Transients at Stop-Consonant Releases

by

Noel Steven Massey

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

© Noel Steven Massey, MCMXCIV. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part, and to grant others the right to do so.

Author ....
Department of Electrical Engineering and Computer Science
May 6, 1994

Certified by ....
Kenneth N. Stevens
Clarence J. LeBel Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Frederic R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

Eng.
MASSACHUSETTS INSTITUTE
JUL 13 1994
LIBRARIES

Transients at Stop-Consonant Releases

by

Noel Steven Massey

Submitted to the Department of Electrical Engineering and Computer Science
on May 6, 1994, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

The acoustic properties and perceptual significance of acoustic transients at the release of stop consonants and affricates was studied. The theory of transient production is based on a model in which pressure is built up in a closed tube and the radiated sound is calculated following an abrupt release of the pressure. Acoustic analysis of the amplitude and frequency content of transients in utterances from ten speakers was used as a guide in setting the parameters of the model. A speech synthesizer was modified to permit the generation of transients with frequency content similar to that observed in natural speech. Utterances consisting of a stop consonant or affricate followed by the vowel /ɑ/, with various amplitudes of an initial transient, were synthesized, and listeners were asked to judge the naturalness of the stimuli. Listeners tended to prefer the presence of a transient with an amplitude approximating the measured amplitude in natural speech, although there was some variability in the responses.

Thesis Supervisor: Kenneth N. Stevens
Title: Clarence J. LeBel Professor of Electrical Engineering

Acknowledgments

I would like to dedicate this thesis to my parents, Bob and Lois Massey, who continue to give me their support and love. Even the knowledge gained at MIT does not compare to what you have taught me.

Thank you Ken Stevens for your guidance and council. Also, thank you for funding most of my graduate study while still allowing me to spend time outside of the lab. Your generosity is appreciated. I would also like to thank the other members of the Speech Communication Group for helping with random questions, and a special thanks goes to my office mates, who have helped sustain public relations by answering many of my phone calls and who have put up with an unsightly desk.

Thanks Kent, Jack, Andy, Steve, and Pat for caring enough to get involved in my life at least once a week and for being a constant support in an occasionally chaotic life. Thanks Bryan, you have been an example that will not be forgotten. We are very different people in some ways, but you have taught me a lot. I admire your character and I look forward to even more growth that results from spurring each other on in life. Thank you Tammy for making life a little more interesting. You are a gift from God. Thanks Jenny for the many late night walks and your insightful second opinions. Thanks to my housemates for putting up with my mess (yes I am taking my stereo with me). Thanks Karl, for your willingness to lose major body parts with me for a little fun, let's make sure that they put us in the same hospital room. I would also like to thank Scott and Christine for keeping me up-to-date with new music which has often been a great stress reliever. Thanks Shana for never giving up on me, maybe now I can write back a little more often. Thank you also to Seekers and Crusade especially for providing an avenue to vent my musical abilities by allowing me to beat the skins (or plastic) each week.

Most importantly, I would like to thank Jesus for giving my life a meaning and purpose. Even the best of my peers are looking for more in life. I hope that my life encourages them to let you fill that void. May you receive all the glory for my accomplishments.

Contents

1 Background **14**

1.1 Description of Transient Bursts 14

1.2 Limited Transient Burst Knowledge 16

1.3 Importance of Transient Bursts 18

1.4 Difficulties in Studying Transient Bursts 18

1.5 Summary of Related Research 19

1.6 Preliminary Research 21

1.7 Proposed Research 21

2 Theory **23**

2.1 Vocal Tract Model 23

2.2 Pressure Equations 24

2.3 Equivalent Circuit 27

2.4 MATLAB Model - Rectangular Constriction 29

2.5 MATLAB Model - Circular Constriction 30

2.6 Filtering of Transient due to Anterior Cavity 33

3 Acoustic Analysis **36**

3.1 Data Collection 36

3.2 Amplitude Analysis 37

3.3 Spectral Analysis 58

4 Modifications of Synthesizer to Produce Transient Sources **69**

4.1 Synthesizer Background 69

4.2 Synthesizer Modifications 71

4.3 Synthetic Transient Waveform Shape 75

4.4 Analysis of Synthetic Transient Burst 76

5 Perception Tests **78**

5.1 Pilot Test 78

5.2 Description of Stimuli, Subjects, and Experiment 79

5.3 Perception Test Results 80

5.4 Comparison of Synthetic and Natural Transient Magnitudes 83

5.5 Problems with Synthesized Utterances 83

6 Conclusions **86**

6.1 Further Work 86

A Most Preferred Synthesized Stimuli - .DOC Files **88**

A.1 Most Preferred /pa/ Synthesis 88

A.2 Most Preferred /ta/ Synthesis 92

A.3 Most Preferred /ka/ Synthesis 96

A.4 Most Preferred /ba/ Synthesis 100

A.5 Most Preferred /da/ Synthesis 104

A.6 Most Preferred /ga/ Synthesis 108

A.7 Most Preferred /ča/ Synthesis 112

A.8 Most Preferred /ǰa/ Synthesis 116

B MATLAB Script Files **120**

B.1 Analysis Programs 120

B.1.1 abs_rms.m 120

B.1.2 analysis.m 120

B.1.3 analysis3.m 122

B.1.4 avg_spectrum.m 124

B.1.5 avg_spectrum_all.m 125

B.1.6 dftmarks.m 126

B.1.7 draw_all_marks.m 127
B.1.8 imp_syn.m 128
B.1.9 mark_dft.m 128
B.1.10 model1.m 130
B.1.11 model2.m 132
B.1.12 scale_speech_noel.m 134
B.1.13 segment.m 137
B.1.14 spect.m 138
B.1.15 spect_from_mag.m 139
B.1.16 transient_mag.m 141
B.2 MATLAB Window Management Programs 144
B.2.1 center_window.m 144
B.2.2 clear_mark.m 144
B.2.3 compress_speech.m 144
B.2.4 dft_active.m 145
B.2.5 expand_speech.m 145
B.2.6 ham.m 146
B.2.7 mark2_down.m 147
B.2.8 mark2_up.m 147
B.2.9 mark_speech.m 148
B.2.10 mouse_down.m 148
B.2.11 mouse_up.m 149
B.2.12 window2_update.m 149
B.2.13 window_left_move.m 150
B.2.14 window_move.m 150
B.2.15 window_position.m 151
B.2.16 window_right_move.m 151
B.3 Synthesized Transient Programs 152
B.3.1 trans_length.m 152
B.3.2 transient4.m 152

B.4 Perception Test Analysis Programs 154
B.4.1 ltr2hist.m 154
B.4.2 max_compare.m 156
Bibliography **157**

List of Figures

1-1 Waveform of utterance 'cha'. 14

2-1 Models of the vocal tract before (a) and after (b) the release of a stop consonant. 24

2-2 Equivalent circuit used to determine airflow. 27

2-3 Estimated spectrum of transient. 29

2-4 Transient pressure components (a), volume velocity (b), pressure at 20 cm (c), and spectra (d) of rectangular constriction model. 31

2-5 Acoustic mass term (a) and its derivative (b) for rectangular constriction model. 31

2-6 Transient pressure components (a), volume velocity (b), pressure at 20 cm (c), and spectra (d) of circular constriction model. 32

2-7 Approximate cavity for affricate transient filtering. 34

2-8 Transfer functions between constriction and mouth opening for alveolar and velar stop consonant releases.(Stevens, in preparation) 34

3-1 Measurement of vowel magnitude. 38

3-2 Measurement of transient magnitude. 39

3-3 Transient and Vowel magnitudes - 'peak'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 42

3-4 Transient and Vowel magnitudes - 'pot'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 43

3-5 Transient and Vowel magnitudes - 'teach'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 44

3-6 Transient and Vowel magnitudes - 'top'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 45

3-7 Transient and Vowel magnitudes - 'keep'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 46

3-8 Transient and Vowel magnitudes - 'cot'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 47

3-9 Transient and Vowel magnitudes - 'beat'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 48

3-10 Transient and Vowel magnitudes - 'bought'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 49

3-11 Transient and Vowel magnitudes - 'deep'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 50

3-12 Transient and Vowel magnitudes - 'dock'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 51

3-13 Transient and Vowel magnitudes - 'geek'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 52

3-14 Transient and Vowel magnitudes - 'got'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 53

3-15 Transient and Vowel magnitudes - 'jeep'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 54

3-16 Transient and Vowel magnitudes - 'job'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 55

3-17 Transient and Vowel magnitudes - 'cheap'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 56

3-18 Transient and Vowel magnitudes - 'chalk'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom). 57

3-19 Spectral content of transient bursts in /p/ followed by /ɑ/ or /i/ spoken by males or females. 61

3-20 Spectral content of transient bursts in /b/ followed by /ɑ/ or /i/ spoken by males or females. 62

3-21 Spectral content of transient bursts in /t/ followed by /ɑ/ or /i/ spoken by males or females. 63

3-22 Spectral content of transient bursts in /d/ followed by /ɑ/ or /i/ spoken by males or females. 64

3-23 Spectral content of transient bursts in /k/ followed by /ɑ/ or /i/ spoken by males or females. 65

3-24 Spectral content of transient bursts in /g/ followed by /ɑ/ or /i/ spoken by males or females. 66

3-25 Spectral content of transient bursts in /ch/ followed by /ɑ/ or /i/ spoken by males or females. 67

3-26 Spectral content of transient bursts in /j/ followed by /ɑ/ or /i/ spoken by males or females. 68

4-1 Simplified block diagram of formant synthesizer.(Klatt,1980) 70
4-2 Block diagram of formant synthesizer. 71
4-3 Block diagram of KLSYN93.(adapted from Klatt and Klatt, 1990) . . 72
4-4 KLSYN93 control parameter list. 73
4-5 Frequency content of transient approximations (length = 2,3,4,and 5 samples). 76
4-6 Frequency content of transient approximation (length=4 samples). . . 77
4-7 Transient approximation error (approximation - 1/f). 77
4-8 Best-fit transient waveform (length = 4 samples). 77

5-1 Results of pilot test for /pɑ/. 79
5-2 Perception test results. 82

List of Tables

3.1 Words used to analyze transient bursts. 37

5.1 Perception test results. Each number is the percent of times the stimuli was chosen when it was paired with the most preferred transient (marked by $\oplus$). 83

5.2 Comparison of transient magnitudes from most preferred perception test results and observed natural speech. 85

Chapter 1

Background

1.1 Description of Transient Bursts

The presence of transient bursts in speech has been recognized for many years [1]. All speech is generated by modulating airflow through a constriction in the vocal tract somewhere between the trachea and the lips. In some classes of sounds the constrictor forms a complete closure. When this constriction is released the initial increase in cross-sectional area of the constriction can produce a transient pulse in the waveform [2] as can be seen in Figure 1.1.

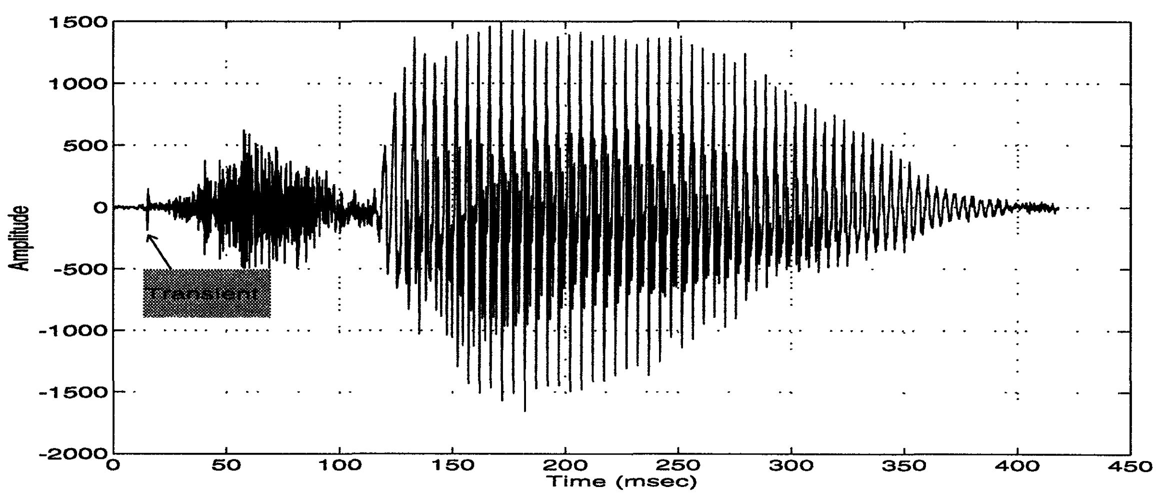

Figure 1-1: Waveform of utterance ‘cha’.

The stop consonants are one class of sounds produced by a series of articulatory movements which includes a complete closure of the vocal tract at some point in time. The first stage in the production of a stop consonant is often the movement of the tongue body, tongue blade, or lips to a position which allows for the effective closure of the vocal tract. Once in position, the major articulator forms a complete constriction. Pressure is then built up behind the constriction and this increased pressure can cause an outward displacement of the walls of the vocal tract. Following an interval of closure there is a rapid increase in the cross-sectional area of the constriction and a rapid decrease in intraoral pressure. This action causes an initial transient of airflow to be produced. This transient source is usually followed by a brief interval in which turbulence noise is generated in the vicinity of the constriction. Aspiration noise is then usually generated at the glottis, and voicing may follow if the vocal folds are in a configuration that permits vibration.

The production of affricates consists of a series of articulatory movements which are slightly more complicated than those required to produce a stop consonant. As with the stops, a complete closure is made at some point along the vocal tract. The production of an affricate differs from that of a stop in that the length of the affricate constriction is much greater and there is separate control over the anterior and posterior portions of the constriction [3]. The initial release following the closure is made by moving only the anterior portion of the constriction. A transient is still generated as with the stops but here the posterior portion of the constriction remains nearly unchanged for the next few tens of milliseconds and is then released. This produces an interval of silence which is followed by frication and aspiration, and finally voicing may occur if the vocal folds are allowed to vibrate.

Affricates and stop consonants are the only two sound classes in English which produce transient bursts. However, other languages have members in their inventory of sounds called clicks, and people communicate using this form of transient burst. Unlike the stop consonants in English, clicks are transient bursts which are produced by building up a negative pressure behind a constriction [2]. The first step in producing this negative pressure is to form a complete closure of the vocal tract with the tongue body and also to form another constriction at some point anterior to the first constriction. The volume between the two constrictions is then expanded to create the negative pressure. The release of the anterior constriction creates a rapid flow of air into the mouth which generates a transient burst. Several African clicks have been recorded and described by Anthony Traill [4]. Analyzing another class of sounds which contains transient bursts, such as these clicks, could enhance the theory and models of transients. This work has not yet been performed, and clicks were not investigated in this thesis due to the unavailability of a subjects who speak a language which uses clicks, and due to the limited number of clicks in the data base generated by Traill (at the Speech Communication Group at MIT). Without clicks to analyze and people to judge whether the changes in the clicks are perceived, the important features of clicks can not be determined.

1.2 Limited Transient Burst Knowledge

Transient bursts have not been studied in as much depth as other sound sources. A number of studies of turbulence noise have been performed which describe the airflow and pressures during the production of utterances containing various classes of speech segments. (An attempt to tie all the turbulent noise information together was presented by Stevens [5].) Transient noise, however, is generated by a different mechanism. Maeda [6] pointed out that the transient bursts are generated by a "coherent source", which means that the waveform and corresponding spectrum can be predicted exactly if the articulator positions and aerodynamic conditions are known. This coherent source contrasts the sound generated by aspiration or frication, which are stochastic noise sources. The output of these noise sources can not be predicted exactly, and only probabilistic events can be determined. The exact calculation of the transient burst can be performed, but has been limited primarily by the unavailability of articulator and pressure data.

There are several reasons for the lack of transient burst research, but the primary reason is probably the assumed perceptual insignificance of transient bursts. There is a large variability in the amplitudes of transient bursts, and sometimes it is difficult to see a transient burst in the waveform because its amplitude is no greater than that of the surrounding noise. These amplitude variations show that transient bursts may not provide the primary cue used to distinguish stop consonants, but this observation does not prove that listeners do not use the information present in transient bursts when the transient is loud enough to be perceived. Some researchers would also argue that transient bursts are not perceptually important since speakers are not intending to generate transient bursts. It is true that transient bursts are a by-product of a rapid opening of a constriction, but transient bursts do contain information about the place of articulation, and it would make sense that a listener would somehow use this information.

Current speech synthesizers contribute to the assumed irrelevance of transient bursts because the synthesizers do not incorporate transient bursts and yet the synthesizers are intelligible [7]. Synthesized stop consonants can be accurately classified without transient bursts, and even the highest quality speech synthesizers ignore the generation of these transients. One exception is the synthesizer designed by Fujisaki, Hirose, and Asano [8], which includes an impulse generator and path for synthesizing stops. The improvement to stops was not quantitatively measured, though it was mentioned that the new synthesizer generated stops of much higher quality. Most synthesizers, however, are not even capable of generating transients. If the transients are perceived, then the inclusion of transient bursts in synthesized speech would enhance the naturalness of synthesized speech.

The inability of synthesizers to generate transient bursts has also been a major inhibitor to our knowledge of transients. The importance of synthetic speech in research is that one parameter can be varied while the other parameters are kept exactly the same. Therefore, the perception of that one property can be closely examined. The limitations of current synthesizers have hindered our ability to study transients by forcing the primary method of investigation to be the observation of natural transient bursts, which are extremely difficult to control. The understanding of transient bursts will be rapidly accelerated by the ability to synthesize transient bursts, and the assumed irrelevance of transients can then be proved or disproved.

1.3 Importance of Transient Bursts

Though transient bursts are commonly observed, the amount of information that is conveyed by these bursts is still unknown. There are two major reasons why it would be valuable to know how extensively these transients contribute to the perception of speech. First, there is a high demand for verbal communication between humans and computers. Most speech synthesizers do not produce speech which includes these transient bursts. If the transient bursts are perceived, then their absence could degrade the naturalness of the synthesized speech. Many work environments and tasks could benefit from synthesized speech, but currently many people are not willing to listen this speech because it sounds too mechanical. Generating the transient bursts could make synthesized speech more pleasant to listen to and allow synthesized speech to be used in more applications. Also, the information contained in the transient bursts could potentially be used in speech recognition systems. Any improvement in the accuracy or speed of speech recognition would make interfaces with machines more natural. The second value in learning about transient bursts is that more accurate models can enhance our understanding of how humans communicate. The human mind and body is a complicated yet efficient system, and scientists and engineers are still striving to match the abilities of the human body. An understanding of humans can assist in developing other complicated systems.

1.4 Difficulties in Studying Transient Bursts

Analysis of transient bursts is difficult due to the fact that the bursts only exist for a very short period of time. The transients are generated immediately after the closure is released, and there is no way to release the closure in such a way as to make a substantial change in the duration of the transient burst. Furthermore the English stop consonants can not be produced in such a way as to allow a time of silence between the transient burst and the onset of frication.

In order to perform high resolution frequency analysis it is necessary to take a window that covers a large number of samples, much longer than the 1-3 ms during which the transient is present (for typical sampling rates of 10-20 kHz). A detailed analysis of the transient burst alone is not possible since the frequency analysis using a window which gives detailed frequency resolution will likely show effects from the following frication and aspiration noise. Likewise, the amplitude of the transient is often so small and the frication and aspiration noise so large that it is difficult to see the transient burst in the waveform itself. Therefore the nature of the transient burst makes it difficult to analyze the speech in both the time and the frequency domains.

1.5 Summary of Related Research

The burst portions of stop releases, which include the transient burst, frication, and aspiration have been extensively studied. It has been shown that there are indeed invariant acoustic properties intrinsic to the place of articulation. Blumstein and Stevens performed an experiment in 1978 [9] which classified the place of articulation using a rough template matching procedure. The place of articulation was correctly identified for over 80% of the stop consonants, which was a strong support for the theory of invariance. Winitz, Scheib, and Reeds [10] conducted a study which showed that listeners could identify /p/, /t/, or /k/ from the burst portion alone. In many instances the vowel preceding or following the stop could also be identified by only the burst portion, giving perceptual evidence for coarticulation. It is not known if these same types of characteristics that exist for the stop consonant burst also exist for the transient burst alone.

Transient bursts are the response of the vocal tract to the sudden pressure release. Due to their short duration in time (nearly impulsive) and the relatively slow movements of the articulators, the transient bursts are essentially acoustic snapshots of the vocal tract. Stevens and Blumstein's [9] claim that the cues for place of articulation can be perceived by these static snapshots was challenged by Kewley-Port's claim that listeners use time-varying features to determine place of articulation [11]. Kewley-Port, Pisoni, and Kennedy [12] later conducted an experiment to test which of these theories is correct. They confirmed that the cues for place of articulation are located in the initial 20-40 ms of natural stop-vowel syllables and performed additional experiments using synthetic speech. The data suggest that the static snapshots in general not did not provide enough information to determine place of articulation, though in some cases it was sufficient. The experiment should be questioned though because the synthetic snapshots may have lacked important information contained in natural speech.

Experiments with natural speech are difficult since speakers have little control over the transient burst. However, Repp conducted several experiments to investigate stop consonant release transients that were produced in isolation [13]. These experiments required a speaker to recite stop consonants in a very quiet whisper so that "the frication and aspiration sources were attenuated to virtual silence and only the initial transient remained" [13, page 381]. Repp analyzed the transients using a 12.8 ms Hamming window which started at the transient release. A high degree of influence of the following vowel was noted. It was also concluded that the transient bursts do contain precise information about the vocal tract, but listeners were not very good at identifying place of articulation or vowel configuration from the transient burst alone. Listeners who were presented with transients alone, however, did not seem to perform any worse than when they were presented with the entire release burst (transient, frication, and aspiration). The information contained in the transient appears to be perceptually equivalent to the information contained in the entire release burst.

Maeda [6] constructed a transient burst model which calculated the airflows and pressures due to the "coherent", or non-stochastic, source. Various model parameters were changed and simulations were run to show how the model behaved for different values of the parameters. The effects of these variations on the time waveform were described but no frequency analysis was performed and only a few comparisons were made between the simulations and natural speech. As a result, it is difficult to conclude whether this model captures the salient features of a transient burst. The research presented in this thesis will approach this problem from a spectral point of view and will develop a theory and model for the generation of transient bursts which will be rigorously compared to transients in natural speech.

1.6 Preliminary Research

As a final project in the Speech Communications class here at MIT, some preliminary research on transient bursts was performed. An impulse, lasting only one sample, was added to the beginning of six synthesized consonant-vowel utterances. The utterances consisted of a stop consonant (/ptkbdg/) followed by the vowel /ɑ/. The level of the impulse was varied, so that a total of six levels for each stop consonant was generated. These stimuli were combined in every possible pair (within the same stop consonant) and presented to six listeners as a two-alternative forced choice test. The task of the listeners was to judge which element of the pair, and hence which transient amplitude, sounded more natural.

The results of the perception test showed that people preferred a moderate level transient to no transient at all, but also that people preferred no transient to a very loud transient. The general results were similar for each of the stop consonants. These results, however, have limited significance to speech because the transient burst that was added does not resemble transient bursts found in natural speech. The high frequency components of the synthesized transient were much too strong. This thesis describes experiments which correct for the shortcomings of the preliminary experiments. Any experimental evidence or theory resulting from this thesis will supersede any previous conclusions, and therefore the aforementioned experiments have limited value in understanding transient bursts and will not be described any further.

1.7 Proposed Research

The research presented in this thesis consists of four major components; theory of stop-consonant releases, analysis of natural transient bursts, synthesis of transient bursts, and testing of transient burst perception. First, a theory of transient bursts was developed which predicts the sound pressure at a distance given the pressures and articulator positions surrounding the stop-consonant release. Natural speech was then analyzed to determine properties of transient bursts, and these transients were compared to those predicted by the theory. Synthetic transient bursts were then generated and compared to natural transients. Perception tests involving synthesized speech were administered and the resulting most preferred transients were compared to natural transients.

Chapter 2

Theory

2.1 Vocal Tract Model

To understand the origin of the transient burst we begin by modeling the vocal tract as a uniform tube of length l and cross sectional area A as seen in Figure 2.1a. Before the stop consonant is released one end of the tube is closed, corresponding to the complete closure made by the lips, tongue blade, or tongue body. The closure causes pressure to build up in the tube, eventually reaching a pressure close to the subglottal pressure. When the constriction is released our model changes as shown in Figure 2.1b where the constriction now has cross-sectional area $A_c(t)$ and length $l_c(t)$. The time varying length of the constriction is not known since it is very difficult to measure the length of the constriction immediately following the release. As a result of our lack of physical data, the theory assumes that the length is not varying with time but is actually fixed (at a length of approximately 0.1 cm) for the first two milliseconds after the stop consonant release. The air passing through this narrow constriction will cause sound to be produced. The sound is filtered by the cavity anterior to the constriction, except in the case of labial constrictions where this sound is directly radiated into air.

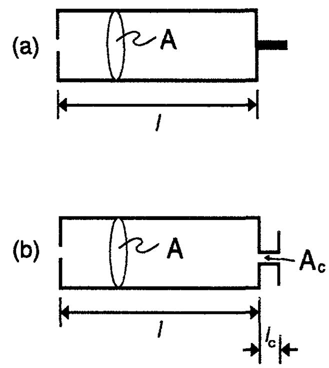

Figure 2-1: Models of the vocal tract before (a) and after (b) the release of a stop consonant.

2.2 Pressure Equations

We are ultimately interested in the sound pressure p_r at some distance r from the lips. If the output of the vocal tract is taken as the volume velocity U_0 at the lips then the sound pressure can be approximated [2] by:

$$p_r(t) = \frac{\rho}{4\pi r}\frac{\partial U_0(t - r/c)}{\partial t}. \tag{2.1}$$

where c is the velocity of sound and ρ is the density of air. Thus in order to interpret the features of the transient pulse as it relates to the sound pressure reaching the ears of the listeners we should examine the time derivative of the volume velocity at the lips.

The volume velocity at the lips is the volume velocity of the transient source after it has been filtered primarily by the anterior cavity, and to a lesser extent by the posterior cavity. The back-cavity resonances will not affect the transient burst if the constriction has a very small cross-sectional area and the vocal tract right behind the constriction has a large cross-sectional area. This configuration may not be realized due to the influence of surrounding phonemes. Therefore the back-cavity resonances, in addition to the front-cavity resonances, can affect the transient bursts. In order to keep the equations simple, the volume velocity will be calculated first without any filtering. The effects of the filtering will be discussed after the volume velocity has been determined. Ignoring any type of filtering results in a relation between pressure and volume velocity through a constriction that consists of four primary components.

The first component is due to the characteristic impedance of a uniform tube. The solutions to the one dimensional wave equation show that pressure and volume velocity are not independent. As a result of Newton's Law for a 'pocket' of air, a one dimensional analysis shows that

$$\frac{\partial p}{\partial x} = -\rho \frac{\partial u}{\partial t}. \tag{2.2}$$

where u is the linear velocity of the wave. Taking partial derivatives of the equations for pressure and particle velocity of a wave traveling in the forward direction in a uniform tube, and substituting them in the above equation results in the following relation:

$$P = \frac{\rho c}{A} U \tag{2.3}$$

where A is the cross-sectional area of the tube, U is the volume velocity of the air, and ρ is the density of air. The term $\rho c/A$ is known as the characteristic impedance of the tube.

The second pressure component which should be examined is the resistance of the constriction to the flow of air. Pressure and volume velocity at a constriction are not necessarily linearly related, and this relation is highly dependent on the shape of the constriction. Unfortunately the shape of the constriction immediately following the release is not known. Two different shapes, one circular and one rectangular, will be considered.

The cross-sectional area of the opening can be approximated as a rectangular slit with larger dimension b and smaller dimension d. The resistance of a rectangular constriction consists of two parts: a viscous component and a kinetic component [2, page 2.22]. If the flow is laminar and if one dimension is much less than the other dimension, then

$$P_{rect} = \frac{12\mu l_c}{bd^3} U \tag{2.4}$$

where μ is the viscosity and l_c is length of the constriction.

Similarly, for a circular constriction of length l_c the relation between pressure and volume velocity is

$$P_{circ} = \frac{128\mu l_c}{\pi D^4} U \tag{2.5}$$

where D is the diameter of the opening.

In addition to the effects of the viscosity of air, there is energy loss due to both the transition from a wide tube to a narrow tube and again from the transition from the narrow tube back to a wide tube. These losses are due to eddies that form on either side of the constriction. These pockets of circular airflow dissipate energy as heat. Experimental data have shown that this effect can be approximated by

$$P_{kinetic} = \frac{k\rho U^2}{2A_c^2} \tag{2.6}$$

where k is a constant close to 1.

A final addition to the pressure equation is a term which takes into account the acoustic mass of the air in the constriction. If we approximate the constriction as a tube of length l_c and area A_c which is open at both ends, the effect of the acoustic mass is simply

$$P_{mass} = \frac{d}{dt}\left(\frac{\rho l_c U}{A_c}\right). \tag{2.7}$$

The final equation used to approximate the pressure at the lips therefore consists of four terms. Assuming a rectangular constriction, the pressure can be approximated as

$$P_{i_{rect}} = \frac{\rho c}{A} U + \frac{12\mu l_c}{bd^3} U + \frac{\rho}{2A_c^2} U^2 + \frac{d}{dt}\left(\frac{\rho l_c U}{A_c}\right). \tag{2.8}$$

2.3 Equivalent Circuit

Analysis and understanding of the transient burst may be enhanced by constructing an equivalent circuit from the acoustic model. Figure 2.2 shows the circuit using a method where pressure is analogous to voltage and volume velocity is analogous to current. The equation which is represented by this circuit is:

$$P_i = \frac{\rho c}{A} U + R_c U + \frac{d}{dt}(M_c U) \tag{2.9}$$

where ρ is the density of air, R_c is the resistance of the constriction (consisting of both the viscous and kinetic components), and M_c is the acoustic mass of the air $(\rho l_c / A_c)$.

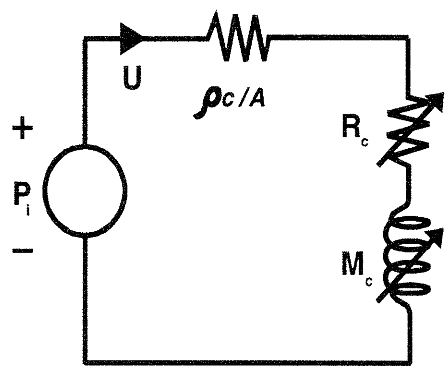

Figure 2-2: Equivalent circuit used to determine airflow.

Using typical experimental data for the parameters above [2], we can simplify Equation 2.8. At time t=0.2 ms following the release of a stop consonant A_c=0.02 cm^2 for a typical labial or alveolar release rate of 100 cm^2/s. If the tube's cross-sectional area A=3 cm^2 then the coefficient of the first term in Equation 2.8 is 13.5 gm/(s·cm^4). Using l_c=0.1 cm, and selecting typical dimensions d=0.1 cm and b=0.2 cm we find that the coefficient of the second term (viscous part of the constriction resistance) is 1.16 gm/(s·cm^4). The coefficient of the third term (kinetic part of the constriction resistance) would be 1.4 gm/cm^7. The last term is approximately 0 which we will show later. Comparing the magnitude of each of the terms in Equation 2.8, and noticing that the first two terms are proportional to U while the third is proportional to U^2, we can show that for typical parameter values, the dominant term is the kinetic resistance term. Therefore the inductance and top resistance are insignificant in our circuit. The insignificance of the other terms allows us to approximate the pressure as:

$$P_i \approx \frac{\rho U^2}{2A_c^2} \tag{2.10}$$

which can be rewritten as

$$U(t) \approx A_c(t)\sqrt{\frac{2P_i}{\rho}}. \tag{2.11}$$

From transmission line theory [14], the pressure in the tube will remain constant for time less than the time it takes for sound to travel twice the length of the tube. For a uniform tube, this means that for times less than about 1 millisecond the pressure will be constant (and can be approximated as constant for short times after this). Since P_i is a constant for the first millisecond or so, the volume velocity is just a linear function of $A_c(t)$. Since Fujimura [15] predicted a linear increase in the cross-sectional are of the opening, $U(t)$ is simply proportional to time. Therefore the time derivative of $U(t)$, which is a multiple of the pressure at some distance r from the lips, is a constant for the first millisecond or so. The spectrum of this pressure waveform falls off as 1/f. There are a few corrections to this spectrum which should be made due to the extensive approximations. The first correction is that the constriction opening will not increase as rapidly as predicted. This is primarily due to the Bernoulli forces which tend to pull the opening closed. The second correction is that the negative traveling wave in the tube will cause a reversal in the sign of $U(t)$ which will limit the amount of low frequency energy. If the tube is uniform this effect will not take place until the first millisecond or so after the stop release, but reflections from non-uniform vocal tract configurations could affect the pressure at times very close to the release of the stop. The combination of these effects will result in a spectrum which approximately falls off as 1/f in the middle frequencies but is limited at high and low frequencies as shown in Figure 2.3.

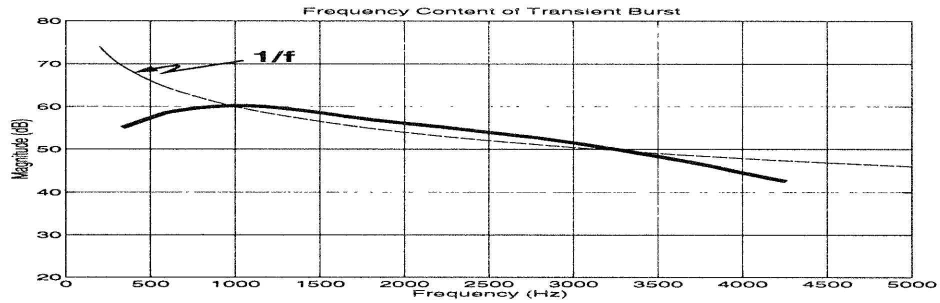

Figure 2-3: Estimated spectrum of transient.

2.4 MATLAB Model - Rectangular Constriction

In order to determine the validity of ignoring everything except for the kinetic component of the pressure term, a model of this theory was created using MATLAB. The model assumed constant pressure P_i for the first two milliseconds (which is a bit too long) and also ignored the last term in Equation 2.8 as a first approximation, since this term is expected to have a small effect for times and areas of typical stop consonant releases. The MATLAB model calculated the pressure resulting from a labial stop release, as measured at a distance of 20 cm from the lips. The model used typical values of $P_{i_{rect}}$ = 9 cm H_2O, l_c=0.1 cm, A=3 cm^2, and used a constriction whose length stays constant at 0.1 cm and whose area increases at 100 cm^2/s. These values are not experimental and are simply reasonable guesses, except for the rate of increase in the area which was determined by Fujimura [15] (though these numbers were determined for times much greater than 2 milliseconds after the release). The MATLAB results which model a rectangular constriction are displayed in Figure 2.4. The first graph, Figure 2.4a, shows the magnitude of each of the three pressure components (ignoring the d/dt term for now) corresponding to the first three terms on the right side of Equation 2.8. It can be seen from this graph that the term due to the kinetic resistance, $P_{kinetic} = \frac{\rho U^2}{2A_c^2}$, is the dominant term for about the first millisecond or so. It therefore seems reasonable to ignore the other terms in Equation 2.8 and approximate this with Equation 2.10, which results in the same conclusions as we arrived at without the detailed MATLAB model.

Figure 2.4b shows the differences between using this approximation and keeping the other terms (though we are still ignoring the effect due to the acoustic mass of the air in the constriction). We are ultimately interested in the pressure at some distance r from the lips. In order to interpret the features of the transient pulse as it relates to the sound pressure reaching the ears of the listeners, the time derivative of the volume velocity should be examined, as can be seen from Equation 2.1. This pressure was calculated at a distance of 20 cm using both the approximation and the calculated value (Figure 2.4c). The frequency content of both of these pressures is displayed in Figure 2.4d. Also displayed on this graph is a thick line corresponding to 1/f. A comparison of this line with the model shows that the frequency content of the transient burst (for labials) should fall off approximately as 1/f.

To verify the insignificance of the last term in Equation 2.8 corresponding to the acoustic mass of the air in the constriction, its value and derivative were calculated and displayed in (Figure 2.5). It is clear that this term only has an effect for time less than about 0.1 ms. Since this value is smaller than typical sampling intervals, it is clear that this term cannot have a consistent effect on the transient level, though it might occasionally affect transient amplitudes. It therefore seems reasonable to neglect this term in the development of transient theory, at least for short constrictions.

2.5 MATLAB Model - Circular Constriction

A similar model for calculating the pressure was created assuming a circular constriction. The pressure is again composed of four components as shown in the following equation:

$$P_{i_{circ}} = \frac{\rho c}{A} U + \frac{128\mu}{\pi D^4} U + \frac{\rho U^2}{2A_c^2} + \frac{d}{dt}\left(\frac{\rho l_c U}{A_c}\right). \tag{2.12}$$

MATLAB was used to model this equation and calculate the pressure components (Figure 2.6a), volume velocity (Figure 2.6b), pressure at a distance of 20 cm (Figure 2.6c), and spectral content of the pressure (Figure 2.6d). Again the kinetic resis-

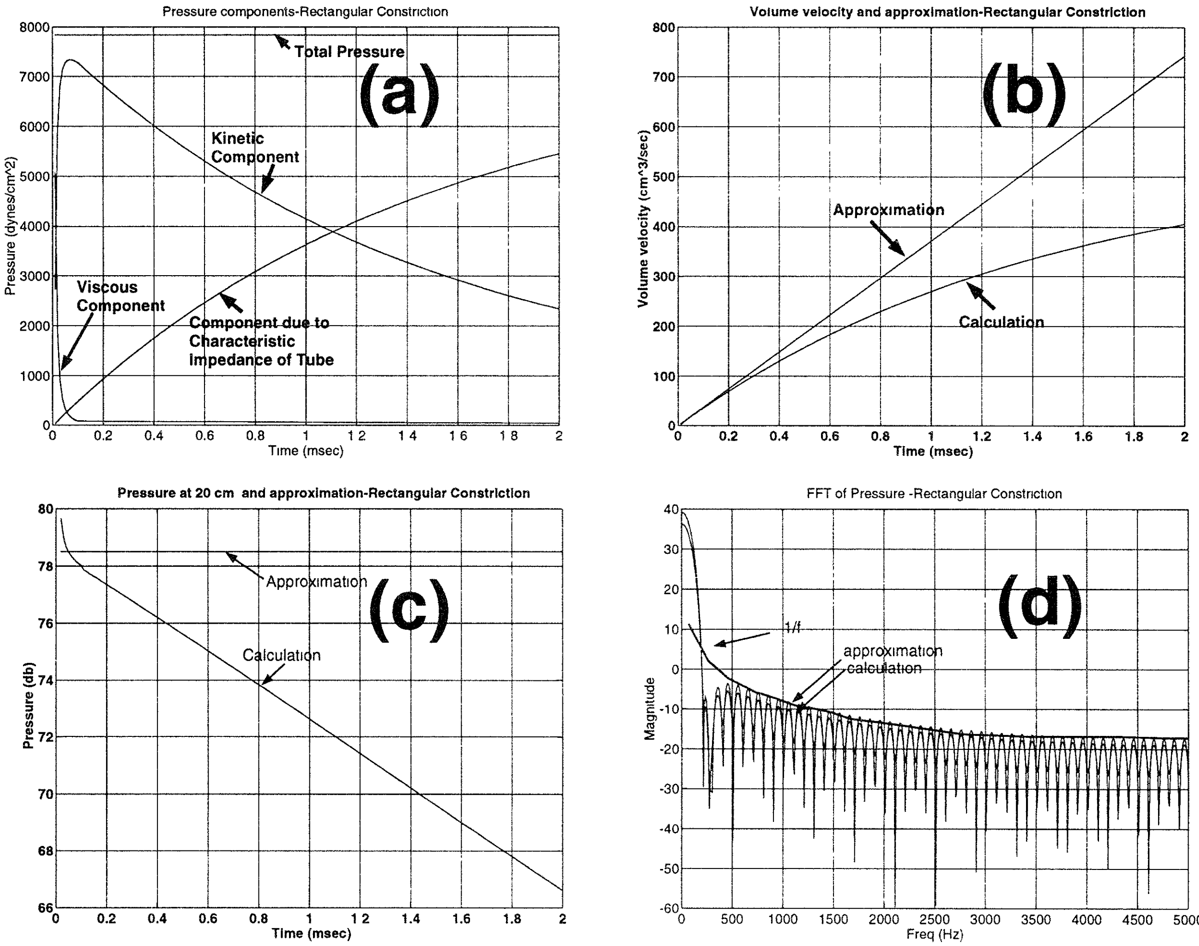

Figure 2-4: Transient pressure components (a), volume velocity (b), pressure at 20 cm (c), and spectra (d) of rectangular constriction model.

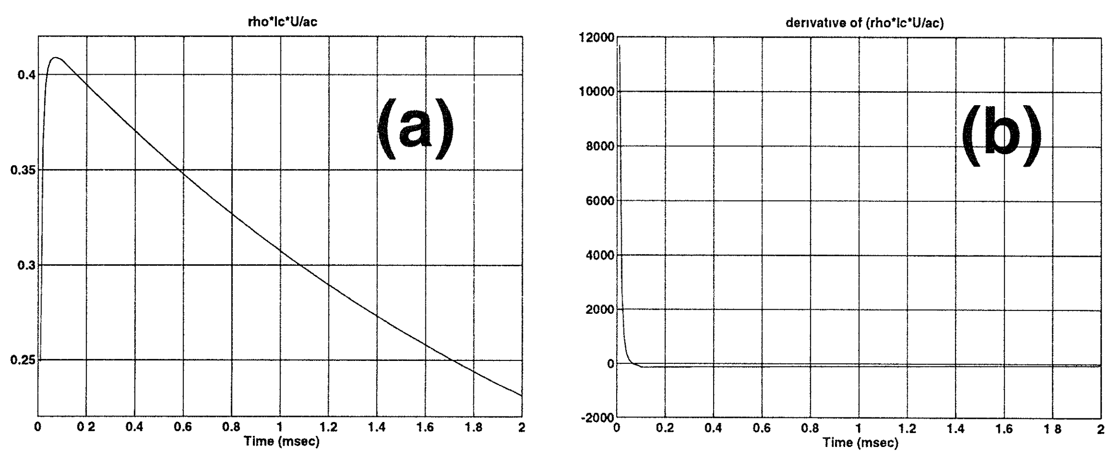

Figure 2-5: Acoustic mass term (a) and its derivative (b) for rectangular constriction model.

tance term dominates, and the same approximation can be made which neglects the remaining terms leaving us with the approximation

$$P_{i_{circ}} \approx \frac{\rho U^2}{2A_c^2} \tag{2.13}$$

which is identical to Equation 2.10. Since this is the same equation as for a rectangular opening, the analysis and theory from this point on is exactly the same as described for the rectangular opening. Therefore, the actual shape of the constriction does not seem to be important in the generation of the transient burst except during the first 0.1 to 0.2 ms after the release of a stop consonant.

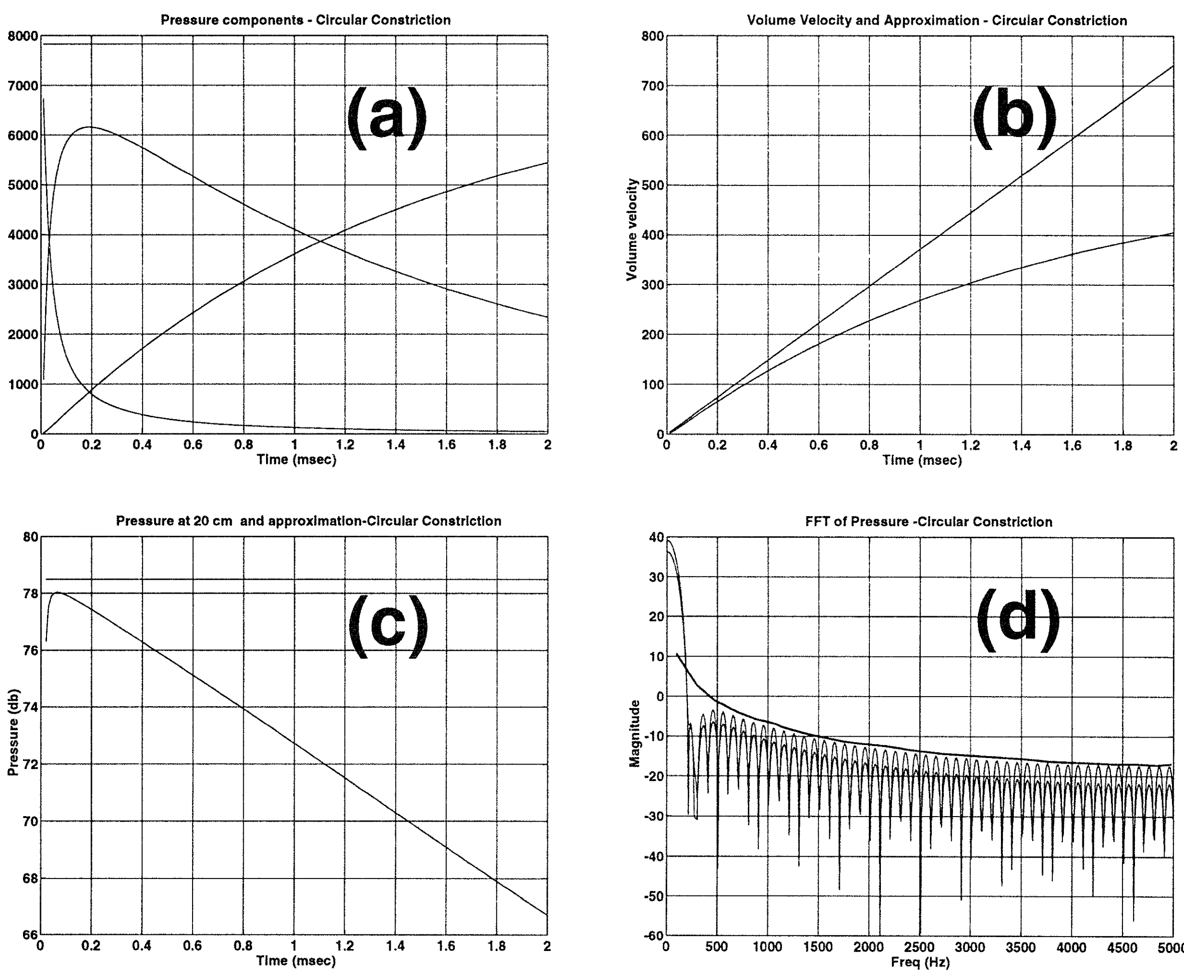

Figure 2-6: Transient pressure components (a), volume velocity (b), pressure at 20 cm (c), and spectra (d) of circular constriction model.

