in speech synthesis. Since the acoustic shape of formant transitions varies as a function of the following vowel, formant transitions *are* necessarily context dependent cues for stop consonants. The same is true for Velar bursts.

Hoffman found that, while bursts centered at frequencies above 3OOO hz acted as cues for /d/, burst cue for /g/ lay near the second formant of the vowel and were therefore, context dependent (of. Liberman, Delattre and Cooper, 1952) Hoffman could find no burst that would serve as a powerful cue for /b/ but this may have reflected, in part, the efficiencies of his synthesizer, rather than those of natural speech.

Infact, attention has recently turned to the question of how cues isolated in synthetic speech experiments act end interact in naturally produced speech. Cole and Scott (1974b) have argued that, while dormant transitions do provide essential phonetic information for the consonant phonemes, the major role of transitional cues is to provide information about the temporal order of phonetic segment© within a syllable. Cole and scott (1974a,b) go further to suggest that for / b,d,g / or / b,d / in stressed syllable initial position, the invariant place cue lies in the initial noise energy (burst and aspiration) before the outset of laryngeal vibration.

Release burst energy:- The energy (duration x intensity) in the transient release and it's following frication varies as a function of several factors, including the cross sectional area of the constriction just after release, the resonant cavity in front of the point of release, and perhaps, the release gesture itself. Thus, /b/, far which, there is essentially no front cavity and for which the release gesture is rapid (Fujimura,1961; Kuehn,1973), usually displays a weak transient and vertually no frication. While /g/ for which the cross sectional area between tongue and palate is relatively large , for which the front cavity is narrowly tuned, and for which tongue release is relatively slow, displays the longest burst of the three stops, including on occasion as Fischer-Jorgensen (1954) noted, a " double " release transient [perhaps due to a suction effect (Fant, note 1)]. Burst energy for /d/, with a smaller cross sectional *area* between tongue and alveolar ridge and a more broadly tuned front cavity than for /b/, falls midway. It could thus be predicted that increasing energy in *and* therefore perceptual importance of the burst, as the point of occlusion moves back in the mouth.

Burst enegy also varies with the vowel following. A major contrast is between front unrounded vowels, such as /u/. For /b/, increased cross-sectional area of the constriction just after release may give rise to a longer, and so more effective, release burst For /g/, the effect of front cavity elongation before rounded vowels may be counteracted by increased cross-sectional area of the palato-lingual constriction and narrower front cavity tuning than before unrounded vowels. Thus, if it is assumed that acoustic energy atleast partially determines auditory salience and perceptual weight, one might expect the release burst to play a more important, rote before unrounded vowels for /b/ and /g/, but exactly the reverse for /d/.

Release burst spectrum: - Spectral sections taken through the release burst of /b/ in nine vocalic environments generally show a broad curve with peaks over low frequencies, the low frequency peaks tend to be stronger before rounded than before unrounded vowels. For /d/, the spectral curve is broad and of a relatively high intensity, with peaks generally over high frequencies, above approximately 2000 Hz; the peaks tend to shift upward before unrounded vowels and be somewhat stronger than before rounded vowels. Apart from these minor rounding dependencies, /b/ and /d/ bursts *are* relatively uneffected by the following vowels. However, these bursts donot occupy invariant positions on the frequency scale in relation to their following vowels; the apical burst is spectrally continuous with F2/F3 of the high front vowels, but spectrally distinct from F2 of the back rounded vowels; for the labial bursts these relations tend to be reversed. The spectrum of the velar burst, on the other hand, is generally narrower and of a relatively high intensity, with it's main peaks close to F3 of a following back vowel, reflecting the changes from the front articulation of /gi/ to the articulation of /gu/. Thus, while labial and apical bursts *are* largely invariant on the frequency scale, but variable in relation to following vowel, velar bursts are *more* or less invariant in relation to the? fallowing vowel but variable on the frequency scale?.

Formant-trasition range and energy: - At least three articulatory factors underlie variations in formant-transition structure. First is variations in the extent of transitions as a function of place of articulation and

* following vowel, *for* bilabials, transitions *are* longer before unrounded than before rounded vowels. For apical stops, the distance between point of occlusion and vowel- target configuration varies, so that one might expect both devoiced and voiced transitions to be more effective cues to /d/ before back vowels, where transitions *are* relatively long, than before front vowels, where they are relatively short. Similarly, for velars the determining factor is a degree of similarity between the velar tongue constriction and that of the following vowel, in general, close vowels (such as /i/) will have relatively little transition, and open vowel (such as /a/), a more marked transitio.

is the onset of voicing relative to onset of release burst

i. e. VOT for consonantal release leads to an increase in the time taken for consonantal release leads to an increase in time taken for consonantal release leads t.o an increase in the t.inie taken for development of a transqlnttal pressure drop sufficient to initiate voicing, and to an increase ifi VG'f . if VtJ'f xa .increased, transitions into the following vowul Hi.ty he iaiMieiv ' ' xiil.) 1 e t .< • at, viiii, i riij uobi'l, , so that the dura11ari(ifdev<)i(.eijtr'.<11>itio n'>r(•1at.i.v<•tovoiced transitions is increased, since release burst duration (and so VOX) typically increases from labial to (apical to velar points of articulation •; Lislier and Abranisoin, 1964), it may be reasonably predicted that, the perceptual weight attached to devoiced transitions correspondingly increases". (Dorman, **Studdert-** Kennedy «< Raphael , 1976)

Finally, speakers differ in vocal-tract. shape and dimensions, as well as in articulatory habits and even two phonetically identical utterances of the same speaker *are* probably *never* identical acoustically. If chance variation is added in relative effectiveness of bursts and transitions, due to such factors as distance between speaker and listener (or between speaker and microphone), it might be concluded that predictions of perceptual weight attached to the several acoustic cues to place of articulation can be, at best, statistical, and that the likelyhood of any single cue being the sole determinant. of the percept in all contexts is extremely low. In this chapter a review of the cues for place, voicing and manner of articulation of stop consonants and the cues that distinguish stops from other consonants *are* presented.

CUES FOR
PLACE OF ARTICULATION

Many features cue the place of articulation of stop consonants of which VOT, closure duration, onset spectra, formant transitions and release bursts are a few. Voicing as a distinctive feature of English stop consonants in initial position is centered on the measure of ' VOT ' , the time of onset of laryncgeal signal relative to the noise pulse generated by the stop release. However, it has been shown that listener's selection of / b,d,g / Vs / p,t,k / responses to synthetic stop vowel stimuli is not determined entirely by VOT but also by the behavior of Fl of the following vowel.

The relative importance of VOT as against the presence Vs absence of Fl frequency shift after voice onset was assessed in several synthesis experiments in which VOT and Fi configurations are systematically varied. Varying VOT regularly effects a significant change in listeners judgements, and that varying Fl has some effect too, but this latter variation is neither necessary nor sufficient generally to shift judgements decisively from one stop category to the other. The data further suggested that the presence of an F1 rising transition after voice onset serves as a voiced-stop cue (Lisker, 1975).

Repp (1977) investigated the dependence of voicing boundary on place cues by varying F2 and F3 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 dependancy of the place boundary on VQT:- labial-alveolar-velar boundaries converge as VOT increases, resulting in a reduction of the size of the alveolar category;- was reported.

In Danish stop consonants (Fisher-Jorgensen, 1979), it was found that the closure increased as one proceeded from /k/ to /b/ and the open interval increased in the reverse; order for /p,t,k/. The relations are, for the closure p>t>k and for the open interval t>k>p.

Measurements of the onset spectra from real speech has revealed distinctively different gross shapes for different places of articulation. Diffuse spectra for labials and alveolars and spectra with relatively narrow energy concentrations for velars were reported by Stevens (1977).

In a series of experiments conducted by Blumstein and Stevens (1976), subjects were randomly presented with synthetically produced stimuli consisting of a 5 or 10 Ms noise burst followed by a brief voiced interval containing three formant transitions with onset and offset characteristics appropriate to the consonants /b,d,g/ in the environment of the vowels /a,i,u/ as well as stimuli with steady F2 and F3 transitions. The length of the voiced interval was systematically varied from 40 to 5 Ms. The results indicated that an onset spectrum consisting of the burst and the initial 5 to 10 Ms of voicing provide sufficient cues for the identification of the stop consonant and that vocalic information can be reliably devoiced from these brief stimuli containing only one or two glottal pulses (Blumstein and Stevens, 1976).

Fisher-Jorgensen (1979) observed that higher F2 values for stops were associated with front vowels and lower F2 values with a following back vowel. Variability of initial F2 values was found to be more for /k/ and /g/. Also, variability of /b/ was greater than F2 of /g/.

Liberman (1967) concluded that formant transitions were "important acoustic cues for the perception of consonants". Kewley-Port (1979) reported that spectral continuity of burst and formant-transitions served as cues to place of articulation in stop consonants. Further, Kewley-Port, Pisoni and Studdert-Kennedy (1983) investigated the perception of static and dynamic acoustic cues to place of articulation in initial consonants. Their results suggested that the listeners identify place better from stimuli with preserved dynamic acoustic properties than from those based on static onset spectra.

Dorman, Studdert-Kennedy and Raphael (1976, 1977) conducted three experiments to assess the role of release bursts and formant transitions as acoustic cues to place of articulation in syllable initial voiced stop consonants by systematically removing them from American English /b,d,g/, spoken before nine different vowels by two speakers, and by transposing the bursts across all vowels for each class of stop consonants. The results showed that bursts and transitions tended to be reciprocally related; where the perceptual weight of one increases, the weight of the other declined. They were thus shown to be functionally equivalent, context-dependent cues, each contributing to the rapid spectral changes that follow consonantal release. The results are interpreted as pointing to the possible role of the front cavity resonance in signalling place of articulation.

In a series of experiments Dorman and Raphael (1977) presented listeners signals consisting of an onset spectrum appropriate for one place of articulation followed at silent intervals from 0 to 150 Ms by transition cues appropriate intervals from 0 to 150 Ms by transition cues appropriate for a different place of articulation. They opined that the onset spectrum determines place? identification in some vocalic environments hut. not. others and that when the onset spectrum does determine place judgements it 'overrides' the place of articulation signal1ed by the transition cues over the salient intervals of upto 60 Ms.

Dorman and Raphael (1980) conducted a series of experiments to examine the identification of place of VCV syllables. They excised VCVs from the sentence, the little VCV dog spoken by a male and stored it in computer memory. To create stimuli with conflicting burst and transition cues. CV' s were first extracted fnmi the 5yl.lab.le stored separately. The stimuli were then recombined in a manner such that each burst was paired with each of the other vocalic portions. To create VCV stimuli in which the transitions cued one ploee of articulation and the burst another, the original VCVs were edited and the preclosure, vocalic portions of the utterances were stored in memory. The preclosure, vocalic sections were then combined with CV stimuli so that in each stimulus the closing and opening transitions specified one place of articulation and the burst a different place. Ten listeners participated in the perceptual experiments.

They found that not only do the burst and opening transitions affect the judgement of place, but also do the closing transitions and the duration of the closure Interval, The results suggest that there are multiple acoustic events which bear *on* the identification of a given phone and that those acoustic events are distributed over time (Dorman and Raphael. 1980 *).*

Erickson. Fitch, Hallwes and Lieberman (1977) investigated the trading relation in perception between silence and spectrum (Presence or absence of medial stop consonant in /split/ Vs /slit/). The duration of silence between the 's' noise and the vocalic portion of the syllable {temporal cue) and the presence or absence of those formant transitions that distinguish /plit/ and /lit/ (spectral cue) were the cues, The result was that the amount of temporal cue was 25 Ms less/more when the spectral cue was absent.

Keating and Blumstein (1978) conducted two experiments to explore the effects of lengthened transitions on the perception of stop consonants. They used synthetic /da~ga/ continuum with F2 and F3 transition lengths of 45, 95 and 145 Ms duration and ran a perceptual test with 20 untrained listeners. The results suggest that lengthened transitions minimally affect the perception of stop consonants. The identification and discrimination scores didn't vary systematically with increase in the length of transition duration.

Winits. Schieh and Reeds (1972) isolated stimuli *Segments* burst/ and burst + 100 Ms vowel) from words which initial and final /p.t.k/ and presented to listeners for the burst alone. Initials were identified better than finals and /t/ exhibited the highest level of identification. Also, in many instances, vowels could be identified on the burst portions, giving perceptual evidence for co articulations are with the additional. 100 Ms, the level of identification was high. The results of the experiments by Blumtein (1976) indicated that an onset spectrum consisting of the burst and the initial b-10 Ms of voicing provide sufficient cues for the identification of the stop consonants and that vocalic information can be reliably derived from those brief stimulus containing only one or two glottal pulses.

Repp (1984) studied the role of release burst as cue to the perception of stop consonants following /S/. His experiments demonstrated that silent closure duration and burst duration; closure duration and burst amplitude can be treated as cues for the "say-stay" distinction. Also, absolute not relative, burst amplitude was found to be important. The results suggested that listeners' sensitivity to burst in a labeling task was equal to their sensitivity in a burst detection task. All the experiments revealed that listeners were remarkably sensitive to the presence of even very week release bursts.

Mann and Repp (1978) investigated the category boundary shift in synthetic /da/-/ga/ and /du/-/gu/ continuum and found that the boundaries are shifted substantially when the syllables are preceded by one of the two synthetic fricatives /s/ or / ~ /. Also, more velar stops were heard following; /s/ than following / ~ /.

Dorman and Dougherty (1981) assessed the identification functions for stimuli from a two formant /bdg/ continuum at three levels of signal presentation, 55, 70, and 90 dB SPL. At 90 dB, the burst category was narrowed, the /d/ category virtually eliminated and the /g/ category greatly enlarged. They concluded that if high SPLs independent of cochlear damage can aJter identification functions, then perceptual experiments roust be conducted at equal SPLs rather than at equal SLs.

Miller and Eimas (1977), Kimas etal (1978) found that in the syllable-initial position both place and manner features underwent very similar forms of processing, though the acoustic information specifying the two features were not processed independently of one another, in that processing of one feature was infiuenced by the value assigned to the other feature.

VOICING AND MANNER CUES

Lisker (1977) suggests 16 parameters that cue voicing in /rapid/ Vs /rabid/. They are:

1. Presence/absence of low frequency buzz during the closure interval.
2. duration of closure.
3. Fl offset frequency before closure.
4. Fl offset transition duration.
5. Fl onset frequency following closure.
6. Fl onset transition duration.
7. /a/ duration.
8. Fl cutback following closure.
9. Fl cutback before closure.
10. VOT cutback before closure.
11. VOT delay after closure.
12. F0 contour before closure,
13. Fo contour after closure.
14. amplitude of /i/ relative to /a/,
15. decay time of glottal signal preceding closure and
16. intensity of burst following closure.

Voice onset time:- 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 conversations among Korean adults and in mothers speech to children were used. VOT values for word-initial apical stops in the splash of all four speakers showed considerable overlap of the weak and aspirated categories. The data suggested that although VOT is sufficient to distinguish the strong from the aspirated. stops, it cannot effectively distinguish either of these from weak stops.

Weismer (.1977) studied VOT in CVC target words and observed that longer VOTs were associated with tense as compared to lax consonants and that longer VOTs were associated with voiced as compared with voiceless final consonants. Klee. Weismer and Ingrisano (1976) compared glottal-supraglottal timing across consonant manner of articulation by normalising the duration of selected target utterances spoken in a sentence frame by 15 subjects and deriving a measure which expresses the timing of glottal supraglottal events relative to the duration of a potential unit of articulatory encoding namely the syllable. Thair data suggests that stop closure duration and VOT are not independent in connected speech.

Diehl (1977) using synthetic CV syllables, reported that the subjects tended to identify /p/ when the test syllables were preceded by a single clear /b/ (VOT = -100 ms) and they tended to identify /b/ when they were preceded by an unambiguous /p/ (VOT = +100 ms). This contrast effect occured even when the contextual stimuli were velar and the test stimuli were bilabial suggesting a featural than a phonemic basis for the effect.

Ohde (1978) using different VOTs examined the perceptual strategies and reported that 55 ms VOT was rated as p- like and 5 and 25 ms VOTs were rated as b-like. Keating (1979) measured VOT differences in production and perception of Polish and English stops in minimal pairs, sentences and conversation, the VOT distribution for voiced and voiceless stops were clearly separated in polish but not always in English, especially in casual speech. In contrast, the Polish VOT perceptual categories were somewhat unstable and did not always match the VOT production categories and the English perceptual categories are quite stable.

Lahiri (1980) acoustically analysed word initial 5 .ops of Hindi. Punjabi and Bengali and reported that the feature interrupted voicing differentiates the so-called voiced aspirates from the other three categories. These stops were reported to be characterised by a pattern of pre-voicing followed by approximately 100 ms of silence and then resumed phonation. The other stop categories of these languages were reported to be characterised by lead, coincident and lag VOT

In a series of experiments conducted by Keating, Mikos and Ganong (1981) it was noticed that Polish speakers' perceptual boundaries fall in the gap betitween their production categories and that for ranges of VOT which include a few pre-voiced stimuli, their boundaries are substantially shifted. Americans show no shifts of this type, although they do show some small shifts.

Preceding vowel duration:- Experiments done at Haskins laboratory suggest that shortening the duration of vowels preceding final stops and fricatives causes them to be perceived as voiceless.

Raphael (1972) synthesized a variety of minimal CVC (C) pairs (with word final stops, fricatives with clusters) and varied the length of each vowel. *Be* found that regardless of cues for voicing or voicelessness used in the synthesis of the final consonant or cluster, the listeners perceived he final segments as voiceless when they were preceded by vowels of short duration and as voiced when they were preceded by vowels of long duration. Thus, cue to voicing characteristic of a final consonant was not within the articulatory period of the consonant itself, but within the duration of the preceding vowel.

Presence of voicing during closure period of & final consonant or cluster was found to have a miner volue compared to that of vowel duration. Preceding vowel duration was found to be sufficient and consistent cue to the perception of voicing characteristic of word final stop, fricative or cluster. It was also found that perception cued by preceding vowel duration was continuous than categorical.

Vowel duration was found not to be equally effective cue before each of the consonant type tested. It was greater for the vowels preceding fricatives. The cue of preceding vowel duration was found to be more effective before stops and clusters than before fricatives (Raphael, 1972). In naturally produced speech the vowel durations in CVC condition were measured and no clear relation was seen between the length of the preceding vowel and the perception of the voicing characteristic of the final consonant (Frain and Bischoff, 1976).

Liberman (1977) found that the introduction of the long noise in the second syllable of the phrase "Gray ship" disposed the listeners to displace the stop to the first syllable so that *they* hear a syllable final stop -Great ship. Using synthetic stimuli, Raphael (1380) reported that the /d/-/t/ phoneme boundaries plotted as a function of vocalic duration, fell within 3 ms of each other indicating that initial CV transitions contribute equally with steady state formants to the perception of that vowel duration which cues final consonant voicing. A difference in duration characterises the distinction between /b-d/ or / d/ on one hand and /b-t/ or / t/ on the other hand and this difference in duration has been attributed to the vowel, *i*, e., vowel duration does affect the perception of a syllable final stop as voiced or voiceless.

In an attempt to determine the effective vowel duration, Raphael et al (1980) considered VC and CVC syllables and found that the duration of the initial transitions (as in CVC syllable) is not quite as effective as the duration 01 the steady-state vowel in cueing the voicing contrast or stops in final position. They suggest that the transitions of syllable-initial stop consonants are almost fully incorporated into the durational estimate that listeners make in determining the voicing class of syllable final stops; since these transitions contain information about the vowel. Thus, the effective { and perceived) duration of the vowel is taken as that span of the signal that contains information about the vowel and this includes any portion that shows the acoustic results of coarticulation between consonants and vowels.

There is evidence that FO varies systematically as a function of laryngeal timing. (1) FO at voicing is highest in utterances beginning with voiceless aspirate stops than in utterances beginning with voiced stops? (Houses and Fairbanks, 1953; Lofquist, 1975; Umeda, 1981). (2) FO reduces after voicing onset in the former but increases In the latter (Lea, 1973). Other studies { Umeda, 1981; House and Fairbanks, 1953) have shown that FO tends to reduce in both voiced and voiceless environments, though it is more in the former than in the latter.

Fundamental frequency:- Experiments on the identification of ambiguous syllable differing only in the pitch curve at voicing onset showed that a low rising pitch leads to perception of an initial stop consonant as voiced while a high falling pitch leads to the perception as voiceless. Although not used reliably by all subjects, high performance from some subjects (90% of subjects) leac the authors to suggest that this pitch dip/rise is more than a naturalness artifact.

The average; FO contours on either side of the stops obtained from Hindi stops (p, ph., b, bh) showed that the breathy voiced stops lowered FO a great deal. However, no consistent effect on FO by other stop types were observed (Ohala, 1978). Gruenfelder's (1979) experiments indicated that although FO contour was capable of functioning as a cue to voicing of a post-vocalic consonant in perception, its generality as a perceptual cue to voicing may be extremaiy limited given that the distinction is not reliably made in speech production as well.

Ohde (1982) studied the effects of consonant environment on F® of vowel in children. Three girls and three boys between eight and nine years of age recorded five repetitions each of voiceless aspirated /ph, th, kh/, voiceless unaspirated /sp, st, sk/ and voiced /b,d,g/ stops in combination with vowels /i,e,u,o,a/. Fo for the first five glottal periods, FO of vowel target and VOT were determined for each utterance. Nearly Identical results of VOT were obtained for voiceless aspirated and voiced stops. Differences were found between voiceless unaspirated and voiceless aspirated stops. Despite the similarity in VOT between voiceless unaspirated and voiced stops, FO contours of voiceless unaspirated stops were more similar to voiceless aspirated than voiced stops. In general, there was a decrease in FO from the first glottal period to the second glottal period for all voicing conditions. Also, F0 was usually higher in the environment of voiceless stops than voiced stops.

Ohde (1982) studied the influence of contexts on temporal and *F0* properties of speech. He measured VOT. Percent decreases in F0 from the first to second glottal period and absolute value of F0 for the first five glottal periods. Results revealed greatest differences in VOT between isolated syllables and syllables produced in carrier phrase for voiceless stops and F0 variation between these contexts was greatest for voiced stops.

In 1984, Ohde experimented with speech sample consisting of five vowels /i,e,u,o,a/ paired with voiceless aspirased /ph, th, kh/ and voiced stops /b, d, *g/* in CVC syllab.es spoken *by* three adult males, both in isolation as well as in context. FO fell substantially in both contexts, for both voiceless and voiced stops after voicing onset.

In another experiment, Ohde (1984) obtained similar findings for voiceless aspirated, voiceless unaspiratted and voiced unasplrated stops in isolation. Although a large proportion of the utterances manifested a reduction in FO from first to second glottal period, magnitude of reduction varied across voicing categories according to voicing category, vowel and consonant place of articulation. The changes in FO were consistently greater for voiceless unasplrated sounds than for voiceless aspirated and voiced stops. These differences were generally greatest for /i/ and least for /a/.

Only with voiced stops, some systematic differences in vowel effects across place of articulation was found, which may be a function of different pattern of tongue movements from stop place of articulation to the vowel during onset of glottal excitation for voiced syllables. For /d/ and /g/, changes in FO were consistently greater for /a/ and. least for /i/ and /u/ (unlike voiceless aspirated stops). FO was higher for voiceless aspirated and voiceless unaspirated stops than voiced stops at all periods. Ohde speculates that both aerodynamic and vocal cord tension factors continue to influence F0 at voicing onset. Also, aerodynamic influences appear less important than other factors as he obtained higher FO at period one for voiceless unaspirated stops compared to voiceless aspirated stops. However this contradicts the aerodynamic hypothesis.

Ohde attributes the FO perturbations to vocal cord tension hypothesis putforth by Halle and Stevens (1871), and Stevens (1975) which states that the vocal cords are slack in order to facilitate voicing during voiced stops and stiff in order to inhibit voicing during voiceless stops (both aspirated and unaspirated) and these states spread to adjacent vowels. Further support to this study comes from EMG findings by Honda (1981) and cineflurographic studies by Perkell (1969), Honda (1981) hypothesised that a forward position of the hyoid results in increase in vertical tension in the larynx which in turn increases the F0 (as in the case of /i/). Perkell (1969) found that the distance between the dorsum of the tongue and cervical vertebrae (C2 and S3) was large during the production of alveolar and velar stops, indicating a forward tongue position. As genioglossus is involved, it would also produce a forward shift of hyoid bone resulting in a high FO (Honda, 1981). Since these mechanisms apply to both alveolar and velar consonants and high vowels, there should be less of change in FO from voicing onset to the vowel in these cases than in the environment of low back vowels. The results obtained in this study for voiced alveolar and velar consonants are consistent with this finding.

Since the tongue is free to anticipate the following sound during the production of labial stops, smaller differences between /i/ and /a/ in F® perturbations from voicing onset to the vowel can be expected for labials than alveolars and velars. A left-right coarticulatory effect found across two phones /s/ + stop + vowel (i.e., vocal cord tension during /s/ spreading to stop and vowel) also support the CV syllable model of speech production by Kozhevnikov and Chistovich (1965) according to which, articulatory movements are organized as a series of consonants preceding a vowel.

Representation of cross language voicing contrasts has bean a problem since the mapping between phonological categories and their physical phonetic realizations is not one-to-one. Keating (1984) has argued that toe representation of such contrasts for stop consonants must involve purely abstract features (+voice) and (voice), which map onto phonetic categories for stops based on VOT in different ways for different languages, However, the author argues that an articulatory analysis of voicing contrasts based on the presence or absence of glottal opening and closing gestures, may provide a more nearly one-to-one mapping between phonological and physical categories.

The basic phonological units in their articulatory gesture approach are articulatory gestures; organised. patterns of movement within the oral, laryngeal and nasal articulatory systems. Thus, according to Browman and Goldstein (1986), voiced stops can be represented as a constellation of two gestures (an oral constriction gesture tightly coordinated with a glottal opening and closing gesture), while voiceless stops can be represented as single oral constriction gesture. Differences between aspirated and unaspirated voiceless stops can be captured directly by this timing between the two gestures in the constellation (Lisked and Abramson, 1984; Browman and Goldstein, 1988).

Voicing contrasts: Although the timing and size of the glottal gestures in English and French differ, categorization of stops as (-voice or +voice) in utterance-medial position correlates well in both languages with the presence Vs absence of a glottal opening and closing gesture. In absolute initial position, the glottis is already open (for breathing) and the opening portion of the glottal opening and-closing gesture is therefore not actually observed. Thus, the relevant difference between +voice and voice stops in this position is in the relative timing of the adduction of the vocal folds. Both French /d/ (Bengueral etal 1978) and English /b/ show the same pattern of glottal adduction, Thus, a physical characterization using articulatory gestures captures the voicing contrast in English and French for utterance-initialas well as utterance-medial position.

In Danish there is a contrast in initial position between aspirated and unaspirated stops both of which show glottal opening gestures (unaspirated stops have smaller glottal gestures which are timed differently). While Keating's abstract analysis predicts that Danish should behave like English and French in showing failing FO pattern following -voice stops and a low rising pattern following +voice stops, gestural analysis predicts that Danish, stops, both of which have glottal gestures, should show high falling F0 patterns. The latter also predicts that Danish will be unlike English and French, which contrast presence Vs absence of glottal gestures, which is supported by Pterson's (1983) study showing FO patterns following aspirated and unaspirated stops to be the same (high and failing). Thus there is correlation between glottal gestures and FO patterns rather than between voicing categories and FO patterns. Thus, analysis of cross linguistic voicing contrasts in terms of glottal opening and closing gestures accounts for similarities between languagoi as well as or, in the case *of* FO patterns, better than the purely abstract analysis posites by Keating.

Source characteristics of a vowel may differ according to the voiced/voiceless nature of adjacent consonants. The post-vocalic consonants could be particularly crucial as vocal fold abduction for a voicless consonan*t* pay be initiated considerably before oral occlusion. Chasaide and Gobl (1987) studied CV(:) C utternaces (where C voiced/voiceless labial stop/ fricative) of female speaker in English and French. Results indicate that the latter par: of the vowel preceding a voiceless consonant shows a marked drop in excitation strength and a steeper spectral slope an see*n* when the vocal cords are opening but vibrating. A spectral consequence of this abducting gesture is a widenies of the F1 bandwidth, and an upward shift in formans frequencies. These effects were much less in French than in English and Swedish. The preceding consonant had comparatively little effect. Full excitation was achieved. almost immediately.

Closure duration;- Measurements of natural speech by Lisker (1978) indicated that stop closure duration does not seen to seperate /b/ from /p/ across speakers and that the phonetic effect of manipulating silent 'closure' differs greatly for different tokens of the source word produced by a single speaker. They concluded that "Neither duration nor the acoustic nature of closure was sufficient to predict listeners phonetic interpretations of the acoustic segment corresponding to the interval of articulately occlusion" (Lisker, 1978). Price and Lisker (1079) reported that shortening the closure of /p/ had relatively little effect, while lengthening the silenced /b/ closure produces a dicisive labeling.

Port (1980) generated synthetic stimuli of the words /dipper/ and /dibber/ with five durations of /dib/ (140-260 ms) and the data of the perceptual teat (16 listeners) implied that the voicing effect on preceding vowels cannot be handled insightfully with a postsegmental temporal implementation rule that modifies the vowel.

Fitch (1980) used slowly spoken /dab/ syllables, shortened to match duration of medium and fast rate /dab/ syllables by deleting pitch pulses from the steady state region of the vowel, A variable silent interval and a constant /b1/ were appended to each of the /dab/ syllables making /dabi/ to /dapi/ continua. Voicing boundary was found to be at a nearly identical ratio of closure duration to preceding syllable duration at all speaking rates. With synthetic speech, it was found that longer syllables with proportionately short steady-state sections needed less silence than shorter syllables with proportionately *long* steady-state sections. He concluded that perceptual voicing boundary is sensitive to the dynamic structure of the preceding syllable and not simply it's duration.

Ingrisano, Hillenbrand, Smith and Flege (1982) in a perceptual study of syllable-final stops reported that very large amounts of voicing had to be removed from the closure intervals before perception changed from voiced to voiceless. Repp (1982) opines that stop perception (in case of /s-lit/) seem to depend as an integrative perceptual strategy and that silence is not always sufficient for stop perception.

The acoustic cues that underlie the voicing features in stop consonants in initial position are reflected interm of the relative onset of the Fl transition (i,e Fl cutback) the presence of aspiration in the higher forroants (Libermans. Delattre and Cooper, 1958).

VOT Vs spectral cues;- Stevens and Klatt (1974) reported a significant trading relationship between the VOT and presence Vs absence of a significant formant transition. The presence or absence of a rapid spectral change following voice onset produces up to 15 ms change in the location of the perceived. phoneme boundary as measured interms of VOT.

First formant and first formant frequency:- Summerfield and Haggard (1977) ran three experiments to assess (l) the role of Fl onset as a ewe to the voiced percept. (2) whether spectral influences on tit© perception of voicing was a function only of the frequency of Fl or the distribution of energy in both Fl and the higher formants and (3) whether a rising Fl transition was a positive cue to voicing independent of its onset frequency. They used synthes used /CV/ syllables and manipulated Fl onset frequency and Fl transition duration/ extent independently. They reported that the major effect of Fl in initial voicing contrasts was determined by its perceived frequency at the onset of voicing. Also they show that a periodically excited Fl transition is not, per se. a positive cue to voicing. In further experiments, the relative levels and the frequencies at the onset of voicing of both Fl and F2 were manipulated. The influences on the perception of stop consonant voicing that resulted were determined specifically by the frequency of Fl and not *by* its absolute or relative level or by the overall distribution of energy in the spectrum. The results demonstrate a complementary relationship between perceptual cue sensitivity and production constraints. In production, the VOT characterising a particular stop consonant varaes inversely with the degree of vocal-tract constriction, and hence with the frequency of Fl, required by the phoneme following the .stop; in perception, the lower the frequency of Fl at the onset of voicing, the longer the VOT that i- required to cue voicelessness. In this way, the inclusion of Fl onset frequency in the cue repertoire for voicing and the establishment of the cue trading relationship reduce the problem of contextual variation that would be met were VOT alone or some other amalgum of cues are the *only* basis of the voicing distinction (Summerfield and Haggard, 1977).

Aspiration:- Reeds and Wang (1961) suggest that aspiration seems to be a more dominant cue than voicing in the perceptual separation of the two classes of stops namely initial voiced and initial voiceless. This was fully supported by the results of a tape splicing experiment. Two experiments conducted by Repp (1979) demonstrated that the amplitude of aspiration noise is a cue for the distinction between voiced and voiceless syllable initial stop consonants in English and that can be traded for VOT. In a spectrographic analysis of CVCCC monosyllables produced fay six speakers, Walsh and Parker (1080) observed that the. voiceless stops typically displayed a coarticulated glottal stop. Also they observed that a glottal stop coarticulated with a following nonvelar voiceless stop produced vocal fry. which not only served as a secondary cue for voicelessness, but also helped in explaining the differences in vowel length and perceptibility exhibited by velar and nonvelar post-vocalic stops.