The conclusion that can be drawn from this model is that the spectrum for a transient burst (radiated directly into air without filtering) should fall off approximately as 1/f. The magnitude of the transient is linearly dependent on the cross-sectional area of the opening, $A_c(t)$, and is proportional to the square root of the intraoral pressure P_i. Since there are no articulation data which can determine the cross-sectional areas (and to a lesser extent intraoral pressures), the actual magnitude can not be predicted very accurately, probably to within only 5 to 10 dB.

2.6 Filtering of Transient due to Anterior Cavity

For labial stop consonants the transient burst is radiated directly from the lips; however, alveolar and velar bursts are filtered by the cavity which is anterior to the constriction. For alveolar stop consonants (excluding affricates) the length of this cavity is estimated to be in the range of 1.5 cm to 2.5 cm for an adult speaker. The transfer function between the volume-velocity source and mouth opening for this cavity (using a male length of 2 cm) will have a peak at about 4.5 kHz as shown in Figure 2-8. The affricate transient will be filtered by a cavity similar to the one shown in Figure 2-7. The total length of the front cavity is approximately 2.5 cm, including the length of the sublingual cavity, leading to a resonance of about 3500 Hz [3]. In addition to this resonance there is a decrease in the amplitude of the transfer function at around 4.5 kHz due to the zero introduced by the side branch of approximately 2 cm, also shown in Figure2-8. The front cavity for a velar stop consonant is much longer, somewhere in the range of 5-6 cm (though the length of this cavity is highly dependent on whether the following vowel is front or back). This transfer function will have a peak at a lower frequency around 1.7 kHz, and also one at about 5.0 kHz. A sketch of this transfer function is also shown in Figure 2-8.

The velar stop consonant differs from the alveolar or labial stop consonant in two primary ways which lead to a significantly different transient burst. The first difference is that the length of the constriction made with the tongue body is much greater than constrictions made with the lips or the tongue blade. The second difference is that the rate of increase of the cross-sectional area is slower, typically about 25 cm^2/s

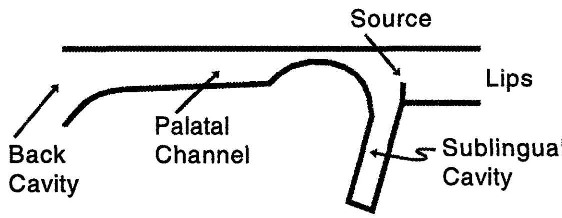

Figure 2-7: Approximate cavity for affricate transient filtering.

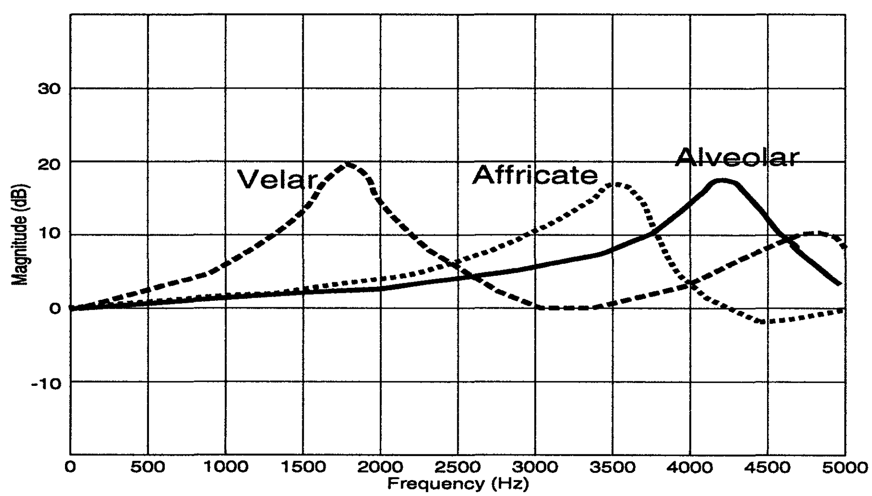

Figure 2-8: Transfer functions between constriction and mouth opening for alveolar and velar stop consonant releases.(Stevens, in preparation)

(using the data collected by Fujimura [15] and Kent and Moll [16]).

The slower rate of increase in the cross-sectional area leads to a initial transient burst which is weaker in amplitude than labial or alveolar transient bursts. As in labial and alveolar stop consonants, the spectrum of velar transient bursts is still directly proportional to the rate of change in area of the constriction, $A_c(t)$, which in turn increases linearly with time. However when the proportionality constant is decreased by a factor of 4, the spectrum will have an amplitude which is also reduced by a factor of 4 (which is -12 dB). Therefore the amplitude of the transient source for velar stop consonants should be approximately 12 dB weaker than the source for alveolar or labial stop consonants. However the resonances of the anterior cavity can increase the magnitude of the output, as shown in Figure 2-8. The overall effect, therefore, is that velar transient bursts which are measured at a distance should be approximately 8 dB stronger at the frequency of the formant than the corresponding magnitude of labial transient bursts.

The longer length and slower increase in the cross-sectional area of the constriction in velar stops and affricates are responsible for permitting a series of transient pulses to be generated at the release. The mechanism by which multiple transient pulses are generated is similar to the mechanism of vibration of the vocal folds. The pressure behind the constriction causes the most anterior part of the closure to separate. This separation decreases the pressure within the constriction causing the walls to displace back to their closed position which in turn causes the flow of air to be cut off. This process then repeats itself and, depending on the rate of opening, two or more oscillations might occur before the separation becomes too large to permit complete closure. The spacing between multiple transients is on the order of a few milliseconds.

Chapter 3

Acoustic Analysis

3.1 Data Collection

A number of naturally-produced consonant-vowel syllables were analyzed to determine the properties of transients. Ten speakers, 5 male and 5 female, were recorded in a soundproof room using an Electro-Voice D054 microphone and Nakamichi LX-5 tape recorder. Before each speaker commenced, a pure 500 Hz sinusoid was amplified and the level at the microphone was recorded by a sound level meter which was calibrated to measure sound levels relative to 0.0002 dyne/cm^2. This measurement and the attempt to set the microphone at a distance of 20 cm from the speakers' mouth permitted the calculation of the absolute sound pressure at 20 cm, and hence the magnitude of the transient, which will be explained below. The recorded signals were later passed through an antialiasing filter with cutoff at 4800 Hz and sampled at a rate of 10 kHz using a Decstation. The resulting files therefore consisted of a 12 bit digital number which was recorded every 0.1 ms.

The utterances were 16 real words spoken in isolation. The list was randomized and repeated six times. Half of the words began with one of the 8 stop consonants (including affricates) and were followed by the vowel /ɑ/. The other half of the words began with one of the 8 stop consonants and were followed by the vowel /i/. All words ended with a stop consonant. (See Table 3.1 for a list of the words.) Two additional words, "judge" and "church", were added to the end of each list, creating a list of 18 words. These last two words were not analyzed but acted as a buffer. People have a tendency to say the last word (or two) in a list differently, so these two additional words prevented this phenomenon from affecting the words that were to be analyzed. The intention was to collect data from five repetitions, but a sixth repetition of the list served as a backup in case there was a problem with one of the lists, such as a speaker mispronouncing a word. No such problems were encountered, so the sixth list was never used.

pot	**beat**	**dock**	**cot**	**teach**	**top**
peak	**geek**	**bought**	**keep**	**got**	**jeep**
chalk	**deep**	**job**	**cheap**		

Table 3.1: Words used to analyze transient bursts.

Once the utterances were digitized they were separated according to word. Sixteen files were created for each of the ten speakers, each file containing five repetitions of the same word. These files were then converted to an ASCII file which could be read by MATLAB. (See program in Appendix B.)

The first task in analysis was to convert the amplitude of each speaker to the same absolute level. The digitized pure 500 Hz tone was measured and converted to a dB scale (relative to 0.0002 dyne/cm^2). This number was compared to the value measured by the sound level meter at the microphone. All waveforms from that speaker were then scaled so that the rms amplitude of the pure tone matched the rms amplitude recorded by the sound level meter. After using this procedure on the files from each of the ten speakers, all the files contained waveforms whose amplitudes corresponded to the same magnitude scale, which was calibrated to absolute sound pressure.

3.2 Amplitude Analysis

The first analysis of these waveforms resulted in the tabulation of the amplitude of each of the transient bursts. The amplitude of a transient is not well defined since the transient does not always have a consistent shape predicted by theory. It is extremely difficult to isolate the transient from background noise and also from the following frication or aspiration noise, so it seems reasonable to anchor the ambiguous transient amplitude to some well defined parameter such as vowel amplitude. The amplitude of the following vowel is much easier to determine, and was measured by calculating the rms value of an integer number of glottal pulses during the steady state portion of the vowel. The rms value is a number which is obtained by taking the square root of the average of the square of all the sample values over a particular interval. An example of the steady state interval of the vowel is labeled 'vowel bin' in Figure 3.1. The absolute maximum amplitude of the steady-state vowel is also measured (labeled 'peak vowel magnitude' in Figure 3.1) for later use. Figure 3.2 shows an example of the measurement of the maximum absolute amplitude of a transient burst. The rms value for the transient is difficult to calculate due to the inability to always separate the transient from the surrounding noise. Therefore, another statistic was chosen. The 'rms' value of the transient was calculated by taking the rms value of the vowel and multiplying this by the ratio of the absolute peak transient magnitude to absolute peak vowel magnitude. The transient magnitude in dB is calculated by taking $20 * log_{10}$ of the above number. The corresponding formula is

$$Transient\ Magnitude_{dB} = 20 * log_{10}\left(Vowel_{rms} * \frac{Transient_{peak}}{Vowel_{peak}}\right). \qquad (3.1)$$

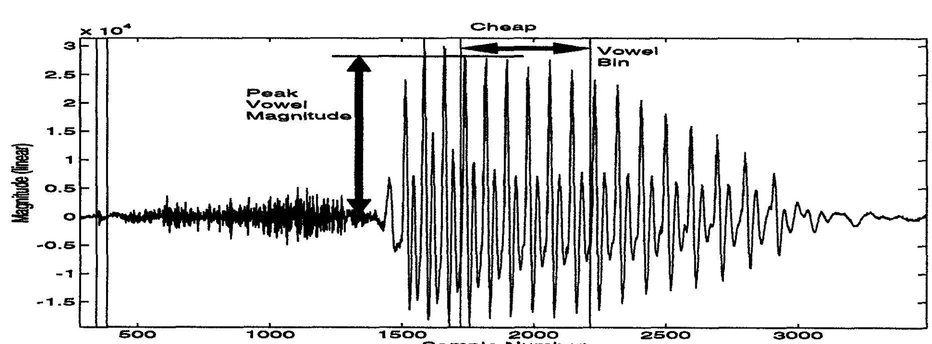

Figure 3-1: Measurement of vowel magnitude.

To implement the amplitude finding procedure described above, a program was

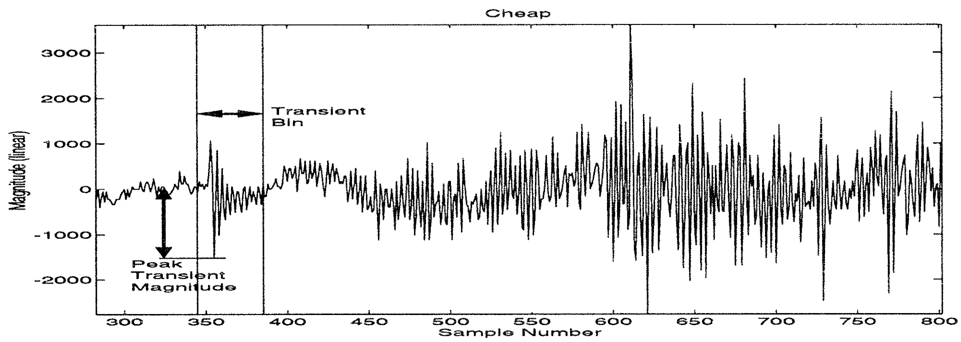

Figure 3-2: Measurement of transient magnitude.

written in MATLAB which allows users to create a histogram of the amplitudes of the vowels and transients in any arbitrary number of waveforms. Unfortunately it was beyond the scope of this research to write a program which automatically detects the times where the transients and vowels begin and end. This task of bin marking was left to humans so a separate program was written which allows the user to maneuver the mouse and place marks at arbitrary places in the waveform. The program, named analysis3.m (see Appendix B), displays an entire waveform in one window and allows the user to zoom in on any section of the waveform and display this section in another window. The user presses a mouse button at the place where he or she wants the zoom window to start and releases the button where the zoom window should end. Once this operation has been performed, the window right above the main waveform displays the expanded view of the desired waveform section. The user can then push the button at the beginning of the transient (or vowel) and release it at the end of the transient (or vowel) which is displayed in the zoom window. The final action will segment the waveform by placing vertical lines on the waveform thereby allowing programs to analyze the transients and vowels.

Once the transients and steady state portions of the waveform have been marked, tabulation of the vowel and transient magnitudes is simple. The program transient_magnitude.m (see Appendix B) calculates the magnitudes of the vowels and transients. Five graphs are displayed for each of the sixteen words (Figures 3.3 to 3.18). The top graph shows the transient and vowel magnitudes for each of the 50 utterances (five repetitions by ten speakers). This graph allows comparisons to be made between vowel and transient magnitudes for each of the five utterances by the same speaker, and also comparison of utterances by different speakers. The middle two histograms show the rms amplitude of the vowel and peak amplitude of the transient. Note that all information about the correlation of transient amplitudes to vowel amplitudes has been lost in these histograms. The lower two histograms try to recapture this lost information by displaying the peak vowel amplitudes and the level of the maximum transient amplitude relative to the peak magnitude of the following vowel (as explained in Equation 3.1).

The magnitude graphs show that there is considerable variability in transient burst magnitudes, though their amplitudes do tend to follow a normal, or bell-shaped curve. The average value of the transient magnitude is typically around 15-20 dB weaker than the following vowel for stops and 20-25 dB weaker than the following vowel for affricates. There seems to be little variability with voicing or place of articulation, though the velar stop releases do produce a slightly weaker transient, probably due to the slower movement of the tongue body as compared with the tongue blade or lips.

Another conclusion that can be drawn from the magnitude graphs is that there does appear to be a correlation between the magnitude of the transient and the magnitude of the following vowel. This can be noticed by looking at the top graph in each of these figures and comparing the peaks and valleys of the vowels with the peaks and valleys of the transients. There tends to be a peak or valley in the vowel amplitude when there is a peak or valley in the transient amplitude. A more quantitative analysis can be performed by noticing that the standard deviation of the absolute transient peak magnitude is slightly more than the standard deviation of the transient magnitude relative to the following vowel. This observation agrees with the theory presented in Chapter 2 which predicts that the amplitude of the transient should increase with the square root of intraoral pressure. The vowel amplitude could change according to subglottal pressure, but the vowel amplitude is also a function of tongue-body height. High vowels tend to be about 4-5 dB weaker than low vowels [2], and there is also a slight increase in amplitude of front vowels. The relative transient amplitude graphs do not take these vowel effects into account, so the relative standard deviation might be slightly lower if these effects are taken into consideration.

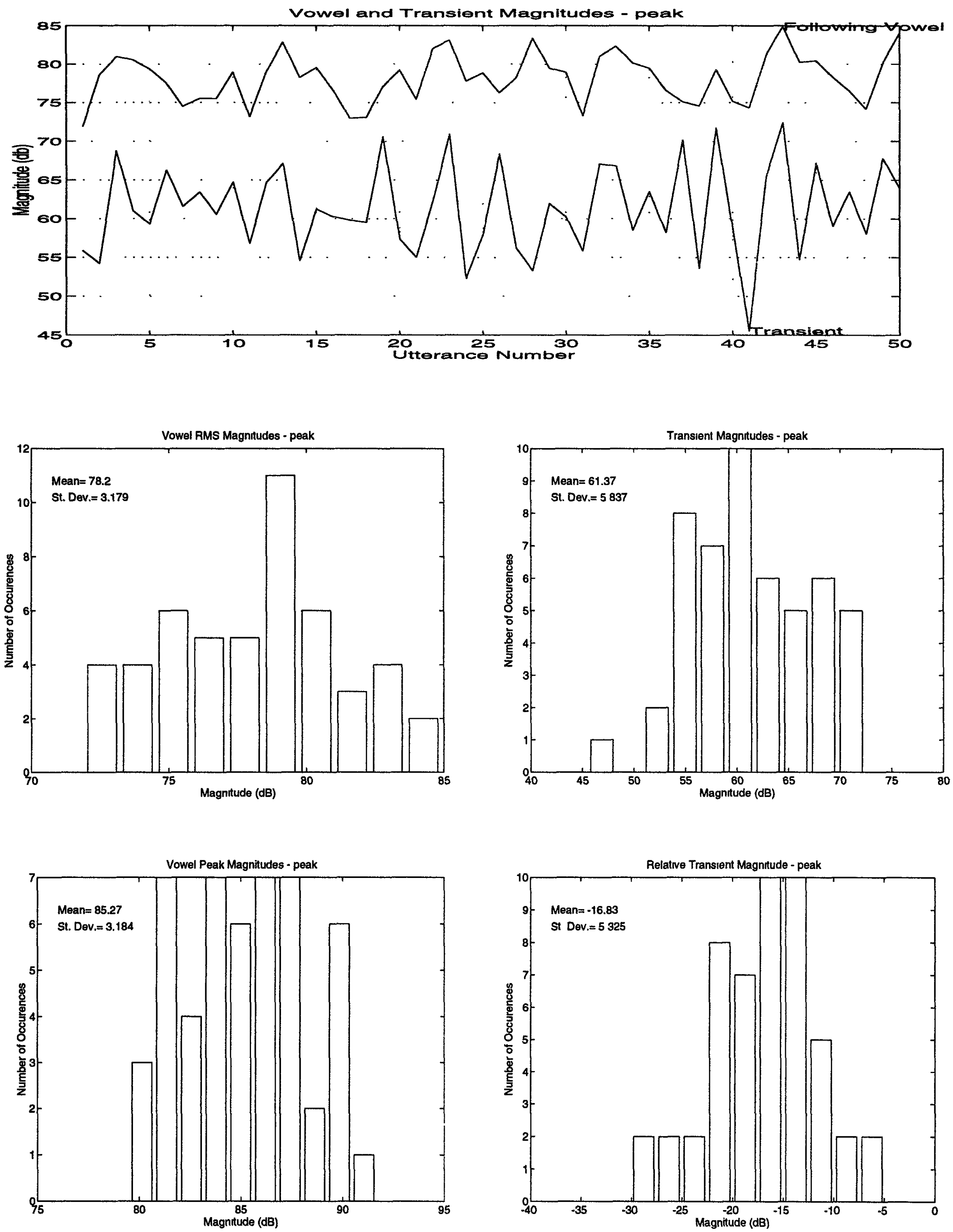

Figure 3-3: Transient and Vowel magnitudes - 'peak'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

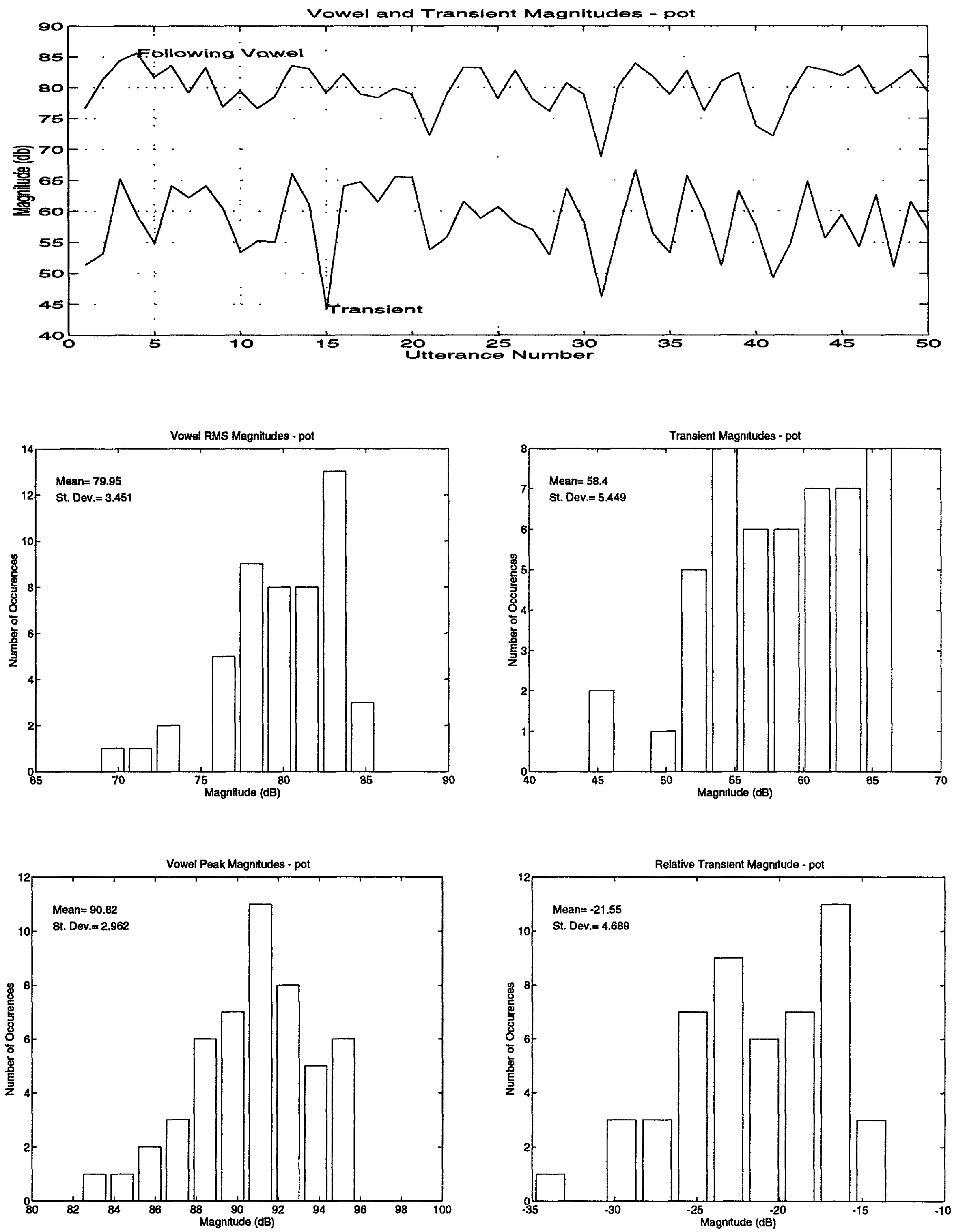

Figure 3-4: Transient and Vowel magnitudes - 'pot'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

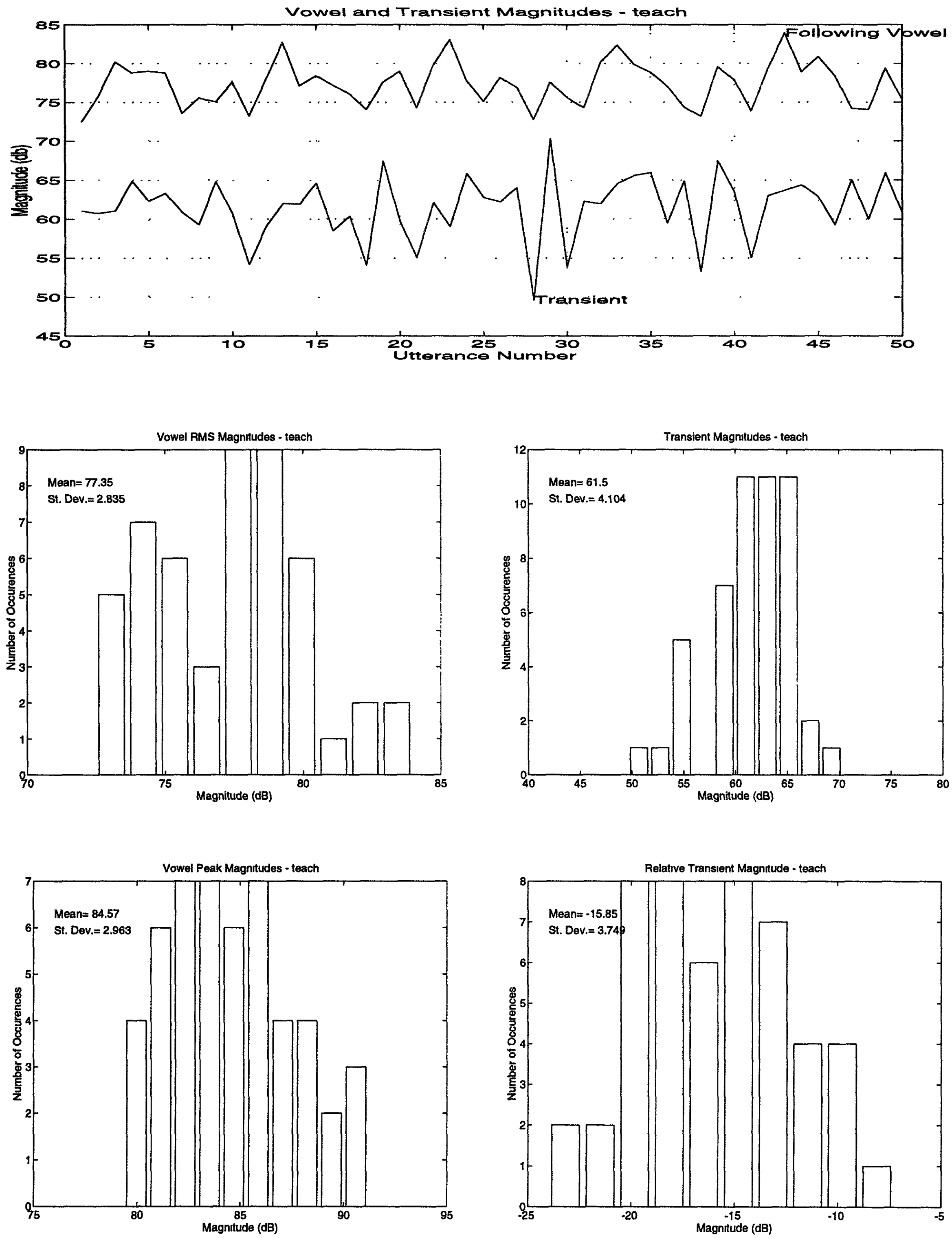

Figure 3-5: Transient and Vowel magnitudes - 'teach'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

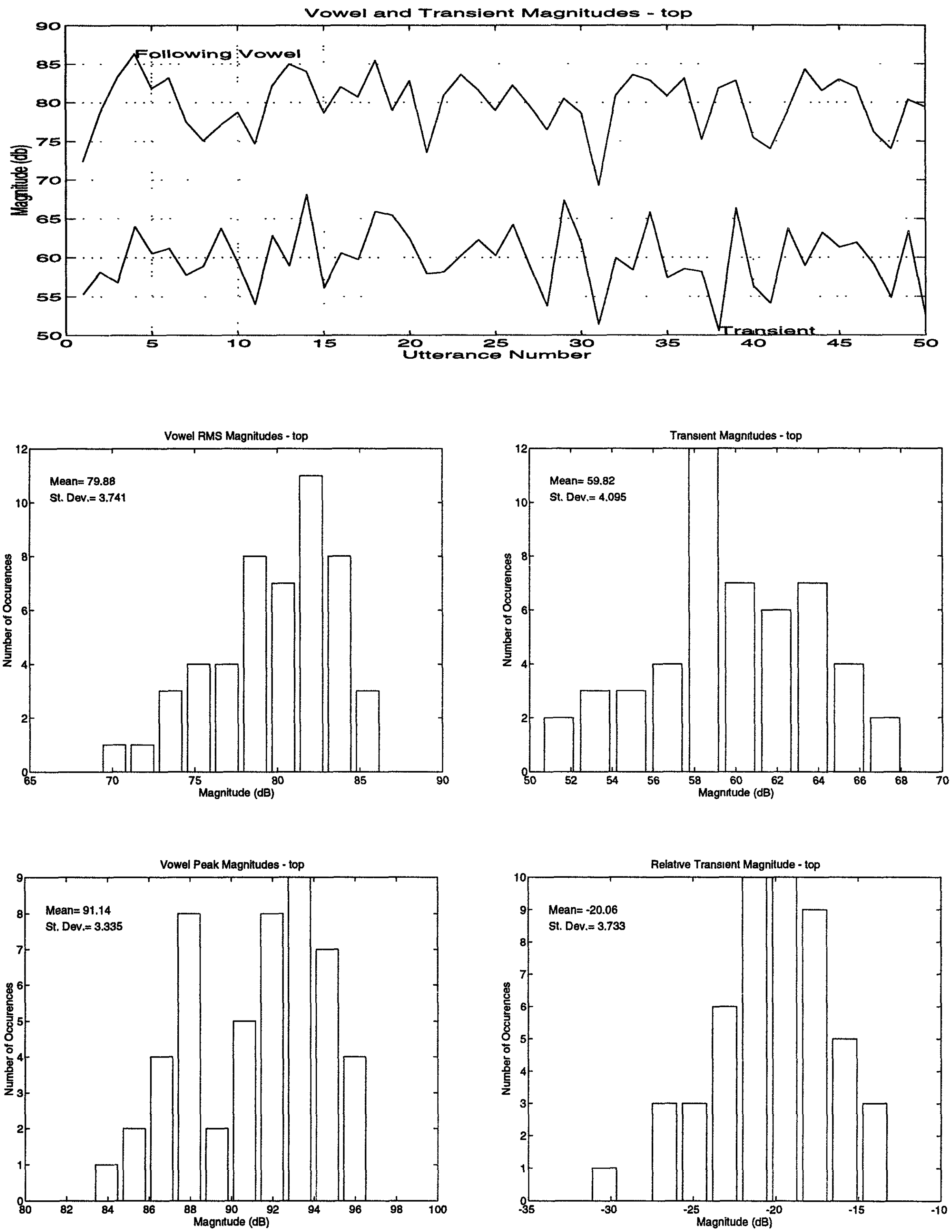

Figure 3-6: Transient and Vowel magnitudes - 'top'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

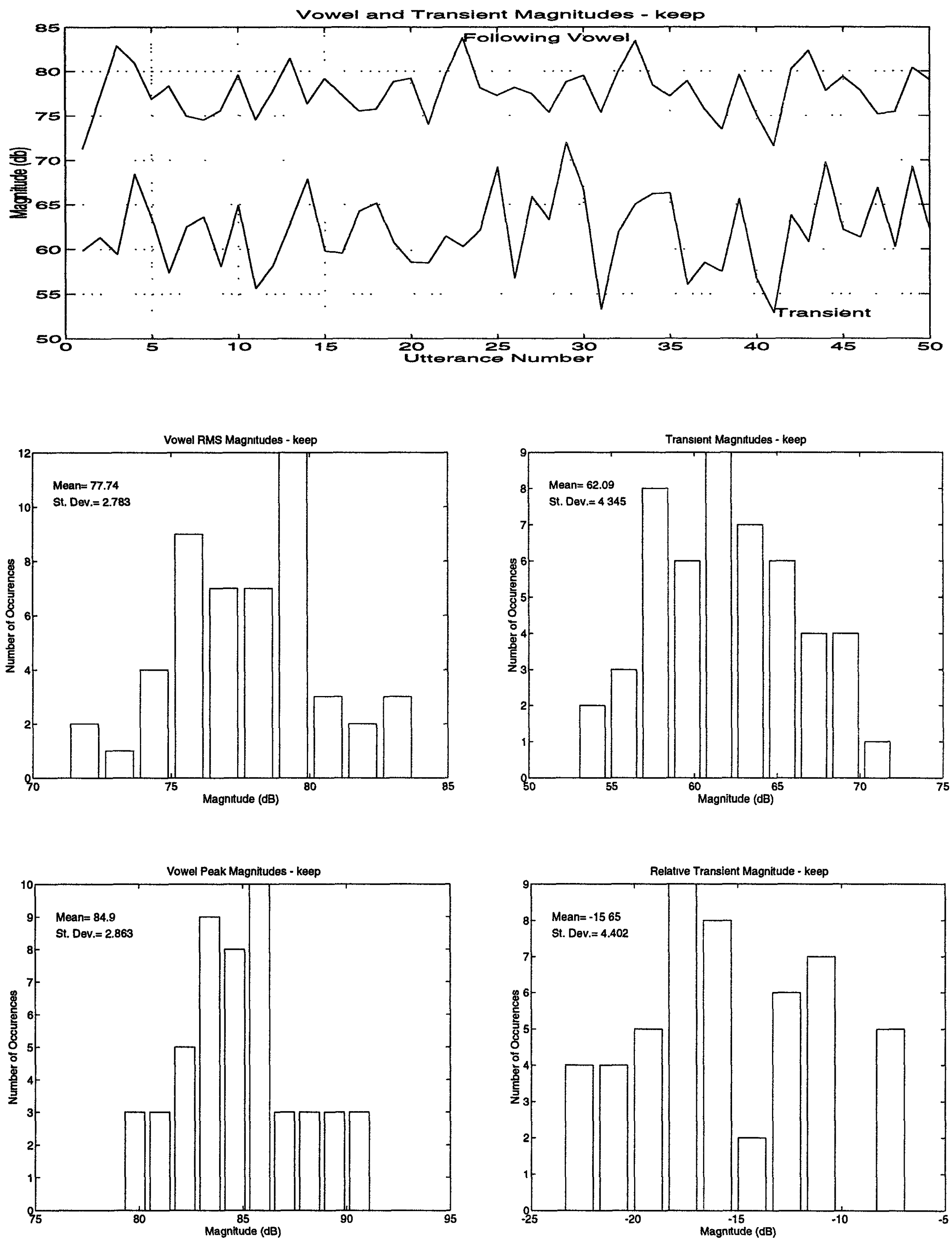

Figure 3-7: Transient and Vowel magnitudes - 'keep'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

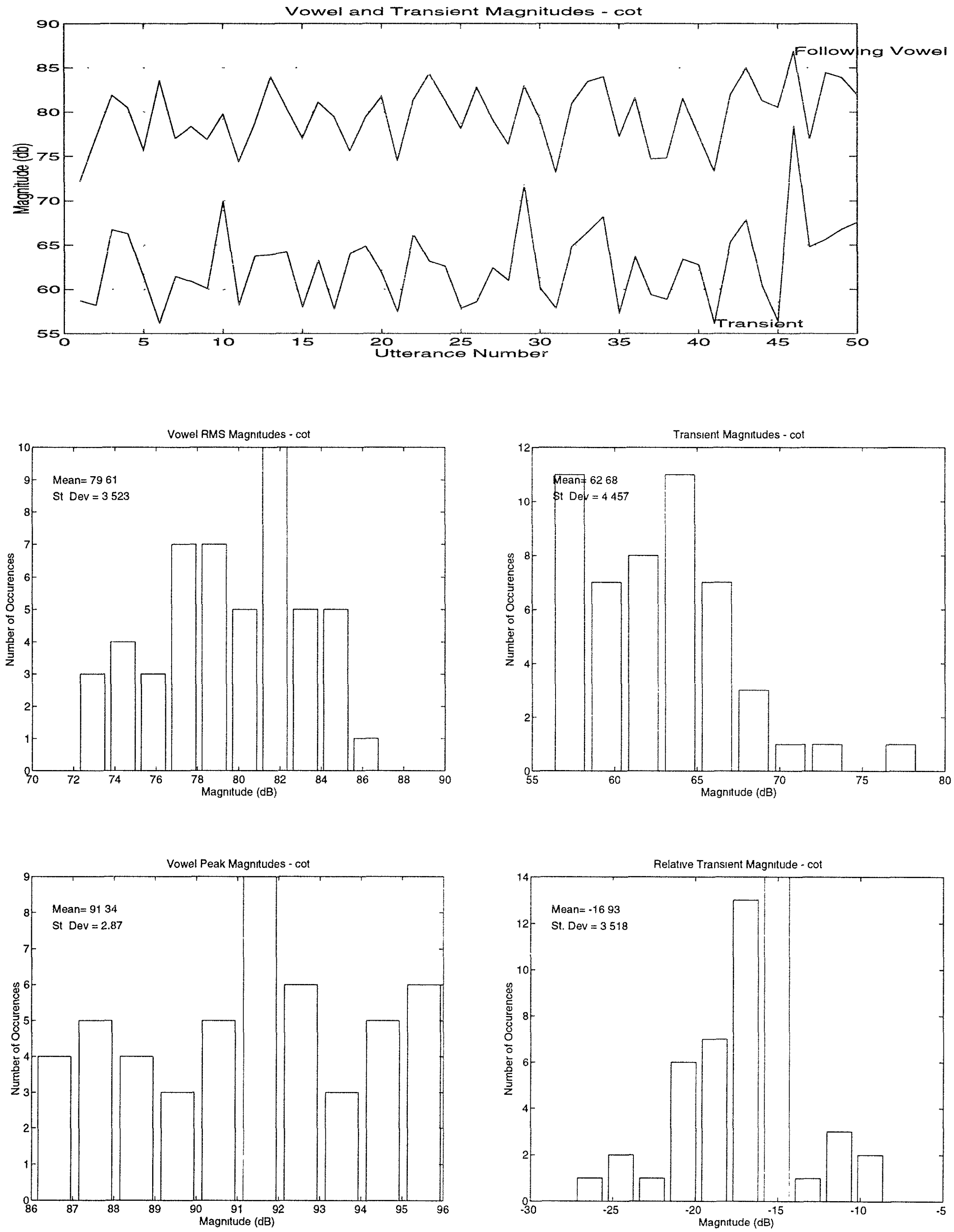

Figure 3-8: Transient and Vowel magnitudes - 'cot'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

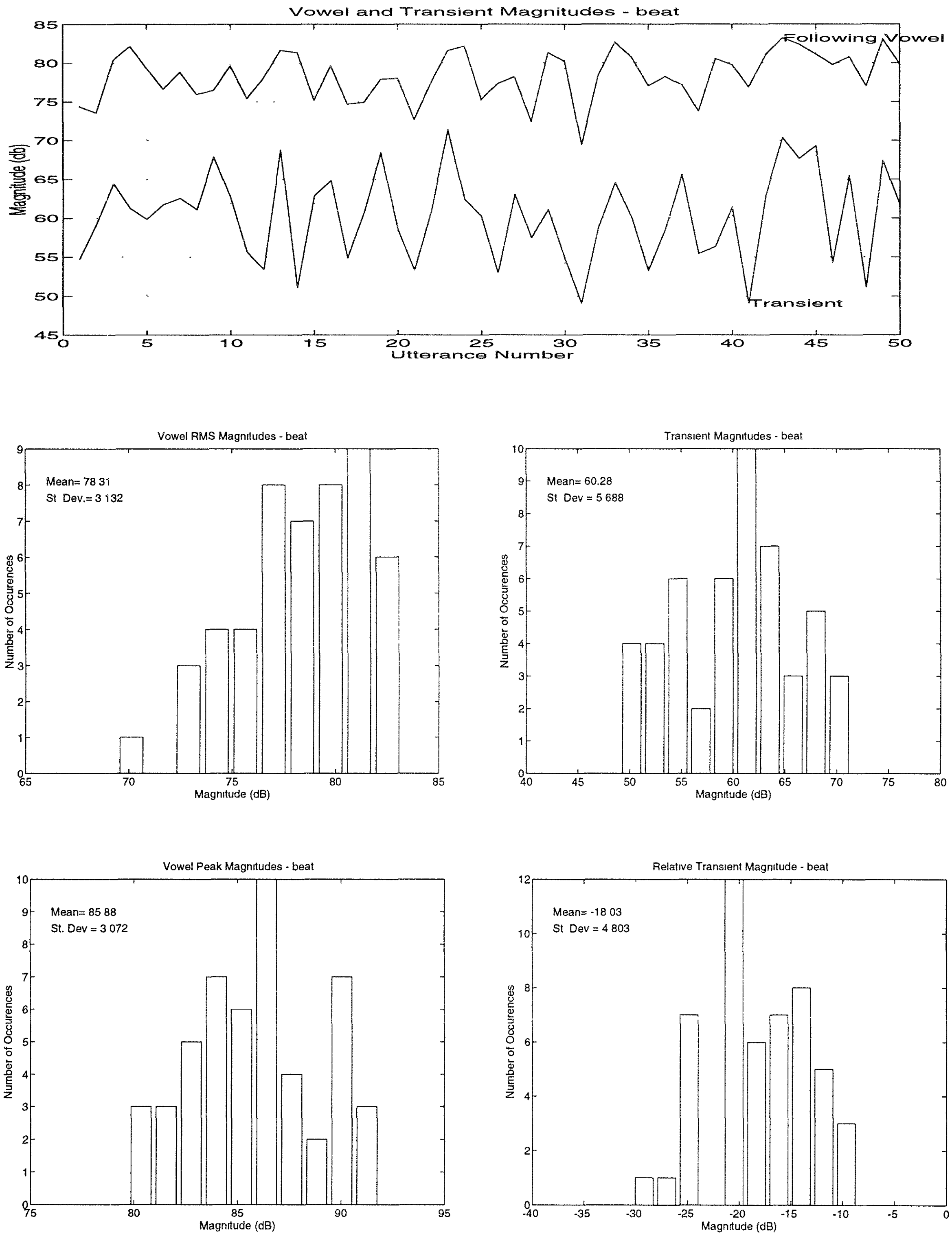

Figure 3-9: Transient and Vowel magnitudes - 'beat'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

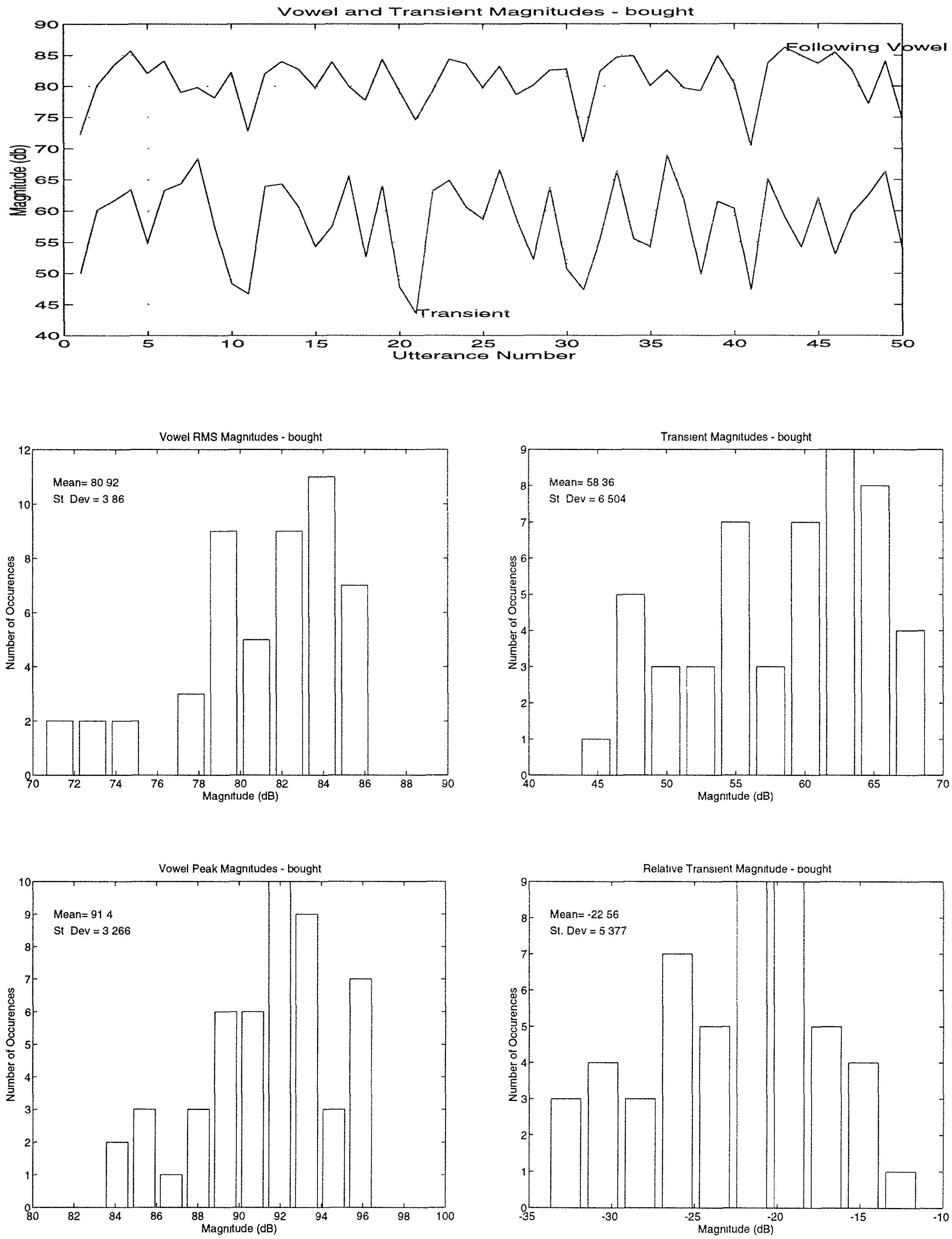

Figure 3-10: Transient and Vowel magnitudes - 'bought'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