It has been found that voiceless stops have greater duration than voiced stops only in the intervocalic post-stressed position (Stathopoulos and Weismer, 1979). The perception of voicing in final stops was investigated by Wolf (1976). He truncated syllables at various points and presented it to subjects under two response conditions (one voiced Vs voiceless, other C Vs C). He found that the formant transitions, closure, burst and vowel duration are important in determining voicing.

STOPS VS OTHER CONSONANTS

In an attempt to find out the effect of preceding liquids on stop consonant, Mann (1980) excised CV portions from natural tokens of /alda/, /alga/, /arda/ and /arga/ any replaced with closely matched synthetic stimuli forming a /da~ga/ continuum. On a perceptual test, it was observed that the listeners gave more /g/ responses to stops which replaced /ga/ than to those which replaced /da/, indicating that the preceding VC portion (/al/ or /ar/) contained cues to the following stop consonant.

In an experiment to identify an invariant acoustic property which can accurately distinguish stops and glides.

Mack and Blumstein (1983), analysed natural utterances of /b/ and /w/ in the environment of 5 vowels /i.e.a.o.u/. They calculated the degree of relative amplitude change occurring in the vicinity of stop and glide release and. found reliably larger changes In energy associated with the stop release than the glide release across vowel contexts and speakers, These changes seemed to provide an invariant property characterizing the stop-glide contrast. These results were also extended to stops /d/ and /g/ and the glide */y/.* Also, they opined that absolute vowel duration and frequency measurements were unlikely candidate for an invariant property.

Miller and Baer (1983) examined changes in transition durations of /b/ and /w/ as a function of speaking rate and found that, in general, the transition duration increased with syllable duration. The transition duration that could optimally distinguish /ba/ from */Wa/* was not constant. Although, at any given rate, /wa/ transitions were longer than the /ba/ transitions, When pooled across rate. transition durations for /ba/ and /wa/ were overlapping. Also magnitude of difference, between average /ba/ and / wa/ transition durations increased with reductions in speaking rate.

In a series of identification experiments, Walsh and Diehl (1987) created sets of synthetic */tea/* and /wa/ stimuli in which transition duration and rise time varied orthogonally. They found that both variables affected labeling performance in the expected direction, but transition duration was by far the more important factor,

Zakia and Kingston (1987) from their experiment on the role of transition rate in place perception concluded that there was a significant interaction within vowel and place of articulation which implied that the magnitude, if not the rate of the transition was a significant cue to the place.

Repp and Mann (1980) in a series of experiments using CV portions excerpted from natural utterances (ta, ka, sta, ska) both in isolation and in combination with synthetic stimulus reported that listeners report more velar stops following /s/. Also, the natural fricative noises contained cues to the following stop consonants.

CONCLUSIONS

The review of literature highlights interesting findings suggesting possibilities for future research. First of all, the stop consonants have several temporal and spectral cues, the information of each of which, and the way in which they combine to produce a percept is not yet known. Secondly, some of the temporal parameters seems to vary across languages suggesting that the perception by non-native speakers may be different from the native speakers and that the selectivity of cues may vary across languages.

Also, the outcome of speech production data of children. is suggestive of the fact that children's perception of adult forms and child forms may be entirely different if perception is based on production. Further, if motor theory is accepted the reference for perception in speech handicapped may be entirely different than in normals. And, finally, the perceptual process of suprasegmentals is least attempted with. These warrant research in the area of infant perception, perception in speech handicapped and perception, of Suprasegmentals.

APPENDIX II

WORD LIST USED FOR THE EXPERIMENT

WORDS WITH STOPS/ AFFRICATES IN MEDIAL POSITION

1. aka :la
2. akhila
3. agi
4. aghora
5. acala
6. ajji
7. atavi
8. adi
9. adha:la
10. ati
11. athava
12. adu
13. adhar a
14. apa:ra
15. abba
16. abhaya
17. ukku
18. ugi
19. ucita
20. ujju
21. uda
22. uttama
23. uttha:na
24. udara
25. uddha: ra
26. upa: ya
27. ubbu
28. ubhaya
29. a ke
30. a *ga*
31. a C6
32. *a* ji:vii
33. *a* ta
34. a du
35. a ta
36. a di
37. a dha:ra
38. a.pattu
39. a bha:ra
40. i ke
41. i ga
42. i ce
43. i ju
44. i *ti*
45. i du
46. i ta
47. o kuli
48. o ta

WORDS WITH /K/ IN INITIAL POSITION

49. *o:* du
50. o:tika:ta
51. o:du
52. kaccu
53. katuka
54. kadi
55. katte
56. kabadi
57. kanaka
58. kara
59. kali
60. kavi
61. kasa
62. kahi
63. kali
64. kiccu
65. kitaki
66. kidi
67. kittu
68. kibbotte
69. kinnara
70. kiruba
71. kilubu
72. kivi
73. kise
74. kuccu
75. kuti:ra
76. kudi
77. kuttu
78. kube:ra
79. kuntala
80 kuri
81. kula
81 kuvara
83. kusuma
84. kuhaka
85. kulla
86. keccu
87. ketta
88. kedu
89. ketta
90. kene
91. keri
92. kelasa
93. kesuvu
94. kelage
95. koccu
96. kottu
97. kodu
98. kobbu
39. kone
100. kori
101. kole
102. kosaru
103. kolaku
104. ka: ta
105. ka: du
106. ka:tura
107. ka:nana
108. ka:raija
109. ka:la
110. ka: vu
111. ka:su
112. ka:li
113. ki : ta
114. ki :re
115. ki: lu
116. ki:vu
117. ki :lu
118. ke:g*u.*
119. ke:tu
120. ke:ru
121. ke:vala
122. ke:sari
123. ke:lu
124. klefsa
125. sulka
126. namaska:ra

127. WITH /KH/ *IN IN*ITIAL POSITION

127. khaga
128. khaja:ne
129. khadga
130. khanda
131. khadi:ma
132. kharidi
133. khala
134. khuddu
135. khusi
136. khedda
137. kha:ki
138. kha:ji
139. kha:te

140. kha:di
141. kha:sa
142. kha:ra
14 3. khi:ru
144. khe:cara

WORDS WITH /G/ IN INITIAL POSITION

145. gagana
146. gaja
147. gatti
148. gadi
149. gati
150. gandha
151. gamaka
152. gari
153. gala:te
154. gavi
155. gittu
156. gida
157. giri
158. gilaki
159. guggu
160. guccha
161. guttu
182. gudi
163. guttige
184. gunugu
165. gummma
166. guru
167. gula. ma
166. guli
169. gejje
170. gedde
171. gere
172. geluvu
173. gelati
174. gojju
175. goddu
176. gottu
177. gone
178. gombe
179. gorake
180. golasu
181. ga:ju
182. ga:tu
183. ga:di
184. ga:tra
185. ga:na
186. ga:mpa
187. ga:re
188. ga:li
189. ga:vila
190. ga:la
191. gi:cU
192. gi:jaga
193. gi:tu
194. gi:te
195. gi:ru
196. ge:li
197. gra:sa
198. varga

WORDS WITH /GH/ IN INITIAL POSITION

199. ghatta
200. ghante
201. gha:tu
202. gha:ta
203. ghe:nda
204. gho:ra
205. gho:sa
206. *vya:*ghra

WORDS WITH /C/ IN INITIAL POSITION

207. cako: ri
208. caccu
209. cata
210. caddi
211. cadara
212. canda
213. capa:ti
214. carma
215. calana
216. cali
217. cikka
218. citike
219. cinna.
220. cippu
221. cirate
222. cilaka
223. civutu
224. cukki
225. cuccu
226. curuku
227. culuku
228. celuvu
229. cokka
230. ca:ku
231. ca:cu
232. ca:di
233. ca:pe
234. ca:ra
235. ca:li
236. ca :vadi.
237. ci:ti
238. ci:ne
239. ci: pu
240. ci:ru
341. *ci:la*
242. ce:ia
243. ce:la
244. ce *:* sts
i 5

WORDS WITH /CH/ IN INITIAL POSITION

245. chadi
246. chatri
247. cha:ti
248. cha:pa
249. chala
250. cbali
Pi

WORDS WITH /J/ IN INITIAL POSITION

251. jakana
252. ji:ku
253. jinke
254. jajju
255. ja:ji
256. jataka
257. juttu
258. jetti
259. jade
260. ja:du
261. je:da
262. jote
263. ja:ti
264. ji:ta
285. jnna
266. jina
267. ja:naki
268. jambha
269. ji:nu
270. je:nu
271. jurnuki
272. jompe
273. ja:mi:nu
274. Jala
275. jille
276. julma :ne

277. jollu
278. ja:J.a
279. javali
280. ja:va
281. ji.:va
282. jalaka
283. ja:lu
284. jaha.iu
285. ja:sti

WORDS WITH /JH/ IN INITIAL POSITION

286. Jhagala
287. jhari
288. jhala

WORDS WITH /T/ IN INITIAL POSITION

289. tagaru
290. torige
291. ta:riga
292. tasse
293. tollu
294. kasta

WORDS WITH /TH/ IN INITIAL POSITION

295. thakka
296. thikami
297. tha:ne
298. tha:vu
299. the:vani
300. sasthi

WORDS WITH /D/ IN INITIAL POSITION

301. darigura
302. doriku
303. dabbi
304. dubba
305. da:bu
308. damaru
307. dumma
308. da:maru

WORDS WITH /DH/ IN INITIAL POSITION

309. khadga
310. hegde

WORDS WITH /DH/ IN INITIAL POSITION

311. dhakke
312. dha:ku

WORDS WITH /T/ IN INITIAL POSITION

313. tagalu
314. tegi
315. togalu
316. ta:gu
317. te:gu
318. tucca
319. tijori
320. ta:ja
321. te:jassu
322. tatt.e
323. tittu
324. tuti
325. tottu
326. ta:taki
327. ti:te
328. tada
329. tode
330. ti:du
331. tadige
332. tapa
333. tippe
334. tupa:ki
335. teppa
336. toppe
337. ta:pa
338. te:pe
339. tampu
340. timingila
341. tumba
342. ta:mra
343. tale
344. tilaka
345. tula
348. telugu
347. tola
348. ta:lluku
349. te:]u
350. tava
351. tivi
352. teyalu
353 tovve
354. te:vu
355. ti:yra
356. te:va
357. taia
358. tili
359. tuli
360. telu
361. to^i
362. ta:lu
363 rakta
364. utkata
365. spu:rti
368. ra:tri
367. hasta
368. vatsa

WORDS WITH /TH/ IN INITIAL POSITION

389. thadi
370. thand i
371. artha

WORDS WITH /D/ IN INITIAL POSITION

372. dakku
373. dikku
374. di:kse
375. darige
376. datta
377. ditta
378. da:tu
379. dada
380. duddu
381. dodda
382. dattu
383. da:ta
384. dana
385. dina
386. donne
387. da:na
388. di:na

389. dabhu
390. dibba
391. duba:ri
392. dara
393. duranta
394. dore
395. da:ra
396. di:rgha
397. dalla:li
398. dula:yi
399. davana
400. divasa
401. devva
402. da:ve
403. di:vige
404. de:vi
405. dasara
406. dese
407. da:sa
408. dahana
409. da:ha
410. de:ha
411. da:la
412. udga:ra
413. va:gdana
414. druda

WORDS WITH /DH/ IN INITIAL POSITION

415. dhikka:ra
416. dhana
417. dha:nya
418. dhamani
419. dhima:ku
420. dha:di
421. dhavala
422. dha:li
423. randhra
424. ardha

WORDS WITH /P/ IN INTIAL POSITION

425. pa:ti
426. pagade
427. pogaru
428. pacce
429. peccu
430. pa:ci
431. pe:cu
432. pata:ki
433. piti:lu
434. puta
435. pettige
436. pottana
437. pe:ta
438. pade
439. pidice
440. pudi
441. pa:du
442. pi:de
443. pana
444. pindi
445. pada
446. podaru
447. pa:da
448. pe:de
449. pampa
450. palya
451. pa:lu
452. po:lu
453. pavanu
454. pa:vu
456. pase
457. pustaka
458. pa:se
459. pahare
460. pulaka
461. pa:lu
462. pi:lige
463. pra:na
464. arpane
465. bha:spa

WORDS WITH /PH/ IN INITIAL POSITION

466. phani
467. phaji:ti
468. phe:da
469. phe:ni
470. phala
471. phara:ri
472. pha:1a:ksba
473. sphuarbi

WORDS WITH /B/ IN INITIAL POSITION

474. baka
475. bikku
476. bekku
477. bokke
478. ba:ku
479. be:ku
480. baca:vu
481. biccu
482. beccu
483. boccu
484. ba:cu
485. bata:ni
486. bitti
487. butti
488. betta
489. bottu
490. be:te
491. badi
492. bidi
493. buda
494. bedagu
495. ba:du
496. bi:du
497. bo:du
498. batta
499. bittu
500. betta
501. ba:tu
502. be:ta:la
503. bana
504. bindu
505. bennu
506. ba:nu
507. be:ne
508. bembattu
509. bombu
510. bimba
511. bale
512. bila
513. bele
514. ba:lu
615. be:la
516. bavane
517. bevaru
518. ba:vi
519. be:vu
520. basadi
521. bisi
522. besa
523. bi:su
524. be:sige
525. bahu
526. ba:hu
527. bale
528. bilalu
529. bele
530. ba:lu
531. bi:lu

532. be:le
533. kubja

WORDS WITH /BH/
IN INITIAL, MEDIAL
POSITION

534. bhakta
535. bhika:ri
536. bhi:kara
537. bhariga
538. bhajane
539. bhuja
540. bhatta
541. bhe:ti
542. bharida
543. bhatta
544. bhi:ti
545. bha:me
546. bhi:ma
547. bhalle
548. bha:va
549. bhuvana
550. bhava
551. bha:se

APPENDIX III

ARTICLES PUBLISHED/ SENT FOR PUBLICATION

ON THE PROPERTIES OF VELAR /K/ IN KANNADA

U.S.Sridevi

S.*R.Savithri*

Deptt of $peech Sciemces

All India Institute of Speech & Hearing

Mysore- 570 006

PRESENTED AT ISHA, 1989, MADRAS & SUBMITTED TO THE JOURNAL OF THE ALL INDIA INSTITUTE OF SPEECH AND HEARING.

1. Intraducti*an:*

The stop consonants are produced by occluding the oral cavity by an articulator, behind which air is held for sometime and released. Acoustic analysis of s t o p consonants has showed five distinct sygments namely occlusipn, burst, frication, aspiration, and voiced formant transitions. Stop consonant are unique in that they represent the noalinearity of the speech production and perceptual system. The demonstrate theredundancy of acoustic cues avai1ble to distinguish speech Bounds. They provide the best example *of* listener use of acoustic overlapping of phonemes in the speech stream. They have consistent1y produced evidence of phonetic level processing. They appear to be the most highly encoded speech sounds. In word initial position, they provide the most important and reliable phonetic information about a word's identity in fluent speech. Also, they are the most information bearing elements of speech.

Over the past few decades, seve*ral* investigators (Fischer-Jorgensen, 1954; Fujimura, 1961; Raphael, 1972 Resting, 1979; Lahiri, 1980; Stevens, 1975) have studied stop consonants in various languages of the world and has attempted at explaining the manner, place and voicing using ©caustic; parameters. This study aims at extracting various temporal and spectral parameters of the velar unvoiced, unaspirated stop consonant /k/ in Kannada language.

2. **METHOLODOLOGY:**

STIMULUS: Meaningful Kannada words with /k/ in various positions and in various phonetic contexts were selected for this study. They were so selected to study the effect of (1) preceding vowel (2) following vowel (3) following consonant (4) clustering and (5) embedding in a sentence on the temporal land (spectral aspects of /k/.

Short vowels /a/, /i/, /u/, /e/ and */a/* and their long counterparts /at/, /is/, /u;/, e:/ and /o:/ were seiected. Also, amoncs the consonants, the affricates /c./, /j/ (palatal), the stops /t,/,/d/ (retroflexes) . /t/, /t/ (dentals), /p/, /b/ (bilabials), the nasals /n/ (retroflex), /n/ (dental), /n/ (palatal), /m/ (bilabial) and /r/, /l/ , /v/, /s/, /s/, /h/, and /l/ were selected. All the combinations selected were in meaningful Kannada words.
The *carrier* phrase /i:ga na:n he:lti:ni WORD/ was used for embedding in sentences. Totally there were 56 words.

Subject: One young male adult Kannada speaker aged *26* year is served as a subject. He was a trained speech pathologist with normal hearing and speech.