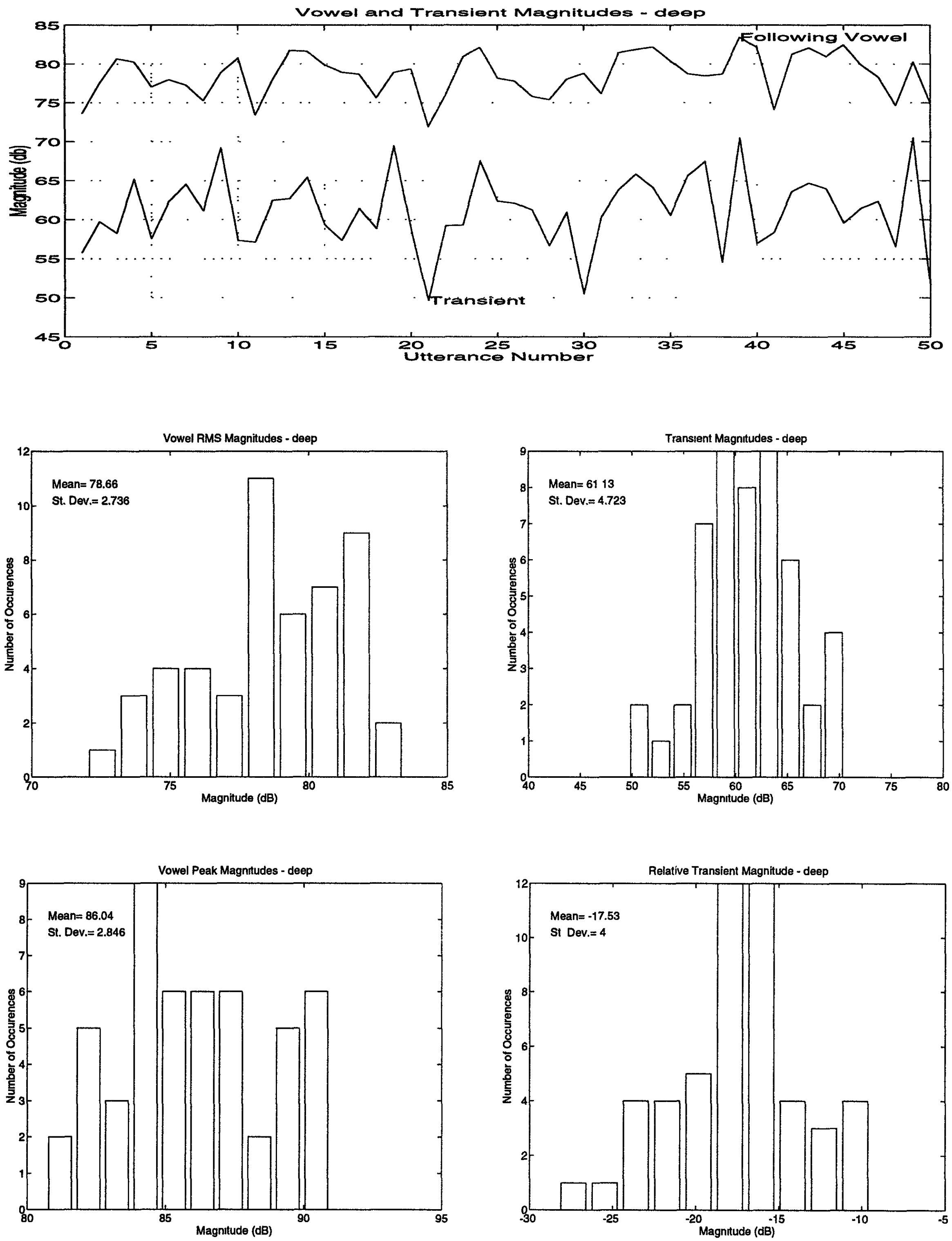

Figure 3-11: Transient and Vowel magnitudes - 'deep'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

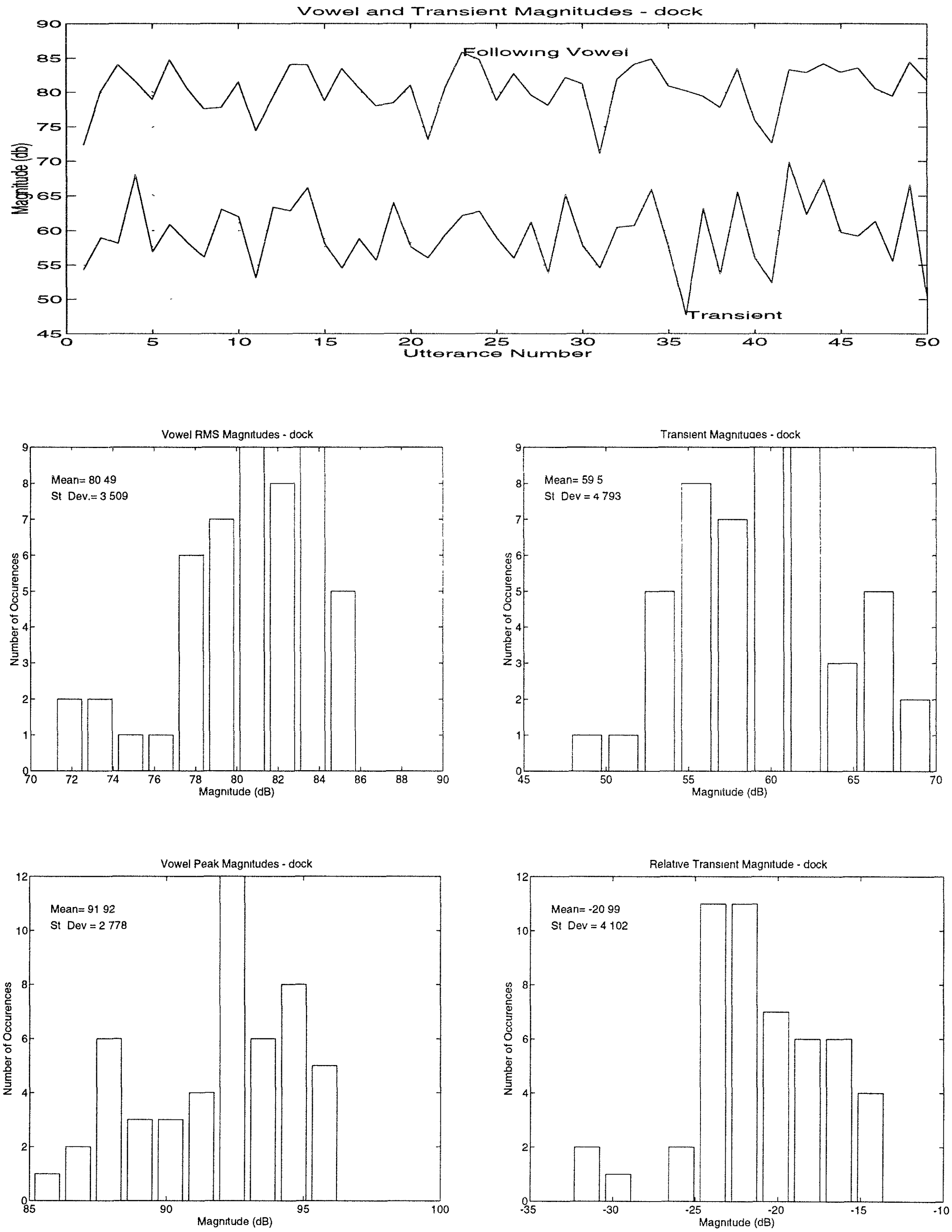

Figure 3-12: Transient and Vowel magnitudes - 'dock'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

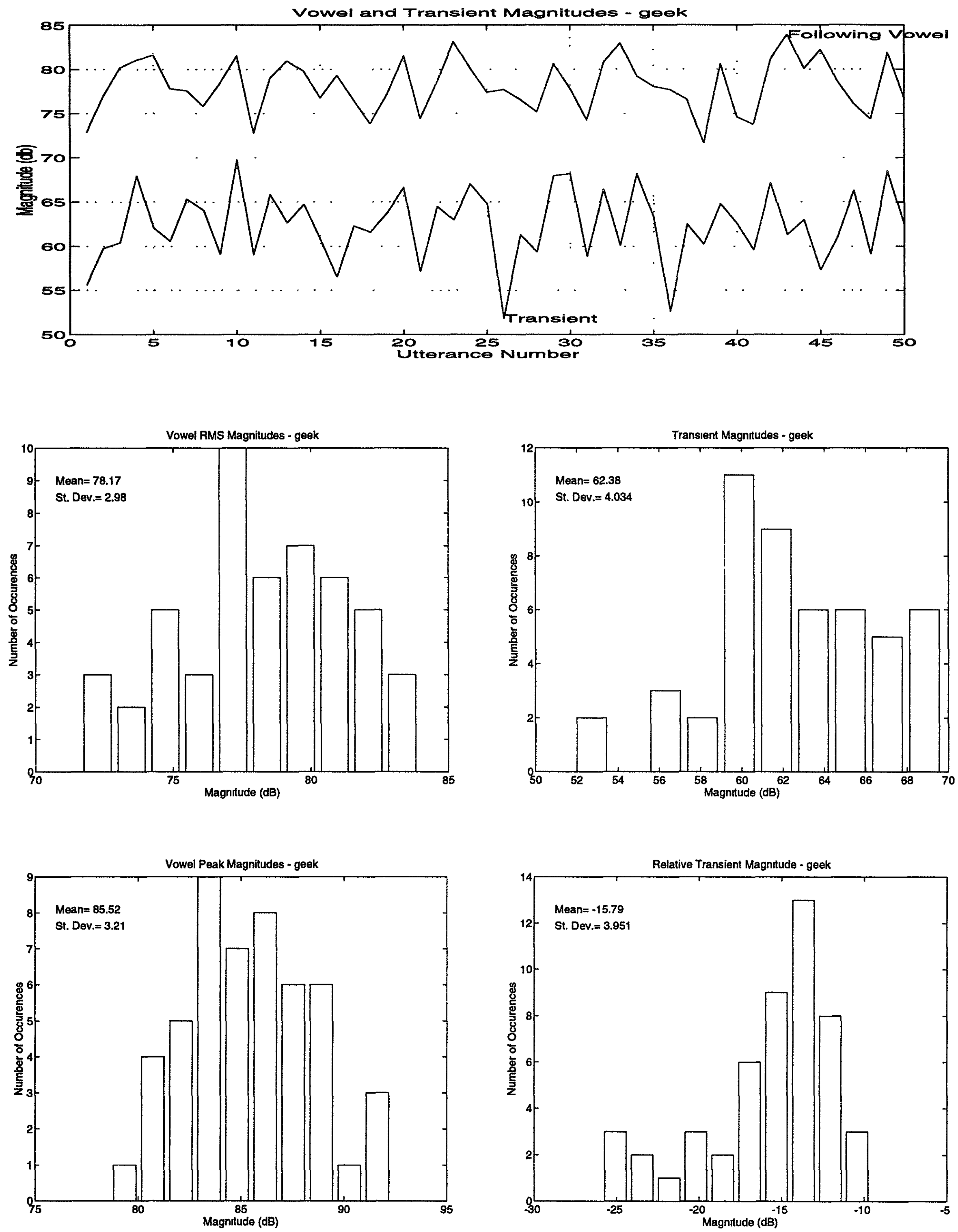

Figure 3-13: Transient and Vowel magnitudes - 'geek'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

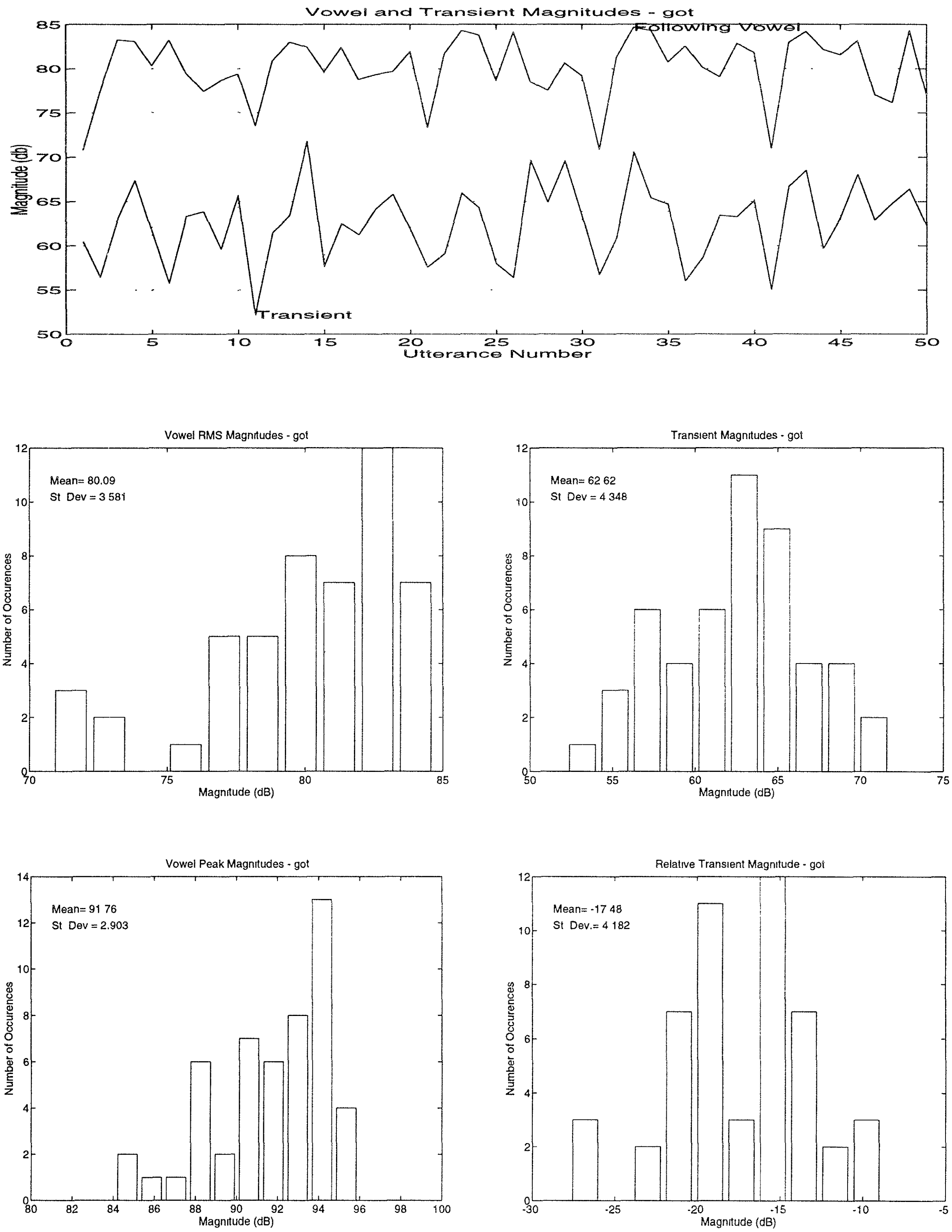

Figure 3-14: Transient and Vowel magnitudes - 'got'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

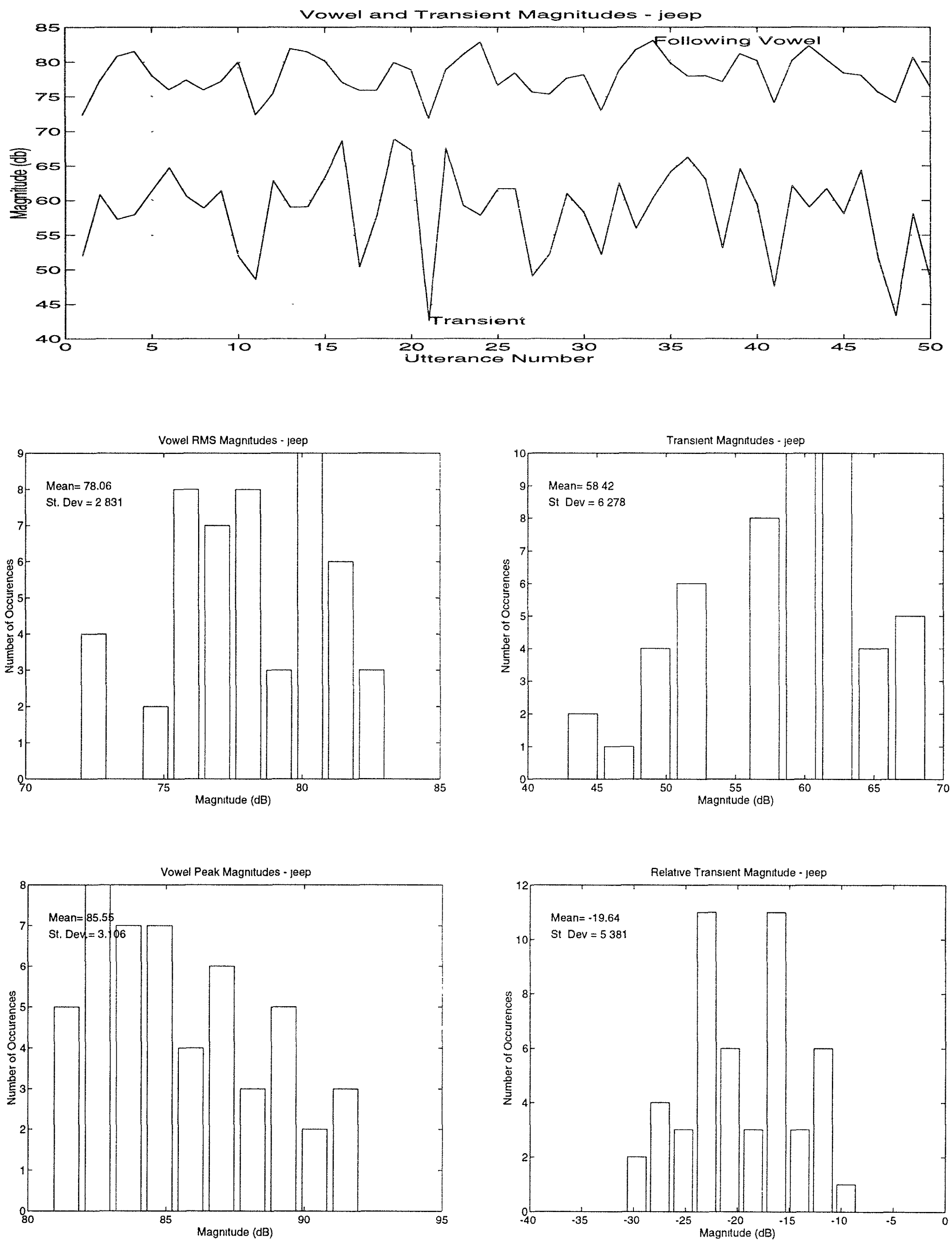

Figure 3-15: Transient and Vowel magnitudes - 'jeep'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

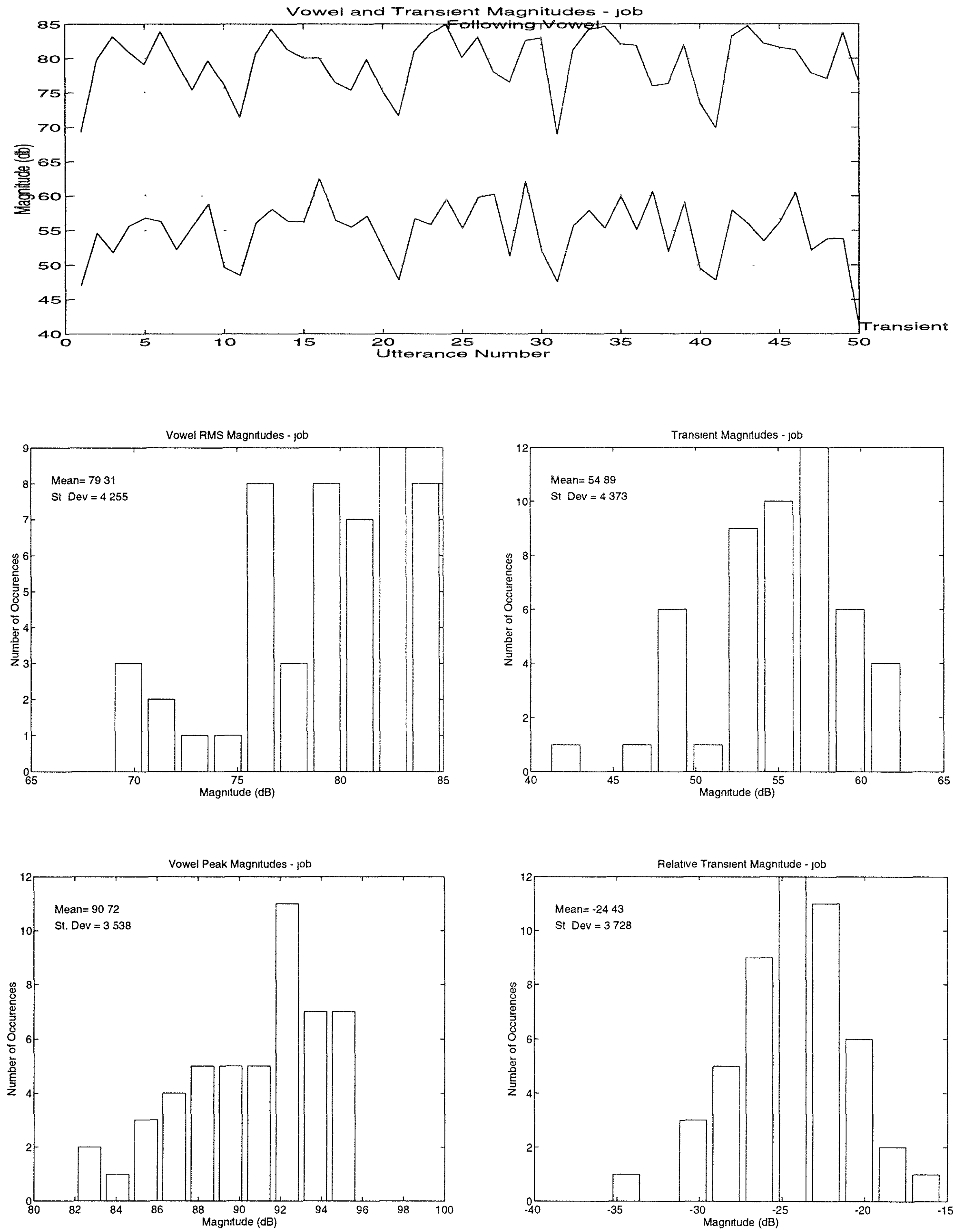

Figure 3-16: Transient and Vowel magnitudes - 'job'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

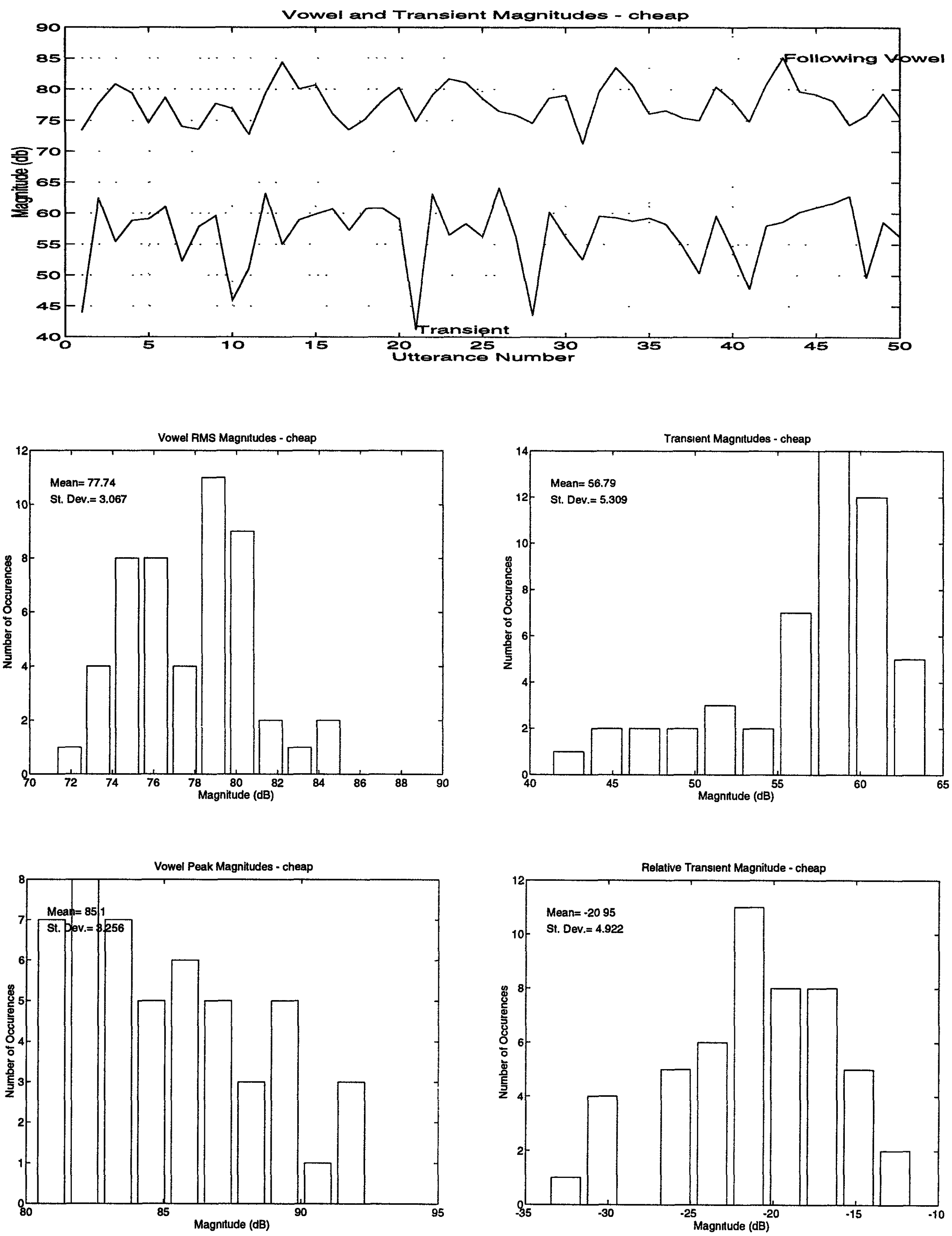

Figure 3-17: Transient and Vowel magnitudes - 'cheap'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

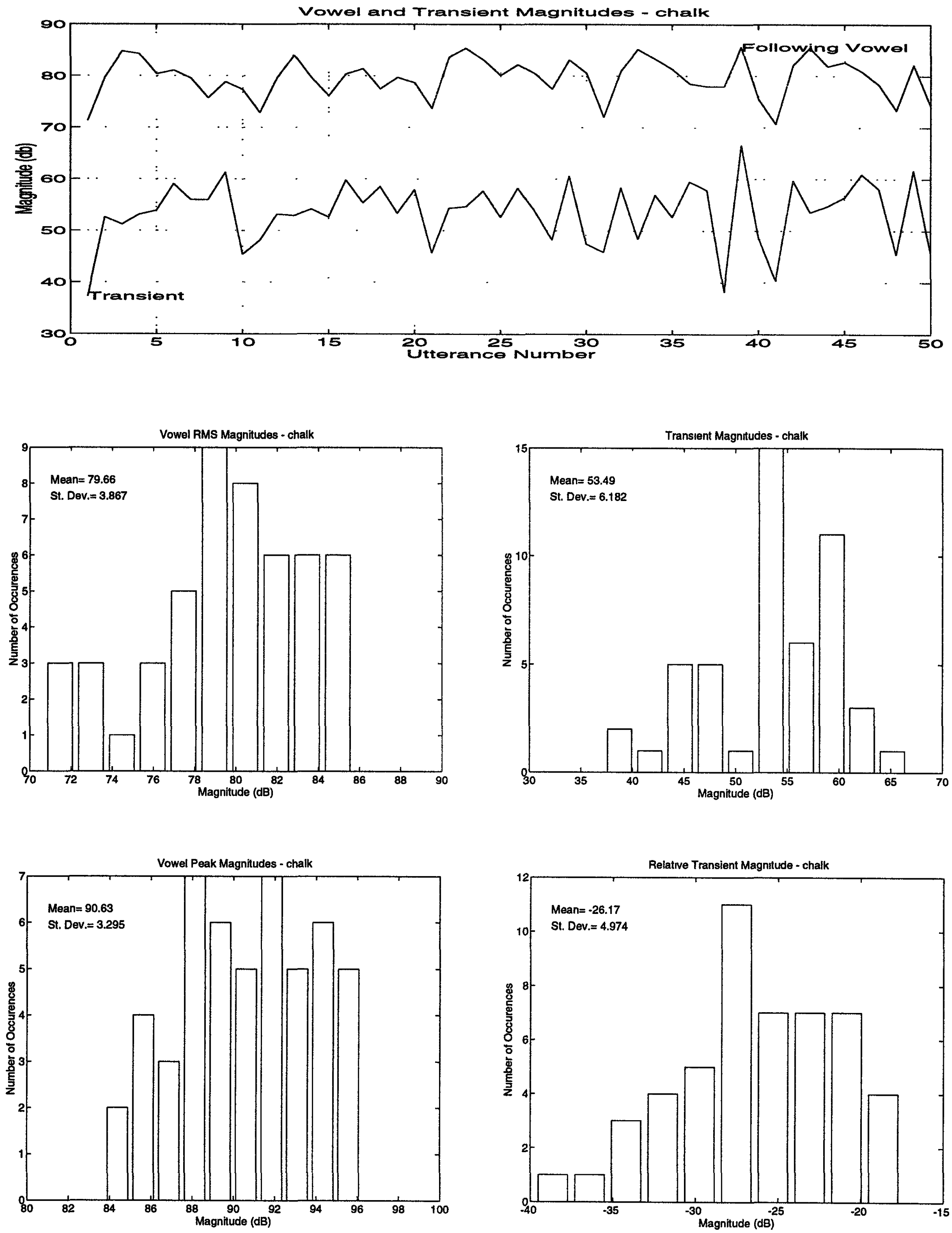

Figure 3-18: Transient and Vowel magnitudes - 'chalk'. Top graph - transient and vowel magnitudes for each utterance. Left histograms - vowel rms (top) and peak (bottom) magnitudes. Right histograms - transient peak magnitudes (top) and transient magnitudes relative to the following vowel (bottom).

3.3 Spectral Analysis

The frequency content of the transient pulses was also tabulated. A 6.4 ms DFT was taken using a Hamming window which was centered at the absolute maximum in the transient waveform. The spectrum was then smoothed by taking a moving average within a band of about 300 Hz. This process was conducted for each of the five repetitions of each stop consonant. The spectra for each stop consonant were then averaged for each of the five male or five female speakers. The final result for each gender is an average of 25 spectra (five repetitions by five speakers). A problem which arises due to this averaging is the effect of bandwidth widening. If two speakers have a slightly different place of articulation the spectra for the stop releases will have different center frequencies for the formants, though they could have exactly the same bandwidths. If these two spectra are averaged, then the resulting spectrum will most likely have a single prominence with center frequency somewhere between the two original center frequencies. The problem is that the bandwidth of the formant will be greater than either of the original formants and the spectrum amplitude at the peak will be reduced. Therefore the results of the averaging should only be considered as general areas of formant peaks. The bandwidth of any single transient should not be expected to have the same bandwidth as the average results.

The following pages show the tabulation of transient and vowel spectra. There are four graphs for each stop consonant: one for male speakers and vowel /ɑ/, one for female speakers and vowel /ɑ/, one for male speakers and vowel /i/, and one for female speakers and vowel /i/. The pages are ordered according to place of articulation (starting with the most anterior constriction), and alternate between voiced and voiceless consonants. The stops are presented first followed by the affricates. Superimposed on each of the graphs are two lines. The steeper sloped curve is the plot of 1/frequency vs. frequency. The second, less steeply sloped curve is a plot of 1/frequency^2 vs. frequency. The magnitude scales for these curves were arbitrarily chosen.

The first observation is the difference between the spectra for the stop consonant followed by /ɑ/ and for the same stop consonant followed by /i/, particularly for labial and velar stops. This difference is seen by comparing the top two spectra on each page with the bottom two spectra. The spectra show that the production of transients is not isolated from the phonetic environment. The following vowel has two primary effects on the spectra; first, the formant peaks are slightly higher in frequency due to anticipation of a front vowel (F2 and F3 are higher for the vowel), and second, the spectra for the stop consonants which are followed by the vowel /i/ typically have an additional small peak somewhere between about 1.5 kHz and 2.0 kHz. A good example of this effect is demonstrated in Figures 3.19 and 3.20 where there is a clear peak at around 1.7 kHz. This is due to the narrower constriction of the tongue right behind the stop constriction in anticipation of a high front vowel. This constriction due to the anticipation of the vowel can be modeled as a fairly short narrow tube which opens into a wide tube at the posterior portion of the vocal tract. The negative traveling wave due to the stop-consonant release will be reflected at the boundary between the small tube and the larger posterior tube. The reflected wave will be radiated through the lips and this will show up as an additional resonance in the transient spectrum.

The next observation is concerned with the frequency content for labial stop consonants. The source for all stop consonants is the same though these sources get filtered in some way by the cavity anterior to the source, except for labial stop consonants. The sound from labial stop consonants is radiated directly and therefore the theory predicts that the spectra should fall off as 1/f. As can be seen in the first two pages (Figures 3.19 and 3.20), the spectra fall off faster than the 1/f curve, some of them even approach the $1/f^2$ curve. This result shows that the theory made some approximations which are not completely true. The estimated spectrum displayed in Figure 2.3 comes close to predicting the transient bursts found in natural speech, though the match is not exact. The primary control of the theoretical spectra is the cross-sectional area of the constriction as a function of time. Since natural transients appear to fall off faster than $1/f$, the linear increase in the cross-sectional area as a function of time should be questioned.

The remaining spectra (Figures 3.21 through 3.26) generally agree with the theory which predicts that the spectra for non-labial stop consonants should be similar to the labial spectra with the addition of formants due to filtering by the cavity anterior to the constriction. The effect of front cavity filtering is best shown by the affricates (Figures 3.25 to 3.26). Here one can observe a consistent peak between 2.8 kHz and 3.2 kHz for males (corresponding to a length of 2.8 to 3.2 cm) and 3.2 kHz and 3.5 kHz for females (corresponding to a length between 2.5 and 2.8 cm). The velar stops (Figures 3.23 and 3.24) also show a distinct formant, though this formant varies in position according to the following vowel. The formant is between 1.5 kHz and 1.8 kHz for the stop followed by /ɑ/ (corresponding to an anterior cavity length between 4.9 and 5.9 cm). The formant center frequency is considerably higher for the stop consonants followed by /i/. Here the formant is between 2.6 kHz and 3.2 kHz (corresponding to a cavity length between 2.8 and 3.4 cm). The remaining class of stop consonants studied, alveolars (Figures 3.21 to 3.22), did not show a clear peak near the expected frequencies of 4.5 kHz for males and 6 kHz for females. Though we were limited to frequencies less than 5 kHz due to the slow sampling rate, a distinct formant is not noticed. One can observe only a very wide prominence which covers the 3 to 5 kHz frequency range and whose magnitudes are typically 10 to 20 dB larger than the corresponding magnitudes for equivalent frequencies of the labial spectra. Review of the individual spectra for each speaker showed that speakers had a wide variability in the place of articulation and that the averaging smoothed this out so much that a formant peak is barely noticeable. Also the speech was sampled at 10 kHz which means that the formant frequency for females could be beyond the Nyquist frequency of 5 kHz. In addition to the formants being at the predicted frequencies, the amplitude of the formants all show an increase of approximately 20 dB, just as was predicted in Figure 2.8. In general, the spectra measured from natural utterances agree with the theoretical predictions, though the variability in these acoustic data makes precise comparisons difficult.

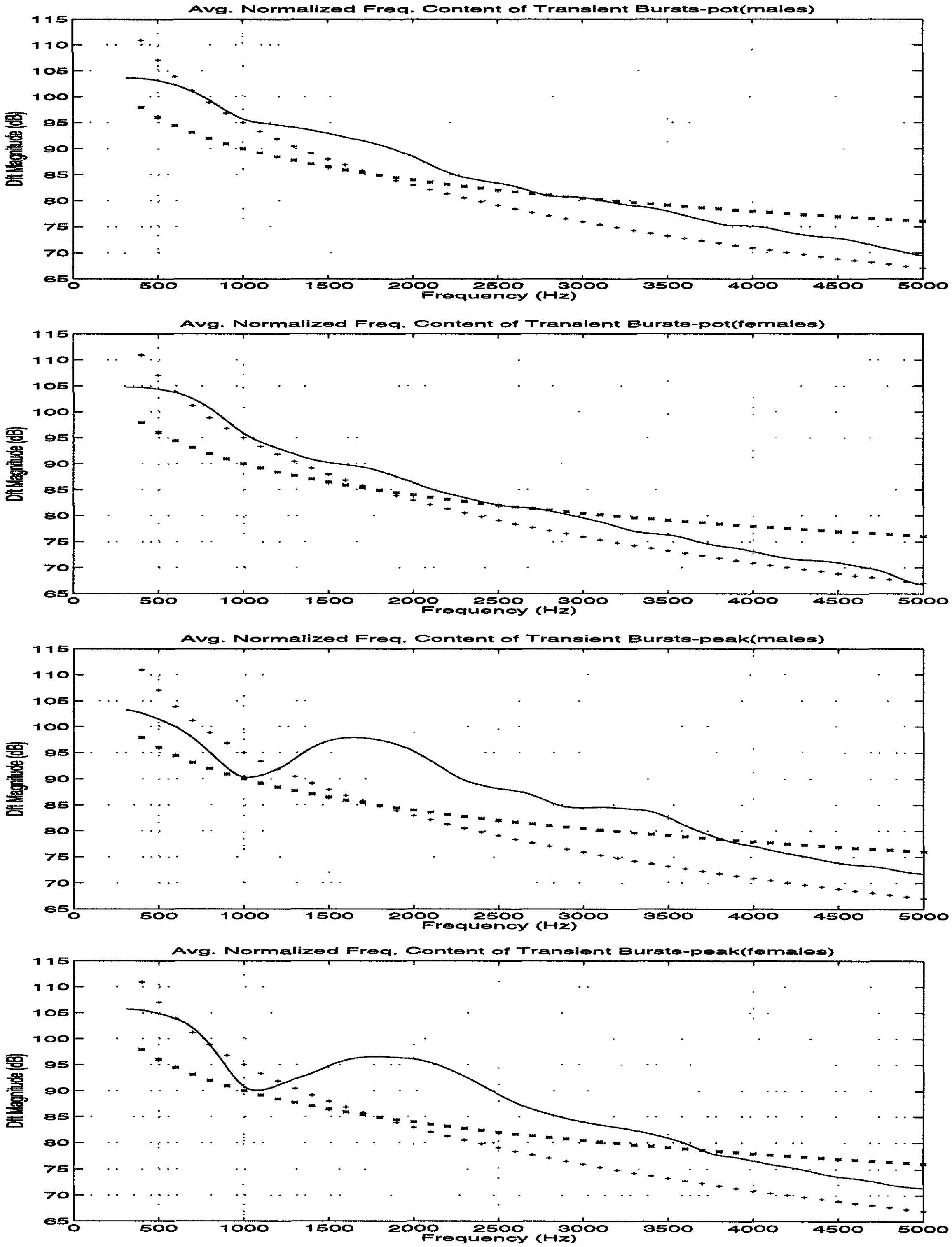

Figure 3-19: Spectral content of transient bursts in /p/ followed by /ɑ/ or /i/ spoken by males or females.

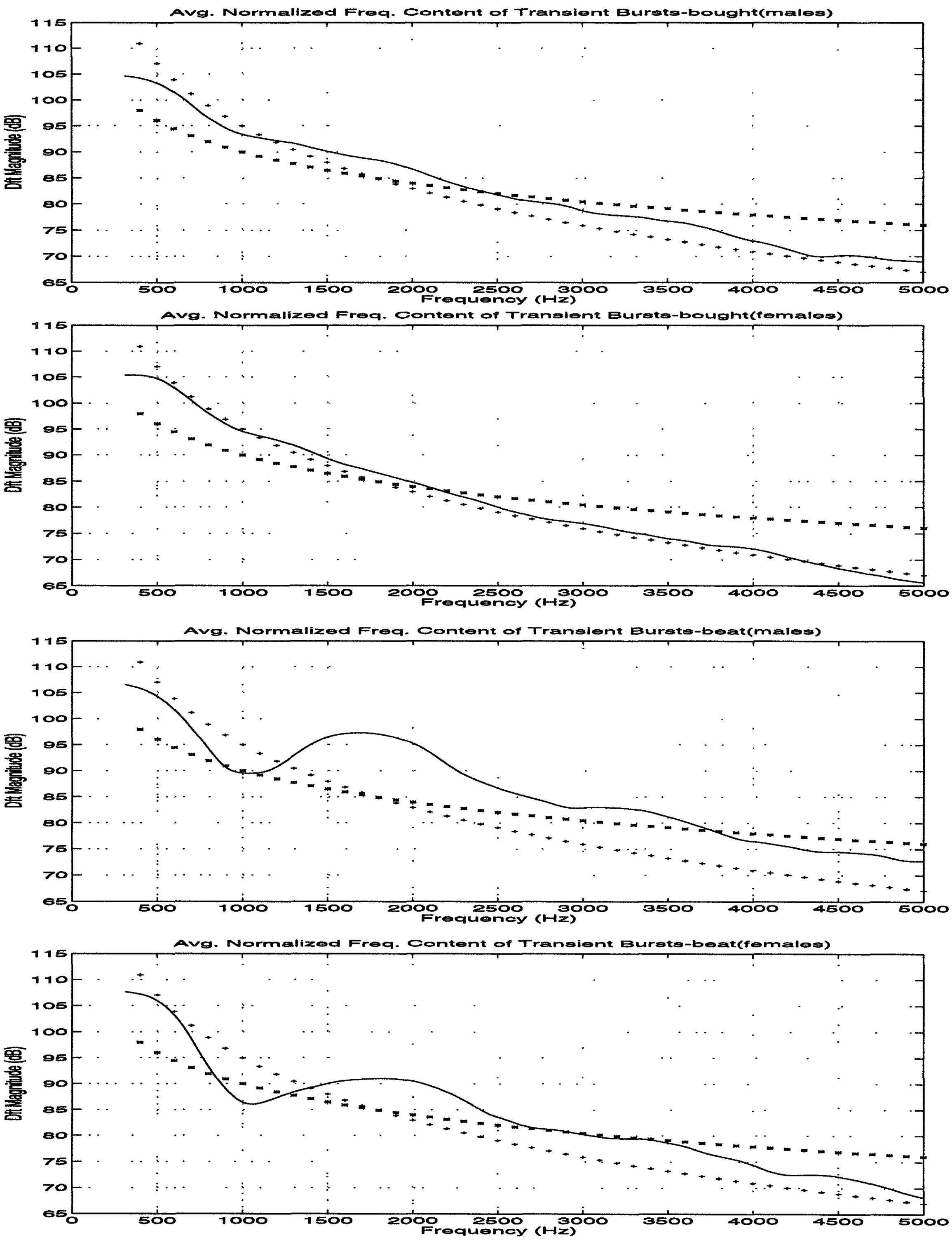

Figure 3-20: Spectral content of transient bursts in /b/ followed by /ɑ/ or /i/ spoken by males or females.

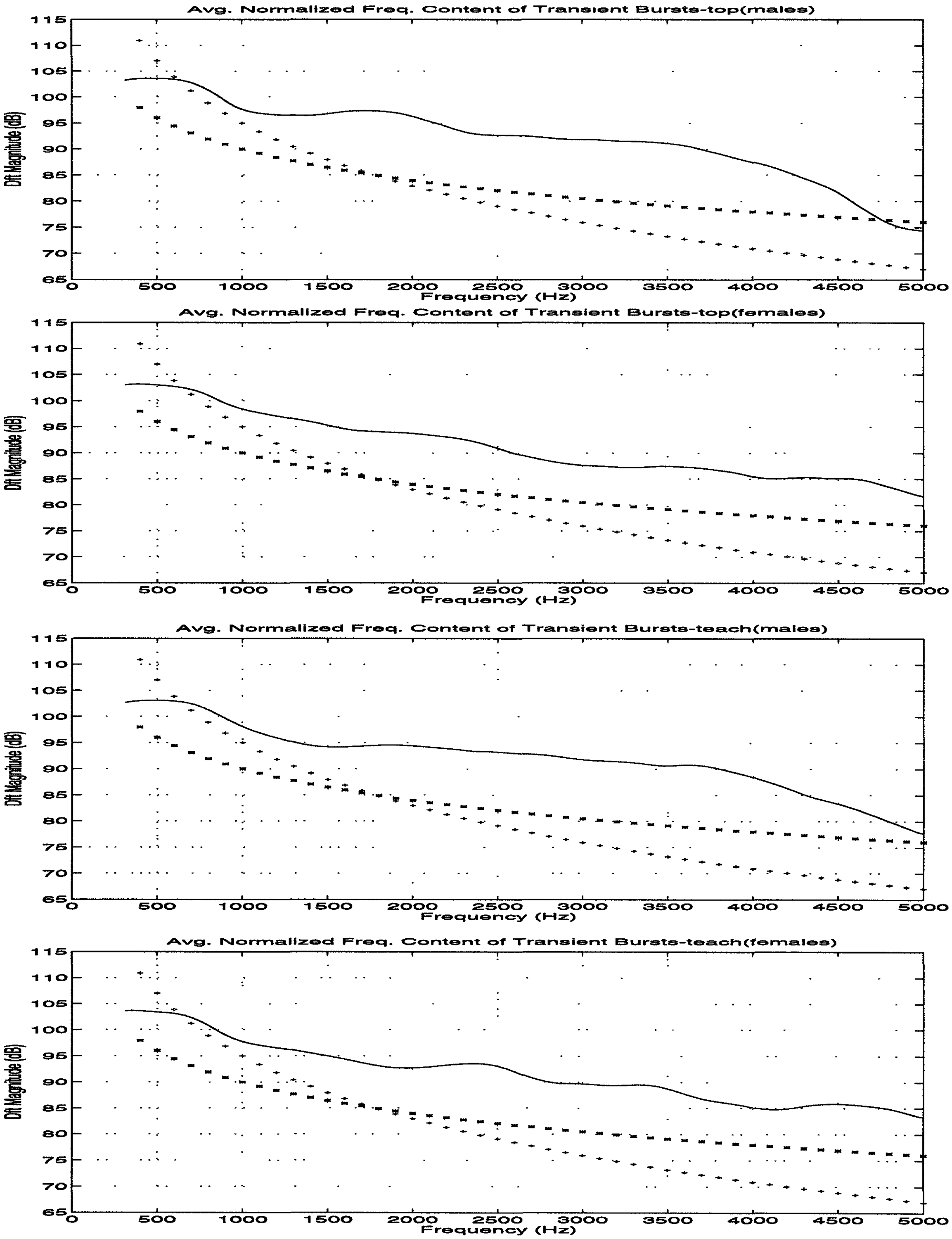

Figure 3-21: Spectral content of transient bursts in /t/ followed by /ɑ/ or /i/ spoken by males or females.

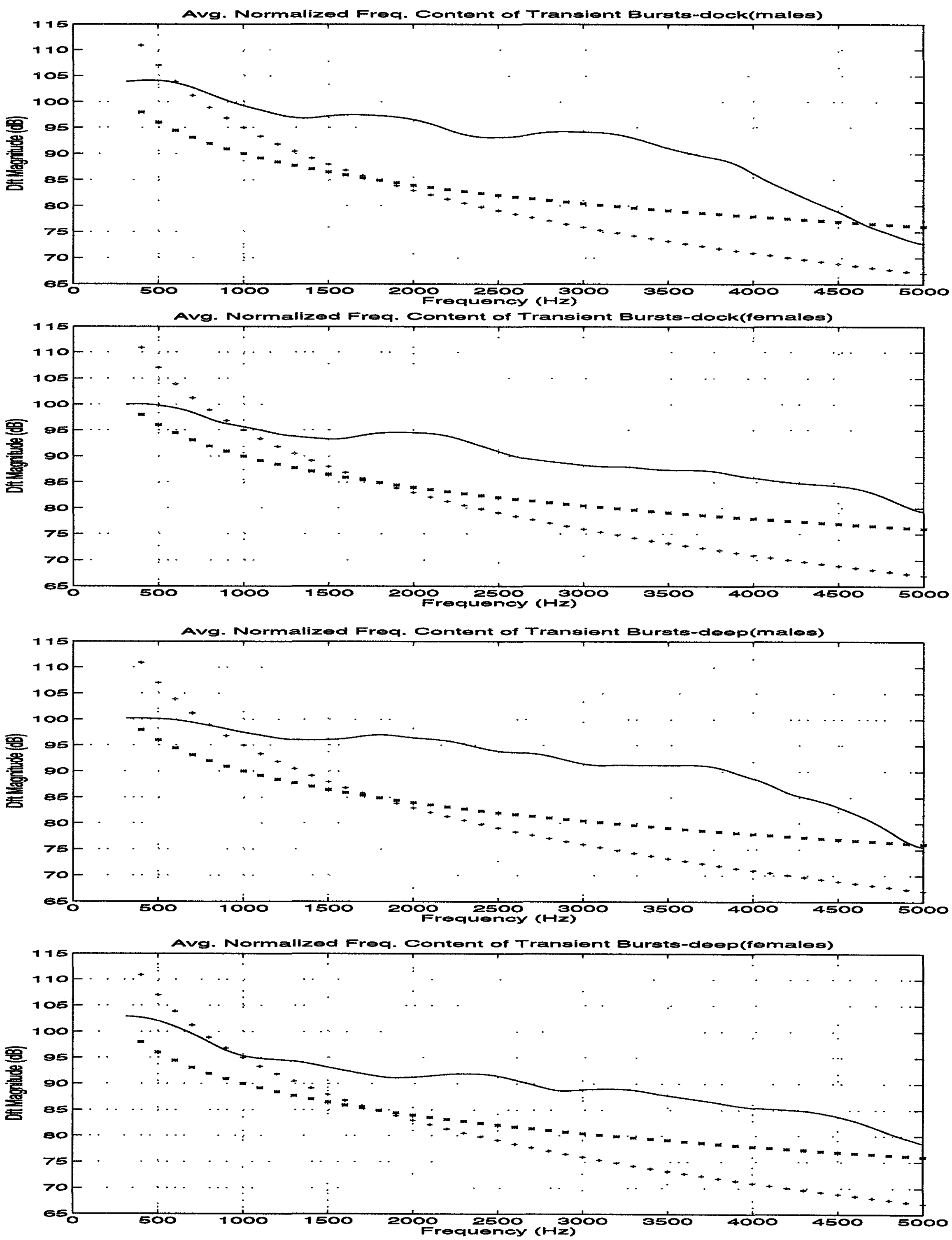

Figure 3-22: Spectral content of transient bursts in /d/ followed by /ɑ/ or /i/ spoken by males or females.

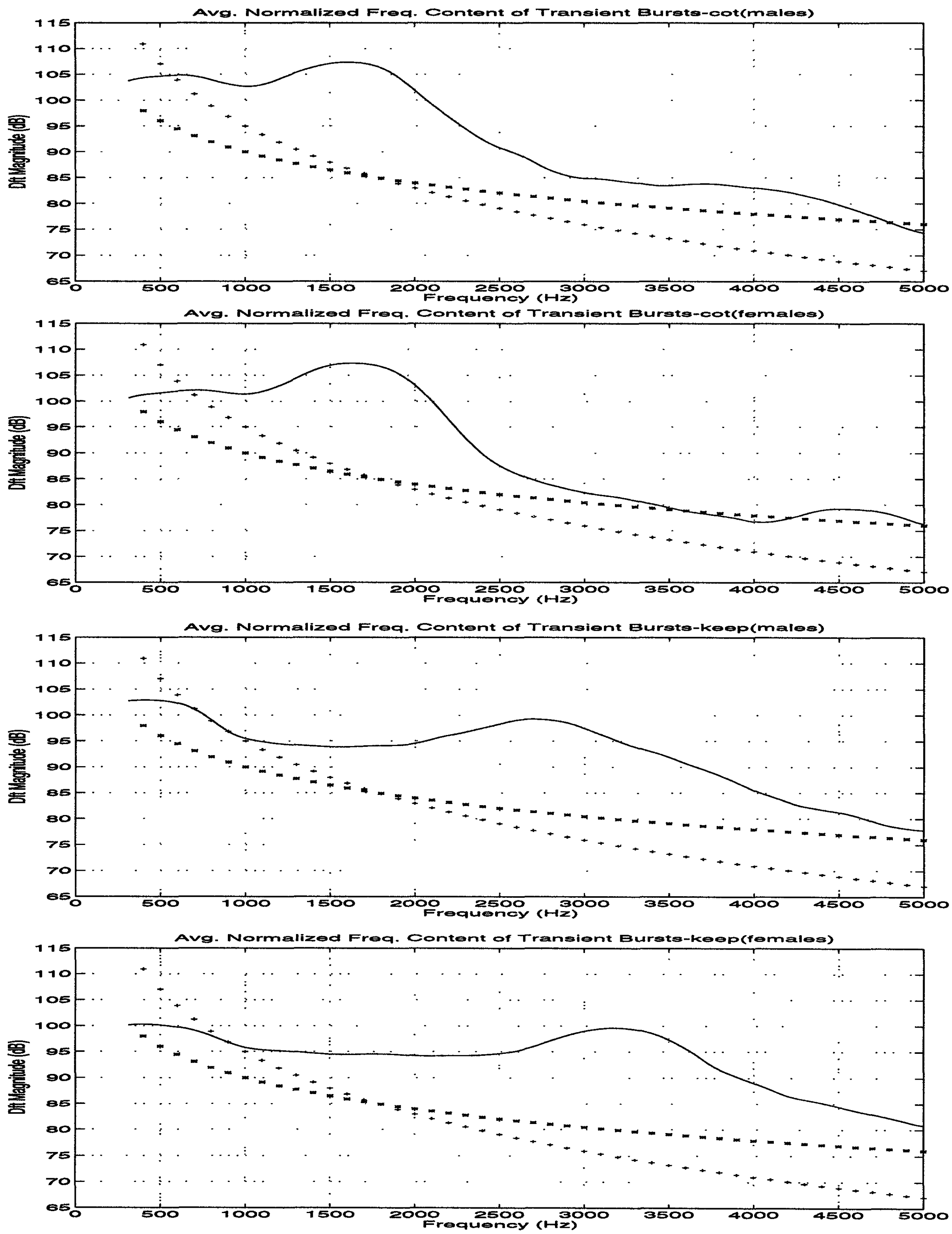

Figure 3-23: Spectral content of transient bursts in /k/ followed by /ɑ/ or /i/ spoken by males or females.

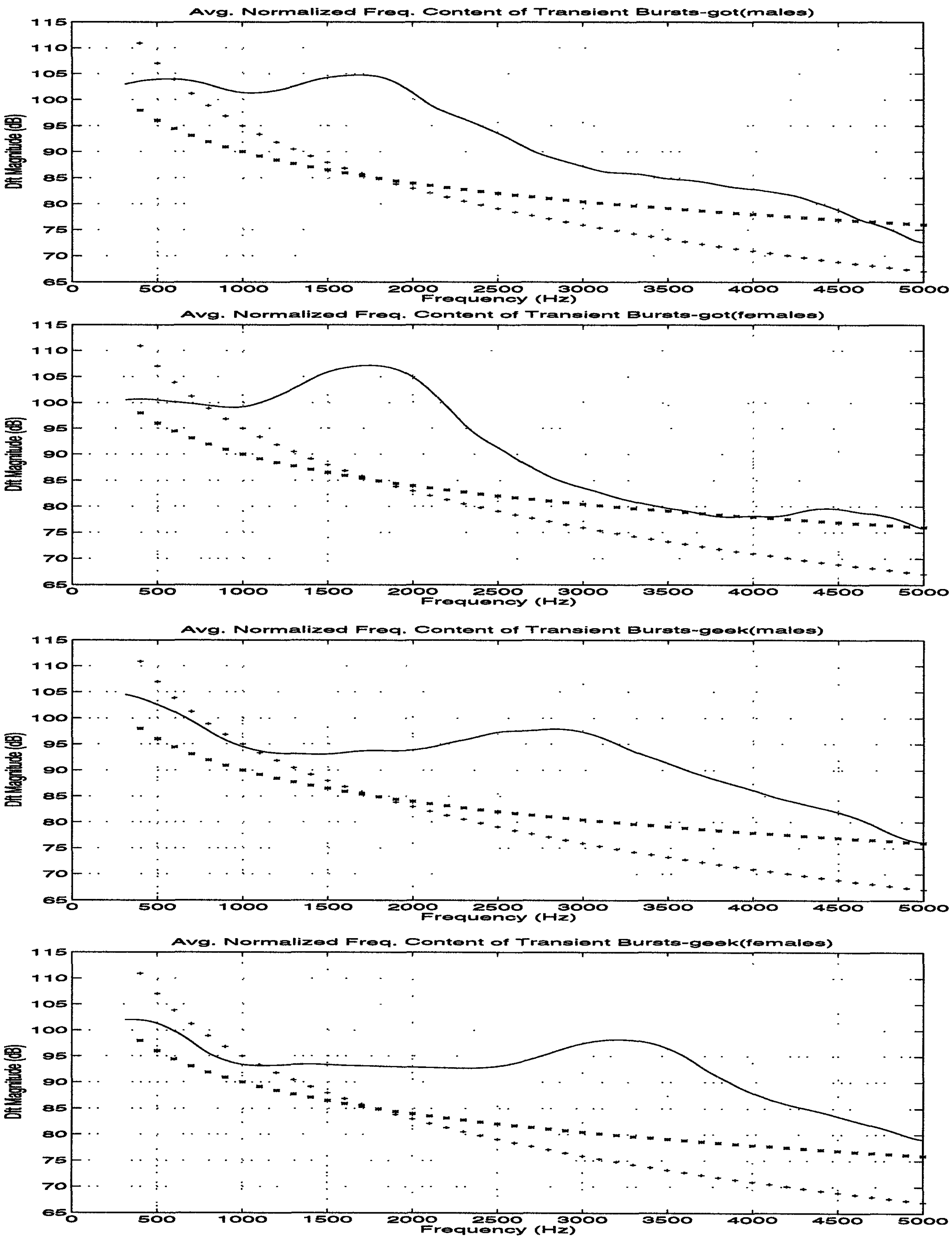

Figure 3-24: Spectral content of transient bursts in /g/ followed by /ɑ/ or /i/ spoken by males or females.

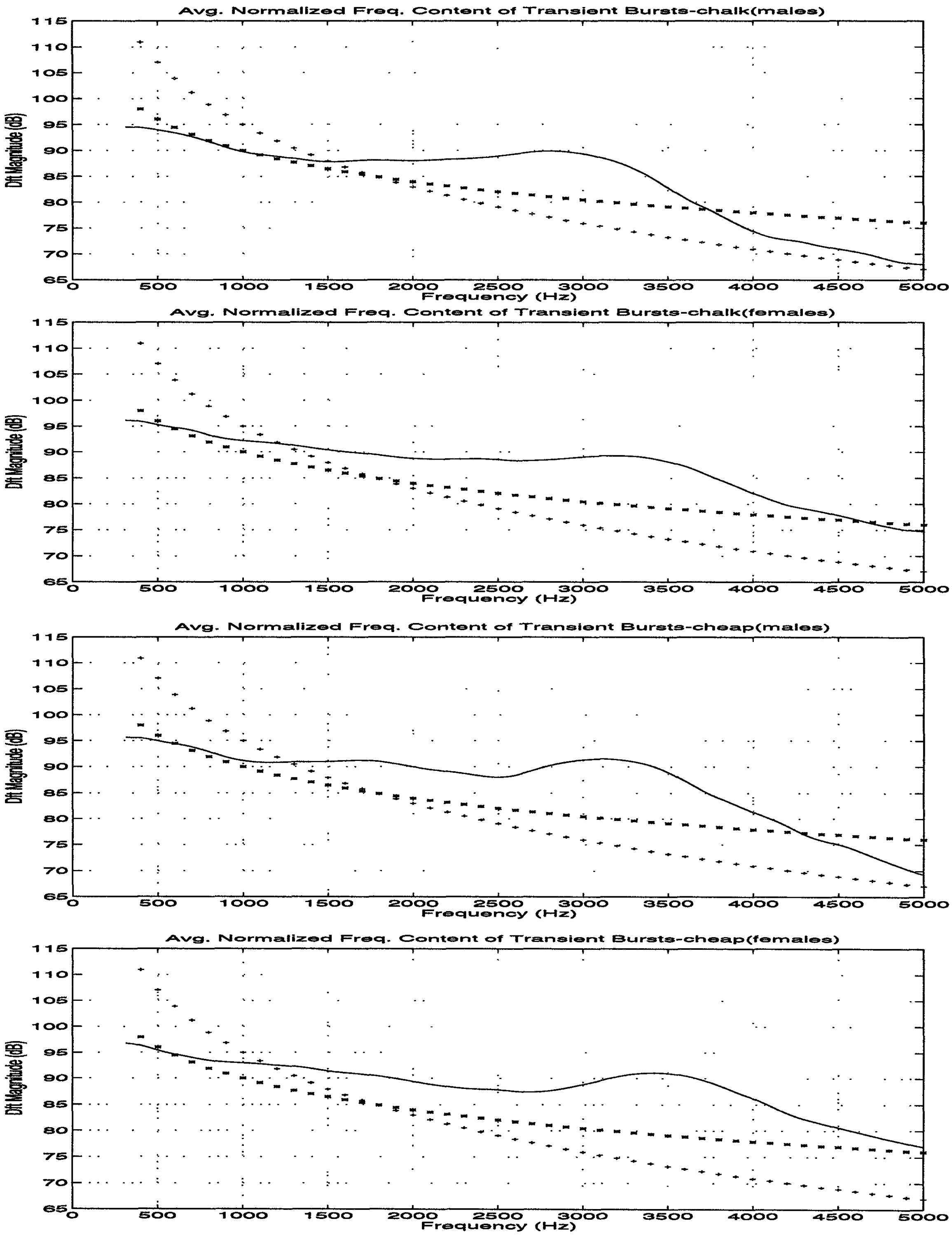

Figure 3-25: Spectral content of transient bursts in /ch/ followed by /ɑ/ or /i/ spoken by males or females.

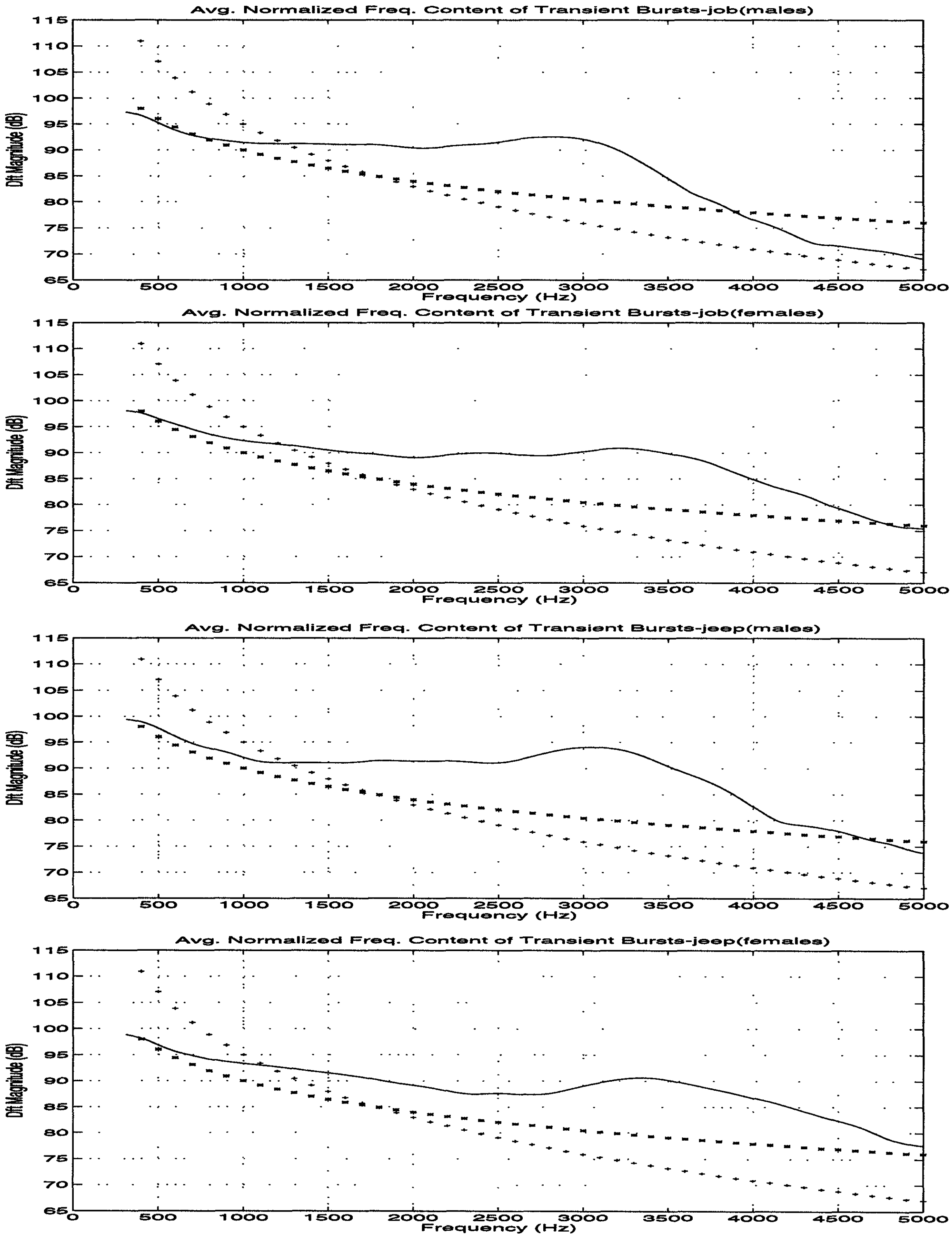

Figure 3-26: Spectral content of transient bursts in /j/ followed by /ɑ/ or /i/ spoken by males or females.

Chapter 4

Modifications of Synthesizer to Produce Transient Sources

4.1 Synthesizer Background

Speech synthesizers historically fall into two broad categories: articulatory synthesizers, and formant synthesizers. Articulatory synthesizers attempt to model the mechanical motions of the articulators, or at least the shapes of the airways. From these configurations the volume velocity and pressure in the lungs, larynx, vocal tract, and nasal tract can be calculated. The second category of synthesizers attempts to approximate the speech waveform directly by using a model formulated in the acoustic domain. These synthesizers are categorized as formant synthesizers.

The synthesizer used for this research is a modification of KLSYN88, a formant synthesizer originally created by Dennis Klatt. The major components of this synthesizer will be described but the reader should refer to the paper written by Dennis Klatt [17] for complete detailed description of the original synthesizer. A simplified block diagram is shown in Figure 4.1. One or more sources of sound energy is activated due to a pressure drop across some constriction in the airway and consequent flow of air. Each sound source excites the vocal tract, which acts as a resonating system, much like an organ pipe. The vocal tract shapes the sound source and the corresponding output, in the form of a volume velocity at the lips, is radiated to yield a sound pressure at a distance.

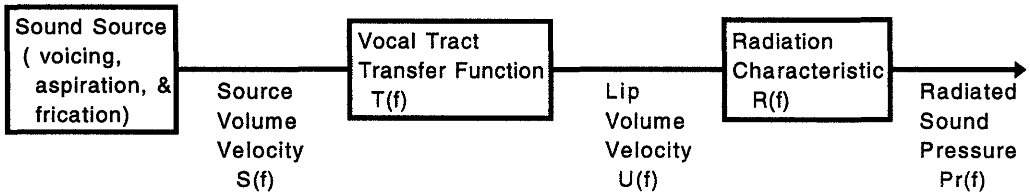

Figure 4-1: Simplified block diagram of formant synthesizer.(Klatt,1980)

KLSYN88 uses three separate sources to excite the vocal tract filter functions. A glottal sound source creates a waveform which imitates the sound caused by the airflow modulations resulting from opening and closing of the vocal folds. Aspiration noise can be added to this waveform to account for the turbulence of the air as it passes through the glottis. The third sound source can be generated by forming a constriction somewhere along the vocal tract. This type of noise is called frication and can be produced independently of the other noise sources.

The three types of sound sources do not excite the vocal tract in the same manner. First, the spectra of the sources are different. For example, frication does not have nearly the same amount of low frequency energy as voicing. Second, the sound sources can be generated at different locations along the vocal tract. Aspiration and glottal pulses are generated at the glottis so the filtering due to the vocal tract shape is about the same for both of these sources. The frication noise on the other hand does not always excite the same section of the vocal tract since this noise is generated by a constriction that can be made at varying places along the vocal tract. This means that the filtering of the frication noise is not just a function of the formant positions but also of the location of the constriction.

As a result of the varying location for the vocal tract constriction, the frication noise does not always affect the formant amplitudes in the same manner. The frication source is always filtered by the cavity anterior to the constriction, except for labials for which there is no anterior cavity. The total vocal tract resonances are the sum of resonances of the anterior cavity and the posterior cavity, noting that zeros in the transfer function can cancel or reduce the amplitudes of some resonances. In some vocal tract configurations the resonances of the anterior cavity may be responsible for only F4, while for other configurations the anterior cavity may be responsible for F3 and F5, for example. The formant filters for the frication source can be contrasted to the filters used to implement the transfer function from the glottis to the mouth. The relative strengths of the formant peaks with a glottal source can be calculated if the formant frequencies and bandwidths are known. As a result, the synthesizer does not need to provide users with the ability to control the individual formant amplitudes for the filtering of the glottal sources. On the other hand, the variability in the location of the constriction requires that the user have individual control of the formant amplitudes for frication filtering. As a result, a separate filter bank is used to filter the frication source, as shown in Figure 4.2. The user must set the amplitudes and bandwidths of each of the formants in accordance to the place of articulation.

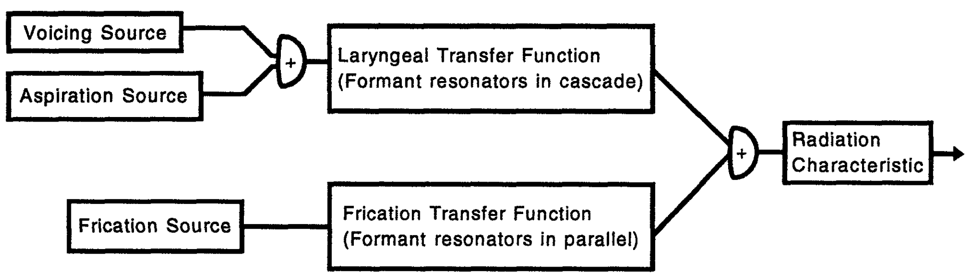

Figure 4-2: Block diagram of formant synthesizer.

4.2 Synthesizer Modifications

The transient burst in stop consonants is the sound caused by the release of the articulator forming the constriction and is therefore generated at various places along the vocal tract. Frication can also be generated in the vicinity of the constriction. Therefore, implementation of the transient burst could also make use of the previously mentioned filters used to shape the frication. The synthesizer has been modified by the addition of a new transient generator placed in parallel with the frication generator and the sum of these sounds is passed through the parallel vocal tract filter bank. The new synthesizer, with the ability to synthesize transients, is called KLSYN93.

The overall block diagram of KLSYN93, including the addition of the transient generator, is displayed in Figure 4.3. The cascade filter bank at the top of the figure is used to filter the glottal waveform and aspiration noise. The parallel filter bank in the lower portion of the figure is used to filter the frication and transient sources. In addition to these primary filter banks a secondary parallel filter bank (located in the center of the figure) exists which can be used to approximate laryngeal sounds. This secondary filter is only used for specialized synthesis applications and was not used in the research associated with this thesis.

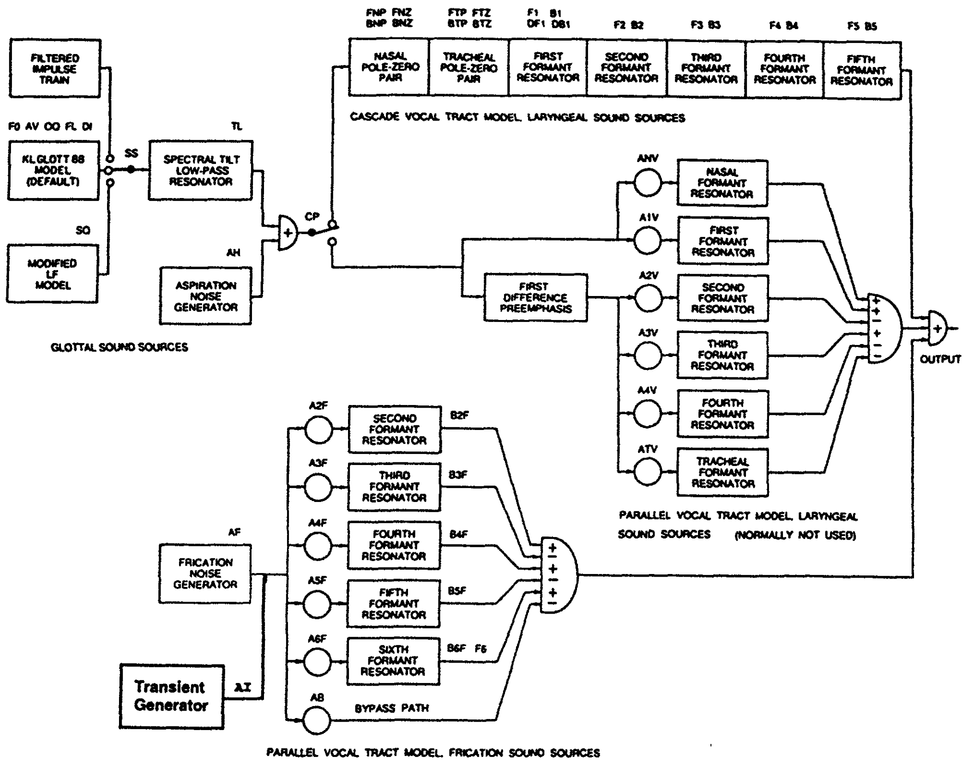

Figure 4-3: Block diagram of KLSYN93.(adapted from Klatt and Klatt, 1990)

The output of the synthesizer is altered by changing variables called control parameters. Each control parameter is given a two- or three-symbol name, a minimum value, a maximum value, a default value which is used if no changes are made, and an English description of the parameter's effect on the synthesis. A list of the control parameters is given in Figure 4.4.

Two new variables were added to the synthesizer parameter list in order to control the new transient generator. Parameter AI, standing for amplitude of the impulsive source, is a time varying parameter. Like the amplitude of frication, AF, this new parameter can be changed every update interval (UI) by using the synthesizer's draw

Synthesis specification for file: 'temp.wav' Mon Mar 28 16:59:53 1994

KLSYN93 Version 2.0 April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is -8.6 dB
Total number of waveform samples = 5000

CURRENT CONFIGURATION:
63 parameters

SYM	V/C	MIN	VAL	MAX	DESCRIPTION
DU	C	30	500	5000	Duration of the utterance, in msec
UI	C	1	5	20	Update interval for parameter reset, in msec
SR	C	5000	10000	20000	Output sampling rate, in samples/sec
NF	C	1	5	6	Number of formants in cascade branch
SS	C	1	2	3	Source switch (1=impulse, 2=natural, 3=LF model)
RS	C	1	8	8191	Random seed (initial value of random # generator)
SB	C	0	1	1	Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
CP	C	0	0	1	0=Cascade, 1=Parallel tract excitation by AV
OS	C	0	0	20	Output selector (0=normal,1=voicing source,...)
GV	C	0	60	80	Overall gain scale factor for AV, in dB
GH	C	0	60	80	Overall gain scale factor for AH, in dB
GF	C	0	60	80	Overall gain scale factor for AF, in dB
GI	C	0	60	80	Overall gain scale factor for AI, in dB
F0	v	0	1000	5000	Fundamental frequency, in tenths of a Hz
AV	v	0	60	80	Amplitude of voicing, in dB
OQ	v	10	50	99	Open quotient (voicing open-time/period), in %
SQ	v	100	200	500	Speed quotient (rise/fall time, LF model), in %
TL	v	0	0	41	Extra tilt of voicing spectrum, dB down @ 3 kHz
FL	v	0	0	100	Flutter (random fluct in f0), in % of maximum
DI	v	0	0	100	Diplophonia (alt periods closer), in % of max
AH	v	0	0	80	Amplitude of aspiration, in dB
AF	v	0	0	80	Amplitude of frication, in dB
F1	v	180	500	1300	Frequency of 1st formant, in Hz
B1	v	30	60	1000	Bandwidth of 1st formant, in Hz
DF1	v	0	0	100	Change in F1 during open portion of period, in Hz
DB1	v	0	0	400	Change in B1 during open portion of period, in Hz
F2	v	550	1500	3000	Frequency of 2nd formant, in Hz
B2	v	40	90	1000	Bandwidth of 2nd formant, in Hz
F3	v	1200	2500	4800	Frequency of 3rd formant, in Hz
B3	v	60	150	1000	Bandwidth of 3rd formant, in Hz
F4	v	2400	3250	4990	Frequency of 4th formant, in Hz
B4	v	100	200	1000	Bandwidth of 4th formant, in Hz
F5	v	3000	3700	4990	Frequency of 5th formant, in Hz
B5	v	100	200	1500	Bandwidth of 5th formant, in Hz
F6	v	3000	4990	4990	Frequency of 6th formant, in Hz (applies if NF=6)
B6	v	100	500	4000	Bandwidth of 6th formant, in Hz (applies if NF=6)
FNP	v	180	280	500	Frequency of nasal pole, in Hz
BNP	v	40	90	1000	Bandwidth of nasal pole, in Hz
FNZ	v	180	280	800	Frequency of nasal zero, in Hz
BNZ	v	40	90	1000	Bandwidth of nasal zero, in Hz
FTP	v	300	2150	3000	Frequency of tracheal pole, in Hz
BTP	v	40	180	1000	Bandwidth of tracheal pole, in Hz
FTZ	v	300	2150	3000	Frequency of tracheal zero, in Hz
BTZ	v	40	180	2000	Bandwidth of tracheal zero, in Hz
A2F	v	0	0	80	Amp of fric-excited parallel 2nd formant, in dB
A3F	v	0	0	80	Amp of fric-excited parallel 3rd formant, in dB
A4F	v	0	0	80	Amp of fric-excited parallel 4th formant, in dB
A5F	v	0	0	80	Amp of fric-excited parallel 5th formant, in dB
A6F	v	0	0	80	Amp of fric-excited parallel 6th formant, in dB
AB	v	0	0	80	Amp of fric-excited parallel bypass path, in dB
B2F	v	40	250	1000	Bw of fric-excited parallel 2nd formant, in Hz
B3F	v	60	300	1000	Bw of fric-excited parallel 3rd formant, in Hz
B4F	v	100	320	1000	Bw of fric-excited parallel 4th formant, in Hz
B5F	v	100	360	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	v	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

No parameters are varied

Figure 4-4: KLSYN93 control parameter list.

command. However, unlike AF, the parameter AI does not remain at the same amplitude for the entire update interval. The parameter AI generates a transient pulse at the beginning of the update interval, and the transient pulse lasts four samples regardless of the update interval. The second new parameter is named GI, standing for the overall gain of the impulsive source. This parameter is a constant for the entire duration of the synthesis. Its value would be changed if the relative amplitudes of the transient pulses, set by using the parameter AI, were satisfactory, but each transient seemed a little too loud or a little too soft. In this case it is much easier to change the parameter GI than to change the individual levels of each of the transients.

The transient burst precedes any frication noise in time. According to theory, the filtering of these two different noise sources should be identical. It is, however, conceivable that this transient burst could be used to produce different sounds such as clicks, as mentioned in section 1.1. In these cases it is possible that the transient noise burst should be filtered differently than the frication noise. For the research with stops and affricates, the transient burst and frication noise is turned on during the same update interval, and it is therefore required that they both be filtered identically. It is possible to change the update interval to 1 ms and then turn on the frication noise 1 ms later than the transient burst. In this case the transient burst and frication could be filtered differently by changing the filter parameters when the frication is turned on. Therefore it should not be a problem that the transient burst and frication are sent through the same parallel filter bank.

There are potential problems that could arise due to the synthesizer source approximations. It could be possible that the frication source does not actually generate noise which matches the frication generated in natural speech. Since Klatt did not intend for any other source to be filtered by the parallel filter bank, he may have built a source whose imperfections were corrected by changing some of the filter parameters. If this is the case then the filtering of the transient source and the frication source will not be the same. An attempt to compensate for the inadequacies of one source will cause improper filtering of the other source. Also, since the transient source is a volume velocity source, the actual source of noise will be generated in the vocal tract slightly downstream from the constriction [2]. The transient source will couple differently to the vocal tract than the frication source, causing different relative amplitudes of the formants. Therefore the formant amplitude parameters may have to be changed when the transient source is turned on. Experience may show that it is too difficult to determine appropriate synthesizer parameters to filter these two sources with the same filter bank, so another parallel bank of filters may be needed to separately filter the transient source.

4.3 Synthetic Transient Waveform Shape

The shape of the transient burst was chosen according to the theory presented in Chapter 2. This theory predicts that the transient pulse measured at the lips should have a spectrum which falls off as $1/f^2$ (12 dB per octave). The synthesizer generates sound pressure that would be present at a listener's ear, which is approximately proportional to the derivative of volume velocity at the lips. In the frequency domain this translates to an addition of 6dB/oct in the spectrum. The overall result is a waveform whose spectrum falls off at 6 dB/oct (proportional to $1/f$).

The Fourier transform of a step function gives this desired spectrum; however, it is not reasonable to generate a step function. This function lasts forever and it is not necessary to represent the step function for very low or very high frequencies. Speech does not contain these frequencies, so to generate a waveform which contained these frequencies would produce an unnatural sound. As a result of these observations, an approximation to the spectrum of a step function was used.

MATLAB was used to determine the transient pulse which best fit the theoretical data. The first question that needs to be asked is what is the best frequency response of the transient. Since we are approximating the frequency response there will obviously be some deviations from the theoretical transient. The MATLAB command 'invfreqs' allows the user to specify a weighting function which would force the approximation to be very close in some frequency regions while allowing it to deviate from the ideal in other regions. The important regions seem to be above about 1 kHz, and so the weighting function was chosen to restrict the error over this region and allow the function to do whatever it wants to for low frequencies (though the general shape is constrained by the order of the filter). The best-fit approximation using FIR filters of length 2,3,4, and 5 samples are displayed in Figure 4.5.

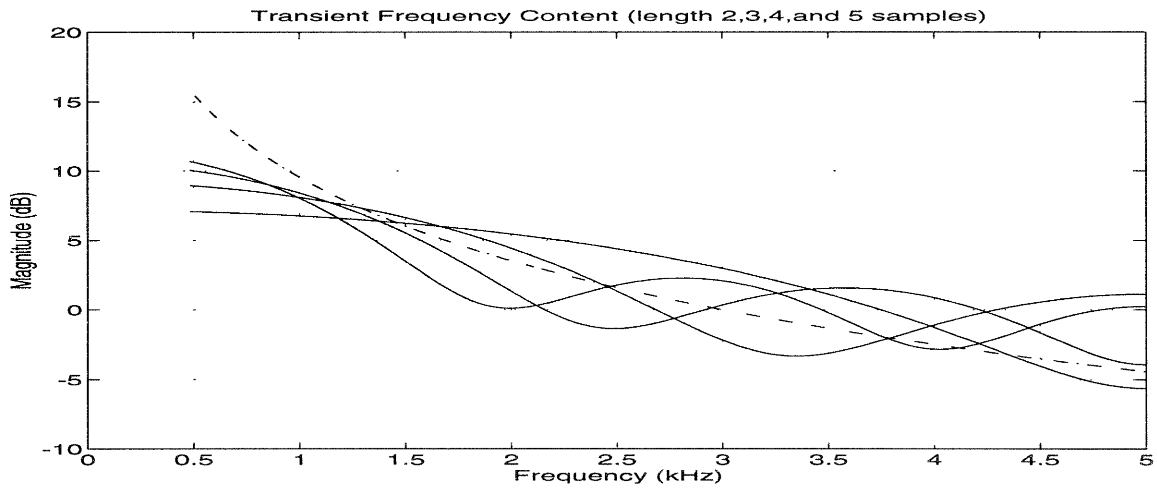

Figure 4-5: Frequency content of transient approximations (length = 2,3,4,and 5 samples).

4.4 Analysis of Synthetic Transient Burst

It seems reasonable for the approximation to change very slowly with frequency, yet the more 'bumps' that are allowed in the spectrum the closer is the approximation to the desired spectrum. A transient length of 4 samples was chosen as a good compromise between these two constraints. The approximation and 1/f curves are displayed in Figure 4.6 and the error in this approximation, or the difference between the approximation and a 1/f curve is displayed in Figure 4.7. Notice that the error is less than 3 dB for frequencies greater than 1 kHz. These frequency requirements produce a waveform displayed in Figure 4.8. This waveform is generated each time the synthesizer parameter AI is nonzero.

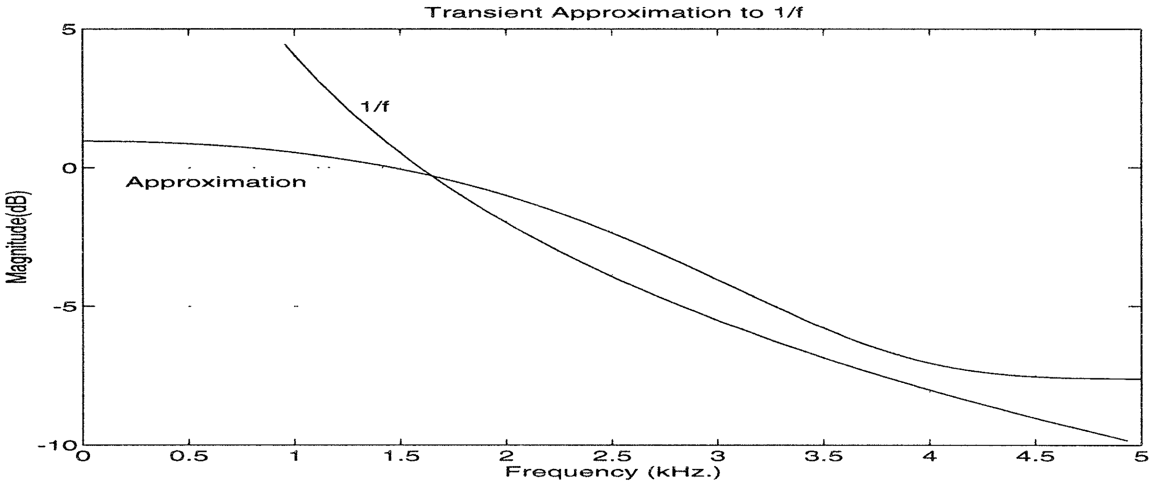

Figure 4-6: Frequency content of transient approximation (length=4 samples).

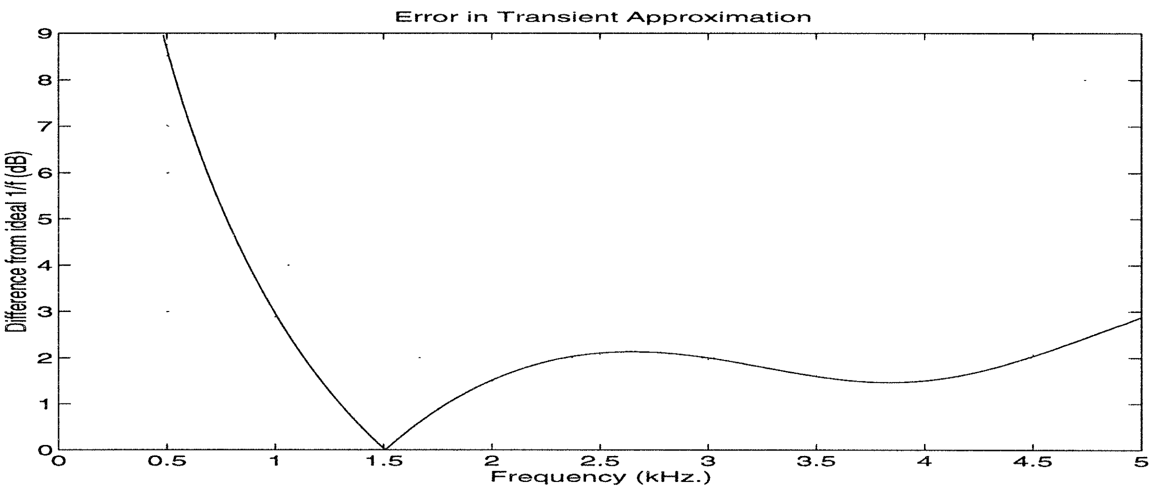

Figure 4-7: Transient approximation error (approximation - 1/f).

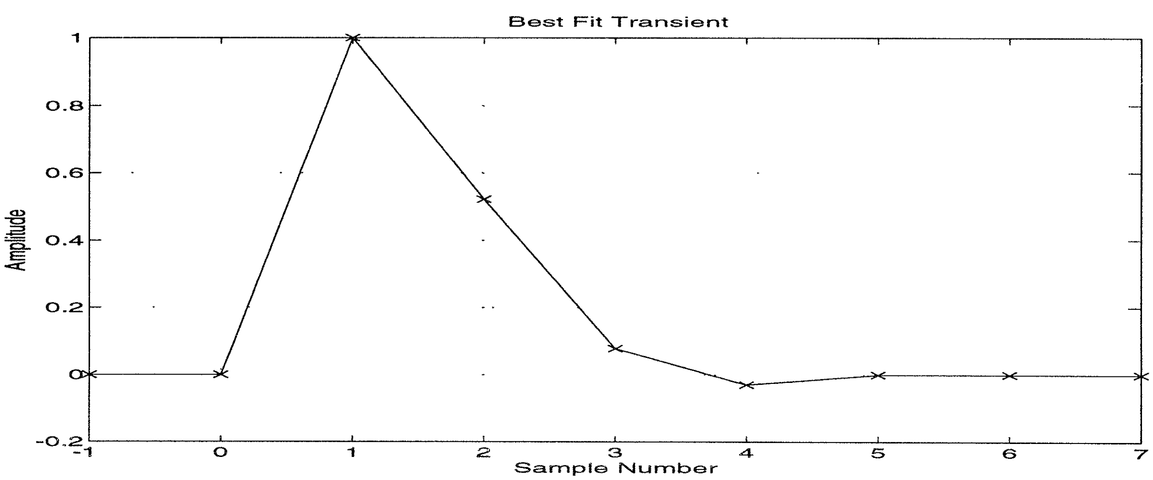

Figure 4-8: Best-fit transient waveform (length = 4 samples).

Chapter 5

Perception Tests

The development of theory and the analysis of transient bursts is futile if these transients are not perceived. The purpose of studying speech is to determine how words get from one person's brain to another person's brain. Discovering and quantifying the type of toothpaste which is most likely to have been used prior to a conversation would probably not enhance our understanding of speech communication. Likewise, the description of transient bursts is valuable only if the listeners use transients to determine what was spoken. In order to ascertain the perception of transient bursts, a series of listening tests was designed.

5.1 Pilot Test

A pilot perception test was performed using the utterance /pɑ/. Eight different transient levels were synthesized. The lowest amplitude transient had a magnitude of 0 (meaning no transient was present), the quietest transient had a magnitude of 30 dB (in standard synthesizer amplitudes, with the following vowel at AV=60 dB) and the remaining six were simply 5 dB additions to the previous magnitude. Each stimulus was paired with each of the remaining stimuli resulting in 28 pairs. The order of each pair was reversed, resulting in 28 more pairs. A total of 56 pairs were repeated twice to the listeners. Five listeners were asked to judge which element of the pair sounded more natural. Three of the listeners worked in the Speech Communication

Group at MIT, and the remaining two were college students not involved in any type of speech research. The results of the perception tests are shown in Figure 5.1. The important information that this test showed was that listeners definitely disliked of large amplitude transients. Also the histogram does not change rapidly between two consecutive bars; the same general trends could be shown by looking only at every other amplitude of the stimuli (in other words, by looking at transients spaced every 10 dB instead of every 5 dB). This means that listeners are not extremely sensitive to small changes in the magnitude of the transient over the area of interest. Therefore, subsequent tests were conducted using only five different transient magnitudes. This allows the tests to be shorter thereby allowing the same information to be obtained in substantially less time.