Method: The subject was familiarized with the material which was visually presented with the words/sentences, one at a time. He was instructed to speak the words sentences into a microphone which was placed at a distance of 10 cms from the mouth. All the utterances were recorded on a high fidelity magnetic tape and these were subjected to spectreggraphic analysis, wide bandbar and section type of spectraggrams were obtained for all the stimuli. From the spectrggrams. 17 temporal parameters, namely, closure duration, burst duration, VOT, consonant duration, precending vowel duration. transition duration of Fl, F2 and F3 of the precending and fallowing vowel, Speed of transition of F1. F2 and F3 of the preceding and following vowel and 18 spectral parameters namely Terminal Fl, F2 and F3 of the precending vowel, stead FO of the preceding and the following vowel , terminal Fo of the preceding vowel, initial FO of the following vowel, burst amplitude, amplitude of the consonant, precending and following vowel, Fl, F2, F3, Bl, B2, 83, LI, L2 *and L3 of* the consonant were extracted. The signal was fed in to the Visipitch to obtain FO measurements.

Results: The data was tabulated and a principal component analvsis was performed. The results of these are in table *1* and 2.

In general, the duration of /k/ was longest in the geminate-cluster condition and shortest in singleton condition. The transition durations and speed of transitions of Fl, F2 and F3 of the preceding vowel decreased successively. FO patterns of the preceding vowels were falling no and that of the vowel following was either falling or raising. The bandwidths and the intensities of the Fl and F2 declined respectively. The principal component analysis depicted the effects of various conditions on the tempora.1 and spectral properties of /k/.

3.1. Effect of precending vowel on /k/:- It seemed that the closure duration of /k/increased when preceded by the long vowels /a:/, /i:/and */e/* and VOT increased when preceded py /a/. The transition durations were longer for the vowel /a/ and /i/ and maximum speed of transition was achieved in vowels /a/ and /a;/. F1 and F2 of /k/ was lowered when preceded by high back vowels /u/ and /o/ respectively. F2 was raised when preceded by front vowels and lowered when preceded by back vowels. The amplitude of the burst, was maximum when /k/ was preceded by */i/* and minimum when preceded by /e/.

		Noncluster condition	6emminate-cluster condition	Non-geminat cluster condition
Consonant durallon		164	309	215
Closure duration		137	281	165
vor		16	14	21
Burst duration		9		
Iransition duratxon Preceding	vowel			
	Fl	27		
	F2	25		
	F3	13		
Speed of transition Preceding (Hs/ms)	vowel			
	Fl			
	F2			
	F3	3		
Transition duration Following	vowel			
	Fl	14		
	F2	19		
	F3	*3*		
Speed of transition Following (Hz/ms)	vowel			
	Fl	8		
	F2			
	F3			
Preceding vowel duration	short		74	70
	long	160		

Table 1: Average durations temporal parameters of */k/ in* m.secs

	Preceding vowe1	Following vowel
Steady FC(Hz)	117	119
Termnal /in itial FO		116
FO dip (Hz)	0.89	3.1
Burst amplitude (dBR)	10	
Fl (Hz)		
F2 (Hz)		
F3 *(Hz)*		
81 (Hz)	364	
B2 (Hz)	341	
L1	S/	
L2	30	
Terminal F1 (Hz)		
Terminal **F2**: (Hz)		
Terminal F3 (Hz)		

Table 2: showing the values of various spectral parameters of /k/.

3.2. Effect of. following vowel on */k/:-* *T*he fallowing vowel seemed to affect the VOT of */k/.* VOT increased when /k/ was followed by /a/ and /us/. Transition durations of Fl and F2 were longer when the following vowel was /a/. Speed of transitions of Fl and F2 were maximum for the vowels /a/, /i/ snd /u/ respectively. The burst was weak when followed by the front vowel */e/* and absent when followed by /i/. Bursts at regions of F2 of the vowels /a/, /u/ and /o/ were observed. Burst amplitude was maximum when /k/ was followed by /a/.

3.3. Effect of following osonant (kvc condition) on /k/:- The following consonant mainly seemed to affect the burst duration. Burst duration, VOT and transition duration of Fl were maximum when the following consonant was the dentil aspirated /th/. Longer VOT in this condition might be attributed to the aspiration in /k/ which though unaspirated was uttered as aspirated. The amplitude of the burst increased when followed by palatals and decreased WM-.-n followed by nasals. Initial FO of the following vowels was high (128 Hz) when /k/ was followed by /e/ and low (93 Has when /k/ was followed by /m/. The following vowels exibited two FO patterns- raising when the following consonant was n,n,n,n,m,p,r,1 and s and fa11ing for the others.

3.4. Effect of clustering on /k/:- The following consonant affected closure duration and VOT. Closure duration has lengthened when followed by /s/ (220 ms) and shortened when fallowed by retroflexes /r/ and /s/ (160 ms), when compared to singleton /k/. The intrinsic: duration of /k/ in noncluster non-geminate condition appears to be 150.5 msecs and in geminate condition 345 msecs. VOT increased when /k/ was followed by /r/ and /1/. The following consonant did not seem to affect the spectral aspects of /k/.

3.5......Effect of embedding in a sentence:- There was no significant effect on either temporal or spectral parameters of /k/ when they were embedded in sentence.

It. appeared that the preceding vowel, following vowel fallowing consonant and cluster mainly affected the temporal parameters closure duration and VOT and the spectral parameters Fl, F2 and burst amplitude of /k/.

4.Discussion:- The result af the present study an /k/ duration is in consonance with that of Zue (1976) who reported singleton /k/ duration as 148 msecs. Liberman et al (1952), on the basis of speech production and perception data postulated that VOTs greater than 20 msecs tend to he perceived as voiceless whereas VOTs lesser than 20 msecs wil1 be perceived as voiced. In contrast, Klatt (1975) reported VOT range of 12-39 msecs for voiceless unaspirated plosive. Also, Klatt (1975) and Lisker and Abranison (1967) reported shorter VOTs for /p,t,k/ in post-vocalic and pre-unstrested position. In the present study, the average VOT value for /k/ was found to be 17 msecs which is in contrast with the notion of Liberman et al (1952), This might be attributed to language differpances. Whereas Liberman et al report it for English, the present, data is from Kannada language.

According to Klatt (1975), the average burst duration of /k/ was 37 maecs. However, Fischer-Jorgensen (1954), Halls, Hughes and Radley (1957) opine that the burst or the transient explosion produced by the shock excitation of the vocal tract upon the release of occlusion is usually less than 20 msecs. The average burst duration of /k/ in this study was 9 msecs which is in agreement with Fischer-Joroensen (1954) and Halle, Hughes and Radley (1957). Also. the double/multiple bursts reported to be the characteristic of velar place of articulation were noticed in this study.

Klatt (1975) and Raphael (1972) found an average duration of the vowel preceding a singleton voiceless stop in English to be 164 msecs. The results of the present study which indicates 160 msecs duration for the long vowel preceding */k/* is in par with the above findings.

Fundamental frequency dip (pitch fall) pattern has been found to differentiate voiced from voiceless stops. Voiced are reported to have a rising pitch pattern Kohler, 1985 Ohde, 1984). In this study, the preceding vowels exhibited a falling FO pattern and the following either a falling or rising F0 pattern Also. the bandwidth is reported to increase with the formants which was not observed. Instead, the bandwidth reduced with the formants.

The greater duration and speed of Fl transition in /a/ can be explained an the basis of the extent of movement the articulator need to perform. /a/ has a much higher Fi than /i/ or /u/ and hence the articulator can the presumed to take longer time in transiting from an open tract to complete closure as in the stop consonant.

Longer VOTs when preceded by a high vowel are reported by MacCawley (1968) and Klatt. (1975) and no change in VOT as a function of vowel condition was reported by Lisker and Abramson (1967). The results of the present study contradicts both these findings in that here VOT was found to increase when the preceding vowel was /a/.

Cooper e*t* al (1952) found the bursts located near the F2 regions of the following vowel which was also noticed in the present study. However, as the burst location changes depends on the following vowel, it may not be an invariant cue for the perception of stop consonants.

Many of the spectral and temporal parameters seem to be affected by the preceding and the following phonemes and some results seem to be specific to Kannada language, This warrants research on the other speech sounds in Kannada for possible application in perception of the hearing handicapped.

ACKNOWLEDGEMENTS:
This study is made possible by the grants received from the Deptt. of Science & Technology (SP/YS/L57/87).

BIBLIOGRAPHY

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. H. and Gerstman, L. J. (1952). "Some experiments on the perception of synthetic speech sounds". J, Acoust. Soc. Am. , 24, 597-606.

Fischer-Jorgensen, E. (1954). Cit. in "Stop consonant recognition: Release bursts rand formant transitions as functionally equivalent, context-dependent cues", by Dorman, M. F. , Studdert-Kennedy, M. and Raphael, L. J. (1977). Perception and Psychophysies, 22 (2), 109-122.

Fujimura, O. (1961). "Bilabial stop and nasal consoniants: A motion picture study and its acoustical implications", J.S.H.R., 4, 233-247.

Halle, M., Hughes, G. W. and Radley, J-P. A. (1957). Cit. in Zue, V, W. (1976). "Acoustic characteristics of stop consonants: A controlled study". Cambridge, Mass. M.I.T. Press.

Keating, P. (1979). "Dirterences in production and perception of VOX in Polish and English", *3.* Acoust. Soc, Am., 66, Suppl . No. 1, £387.

Klatt, D. H. (1975). "Voice onset time, frication and aspiration in word-initial consonant, clusters". J. S. H. R. 18, 686-706.

Kohler, K. J. (1985). "FO in the perception of lenis and fortis plosives". J. Acoust. Soc. Am., 78 (1), 21-32.

Lahiri, A. (1980). "Voicing contrast, in four category stop consonants", J. Acoust, Soc. Am., 67, Suppl., NO, 1, S51.

Laberman, A. M., Delattre, P. C. and Cooper, F. S. (1952). "The role of selected stimulus variables in the perception of the unvoiced stop consonants" , Am. J. Psychol., 55, 497-516.

Lisker, L.. and Abramson. A. S. (1967). "Some effe*cts* of context on voice onset time in English stops". Lang. Speech. 10, 1-28.

Ohde, R. N. (1984). "FO as an acoustic correlate of stop consonant voicing". J. Acoust. Sac. Am. 75 (1)

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

Stevens, K. N. (1975). Cit. in "Distribution of acoustic cues for stop consonant of articulation in VCV syllable", by Dorman, M. F. and Raphael, L. J. (1980), J. Acoust. Soc. Am., 67 (4), 1333-1335.

Zue, V. W. (1976). "Acoustic characteristics of stop consonants! A controlled study. Cambridge, Mass,. M. I. T. Press.

VOT AS VOICING,. CUE.. TO STOP. CONSONANTS IN. KANNADA

H.S.SRIDEVI

(SUBMITTED TO NSA CONFERENCE TO BE HƐLD IN NOVEMBER 1990)

In Speech, the many to one relationsship between stimulus and percept has two aspects; several phonetic contrasts can be produced by the same acoustic cue: conversely, several acoustic cues can produce the same phonetic contrast. In the latter aspect, the cues can be radically different. *Eg*. in stop consonants, voicing contrast can be cued by VGT (Lisker and Abramson, 1964, 1967), preceding vowel duration (Raphael 1972), Fl onset frequency (Summerfield and Haggard, 1977).

In the production of stop consonants, the oral closure release and onset of glottal pulsing define important articulatory aspects of voicing as they lead to bath temporal and spectral consequences. Of these constquonces, the interval from the burst to the first glottal pulse, which has been defined as the VOT of the stop consonants (Lisker and Abramson. 1964) has received maximum attention, This VOT characterizes the voicing contrast for initial step consonants in most languages (Lisker & Abramson, 1964).

Lisker & Abramson (1964) proposed that languages which have stop voicing contrasts, have chosen among thres VOT categories namely;

a) Voicing lead (-ve VOT *or* prevoicing) in which voicing onset precedes the release burst.

b) Coincident or short lag VOT (with *zero* or low +ve VOT values) in which voicing onset, is simultaneous or briefly 1ags behind the release burst.

c) Long lag VOT (with high +ve VOT values) in which the voicinq onset lags behind the release burst.

Systematic variation of VOT with respect to place of articulation (Keating et al. 1980; Lisker and Abramson, 1964) and following vowel (Klatt, 1975; Weisma.G, 1977, Summerfield, 1975) have been reported which is limited to Western literature; Lahiri (1980), Lisker & Abramson (1964), opine that the four category language present clear cut cases in which the measure of VOT is insufficient for distinguishing among all the stop categories of a language. However, Shukla (1990) reports that. VOT can effectively distinguish four category stops in kannada.

In the present study the role of VOT in cueing vpicinr contrast in Kannada which has 4 categories of stops. variation of VOT with respect of place of articulation and following vowel are evaluated.

Procedure: 248 meaningful words with 16 stop consonants. (k, kh, g, ah, t, th, d, dh, t, th, d, dh, p, ph,, b, ph) in the initial position precedinq 9 vowels (a, i, u, e, o, a, i, e, o) were considered. The words had either CVCV *or* CVCVCV Comibination. One young adult male kannada (speaker spoke these words embedding them *in* sentences (word and he:iti:ni) which were recorded *an* a high -fidelity magnetic spool with mic-to-mouth distance of 10 cms. Broad hand spectoorams of these words were analysed for VOT measures.

Results:—

	Maner			
Place of Articulation	VL Unasp	VL asp	VI) unasp	VD asp
Bi1abial	14	*70*	- 134	- 64
range	9--25	70	-83 -200	-SO -93
Dental	18	30	- 107	-55
range	9-28	30	-68 -133	-45 -60
Retroflex:		70	- 108	- 55
range	2 0--5	0 - 100	--92 -123	-5O - 50
Velar	31	83	-- 112	48
rancje	19-58	43-11.5	--95 -129	-5 -90
Average	16	63	- 115	

Table Is Representing averages and ranae of VOT values for each place of articulation across 4 categories *of* stops.

Kannada has four categories of stops, voiceless unaspirate voice1ess aspirated voiced usnaspirated and voiced aspirated. It seems that kannada contrasts voicing lead with voicing lag. Both aspirated and unaspirated voiced stops have voicing lead while voiceless aspirated has a long lag VOT and voiceless uridspirsted have a short lag VOT. Average values Of VOT for each category is clearly seperated. especialy in the voiced category.

Lisker and Abramson (1964) report that though VOT a capable of distinniushing various cateaoruis of stops, it fails to distinguish between voiced unaspirated and aspirated stops. Lahiri (1980) reported that the feature interrupted voicing differentiates voiced aspirates from the other thres categories, wherein there is prevoicing followed by approximately 100 msec of silence *and* then resumed phonation. Present findings do not lend support to these reports. The voiced aspirates had a significantly smaller VOT values compared to their voiced unaspirated counterparts. ALSO, prevoicinq followed by a short interval of silence (3-30 Msec) was found in both voiced aspirates and unaspirates. It has been stressed in the earlier studies that VOT of greater 20 msec will be perceived as voiceless and VOTs up to 20 msec will be perceived as voiced. However, the average VOT and VOT for different places of articulation of voiceless unaspirated stops in the present study was less than 20 msec except for velars. It has been reported that VOTs for /p,t,k/ *are* much shorter in post vocalic and preunstressed conditions (Klatt, 1975; Lisker and Abramson, 1967), the basis being that unstressed syllables can be produced with lest articulatory effort. This can be extended to the voiceless unaspirated stops wherein the glottal aperture is somewhat greater than in voiceless aspirated conditions.

It has been constantly reported that VOT following release of stop increases as the stop place of articulation moves further back in the oral cavitv i.e, VOT is somewhat greater in velars than alveolars and in alveolars than labials (Borders & Harris, 1980s Keating et al. 1980; Lisker and Abramson, 1964). The basis for this is that VOT varies inversely with the rate at. which oral release gesture is made (Summerfield and Haggard, 1977). The duration of the movement of the articulation that forms the closure is greatest for the tongue body, less for the tongue lip and least for the lips (Pant, 1960; Stevens and Klatt, 1974). The present study partly agrees with the earlier findinqs in that voiceless velars were found to have greater VOT compared to voiceless stops in other places of articulation. However, the relation velar alvealar bilateral was not found in this study. Among the voiced stops, bilabials were found to have oreater V0T compared to other voiced counterparts. VOT does not serve to be a cue to place of articulation at least in Kannada. Place of articulation may be cued by other consequences of VOT such as rate of Fl transition. F1 onset frequency or F2 transition. This calls for further investigation of VOT in relation to these features.