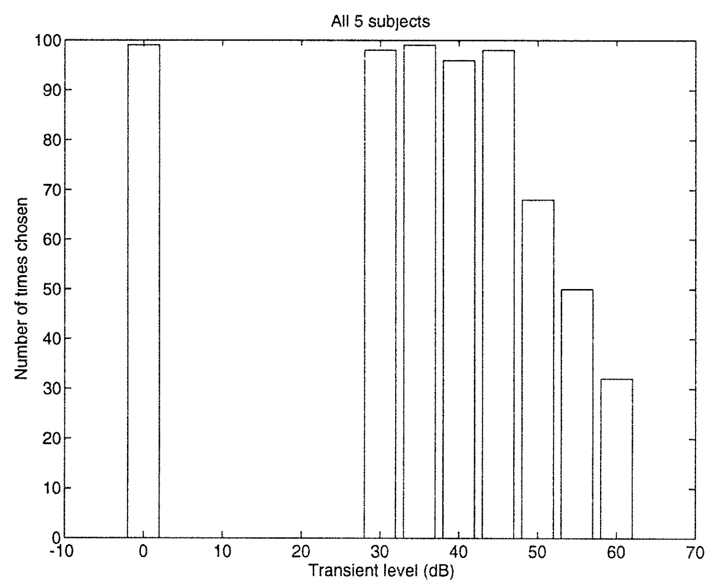

Figure 5-1: Results of pilot test for /pɑ/.

5.2 Description of Stimuli, Subjects, and Experiment

Each of the stop consonants followed by the vowel /ɑ/ was synthesized with varying transient amplitudes. Five different transient levels were chosen for each stop consonant, ranging from no transient present to a level which was chosen by the researcher to sound much louder than those present in natural speech. Each of the five synthesized utterances was paired with the other four utterances resulting in ten pairs. The order of the elements in each of the pairs was reversed, resulting in another ten pairs. These were combined with the previous ten pairs and each pair of this set of 20 pairs was presented 4 times in a random order to the listeners. The final test therefore consisted of 80 stimulus pairs.

Seven listeners were selected for the perception tests. The pilot perception test showed that subjects who were not experienced in perception tests did not score significantly different from those who are involved in speech research. As a result, it seemed fair to take advantage of the convenience of choosing subjects from the Speech Communication Group at MIT. Six of the seven subjects were native English speakers. The only non-native English speaker has been living in the United States and speaking English for 22 years.

The listeners were presented with a test involving pairs of transient bursts of various amplitudes. Each of the seven listeners was presented with an 80-item test for each of the eight stop consonants. The synthesized utterances were recorded onto a cassette tape and played for the listeners using a Nakamichi LX-5 tape deck and Sennheiser HD430 headphones. The tests were administered in a soundproof room to reduce distracting noises. The listeners were asked to choose which element of the pair sounded most natural. The responses were recorded on paper and later entered into a computer for analysis.

5.3 Perception Test Results

One method of analyzing the data consists of simply keeping track of how many times a particular stimuli was chosen. The perception tests consisted five stimuli being presented 32 times, resulting in 160 stimuli (80 pairs). Since every stimulus is presented 32 times to each listener, if the seven listeners all agree that a particular stimuli is the most natural sounding then the total number of times this element would be chosen is 7*32=224 times. Therefore a useful statistic is the total number of times each element was chosen by the listeners. This record keeping was performed by a MATLAB program named ltr2hist.m (Appendix B). This program reads the response sheets and scores them according to the answer key generated by the MAKETAPE command. The scores for all of the subjects were combined to generate histograms, which are shown in Figure 5-2.

For each of the stop consonants one should notice that the most preferred transient level is non-zero. Further analysis of the perception test results should be performed on these most preferred data to determine their significance. It is quite clear that listeners do not think that an extremely loud transient sounds natural. Therefore a more detailed statistical analysis of these data does not need to be performed.

An additional statistical analysis of the most preferred stimulus was performed. The results in Table 5.1 summarize what happened every time the most preferred stimulus, as determined by the largest bar in Figure 5-2, was presented as an element in the pair. The number in each column is the percent of times that the stimulus indicated at the top of the column was chosen when the stimulus was presented with the other element of the pair being the most preferred transient (marked in the table with a $\oplus$). For example, the number in the last column of the /p/ row means that the /pɑ/ stimulus with a 60 dB transient was preferred 8.9% of the time when listeners were presented with the 60 dB stimulus as one element of the pair and the most preferred stimulus (40 dB for /pɑ/) as the other element of the pair. Each percentage is the result of eight tests presented to each of the seven listeners. The number in parenthesis in column 1 of Table 5.1 is the confidence level of the most preferred transient level when it was compared to the stimuli with no transient present (0 dB). This number is the chance that the most preferred transient was actually a result of random guessing, not the fact that the listeners preferred the presence of the transient. These results show that listeners definitely prefer the presence of a transient bursts in affricates. The conclusion is probably the same for stops, though the data are not conclusive that this is the case.

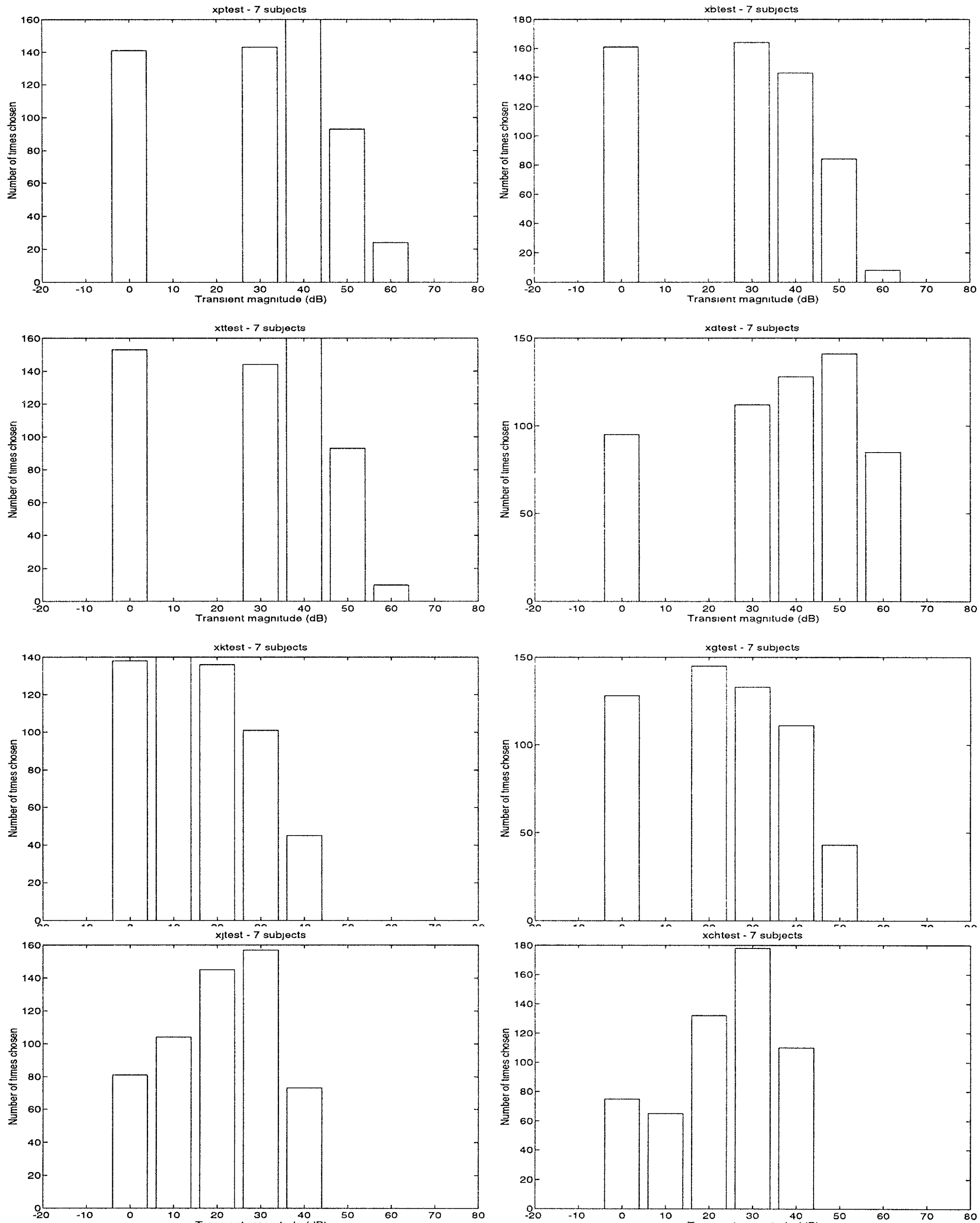

Figure 5-2: Perception test results.

phoneme	0 dB (conf)	10 dB	20 dB	30 dB	40 dB	50 dB	60 dB
p	44.6 (.25)			42.9	$\oplus$	17.9	8.9
t	50.0 (.50)			35.7	$\oplus$	23.2	5.4
k	39.3 (.07)	$\oplus$	55.4	35.7	19.6		
b	57.1 (.83)			$\oplus$	32.1	14.3	3.6
d	37.5 (.04)			35.7	44.6	$\oplus$	30.4
g	42.9 (.17)		$\oplus$	41.1	41.1	16.1	
ǰ	23.2 (<.001)	35.7	37.5	$\oplus$	23.2		
č	10.7(<.001)	10.7	30.4	$\oplus$	30.4		

Table 5.1: Perception test results. Each number is the percent of times the stimuli was chosen when it was paired with the most preferred transient (marked by $\oplus$).

5.4 Comparison of Synthetic and Natural Transient Magnitudes

A final analysis of the perception test data involves the comparison of the most preferred transient level to the transient levels found in natural speech. The same analysis that was performed on natural speech was also performed on the synthesized speech that was judged by listeners to be the most preferred. Table 5.2 shows the vowel, transient, and relative transient magnitudes found in the synthesized speech. The average relative transient magnitude measured in natural speech for each of the stop consonants followed by the vowel /ɑ/ was extracted from Figures 3.4 to 3.18 and is displayed in the last column. A comparison of the last two columns shows that the most preferred synthetic transient was within 5 dB of the observed natural transients for 4 of the 8 cases, and within 9 dB for 7 of the 8 cases.

5.5 Problems with Synthesized Utterances

A couple of comments should be made about the synthesized utterances (see Appendix A for the complete .DOC files for these utterances). First, the synthesizer is designed so that the frication and aspiration are turned on abruptly. This is not what happens in natural speech. The synthesizer generates a pseudo-random noise whose average value is determined by the appropriate synthesizer parameter, AF for example. The noise is nevertheless random, and the first few samples of this noise could be very large. The control parameter RS allows the pseudo-random number generator to start with a different value. This parameter was not varied in the synthesis of the utterances used for the perception tests, but this parameter could change the perceived abruptness of the noise. Due to the abrupt onset of noise, an argument can be made that the old synthesizer already had a type of noise source which generated a transient-like onset. This would mean that the preferred magnitude of the transient generated using the new parameter AI would be less than the magnitude predicted by theory or observed in natural speech.

Second, there is a lack of consistency in the synthesis quality of the different stop consonants. The utterances were synthesized according to the researcher's taste; the difference between /pɑ/ and /tɑ/ was not just a change in formant positions. Other parameters were changed to make the utterance sound more natural. Therefore the frication and aspiration noise for some of the stop consonants was larger than for others. In the past, the amplitude of the frication noise source was increased during the first update interval to imitate the generation of a transient. This shows that the higher amplitude of the frication noise could add to the perception of a transient burst and thereby cause the preferred magnitude of the transient generated using the new parameter AI to be less than the magnitude predicted by theory or observed in natural speech. Therefore, the variation in frication and aspiration noise across synthesized stop consonants could also have changed the perceived magnitude of the transient burst even though the parameter AI was the same.

phoneme	Synthesized				Natural
	avg vowel	avg_vowel peak	avg trans	avg relative	avg relative
p	52 dB	61 dB	26 dB	-26 dB	-22 dB
t	52 dB	61 dB	24 dB	-28 dB	-20 dB
k	52 dB	61 dB	7 dB	-44 dB	-17 dB
b	52 dB	61 dB	19 dB	-32 dB	-23 dB
d	52 dB	61 dB	31 dB	-21 dB	-21 dB
g	52 dB	61 dB	27 dB	-24 dB	-17 dB
ǰ	51 dB	60 dB	30 dB	-21 dB	-24 dB
č	51 dB	60 dB	30 dB	-21 dB	-26 dB

Table 5.2: Comparison of transient magnitudes from most preferred perception test results and observed natural speech.

Chapter 6

Conclusions

The theory developed in this research predicts the properties of transient bursts observed in natural speech. The intraoral pressure before the stop-consonant release and cross-sectional area of the opening following the release is not known. This information is necessary to predict the amplitude of the of the transient burst, thereby limiting the accuracy of the results predicted by this theory. A modification to the Klatt-based formant synthesizer now enables researchers to generate synthetic transients with properties similar to natural transients. Though the results are not conclusive, listeners generally prefer the presence of transient bursts in synthesized speech, especially for synthetic affricates.

6.1 Further Work

More studies similar to the perception tests described above need to be conducted. There are not enough data yet to make any solid conclusions as to the transient magnitude that is most preferred by listeners. It is also not known to what degree listeners can distinguish variations in the amplitude of transient bursts in speech. Experiments could be designed to present subjects with a pair of stop consonants in which the second element of the pair sometimes differed only in its transient burst amplitude. If the subjects were asked to judge whether the two elements were the same or different, the sensitivity of listeners to amplitude variations in transient bursts could be determined.

The analysis and synthesis of clicks could also provide valuable insight to the perception of transient bursts. The analysis of clicks (particularly abrupt clicks without frication noise) should be fairly easy, due to the fact that these clicks should not be difficult to separate from the surrounding noise due to their larger amplitudes.

Studies of stop releases have shown that the place of articulation can be determined from the burst portion alone [10]. In many instances the vowel preceding or following the stop could also be identified by only the burst portion, giving perceptual evidence for coarticulation. It would be interesting to determine if these same types of characteristics that exist for the stop consonant burst also exist for the transient burst alone. Now that a synthesizer is able to generate transients, studies should be performed in which listeners are presented with a transient in isolation. Listeners' ability to detect place of articulation from the transient alone could then be measured.

Appendix A

Most Preferred Synthesized Stimuli - .DOC Files

A.1 Most Preferred /pa/ Synthesis

```
Synthesis specification for file: 'xp40.wav'  Sat Feb  5 12:14:34 1994

KLSYN93   Version 2.0  April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is  -4.8 dB
Total number of waveform samples = 7000


  CURRENT CONFIGURATION:
    63 parameters

  SYM V/C  MIN  VAL   MAX  DESCRIPTION
----------------------------------------------
  DU   C   30   700  5000  Duration of the utterance, in msec
  UI   C    1     5    20  Update interval for parameter reset, in msec
  SR   C 5000 10000 20000  Output sampling rate, in samples/sec
  NF   C    1     5     6  Number of formants in cascade branch
  SS   C    1     2     3  Source switch (1=impulse, 2=natural, 3=LF model)
  RS   C    1     8  8191  Random seed (initial value of random # generator)
  SB   C    0     1     1  Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
  CP   C    0     0     1  0=Cascade, 1=Parallel tract excitation by AV
  OS   C    0     0    20  Output selector (0=normal,1=voicing source,...)
  GV   C    0    60    80  Overall gain scale factor for AV, in dB
  GH   C    0    60    80  Overall gain scale factor for AH, in dB
  GF   C    0    60    80  Overall gain scale factor for AF, in dB
  GI   C    0    60    80  Overall gain scale factor for AI, in dB
  FO   V    0  1000  5000  Fundamental frequency, in tenths of a Hz
  AV   V    0    60    80  Amplitude of voicing, in dB
  OQ   V   10    50    99  Open quotient (voicing open-time/period), in %
  SQ   v  100   200   500  Speed quotient (rise/fall time, LF model), in %
  TL   V    0     0    41  Extra tilt of voicing spectrum, dB down @ 3 kHz
  FL   v    0     0   100  Flutter (random fluct in f0), in % of maximum
  DI   v    0     0   100  Diplophonia (alt periods closer), in % of max
  AH   V    0     0    80  Amplitude of aspiration, in dB
  AF   V    0     0    80  Amplitude of frication, in dB
  F1   V  180   500  1300  Frequency of 1st formant, in Hz
```

B1	V	30	60	1000	Bandwidth of 1st formant, in Hz
DF1	v	0	0	100	Change in F1 during open portion of period, in Hz
DB1	v	0	0	400	Change in B1 during open portion of period, in Hz
F2	V	550	1500	3000	Frequency of 2nd formant, in Hz
B2	v	40	100	1000	Bandwidth of 2nd formant, in Hz
F3	V	1200	2500	4800	Frequency of 3rd formant, in Hz
B3	v	60	150	1000	Bandwidth of 3rd formant, in Hz
F4	v	2400	3500	4990	Frequency of 4th formant, in Hz
B4	v	100	300	1000	Bandwidth of 4th formant, in Hz
F5	v	3000	4400	4990	Frequency of 5th formant, in Hz
B5	v	100	400	1500	Bandwidth of 5th formant, in Hz
F6	v	3000	4990	4990	Frequency of 6th formant, in Hz (applies if NF=6)
B6	v	100	500	4000	Bandwidth of 6th formant, in Hz (applies if NF=6)
FNP	v	180	280	500	Frequency of nasal pole, in Hz
BNP	v	40	90	1000	Bandwidth of nasal pole, in Hz
FNZ	v	180	280	800	Frequency of nasal zero, in Hz
BNZ	v	40	90	1000	Bandwidth of nasal zero, in Hz
FTP	v	300	2150	3000	Frequency of tracheal pole, in Hz
BTP	v	40	180	1000	Bandwidth of tracheal pole, in Hz
FTZ	v	300	2150	3000	Frequency of tracheal zero, in Hz
BTZ	v	40	180	2000	Bandwidth of tracheal zero, in Hz
A2F	v	0	35	80	Amp of fric-excited parallel 2nd formant, in dB
A3F	v	0	25	80	Amp of fric-excited parallel 3rd formant, in dB
A4F	v	0	0	80	Amp of fric-excited parallel 4th formant, in dB
A5F	v	0	0	80	Amp of fric-excited parallel 5th formant, in dB
A6F	v	0	0	80	Amp of fric-excited parallel 6th formant, in dB
AB	v	0	45	80	Amp of fric-excited parallel bypass path, in dB
B2F	v	40	250	1000	Bw of fric-excited parallel 2nd formant, in Hz
B3F	v	60	300	1000	Bw of fric-excited parallel 3rd formant, in Hz
B4F	v	100	320	1000	Bw of fric-excited parallel 4th formant, in Hz
B5F	v	100	700	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	B1	F2	F3	AI
0	0	0	65	15	0	0	300	150	800	2500	0
5	0	0	65	15	0	0	300	150	800	2500	0
10	0	0	65	15	0	0	300	150	800	2500	0
15	0	0	65	15	0	0	300	150	800	2500	0
20	0	0	65	15	0	0	300	150	800	2500	0
25	0	0	65	15	0	0	300	150	800	2500	0
30	0	0	65	15	0	0	300	150	800	2500	0
35	0	0	65	15	0	0	300	150	800	2500	0
40	0	0	65	15	0	0	300	150	800	2500	0
45	0	0	65	15	0	0	300	150	800	2500	0
50	0	0	65	15	0	0	300	150	800	2500	0
55	0	0	65	15	0	0	300	150	800	2500	0
60	0	0	65	15	0	0	300	150	800	2500	0
65	0	0	65	15	0	0	300	150	800	2500	0
70	0	0	65	15	0	0	300	150	800	2500	0
75	0	0	65	15	0	0	300	150	800	2500	0
80	0	0	65	15	0	0	300	150	800	2500	0
85	0	0	65	15	0	0	300	150	800	2500	0
90	0	0	65	15	0	0	300	150	800	2500	0
95	0	0	65	15	0	0	300	150	800	2500	0
100	0	0	65	15	0	0	300	150	800	2500	0
105	0	0	65	15	0	0	300	150	800	2500	0

110	0	0	65	15	0	0	300	150	800	2500	0
115	0	0	65	15	0	0	300	150	800	2500	0
120	0	0	65	15	0	0	300	150	800	2500	0
125	0	0	65	15	0	0	300	150	800	2500	0
130	0	0	65	15	0	0	300	150	800	2500	0
135	0	0	65	15	0	0	300	150	800	2500	0
140	0	0	65	15	0	0	300	150	800	2500	0
145	0	0	65	15	0	0	300	150	800	2500	0
150	0	0	65	15	0	0	300	150	800	2500	0
155	0	0	65	15	0	0	300	150	800	2500	0
160	0	0	65	15	0	0	300	150	800	2500	0
165	0	0	65	15	0	0	300	150	800	2500	0
170	0	0	65	15	0	0	300	150	800	2500	0
175	0	0	65	15	0	0	300	150	800	2500	0
180	0	0	65	15	0	0	300	150	800	2500	0
185	0	0	65	15	0	0	300	150	800	2500	0
190	0	0	65	15	0	0	300	150	800	2500	0
195	0	0	65	15	0	0	300	150	800	2500	0
200	0	0	65	15	0	0	300	150	800	2500	0
205	0	0	65	15	0	0	300	150	800	2500	0
210	0	0	65	15	0	0	300	150	800	2500	0
215	0	0	65	15	0	0	300	150	800	2500	0
220	0	0	65	15	0	0	300	150	800	2500	0
225	0	0	65	15	0	0	300	150	800	2500	0
230	0	0	65	15	0	0	300	150	800	2500	0
235	0	0	65	15	0	0	300	150	800	2500	0
240	0	0	65	15	0	0	300	150	800	2500	0
245	0	0	65	15	0	65	300	150	800	2500	40
250	0	0	65	15	57	0	450	150	875	2500	0
255	0	0	65	15	57	0	500	150	950	2500	0
260	0	0	65	15	57	0	550	150	1025	2500	0
265	0	0	65	15	57	0	570	150	1050	2500	0
270	0	0	65	15	57	0	575	150	1052	2500	0
275	0	0	65	15	57	0	580	150	1054	2500	0
280	0	0	65	15	57	0	585	150	1056	2500	0
285	0	0	65	15	57	0	590	150	1059	2500	0
290	1300	55	62	13	54	0	595	150	1061	2500	0
295	1289	56	60	12	52	0	600	140	1063	2500	0
300	1278	57	58	10	50	0	605	130	1065	2500	0
305	1267	59	56	9	49	0	610	120	1065	2500	0
310	1256	60	54	7	47	0	615	110	1066	2500	0
315	1245	60	52	6	46	0	620	100	1066	2500	0
320	1235	60	50	5	45	0	623	90	1067	2500	0
325	1246	60	50	5	45	0	626	90	1067	2500	0
330	1257	60	50	5	45	0	629	90	1068	2500	0
335	1268	60	50	5	45	0	633	90	1068	2500	0
340	1279	60	50	5	45	0	636	90	1069	2500	0
345	1289	60	50	5	45	0	639	90	1069	2500	0
350	1300	60	50	5	45	0	643	90	1070	2500	0
355	1290	60	50	5	45	0	646	90	1070	2500	0
360	1280	60	51	5	45	0	649	90	1071	2500	0
365	1270	60	51	5	45	0	652	90	1072	2500	0
370	1260	60	51	5	45	0	656	90	1072	2500	0
375	1250	60	51	5	45	0	659	90	1073	2500	0
380	1240	60	51	5	45	0	662	90	1073	2500	0
385	1230	60	51	5	45	0	666	90	1074	2500	0
390	1220	60	51	5	45	0	669	90	1074	2500	0
395	1210	60	51	5	45	0	672	90	1075	2500	0
400	1200	60	51	5	45	0	675	90	1075	2500	0
405	1190	60	51	5	45	0	675	90	1075	2500	0
410	1180	60	51	5	45	0	675	90	1075	2500	0
415	1170	60	51	5	45	0	675	90	1075	2500	0
420	1160	60	51	5	45	0	675	90	1075	2500	0
425	1150	60	51	5	45	0	675	90	1075	2500	0
430	1140	60	52	6	45	0	675	90	1075	2500	0
435	1130	60	52	6	45	0	675	90	1075	2500	0
440	1120	60	52	6	45	0	675	90	1075	2500	0
445	1110	60	52	6	45	0	675	90	1075	2500	0

450	1100	60	52	6	45	0	675	90	1075	2500	0
455	1090	59	52	6	45	0	675	90	1075	2500	0
460	1080	59	52	6	45	0	675	90	1075	2500	0
465	1070	58	52	6	45	0	675	90	1075	2500	0
470	1060	58	52	6	45	0	675	90	1075	2500	0
475	1050	58	52	6	45	0	675	90	1075	2500	0
480	1040	57	52	6	45	0	675	90	1075	2500	0
485	1030	57	52	6	45	0	675	90	1075	2500	0
490	1020	57	52	6	45	0	675	90	1075	2500	0
495	1010	56	52	6	45	0	675	90	1075	2500	0
500	1000	56	53	6	45	0	675	90	1075	2500	0
505	995	56	53	6	45	0	675	90	1075	2500	0
510	990	55	53	6	45	0	675	90	1075	2500	0
515	985	55	53	6	45	0	675	90	1075	2500	0
520	980	55	53	6	45	0	675	90	1075	2500	0
525	975	54	53	6	44	0	675	90	1075	2500	0
530	970	54	53	6	44	0	675	90	1075	2500	0
535	965	53	53	6	44	0	675	90	1075	2500	0
540	960	53	53	6	44	0	675	90	1075	2500	0
545	955	53	53	6	44	0	675	90	1075	2500	0
550	950	52	53	6	44	0	675	90	1075	2500	0
555	945	52	53	6	44	0	675	90	1075	2500	0
560	940	52	53	6	43	0	675	90	1075	2500	0
565	935	51	54	6	43	0	675	90	1075	2500	0
570	930	51	54	7	43	0	675	90	1075	2500	0
575	925	51	54	7	43	0	675	90	1075	2500	0
580	920	50	54	7	43	0	675	90	1075	2500	0
585	915	50	54	7	43	0	675	90	1075	2500	0
590	910	50	54	7	43	0	675	90	1075	2500	0
595	905	45	54	7	40	0	675	90	1075	2500	0
600	900	40	54	7	35	0	675	90	1075	2500	0
605	900	35	54	7	30	0	675	90	1075	2500	0
610	900	25	54	7	25	0	675	90	1075	2500	0
615	0	0	54	7	20	0	675	90	1075	2500	0
620	0	0	54	7	17	0	675	90	1075	2500	0
625	0	0	54	7	15	0	675	90	1075	2500	0
630	0	0	54	7	7	0	675	90	1075	2500	0
635	0	0	55	7	0	0	675	90	1075	2500	0
640	0	0	55	7	0	0	675	90	1075	2500	0
645	0	0	55	7	0	0	675	90	1075	2500	0
650	0	0	55	7	0	0	675	90	1075	2500	0
655	0	0	55	7	0	0	675	90	1075	2500	0
660	0	0	55	7	0	0	675	90	1075	2500	0
665	0	0	55	7	0	0	675	90	1075	2500	0
670	0	0	55	7	0	0	675	90	1075	2500	0
675	0	0	55	7	0	0	675	90	1075	2500	0
680	0	0	55	7	0	0	675	90	1075	2500	0
685	0	0	55	7	0	0	675	90	1075	2500	0
690	0	0	55	7	0	0	675	90	1075	2500	0
695	0	0	55	7	0	0	675	90	1075	2500	0

A.2 Most Preferred /ta/ Synthesis

Synthesis specification for file: 'xt40.wav' Fri Apr 15 14:06:24 1994

KLSYN93 Version 2.0 April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is -5.2 dB
Total number of waveform samples = 7000

CURRENT CONFIGURATION:
63 parameters

SYM	V/C	MIN	VAL	MAX	DESCRIPTION
DU	C	30	700	5000	Duration of the utterance, in msec
UI	C	1	5	20	Update interval for parameter reset, in msec
SR	C	5000	10000	20000	Output sampling rate, in samples/sec
NF	C	1	5	6	Number of formants in cascade branch
SS	C	1	2	3	Source switch (1=impulse, 2=natural, 3=LF model)
RS	C	1	10	8191	Random seed (initial value of random # generator)
SB	C	0	1	1	Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
CP	C	0	0	1	0=Cascade, 1=Parallel tract excitation by AV
OS	C	0	0	20	Output selector (0=normal,1=voicing source,...)
GV	C	0	60	80	Overall gain scale factor for AV, in dB
GH	C	0	60	80	Overall gain scale factor for AH, in dB
GF	C	0	60	80	Overall gain scale factor for AF, in dB
GI	C	0	60	80	Overall gain scale factor for AI, in dB
FO	V	0	1000	5000	Fundamental frequency, in tenths of a Hz
AV	V	0	60	80	Amplitude of voicing, in dB
OQ	V	10	50	99	Open quotient (voicing open-time/period), in %
SQ	v	100	200	500	Speed quotient (rise/fall time, LF model), in %
TL	V	0	0	41	Extra tilt of voicing spectrum, dB down @ 3 kHz
FL	v	0	0	100	Flutter (random fluct in f0), in % of maximum
DI	v	0	0	100	Diplophonia (alt periods closer), in % of max
AH	V	0	0	80	Amplitude of aspiration, in dB
AF	V	0	0	80	Amplitude of frication, in dB
F1	V	180	500	1300	Frequency of 1st formant, in Hz
B1	V	30	60	1000	Bandwidth of 1st formant, in Hz
DF1	v	0	0	100	Change in F1 during open portion of period, in Hz
DB1	v	0	0	400	Change in B1 during open portion of period, in Hz
F2	V	550	1500	3000	Frequency of 2nd formant, in Hz
B2	v	40	100	1000	Bandwidth of 2nd formant, in Hz
F3	V	1200	2500	4800	Frequency of 3rd formant, in Hz
B3	v	60	150	1000	Bandwidth of 3rd formant, in Hz
F4	v	2400	3500	4990	Frequency of 4th formant, in Hz
B4	v	100	300	1000	Bandwidth of 4th formant, in Hz
F5	v	3000	4400	4990	Frequency of 5th formant, in Hz
B5	v	100	400	1500	Bandwidth of 5th formant, in Hz
F6	v	3000	4990	4990	Frequency of 6th formant, in Hz (applies if NF=6)
B6	v	100	500	4000	Bandwidth of 6th formant, in Hz (applies if NF=6)
FNP	v	180	280	500	Frequency of nasal pole, in Hz
BNP	v	40	90	1000	Bandwidth of nasal pole, in Hz
FNZ	v	180	280	800	Frequency of nasal zero, in Hz
BNZ	v	40	90	1000	Bandwidth of nasal zero, in Hz
FTP	v	300	2150	3000	Frequency of tracheal pole, in Hz
BTP	v	40	180	1000	Bandwidth of tracheal pole, in Hz
FTZ	v	300	2150	3000	Frequency of tracheal zero, in Hz
BTZ	v	40	180	2000	Bandwidth of tracheal zero, in Hz
A2F	v	0	0	80	Amp of fric-excited parallel 2nd formant, in dB
A3F	v	0	35	80	Amp of fric-excited parallel 3rd formant, in dB
A4F	v	0	25	80	Amp of fric-excited parallel 4th formant, in dB
A5F	v	0	0	80	Amp of fric-excited parallel 5th formant, in dB
A6F	v	0	0	80	Amp of fric-excited parallel 6th formant, in dB
AB	v	0	35	80	Amp of fric-excited parallel bypass path, in dB
B2F	v	40	250	1000	Bw of fric-excited parallel 2nd formant, in Hz
B3F	v	60	300	1000	Bw of fric-excited parallel 3rd formant, in Hz
B4F	v	100	320	1000	Bw of fric-excited parallel 4th formant, in Hz

B5F	v	100	700	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	B1	F2	F3	AI
0	0	0	65	20	0	0	300	150	1800	3100	0
5	0	0	65	20	0	0	300	150	1800	3100	0
10	0	0	65	20	0	0	300	150	1800	3100	0
15	0	0	65	20	0	0	300	150	1800	3100	0
20	0	0	65	20	0	0	300	150	1800	3100	0
25	0	0	65	20	0	0	300	150	1800	3100	0
30	0	0	65	20	0	0	300	150	1800	3100	0
35	0	0	65	20	0	0	300	150	1800	3100	0
40	0	0	65	20	0	0	300	150	1800	3100	0
45	0	0	65	20	0	0	300	150	1800	3100	0
50	0	0	65	20	0	0	300	150	1800	3100	0
55	0	0	65	20	0	0	300	150	1800	3100	0
60	0	0	65	20	0	0	300	150	1800	3100	0
65	0	0	65	20	0	0	300	150	1800	3100	0
70	0	0	65	20	0	0	300	150	1800	3100	0
75	0	0	65	20	0	0	300	150	1800	3100	0
80	0	0	65	20	0	0	300	150	1800	3100	0
85	0	0	65	20	0	0	300	150	1800	3100	0
90	0	0	65	20	0	0	300	150	1800	3100	0
95	0	0	65	20	0	0	300	150	1800	3100	0
100	0	0	65	20	0	0	300	150	1800	3100	0
105	0	0	65	20	0	0	300	150	1800	3100	0
110	0	0	65	20	0	0	300	150	1800	3100	0
115	0	0	65	20	0	0	300	150	1800	3100	0
120	0	0	65	20	0	0	300	150	1800	3100	0
125	0	0	65	20	0	0	300	150	1800	3100	0
130	0	0	65	20	0	0	300	150	1800	3100	0
135	0	0	65	20	0	0	300	150	1800	3100	0
140	0	0	65	20	0	0	300	150	1800	3100	0
145	0	0	65	20	0	0	300	150	1800	3100	0
150	0	0	65	20	0	0	300	150	1800	3100	0
155	0	0	65	20	0	0	300	150	1800	3100	0
160	0	0	65	20	0	0	300	150	1800	3100	0
165	0	0	65	20	0	0	300	150	1800	3100	0
170	0	0	65	20	0	0	300	150	1800	3100	0
175	0	0	65	20	0	0	300	150	1800	3100	0
180	0	0	65	20	0	0	300	150	1800	3100	0
185	0	0	65	20	0	0	300	150	1800	3100	0
190	0	0	65	20	0	0	300	150	1800	3100	0
195	0	0	65	20	0	0	300	150	1800	3100	0
200	0	0	65	20	0	0	300	150	1800	3100	0
205	0	0	65	20	0	0	300	150	1800	3100	0
210	0	0	65	20	0	0	300	150	1800	3100	0
215	0	0	65	20	0	0	300	150	1800	3100	0
220	0	0	65	20	0	0	300	150	1800	3100	0
225	0	0	65	20	0	0	300	150	1800	3100	0
230	0	0	65	20	0	0	300	150	1800	3100	0
235	0	0	65	20	0	0	300	150	1800	3100	0
240	0	0	65	20	0	0	300	150	1800	3100	0
245	0	0	65	20	0	70	300	150	1800	3100	40
250	0	0	65	20	40	0	450	150	1720	3020	0
255	0	0	65	20	40	0	500	150	1640	2940	0

260	0	0	65	20	40	0	550	150	1560	2860	0
265	0	0	65	20	40	0	570	150	1480	2780	0
270	0	0	65	20	40	0	575	150	1400	2700	0
275	0	0	65	20	40	0	580	150	1366	2683	0
280	0	0	65	20	40	0	585	150	1333	2666	0
285	0	0	65	20	40	0	590	150	1300	2650	0
290	1300	55	62	18	40	0	595	150	1266	2633	0
295	1289	56	60	16	40	0	600	140	1233	2616	0
300	1278	57	58	15	40	0	605	130	1200	2600	0
305	1267	59	56	13	40	0	610	120	1187	2590	0
310	1256	60	54	12	40	0	615	110	1174	2580	0
315	1245	60	52	11	40	0	620	100	1161	2570	0
320	1235	60	50	10	40	0	623	90	1148	2560	0
325	1246	60	50	10	42	0	626	90	1135	2550	0
330	1257	60	50	10	45	0	629	90	1122	2540	0
335	1268	60	50	10	45	0	633	90	1109	2530	0
340	1279	60	50	10	45	0	636	90	1096	2520	0
345	1289	60	50	10	45	0	639	90	1083	2510	0
350	1300	60	50	10	45	0	643	90	1070	2500	0
355	1290	60	50	10	45	0	646	90	1070	2500	0
360	1280	60	51	10	45	0	649	90	1071	2500	0
365	1270	60	51	10	45	0	652	90	1072	2500	0
370	1260	60	51	10	45	0	656	90	1072	2500	0
375	1250	60	51	10	45	0	659	90	1073	2500	0
380	1240	60	51	10	45	0	662	90	1073	2500	0
385	1230	60	51	10	45	0	666	90	1074	2500	0
390	1220	60	51	10	45	0	669	90	1074	2500	0
395	1210	60	51	10	45	0	672	90	1075	2500	0
400	1200	60	51	10	45	0	675	90	1075	2500	0
405	1190	60	51	10	45	0	675	90	1075	2500	0
410	1180	60	51	10	45	0	675	90	1075	2500	0
415	1170	60	51	10	45	0	675	90	1075	2500	0
420	1160	60	51	10	45	0	675	90	1075	2500	0
425	1150	60	51	10	45	0	675	90	1075	2500	0
430	1140	60	52	10	45	0	675	90	1075	2500	0
435	1130	60	52	10	45	0	675	90	1075	2500	0
440	1120	60	52	10	45	0	675	90	1075	2500	0
445	1110	60	52	10	45	0	675	90	1075	2500	0
450	1100	60	52	11	45	0	675	90	1075	2500	0
455	1090	59	52	11	45	0	675	90	1075	2500	0
460	1080	59	52	11	45	0	675	90	1075	2500	0
465	1070	58	52	11	45	0	675	90	1075	2500	0
470	1060	58	52	11	45	0	675	90	1075	2500	0
475	1050	58	52	11	45	0	675	90	1075	2500	0
480	1040	57	52	11	45	0	675	90	1075	2500	0
485	1030	57	52	11	45	0	675	90	1075	2500	0
490	1020	57	52	11	45	0	675	90	1075	2500	0
495	1010	56	52	11	45	0	675	90	1075	2500	0
500	1000	56	53	11	45	0	675	90	1075	2500	0
505	995	56	53	11	45	0	675	90	1075	2500	0
510	990	55	53	11	45	0	675	90	1075	2500	0
515	985	55	53	11	45	0	675	90	1075	2500	0
520	980	55	53	11	45	0	675	90	1075	2500	0
525	975	54	53	11	44	0	675	90	1075	2500	0
530	970	54	53	11	44	0	675	90	1075	2500	0
535	965	53	53	11	44	0	675	90	1075	2500	0
540	960	53	53	11	44	0	675	90	1075	2500	0
545	955	53	53	11	44	0	675	90	1075	2500	0
550	950	52	53	11	44	0	675	90	1075	2500	0
555	945	52	53	11	44	0	675	90	1075	2500	0
560	940	52	53	11	43	0	675	90	1075	2500	0
565	935	51	54	11	43	0	675	90	1075	2500	0
570	930	51	54	11	43	0	675	90	1075	2500	0
575	925	51	54	11	43	0	675	90	1075	2500	0
580	920	50	54	12	43	0	675	90	1075	2500	0
585	915	50	54	12	43	0	675	90	1075	2500	0
590	910	50	54	12	43	0	675	90	1075	2500	0
595	905	45	54	12	40	0	675	90	1075	2500	0

600	900	40	54	12	35	0	675	90	1075	2500	0
605	900	35	54	12	30	0	675	90	1075	2500	0
610	900	25	54	12	25	0	675	90	1075	2500	0
615	0	0	54	12	20	0	675	90	1075	2500	0
620	0	0	54	12	17	0	675	90	1075	2500	0
625	0	0	54	12	15	0	675	90	1075	2500	0
630	0	0	54	12	7	0	675	90	1075	2500	0
635	0	0	55	12	0	0	675	90	1075	2500	0
640	0	0	55	12	0	0	675	90	1075	2500	0
645	0	0	55	12	0	0	675	90	1075	2500	0
650	0	0	55	12	0	0	675	90	1075	2500	0
655	0	0	55	12	0	0	675	90	1075	2500	0
660	0	0	55	12	0	0	675	90	1075	2500	0
665	0	0	55	12	0	0	675	90	1075	2500	0
670	0	0	55	12	0	0	675	90	1075	2500	0
675	0	0	55	12	0	0	675	90	1075	2500	0
680	0	0	55	12	0	0	675	90	1075	2500	0
685	0	0	55	12	0	0	675	90	1075	2500	0
690	0	0	55	12	0	0	675	90	1075	2500	0
695	0	0	55	12	0	0	675	90	1075	2500	0

A.3 Most Preferred /ka/ Synthesis

Synthesis specification for file: 'xk10.wav' Fri Apr 15 14:06:53 1994

KLSYN93 Version 2.0 April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is -4.8 dB
Total number of waveform samples = 7000

CURRENT CONFIGURATION:
63 parameters

SYM	V/C	MIN	VAL	MAX	DESCRIPTION
DU	C	30	700	5000	Duration of the utterance, in msec
UI	C	1	5	20	Update interval for parameter reset, in msec
SR	C	5000	10000	20000	Output sampling rate, in samples/sec
NF	C	1	5	6	Number of formants in cascade branch
SS	C	1	2	3	Source switch (1=impulse, 2=natural, 3=LF model)
RS	C	1	8	8191	Random seed (initial value of random # generator)
SB	C	0	1	1	Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
CP	C	0	0	1	0=Cascade, 1=Parallel tract excitation by AV
OS	C	0	0	20	Output selector (0=normal,1=voicing source,...)
GV	C	0	60	80	Overall gain scale factor for AV, in dB
GH	C	0	60	80	Overall gain scale factor for AH, in dB
GF	C	0	60	80	Overall gain scale factor for AF, in dB
GI	C	0	60	80	Overall gain scale factor for AI, in dB
FO	V	0	1000	5000	Fundamental frequency, in tenths of a Hz
AV	V	0	60	80	Amplitude of voicing, in dB
OQ	V	10	50	99	Open quotient (voicing open-time/period), in %
SQ	v	100	200	500	Speed quotient (rise/fall time, LF model), in %
TL	V	0	0	41	Extra tilt of voicing spectrum, dB down @ 3 kHz
FL	v	0	0	100	Flutter (random fluct in f0), in % of maximum
DI	v	0	0	100	Diplophonia (alt periods closer), in % of max
AH	V	0	0	80	Amplitude of aspiration, in dB
AF	V	0	0	80	Amplitude of frication, in dB
F1	V	180	500	1300	Frequency of 1st formant, in Hz
B1	V	30	60	1000	Bandwidth of 1st formant, in Hz
DF1	v	0	0	100	Change in F1 during open portion of period, in Hz
DB1	v	0	0	400	Change in B1 during open portion of period, in Hz
F2	V	550	1500	3000	Frequency of 2nd formant, in Hz
B2	v	40	100	1000	Bandwidth of 2nd formant, in Hz
F3	V	1200	2500	4800	Frequency of 3rd formant, in Hz
B3	v	60	150	1000	Bandwidth of 3rd formant, in Hz
F4	v	2400	3500	4990	Frequency of 4th formant, in Hz
B4	v	100	300	1000	Bandwidth of 4th formant, in Hz
F5	v	3000	4400	4990	Frequency of 5th formant, in Hz
B5	v	100	400	1500	Bandwidth of 5th formant, in Hz
F6	v	3000	4990	4990	Frequency of 6th formant, in Hz (applies if NF=6)
B6	v	100	500	4000	Bandwidth of 6th formant, in Hz (applies if NF=6)
FNP	v	180	280	500	Frequency of nasal pole, in Hz
BNP	v	40	90	1000	Bandwidth of nasal pole, in Hz
FNZ	v	180	280	800	Frequency of nasal zero, in Hz
BNZ	v	40	90	1000	Bandwidth of nasal zero, in Hz
FTP	v	300	2150	3000	Frequency of tracheal pole, in Hz
BTP	v	40	180	1000	Bandwidth of tracheal pole, in Hz
FTZ	v	300	2150	3000	Frequency of tracheal zero, in Hz
BTZ	v	40	180	2000	Bandwidth of tracheal zero, in Hz
A2F	v	0	50	80	Amp of fric-excited parallel 2nd formant, in dB
A3F	v	0	0	80	Amp of fric-excited parallel 3rd formant, in dB
A4F	v	0	0	80	Amp of fric-excited parallel 4th formant, in dB
A5F	v	0	25	80	Amp of fric-excited parallel 5th formant, in dB
A6F	v	0	0	80	Amp of fric-excited parallel 6th formant, in dB
AB	v	0	30	80	Amp of fric-excited parallel bypass path, in dB
B2F	v	40	100	1000	Bw of fric-excited parallel 2nd formant, in Hz
B3F	v	60	300	1000	Bw of fric-excited parallel 3rd formant, in Hz
B4F	v	100	320	1000	Bw of fric-excited parallel 4th formant, in Hz

B5F	v	100	700	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	B1	F2	F3	AI
0	0	0	65	30	0	0	300	110	1500	1900	0
5	0	0	65	30	0	0	300	110	1500	1900	0
10	0	0	65	30	0	0	300	110	1500	1900	0
15	0	0	65	30	0	0	300	110	1500	1900	0
20	0	0	65	30	0	0	300	110	1500	1900	0
25	0	0	65	30	0	0	300	110	1500	1900	0
30	0	0	65	30	0	0	300	110	1500	1900	0
35	0	0	65	30	0	0	300	110	1500	1900	0
40	0	0	65	30	0	0	300	110	1500	1900	0
45	0	0	65	30	0	0	300	110	1500	1900	0
50	0	0	65	30	0	0	300	110	1500	1900	0
55	0	0	65	30	0	0	300	110	1500	1900	0
60	0	0	65	30	0	0	300	110	1500	1900	0
65	0	0	65	30	0	0	300	110	1500	1900	0
70	0	0	65	30	0	0	300	110	1500	1900	0
75	0	0	65	30	0	0	300	110	1500	1900	0
80	0	0	65	30	0	0	300	110	1500	1900	0
85	0	0	65	30	0	0	300	110	1500	1900	0
90	0	0	65	30	0	0	300	110	1500	1900	0
95	0	0	65	30	0	0	300	110	1500	1900	0
100	0	0	65	30	0	0	300	110	1500	1900	0
105	0	0	65	30	0	0	300	110	1500	1900	0
110	0	0	65	30	0	0	300	110	1500	1900	0
115	0	0	65	30	0	0	300	110	1500	1900	0
120	0	0	65	30	0	0	300	110	1500	1900	0
125	0	0	65	30	0	0	300	110	1500	1900	0
130	0	0	65	30	0	0	300	110	1500	1900	0
135	0	0	65	30	0	0	300	110	1500	1900	0
140	0	0	65	30	0	0	300	110	1500	1900	0
145	0	0	65	30	0	0	300	110	1500	1900	0
150	0	0	65	30	0	0	300	110	1500	1900	0
155	0	0	65	30	0	0	300	110	1500	1900	0
160	0	0	65	30	0	0	300	110	1500	1900	0
165	0	0	65	30	0	0	300	110	1500	1900	0
170	0	0	65	30	0	0	300	110	1500	1900	0
175	0	0	65	30	0	0	300	110	1500	1900	0
180	0	0	65	30	0	0	300	110	1500	1900	0
185	0	0	65	30	0	0	300	110	1500	1900	0
190	0	0	65	30	0	0	300	110	1500	1900	0
195	0	0	65	30	0	0	300	110	1500	1900	0
200	0	0	65	30	0	0	300	110	1500	1900	0
205	0	0	65	30	0	0	300	110	1500	1900	0
210	0	0	65	30	0	0	300	110	1500	1900	0
215	0	0	65	30	0	0	300	110	1500	1900	0
220	0	0	65	30	0	0	300	110	1500	1900	0
225	0	0	65	30	0	0	300	110	1500	1900	0
230	0	0	65	30	0	0	300	110	1500	1900	0
235	0	0	65	30	0	55	300	110	1500	1900	10
240	0	0	65	30	0	50	300	110	1500	1900	0
245	0	0	65	30	0	45	300	110	1500	1900	0
250	0	0	65	30	52	0	450	110	1459	1974	0
255	0	0	65	30	52	0	500	110	1418	2048	0

260	0	0	65	30	52	0	550	110	1377	2122	0
265	0	0	65	30	52	0	570	110	1350	2181	0
270	0	0	65	30	52	0	575	110	1323	2241	0
275	0	0	65	30	52	0	580	110	1296	2300	0
280	0	0	65	30	51	0	585	110	1269	2323	0
285	0	0	65	30	51	0	590	110	1242	2346	0
290	1300	55	62	26	50	0	595	110	1216	2369	0
295	1289	56	60	22	50	0	600	106	1205	2392	0
300	1278	57	58	19	50	0	605	103	1194	2415	0
305	1267	59	56	15	49	0	610	100	1183	2438	0
310	1256	60	54	12	47	0	615	96	1172	2461	0
315	1245	60	52	8	46	0	620	93	1169	2465	0
320	1235	60	50	5	45	0	623	90	1166	2470	0
325	1246	60	50	5	45	0	626	90	1163	2475	0
330	1257	60	50	5	45	0	629	90	1161	2480	0
335	1268	60	50	5	45	0	633	90	1158	2485	0
340	1279	60	50	5	45	0	636	90	1155	2490	0
345	1289	60	50	5	45	0	639	90	1152	2495	0
350	1300	60	50	5	45	0	643	90	1150	2500	0
355	1290	60	50	5	45	0	646	90	1070	2500	0
360	1280	60	51	5	45	0	649	90	1071	2500	0
365	1270	60	51	5	45	0	652	90	1072	2500	0
370	1260	60	51	5	45	0	656	90	1072	2500	0
375	1250	60	51	5	45	0	659	90	1073	2500	0
380	1240	60	51	5	45	0	662	90	1073	2500	0
385	1230	60	51	5	45	0	666	90	1074	2500	0
390	1220	60	51	5	45	0	669	90	1074	2500	0
395	1210	60	51	5	45	0	672	90	1075	2500	0
400	1200	60	51	5	45	0	675	90	1075	2500	0
405	1190	60	51	5	45	0	675	90	1075	2500	0
410	1180	60	51	5	45	0	675	90	1075	2500	0
415	1170	60	51	5	45	0	675	90	1075	2500	0
420	1160	60	51	5	45	0	675	90	1075	2500	0
425	1150	60	51	5	45	0	675	90	1075	2500	0
430	1140	60	52	6	45	0	675	90	1075	2500	0
435	1130	60	52	6	45	0	675	90	1075	2500	0
440	1120	60	52	6	45	0	675	90	1075	2500	0
445	1110	60	52	6	45	0	675	90	1075	2500	0
450	1100	60	52	6	45	0	675	90	1075	2500	0
455	1090	59	52	6	45	0	675	90	1075	2500	0
460	1080	59	52	6	45	0	675	90	1075	2500	0
465	1070	58	52	6	45	0	675	90	1075	2500	0
470	1060	58	52	6	45	0	675	90	1075	2500	0
475	1050	58	52	6	45	0	675	90	1075	2500	0
480	1040	57	52	6	45	0	675	90	1075	2500	0
485	1030	57	52	6	45	0	675	90	1075	2500	0
490	1020	57	52	6	45	0	675	90	1075	2500	0
495	1010	56	52	6	45	0	675	90	1075	2500	0
500	1000	56	53	6	45	0	675	90	1075	2500	0
505	995	56	53	6	45	0	675	90	1075	2500	0
510	990	55	53	6	45	0	675	90	1075	2500	0
515	985	55	53	6	45	0	675	90	1075	2500	0
520	980	55	53	6	45	0	675	90	1075	2500	0
525	975	54	53	6	44	0	675	90	1075	2500	0
530	970	54	53	6	44	0	675	90	1075	2500	0
535	965	53	53	6	44	0	675	90	1075	2500	0
540	960	53	53	6	44	0	675	90	1075	2500	0
545	955	53	53	6	44	0	675	90	1075	2500	0
550	950	52	53	6	44	0	675	90	1075	2500	0
555	945	52	53	6	44	0	675	90	1075	2500	0
560	940	52	53	6	43	0	675	90	1075	2500	0
565	935	51	54	6	43	0	675	90	1075	2500	0
570	930	51	54	7	43	0	675	90	1075	2500	0
575	925	51	54	7	43	0	675	90	1075	2500	0
580	920	50	54	7	43	0	675	90	1075	2500	0
585	915	50	54	7	43	0	675	90	1075	2500	0
590	910	50	54	7	43	0	675	90	1075	2500	0
595	905	45	54	7	40	0	675	90	1075	2500	0

600	900	40	54	7	35	0	675	90	1075	2500	0
605	900	35	54	7	30	0	675	90	1075	2500	0
610	900	25	54	7	25	0	675	90	1075	2500	0
615	0	0	54	7	20	0	675	90	1075	2500	0
620	0	0	54	7	17	0	675	90	1075	2500	0
625	0	0	54	7	15	0	675	90	1075	2500	0
630	0	0	54	7	7	0	675	90	1075	2500	0
635	0	0	55	7	0	0	675	90	1075	2500	0
640	0	0	55	7	0	0	675	90	1075	2500	0
645	0	0	55	7	0	0	675	90	1075	2500	0
650	0	0	55	7	0	0	675	90	1075	2500	0
655	0	0	55	7	0	0	675	90	1075	2500	0
660	0	0	55	7	0	0	675	90	1075	2500	0
665	0	0	55	7	0	0	675	90	1075	2500	0
670	0	0	55	7	0	0	675	90	1075	2500	0
675	0	0	55	7	0	0	675	90	1075	2500	0
680	0	0	55	7	0	0	675	90	1075	2500	0
685	0	0	55	7	0	0	675	90	1075	2500	0
690	0	0	55	7	0	0	675	90	1075	2500	0
695	0	0	55	7	0	0	675	90	1075	2500	0

A.4 Most Preferred /ba/ Synthesis

```
Synthesis specification for file: 'xb30.wav'  Fri Apr 15 14:07:15 1994

KLSYN93   Version 2.0  April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is  -4.9 dB
Total number of waveform samples = 7000


  CURRENT CONFIGURATION:
    63 parameters

  SYM V/C  MIN  VAL   MAX  DESCRIPTION
----------------------------------------------
  DU   C    30   700  5000  Duration of the utterance, in msec
  UI   C     1     5    20  Update interval for parameter reset, in msec
  SR   C  5000 10000 20000  Output sampling rate, in samples/sec
  NF   C     1     5     6  Number of formants in cascade branch
  SS   C     1     2     3  Source switch (1=impulse, 2=natural, 3=LF model)
  RS   C     1     8  8191  Random seed (initial value of random # generator)
  SB   C     0     1     1  Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
  CP   C     0     0     1  0=Cascade, 1=Parallel tract excitation by AV
  OS   C     0     0    20  Output selector (0=normal,1=voicing source,...)
  GV   C     0    60    80  Overall gain scale factor for AV, in dB
  GH   C     0    60    80  Overall gain scale factor for AH, in dB
  GF   C     0    60    80  Overall gain scale factor for AF, in dB
  GI   C     0    60    80  Overall gain scale factor for AI, in dB
  FO   V     0  1000  5000  Fundamental frequency, in tenths of a Hz
  AV   V     0    60    80  Amplitude of voicing, in dB
  OQ   V    10    50    99  Open quotient (voicing open-time/period), in %
  SQ   v   100   200   500  Speed quotient (rise/fall time, LF model), in %
  TL   V     0     0    41  Extra tilt of voicing spectrum, dB down @ 3 kHz
  FL   v     0     0   100  Flutter (random fluct in f0), in % of maximum
  DI   v     0     0   100  Diplophonia (alt periods closer), in % of max
  AH   V     0     0    80  Amplitude of aspiration, in dB
  AF   V     0     0    80  Amplitude of frication, in dB
  F1   V   180   500  1300  Frequency of 1st formant, in Hz
  B1   V    30    60  1000  Bandwidth of 1st formant, in Hz
  DF1  v     0     0   100  Change in F1 during open portion of period, in Hz
  DB1  v     0     0   400  Change in B1 during open portion of period, in Hz
  F2   V   550  1500  3000  Frequency of 2nd formant, in Hz
  B2   v    40   100  1000  Bandwidth of 2nd formant, in Hz
  F3   V  1200  2500  4800  Frequency of 3rd formant, in Hz
  B3   v    60   150  1000  Bandwidth of 3rd formant, in Hz
  F4   v  2400  3500  4990  Frequency of 4th formant, in Hz
  B4   v   100   300  1000  Bandwidth of 4th formant, in Hz
  F5   v  3000  4400  4990  Frequency of 5th formant, in Hz
  B5   v   100   400  1500  Bandwidth of 5th formant, in Hz
  F6   v  3000  4990  4990  Frequency of 6th formant, in Hz (applies if NF=6)
  B6   v   100   500  4000  Bandwidth of 6th formant, in Hz (applies if NF=6)
  FNP  v   180   280   500  Frequency of nasal pole, in Hz
  BNP  v    40    90  1000  Bandwidth of nasal pole, in Hz
  FNZ  v   180   280   800  Frequency of nasal zero, in Hz
  BNZ  v    40    90  1000  Bandwidth of nasal zero, in Hz
  FTP  v   300  2150  3000  Frequency of tracheal pole, in Hz
  BTP  v    40   180  1000  Bandwidth of tracheal pole, in Hz
  FTZ  v   300  2150  3000  Frequency of tracheal zero, in Hz
  BTZ  v    40   180  2000  Bandwidth of tracheal zero, in Hz
  A2F  v     0    35    80  Amp of fric-excited parallel 2nd formant, in dB
  A3F  v     0    25    80  Amp of fric-excited parallel 3rd formant, in dB
  A4F  v     0     0    80  Amp of fric-excited parallel 4th formant, in dB
  A5F  v     0    25    80  Amp of fric-excited parallel 5th formant, in dB
  A6F  v     0     0    80  Amp of fric-excited parallel 6th formant, in dB
  AB   v     0    50    80  Amp of fric-excited parallel bypass path, in dB
  B2F  v    40   250  1000  Bw of fric-excited parallel 2nd formant, in Hz
  B3F  v    60   300  1000  Bw of fric-excited parallel 3rd formant, in Hz
  B4F  v   100   320  1000  Bw of fric-excited parallel 4th formant, in Hz
```

B5F	v	100	700	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	B1	F2	F3	AI
0	0	0	50	15	0	0	300	30	800	2500	0
5	0	0	50	15	0	0	300	30	800	2500	0
10	0	0	50	15	0	0	300	30	800	2500	0
15	0	0	50	15	0	0	300	30	800	2500	0
20	0	0	50	15	0	0	300	30	800	2500	0
25	0	0	50	15	0	0	300	30	800	2500	0
30	0	0	50	15	0	0	300	30	800	2500	0
35	0	0	50	15	0	0	300	30	800	2500	0
40	0	0	50	15	0	0	300	30	800	2500	0
45	0	0	50	15	0	0	300	30	800	2500	0
50	0	0	50	15	0	0	300	30	800	2500	0
55	0	0	50	15	0	0	300	30	800	2500	0
60	0	0	50	15	0	0	300	30	800	2500	0
65	0	0	50	15	0	0	300	30	800	2500	0
70	0	0	50	15	0	0	300	30	800	2500	0
75	0	0	50	15	0	0	300	30	800	2500	0
80	0	0	50	15	0	0	300	30	800	2500	0
85	0	0	50	15	0	0	300	30	800	2500	0
90	0	0	50	15	0	0	300	30	800	2500	0
95	0	0	50	15	0	0	300	30	800	2500	0
100	0	0	50	15	0	0	300	30	800	2500	0
105	0	0	50	15	0	0	300	30	800	2500	0
110	0	0	50	15	0	0	300	30	800	2500	0
115	0	0	50	15	0	0	300	30	800	2500	0
120	0	0	50	15	0	0	300	30	800	2500	0
125	0	0	50	15	0	0	300	30	800	2500	0
130	0	0	50	15	0	0	300	30	800	2500	0
135	0	0	50	15	0	0	300	30	800	2500	0
140	0	0	50	15	0	0	300	30	800	2500	0
145	0	0	50	15	0	0	300	30	800	2500	0
150	0	0	50	15	0	0	300	30	800	2500	0
155	0	0	50	15	0	0	300	30	800	2500	0
160	0	0	50	15	0	0	300	30	800	2500	0
165	990	30	50	15	20	0	300	30	800	2500	0
170	997	32	50	15	22	0	300	30	800	2500	0
175	1004	34	50	15	25	0	300	30	800	2500	0
180	1011	34	50	15	25	0	300	30	800	2500	0
185	1018	34	50	15	25	0	300	30	800	2500	0
190	1025	34	50	15	25	0	300	30	800	2500	0
195	1031	34	50	15	25	0	300	30	800	2500	0
200	1038	34	50	15	25	0	300	30	800	2500	0
205	1045	34	50	15	25	0	300	30	800	2500	0
210	1052	34	50	15	25	0	300	30	800	2500	0
215	1059	34	50	15	25	0	300	30	800	2500	0
220	1066	34	50	15	25	0	300	30	800	2500	0
225	1073	34	50	15	25	0	300	30	800	2500	0
230	1080	34	50	15	25	0	300	30	800	2500	0
235	1086	34	50	15	25	0	300	30	800	2500	0
240	1093	34	50	15	25	0	300	30	800	2500	0
245	1100	34	50	15	25	60	300	30	800	2500	30
250	1107	60	50	10	45	0	450	120	875	2500	0
255	1114	60	50	5	45	0	500	90	950	2500	0

260	1121	60	50	5	45	0	550	90	1025	2500	0
265	1128	60	50	5	45	0	570	90	1050	2500	0
270	1135	60	50	5	45	0	575	90	1052	2500	0
275	1144	60	50	5	45	0	580	90	1054	2500	0
280	1153	60	50	5	45	0	585	90	1056	2500	0
285	1162	60	50	5	45	0	590	90	1059	2500	0
290	1171	60	50	5	45	0	595	90	1061	2500	0
295	1180	60	50	5	45	0	600	90	1063	2500	0
300	1191	60	50	5	45	0	605	90	1065	2500	0
305	1202	60	50	5	45	0	610	90	1065	2500	0
310	1213	60	50	5	45	0	615	90	1066	2500	0
315	1230	60	50	5	45	0	620	90	1066	2500	0
320	1235	60	50	5	45	0	623	90	1067	2500	0
325	1246	60	50	5	45	0	626	90	1067	2500	0
330	1257	60	50	5	45	0	629	90	1068	2500	0
335	1268	60	50	5	45	0	633	90	1068	2500	0
340	1279	60	50	5	45	0	636	90	1069	2500	0
345	1289	60	50	5	45	0	639	90	1069	2500	0
350	1300	60	50	5	45	0	643	90	1070	2500	0
355	1290	60	50	5	45	0	646	90	1070	2500	0
360	1280	60	51	5	45	0	649	90	1071	2500	0
365	1270	60	51	5	45	0	652	90	1072	2500	0
370	1260	60	51	5	45	0	656	90	1072	2500	0
375	1250	60	51	5	45	0	659	90	1073	2500	0
380	1240	60	51	5	45	0	662	90	1073	2500	0
385	1230	60	51	5	45	0	666	90	1074	2500	0
390	1220	60	51	5	45	0	669	90	1074	2500	0
395	1210	60	51	5	45	0	672	90	1075	2500	0
400	1200	60	51	5	45	0	675	90	1075	2500	0
405	1190	60	51	5	45	0	675	90	1075	2500	0
410	1180	60	51	5	45	0	675	90	1075	2500	0
415	1170	60	51	5	45	0	675	90	1075	2500	0
420	1160	60	51	5	45	0	675	90	1075	2500	0
425	1150	60	51	5	45	0	675	90	1075	2500	0
430	1140	60	52	6	45	0	675	90	1075	2500	0
435	1130	60	52	6	45	0	675	90	1075	2500	0
440	1120	60	52	6	45	0	675	90	1075	2500	0
445	1110	60	52	6	45	0	675	90	1075	2500	0
450	1100	60	52	6	45	0	675	90	1075	2500	0
455	1090	59	52	6	45	0	675	90	1075	2500	0
460	1080	59	52	6	45	0	675	90	1075	2500	0
465	1070	58	52	6	45	0	675	90	1075	2500	0
470	1060	58	52	6	45	0	675	90	1075	2500	0
475	1050	58	52	6	45	0	675	90	1075	2500	0
480	1040	57	52	6	45	0	675	90	1075	2500	0
485	1030	57	52	6	45	0	675	90	1075	2500	0
490	1020	57	52	6	45	0	675	90	1075	2500	0
495	1010	56	52	6	45	0	675	90	1075	2500	0
500	1000	56	53	6	45	0	675	90	1075	2500	0
505	995	56	53	6	45	0	675	90	1075	2500	0
510	990	55	53	6	45	0	675	90	1075	2500	0
515	985	55	53	6	45	0	675	90	1075	2500	0
520	980	55	53	6	45	0	675	90	1075	2500	0
525	975	54	53	6	44	0	675	90	1075	2500	0
530	970	54	53	6	44	0	675	90	1075	2500	0
535	965	53	53	6	44	0	675	90	1075	2500	0
540	960	53	53	6	44	0	675	90	1075	2500	0
545	955	53	53	6	44	0	675	90	1075	2500	0
550	950	52	53	6	44	0	675	90	1075	2500	0
555	945	52	53	6	44	0	675	90	1075	2500	0
560	940	52	53	6	43	0	675	90	1075	2500	0
565	935	51	54	6	43	0	675	90	1075	2500	0
570	930	51	54	7	43	0	675	90	1075	2500	0
575	925	51	54	7	43	0	675	90	1075	2500	0
580	920	50	54	7	43	0	675	90	1075	2500	0
585	915	50	54	7	43	0	675	90	1075	2500	0
590	910	50	54	7	43	0	675	90	1075	2500	0
595	905	45	54	7	40	0	675	90	1075	2500	0

600	900	40	54	7	35	0	675	90	1075	2500		0
605	900	35	54	7	30	0	675	90	1075	2500		0
610	900	25	54	7	25	0	675	90	1075	2500		0
615	0	0	54	7	20	0	675	90	1075	2500		0
620	0	0	54	7	17	0	675	90	1075	2500		0
625	0	0	54	7	15	0	675	90	1075	2500		0
630	0	0	54	7	7	0	675	90	1075	2500		0
635	0	0	55	7	0	0	675	90	1075	2500		0
640	0	0	55	7	0	0	675	90	1075	2500		0
645	0	0	55	7	0	0	675	90	1075	2500		0
650	0	0	55	7	0	0	675	90	1075	2500		0
655	0	0	55	7	0	0	675	90	1075	2500		0
660	0	0	55	7	0	0	675	90	1075	2500		0
665	0	0	55	7	0	0	675	90	1075	2500		0
670	0	0	55	7	0	0	675	90	1075	2500		0
675	0	0	55	7	0	0	675	90	1075	2500		0
680	0	0	55	7	0	0	675	90	1075	2500		0
685	0	0	55	7	0	0	675	90	1075	2500		0
690	0	0	55	7	0	0	675	90	1075	2500		0
695	0	0	55	7	0	0	675	90	1075	2500		0

A.5 Most Preferred /da/ Synthesis

Synthesis specification for file: 'xd50.wav' Fri Apr 15 14:07:40 1994

KLSYN93 Version 2.0 April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is -4.9 dB
Total number of waveform samples = 7000

CURRENT CONFIGURATION:
63 parameters

SYM	V/C	MIN	VAL	MAX	DESCRIPTION
DU	C	30	700	5000	Duration of the utterance, in msec
UI	C	1	5	20	Update interval for parameter reset, in msec
SR	C	5000	10000	20000	Output sampling rate, in samples/sec
NF	C	1	5	6	Number of formants in cascade branch
SS	C	1	2	3	Source switch (1=impulse, 2=natural, 3=LF model)
RS	C	1	8	8191	Random seed (initial value of random # generator)
SB	C	0	1	1	Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
CP	C	0	0	1	0=Cascade, 1=Parallel tract excitation by AV
OS	C	0	0	20	Output selector (0=normal,1=voicing source,...)
GV	C	0	60	80	Overall gain scale factor for AV, in dB
GH	C	0	60	80	Overall gain scale factor for AH, in dB
GF	C	0	60	80	Overall gain scale factor for AF, in dB
GI	C	0	60	80	Overall gain scale factor for AI, in dB
FO	V	0	1000	5000	Fundamental frequency, in tenths of a Hz
AV	V	0	60	80	Amplitude of voicing, in dB
OQ	V	10	50	99	Open quotient (voicing open-time/period), in %
SQ	v	100	200	500	Speed quotient (rise/fall time, LF model), in %
TL	V	0	0	41	Extra tilt of voicing spectrum, dB down @ 3 kHz
FL	v	0	0	100	Flutter (random fluct in f0), in % of maximum
DI	v	0	0	100	Diplophonia (alt periods closer), in % of max
AH	V	0	0	80	Amplitude of aspiration, in dB
AF	V	0	0	80	Amplitude of frication, in dB
F1	V	180	500	1300	Frequency of 1st formant, in Hz
B1	V	30	60	1000	Bandwidth of 1st formant, in Hz
DF1	v	0	0	100	Change in F1 during open portion of period, in Hz
DB1	v	0	0	400	Change in B1 during open portion of period, in Hz
F2	V	550	1500	3000	Frequency of 2nd formant, in Hz
B2	v	40	100	1000	Bandwidth of 2nd formant, in Hz
F3	V	1200	2500	4800	Frequency of 3rd formant, in Hz
B3	v	60	150	1000	Bandwidth of 3rd formant, in Hz
F4	v	2400	3500	4990	Frequency of 4th formant, in Hz
B4	v	100	300	1000	Bandwidth of 4th formant, in Hz
F5	v	3000	4400	4990	Frequency of 5th formant, in Hz
B5	v	100	400	1500	Bandwidth of 5th formant, in Hz
F6	v	3000	4990	4990	Frequency of 6th formant, in Hz (applies if NF=6)
B6	v	100	500	4000	Bandwidth of 6th formant, in Hz (applies if NF=6)
FNP	v	180	280	500	Frequency of nasal pole, in Hz
BNP	v	40	90	1000	Bandwidth of nasal pole, in Hz
FNZ	v	180	280	800	Frequency of nasal zero, in Hz
BNZ	v	40	90	1000	Bandwidth of nasal zero, in Hz
FTP	v	300	2150	3000	Frequency of tracheal pole, in Hz
BTP	v	40	180	1000	Bandwidth of tracheal pole, in Hz
FTZ	v	300	2150	3000	Frequency of tracheal zero, in Hz
BTZ	v	40	180	2000	Bandwidth of tracheal zero, in Hz
A2F	v	0	0	80	Amp of fric-excited parallel 2nd formant, in dB
A3F	v	0	35	80	Amp of fric-excited parallel 3rd formant, in dB
A4F	v	0	25	80	Amp of fric-excited parallel 4th formant, in dB
A5F	v	0	0	80	Amp of fric-excited parallel 5th formant, in dB
A6F	v	0	0	80	Amp of fric-excited parallel 6th formant, in dB
AB	v	0	35	80	Amp of fric-excited parallel bypass path, in dB
B2F	v	40	250	1000	Bw of fric-excited parallel 2nd formant, in Hz
B3F	v	60	300	1000	Bw of fric-excited parallel 3rd formant, in Hz
B4F	v	100	320	1000	Bw of fric-excited parallel 4th formant, in Hz

B5F	v	100	700	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	B1	F2	F3	AI
0	0	0	50	15	0	0	250	30	1800	3100	0
5	0	0	50	15	0	0	250	30	1800	3100	0
10	0	0	50	15	0	0	250	30	1800	3100	0
15	0	0	50	15	0	0	250	30	1800	3100	0
20	0	0	50	15	0	0	250	30	1800	3100	0
25	0	0	50	15	0	0	250	30	1800	3100	0
30	0	0	50	15	0	0	250	30	1800	3100	0
35	0	0	50	15	0	0	250	30	1800	3100	0
40	0	0	50	15	0	0	250	30	1800	3100	0
45	0	0	50	15	0	0	250	30	1800	3100	0
50	0	0	50	15	0	0	250	30	1800	3100	0
55	0	0	50	15	0	0	250	30	1800	3100	0
60	0	0	50	15	0	0	250	30	1800	3100	0
65	0	0	50	15	0	0	250	30	1800	3100	0
70	0	0	50	15	0	0	250	30	1800	3100	0
75	0	0	50	15	0	0	250	30	1800	3100	0
80	0	0	50	15	0	0	250	30	1800	3100	0
85	0	0	50	15	0	0	250	30	1800	3100	0
90	0	0	50	15	0	0	250	30	1800	3100	0
95	0	0	50	15	0	0	250	30	1800	3100	0
100	0	0	50	15	0	0	250	30	1800	3100	0
105	0	0	50	15	0	0	250	30	1800	3100	0
110	0	0	50	15	0	0	250	30	1800	3100	0
115	0	0	50	15	0	0	250	30	1800	3100	0
120	0	0	50	15	0	0	250	30	1800	3100	0
125	0	0	50	15	0	0	250	30	1800	3100	0
130	0	0	50	15	0	0	250	30	1800	3100	0
135	0	0	50	15	0	0	250	30	1800	3100	0
140	0	0	50	15	0	0	250	30	1800	3100	0
145	0	0	50	15	0	0	250	30	1800	3100	0
150	0	0	50	15	0	0	250	30	1800	3100	0
155	0	0	50	15	0	0	250	30	1800	3100	0
160	0	0	50	15	0	0	250	30	1800	3100	0
165	990	30	50	15	20	0	250	30	1800	3100	0
170	997	32	50	15	22	0	250	30	1800	3100	0
175	1004	33	50	15	25	0	250	30	1800	3100	0
180	1011	34	50	15	25	0	250	30	1800	3100	0
185	1018	34	50	15	25	0	250	30	1800	3100	0
190	1025	34	50	15	25	0	250	30	1800	3100	0
195	1031	34	50	15	25	0	250	30	1800	3100	0
200	1038	34	50	15	25	0	250	30	1800	3100	0
205	1045	34	50	15	25	0	250	30	1800	3100	0
210	1052	34	50	15	25	0	250	30	1800	3100	0
215	1059	34	50	15	25	0	250	30	1800	3100	0
220	1066	34	50	15	25	0	250	30	1800	3100	0
225	1073	34	50	15	25	0	250	30	1800	3100	0
230	1080	34	50	15	25	0	250	30	1800	3100	0
235	1086	32	50	15	25	0	250	30	1800	3100	0
240	1093	29	50	15	25	0	250	30	1800	3100	0
245	1100	28	50	15	25	60	250	30	1800	3100	50
250	1107	45	50	10	45	0	400	120	1720	3020	0
255	1114	60	50	5	45	0	500	90	1640	2940	0

260	1121	60	50	5	45	0	550	90	1560	2860	0
265	1128	60	50	5	45	0	570	90	1480	2780	0
270	1135	60	50	5	45	0	575	90	1400	2700	0
275	1144	60	50	5	45	0	580	90	1366	2683	0
280	1153	60	50	5	45	0	585	90	1333	2666	0
285	1162	60	50	5	45	0	590	90	1300	2650	0
290	1171	60	50	5	45	0	595	90	1266	2633	0
295	1180	60	50	5	45	0	600	90	1233	2616	0
300	1191	60	50	5	45	0	605	90	1200	2600	0
305	1202	60	50	5	45	0	610	90	1188	2590	0
310	1213	60	50	5	45	0	615	90	1176	2580	0
315	1230	60	50	5	45	0	620	90	1164	2570	0
320	1235	60	50	5	45	0	623	90	1152	2560	0
325	1246	60	50	5	45	0	626	90	1140	2550	0
330	1257	60	50	5	45	0	629	90	1129	2540	0
335	1268	60	50	5	45	0	633	90	1117	2530	0
340	1279	60	50	5	45	0	636	90	1105	2520	0
345	1289	60	50	5	45	0	639	90	1093	2510	0
350	1300	60	50	5	45	0	643	90	1081	2500	0
355	1290	60	50	5	45	0	646	90	1070	2500	0
360	1280	60	51	5	45	0	649	90	1071	2500	0
365	1270	60	51	5	45	0	652	90	1072	2500	0
370	1260	60	51	5	45	0	656	90	1072	2500	0
375	1250	60	51	5	45	0	659	90	1073	2500	0
380	1240	60	51	5	45	0	662	90	1073	2500	0
385	1230	60	51	5	45	0	666	90	1074	2500	0
390	1220	60	51	5	45	0	669	90	1074	2500	0
395	1210	60	51	5	45	0	672	90	1075	2500	0
400	1200	60	51	5	45	0	675	90	1075	2500	0
405	1190	60	51	5	45	0	675	90	1075	2500	0
410	1180	60	51	5	45	0	675	90	1075	2500	0
415	1170	60	51	5	45	0	675	90	1075	2500	0
420	1160	60	51	5	45	0	675	90	1075	2500	0
425	1150	60	51	5	45	0	675	90	1075	2500	0
430	1140	60	52	6	45	0	675	90	1075	2500	0
435	1130	60	52	6	45	0	675	90	1075	2500	0
440	1120	60	52	6	45	0	675	90	1075	2500	0
445	1110	60	52	6	45	0	675	90	1075	2500	0
450	1100	60	52	6	45	0	675	90	1075	2500	0
455	1090	59	52	6	45	0	675	90	1075	2500	0
460	1080	59	52	6	45	0	675	90	1075	2500	0
465	1070	58	52	6	45	0	675	90	1075	2500	0
470	1060	58	52	6	45	0	675	90	1075	2500	0
475	1050	58	52	6	45	0	675	90	1075	2500	0
480	1040	57	52	6	45	0	675	90	1075	2500	0
485	1030	57	52	6	45	0	675	90	1075	2500	0
490	1020	57	52	6	45	0	675	90	1075	2500	0
495	1010	56	52	6	45	0	675	90	1075	2500	0
500	1000	56	53	6	45	0	675	90	1075	2500	0
505	995	56	53	6	45	0	675	90	1075	2500	0
510	990	55	53	6	45	0	675	90	1075	2500	0
515	985	55	53	6	45	0	675	90	1075	2500	0
520	980	55	53	6	45	0	675	90	1075	2500	0
525	975	54	53	6	44	0	675	90	1075	2500	0
530	970	54	53	6	44	0	675	90	1075	2500	0
535	965	53	53	6	44	0	675	90	1075	2500	0
540	960	53	53	6	44	0	675	90	1075	2500	0
545	955	53	53	6	44	0	675	90	1075	2500	0
550	950	52	53	6	44	0	675	90	1075	2500	0
555	945	52	53	6	44	0	675	90	1075	2500	0
560	940	52	53	6	43	0	675	90	1075	2500	0
565	935	51	54	6	43	0	675	90	1075	2500	0
570	930	51	54	7	43	0	675	90	1075	2500	0
575	925	51	54	7	43	0	675	90	1075	2500	0
580	920	50	54	7	43	0	675	90	1075	2500	0
585	915	50	54	7	43	0	675	90	1075	2500	0
590	910	50	54	7	43	0	675	90	1075	2500	0
595	905	45	54	7	40	0	675	90	1075	2500	0

600	900	40	54	7	35	0	675	90	1075	2500	0
605	900	35	54	7	30	0	675	90	1075	2500	0
610	900	25	54	7	25	0	675	90	1075	2500	0
615	0	0	54	7	20	0	675	90	1075	2500	0
620	0	0	54	7	17	0	675	90	1075	2500	0
625	0	0	54	7	15	0	675	90	1075	2500	0
630	0	0	54	7	7	0	675	90	1075	2500	0
635	0	0	55	7	0	0	675	90	1075	2500	0
640	0	0	55	7	0	0	675	90	1075	2500	0
645	0	0	55	7	0	0	675	90	1075	2500	0
650	0	0	55	7	0	0	675	90	1075	2500	0
655	0	0	55	7	0	0	675	90	1075	2500	0
660	0	0	55	7	0	0	675	90	1075	2500	0
665	0	0	55	7	0	0	675	90	1075	2500	0
670	0	0	55	7	0	0	675	90	1075	2500	0
675	0	0	55	7	0	0	675	90	1075	2500	0
680	0	0	55	7	0	0	675	90	1075	2500	0
685	0	0	55	7	0	0	675	90	1075	2500	0
690	0	0	55	7	0	0	675	90	1075	2500	0
695	0	0	55	7	0	0	675	90	1075	2500	0

A.6 Most Preferred /ga/ Synthesis

```
Synthesis specification for file: 'xg20.wav'  Fri Apr 15 14:08:03 1994

KLSYN93   Version 2.0  April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is  -4.9 dB
Total number of waveform samples = 7000


  CURRENT CONFIGURATION:
    63 parameters

  SYM V/C  MIN  VAL   MAX  DESCRIPTION
----------------------------------------------
  DU   C    30   700  5000  Duration of the utterance, in msec
  UI   C     1     5    20  Update interval for parameter reset, in msec
  SR   C  5000 10000 20000  Output sampling rate, in samples/sec
  NF   C     1     5     6  Number of formants in cascade branch
  SS   C     1     2     3  Source switch (1=impulse, 2=natural, 3=LF model)
  RS   C     1     8  8191  Random seed (initial value of random # generator)
  SB   C     0     1     1  Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
  CP   C     0     0     1  0=Cascade, 1=Parallel tract excitation by AV
  OS   C     0     0    20  Output selector (0=normal,1=voicing source,...)
  GV   C     0    60    80  Overall gain scale factor for AV, in dB
  GH   C     0    60    80  Overall gain scale factor for AH, in dB
  GF   C     0    60    80  Overall gain scale factor for AF, in dB
  GI   C     0    60    80  Overall gain scale factor for AI, in dB
  FO   V     0  1000  5000  Fundamental frequency, in tenths of a Hz
  AV   V     0    60    80  Amplitude of voicing, in dB
  OQ   V    10    50    99  Open quotient (voicing open-time/period), in %
  SQ   v   100   200   500  Speed quotient (rise/fall time, LF model), in %
  TL   V     0     0    41  Extra tilt of voicing spectrum, dB down @ 3 kHz
  FL   v     0     0   100  Flutter (random fluct in f0), in % of maximum
  DI   v     0     0   100  Diplophonia (alt periods closer), in % of max
  AH   V     0     0    80  Amplitude of aspiration, in dB
  AF   V     0     0    80  Amplitude of frication, in dB
  F1   V   180   500  1300  Frequency of 1st formant, in Hz
  B1   V    30    60  1000  Bandwidth of 1st formant, in Hz
  DF1  v     0     0   100  Change in F1 during open portion of period, in Hz
  DB1  v     0     0   400  Change in B1 during open portion of period, in Hz
  F2   V   550  1500  3000  Frequency of 2nd formant, in Hz
  B2   v    40   100  1000  Bandwidth of 2nd formant, in Hz
  F3   V  1200  2500  4800  Frequency of 3rd formant, in Hz
  B3   v    60   150  1000  Bandwidth of 3rd formant, in Hz
  F4   v  2400  3500  4990  Frequency of 4th formant, in Hz
  B4   v   100   300  1000  Bandwidth of 4th formant, in Hz
  F5   v  3000  4400  4990  Frequency of 5th formant, in Hz
  B5   v   100   400  1500  Bandwidth of 5th formant, in Hz
  F6   v  3000  4990  4990  Frequency of 6th formant, in Hz (applies if NF=6)
  B6   v   100   500  4000  Bandwidth of 6th formant, in Hz (applies if NF=6)
  FNP  v   180   280   500  Frequency of nasal pole, in Hz
  BNP  v    40    90  1000  Bandwidth of nasal pole, in Hz
  FNZ  v   180   280   800  Frequency of nasal zero, in Hz
  BNZ  v    40    90  1000  Bandwidth of nasal zero, in Hz
  FTP  v   300  2150  3000  Frequency of tracheal pole, in Hz
  BTP  v    40   180  1000  Bandwidth of tracheal pole, in Hz
  FTZ  v   300  2150  3000  Frequency of tracheal zero, in Hz
  BTZ  v    40   180  2000  Bandwidth of tracheal zero, in Hz
  A2F  v     0    50    80  Amp of fric-excited parallel 2nd formant, in dB
  A3F  v     0     0    80  Amp of fric-excited parallel 3rd formant, in dB
  A4F  v     0     0    80  Amp of fric-excited parallel 4th formant, in dB
  A5F  v     0    25    80  Amp of fric-excited parallel 5th formant, in dB
  A6F  v     0     0    80  Amp of fric-excited parallel 6th formant, in dB
  AB   v     0    30    80  Amp of fric-excited parallel bypass path, in dB
  B2F  v    40   250  1000  Bw of fric-excited parallel 2nd formant, in Hz
  B3F  v    60   300  1000  Bw of fric-excited parallel 3rd formant, in Hz
  B4F  v   100   320  1000  Bw of fric-excited parallel 4th formant, in Hz
```

B5F	v	100	700	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	B1	F2	F3	AI
0	0	0	50	15	0	0	250	30	1500	1900	0
5	0	0	50	15	0	0	250	30	1500	1900	0
10	0	0	50	15	0	0	250	30	1500	1900	0
15	0	0	50	15	0	0	250	30	1500	1900	0
20	0	0	50	15	0	0	250	30	1500	1900	0
25	0	0	50	15	0	0	250	30	1500	1900	0
30	0	0	50	15	0	0	250	30	1500	1900	0
35	0	0	50	15	0	0	250	30	1500	1900	0
40	0	0	50	15	0	0	250	30	1500	1900	0
45	0	0	50	15	0	0	250	30	1500	1900	0
50	0	0	50	15	0	0	250	30	1500	1900	0
55	0	0	50	15	0	0	250	30	1500	1900	0
60	0	0	50	15	0	0	250	30	1500	1900	0
65	0	0	50	15	0	0	250	30	1500	1900	0
70	0	0	50	15	0	0	250	30	1500	1900	0
75	0	0	50	15	0	0	250	30	1500	1900	0
80	0	0	50	15	0	0	250	30	1500	1900	0
85	0	0	50	15	0	0	250	30	1500	1900	0
90	0	0	50	15	0	0	250	30	1500	1900	0
95	0	0	50	15	0	0	250	30	1500	1900	0
100	0	0	50	15	0	0	250	30	1500	1900	0
105	0	0	50	15	0	0	250	30	1500	1900	0
110	0	0	50	15	0	0	250	30	1500	1900	0
115	0	0	50	15	0	0	250	30	1500	1900	0
120	0	0	50	15	0	0	250	30	1500	1900	0
125	0	0	50	15	0	0	250	30	1500	1900	0
130	0	0	50	15	0	0	250	30	1500	1900	0
135	0	0	50	15	0	0	250	30	1500	1900	0
140	0	0	50	15	0	0	250	30	1500	1900	0
145	0	0	50	15	0	0	250	30	1500	1900	0
150	0	0	50	15	0	0	250	30	1500	1900	0
155	0	0	50	15	0	0	250	30	1500	1900	0
160	0	0	50	15	0	0	250	30	1500	1900	0
165	990	30	50	15	20	0	250	30	1500	1900	0
170	997	32	50	15	22	0	250	30	1500	1900	0
175	1004	34	50	15	25	0	250	30	1500	1900	0
180	1011	34	50	15	25	0	250	30	1500	1900	0
185	1018	34	50	15	25	0	250	30	1500	1900	0
190	1025	34	50	15	25	0	250	30	1500	1900	0
195	1031	34	50	15	25	0	250	30	1500	1900	0
200	1038	34	50	15	25	0	250	30	1500	1900	0
205	1045	34	50	15	25	0	250	30	1500	1900	0
210	1052	34	50	15	25	0	250	30	1500	1900	0
215	1059	34	50	15	25	0	250	30	1500	1900	0
220	1066	34	50	15	25	0	250	30	1500	1900	0
225	1073	34	50	15	25	0	250	30	1500	1900	0
230	1080	34	50	15	25	0	250	30	1500	1900	0
235	1086	34	50	15	25	62	250	30	1500	1900	20
240	1093	34	50	15	25	62	250	30	1500	1900	0
245	1100	34	50	15	25	60	250	30	1500	1900	0
250	1107	60	50	10	45	0	450	120	1459	1974	0
255	1114	60	50	5	45	0	500	90	1418	2048	0

260	1121	60	50	5	45	0	550	90	1377	2122		0
265	1128	60	50	5	45	0	570	90	1350	2181		0
270	1135	60	50	5	45	0	575	90	1323	2241		0
275	1144	60	50	5	45	0	580	90	1296	2300		0
280	1153	60	50	5	45	0	585	90	1269	2323		0
285	1162	60	50	5	45	0	590	90	1242	2346		0
290	1171	60	50	5	45	0	595	90	1216	2369		0
295	1180	60	50	5	45	0	600	90	1205	2392		0
300	1191	60	50	5	45	0	605	90	1200	2415		0
305	1202	60	50	5	45	0	610	90	1195	2438		0
310	1213	60	50	5	45	0	615	90	1190	2461		0
315	1230	60	50	5	45	0	620	90	1185	2465		0
320	1235	60	50	5	45	0	623	90	1180	2470		0
325	1246	60	50	5	45	0	626	90	1175	2475		0
330	1257	60	50	5	45	0	629	90	1170	2480		0
335	1268	60	50	5	45	0	633	90	1165	2485		0
340	1279	60	50	5	45	0	636	90	1160	2490		0
345	1289	60	50	5	45	0	639	90	1155	2495		0
350	1300	60	50	5	45	0	643	90	1150	2500		0
355	1290	60	50	5	45	0	646	90	1070	2500		0
360	1280	60	51	5	45	0	649	90	1070	2500		0
365	1270	60	51	5	45	0	652	90	1071	2500		0
370	1260	60	51	5	45	0	656	90	1072	2500		0
375	1250	60	51	5	45	0	659	90	1072	2500		0
380	1240	60	51	5	45	0	662	90	1073	2500		0
385	1230	60	51	5	45	0	666	90	1074	2500		0
390	1220	60	51	5	45	0	669	90	1075	2500		0
395	1210	60	51	5	45	0	672	90	1075	2500		0
400	1200	60	51	5	45	0	675	90	1075	2500		0
405	1190	60	51	5	45	0	675	90	1075	2500		0
410	1180	60	51	5	45	0	675	90	1075	2500		0
415	1170	60	51	5	45	0	675	90	1075	2500		0
420	1160	60	51	5	45	0	675	90	1075	2500		0
425	1150	60	51	5	45	0	675	90	1075	2500		0
430	1140	60	52	6	45	0	675	90	1075	2500		0
435	1130	60	52	6	45	0	675	90	1075	2500		0
440	1120	60	52	6	45	0	675	90	1075	2500		0
445	1110	60	52	6	45	0	675	90	1075	2500		0
450	1100	60	52	6	45	0	675	90	1075	2500		0
455	1090	59	52	6	45	0	675	90	1075	2500		0
460	1080	59	52	6	45	0	675	90	1075	2500		0
465	1070	58	52	6	45	0	675	90	1075	2500		0
470	1060	58	52	6	45	0	675	90	1075	2500		0
475	1050	58	52	6	45	0	675	90	1075	2500		0
480	1040	57	52	6	45	0	675	90	1075	2500		0
485	1030	57	52	6	45	0	675	90	1075	2500		0
490	1020	57	52	6	45	0	675	90	1075	2500		0
495	1010	56	52	6	45	0	675	90	1075	2500		0
500	1000	56	53	6	45	0	675	90	1075	2500		0
505	995	56	53	6	45	0	675	90	1075	2500		0
510	990	55	53	6	45	0	675	90	1075	2500		0
515	985	55	53	6	45	0	675	90	1075	2500		0
520	980	55	53	6	45	0	675	90	1075	2500		0
525	975	54	53	6	44	0	675	90	1075	2500		0
530	970	54	53	6	44	0	675	90	1075	2500		0
535	965	53	53	6	44	0	675	90	1075	2500		0
540	960	53	53	6	44	0	675	90	1075	2500		0
545	955	53	53	6	44	0	675	90	1075	2500		0
550	950	52	53	6	44	0	675	90	1075	2500		0
555	945	52	53	6	44	0	675	90	1075	2500		0
560	940	52	53	6	43	0	675	90	1075	2500		0
565	935	51	54	6	43	0	675	90	1075	2500		0
570	930	51	54	7	43	0	675	90	1075	2500		0
575	925	51	54	7	43	0	675	90	1075	2500		0
580	920	50	54	7	43	0	675	90	1075	2500		0
585	915	50	54	7	43	0	675	90	1075	2500		0
590	910	50	54	7	43	0	675	90	1075	2500		0
595	905	45	54	7	40	0	675	90	1075	2500		0