Change in VQT as a function of following vowel:

Klatt (1975), Macawley (1968), Summerfield (1975a) have reported an increase in VOT of a stop when it preceded a high vowel. The explanation offered for this is that high vowels influence the behaviour of the larynx such that the laryngeal fundamental frequency is higher and voicing is less easy to initiate or sustain than in other vowels.

Table 2s Representing average VOT values of stops as a function of vowels:

	a	i	u	e	a	a	i		
VL Unasp	16	20	21	24	12	11	34	19	
VL Asp	69	0	58	45		101	115	83	
VD Unasp	-102	-97	-134	-127	-114	-109	-127	-130	
VD ASP	- 58	-55	- 93		— —	55	- 50	- 73	-5

The result of this revealed that VOT was more for all categories except voiced aspirates while preceding high front vowel /i/ Also, VOT was higher for both categories of Voiced stops when they preceded /e/. However, no clear vowel conditioned effects on VOT could be seen in the present study which agres with the similar findings of Lisker and Abramson (1967). However, an analysis of VOT in relation laryngeal F0 is warranted before any conclusion is reached.

In general, the findings of the present, study can be summarized as (1) VOT can distinguish all the 4 categories of stop consonants in Kannada. (2) Kannada uses voicing lead to contrast voiced stops against voiceless stops which have either short or long a VOT lag. (3) VOT does not cue place of articulation in Kannada. (4) There is no vowel conditioned effects of vowels on stop consonants in kannada.

Discussions The results reveal that VQT does distinguish all four categories of stop consonants in Kannada. Though it seems to be a simple measure of onset of glottal vibration in time relative to the release of the stop closure, son . observations do require explanation.

Firstly, VOT has been reported to increase as the place of articulation moves back in the vocal tract (Borden and Harris, 1980s Keating et al., 1980; Lisker and Abramson, 1964). However, the results of this study do not support this. Though the voiceless velars had maximum VOT, the VOT at, other places of articulation was not linear with movement.

of place of articulation from lips to velum. Moreover, the distribution of VOTs for voiceless aspirated and unaspirated at the same place of articulation was not the same. Though the larger VOT values for velar place of articulation can he attributed for the involvement of tongue mass for articulation, VOTs for other places do not yield for such description. Among the voiced stops, except for the finding that both aspirated and unaspirated bilabials had maximum VOT, no systematic variation of VOT across place of articulation was found. Voiced dentals and retrof1exes had similar VOT values. Though larger area behind the occlusion can be attributed to larger VOT values in voiced bilabials, such generalizations cannot be drawn for other places of articulation.

Secondly, voicing was not present for the entire closure period for voiced stops. Cessation of voicing was found 3 to 30 msec prior to stop release. Similar tokens were rejected in earlier studies as being faulty. However, studies on aerogvnamics of stop consonants (Muller and Brown, 1980) reveal that cessation of voicing in voiced stops occures when the subglottic pressure equals that of the supraglottic pressure. The equalization of pressure can be brought about by several mechanisms which may act independently or collectively. The mecahnisras which either facilitate voicing or devoieing are of two types (a) glottal and supraglottal articulatory adjustments which may either sustain or diminish the aerodynamic driving force which maintains the vocal fold vibration and (b) internal laryngeal adjustments which affect the biomechanical dynamics of the folds thereby mode, frequency of vibration and their susceptibility to oscillation. Thus cessation of voicing in voiced stops has physiological basis. If Lisker and Abramson's criteria of VOT is considered in such cases where there is cessation of voicing, then onset of voicing should be considered after the release and the negative VOT should not be considered in such cases. Thirdly, voiced aspirates had significantly lower VOT values than their voiced unaspirated counterparts at all places of articulation. Though no explanation can be offered for this finding at this juncture with only the VOT values, measures of burst, aspiration and murmur may shed more light on this aspect.

The four category stops thus seem to be different from the two or three category stops. Detailed acoustical analysis of four category stops involving simultaneous measurements of various parameters (eg. durations of closure, voicing, burst, aspiration and murmur) will not only shed more light on production of these stops but also carries interest to their perception.

ACKNOWLEDGEMENTS: This study was made possible by the grants received from the DST (NO. SP/YS/L37/87).

REFERENCE:

1. Border), G.J and Harris, K,S (1980) "Speech Science Frimer - Physioloqy, acoustics and Perception of Speech", Williams and Wi1kins Company, U.S.A.

2. Fant, G. (1960) "Acoustic theory of Speech production", Monton, The Hague.

3. Fitch, H.L, Maleves,T., Erickson, D.A and Liberman, A.M. (1980) "Perceptual equivalence of two acoustic cues for step consonant manner, perception and Psychophysics", 27 (4) 343 350.

4. Keatinq, P.A, Westbury, J.S, and Stevens, K.N (1980, "Mechanisms of stop consonant release for different, places of articulation" J. Acoustical Society AM., 67, Suppl.No.l, 893

5. Klatt, D.H (1975). "Voice onst time, frication and aspiration in word initial consonant clusters, J.S.H.R, 18, 686-706

6. Lahiri.A.(1980) "voicing contrast in 4 category stop consonants", J.Acoust.Society,Am.,67, Suppl.No.l, 851

7. Liberman.A.M, Cooper, F.S., Shankineiler,D. and Studdert Kennedy, M. (1967) "Perception of the speech code', Psychological review, 84, 452 - 471

8. Lisker, L and Abramson, A (1964) " A cross study of voicing in initial stops: acoustical measurements, word, 20 (3), 384 - 422.

9. Lisker.L and Abramson, A. (1967) "some effects of context on VOT in English stops" Language and speech, 1.0, Part 1, 1 - 28.

10. Muller, E. M. and Brown, W. S. (1980) "Variations in the subglottal air pressure waveform and their articulatory interpretation" Speech and Language, Advances in basic research and practice, Vol 4, 317-387.

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

12. Shukla.R.S, (1990), "Closed mouth technique -Unpublished article presented at ISHA, 1990.

13. Sati.S.D (1983), "The role of spectral cues in discrimination of VOT differences" J.Acoust.Soc.Am.,73 (6), 2150-2165.

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

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

16. Weismer ,G (1977) "Some context effects on voice onset time". J, Acoust. Soc, Am., 62, Suppl.No.1, 878, HH10.

PERCEPTUAL CUE'S OF STOP CONSONANTS IN KANNADA

S.R. Savithri . (Ph.D. Speech & Hearing)
H. B. Sridevi . (M.Sc Speech & Hearing)
Deptt. of Speech Sciences
All India Institute of Speech & Hearing
Manasaqagothri , Mysore- 570 006. India

Received:

(submitted to journal of the acoustical soceity of America)

It is commonly believed that the perceptual representation of Speech sounds is organized atleast in part, in terms of intersecting properties or dimensions Several investigators (Reeds and Wang, 1961 Raphael, .1972, Winitz, Schieb and Reeds, 1972, Stevens and Klatt, 1974, Darwin and Brady, 1975, Lisker, 1975, Blumstein and Stevens, 1976, Mas1in and John, 1976, Siuddert Kennedy and Rapheal , 1976. Dieh1, 1977, Dorman and Raphael, 1977, Summerfield, and Haggard, 1977, Stevens. 1977, Keating and Blumstein, 1978, Ohala, 1978, Ohde, 1978. Dorman, Raphael, and Liberman, 1979a, Fisher-JOrgensen, 1979, Gruenfelder, 1979, Lisker and Price, 1979, Price and Lisker, 1979, Ahmed and Gupta, 1980, Bailey and Summerfield. 1980, Fitch, Hallwes. Ericksan and Liberman, 198o. Port, 1980, Raphael, 1980, Walsh and Parker, 1980, Keating, Mikos and Ganong , 1981, Ohde, 1982, Repp, 1974, Chasaide and Gobl, 1987, Walsh and Diehl, 1987, Zakia and Kingsten, 1987, Datta,1989, LJsha, 1989, Vinay, 1990) have attempted to determine the nature and number of the perceptual dimensions or features that listeners employ in the identification of Speech sounds. Most of the studies *are* limited to the use of isolated monosyllables or disyllables. However, inspite of the extensive research, human Speech perception remains unanswered. Further, the use of voice prints in Forensic medicine has become more difficult because of the capacity of the human being to imitate, " Imitation is a remarkable skill that requires one to parse the behaviour of another in to components and then activate his own corresponding motor controls to reproduce the behaviour " (Studdert- Kennedy, 1981). This artcle adresses at. the perceptual cues of Kannada stop consonants used in imitation.

Method:

Material: 467 meaningful Kannada words with stop consonants (k . kh . g . gh , t , th , d , dh , t, t.h , d , dh, p ,ph, b , bh) intial and madial positions were selected for the study with all passible vowel and consonant. environments. These words as embedded in sentences 'word anta he:ltini lurmed the test material.

Subjects: One 17 year old Kannada speaking male served as the mode1 subject and a 26 year old Kannada speaking male (henceforth SI) served as imitator. He was selected on the basis of his performance in mimicry.

Recordings Kannada words were visually presented (as written on cards) one at a time and the model subject was instructed to utter them embedding in sentence in a natural manner into the microphone (cardiode-unidirectianal kept at a distance of 10 cms from the mouth. All these were recorded on high fidelity audio-tapes. These sentence were further transferred to audio-cassettes and were given to the imitator for practice. Once the imitator felt confident about his proficiency, he was audio-presented with the sentences uttered by the model through earphones and was instructed to imitate the sentences with special attention to the stop consonants. These were also recorded on high-fideliy audio-tapes. The key words containing the stop consonants of all the subjects were extracted from the sentence an stored in an audio-casette. The words were further digitally stored in magnetic diskettes using a PC/XT with a 12 bit A/D and D/A converter at a sampling rate of 8OOOHz.

Acoustic analysis: Stop consonants in the words which were transformed digitally were computer analysed using LPC autocorrelation method to extract F ,F , B1,B2,B3,L1,L2,L3." Further, using the sound spectrograph (VII 700), wideband bar-type and average amplitude type of spectrograms were obtained for all the words with stop consonants and 24 parameters were measured using these spectrograms, viz., Closure duration(C.D), VOT, Burst
duration (B.D), Aspiration duration(A.D), Voicing duration (v.D , Marmur duration(M.D), Stop-consonant duration, Word duration(W.D), Transition duration of Fl , F2, F3 of the preceding and the following vowel (TDF1, TDF2, TDF3) , Speed of transition of Fl , F2, F3 of the preceding and the following vowel (STF1, STF2„ STF3) , Burst amplitude. Amplitude of the following vowel and consonant., and Overall amplitude. Totally 33 parameters were extracted for each consonant, and these parameters as used by the model were compared with those of the imitator.

Perceptual analysis: The words as uttered by the model and the imitator were transferred in order on to an audio-casette and a booklet was provided to two Kannada speakers (one mals and one female) for perceptual evaluation. They were instructed to judge the imitation ay very good/ fair/ poor/ bad and indicate the cues for their judgments The perceptual judgements thus obtained were correlated with the acoustic parameters imitated to arrive at the perceptual cues of Kannada stops.

Definition of the parameters used:
F , F , F (Hz): Frequency of the first second and third formants as obtained through acoustic analysis by computer.
Bl, B2, B3: Bandwidths of the Fl, F2 and F3 respectively as obtained through acaustic ana1ysis by camputer.
L1, L2 L3: Levels of the F1, F2 and F3 respectively as obtained through acoustic analysis by computer.

The following parameters were measured from the wide-band bar and average amplitude spectrograms:
Closure duration (msecs): Duration between the offset of resonance for the preceding vowel and the onset of burst for the stop consonant in intervocalic condition and duration from the regular vertical striation on the baseline to the onset of burst for the voiced stop in the initial position. Burst duration (msecs): Duration of the vertical irregular striations depicting the art.icul atory release.
Aspiration duration (msecs): Duration of the low frequency energy measured as the time between the onset of irregular vertical striation in the low freauency region to the end of the same between the burst and the onset of resonances for the following vowel.
Voicing duration (msecs) : The time between the onset and offset of regular vertical striations within the closure duration.
VOT (msecs); The duration between the burst and the onset of the resonances for the following vowel.
Murmur duration (msecs): The duration of partial voicing depicted as regular vertical striations superimposed by irregular vertical striations on the spectrogram between the burst and the onset of resonances for the following vowel.

Transition duration of F1, F2 and F3 of the vowel (msecs): Duration between the onset of resonances changing to the onset of steady resonances for the following vowel and the duration between the onset/offset of resonances shifting for the preceding vowal.
Speed of transition (Hz/msec): The ratio erf the frequenty shift during transition to the duration of transition.
Stop duration (msecs) : The duration including the closure, burst and aspiration.
Word duration (imsecs): The time between the onset*/offset* of regular vertical striations on the baseline/resonance for a word Btartino with voiced stop/vowel and as the time between the burst ancf the offset of regular vertical striations on the baseline/resonance for a word starting with voiceless stop).
Burst amplitude (R dB) : Peak amplitude at the region of the burst as on the average amplitude spectrogram.
Overall amplitude (R dB): Peak amplitude (except that of the burst) in the region of ths stop as *on* the average amplitude spectrogram.
Amplitude of the following vowel/consonsnt (R *dB): Peak* amplitude within the duration of vowel /consonant as on the average amplitude spectrogram.

Results:
Acoustic analysis: All the temporal and spectral parameters were measured and are presented in table 1. Visual inspection of the spectrograms (fig.l) revealed that SI imitated the articulatory releases of the model for stops which was evident from strong bursts. Also, as in the model, the imitator showed multiple bursts and, of tmi stops, the aspirated and murmered were the best imitated.

INSERT TABLE 1, FIGURE 1 HERE

On a T-test, among the temporal measures, it seemed that SI used transition duration and speed of transition of Fi, F2 an*d* F3 as the major cues (Table.2). However the temporal parameters were also used though not to a great extent. On the T-test, significant differences were noticed in the closure duration of /g,t,b/, voicing duration *of* /g,bb/, burst duration of /k,g,t,t/, aspiration duration of /kh,bh/,murmer duration of /gh/, VOT of /k,kh,g,th,t & bh/. Word durations of SI were longer in all the instances. Of the spectral parameters, L2, B2, burst amplitude and B3 were the features imitated. However, as the word durations of SI were longer, a scaling of various temporal parameters (closure duration, burst duration, VOT, aspiration duration, murmer duration and voicing duration) was performed as a percent of the duration of the key Bound. Also, the distances between F F and F F were calculated. The results indicated that SI used all the temporal measures as cues in a scaled way (Fig.2) , F F was used 20% of the time and F F 80% of the time (Table 3, Fig, 3) .

INSERT TABLE 2,3, FIGURE 2,3 HERE

On a grading of the temporal and spectral features (table 4), it seemed that the relative timings of the various parameters of the stop consonant, speed at transition and transition durations, burst amplitude, B2, B3 and L2 were maintained by SI for more than 75% of the time. However, it was apparent that the formant structure *in* the speech of the model and that of the imitator, in general, did not agree.

INSERT TABLE 4 HERE

This is in agreement with Endress, Bombach & Flosser (1971) who studied the impersoanations of five public figures by two well known German imitators. The concluded that the imitators succeeded in varying the formant structure and fundamental frequency *of their* voices, but they were not able to adopt these parameters to match or even be similar to those of the imitated persons.

Also, the same opinion was held by several researchers including Kersta (1962 b), who disclosed spectrograms of a Speech made by President Kennedy and copied by his imitator, Elliot. Reed. There were obvious differences in the spectrograms despite the -fact that, according to Kersta, the voices sounded extremely similar.

Perceptual analysis: Of the 467 imitations, 219 were judged to be very good (imitated sample is exactly like the model), 208 to be fair (imitated sample resembles the model but not exactly), 40 to be poor (imitated sample resembles the model to a minimal extent) and zero as bad (imitated sample completely deviates from the model). The perceptual features used to judge the imitation as very good were similarity in pitch, syllable duration, inflection and rate. Whenever, the imitation changed in pitch/ syllable duraton/ inflection/ rate/ stress/ aspiration, it was judged 'fair'. In addition to this, any nasalization or substitution was judged to be a poor imitation.

Discussions Within the limitations of this experiment a few general points mav be made. First of all, the imitator used the features transition duration of Fl, F2 & F3 and speed of transition of Fl, F2, F3 and scaled temporal parameters as cues. Word durations were sign i f i cant 1 y longer in SI and SI used *deep* inspirations before uttering each sentence, both of which suggest increased effort in imitation. This suggests that, inspite of the voices sounding extremely similar, greater effort. is required to imitate, thus lengthening the utterances, which would be a point of consideration in voice print forensic applications.