600	900	40	54	7	35	0	675	90	1075	2500	0
605	900	35	54	7	30	0	675	90	1075	2500	0
610	900	25	54	7	25	0	675	90	1075	2500	0
615	0	0	54	7	20	0	675	90	1075	2500	0
620	0	0	54	7	17	0	675	90	1075	2500	0
625	0	0	54	7	15	0	675	90	1075	2500	0
630	0	0	54	7	7	0	675	90	1075	2500	0
635	0	0	55	7	0	0	675	90	1075	2500	0
640	0	0	55	7	0	0	675	90	1075	2500	0
645	0	0	55	7	0	0	675	90	1075	2500	0
650	0	0	55	7	0	0	675	90	1075	2500	0
655	0	0	55	7	0	0	675	90	1075	2500	0
660	0	0	55	7	0	0	675	90	1075	2500	0
665	0	0	55	7	0	0	675	90	1075	2500	0
670	0	0	55	7	0	0	675	90	1075	2500	0
675	0	0	55	7	0	0	675	90	1075	2500	0
680	0	0	55	7	0	0	675	90	1075	2500	0
685	0	0	55	7	0	0	675	90	1075	2500	0
690	0	0	55	7	0	0	675	90	1075	2500	0
695	0	0	55	7	0	0	675	90	1075	2500	0

A.7 Most Preferred /ča/ Synthesis

```
Synthesis specification for file: 'xch30.wav'  Thu Apr 21 23:29:06 1994

KLSYN93   Version 2.0  April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is  -5.7 dB
Total number of waveform samples = 7000


  CURRENT CONFIGURATION:
    63 parameters

  SYM V/C  MIN  VAL   MAX  DESCRIPTION
----------------------------------------------
  DU   C    30   700  5000  Duration of the utterance, in msec
  UI   C     1     5    20  Update interval for parameter reset, in msec
  SR   C 5000 10000 20000  Output sampling rate, in samples/sec
  NF   C     1     5     6  Number of formants in cascade branch
  SS   C     1     2     3  Source switch (1=impulse, 2=natural, 3=LF model)
  RS   C     1     8  8191  Random seed (initial value of random # generator)
  SB   C     0     1     1  Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
  CP   C     0     0     1  0=Cascade, 1=Parallel tract excitation by AV
  OS   C     0     0    20  Output selector (0=normal,1=voicing source,...)
  GV   C     0    60    80  Overall gain scale factor for AV, in dB
  GH   C     0    60    80  Overall gain scale factor for AH, in dB
  GF   C     0    60    80  Overall gain scale factor for AF, in dB
  GI   C     0    60    80  Overall gain scale factor for AI, in dB
  FO   V     0  1000  5000  Fundamental frequency, in tenths of a Hz
  AV   V     0    60    80  Amplitude of voicing, in dB
  OQ   V    10    50    99  Open quotient (voicing open-time/period), in %
  SQ   v   100   200   500  Speed quotient (rise/fall time, LF model), in %
  TL   V     0     0    41  Extra tilt of voicing spectrum, dB down @ 3 kHz
  FL   v     0     0   100  Flutter (random fluct in f0), in % of maximum
  DI   v     0     0   100  Diplophonia (alt periods closer), in % of max
  AH   V     0     0    80  Amplitude of aspiration, in dB
  AF   V     0     0    80  Amplitude of frication, in dB
  F1   V   180   500  1300  Frequency of 1st formant, in Hz
  B1   v    30    90  1000  Bandwidth of 1st formant, in Hz
  DF1  v     0     0   100  Change in F1 during open portion of period, in Hz
  DB1  v     0     0   400  Change in B1 during open portion of period, in Hz
  F2   V   550  1500  3000  Frequency of 2nd formant, in Hz
  B2   v    40    90  1000  Bandwidth of 2nd formant, in Hz
  F3   V  1200  2500  4800  Frequency of 3rd formant, in Hz
  B3   v    60   150  1000  Bandwidth of 3rd formant, in Hz
  F4   V  2400  3250  4990  Frequency of 4th formant, in Hz
  B4   v   100   200  1000  Bandwidth of 4th formant, in Hz
  F5   V  3000  3700  4990  Frequency of 5th formant, in Hz
  B5   v   100   200  1500  Bandwidth of 5th formant, in Hz
  F6   v  3000  4990  4990  Frequency of 6th formant, in Hz (applies if NF=6)
  B6   v   100   500  4000  Bandwidth of 6th formant, in Hz (applies if NF=6)
  FNP  v   180   280   500  Frequency of nasal pole, in Hz
  BNP  v    40    90  1000  Bandwidth of nasal pole, in Hz
  FNZ  v   180   280   800  Frequency of nasal zero, in Hz
  BNZ  v    40    90  1000  Bandwidth of nasal zero, in Hz
  FTP  v   300  2150  3000  Frequency of tracheal pole, in Hz
  BTP  v    40   180  1000  Bandwidth of tracheal pole, in Hz
  FTZ  v   300  2150  3000  Frequency of tracheal zero, in Hz
  BTZ  v    40   180  2000  Bandwidth of tracheal zero, in Hz
  A2F  v     0     0    80  Amp of fric-excited parallel 2nd formant, in dB
  A3F  V     0    60    80  Amp of fric-excited parallel 3rd formant, in dB
  A4F  V     0    63    80  Amp of fric-excited parallel 4th formant, in dB
  A5F  v     0     0    80  Amp of fric-excited parallel 5th formant, in dB
  A6F  v     0     0    80  Amp of fric-excited parallel 6th formant, in dB
  AB   v     0    30    80  Amp of fric-excited parallel bypass path, in dB
  B2F  v    40   250  1000  Bw of fric-excited parallel 2nd formant, in Hz
  B3F  v    60   300  1000  Bw of fric-excited parallel 3rd formant, in Hz
  B4F  v   100   320  1000  Bw of fric-excited parallel 4th formant, in Hz
```

B5F	v	100	360	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	F2	F3	F4	F5	A3F	A4F	AI
0	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
5	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
10	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
15	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
20	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
25	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
30	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
35	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
40	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
45	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
50	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
55	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
60	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
65	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
70	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
75	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
80	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
85	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
90	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
95	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
100	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
105	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
110	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
115	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
120	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
125	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
130	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
135	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
140	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
145	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
150	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
155	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
160	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
165	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
170	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
175	0	0	65	20	0	0	400	1700	2600	3500	4300	58	0	0
180	0	0	65	20	45	40	400	1700	2600	3500	4300	58	0	30
185	0	0	65	20	45	41	405	1691	2590	3485	4300	58	0	0
190	0	0	65	20	45	42	410	1682	2580	3471	4300	58	0	0
195	0	0	65	20	45	44	415	1673	2570	3457	4300	58	0	0
200	0	0	65	20	45	45	420	1664	2560	3442	4300	59	0	0
205	0	0	65	20	45	46	425	1655	2550	3428	4300	61	0	0
210	0	0	65	20	45	46	430	1647	2540	3414	4300	63	0	0
215	0	0	65	20	45	46	435	1638	2530	3400	4300	65	0	0
220	0	0	65	20	45	46	440	1629	2520	3385	4300	67	0	0
225	0	0	65	20	45	46	445	1620	2510	3371	4300	68	0	0
230	0	0	65	20	45	46	450	1611	2500	3357	4300	68	0	0
235	0	0	65	20	46	43	455	1602	2490	3342	4300	68	0	0
240	0	0	65	20	48	40	460	1594	2480	3328	4300	68	0	0
245	0	0	65	20	51	38	465	1585	2470	3314	4300	68	0	0
250	0	0	65	20	54	35	470	1576	2460	3300	4300	68	0	0
255	0	0	65	20	57	30	475	1567	2450	3300	4300	68	0	0

260	0	0	65	20	60	25	480	1558	2440	3300	4300	68	0	0
265	0	0	65	20	60	22	485	1550	2430	3300	4300	68	0	0
270	0	0	65	20	60	14	490	1523	2420	3300	4300	68	0	0
275	1299	50	65	18	55	5	495	1497	2410	3300	4300	68	0	0
280	1294	55	65	16	50	0	500	1470	2400	3300	4300	68	0	0
285	1288	57	65	14	45	0	500	1444	2400	3300	4300	68	0	0
290	1300	60	62	12	45	0	550	1417	2400	3300	4300	68	0	0
295	1289	60	60	10	45	0	564	1391	2400	3300	4300	68	0	0
300	1278	60	58	10	45	0	578	1364	2400	3300	4300	68	0	0
305	1267	60	56	10	45	0	593	1338	2400	3300	4300	68	0	0
310	1256	60	54	10	45	0	607	1311	2400	3300	4300	68	0	0
315	1245	60	52	10	45	0	622	1285	2400	3300	4300	68	0	0
320	1235	60	50	10	45	0	636	1258	2400	3300	4300	68	0	0
325	1246	60	50	10	45	0	650	1232	2400	3300	4300	68	0	0
330	1257	60	50	10	45	0	650	1205	2400	3300	4300	68	0	0
335	1268	60	50	10	45	0	650	1179	2400	3300	4300	68	0	0
340	1279	60	50	10	45	0	650	1152	2400	3300	4300	68	0	0
345	1289	60	50	10	45	0	650	1126	2400	3300	4300	68	0	0
350	1300	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
355	1290	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
360	1280	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
365	1270	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
370	1260	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
375	1250	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
380	1240	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
385	1230	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
390	1220	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
395	1210	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
400	1200	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
405	1190	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
410	1180	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
415	1170	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
420	1160	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
425	1150	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
430	1140	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
435	1130	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
440	1120	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
445	1110	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
450	1100	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
455	1090	59	51	10	45	0	650	1100	2400	3300	4300	68	0	0
460	1080	59	51	10	45	0	650	1100	2400	3300	4300	68	0	0
465	1070	58	51	10	45	0	650	1100	2400	3300	4300	68	0	0
470	1060	58	51	10	45	0	650	1100	2400	3300	4300	68	0	0
475	1050	58	52	10	45	0	650	1100	2400	3300	4300	68	0	0
480	1040	57	52	10	45	0	650	1100	2400	3300	4300	68	0	0
485	1030	57	52	10	45	0	650	1100	2400	3300	4300	68	0	0
490	1020	57	52	10	45	0	650	1100	2400	3300	4300	68	0	0
495	1010	56	52	10	45	0	650	1100	2400	3300	4300	68	0	0
500	1000	56	52	10	45	0	650	1100	2400	3300	4300	68	0	0
505	995	56	52	10	45	0	650	1100	2400	3300	4300	68	0	0
510	990	55	52	10	45	0	650	1100	2400	3300	4300	68	0	0
515	985	55	52	10	45	0	650	1100	2400	3300	4300	68	0	0
520	980	55	52	10	45	0	650	1100	2400	3300	4300	68	0	0
525	975	54	52	10	45	0	650	1100	2400	3300	4300	68	0	0
530	970	54	53	10	45	0	650	1100	2400	3300	4300	68	0	0
535	965	53	53	10	45	0	650	1100	2400	3300	4300	68	0	0
540	960	53	53	10	45	0	650	1100	2400	3300	4300	68	0	0
545	955	53	53	10	45	0	650	1100	2400	3300	4300	68	0	0
550	950	52	53	10	45	0	650	1100	2400	3300	4300	68	0	0
555	945	52	53	10	45	0	650	1100	2400	3300	4300	68	0	0
560	940	52	53	10	45	0	650	1100	2400	3300	4300	68	0	0
565	935	51	53	10	45	0	650	1100	2400	3300	4300	68	0	0
570	930	51	53	10	45	0	650	1100	2400	3300	4300	68	0	0
575	925	51	53	10	45	0	650	1100	2400	3300	4300	68	0	0
580	920	50	53	10	45	0	650	1100	2400	3300	4300	68	0	0
585	915	50	53	10	45	0	650	1100	2400	3300	4300	68	0	0
590	910	50	54	10	45	0	650	1100	2400	3300	4300	68	0	0
595	905	45	54	10	45	0	650	1100	2400	3300	4300	68	0	0

600	900	40	54	10	45	0	650	1100	2400	3300	4300	68	0	0
605	900	35	54	10	35	0	644	1100	2400	3300	4300	68	0	0
610	900	25	54	10	30	0	638	1100	2400	3300	4300	68	0	0
615	0	0	54	10	25	0	633	1100	2400	3300	4300	68	0	0
620	0	0	54	10	20	0	627	1100	2400	3300	4300	68	0	0
625	0	0	54	10	15	0	622	1100	2400	3300	4300	68	0	0
630	0	0	54	10	5	0	616	1100	2400	3300	4300	68	0	0
635	0	0	54	10	0	0	611	1100	2400	3300	4300	68	0	0
640	0	0	54	10	0	0	605	1100	2400	3300	4300	68	0	0
645	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
650	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
655	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
660	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
665	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
670	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
675	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
680	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
685	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
690	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
695	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0

A.8 Most Preferred /ǰa/ Synthesis

Synthesis specification for file: 'xj30.wav' Fri Apr 15 14:08:22 1994

KLSYN93 Version 2.0 April 2,1993 N.M.(original program by D.H. Klatt)

Max output signal (overload if greater than 0.0 dB) is -5.9 dB
Total number of waveform samples = 7000

CURRENT CONFIGURATION:
63 parameters

SYM	V/C	MIN	VAL	MAX	DESCRIPTION
DU	C	30	700	5000	Duration of the utterance, in msec
UI	C	1	5	20	Update interval for parameter reset, in msec
SR	C	5000	10000	20000	Output sampling rate, in samples/sec
NF	C	1	5	6	Number of formants in cascade branch
SS	C	1	2	3	Source switch (1=impulse, 2=natural, 3=LF model)
RS	C	1	8	8191	Random seed (initial value of random # generator)
SB	C	0	1	1	Same noise burst, reset RS if AF=AH=0, 0=no,1=yes
CP	C	0	0	1	0=Cascade, 1=Parallel tract excitation by AV
OS	C	0	0	20	Output selector (0=normal,1=voicing source,...)
GV	C	0	60	80	Overall gain scale factor for AV, in dB
GH	C	0	60	80	Overall gain scale factor for AH, in dB
GF	C	0	60	80	Overall gain scale factor for AF, in dB
GI	C	0	60	80	Overall gain scale factor for AI, in dB
F0	V	0	1000	5000	Fundamental frequency, in tenths of a Hz
AV	V	0	60	80	Amplitude of voicing, in dB
OQ	V	10	50	99	Open quotient (voicing open-time/period), in %
SQ	v	100	200	500	Speed quotient (rise/fall time, LF model), in %
TL	V	0	0	41	Extra tilt of voicing spectrum, dB down @ 3 kHz
FL	v	0	0	100	Flutter (random fluct in f0), in % of maximum
DI	v	0	0	100	Diplophonia (alt periods closer), in % of max
AH	V	0	0	80	Amplitude of aspiration, in dB
AF	V	0	0	80	Amplitude of frication, in dB
F1	V	180	500	1300	Frequency of 1st formant, in Hz
B1	v	30	90	1000	Bandwidth of 1st formant, in Hz
DF1	v	0	0	100	Change in F1 during open portion of period, in Hz
DB1	v	0	0	400	Change in B1 during open portion of period, in Hz
F2	V	550	1500	3000	Frequency of 2nd formant, in Hz
B2	v	40	90	1000	Bandwidth of 2nd formant, in Hz
F3	V	1200	2500	4800	Frequency of 3rd formant, in Hz
B3	v	60	150	1000	Bandwidth of 3rd formant, in Hz
F4	V	2400	3250	4990	Frequency of 4th formant, in Hz
B4	v	100	200	1000	Bandwidth of 4th formant, in Hz
F5	V	3000	3700	4990	Frequency of 5th formant, in Hz
B5	v	100	200	1500	Bandwidth of 5th formant, in Hz
F6	v	3000	4990	4990	Frequency of 6th formant, in Hz (applies if NF=6)
B6	v	100	500	4000	Bandwidth of 6th formant, in Hz (applies if NF=6)
FNP	v	180	280	500	Frequency of nasal pole, in Hz
BNP	v	40	90	1000	Bandwidth of nasal pole, in Hz
FNZ	v	180	280	800	Frequency of nasal zero, in Hz
BNZ	v	40	90	1000	Bandwidth of nasal zero, in Hz
FTP	v	300	2150	3000	Frequency of tracheal pole, in Hz
BTP	v	40	180	1000	Bandwidth of tracheal pole, in Hz
FTZ	v	300	2150	3000	Frequency of tracheal zero, in Hz
BTZ	v	40	180	2000	Bandwidth of tracheal zero, in Hz
A2F	v	0	0	80	Amp of fric-excited parallel 2nd formant, in dB
A3F	V	0	60	80	Amp of fric-excited parallel 3rd formant, in dB
A4F	V	0	63	80	Amp of fric-excited parallel 4th formant, in dB
A5F	v	0	0	80	Amp of fric-excited parallel 5th formant, in dB
A6F	v	0	0	80	Amp of fric-excited parallel 6th formant, in dB
AB	v	0	30	80	Amp of fric-excited parallel bypass path, in dB
B2F	v	40	250	1000	Bw of fric-excited parallel 2nd formant, in Hz
B3F	v	60	300	1000	Bw of fric-excited parallel 3rd formant, in Hz
B4F	v	100	320	1000	Bw of fric-excited parallel 4th formant, in Hz

B5F	v	100	360	1500	Bw of fric-excited parallel 5th formant, in Hz
B6F	v	100	1500	4000	Bw of fric-excited parallel 6th formant, in Hz
ANV	v	0	0	80	Amp of voice-excited parallel nasal form., in dB
A1V	v	0	60	80	Amp of voice-excited parallel 1st formant, in dB
A2V	v	0	60	80	Amp of voice-excited parallel 2nd formant, in dB
A3V	v	0	60	80	Amp of voice-excited parallel 3rd formant, in dB
A4V	v	0	60	80	Amp of voice-excited parallel 4th formant, in dB
ATV	v	0	0	80	Amp of voice-excited par tracheal formant, in dB
AI	V	0	0	80	Amp of impulse, in dB
FSF	v	0	0	1	Formant Spacing Filter (1=on, 0=off)

Varied Parameters:

time	F0	AV	OQ	TL	AH	AF	F1	F2	F3	F4	F5	A3F	A4F	AI
0	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
5	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
10	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
15	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
20	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
25	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
30	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
35	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
40	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
45	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
50	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
55	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
60	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
65	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
70	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
75	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
80	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
85	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
90	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
95	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
100	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
105	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
110	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
115	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
120	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
125	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
130	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
135	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
140	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
145	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
150	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
155	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
160	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
165	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
170	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
175	0	0	50	15	0	0	300	1800	2600	3500	4300	58	0	0
180	0	0	50	15	42	40	300	1800	2600	3500	4300	58	0	30
185	0	0	50	15	42	41	309	1787	2590	3485	4300	58	0	0
190	0	0	50	15	42	42	319	1775	2580	3471	4300	58	0	0
195	0	0	50	15	42	44	328	1762	2570	3457	4300	58	0	0
200	0	0	50	15	42	45	338	1750	2560	3442	4300	59	0	0
205	1180	0	50	15	42	46	347	1737	2550	3428	4300	61	0	0
210	1188	38	50	15	42	46	357	1725	2540	3414	4300	63	0	0
215	1197	41	50	14	42	46	366	1712	2530	3400	4300	65	0	0
220	1206	45	50	14	42	46	376	1700	2520	3385	4300	67	0	0
225	1215	48	50	14	42	46	386	1687	2510	3371	4300	68	0	0
230	1224	51	50	13	42	46	395	1675	2500	3357	4300	68	0	0
235	1233	55	50	13	42	46	405	1662	2490	3342	4300	68	0	0
240	1242	57	50	13	43	46	414	1650	2480	3328	4300	68	0	0
245	1250	58	50	12	45	44	424	1637	2470	3314	4300	68	0	0
250	1258	60	50	12	47	42	433	1625	2460	3300	4300	68	0	0
255	1267	60	50	12	49	40	443	1612	2450	3300	4300	68	0	0

260	1275	60	50	12	50	36	453	1600	2440	3300	4300	68	0	0
265	1284	60	50	11	50	32	462	1587	2430	3300	4300	68	0	0
270	1292	60	50	11	50	14	472	1575	2420	3300	4300	68	0	0
275	1300	60	50	11	48	5	481	1562	2410	3300	4300	68	0	0
280	1294	60	50	10	46	0	491	1550	2400	3300	4300	68	0	0
285	1288	60	50	10	45	0	500	1500	2400	3300	4300	68	0	0
290	1300	60	50	10	45	0	518	1450	2400	3300	4300	68	0	0
295	1289	60	50	10	45	0	537	1400	2400	3300	4300	68	0	0
300	1278	60	50	10	45	0	556	1350	2400	3300	4300	68	0	0
305	1267	60	50	10	45	0	575	1300	2400	3300	4300	68	0	0
310	1256	60	50	10	45	0	594	1250	2400	3300	4300	68	0	0
315	1245	60	50	10	45	0	613	1200	2400	3300	4300	68	0	0
320	1235	60	50	10	45	0	632	1150	2400	3300	4300	68	0	0
325	1246	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
330	1257	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
335	1268	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
340	1279	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
345	1289	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
350	1300	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
355	1290	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
360	1280	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
365	1270	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
370	1260	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
375	1250	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
380	1240	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
385	1230	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
390	1220	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
395	1210	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
400	1200	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
405	1190	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
410	1180	60	50	10	45	0	650	1100	2400	3300	4300	68	0	0
415	1170	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
420	1160	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
425	1150	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
430	1140	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
435	1130	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
440	1120	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
445	1110	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
450	1100	60	51	10	45	0	650	1100	2400	3300	4300	68	0	0
455	1090	59	51	10	45	0	650	1100	2400	3300	4300	68	0	0
460	1080	59	51	10	45	0	650	1100	2400	3300	4300	68	0	0
465	1070	58	51	10	45	0	650	1100	2400	3300	4300	68	0	0
470	1060	58	51	10	45	0	650	1100	2400	3300	4300	68	0	0
475	1050	58	52	10	45	0	650	1100	2400	3300	4300	68	0	0
480	1040	57	52	10	45	0	650	1100	2400	3300	4300	68	0	0
485	1030	57	52	10	45	0	650	1100	2400	3300	4300	68	0	0
490	1020	57	52	10	45	0	650	1100	2400	3300	4300	68	0	0
495	1010	56	52	10	45	0	650	1100	2400	3300	4300	68	0	0
500	1000	56	52	10	45	0	650	1100	2400	3300	4300	68	0	0
505	995	56	52	10	45	0	650	1100	2400	3300	4300	68	0	0
510	990	55	52	10	45	0	650	1100	2400	3300	4300	68	0	0
515	985	55	52	10	45	0	650	1100	2400	3300	4300	68	0	0
520	980	55	52	10	45	0	650	1100	2400	3300	4300	68	0	0
525	975	54	52	10	45	0	650	1100	2400	3300	4300	68	0	0
530	970	54	53	10	45	0	650	1100	2400	3300	4300	68	0	0
535	965	53	53	10	45	0	650	1100	2400	3300	4300	68	0	0
540	960	53	53	10	45	0	650	1100	2400	3300	4300	68	0	0
545	955	53	53	10	45	0	650	1100	2400	3300	4300	68	0	0
550	950	52	53	10	45	0	650	1100	2400	3300	4300	68	0	0
555	945	52	53	10	45	0	650	1100	2400	3300	4300	68	0	0
560	940	52	53	10	45	0	650	1100	2400	3300	4300	68	0	0
565	935	51	53	10	45	0	650	1100	2400	3300	4300	68	0	0
570	930	51	53	10	45	0	650	1100	2400	3300	4300	68	0	0
575	925	51	53	10	45	0	650	1100	2400	3300	4300	68	0	0
580	920	50	53	10	45	0	650	1100	2400	3300	4300	68	0	0
585	915	50	53	10	45	0	650	1100	2400	3300	4300	68	0	0
590	910	50	54	10	45	0	650	1100	2400	3300	4300	68	0	0
595	905	45	54	10	45	0	650	1100	2400	3300	4300	68	0	0

600	900	40	54	10	45	0	650	1100	2400	3300	4300	68	0	0
605	900	35	54	10	35	0	644	1100	2400	3300	4300	68	0	0
610	900	25	54	10	30	0	638	1100	2400	3300	4300	68	0	0
615	0	0	54	10	25	0	633	1100	2400	3300	4300	68	0	0
620	0	0	54	10	20	0	627	1100	2400	3300	4300	68	0	0
625	0	0	54	10	15	0	622	1100	2400	3300	4300	68	0	0
630	0	0	54	10	5	0	616	1100	2400	3300	4300	68	0	0
635	0	0	54	10	0	0	611	1100	2400	3300	4300	68	0	0
640	0	0	54	10	0	0	605	1100	2400	3300	4300	68	0	0
645	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
650	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
655	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
660	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
665	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
670	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
675	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
680	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
685	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
690	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0
695	0	0	55	10	0	0	600	1100	2400	3300	4300	68	0	0

Appendix B

MATLAB Script Files

B.1 Analysis Programs

B.1.1 abs_rms.m

```
% this program computes the amplitude of the measured tone
% by computing it's rms value.   The user in prompted for the level
% read off the meter and the speech waveform is scaled appropriately.



figure(1)

if length(mark)~=3   % error check:  user must have only 2 marks
   error('ERROR: you must have only two marks to compute the rms amplitude');
end

mark=sort(mark); % make sure mark(3) > mark(2)

speech_rms_db=20*log10(sqrt(mean(speech(mark(2):mark(3)).^2))) % rms mag.  in dB

reference=input('What was the level measured by the sound meter (in dB)? ');

scale=10^((reference-speech_rms_db)/20)      % linear scaling factor
speech=speech*10^(scale/20);    % multiply each sample in orig.  speech by scale

newmag=speech_rms_db+20.0*log10(scale) % display new scale

clear reference  speech_rms_db scale newmag  % remove unnecessary variables
% plot(speech)
clear ax txt1
```

B.1.2 analysis.m

```
% this m-file sets up the screen for analysis of waveform by
% creating relevant push buttons and sliders, etc.
```

```
% input signal = speech

samples=1:length(speech);           % find length of speech in samples

figure(1)
clg
set(1,'position',[45 650 960 190])   % position window
plot(samples,speech);               % plot all of speech
ax=axis;                       % get current axis

figure(2)
clg
set(2,'position',[45 450 960 145])   % position window 2
plot(samples,speech)                % plot all of speech waveform

figure(3)
clg
set(3,'name', 'hamming windowed fft','position',[585 275 418 175])
                                   % set up window 3 -hamming dft
cb_dft=uicontrol(3,'style','checkbox','position',[.8 .8 .18 .1],'units',...
        'normalized','string','DFT on','callback','dft_active', ...
         'value',1); % set up ability to turn DFT window on or off
dft_on=1;                      % start with dft window on


figure(1)

% create the mark & unmark button, expand button, compress button, window position slider

fram=uicontrol(1,'Style','frame','position',[790 42 165 145]);

markspeech=uicontrol(1,'Style','Pushbutton','Position',[795 155 75 20],...
              'Callback','mark_speech','string','Mark');

centerspeech=uicontrol(1,'Style','Pushbutton','Position',[875 155 75 20],...
              'Callback','clear_mark','string','Clear Mark');

centerspeech=uicontrol(1,'Style','Pushbutton','Position',[835 100 75 20],...
              'Callback','center_window','string','Center');

expand=uicontrol(1,'Style','Pushbutton','Position',[795 130  75 20],...
          'Callback','expand_speech','String','Expand');
compress=uicontrol(1,'Style','Pushbutton','Position',[875 130  75 20],...
          'Callback','compress_speech','String','Compress');
txt1_position=uicontrol(1,'style','text','string', 'Window Position', ...
               'position',[795 80 150 20]);
txt2_position=uicontrol(1,'style','text','string','t=0 end',  'position',[795 67 150 18]);

position=uicontrol(1,'style','slider','position',[795 47  150 20],...
              'callback','window_position','value',0.5);
```

```
clear l1 l2 l3 l4 len
set(1,'WindowButtonDownFcn','ham');

figure(2)                              % initialize window markers
ax=axis;                               % get current axis
left=ax(1);                            % get left edge
right=ax(2);                           % get right edge
zwinleft=line( [left left],[ax(3) ax(4)]);   % draw left vertical line
zwinright=line( [right right],[ax(3) ax(4)]); % draw right vertical line
set(zwinleft,'color','g');             % change color to green
set(zwinright,'color','g');

set(zwinleft,'ButtonDownFcn','window_left_move');
          % call function window_left_move if clicked close to left line
set(zwinright,'ButtonDownFcn','window_right_move');
         % call function window_right_move if clicked close to right line
set(2,'WindowButtonUpFcn','window_move');
         % when the button is released call window_move
if exist('visible')==1     % flag if we want the windows visible
     set(1,'visible','off') % if so, turn window 1 off
     set(2,'visible','off') % alfo turn window 2 off
end

if exist('marktot')==0        % if marked waveform already
 marktot=0;                  % clear marks
 mark=0;                     % start with mark at 0
else
 initialize=1; % next line assumes you just deleted a mark, set flag that says
          % we got there because of initialization
 draw_all_marks;   % draw marks saved in file
 clear initialize;
end

fprintf('\n');

figure(1)
```

B.1.3 analysis3.m

```
% this program sets up the windows used to analyze the speech waveform
% this is designed so the user can chose the zoom window by moving the
% green vertical lines in figure 2. When the desired waveform section is
% displayed, the mouse button shoud be pressed at the start of the transient
% (or vowel) and released at the end of the transient (or vowel) to mark
% the speech for later magnitude and frequency analysis.

% input signal = speech
```

```
samples=1:length(speech);          % find total length of waveform

figure(1)
clg
set(1,'position',[45 650 960 190]) % set up window 1 (zoom view)
plot(samples,speech);              % plot entire speech waveform
ax=axis;                           % get current axis

figure(2)
clg
set(2,'position',[45 450 960 145]) % set up window 2 (normal view)
plot(samples,speech)               % plot entire waveform

figure(3)
clg                                % set up window 3 (dft window)
set(3,'name', 'hamming windowed fft','position',[585 275 418 175])
cb_dft=uicontrol(3,'style','checkbox','position',[.8 .8 .18 .1],'units',...
        'normalized','string','DFT on','callback','dft_active', ...
        'value',1);
      % set up push button to turn dft on or off (speed up program)
dft_on=1; % start with dft window on


figure(1)

% create the large options box on the right of figure 1. This box includes
% buttons to mark or unmark speech, center the window over the mark,
% expand or compress speech, and also creates window position slider

fram=uicontrol(1,'Style','frame','position',[790 42 165 145]);

markspeech=uicontrol(1,'Style','Pushbutton','Position',[795 155 75 20],...
             'Callback','mark_speech','string','Mark');

centerspeech=uicontrol(1,'Style','Pushbutton','Position',[875 155 75 20],...
             'Callback','clear_mark','string','Clear Mark');

centerspeech=uicontrol(1,'Style','Pushbutton','Position',[835 100 75 20],...
             'Callback','center_window','string','Center');

expand=uicontrol(1,'Style','Pushbutton','Position',[795 130  75 20],...
          'Callback','expand_speech','String','Expand');
compress=uicontrol(1,'Style','Pushbutton','Position',[875 130  75 20],...
          'Callback','compress_speech','String','Compress');
txt1_position=uicontrol(1,'style','text','string', 'Window Position', ...
               'position',[795 80 150 20]);
txt2_position=uicontrol(1,'style','text','string','t=0end', 'position',[795 67 150 18]);

position=uicontrol(1,'style','slider','position',[795 47  150 20],...
              'callback','window_position','value',0.5);
```

```
clear l1 l2 l3 l4 len                % clear window markers
%set(1,'WindowButtonDownFcn','ham');
set(1,'WindowButtonDownFcn','mark2_down');
          % call mark2_down.m if button was pressed in window 1
set(1,'WindowButtonUpFcn','mark2_up');
          % call mark2_up.m if button was released  in window 1

figure(2)                                % initialize window markers
ax=axis;                                 % get current axis
left=ax(1);                              % find left edge
right=ax(2);                             % find right edge
zwinleft=line( [left left],[ax(3) ax(4)]);   % draw left vertical window line
zwinright=line( [right right],[ax(3) ax(4)]); % draw right vertical window line
set(zwinleft,'color','g');                  % change color of window lines
set(zwinright,'color','g');                 %         to green

%set(zwinleft,'ButtonDownFcn','window_left_move');
%set(zwinright,'ButtonDownFcn','window_right_move');
%set(2,'WindowButtonUpFcn','window_move');

set(2,'WindowButtonDownFcn','mouse_down')
          % call mouse_down.m if button was pressed in window 2
set(2,'WindowButtonUpFcn','mouse_up');
          % call mouse_up.m if button was released in window 2
if exist('marktot')==0    % if marks do not already exist
 marktot=0;              % clear marks
 mark=0;                 % set first mark at 0
else                   % if marks exist
 initialize=1;     % next line assumes just deleted a mark, set flag that says
                   % we got there because of initialization
 draw_all_marks;   % draw marks saved in file
 clear initialize; % we no longer need to initialize
end

fprintf('\n');

figure(1)
```

B.1.4 avg_spectrum.m

```
% this program averages the 5 spectra for a single word

clear spect_avg1            % make sure we don't have data from previous call

ind=0;
for i=2:4:length(allmark);     % for every other pair of marks
ind=ind+1;
vowel1_db(ind)=20*log10(sqrt(mean(speech(allmark(i+2):allmark(i+3)).^2)));
                               % compute rms level in dB of vowel
end

trans_scale=80-vowel1_db;      % multiply each transient so vowel=80 dB
```

```
trans_scale=[0 trans_scale];  % put in same form as mark_ham_avg


for i=2:marktot+1;              % for each mark
spect_avg1(i,:)=trans_scale(i)+10*log10(mark_ham_avg(i,16:256))+20;
                               % scale all spectra to same vowel height
end


spect_avg1=mean(spect_avg1(2:marktot+1,:));   % compute average of spectra

figure(5)
plot([16:256]/256*5000,spect_avg1,'m');       % plot average of spectra

f=400:100:5000;                               % calculate frequency vector
freq_ln=line(f,20*log10(1./f)+150);           % calcultate 1/f in dB
set(freq_ln,'linestyle','*','MarkerSize',3);  % plot f vs 1/f

freq_ln2=line(f,20*log10(1./f.^2)+215);       % calculate 1/f^2
set(freq_ln2,'linestyle','+','MarkerSize',3); % plot f vs 1/f^2
grid
xlabel('Frequency (Hz)')
ylabel('Dft Magnitude (dB)')

txt=['Avg. Normalized Freq. Content of Transient Bursts for ',all_names(q,:)];
title(txt)
% title('Frequency Content of Transient Bursts')
%print temp1.ps
%!  lpr temp1.ps

clear  trans_scale
```

B.1.5 avg_spectrum_all.m

```
% this file averages the spectra for all 5 males or females

avg_spect_5=mean(spect_avg_all);          % average in dB

figure(5)
plot([16:256]/256*5000,spect_avg_all,'c');  % plot all spectra

f=400:100:5000;                          % create frequency vector
freq_ln=line(f,20*log10(1./f)+150);          % plot f vs 1/f
set(freq_ln,'linestyle','*','MarkerSize',3); % change line to *'s

freq_ln2=line(f,20*log10(1./f.^2)+215);      % plot f vs 1/f^2
set(freq_ln2,'linestyle','+','MarkerSize',3);% change line to +'s
grid
xlabel('Frequency (Hz)')
ylabel('Dft Magnitude (dB)')
title(['Normalized Freq.  Content of Transient Bursts-',word1,'(',sex,')'])
%title('Normalized Freq.  Content of Transient Bursts- got (males)')
```

```
%title('Normalized Freq. Content of Transient Bursts- got (females)')


figure(6)
plot([16:256]/256*5000,avg_spect_5,'c');    % plot average of spectra
f=400:100:5000;                         % create freq vector
freq_ln=line(f,20*log10(1./f)+150);       % plot f vs 1/f
set(freq_ln,'linestyle','*','MarkerSize',3); % change line to *'s
freq_ln2=line(f,20*log10(1./f.^2)+215);       % plot f vs 1/f^2
set(freq_ln2,'linestyle','+','MarkerSize',3);% change line to +'s
grid
xlabel('Frequency (Hz)')
ylabel('Dft Magnitude (dB)')


title(['Avg. Normalized Freq. Content of Transient Bursts-',word1,'(',sex,')'])
%title('Avg. Normalized Freq. Content of Transient Bursts- got (males)')
%title('Avg. Normalized Freq. Content of Transient Bursts- got (females)')


%figure(5)
%print temp1.ps
%! lpr temp1.ps


%figure(6)
%eval(['print /usr/users/massey/matlab/graphics/',word1,'_',sex]); % print to file

%fprintf(['wrote to file :  ',word1,'_',sex,'\n']); % tell user which file

%! lpr temp1.ps
```

B.1.6 dftmarks.m

```
% this function computes the total dft's at all the marks

sp1_ham=zeros(marktot,512);      % pre-allocate space to increase speed

hamm_win=hamming(wl);            % compute hamming window of length wl
for i=2:marktot+1                % repeat for all the marks

% sp1_ham=(abs(fft(hamm_win.*speech((x-wl/2):(x+wl/2-1)),512))); % rect window
sp1_ham(i,:)=(abs(fft(hamm_win.*speech((mark(i)-wl/2):(mark(i)+wl/2-1)),512)))';
        % calculate 512 point dft of hamming-windowed speech
sp1_ham_avg(i,:)=zeros(1,512);  % pre-allocate for speed

    for m=11:502;          % for each spectra point
     sp1_ham_avg(i,m)=mean(sp1_ham(i,(m-5):(m+5)));
          % smooth spectra by averaging over 10 points
    end
    clear m
```

```
end %for

figure(4)
plot([1:256]/256*5000,20*log10(sp1_ham(:,1:256)),'r',[1:256]/256*5000,
20*log10(sp1_ham_avg(:,1:256))+20)
  % plot spectra and smoothed spectra
grid
figure(1)

fprintf('\n');

% end  %while

clear hamm_win  len
```

B.1.7 draw_all_marks.m

```
% this file draws the marks already in the variable 'mark'

col=['m' 'c' 'b' 'w']; % circle through remaining colors (not red, green, blk)



 if (exist('markline') ~= 0) & (exist('initialize') ~=1)
            % if markline exists and we are not initializing
        figure(1)
        delete(markline(1,marktot+2));  % delete line in window 1
           figure(2)
        delete(markline(2,marktot+2));  % delete line in window 2
  end % if

if (exist('initialize') ==1)  % draw marks on initialization
for m=1:2;                    % for both windows

figure(m)
ax=axis;                      % get current axis
for i=2:marktot+1;            % for each of the marks

markline(m,i)=line([mark(i) mark(i)],[ax(3) ax(4)]);       % draw vertical line
set(markline(m,i),'color',col(rem(i,4)+1),'lineStyle','--');% in alt. colors
end % for i

end % for m
end % if

figure(1)


clear ax i col
```

B.1.8 imp_syn.m

```
% this program creates an impulse surrounded by zeros
clear y ylog
for i=2:10;                  % try different lenth waveforms
i                            % print the current length
[b,a]=invfreqz(finv,w,i,0)   % compute filter length i that has spect=1/f

filtout=filter(b,a,[0 0 0 0 0 1 zeros(1,50)]); % filter impulse
y(i,:)=fft(filtout,128);     % calculate fft of result
ylog(i,:)=20*log10(abs(y(i,1:64))); % convert to dB
end
```

B.1.9 mark_dft.m

```
% this program computes the dft's at each of the marks



len=length(speech);            % find the length of speech waveform

% wl=input('enter the window length (in samples) > ');
wl=64;                         % length of hamming window
fprintf('Mark_dft window length = %3.0f points \n',wl);
                               % print window length to screen
hamm_win=hamming(wl);          % create hamming window of length wl
curfig=1;                      % keep track of windows
hamwin=3;


if exist('l1')~=0              % if vertical window lines exist
 if (l1 >0)
  delete(l1)                   % delete them
  delete(l2)
  delete(l3)
  delete(l4)
clear l1 l2 l3 l4
 end
end


if exist('mark_dft_win1')~=0   % if dft window exists
 for mm=1:2;                   % for each of the windows
  for nn=2:marktot+1           % for each mark
    delete(mark_dft_win1(mm,nn))% delete hamming windows
    delete(mark_dft_win2(mm,nn))
  end
 end % mm
end % if
```

```
for m=1:2                         % do both windows
figure(m)
ax=axis;                          % get current axis
mx=max(speech(ax(1)+1:min(ax(2),length(speech)))); % find max amp of speech
mn=min(speech(ax(1)+1:min(ax(2),length(speech)))); % find min amp of speech
mx=min(mn,mx); % take smallest (symmetrical window could be too large)

for i=2:marktot+1
mark_dft_win1(m,i)=line([mark(i)-wl/2-1 :mark(i)+wl/2-1+1],[0 mx*hamm_win' 0]);
        % draw top half of hamming window
mark_dft_win2(m,i)=line([mark(i)-wl/2-1 :mark(i)+wl/2-1+1],[0 -mx*hamm_win' 0]);
        % draw bottom half of hamming window
set(mark_dft_win1(m,i),'color','r');    % change color of hamming window to red
set(mark_dft_win2(m,i),'color','r');

end %i

end %m


mark_ham_avg(marktot+1,:)=zeros(1,512); % pre-allocate space for average results

for i=2:marktot+1
mark_dft1(i,:)=(abs(fft(hamm_win.*speech((mark(i)-wl/2):(mark(i)+wl/2-1)),512)))'
       % compute the 512 point dft of the windowed speech
for m=1+15:512-15; % for each of the points in the spectra
  mark_ham_avg(i,m)=mean(mark_dft1(i,(m-15):(m+15)).^2);
             % smooth with linear average over 30 points
end
clear m
end %for i


figure(4)

col=['m' 'c' 'b' 'w']; % circle through remaining colors (not red, green, blk)

for i=2:marktot+1;    % for each of the marks
plot([16:256]/256*5000,10*log10(mark_ham_avg(i,16:256))+20,col(rem(i,4)+1));
                   % plot spectra in different colors
hold on
end

hold off
f=100:100:5000;                            % create frequency axis
freq_ln8=line(f,20*log10(1./f)+150);          % plot 1/f
set(freq_ln8,'linestyle','*','MarkerSize',3); % change style of line to *'s

grid
xlabel('Frequency (Hz)')
ylabel('Dft Magnitude (dB)')
```

```
txt=['Frequency Content of Transient Bursts for ',all_names(q,:)];
title(txt)
%title('Frequency Content of Transient Bursts')

%print temp0.ps
%!lpr temp0.ps

figure(1)

fprintf('\n');

clear hamm_win  len f
```

B.1.10 model1.m

```
% this program computes the expected volume velocity
% using a rectangular constriction with the following parameters

rate=100;               % area opening in cm^2/s

msec=[.01:.01:2];        % time from 0.01 to 2 msec
temp1=rate*10*msec/1000;  % find one dimension of opening
temp2=0.1;               % find other dimension of opening
for k=1:length(msec);     % for each point
d(k)=min([temp1(k) temp2]);% d is smallest dimension
b(k)=max([temp1(k) temp2]);% b is largest dimension
end

a=3;                    % area of large tube in cm^2

ac=b.*d;                 % cross-sectional area of opening (len*width)
rho=.00114;               % density of air
mu=1.94e-4;               % viscosity of air
c=35400;                 % speed of sound in cm/sec
lc=0.1;                 % length of the constriction
press=8*980;              % pressure of 8 cm H2O

u1= (-(rho*c/a +12*mu*lc./b./(d.^3)) + sqrt( (rho*c/a+12*mu*lc./b./(d.^3)).^2+4*
   press*(rho/2./(ac.^2)))))./ (rho./(ac.^2));
      % compute one root of volume vel. using quadratic eq
u2= (-(rho*c/a +12*mu*lc./b./(d.^3)) - sqrt( (rho*c/a +12*mu*lc./b./(d.^3)).^2+4*
   press*(rho/2./(ac.^2)))))./ (rho./(ac.^2));
      % compute other root of volume vel. using quadratic eq
uapprox=rate.*msec/1000*sqrt(2*8*980/.00114); % calcultate using approximation

coeffs=[(rho/2./(ac.^2))' (rho*c/a+ 12*mu*lc./b./(d.^3))' (-press)*ones(1,length(ac))'];
         % calculate root using MATLAB's root command (needs coefs of polynomial)

for k=1:length(ac)
root1(k,:)=roots(coeffs(k,:))';  % calculate roots = volume velocity at each point
end
```

```
figure(1)
% hold on
set(1,'position',[50 525 450 300])      % reduce window size
plot(msec,[root1(:,2) uapprox'])      % plot vol. vel. and approx
grid
title('Volume velocity and approximation-Rectangular Constriction')
xlabel('Time (msec)')
ylabel('Volume velocity (cm^3/sec)')

figure(2)
% hold on
set(2,'position',[535 525 450 300])
p1=rho*c/a*root1(:,2);                 % calculate press due to char Z term
p2=12*mu*lc./b./(d.^3).*root1(:,2)';   % pressure due to viscosity
p3=rho/2./(ac.^2).*(root1(:,2).^2)';  % pressure due to kinetic resistance
p4=p1'+p2+p3;                          % total pressure
plot(msec,[p1 p2' p3' p4' ])           % plot all 4
grid
xlabel('Time (msec)')
title('Pressure components-Rectangular Constriction')
ylabel('Pressure (dynes/cm^2)')



% find pressure derivative term

du=1e5*diff(root1(:,2));               % find derivative of real vol. vel.
duapprox=1e5*diff(uapprox);            % find derivative of vol. vel. approx

r=20;                                  % distance from lips in cm
press=rho/4/pi/r*du;                   % presure component due to derivative
press_approx=rho/4/pi/r*duapprox(1:199); % press due to derivative of approx

figure(3)
% hold on
plot(msec(2:length(msec)),20*log10([press press_approx']/.0002))
                                       % plot pressure and approx
grid
xlabel('Time (msec)')
ylabel('Pressure (db)')
title('Pressure at 20 cm  and approximation—Rectangular Constriction')


Y=20*log10(abs(fft(hamming(100).*press(1:100),1024))); % calculate fft of press
Y2=20*log10(abs(fft(hamming(100)'.*press_approx(1:100),1024)));
                            % calculate fft of pressure approximation
freq=[1:1024]/1024*10000;              % create frequency scale
invfreq=20*log10(1./freq)+55;          % calculate 1/f curv

figure(4)
% hold on
plot(freq(1:512),Y(1:512),freq(1:512),Y2(1:512),freq(1:512),invfreq(1:512),'.')
                  % plot 1/f, fft of pressure, fft of pressure approxi
grid
```

```
xlabel('Freq (Hz)')
ylabel('Magnitude')
title('FFT of Pressure -Rectangular Constriction')


figure(5)
% calculate d/dt term
d_by_dt=[1e5*diff(rho*lc*root1(:,2)./ac')' 0]; % caclulate derivative component
ddddd=[(rho*lc*root1(:,2)./ac')' ];          % find coefficient
plot(msec,ddddd)                            % plot coefficient
title('rho*lc*U/ac')
xlabel('Time (msec)')
grid

figure(6)
plot(msec,d_by_dt)                          % plot derivative of press component
grid
title('derivative of (rho*lc*U/ac)')
xlabel('Time (msec)')

figure(1)

% following loop used to print graphs to file

%for mm=1:6
%figure(mm)
%grid on
%fignum=num2str(mm);
%eval(['print /usr/users/massey/matlab/graphics/ch2/rectmodel',fignum]);
%       %print temp0.ps
%       %!lpr temp0.ps
%end
```