In spite of the absence of similarities in the acoustic features imitated, the imitators voice sounded like the voice of the imitated. Further, similarities in pitch, syllable duration, inflection and rate (suprasegmentals) *are* considered to judge the imitation. The imitator, who happened to be a speech pathologist and linguist, reported that he had to lengthen the vocal tract, open the velopharyngeal port a little, lower the larynx and had to have a firm contact of the articulator for imitating the stop consonants. However, this was not effective in bringing the formants closer to that of the imitator. Thus, it seems " t hat the relative movement of Fo in a word and the articulators precisions are considered to judge the similarity of two voices rather than the indivual acoustic features.

Second, it is well known that neither two speakers can have some temporal features nor a single speaker can maintain same temporal feature for the same word uttered twice. This suggests that, the perceptual mechanism may do a scaling, wherein the relative? durations of the speech sounds *are* considered rather than the absolute durations. However, no comment can he made about the spectral characteristics, as neither the formant frequencies nor the F F, F F of the model and SI coincided.

Acknowledgements: This study was made possible by the grants received from the Deptt. of Science and Technology (grant no. SP/YS/L57/87).

BIBLIOGRAPHY:

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

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

Blumstein..S. E. and Stevens, K.N. (1976), "Perceptual invariance and onset spectra for[1] stop cansonants in different vowel environments", J. Acoust. Soc. Am., 60, Suppl.No.1, 590

Chasaide, A.N. and Gobl,C, (1987), "Cross language study of the effects of voiced-voiceless consonants on the vowel voice source characteristics", J. Acoust. Soc. Am, 82, Suppl.No.1 S 116.

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

Diehl.R.L, (1977), "Phonetic context effects on the identification of stop consonants", J, Acoust. Soc;» Am. 62, Suppl.No.i, S 78.

Dorman, M.F. and Raphael , L.J (1977), "onset spectra and stop consonant recognition, J. Ac oust. Soc. Am., 62, Suppl.No.i, S 78

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

Endress,W., Bambach, W., and Flosser,G., (1971) "Voice spectrograms as function of age, voice disguise, and voice imitation", J. of Acoust. Soc.Am,49, 1842-1848 .

Fischer-Jorgensen, E, (1979) "Temporal relations in consonants based on Danish material", in "Frontiers of speech communication Research", Lindblom, B. and 0hman,S. (Eds) Academic Press, London, 1979, Fitch, H.L, Hallwes, T, Eriekson, D.M, and Liberman, A.M. (1980) "Perceptual equivalence of two acoustic cues *for* stop consonant manner". Perception and Psychophysics, 27(4) : 345-350,

Gruenfelder, T.M. (1979), "Fundamental -frequency as a cue to post vocalic consonantal voicing: some data from perception *and* production", J. Acoust. Soc. Am., 72, Suppl.No.1, 58.

Keating F.A, and Blurmstein, 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 .

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

Lisker *L,* and Price, V. (1979), "Context determined effects of varying closure duration", J. Acoust. Soc. Am.65, Suppl.No.l, 6 7,

Kersta,L.G, (1962 b) cit. in voice identification, Theory and legal applications by Tosi , O, University Park Press, Baltimore, 1979

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

Ohala, M. (1978). "Acoustic and aerodynamic correlates of Hindi stops", *J.* Acoust, Soc. Am., 64, Suppl.No.l S 92.

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 , Supp1 . No. 1 , S 19.

Ohde, R.N. (1982), "The effects of linguistic content on temporal and Fo properties of speech"., J. Acoust. Soc. Am., 72, Suppl. No.l, S 65

Port, R.F. (1980) , "V/C ratio as a post, vocalic voicing cue", J . Acoust. Soc:. Am. 67, Suppl . No. 1 , 551.

Price, P., and Lisker , L. (1979) "(/b/-/p/ but -+(/p/--/b/) , J. Acoust.Soc.Am. 65, Suppl No.1, S 7—8.

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): 1296-1303.

Raphael.L.J. (1980), "Contribution of CV transition duration to the perception of final consonant voicing in natural speech" , J. Acoust. Soc. Am. 67, Suppl No. 1 , S 51.

Reeds, J.A. and Wang, W.S.Y. (1961), " The perception of stops offer "s", Phonetics, 6, 78-81.

Repp, B.H, (1984), "Categorical perceptions Issues, methods, findings" in Speech and Languages Advances in basic research and Practice. Vol. 10, Lass, N.J. (eds.) Academic Press, Inc., U.S.A, 1984.

Stevens, K.N. (1977), "Onset sspeetra ascues for cgnspnantal place of articulation", J. Acoust. Soc. Am., 61, 548

Stevens, K.N. and Klatt, 0.H., (1974), "Role of formant transition in the voiced voiceless distinction for stops", J. Acoust. Am. 55 (3),653 -659

Studdert- Kennedy, M. , (1981) ," Perceiving phonetic, segments", in The cognitive representation of speech, Myers, I., Laver,J., and Anderson,J.,(eds.), North Holland Publishing company.

Summerfield, Q and Haggard,M. (1977), "On the dissociation af spectral and temporal cues to the voicing distinction in initial stop consonants" , J . Acoust.. Soc . Am. Vol . 62 (2) , 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, (1990), "Crass language study of stop perception", Unpublished Master's Dissertation submitted to the University of Mysore.

Walsh, T. and Diehl, R.L. (1987)/ " Auditory factors in the perception of stops and glides", J. Acoust. Soc, Am., 02, Suppl , No. 1, S 80.

Walsh, T. and Parker, F. (1980), "Vocal fry: A cue for voicing in post vocalic stops", J. Acoust. Soc. Am., 82, Suppl. No. 1, S51.

Winitz, H. , Scheib, M.E, and Reeds, J.A. (1972), "Identification of stops and vowels for the burst portion of /p t k/ isolated from conversational speech", J. Acoust. Soc. Am., 51, 1309 - 1317.

Zakia, R.A.E. and Kingston, J. (1987), " The role of rate of transition in place perception" J. Acoust. Sac. Am.,82, Suppl. No, 1, S 81.

B.D		B.D	C.D	V.D	B.D	VOT	TDF1		TDF2			TDF3			STF1		STF2			STF3				A.D		
M	SI	M	SI M	SI	M	SI M	SI	M SI	M	SI M	SI		M	SI	M		SI		M	SI			M	SI	M	SI
K	19	12	612	646					2	14	15	29	16	20	30	26	8	12	3	3	4	2	2	2		
	21	12	582	602					3	19	16	31	11	17	30	18	8	18	3	2	4	2	2	2		
	16	14	661	680					53	83	53	83	17	14	18	12	6	9	6	2	6	2	2	1		
Kh	16	14	658	670					56	81	56	84	18	20	21	11	6	12	4	2	1	1	1	1		
	10	5	624	711	73	112	73	112	.1	0	.1	0	34	21	33	33	13	19	4	4	4	3	2	1		
	10	6	596	673	72	114	72	114	.5	0	.5	6	31	21	29	31	13	20	5	5	5	3	2	1		
gh	15	16	585	753	45	48	45	28	9	28	10	27	18	18	20	21	10	5	5	4	6	1	3	.5	66	23
	18	4	713	764	70	53	65	50	13	15	14	15	15	18	12	25	8	3	5	3	4	2	3	.3	53	17
	3	7	613	707									31	16	55	16			5	5	6	2				
t	2	4	602	707									23	11	58	11			3	6	5	1				
	2	6	731	770					50	74	43	57	26	17	10	17	4	14	2	3	3	2	1	3	5	11
	2	7	666	734					63	95	63	80	32	12	10	8			2	2	4	1			5	10
th			659	778	77	108	77	99					31	13	40	48			5	4	6	5				
	.3	.5	650	758	85	90	85	82					15	10	34	36			4	3	4	1				
			662	745	**2**8	55	28	48																	122	130
d	2	5	599	695					5	5	5	19	14	9	33	33	12	17	2	3	5	3	2	2		
	2	5	615	726					5	4	6	19	15	22	22	9	9	18	2	3	3	3	2	3		
	1	2	730	787	102	111	102	101					18	11	55	25	28	18	3	3	6	5	4	3		
D	1	2	329	773	99	98	99	88					35	14	55	28	14	14	3	4	5	5	2	3		
	2	6	627	671							10	14			37	32	25	15			5	4	4	6		
	1	7	638	664							9	11	15	13	29	36	20	8	3	4	4	4	3	2		
B	2	1	716	767	103	134	103	122					12	13	41	36	23	20	2	3	6	4	4	3		
	2	1	717	750	109	136	109	113					15	13	38	26	24	17	2	3	5	4	4	3		
	2	1	708	742					84	69	84	70	33	28	32	15	14	24	3	1	4	4	2	2	17	16
Ph	1	1	738	774					68	62	62	65	29	34	41	22	13	27	2	1	3	4	2	2	34	25
	1	1	741	736	56	60	52	58	23	1	1	1	2	16	9	9	10	14	0	2	1	1	1	2	81	53
bh	1	1	741	741	53	64	48	70	70	3	3	3	29	30	18	17	9	12	2	4	1	1	1	1	72	68

F1		F2		F3		B1		B2		B3		L1		L2		L3		
M	SI	M	SI	M	SI	M	SI	M	SI	M	SI	M	SI	M	SI	M	SI	
K	413	434	1217	1291	1963	2170	752	794	677	744	689	656	135	131	129	124	34	33
Kh	441	429	1184	1277	1885	2114	730	790	625	785	757	665	139	132	138	113	51	29
G	410	438	1162	1215	1858	2093	731	755	650	680	594	649	143	138	134	136	39	28
Gh	465	444	1194	1265	1879	2072	815	805	676	732	601	673	149	134	149	127	48	28
T	496	459	1267	1553	1447	1993	278	432	582	556	437	502	134	152	170	124	53	26
Th	460	510	1512	1624	1664	1876	237	189	675	272	654	333	129	162	228	123	48	72
D	390	442	1502	1469	1680	1588	229	313	618	409	550	739	103	145	93	139	30	45
Dh	417	470	1509	1530	1670	1865	286	412	720	450	370	350	145	145	104	139	26	61
T	439	482	1405	1487	2629	2872	769	838	785	808	709	735	127	171	50	58	9	12
Th	473	491	1413	1360	2468	2939	752	692	728	628	711	814	124	125	47	68	8	14
D	442	482	1384	1445	2713	2792	769	725	761	661	767	658	120	150	46	49	7	9
Dh	429	459	1290	1376	2947	2645	823	893	776	885	856	747	119	126	48	49	5	9
P	363	466	1095	1292	1882	2121	696	796	680	669	765	625	123	124	142	121	17	31
Ph	380	454	1174	1304	1976	2075	707	694	697	653	678	615	615	139	132	123	117	37
B	360	442	1103	1248	2039	2041	690	820	706	740	718	709	125	126	129	136	14	29
bh	390	440	1161	1311	2070	2087	767	846	772	775	728	723	126	132	157	134	20	40

Table 1: Temporal and Spectral parameters of the model and the imitator.

Features	k	kh	g	gh	t	th	d	t	d	p	ph	B	bh
Closure duration			++	--	+		-+		-+			++	--
Voicing duration			++	--			--		--			--	+-
Burst duration	++	--	++	++	+	--	--	++	--	-+	-	--	--
Aspiration duration	++	++	+-	--	-	--	--	--	--	--	-	--	+-
Murmur duration				++	+	--	--	--	--	--	-	--	--
VOT	++	++	+-	--	-	--	+-	++	--	--	-	--	+-
Word duration	--	-+	++	--	+	--	+-	++	-+	++	-	--	--
Transition duration of F1	--	--	++	--	-	--	+-	--	-+	--	-	-+	--
Transition duration of F2	-+	--	--	--	+	--	-+	--	++	--	-	--	--
Transition duration of F3	--	--	--	--	-	--	--	--	--	--	-	+-	--
Transition duration of F1	--	--	--	--	-	--	--	--	--	--	-	--	--
Transition duration of F2	++	--	--	--	+	--	--	--	--	--	-	--	--
Transition duration of F3	--	--	--	--	-	--	--	--	--	--	-	--	--

Initial +/- Presence/ absence of significant difference in following vowel Condition.

Final +/- Presence/absence of significant difference in following consonant Condition.

Features		K	Kh	G	Gh	D	T	D	Dh	P	Ph	B	bh
F1	-+	--	++	--	++	++	-+	-+	-	++	+	++	+-
F2	--	--	++	++	--	--	--	-+	+	++	-	++	++
F3	++	++	++	++	--	--	++	--	-	++	-	-+	--
B1	--	+-	--	--	-+	-+	--	--	-	++	-	++	+-
B2	--	--	+-	--	+-	+-	--	--	-	--	-	--	--
B3	--	--	+-	-+	--	--	--	+-	+	-+	-	--	--
L1	--	--	--	++	-+	-+	--	-+	+	++	-	+-	+-
L2	--	--	--	--	--	--	--	--	-	+-	-	--	--
L3	--	++	-+	++	--	--	--	-+	+	-+	-	++	++
Burst amplitude	--	--	--	++	--	--	--	++	-	--	-	--	--
Overall amplitude	++	--	+-	++	++	++	--	++	+	--	-	--	++

Amplitude of following Vowel.	++	+-	++	++	-+	-+	--	--	-	++	+	+-	--
Amplitude of following Consonant.	++	++	--	--	-+	-+	++	++	+	--	-	--	--
Initial +/- Presence/absence of significant difference in following vowel Condition. Final+/- Presence/absence of significant difference in following consonant Condition.													

Table 2: Significant difference in temporal and spectral parameters of stop consor model the imitator.

Phonernes	F1	F2 (level)	F2	
k	+	(.004)	+	,012
kh	+-	(.004)	-	.31
9	+.	(.O55)	-	.004
gh	+	(.062)	+	.062
d	-	(062)	+	. 062
t	+	(.043)	-	. 027
d	+	(.043)	-	.048
P	-	(0.012)	-	.045
b	+	(.022)	+	. 043
bh	+	,062	-	.062

Table.3: Significant difference between Model *it* SI.

\+ Sig. cliff. Present - Sig. diff. Absent.

Parameter	Percent
Following vowel transition duration F3	100
Speed of transition F1	100
Relative CD	100
Relative VOT	100
Relative B.D	100
Relative V.D	100
Relative As.D	100
Relative M.D	100
L2	96
Speed of transition F3	92
B2	91
Speed of transition F2	838
Transition duration Fl	83
Burst amp1itude	82
B3	77
Bl	68
F ~ F	60
L1	59
Amplitude of the following consonant	59
Word duration	54
Overall amplitude	46
L3	46
F	46
F	41
F	41
F ~ F	40
Amplitude of the following vowel	37

Table 4: Parameters imitated percent times.

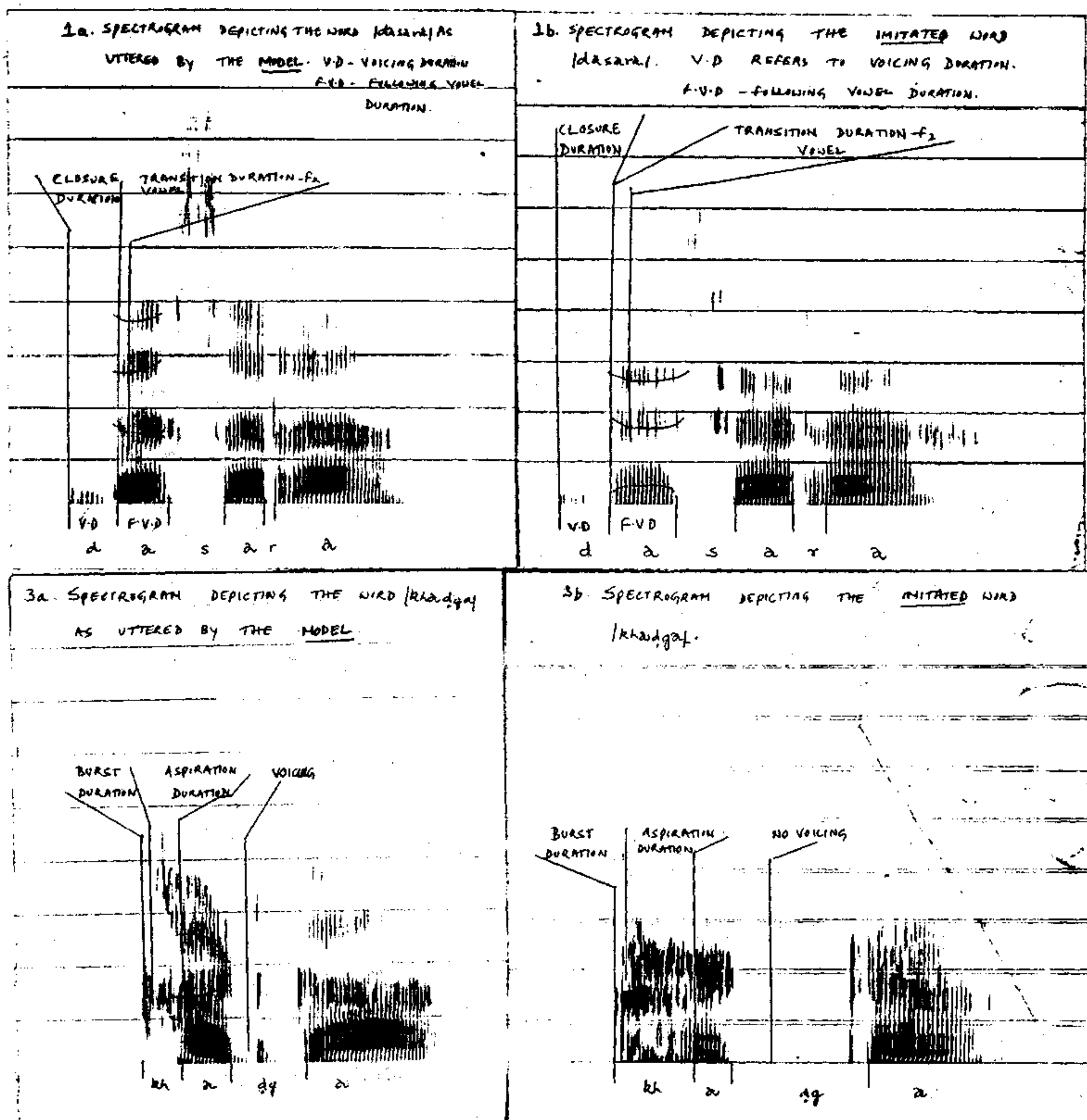
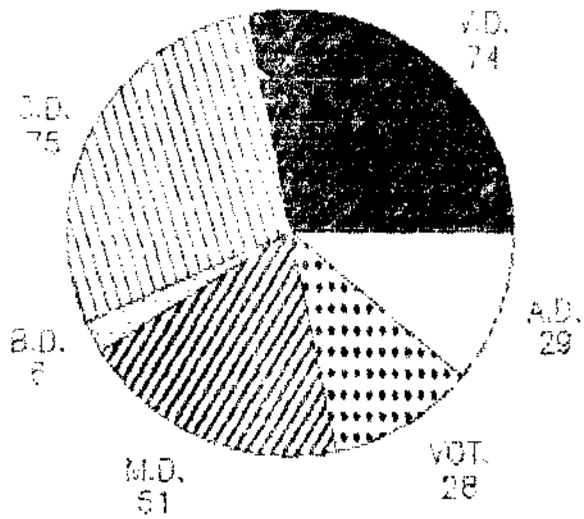
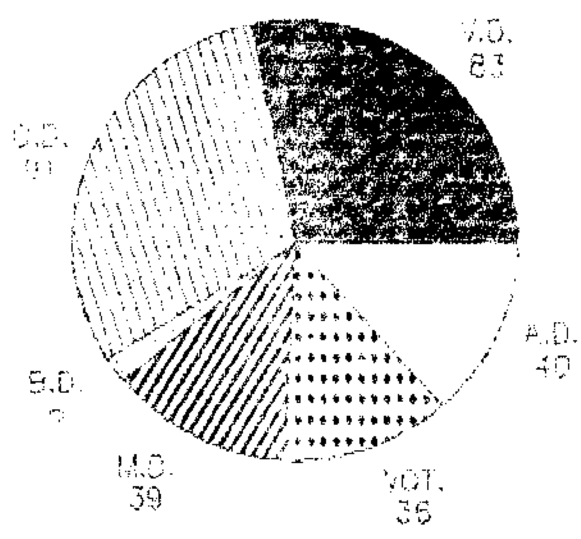


Fig 1: Spectrograms of model and imitator.

I *TO (b)*



(Model)



(Immitator)

Fig.2 Scaled temporal parameters of model and immitator.

Jour. Acoust. Soc, Ind.
Vol. XVII (3&4), December, 1989

PERCEPTUAL CUES OF STOP CONSONANTS IN KANNADA

U.S. Sridevi *&. S*.R.Savithri. Deptt. of Speech sciences. All India Institute of Speech & Hearing, Manasa gangothri. Mysore- 570 006.

This study evaluates the perceptual cues of Kannada stop consonants using an imitation paradigm. The speech of a model *&* An imitator was recorded & was subjected to spectrographic & computer analysis. The duration & speed of transition, closure & aspiration duration, B2, B3, L2 seemed to be the perceptual cues used for imitation.

I.INTRODUCTION:

There are grounds for believing that humans to-be specialised for speech communication, most evidently for speech production by virtue of the shape and flexibility of movement of vocal tract and its articulators. Parallel specialization for speech perception would be possible despite the absence of experimental evidence for perceptual spepcialiaation" (1).

*

It is commonly assumed that the perceptual representation of speech sounds is organized, at leaset in part, in terms of intersecting properties or dimensions. Several studies have attempted to determine the number and nature of these perceptual dimensions or features that listeners employ in the identification of speech sounds (Eg: vowels (3) consonants (5),(8). With few exceptions, most of the earlier studies have been limited to the use of isolated monosyllables or disyllables. Moreover, there has been a suggestion that one cue is primary. Only some studies suggest a specific set of stimuli ({ 5), (7)).

The present article attempts to analyze the dimensions underlying the perception of stop consonants in kannada, using imitation paradigm. Imitation is a remarkable skill that requires one to parse the behavior of another into components and then activate his own corresponding motor controls to reproduqr; the behaviour. The human capacity to imitate the speech of another implies th:s presence of a specialized seneorimotor device and suggests that phonetic structure may be represented in the brain in a form sufficiently abstract, for there to be ready interchange between listening and speaking (16). Hence, during imitation, the imitator will have to attend to those features necessary for the perception of stop consonants {as used in present study) and then activate his motor system to reproduce the stop consonants along with those features he has perceived. Hence, the original and imitated utterances should be similar to each other with respect to perceptual features when measured objectively.

2. METHODOLOGY

2.1. Material: 16 stops (k,kh,g,gh,t,th,d,dh.t.th,d,dh.P.ph.b.hh) were selected for the present study. These stap consonants were considered to study the effect of following vowel (a,i,u,e,o.a:i:&e:) and following consonant (k, g, c , j , t, d, t, d , p,b,bh,ri, n ,n, n.m, r, v, s . h) on temporal and spectral aspects of stops. Meaningful words with stops in initial position were embedded in a carrier phrase- "i:ga na:nu he:ltini word" & this formed the material.

2.2. Subjects: One young adult male Kannada speaker aged 19 years served as a model subject which was imitated by a 24 year old Kannada speaker

2.3. Method: The model subject was presented with the sentences, one at a time. He was instructed to speak the sentences into a microphone which was placed at a distance of 10 cms from his mouth. The sentences of the original speaker were presented binaurally to the imitator who was instructed to imitate the sentences, giving attention to stops. He was allowed to practice and all these were recorded. All the utterances were recorded on a high fidelity magnetic-tape. Only the words (containing the stops) of both the model and imitated sentences were subjected to spectrographic analysis. Bar type wide band and average amplitude spectrograms were obtained for each word. The burst of stop in each word was analyzed using a computer for spectral measures.

Temporal eues extracted included closure duration, voicing duration, burst duration, aspiration duration, murmur duration, VOT, total word duration, transition durations of Fl, F2 and F3 of the following vowel and speed of transition of Fl, F2 and F3 of the following vowel. Spectral parameyters extracted included Fl, F2, F3, Bl, B2, B3, LI, L2, L3 , burst amplitude , overall amplitude, amplitude of following vowel and amplitude of following consonant.

3. RESULTS

Walsh test was performed to find out the significance of difference between means of various spectral and temporal parameters of the model and imitated speech.

TABLE 1: **Significant difference in temporal parameters of stop consonants:**

Features	k	kh	g	gh	ṭ	ṭh	ḍ	t	d	p	ph	b	bh
closure duration			++	--	+		-+		-+			++	--
voicing duration			++	--			--		--			--	+-
Burst duration	++	--	++	++	+	--	--	++	--	-+	-	--	--
Aspiration duration	++	++	--	--	-	--	--	--	--	--	-	--	+-
Murmur duration				++	+	--	--	--	--	--	-		--
VOT	++	++	+-	--	-	+-	--	++	--	--	-	--	+-
Word duration	--	-+	++	--	+	--	+-	++	-+	++	-	-+	--
Transition duration of F1	--	--	++	--	-	--	-+	--	-+	--	-	--	--
Transition duration of F2	-+	--	--	--	+	--	--	--	++	--	-	+-	--
Transition duration of F3	--	--	--	--	-	--	--	--	--	--	-	--	--
Speed of transition of F1	--	--	--	--	-	--	--	--	--	--	-	--	--
Speed of transition of F2	++	--	--	--	+	--	--	--	--	--	-	--	--
Speed of transition of F3	--	--	--	+-	-	--	--	--	--	--	-	-+	--

Initial +/- Presence/absence of significant difference in following vowel condition

Final +/- Presence/absence of significant difference in following consonant condition

Table 2: Significant difference in spectral parameters of stop consonants :

Features	k	kh	g	gh	ḍ	t	d	dh	p	ph	b	bh
F1	-+	--	++	--	++	-+	-+	-	++	+	++	+-
F2	--	--	++	++	--	--	-+	+	++	-	++	++
F3	++	++	++	++	--	++	--	-	++	-	-+	--
B1	--	+-	--	--	-+	--	--	-	++	-	++	+-
B2	--	--	+-	--	+-	--	--	--	--	-	--	--
B3	--	--	+-	--	--	--	+-	+	-+	-	--	--
L1	--	--	--	++	-+	--	-+	+	++	-	+-	+-
L2	--	--	--	--	--	--	--	-	+-	-	--	--
L3	--	++	-+	++	--	--	-+	+	-+	-	++	++
Burst amplitude	--	--	--	++	--	--	++	-	--	-	--	--
Overall amplitude	++	--	+-	++	++	--	++	+	--	-	--	++
Amplitude of following vowel	++	+-	++	++	-+	++	--	-	++	+	+-	--
Amplitude of following consonant	++	++	--	--	-+	-+	++	+	--	-	--	--

Initial +/- Presence/absence of significant difference in following vowel condition.

Final +/- Presence/absence of significant difference in following consonant condition.

Table 3: Percentage of significant difference between two subjects with respect. to temporal parameters.

Parameters	%age of signific ant difference Present	absent
Transition duration of F3	0	100
Speed of transition of Fl	0	100
Speed of transition of F3	8.33	91.67
Speed of transition of F2	12.5	87.5
Transition duration of Fl	16.67	83.33
Murmur duration	18.75	81.25
Transition duration of F2	20.83	79.17
Aspiration duration	20.83	79.17
Voicing duration	25	75
VOT	37.5	62.5
Burst duration	41.67	58.33
Word duration	45.83	54.17
Closure duration	53.84	46.15

Table 4: Percentage of significant difference between two subjects with respect to spectral parameters.

Features	%age of significant difference Present	Absent
L2	4.55	95. 45
B2	9.09	90.81
Burst amplitude	18. 19	81.81
B3	22.73	77. £7
Bl	31.82	68. 18
L1	40.9	58.09
Amplitude of following consonant	40.9	59.09
Overall amplitude	54.55	45. 65
L3	54,55	45. 65
F2	54 . 55	45. 65
Fl	59.09	40.60
F3	59.09	40 . 80
Amplitude of following vowel	63.64	36.88

The results are presented in tables i thro' 4. It was noticed that the temporal parameters, durations and speed of transitions of Fl. F2 and F3, murmur duration, aspiration duration find voicing duration were imitated to an extent of more than 75%. There was no marked similarity across the two subjects with respect to closure duration, burnt duration, VOT and word duration. Among the spectral parameters, bandwidth and level of F2, burst amplitude and bandwidth of F3 were imitated to an extent, of more than 75%. There was no marked similarity across the t.wo sub.iects with respect to Fl , F2. F3, bandwidth of Fl , Levels of Fl and F3, overall amplitude and amplitude of the following consonant.

4. DISCUSSION

Though the spectrograms of the model and imitator resembled each other. there were differences in them. The features (both temporal and spectral can be classified into talker dependent and talker independent features. The talker dependent features for stops in the present study were transition durations and speed of transitions of F1, F2 , F3, aspiration duration, murmur duration and voicing duration among the temporal parameters and bandwidths of F2, F3 level of F2 and burst amplitude among spectral parameters.

The present study is in consonance with several studies with respect to perceptual features of stops (Transition duration (8),(9); aspiration duration (10),(11); closure voicing (11); level of F2 (12));burst amplitude (13).

Also the results are in consonance with several studies in that some features do not aid in stop consonant perception (Closure duration (14), (15); spectra (16) .

Thus it seems that durations and speed of transitions of formants, burst amplitude, closure duration, aspiration duration, B2, B3, L2 and burst amplitude serve as perceptual cues to stop consonants in Kannada. However as thin study has used an imitation paradigm, the perceptual cues could be testified by using synthetic stimuli which may yield a better result.

REFERENCES

I. Repp.B.H.. (1982), cit in Hearing science & Hearing disorder by Academic Press. London, 1933.
3. Ainsworth. W.A., (1972) cit in Speech perception by Studdert-Kennedy, Language & Speech, Vol 23, 1980. 45-66.
5. Raphael. L. J.(1972) preceding vowel duration as *a* cue to the perception of the voicing characterstics of word- final consonants in American English ' JASA, 51(4), 1296- 1303.
6. Lisker, L. *&.* Abramson, A., (1964), 'A cross-language study of voicing in initial stops Acoustic measurements' Word. 20, 384-422.
7. Noli, J.D. (1960) cit in Duration, context as cues to word final cognate opposition in English, by Rapheal, L.J. Phonetica. 38, 1981, 126-147,
8. Walsh, M.A. & Diehl. R.L. (1987), Auditory factors in the perception of stops & glides' . JASA,82 Supp1, S80.
9. Zakia, k.A.E. & Kingston, J. (1907), The role of rate of transition in place perception', JASA, 82 Supp 1, S81.
10. Reeds, J.A. & Wang, W.S.Y.,(1961) , The perception of stops after S, Phonetica, 6 78 - 81.
11. Raphael, L.J. (1981). 'Durations & contexts as cues to word- final cognate. opposition in English ', Phonotica, 38. 126- 147.
12. Ainsworth, W. A. & Miller,J.B. '(1972) , The effect of relative formant amplitude on the perceived identiyt of stnthetic vowels'. Language & Speech, Vol 5 (41), 328-341 .
13. Mack, M. A. Blumstien. S.E (1983), 'Further evidence of acoustic invariance in speech production: The stop glide contrast', JASA, 73 (5). 1739-1760.
14. Lisker. L., (1978), The stability of closure as cue to medial stop voicing in English' JASA. 64. 1, S18.
15. Kepp. B.H. ' Limits on the power of silence as a stop manner cue' JASA, 72, suppl, S16.
16. Endres, W . Bambach. W & Fiusster, G. (1971), Voice spectrograms as a function of age. voice disiguise & voice imitation[1],' . JASA, 49, 6(2), 1842-1848.

ACKNOWLEDGEMENTS:

This study is supported bv the grants received from the Deptt. of Science & Tchnology (SP/ YS/LS7/ 1937).

APPENDIX IV

PHONETIC SYMBOLS USED

	VOICELESS UNASPIRATED	VOICELESS ASPIRATED	VOICED UNASPIRATED	VOICED MURMERED
VELAR	k	kh	g	gh
RETROFUEX	t	th	d	dh
DENTAL	t	th	d	dh
BILABIAL	P	ph	b	bh
PALATAL (AFFRICATE)	c	c h	j	jh

BIBLIOGRAPHY

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

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

Benquerel el al. (1978), Cit. in "Representation of voicing contrasts using articulatary gestures", by Goldstein, L. and Browman, *C.P.* (1986), J. of Phonetics, 14, 339-342.

Blufflstein. S.E. and Stevens, K.N. (1976), "Perceptual invariance and onset spectra -for stop consonants in different vowel environments", J. Acoust.Soc. Am. 60, Suppl. No. 1, S90.

Browman, C.P. and Goldstein, L. (1986), Cit, in "Representation of voicing contrasts using articulatory gestures", by Goldstein, L. and Browman, C.P. (1994?), J. of Phonetics, 14, 339-342.

Chasedde, A.N. and Gobi, C. (1987), "Cross-lanquage study of the effects of voiced-voiceless consonants on the vowel voice source characteristics", J, Acoust. Sac. Am., 82, Suppl , No, 1, S116.