B.1.11 model2.m

```
% this program computes the expected volume velocity
% using a circular constriction with the following parameters

msec=[.01:.01:2];  % time from .01 to 2 ms
rate=100;          % rate of constriction opening in cm^2/sec
ac=rate/1000*msec; % area of constriction opening

d=2*sqrt(ac/pi);   % diameter of circular opening

a=3;               % cross-sectional area of large tube
rho=.00114;          % rho=density
mu=1.94e-4;          % mu=viscosity
c=35400;           % speed of sound
lc=0.1;            % length of constriction
press=8*980;         % pressure = 9 cm H2O
```

```
uapprox=rate.*msec/1000*sqrt(2*8*980/.00114); % approximate vol. velocity

coeffs=[(rho/2./(ac.^2))' ( rho*c/a+ 128*mu/pi./(d.^4))' (-press)*ones(1,length(ac))'];
     % coeffs of volume velocity polynomial (neglecting derivative term)

for k=1:length(ac)                  % calculate for each area
root1(k,:)=roots(coeffs(k,:))';   % find solutions to poly = volume vel.
end

% hold on

set(1,'position',[50 525 450 300]) % reduce size of window
plot(msec,[root1(:,2) uapprox'])    % plot vol.  vel.  and approx
grid on
title('Volume Velocity and Approximation - Circular Constriction')
xlabel('Time (msec)')
ylabel('Volume velocity')

figure(2)
% hold on
set(2,'position',[535 525 450 300])
p1=rho*c/a*root1(:,2);                  % calculate press due to char Z term
p2=128*mu/pi./(d.^4).*root1(:,2)'; % pressure due to viscosity
p3=rho/2./(ac.^2).*(root1(:,2).^2)';% pressure due to kinetic resistance
p4=p1'+p2+p3;                % total pressure
plot(msec,[p1 p2' p3' p4'  ])        % plot all 4

% hold on
% plot(msec,  rho/2./(ac.^2).*uapprox.^2 ,'x')
% hold off
grid on
xlabel('Time (msec)')
title('Pressure components - Circular Constriction')


% find pressure addition of derivative term

du=1e5*diff(root1(:,2));         % calculate real derivative from above
duapprox=1e5*diff(uapprox);      % calculate derivative if using approx.

r=20;                            % distance from lips in cm
press=rho/4/pi/r*du;             % pressure due to derivative term
press_approx=rho/4/pi/r*duapprox;% pressure if using approximation

figure(3)
% hold on
plot(msec(2:length(msec)),20*log10([press press_approx']/.0002))
                    % plot real and approx. pressure components
grid on
xlabel('Time (msec)')
ylabel('Pressure (db)')
title('Pressure at 20 cm  and approximation-Circular Constriction')
```

```
Y=20*log10(abs(fft(hamming(100).*press(1:100),1024))); % calculate fft of p
Y2=20*log10(abs(fft(hamming(100)'.*press_approx(1:100),1024)));
                                                      % calc fft of p approx
freq=[1:1024]/1024*10000;
invfreq=20*log10(1./freq)+55;                         % calculate 1/f

figure(4)
% hold on
plot(freq(1:512),Y(1:512),freq(1:512),Y2(1:512),freq(1:512),invfreq(1:512),'.')
                  % plot 1/f, fft of approx press, and fft of actual pressure

grid on
xlabel('Freq (Hz)')
ylabel('Magnitude')
title('FFT of Pressure -Circular Constriction')

figure(1)


%for mm=1:4
%figure(mm)
%grid on
%fignum=num2str(mm);
%eval(['print /usr/users/massey/matlab/graphics/ch2/circmodel',fignum]);
%       %print temp0.ps
%       %!lpr temp0.ps
%end
%

clear mm
```

B.1.12 scale_speech_noel.m

```
% this file scales the speeech so the rms values match the absolute
% rms values measured with a sound meter in a soundproof room


% person = noel
scale= 41.7339
clear

scale= 41.7339
load ./speech/noel/noel88˙5.txt
noel88˙5=noel88˙5*scale;
clear scale
save ./speech/noel/noel88˙5
clear
```

```
scale= 41.7339
load ./speech/noel/noelbeat.txt
noelbeat=noelbeat*scale;
clear scale
save ./speech/noel/noelbeat
clear

scale= 41.7339
load ./speech/noel/noelbought.txt
noelbought=noelbought*scale;
clear scale
save ./speech/noel/noelbought
clear

scale= 41.7339
load ./speech/noel/noelchalk.txt
noelchalk=noelchalk*scale;
clear scale
save ./speech/noel/noelchalk
clear

scale= 41.7339
load ./speech/noel/noelcheap.txt
noelcheap=noelcheap*scale;
clear scale
save ./speech/noel/noelcheap
clear

scale= 41.7339
load ./speech/noel/noelchurch.txt
noelchurch=noelchurch*scale;
clear scale
save ./speech/noel/noelchurch
clear

scale= 41.7339
load ./speech/noel/noelcot.txt
noelcot=noelcot*scale;
clear scale
save ./speech/noel/noelcot
clear

scale= 41.7339
load ./speech/noel/noeldeep.txt
noeldeep=noeldeep*scale;
clear scale
save ./speech/noel/noeldeep
clear

scale= 41.7339
load ./speech/noel/noeldock.txt
noeldock=noeldock*scale;
clear scale
save ./speech/noel/noeldock
```

```
clear

scale= 41.7339
load ./speech/noel/noelgeek.txt
noelgeek=noelgeek*scale;
clear scale
save ./speech/noel/noelgeek
clear

scale= 41.7339
load ./speech/noel/noelgot.txt
noelgot=noelgot*scale;
clear scale
save ./speech/noel/noelgot
clear

scale= 41.7339
load ./speech/noel/noeljeep.txt
noeljeep=noeljeep*scale;
clear scale
save ./speech/noel/noeljeep
clear

scale= 41.7339
load ./speech/noel/noeljob.txt
noeljob=noeljob*scale;
clear scale
save ./speech/noel/noeljob
clear

scale= 41.7339
load ./speech/noel/noeljudge.txt
noeljudge=noeljudge*scale;
clear scale
save ./speech/noel/noeljudge
clear

scale= 41.7339
load ./speech/noel/noelkeep.txt
noelkeep=noelkeep*scale;
clear scale
save ./speech/noel/noelkeep
clear


scale= 41.7339
load ./speech/noel/noelpeak.txt
noelpeak=noelpeak*scale;
clear scale
save ./speech/noel/noelpeak
clear

scale= 41.7339
load ./speech/noel/noelpot.txt
```

```
noelpot=noelpot*scale;
clear scale
save ./speech/noel/noelpot
clear

scale= 41.7339
load ./speech/noel/noelteach.txt
noelteach=noelteach*scale;
clear scale
save ./speech/noel/noelteach
clear

scale= 41.7339
load ./speech/noel/noeltop.txt
noeltop=noeltop*scale;
clear scale
save ./speech/noel/noeltop
clear
```

B.1.13 segment.m

```
% this file allows the user to mark segments of speech for later analysis

samples=1:length(speech);        % find length of speech
figure(1)
set(1,'position',[45 650 960 190])% set window 1 to be long and narrow
plot(samples,speech);            % plot speech waveform

clear impulse;

numtot=input('Enter number of segments to extract > ');
        % number of interesting segments in speech
   figure('name','Zoom window','position',[180 525 822 145].'visible'.'off'):
          % prepare window but don't make it visible yet
     curfig=gcf;

for i=1:numtot;                  % for each segment
          figure(1);
         [samp,temp]=ginput(2); % define a zoom window
             figure(curfig)
             set(curfig,'visible','on'); % turn window on
             plot(samples(samp(1):samp(2)),speech(samp(1):samp(2))); % zoom in
             [samp,temp]=ginput(2);      % get next mouse click
             impulse=[impulse samp];     % append speech segment to previous segs.
             clg
             set(curfig,'visible','off'); % turn off window

        figure(1)
end;
figure(curfig)
```

```
close                                   % close zoom window
figure(1)
clear temp samp i curfig
```

B.1.14 spect.m

```
% this program computes the average spectrum

spectavg=zeros(5,256);                  % pre-allocate for increased speed
numtot=length(impulse);                  % find length of impulse

for i=1:numtot;                         % do this for each mark
  len(i)=ceil(impulse(2,i))-floor(impulse(1,i)); % find number of samples in marked region
end;
maxlength=max(len);                     % find largest region

 seg=zeros(numtot,maxlength+1);         % pre-allocate for increased speed

for i=1:numtot;                         % for each mark
seg1(i,:)=speech(floor(impulse(1,i)):ceil(impulse(2,i)))'  ; % get spech segment
ham1(i,:)=hamming(length(seg1(i,:)))';          % compute hamming window
seg(i,:)=[ ham1(i,:).*seg1(i,:)  zeros(1,maxlength-len(i))];
                                        % zero pad speech
clear ham1 seg1


spect1(i,:)=abs(fft(seg(i,:),256));            % compute fft

     for mm=11:246;                     % for each point in fft
          spectavg(i,mm)=mean(spect1(i,(mm-10):(mm+10))); % smooth spectrum
     end
     clear mm


spectr(i,:)=20*log10(abs(fft(seg(i,:),256)));     % convert to dB
% spectavg(i,:)=spectavg(i,:)/max(spectavg(i,:)/1000); %normalize averaging

end;


for i=1:256;                            % for each freq. point
avgspec(i)=mean(spectr(:,i));             % compute average of spectra
end;

freq=10000/256*[1:256];                  % convert to Hz
if numtot<5                             % if less than 5 spectra
   spectr(5,255)=0;                     % fill with 0's
% needed make sure matrix is 5 rows so next plot statement will work
end


finv=20*log10(1./[1:256]);                      % compute 1/f
```

```
figure
plot(freq,spectr(1,:),'-.',freq,spectr(2,:),freq,spectr(3,:),':',
freq,spectr(4,:),'o',freq,spectr(5,:),'x',freq(1:128),finv(1:128)+spectr(1,25)+50,
 'wo');
 % plot 5 spectra and 1/f
hold on
 plot(freq,20*log10(spectavg)+20)                % also plot average 20
hold off
grid
xlabel('Frequency (Hz)')
ylabel('Magnitude (dB)')

t1=gcf;
get(t1,'CurrentCharacter');

while 1 ,  % repeat until key pressed
fprintf('Click where you want all magnitudes to be equal (any key to quit)\n')
  rt=ginput(1); % find the location to normalize to
 [temp,pt] =min(abs(rt(1)-freq)); % convert x-coord to frequency
if  (get(t1,'CurrentCharacter')) ~= [],break,end  % exit while loop

%plot(freq(pt),finv(pt)+spectr(1,25)+50,'rx')


spectavg=spectavg./(spectavg(:,pt)*ones(1,256)); % rescale spectra
fprintf('in loop\n')
plot(freq,20*log10(spectavg)+finv(pt)+spectr(1,25)+50,freq(1:128),finv(1:128)+
   spectr(1,25)+50,'wo');
    % and plot spectra and 1/f

grid
xlabel('Frequency (Hz)')
ylabel('Magnitude (dB)')
title('Normalized frequency responses')

end  % while         .

clear numtot len spect1   maxlength i seg avgspec

 clear  spectavg spectr temp pt rt t1
```

B.1.15 spect_from_mag.m

```
% this file marks the transient bursts by centering a window around
% the maximum amplitude of the transient (search area delimited already
% for use in transient'mag.m

% all'names=['/usr/users/massey/synth/xp40_mag'];
%   all_names=['./dave/davekeep_mag'];
```

```
word1='job'          % choose word to be analyzed
sex='females'        % choose males or females to analyze
visible='no'         % calculate without plotting spectra every time



if sex(1:4)=='male'                    % if we hose to analyze males
    all_names=zeros(5,15+length(word1)); % set length of file name
    names=['./dave/dave'; './john/john';
              './ken/ken__';'./mark/mark'; './noel/noel'];
        % file names to load (assumes we are in directory ~/matlab/speech
else      % if analyzing females
    all_names=zeros(5,17+length(word1)); % set file length
    names=['./helen/helen'; './jane/jane__';...
           './jenny/jenny'; './kelly/kelly';'./lorin/lorin'];
        % file names to load (assumes we are in directory ~/matlab/speech
end %if

for nn=1:5 % for each of the five files
all_names(nn,:)=[ names(nn,:) word1 '_mag']; % create names of files to load
end

all_names  % show user names of files used


% the following is the old way of loading files
%all_names=['./dave/davekeep_mag    '; './john/johnkeep_mag    ';...
%        './ken/kenkeep_mag      ';'./mark/markkeep_mag    ';...
%        './noel/noelkeep_mag    ']
%;...


%all_names=['./helen/helenkeep_mag '; './jane/janekeep_mag    ';...
%        './jenny/jennykeep_mag '; './kelly/kellykeep_mag ';...
%        './lorin/lorinkeep_mag ']



%all_names=['./dave/davejeep_mag   '; './john/johnjeep_mag   ';'
          ./ken/kenjeep_mag     ';...
%        './mark/markjeep_mag   ';'./noel/noeljeep_mag   ' ;...
%
%       './helen/helenjeep_mag'; './jane/janejeep_mag   ';...
%        './jenny/jennyjeep_mag'; './kelly/kellyjeep_mag';...
%       './lorin/lorinjeep_mag'];



for q=1:size(all_names,1); % calculate each person
eval(['load ',all_names(q,:)]);


mark=sort(mark);            % sort in ascending order
```

```
allmark=mark;              % save marks for later use

ind=0;
  for i=2:4:marktot+1;       % for each transient/vowel pair (4 marks)
    ind=ind+1;
    [temp, index]=max(abs(speech(mark(i):mark(i+1)))); % find max of transient
    mark(ind)=index+mark(i); % replace bin marks with markers at max amp of transient
  end

marktemp=mark(1:ind);       % temporarily store max transient magnitude marks
clear mark                 % clear old marks
mark=[0 marktemp];           % replace with new transient mag marks
marktot=ind;                % store total number of transients marked

analysis                   % call anlysis program
mark_dft                   % compute the dft at all marks
avg_spectrum                 % average spectra for this word (each speaker)
spect_avg_all(q,:)=spect_avg1(1,:); % also average all speakers


clear i ind temp marktemp    % clear variables for next time through loop

clear mark_dft_win1
end % q  for all names
```

B.1.16 transient_mag.m

```
% after placing marks on either side of the transients and also
% placing marks somewhere in the next vowel, the user can
% find the difference between the max and min amplitude of the transients
% relative to the following vowel

                              .
mark=sort(mark); % sort in ascending order

              % remember mark(1)=0 and is useless to us


word='job';      % choose current word to analyze

%all_names=['./synth/xch30_mag'];

% analyze following files

all_names=['./dave/davejob_mag    '; './john/johnjob_mag    ';...
        './ken/kenjob_mag      ';'./mark/markjob_mag     ';...
        './noel/noeljob_mag    ';...
        './helen/helenjob_mag  '; './jane/janejob_mag    ';...
        './jenny/jennyjob_mag  '; './kelly/kellyjob_mag  ';...
       './lorin/lorinjob_mag  '];
```

```
%  './dave/daveteach_mag  '; './john/johnteach_mag  ';'./ken/kenteach_mag    ';...
%       './mark/markteach_mag  ';'./noel/noelteach_mag  ';...
%  './helen/helenteach_mag'; './jane/janeteach_mag  ';...
%       './jenny/jennyteach_mag'; './kelly/kellyteach_mag';...
%       './lorin/lorinteach_mag'];

for q=1:size(all_names,1);        % calculate each person
eval(['load ',all_names(q,:)]);   % load new file
mark=sort(mark);                  % make sure marks are in ascending order

if (rem(marktot,4)) ~= 0          % make sure there is a multiple of 4 # of marks
  error('ERROR: number of marks must be mult of 4 (2/ea trans.+2/ea vowel)');
end

ind=0;
for i=2:4:marktot+1;              % do for each transient/vowel pair of marks
ind=ind+1;
trans_amp(ind)=max(abs(speech(mark(i):mark(i+1))));   % find max of transient
vowel_amp(ind)=max(abs(speech(mark(i+2):mark(i+3)))); % find max of vowel
vowel1_db(ind)=20*log10(sqrt(mean(speech(mark(i+2):mark(i+3)).^2)));
                    % compute rms level in dB of vowel

end

relative_amp_db(q,:)= 20*log10(trans_amp./vowel_amp); % find trans. relative to vowel
vowel_db(q,:)=vowel1_db;                     % save rms vowel dB for later user
vowel_peak_db(q,:)=20*log10(vowel_amp);          % find vowel peak (dB)
imp_db(q,:)=vowel1_db+relative_amp_db(q,:);      % find trans peak (dB)



end % for each person

avg_vowel=mean(vowel_db(:))               % compute vowel average for all people
std_vowel=std(vowel_db(:))               % compute v standard dev. for all people
avg_vowel_peak=mean(vowel_peak_db(:))   % compute vowel average for all people
std_vowel_peak=std(vowel_peak_db(:))    % compute v standard dev. for all people
avg_transient=mean(imp_db(:))            % compute trans average for all people
std_transient=std(imp_db(:))            % compute trans standard dev. for all people
avg_relative=mean(relative_amp_db(:))   % compute transient average for all people
std_relative=std(relative_amp_db(:))    % compute trans standard dev. for all people


figure(4)
clg
plot([1:5*size(all_names,1)],imp_db(:),'g',[1:5*size(all_names,1)],vowel_db(:),'b')
                 % plot vowel and transient magnitudes for all people
grid
set(4,'position',[12 565 312 250]);

xlabel('Utterance Number')
ylabel('Magnitude (db)')
```

```
title(['Vowel and Transient Magnitudes - ',word])
[a,b]=max(vowel_db(:));
text(b,a,'Following Vowel')
%[a,b]=min(imp_db(1:20)); % only put mark on first portion
[a,b]=min(imp_db(:));
text(b,a,'Transient')

figure(5)	% histogram of transient magnitude
set(5,'position',[348 565 312 250]);
hist(imp_db(:))
xlabel('Magnitude (dB)')
ylabel('Number of Occurences')
title(['Transient Magnitudes - ',word])
text('string',['Mean= ',num2str(avg_transient)],'position',[.05 .9],
'units','normalized');
text('string',['St. Dev.= ',num2str(std_transient)],'position',[.05 .85],
 'units','normalized');


figure(6)	% histogram of vowel magnitude
set(6,'position',[685 565 312 250]);
hist(vowel_db(:))
xlabel('Magnitude (dB)')
ylabel('Number of Occurences')
title(['Vowel RMS Magnitudes - ',word])
text('string',['Mean= ',num2str(avg_vowel)],'position',[.05 .9],
'units','normalized');
text('string',['St. Dev.= ',num2str(std_vowel)],'position',[.05 .85],
  'units','normalized');

figure(7)	% histogram of transient magnitude relative to following vowel
set(7,'position',[348 250 312 250]);
hist(relative_amp_db(:))
xlabel('Magnitude (dB)')
ylabel('Number of Occurences')
title(['Relative Transient Magnitude - ',word])
text('string',['Mean= ',num2str(avg_relative)],'position',[.05 .9],
 'units','normalized');
text('string',['St. Dev.= ',num2str(std_relative)],'position',[.05 .85],
 'units','normalized');


figure(8)	% histogram of vowel magnitudes
set(8,'position',[685 250 312 250]);
hist(vowel_peak_db(:))
xlabel('Magnitude (dB)')
ylabel('Number of Occurences')
title(['Vowel Peak Magnitudes - ',word])
text('string',['Mean= ',num2str(avg_vowel_peak)],'position',[.05 .9],'units','normalized');
text('string',['St. Dev.= ',num2str(std_vowel_peak)],'position',[.05 .85],
 'units','normalized');
```

```
%for mm=4:8
%figure(mm)
%fignum=num2str(mm-3);
%eval(['print /usr/users/massey/matlab/graphics/ch3/',word,fignum]);
     %print temp0.ps
     %!lpr temp0.ps
%end


clear b a
```

B.2 MATLAB Window Management Programs

B.2.1 center_window.m

```
% this file centers the zoom window on the cursor (last clicked point) in ham.m

ax=axis;                                    % get current axis coordinates
len=ax(2)-ax(1);                            % find length of x-axis
left=max(1,lastclick-len/2);                % find left window edge ( >=1 )
right= min(length(speech),lastclick+len/2); % find rt edge (<= length of waveform)
top=max(speech(left:right));                % find max amp of speech in window
bottom=min(speech(left:right));             % find min amp of speech in window

axis([left right  bottom top])      % apply new window axis

window2_update;                             % call m-file to update window 2

clear ax len top bottom left right
```

B.2.2 clear_mark.m

```
% this file clears the last mark

len=length(mark);            % find out how many marks were made
mark=mark(1:(len-1));        % save only valid marks
marktot=marktot-1;           % last mark doesn't count, #=length-1
draw_all_marks;                 % call another m-file to draw marks

clear len
```

B.2.3 compress_speech.m

```
% this file is called by a push button
% it zooms out on the center of the window (i.e. compresses more speech
```

```
% into the same window size)


ax=axis;                                   % get current axis
len=ax(2)-ax(1);                           % calculate length of x-axis
left=max(1,ax(1)-len);             % lower left x-axis coord (min of 1)
right= min(length(speech),ax(2)+len);  % increase right x-axis coord (max=length of speech)
top=max(speech(left:right));               % calculate max amp of speech in window
bottom=min(speech(left:right));      % calculate min amp of speech in window

axis([left right  bottom top])             % draw new window


window2'update;                            % call m-file to update window 2

clear ax len top bottom left right
```

B.2.4 dft_active.m

```
% this function toggles between turning the dft window off or on.
% turning it off speeds up the analysis program

if dft'on ==1   % if the DFT button was pushed
  dft'on=0;    % turn it off now
else          % if the DFT button wasn't pushed
  dft_on=1;      % turn it on now
end
```

B.2.5 expand_speech.m

```
% this file is called by a push button
% it zooms in on the center of the window (expands the waveform)

ax=axis;                              % get current axis
len=ax(2)-ax(1);                      % calculate length of x-axis
axis([ax(1)+len/4   ax(2)-len/4   ax(3)   ax(4)]);
                                      % increase left x-axis coord and
                                      %   decrease right x-axis coord
left=ax(1)+len/4;                     % calculate new left axis
right= ax(2)-len/4;                   % calculate new right axis
top=max(speech(left:right));          % find max amp of speech in window
bottom=min(speech(left:right));    % find min amp of speech in window

axis([left right  bottom top])        % set new window axis coordinates

window2'update;                       % call m-file to update window 2



clear ax left right top bottom len
```

B.2.6 ham.m

```
Fileham.m,00:50,May  5 1994
% this file computes the spectrum of a waveform


len=length(speech);                          % find length of current waveform

% dft window length
wl=128;                                      % DFT window length (in samples)
   if dft_on==1                    % if DFT display is selected
      fprintf('Window length = %3.0f points \n',wl); % print window length
     else                        % if DFT display is not selected
      fprintf('DFT turned off \n');          % tell user that no DFT was calculated
   end

figure(1)
[x,y]=get(gca,'currentPoint');         % get current cursor coordinates
x=x(1,1);                                    % keep only the x-coord.
lastclick=x;                                 % store x-coord for later use

hamm_win=hamming(wl);                        % compute hamming window length=wl
ax=axis;                                     % get current window axis coords
mx=max(speech(ax(1)+1:min(ax(2),length(speech)))); % find max speech amp in window


if (exist('l1') ~= 0)                        % if old hamming window exists
  delete(l1)                     % delete top half of red window
  delete(l2)                     % delete bottom half of red window
end

l1=line ( [ x-wl/2-1 :x+wl/2-1+1],[0 mx*hamm_win' 0]); % draw half of ham wind.
l2=line ( [x-wl/2-1 :x+wl/2-1+1],[0 -mx*hamm_win' 0]); % draw bottom half also
set(l1,'color','r');                         % draw hamming window in red
set(l2,'color','r');               % draw other half of hamming window in red

figure(2)

ax=axis;                                      % get current axis in figure 2
mx=max(speech(ax(1)+1:min(ax(2),length(speech)))); % find max amp of speech in window
mn=min(speech(ax(1)+1:min(ax(2),length(speech)))); % find min amp of speech in window
mx=min(mn,mx); % symmetrical window could waste space (too large)

if (exist('l3') ~= 0)              % if hamming window already exists
  delete(l3)                                 % delete old top half of ham window
  delete(l4)                                 % delete old bottom half of ham window
end

l3=line ( [x-wl/2:x+wl/2-1],mx*hamm_win);     % draw new top half of hamming window
l4=line ( [x-wl/2:x+wl/2-1],-mx*hamm_win);     % draw new bottom half of hamming window
```

```
set(l3,'color','r');                                % draw top half in red
set(l4,'color','r');                                % draw bottom half in red


if dft_on==1                                        % if DFT window is selected

figure(3)

% not windowed     sp1=(abs(fft(speech((x-wl/2):(x+wl/2-1)),512))) ; % old command
        % used before hamming window was added ( this uses rectangular window)
sp1_ham=(abs(fft(hamm_win.*speech((x-wl/2):(x+wl/2-1)),512)));
    % compute fft of segment of length wl centered at cursor using hamming window
sp1_ham_avg=zeros(1,512);  % pre-allocate space for averaging spectra

for m=1+15:512-15;          % make a running average of the magnitude squared
  sp1_ham_avg(m)=mean(sp1_ham((m-15):(m+15)).^2); % over a length of 30 points
end
clear m


plot([1:256]/256*5000,20*log10(sp1_ham(1:256)),'r',[16:256]/256*5000,10*
        log10(sp1_ham_avg(16:256))+20)
                    % plot smoothed curve 20 dB above unsmoothed curve
grid                 % turn on the grid

end % if dft=on

figure(1)

fprintf('\n');          % clear the command line
clear hamm_win  mn len ax m
```

B.2.7 mark2_down.m

```
% this m-file stores the current x-axis coordinate for later use

figure(1)
set(1,'pointer','crosshair'); % change cursor to notify user of action
[x,y]=get(gca,'currentPoint'); % get current cursor coordinates
x=x(1,1);                     % save only the x-axis coordinate
firstclick=x;                 % save this value in separate variable for later use
```

B.2.8 mark2_up.m

```
% this file marks the speech waveform at the place where the mouse
```

```
% button was pressed and also where the mouse button was released

figure(1)
set(1,'pointer','arrow');      % return cursor to normal shape
[x,y]=get(gca,'currentPoint'); % get current cursor coordinates
x=x(1,1);                      % save only the x-axis component
lastclick=firstclick;          % first mark speech at place determined by mark2_down.m
mark_speech                    % call m-file to draw vertical lines at mark

lastclick=x;                   % mark speech at place determined above
mark_speech                    % call m-file to draw vertical lines at mark
```

B.2.9 mark_speech.m

```
% this file simply marks the place in the speech waveform
% and stores the position in variable 'mark'

mark=[mark lastclick]; % x= position of last dft window (from ham.m)
marktot=marktot+1;   % we've now added one more mark to the list

col=['m' 'c' 'b' 'w']; % circle through remaining colors (not red, green, blk)


figure(1)
ax=axis;                  % get current axis for figure 1
markline(1,marktot+1)=line([mark(marktot+1) mark(marktot+1)],[ax(3) ax(4)]);
                          % draw a vertical line from top to bottom of window
set(markline(1,marktot+1),'color',col(rem(marktot+1,4)+1),'lineStyle','--');
                          % alternate colors and draw mark with dashed line


figure(2)
ax=axis;                  % get current axis for figure 2
markline(2,marktot+1)=line([mark(marktot+1) mark(marktot+1)],[ax(3) ax(4)]);
                          % draw a line in figure 2 from top to bottom of window
set(markline(2,marktot+1),'color',col(rem(marktot+1,4)+1),'lineStyle','--');
                          % alternate colors, same as in figure 1


figure(1)                 % return focus to figure 1


clear ax  col
```

B.2.10 mouse_down.m

```
% this m-file is called when the mouse button is pushed down

set(2,'pointer','crosshair');    % change cursor to notify user of action

[x2,y2]=get(gca,'currentPoint');  % get current point of cursor
```

```
x2=x2(1,1);                    % save only the x-axis component
```

B.2.11 mouse_up.m

```
% this function is called after the button is released

[x1,y1]=get(gca,'currentPoint'); % get current cursor position
x1=x1(1,1);                     % save only the x-axis component
right=x1;                       % right window edge place where button released
left=x2;                        % left window edge comes from mouse_down.m

set(2,'pointer','arrow');       % return cursor to normal arrow

%if exist('winmove')
%if winmove ==1;  % move left marker
%       left=x1;
%elseif winmove==2; % move right marker
%       right=x1;
%end

if left==0                      % if left edge is at 0
  left=1;                       % make it start at 1 (point 0 is undefined)
end

figure(1)
left1=min([left right]);        % find the lowest edge
right=max([left right]);        % right=largest edge
left=left1;                     % left= lowest edge
clear left1                     % clear temporary variable
top=max(speech(left:min(right,length(speech))));   % find max amp of speech in window
bottom=min(speech(left:min(right,length(speech)))); % find min amp of speech in window
axis([left right+1 bottom top]); % apply new window axis

window2_update;          .      % call m-file to update marks in window 2

end

clear x1 y1 winmove
```

B.2.12 window2_update.m

```
% this program updates the window marks, etc. in window 2

figure(2)
ax=axis;  % get axis coordinates

if zwinleft >0    % if left window marker exists
 delete(zwinleft) % delete old left window marker
 delete(zwinright)% delete old right window marker
end
```

```
zwinleft=line( [left left],[ax(3) ax(4)]);       % draw new right window line
zwinright=line( [right right],[ax(3) ax(4)]);                % draw new left window line
set(zwinleft,'color','g');                                   % draw line in green
set(zwinright,'color','g');                          % draw line in green
set(zwinleft,'ButtonDownFcn','window_left_move'); % watch for mouse button
                                                     % press while near line
set(zwinright,'ButtonDownFcn','window_right_move'); % watch for mouse button
                                                     % press while near line

figure(1)
set(position,'Value', (right+left)/2/length(speech)); % move fig. 1 to new position


clear ax win
```

B.2.13 window_left_move.m

```
% this deletes the left green window marker in figure 2

winmove=1 ;                   % flag that it was the left marker we want to move
set(2,'pointer','crosshair');  % change cursor so user knows button was pressed
fprintf('Move left window edge...  \n'); % write to screen which option was selected
```

B.2.14 window_move.m

```
% this function is called after the mouse button is released
% window markers are changed in fig 2 and window 1 is scaled appropriately

[x1,y1]=get(gca,'currentPoint');  % get the location of cursor
x1=x1(1,1);                         % keep only the x coordinate

set(2,'pointer','arrow');         % change cursor back to standard arrow

if exist('winmove')                  % if the window has been moved before
    if winmove ==1;              % if it was the left marker that was moved
      left=x1;               % move left marker

     elseif winmove==2;              % if it was the right marker that moved
      right=x1;              % move right marker
    end


if left==0                         % if left marker was moved to 0
  left=1;                          % place marker at point 1 (point 0 undefined)
end
```

```
figure(1)
left1=min([left right]);          % find the minimum of the two markers
right=max([left right]);          % the rightmost marker is called right
left=left1;                       % the leftmost marker is called left
clear left1                       % delete temporary variable
top=max(speech(left:min(right,length(speech)))); % find max amp of speech in window
bottom=min(speech(left:min(right,length(speech)))); % find min amp of speech in window
axis([left right+1 bottom top]);  % set axis according to [left right min max]

window2_update;                   % call m-file to update window 2

end
clear x1 y1 winmove
```

B.2.15 window_position.m

```
% this file is called whenever the window position slider is moved

sli_pos=get(position,'value');              % get new slider position
ax=axis;                                    % get current axis
len=ax(2)-ax(1);                            % find x-axis length
totlen=length(speech);                      % find total length of waveform
left=max(1,sli_pos*totlen-len/2);           % find new left window position (>=1)
right=min(totlen,sli_pos*totlen+len/2);     % find new right window positon (<totlen)
if left==1, right=len;, end                 % keep length =len wven if pushed off left edge
if right==totlen, left=totlen-len;,end      % keep length =len even if pushed off right edge
top=max(speech(left:right));                % find max amp of speech in new window
bottom=min(speech(left:right));             % find min amp of speech in new window

axis([left right bottom top]);              % apply new axis

window2_update;                             % call m-file to update window 2

clear len totlen ax left right top bottom
```

B.2.16 window_right_move.m

```
% this deletes the right  green window marker in figure 2

winmove=2;                        % flag that it was the right marker we want to move
set(2,'pointer','crosshair'); % change cursor so user knows button was pressed

fprintf('Move right window edge...  \n'); % print which option was selected
```

B.3 Synthesized Transient Programs

B.3.1 trans_length.m

```
% this m-file explores how the error decreases as the length of
% the best-fit transient increases
w=0:.1:pi;                    % create a freq vector
winv=1./w;                    % find 1/f
winv(1)=8;                    % don't let 1/f blow up at
winv(2)=7;                    %   very small values of f

finv=(1./(1:1024)./1024*5)'; % convert freq to Hz

plot([100:1024]./1024*5,20*log10(finv(100:1024))+102,'-.') % plot 1/f
grid on
hold on

for i=1:4;                    % for best fit transient of length 2,3,4,5
%weight=[zeros(1,4) ones(1,22) 10*ones(1,6)];
weight=[ones(1,4) ones(1,22) ones(1,6)]; % create even weighting function
[b,a]=invfreqz(winv,w,i,0,weight)     % find best fit trans of length i+1
[h3,w3]=freqz(b,a,1024);              % compute the spectrum of transient
plot(w3(100:1024)./pi*5,20*log10(abs(h3(100:1024)))) % plot result


figure(1)



%plot([1:1024]./1024*5, 20*log10(abs(h3)))
%plot(20*log10(finv)+6, 20*log10(abs(h3)))


mse(i,:)=sqrt(mean(diff([20*log10(finv(203:1024))+90 ; 20*log10(abs(h3(203:1024)))]).^2))
  % calculate the error for transient of length 2,3,4 and 5
end

hold off
title('Transient Frequency Content (length 2,3,4,and 5 samples)')
xlabel('Frequency (kHz)')
ylabel('Magnitude (dB)')


%figure(2)
%plot(mse(:,:))
figure(1)
```

B.3.2 transient4.m

```
% this m-file analyzes the best fit transient
```

```
w=0:.1:pi;                          % create an frequency axis
winv=1./w;                           % find 1/f
winv(1)=8;                                     % don't let the first
winv(2)=7;                                       %     2 samples grow too big
f=w/pi*5;                                        % convert w to Hz
plot(f,winv)                                     % plot f vs 1/f

%weight=[zeros(1,4) ones(1,22) ones(1,6)];
weight=[zeros(1,4) ones(1,22) 10*ones(1,6)];  % create a weighting function
[b,a]=invfreqz(winv,w,3,0,weight);           % find best fit transiet for
                                             % length=4 and above weighting fct
norm=b/b(1)                                % normalize amplitude
figure(4)
plot(-1:7,[0 0 norm 0 0 0 ],-1:7,[0 0 norm 0 0 0 ],'x') % plot waveform
grid
xlabel('Sample Number')
ylabel('Amplitude')
title('Best Fit Transient')

figure(1)
[h3,w3]=freqz(b,a,1024);                    % find spectrum of best fit

plot(f(7:32),20*log10(winv(7:32)),w3/pi*5,20*log10(abs(h3)));
                      % plot best fit spectrum and 1/f
xlabel('Frequency (kHz.)')
ylabel('Magnitude(dB)')
title('Transient Approximation to 1/f')
text(1.3,2.2,'1/f')
text(0.2,-0.5,'Approximation')
grid on


%plot difference
figure(2)

ideal= 20*log10(1./(w3(100:1024)/pi*5));         % 1/f
approx=20*log10(abs(h3(100:1024)))-3.5;          % best fit transient spectrum
plot(w3(100:1024)/pi*5,ideal,w3(100:1024)/pi*5, approx)
                     % plot both ideal and approx spectrum

figure(3)
plot(w3(100:1024)/pi*5,abs(ideal- approx))    % plot error  (ideal-approx)
grid
title('Error in Transient Approximation')
xlabel('Frequency (kHz.)')
ylabel('Difference from ideal 1/f (dB)')


figure(1)
```

B.4 Perception Test Analysis Programs

B.4.1 ltr2hist.m

```
% this file takes the answers from xptest.ans and computes
% histograms


function [percent]=ltr2hist(testfor1,testfor2,letter)

%letter='p'  % stop consonant to study
%testfor1=40  % find which questions has element testfor as a choice
%testfor2=0  % and element testfor 2 as the second choice

all_names=zeros(7,46);           % pre-allocate to increase speed

% for ch      all_names=zeros(7,48);  % all ch's have longer file names

names=[ 'ken_'; 'sharl'; 'helen'; 'mark_'; 'alice'; 'noel_' ; 'joyce'];
                                        % file names of answers
clear question

for nn=1:min(size(all_names))      % for all the names
this_name=[names(nn,:)  letter ];  % create file name
all_names(nn,:)=[ '/usr/users/massey/matlab/results/x' letter '/' this_name '.ans'];
              % create full path names

eval(['load ',all_names(nn,:)]);  % load file
this_name=eval(this_name);         % find name for graph titles

xtest=['x',letter,'test2'];       % all files stored as xptest2.ans or xktest2.ans
eval(['load ./x',letter,'/',xtest,'.ans'])  % load the file

xtest=eval(xtest);  % convert from name to matrix

for n=21:100 % for each of the 80 answers convert from 1 or 2 to dB of transient
xtest(n,2:3)=xtest(xtest(n,2),2:3); % find which levels were presented as options

     if nn==1 & ((xtest(n,2)==testfor1) | (xtest(n,3)==testfor1))
                  % if one  element of the pair is the desired value
             if ((xtest(n,2)==testfor2) | (xtest(n,3)==testfor2))
                  % if other elements of the pair is the desired value
                if (exist('question')) % if not the first pair found
                      question=[question (n-20)]; % append to array
                  else  % if it is the first pair found (array does not exist)
                     question=(n-20); % create first elemet of the array
                end
             end
     end
%
%     if nn==1 & ((xtest(n,2)==testfor2) | (xtest(n,3)==testfor2))
%             question=[question (n-20)];
%       end
```

```
end


for k=1:80                              % for each of th questions
if this_name(k)==1                      % if response = first element of pair
   this_name(k)=xtest(20+k,2);          % find transient amp of first element
elseif this_name(k)==2                  % if response = second element of pair
   this_name(k)=xtest(20+k,3);          % find transient amp of second  element
else
  error('ERROR: answer not 1 or 2');  % error:  response not correct
end


compare1(nn,:)=this_name(question,:)';% save this person's responses for later

end

figure(1)
hist(this_name ,[ 0 10 20 30 40 50 60]) % plot hist.  for this person's resp.

total=[total this_name'];               % add this person to the rest




end % for nn

figure(2)
hist(total ,[ 0 10 20 30 40 50 60])     % plot histogram of all subjects
title(['x',letter,'test - ',num2str(min(size(all_names))),' subjects'])
xlabel('Transient magnitude (dB)')
ylabel('Number of times chosen')


% print /usr/users/massey/matlab/graphics/ch5/xptest;


figure(3)

[junk1,junk2]=hist(compare1(:),[ 0 10 20 30 40 50 60])
            % calculate number of elements in each bar

for ii=1:6
  if junk2(ii)==testfor1
            %  if bar is the desired magnitude
      percent=100*junk1(ii)/(sum(junk1));
            % percent of time known magnitude was chosen
  end
end

hist(compare1(:),[ 0 10 20 30 40 50 60])
            % plot histogram when one element of the pair was known
```

```
title(['x',letter,'test - choices where ',num2str(testfor1),' dB was presented'])
xlabel('Transient magnitude (dB)')
ylabel('Number of times chosen')
```

B.4.2 max_compare.m

```
% this file records the responses in which one element of the pair
% was the most preferred transient level


clear

all˙percent=zeros(1,7); % clear answer matrix

letter='t'            % stop consonant to be studied
testfor1=40;          % most preferred transient level

mags=[0 30 50 60];    % other levels that were presented

for all_mags=1:4      % for each of the other transient magnitudes
testfor2=mags(all_mags) % check this specific magnitude
[percent]=ltr2hist(testfor1,testfor2,letter)  %calculate percent of times chosen
all_percent(mags(all_mags)/10+1)=percent;  % convert to percent not chosen
end



for ijk=1:7                  % for each magnitude level
  if all_percent(ijk)==0     % if the levels were not tested
    all_percent(ijk)=100;    % set so percent chosen = 0 (not chosen=100)
  end
end



figure(4)                    % plot results-- magnitude vs percent of times chosen
bar([0 10 20 30 40 50 60],100*ones(1,7)-all_percent)

total_results=100*ones(1,7)-all_percent  % print numerical results to screen

xlabel('Transient magnitude (dB)')
ylabel('Percent of time chosen')
title(['x',letter,'test - choices where ',num2str(testfor1),' dB was presented'])


%print /usr/users/massey/matlab/graphics/ch5/xchtest_30;
```

Bibliography

[1] G. Fant. *Acoustic Theory of Speech Production.* Mouton, The Hague, 1960.

[2] K.N. Stevens. *Acoustic Phonetics*, in preparation.

[3] K.N. Stevens. "Modelling affricate consonants". *Speech Communication*, Vol. 13, pp 33-43, 1993.

[4] P. Ladefoged and A. Traill."Clicks and their accompaniments".*Journal of Phonetics*, Vol. 22, pp 33-64, 1994.

[5] K.N. Stevens. "Airflow and Turbulence Noise for Fricative and Stop Consonants: Static Considerations". *Journal of the Acoustical Society of America.* Vol. 50 (4), pp 1180-1971, 1971.

[6] S. Maeda. "On generation of sound in stop consonants". Speech Communication Group Working Papers, Research Laboratory of Electronics, MIT, Vol. 5, pp 1-14, 1987.

[7] D.H. Klatt and L.C. Klatt. "Analysis, synthesis, and perception of voice quality variations among female and male talkers". *Journal of the Acoustical Society of America*, Vol. 87 (2), pp 820-857, 1990.

[8] H. Fujisaki, K. Hirose, and Y. Asano. " Proposal and evaluation of a new type of terminal analog speech synthesizer". *Proceedings of 1990 International conference on spoken language processing.*, 1990.

[9] S. Blumstein and K.N. Stevens. "Acoustic Invariance in Speech Production: Evidence from Measurements of the Spectral Characteristics of Stop Consonants".*Journal of the Acoustical Society of America*, Vol. 66 (4). pp 1001-1017, 1979.

[10] H. Winitz, M. Scheib, and J. Reeds. "Identification of Stops and Vowels for the Burst Portion of /p,t,k/ Isolated from Conversational Speech".*Journal of the Acoustical Society of America*, Vol. 51 (4), pp 1309-1317, 1972.

[11] D. Kewley-Port. "Time-varying features as correlates of place of articulation in stop consonants".*Journal of the Acoustical Society of America*, Vol. 73 (1), pp 322-335, 1983.

[12] D. Kewley-Port, D. Pisoni, and M. Studdert-Kennedy. "Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants".*Journal of the Acoustical Society of America*. Vol. 73 (5). pp 1779-1793, 1983.

[13] B. Repp and H. Lin. "Acoustic Properties and Perception of Stop Consonants Release Transients". *Journal of the Acoustical Society of America*. Vol. 85 (1). pp 379-396, 1989.

[14] L.C. Shen and J.A. Kong. *Applied Electromagnetism*. 1987.

[15] O. Fujimura."Bilabial Stop and Nasal Consonants: a Motion Picture Study and its Acoustical Implications".*Journal of Speech and Hearing Research*, Vol. 4 (3), pp 233-247, 1961.

[16] R. Kent and K. Moll."Cineflourographic analyses of selected lingual consonants".*Journal of Speech and Hearing Research*, Vol. 15, pp 453-473, 1972.

[17] D.H. Klatt. "Software for a cascade/parallel formant synthesizer".*Journal of the Acoustical Society of America*, Vol. 67 (3), pp 971-995, 1980.