Cole, R.A. and Scott, B. (1974), "Toward a theory of speech perception", Psychological Review, 81, 348-374.

Danilaff, R. (1977, 1980), in "The physiology of Speech and Hearing An introduction", by Daniloff, R., Schuckes, G. and Feth, L. (1980), Prentice-Hall *Inc., N*.J. U.S.A.

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, 15S6-1558

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

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. (1953), "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, S78.

Dorman, M.F. and Dougherty, K, (1981), "shifts in phonetic identification with changes in signal presentation level", *J.* Acoust. Sac. Am., 69 (5), 1439-1440.

Dorman, M.F. and Raphael, L.J. (1977), "Onset spectra and stop consonant, recognition", J. Acoust. Soc. Am.,62, Supp1. No.1, S78, HH9.

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., 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. 65 (6), 1518-1532.

Dorman, M.F., Studdert-Kennedy, M. and Raphael, L.J. (1976), "Stop bursts and transitions as functionally equivalant context sensitive cues", J. Acoust. Soc. Am., 60, Suppl. No.l, S90.

Dorman, M.F., Studdert-Kennedy, M. and Raphael, L.J. (1977), "Stop consonant. recognition : Release bursts and formant transitions as functionally equivalent, content depends it cues", perception and psychophysics, 22 (2) , 109 - 122.

Eimas, P.D. et, al. (1978), Cit. in "Studies of place and manner of articulation *in* syllable final position" by Miller. J.L., Eimas, P.D. and Zatorre, R.J. <1979), J. Acoust, Soc. Am., 66 (4), 1207-1210.

Endress, W., Bombach, W, and Flosser,G., *i* 1971) " Voice spectrograms as function of age, voice disguise, and voice imitation"-, J. of Acoust. Soc. Am, 49, 1842-1848.

Erickson, D.M., Fitch, H.L., Halwes, T.G. and Liberman, A.M. (1977), "Trading relation in perception between silence and spectrum", J. Acoust,. Soc. Am., 61, Suppl. No. 1, S46-47.

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

Fischer-JOrgensen, E. (1954), Cit. *in* "Stop consonant recognitions Release bursts and formant transitions as functionally equivalent, context-dependant cues", by Dorrman, M.F., Studdert-Kennedy, M. and Raphael, (1977) perception and psychophysics, 22(2), 109-122.

Fischer-JOrgensen, E. (1979), "Temporal relations in consonant vowel syllables with stop consonts based of Danish material", in Lindblom, B. a*nd O*hman, S. "Frontiers of speech communication research", Academic Press, Inc., London

Fitch, H.L., (1980), " The influence of preceding syllablo structure on intervocalic voicing boundary",J. Acoust, stop Amer. 63 (1), S50.

Fitch, H.L, Halwes, T, Erickson, D.M, and Liberman, A.M. (1990) "Perceptual equivalence of two acoustic cues for stop consonant manner", Perception and Psychophysics, 27(4): 343-350.

Flanagan, J.L (1972), "Speech analysis, synthesis and perception", Springer-Verlag, New York.

Flege (1982), Cit. in "Representation of voicing contrats using articulatory gestures", toy Goldstein, L. and Browman *C.P* (1986), J,. of Phonetics, 14, 339-342.

Fruin, C.W. *and* Bischaff, D.M. (1976), "Vowel duratjon as a cue for consonant voicing?". J. Acoust. Soc, Am., 60, Suppl. No. 1, S91.

Fry et. a*1.* (1962), Cit. in "Effects of transition length on the perception of stop consonants", by Keating, *P. a*nd Blumstein, S.E., J. Acoust. Sac. Am., 64 (i), 57-64, 1978.

Fujimura, O. (1961), "Bilabial stop and nasal consonant: A motion picture study and its acoustical implications', J.S.H.R., 4, 233-247.

Fujisaki, H, and Kawashima, T. (1969), cit, in "Effectt of transition length an the perception of stop consonants", by Keating, P. and Blumstein, S.E., J. Acoust.Soc. Am. 64 (1), 57-64, 1978,

Goldstein and Browman (1986), "Representation of voicinn contrasts using articulatory gestures", J. of Phonetics, 14, 339-342.

Gruenfelder , T.M. (1979), "Fundamental frequency as a cue to post vocalic consonantal voicing: Some data from perception and production", J. Acoust. Soc, Am., 65, Suppl. No. 1.S8.

Halle and Stevens. K.N. (1971), Cit. in *"FO* as an acoustic correlate of stop consonant voicing", by Ohde, R.N. (1984), J. Acoust. Soc. Am., 75(1), 224 - 230,

Hall, M. and Tosi , O. (1975), "Spectrographic and aural examination of professionally mimicked voices", J. Acoust. Soc. Am,, (58) , S-107 (A).

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.

Hoffman, H.S. (1958), "Study of some cues *in* the perception of the voiced stop consonants", J. Acoust. Soc, Am., Soc. 1035-1041.

Honda (1981), Cit,, in *"F0* as art acoustic correlate of stop consonant voicing", by Ohde, R.N. (1984), J. Acoust. Soc. Am., 75 (1) , 224-230.

House, A.S. and Fairbanks, G, (1953), Cit. in "*FO as* an acoustic correlate of stop consonant voicinq", by Ohde, R.N. (1984), 75 (1), 224-230.

Ingrisang, D., Hillenbrand, J., Smith, B.L. and Flege, J. E. (1982), "A perceptual study of the VDiced-voiceiess contract *in* syllable final stops", J.Acoust. Soc. Am., 72, Suppl.No. 1, S15 - 16.

Keating, P.A. (1979), "Differences in production and perception of VOT in Polish and English", J, Acoust. Soc. Am.,66, Suppl. No. 1, S87.

Keating, P.A. (1984), Cit. in "Representation of voicing contrasts using articulatory gestures", by Goldstein, L. and Browman, C.P. (1986) , J. of Phonetics, 14, 539-342.

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. III, (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.

K'ersta,L.G, (1962 b) cit. in voice identification, Theory and legal applications by Tosi , 0, University Park Prest. , Baltimore, 1979

Kewley-Port, D. (1979), "Spectral continuity of burst ano formant transitions as cues to place of articulation in stop consonants", *J.* Acoust. Sac. Am., 65, Suppl. No.l, S32.

Kewley-Port, D., Pisoni, D.B. and Studdert-Kennedy, M. (19S3), Perception of static and dynamic acoustic cues to place of articulation *in* initial stop consonants", J. Acoust. Soc. Am., 73 (5), 1779-1793.

Klee, T.M., Weismer, G. and Ingrisano, D.R. (1976), "Laryngeal timing constraints in plosive-vowel fricative vowel syllables *or* : Is VOT really the best measure of glottal supraglottal timing?". J. Acoust. 'Soc, Am., 60, Suppl. No.1, S63.

Kozhevni*k*ov. V.A. and Chistovish, L.A. (1965), "Speech, articulation and perception" CTransl), Washington: Clearing house for Fed. Sc. and Techn. Inf. JPRS, 30, 543.

Kuehn, D.P. (1973), Cit. in "Stop-consonant recognition : Release bursts and form ant transitions as functionally equivalent. context-dependent cues", by Dorman, M.F., Studdert-Kennedy, M. and Raphael, L.J. (1977), *Perception* and Psychophysics, 22 (2*),* 109 - 122.

Lahiri, A. (1980), "Voicing contrast in 4 category stop consonants", J. Acoust, Soc. Am., 67, Suppl. No. 1, S52.

Lea, (1973) , Cit. in *"FQ* as an acoustic correlate of stop consonant voicing", by Ohde, R.W. (1984), *J.* Acoust. Soc. Am., 75 (1), 224 - 230.

Liberman et.. al . (1956), Cit. in "Effects of transition length on the perception of stop consonants",by Keating, P. and Blumstein, S.E., J. Acoust. Soc.Am.,64(1), 57-64, 1978.

Liberman et al (1957),, Cit. in "Effects of transition length on the perception of stop consonants", by Keating, *P.* and Blumstein, S.E., J. Acoust, Soc. Am., 64 (1), S7-64, 1*978.*

Liberman, A.M., Delattre,P.C., & Cooper, F.S. (1958), Cit. in "Role of formant transitions in the voiced-voiceless distinction for stops", by Stevens, K.N. *and* Klatt, D.H. (1974), J. Acoust. Soc Am., 55 (3), 653-659.

Liberman, A.M., Cooper, F.S., Shankwei1er, D.P. and Studdert-Kennedy, M. (1967), Cit. in "Vocalic transitions in the perception of voiceless initial stops" by LaRsviere, C., Winitz, H. and Herriman, E. (1975), J. Acoust. Soc. Am., 57 (2), 470-4 75.

Liberman, A.M., Deiattre, P.C. and Cooper, F.S. (1952), "The role of selected stimulus variables in the perception of the unvoiced stop consonants", Am. J. Psycho1., 55, 497-516.

Liberman, A.M., Delattre, D.C., Cooper, F.S. and Gerstman, L.J. (1954), "The role of consonant-vowel transitions in the perception of the stop and the nasal consonants", Psycho1ogica1 monographs, 68, 1-13.

Liberman, A.M., Delattre, P. C. , Serstman, L.J. and Cooper, F.S. (1956), "Tempo of frequency change *an a* cue for distinguishing classes of speech sounds", J. of experimental psychology, 52 (2), 127-137.

Liberman, Mattingly and Turvey (1972), cit. in "effects of transition length on the perception of stop consonants", by Keating, P. and Blumstein, S.E. J. Acoust. Soc. Am., 64 (1), 57-64, 1978.

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

Lieberman, P. (1977), "Speech physiology and acoustic phonetics; An introduction", MacMillan Publishing Co., Inc., N. V.

Lisker, L. (1975), "Is it. VOT or a first formant transition detector?", J. Acoust. Soc. Am., 57 (6) , 1347-1551.

Lisker, L. (1977), "Closure hiatus: Cue to voicing, manner and place of consonant occlusion", J. Acoust. Soc. Am., 61, Suppl. No.l, S18.

Lisker, L. (1978), "Stability of closure as cue to medial stop voicing in English", J. Acoust. Sac. Am., 64, 1, S18.

Lisker, L. and Abramson, A. (1964), "A cross-language study of voicing in initial taps; Acoustic measurements", Word, 20, 364-422.

Lisker L, and Price, V. (1979), "Context determined effects of varying closure duration", J. Acoust. Soc. Am.65, Suppl.No.1, S 7.

Lofquist, (1975), Cit. in "F$ as an acoustic correlate of stop consonant voicing", by Ohde, R.N. (1984), J. Acoust. Soc. Am., 75 (1), 224-230.

Mack, M. and Blumstein, S.E. (1983), "Further evidence of acoustic invar iance in speech production s The stcn glide? contrast", J. Acoust. Soc. Am., 1739-1750.

Mann, V.A. (1980), "Influence of preceding liquids on stop consonant perception", J. Acoust. Sac, Am., 67, Suppl. No. 1, **S99.**

Mann, V.A. and Repp, B.H. (3.978), "Influence of preceding /a/ or / ~ / on stop consonant perception", J. Acoust. Soc. Am., 64, Suppl . No. 1, S17.

Miller, J.L. and Baer , T. (1983), "Some effects of speaking rate on the production of/b/ and /w/", J. Acoust. Soc. Am., 73 (5) , 1751-1755.

Miller, J.L. and Eimas, P.D. (1977), "Studies on the perception of place and manner of articulations. A comparison of the 1abial-alveol*ar* and nasal-stop distinctions" J. Acoust. Soc. Am., 61 (3), 835-845.

Moslin, B.J. and John, K.S. (1976), "Korean apical stop production : A VOT analysis", J. Acoust. Soc. Am., 60, Suppl. No. 1, S45,

Nolan, F. (1983), "The phonetic basis of speaker recognition", Cambridge University Press, Cambridge.

Ohala, M. (1978), "Acoustic and aerodynamic correlates of Hindi stops", J. Acoust. Soc. Am., 64, Suppl. No. 1, S92.

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, S19.

Ohde, R.N. (1982), "The effects of linguistic context on temporal and F0 properties of speech", J. Acoust. Soc. Am., 72, Suppl. No. 1, S65.

Ohde, R.N. (1984), *"FO* as an acoustic correlate of stop consonant voicing", J. Acoust. Soc. Am., 75 (1),

Perkell, J.S. (1969), "Physiology of speech productions Results and implications of a quantitative cineradiographic study", MIT Press, Cambridge.

Peterson (1983), Cit. in "Representation of voicing contrasts using articulatory gestures", by Goldstein and Browman (1986),J. of Phonetics, 14, 339-342.

Port, R.F. (19(30), "V/'C ratio as a post vocalic voicing cue", J. Acoust. Soc. Am., 67, Suppl. No. 1. 351,

Porter et.al, (1987), "Rate va duration of formant transition as a discriminative cue", J. Acoust. Soc. Am., Suppl. No,. 1, S81.

Price, P. and Lisker, L. (1979), "(/b/ - /p/) but, not-(/p /b/)", J, Acoust. Soc. Am., 65, Suppl. No. 1 , S7-8.

Raphael, L.J. (1972), "Preceding vowel duration as a cue to the perception of the voicing characterstics of word final consonants in American English", J. Aoust. Soc. Am., 51 (4), 1296-1303.

Raphael, L.J. (1980), "Contribution of CV transition duration to the perception of final consonant voicing in natural speech", *J.* Acoust. Soc. Am., 67, Suppl. No.l.

Raphael, L.J. and Liberman, A.M., (1980), "On defining the vowel duration that cues voicing in final position". Language a*nd* Speech, 23 (3), 297-308.

Reeds, J.A. and Wang, W.S.Y. (1961), "The perception of stops after "s", phonetica, 6, 78-81.

Repp, B.H. (1976), "The voicing boundary as a function of F2 *and* F3 transitions and FO", J. Acoust. Soc. Am., 60, Suppl. No, 1, S91.

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. 1, S8.

Repp, B.H. (1982), ''Limits ors the power of silence as a stop manner cue", J- Acoust. Soc. Am., 72, Suppl. No. 1, S16.

Repp, B.H. (1984), "Categorical perception : Issues, Methods, findings", in Speech and Languages Advances in bamic research and practice, Vol.10, Lass, N.J. (Ed), Academic Press, Inc., U.S.A., 1984.

Repp, B.H. and Mann, V.A. (1980), "Perceptual assessment of fricative-stop coarticulation", J. Acoust. Soc. Am., 67, Suppl.No. 1, S100.

Shafer, 6. (1974), Cit. in "Implications of studying reduced consonant, clusters in normal and abnormal child speech" by Kornfeld, J.R. in Campbell, R.N. and Smith, P.T. (Eds), Recent advances in the psychology of language, New York, Plenum Press, 1978.

Stathopoulos, E.T. and Weismer, G. (1979), "The duration of stop consonants", J. Acoust. Sac. Am., 65, Suppl. No. 1, S79.

Stevens, K. (1975), Cit. in "FO as an acoustic correlate of stop consonant voicing", by Ohde , R.N. (1984), *J.* Acoust. Soc. Am., 75 (1), 224-230.

Stevens, K.N. (1977), "Onset spectra as cues for consonantal place of articulation", J. Acoust. Soc. Am., 61, S48.

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

Stevens et. al. (1969), Cit. in "Effects of transition length on the perception of stop consonants", by Keating, P. and Blumstein, S.E., J. Acoust. Soc. Am., 64 (1), 57-64, 1978.

Studdert-Kennedy, M. (1976)," Speech perception " in N.J.Lass (ed) Contemporary issues in experimental phonetics, Academic Press, New York.

Summerfield, Q. and Haggard, M. (1972), cit. in "Role of formant transitions in the voiced-voiceless distinction for stops", by Stevens, k.N. and Klatt, D.H. (1974), J. Acoust. Soc. Am., 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, 433-448.

Umeda,N (1981), Cit. in *"FO* as an acoustic correlate of stop consonant voicing", fay Ohde, R.N. (1984), J. Acoust. Soc. Am., 75 (1), 224-230.

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.

Walsh, M.A. and Diehl, R.L. (1987), "Auditory factors in the perception of stops and glides", J. Acoust. Soc. Am., 67, Suppl. No. 1, S51.

Walsh, T. and Parker, F. (1980), "Vocal fry: A cue for voicing in post vocalic: stops", *J.* Acoust. Soc. Am., 67, Suppl. No. 1, SSO.

Walsh, T., Parker, F. and Miller, C.J. (1987), "The contribution of rate of Fl decline to the perception of [+/-voiced", *J.* of Phonetics, 15 (1), 101-103.

Weismer, B. (1977), "Gome context effects on voice onset time", J. Acoust. Soc, Am., 62, Suppl. No. 1, S 78, HH10.