GESTURAL OVERLAP OF STOP-CONSONANT SEQUENCES

by

Sherry Yi Zhao

Submitted to the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

Massachusetts Institute of Technology

May 20, 2003

Author ____

Department of Electrical Engineering and Computer Science
May 20, 2003

Certified by ____

Kenneth N. Stevens
Thesis Supervisor

Accepted by ____

Arthur C. Smith
Chairman, Department Committee on Graduate Theses

GESTURAL OVERLAP OF STOP-
CONSONANT SEQUENCES

by

Sherry Yi Zhao

Submitted to the Department of Electrical
Engineering and Computer Science

May 20, 2003

In Partial Fulfillment of the Requirements for the
Degree of Master of Engineering and Electrical
Engineering and Computer Science

ABSTRACT

This study used an analysis-by-synthesis approach to discover possible principles governing the coordination of oral and laryngeal articulators in the production of English stop-consonant sequences. Recorded utterances containing stop-consonant sequences were analyzed acoustically, with focus on formant movements, closure durations, release bursts, and spectrum shape at low frequencies. The results of the acoustic analysis were translated into general gestural timing estimates. From these estimates, a set of possible principles was derived. Both the general gestural estimates and the derived principles were verified and refined through quasi-articulatory synthesis using HLsyn. Perception tests composed of synthetic sequences with varying degrees of overlap were administered. From acoustic analysis, synthesis verification, and perception testing, two principles emerged. First, V1C1#C2V2 stop-consonant sequences with front-to-back order of place of articulation have more overlap of articulators than those with back-to-front order; this agrees with past research findings (Chitoran, Goldstein, and Byrd, 2002). The extent of the overlapping usually does not go beyond the obliteration of the C1 release burst. Second, gestural overlap involving laryngeal articulators exists but varies from individual to individual. The voicing of C1 usually affects the voicing of C2 in V1C1#C2V2 sequences.

Thesis Supervisor: Kenneth N. Stevens

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgements

First and foremost I would like to thank my thesis advisor, Professor Kenneth N. Stevens, for giving me the opportunity to work on such an exciting topic. His help, support, and guidance are crucial in the completion of this thesis.

I also thank everyone who took time out of his/her busy schedule to participate in the speech recording and perception testing aspects of this study. Without their help, this study would be impossible.

I must also thank all of my colleagues whom I interacted with over the past year. They have made my experience rich and satisfying.

Last, but certainly not least, very special gratitude is due to my parents and Hanching for their support and encouragement.

Chapter 1: Introduction

The coordination of articulatory movements is essential in speech production. However, the search for possible articulatory and acoustic invariance in the process has proven to be a significant challenge. Past research in the area has found articulatory timing to be highly variable and dependent on prosodic and contextual factors. The question of what the underlying principles that govern the coordination of articulatory movements are, if they exist, still remains.

This study employed the analysis-by-synthesis approach to discover possible principles governing the coordination of oral and laryngeal articulators in the production of English stop-consonant sequences. There were two phases to this study. Phase one consisted of analyzing acoustic data such as formant movements, closure durations, release bursts, and spectrum shape at low frequencies of naturally-spoken stop-consonant sequences. Based on the analysis, general gestural timing estimates were inferred. From those estimates, an initial set of possible principles was derived.

In the second phase of the study, both the general gestural estimates and the derived principles were verified and refined through quasi-articulatory synthesis using HLsyn. Consonant sequences were first generated using inferred gestural timing estimates from acoustic data. The synthetic utterances were acoustically and perceptually compared to the actual utterances in order to verify and refine the articulatory timing estimates; the resulting gestural timing estimates were used to cross-check and refine the general estimates inferred in phase one. Next, in order to verify the derived principles, synthetic sequences with varying degrees of overlap were generated. Perception tests composed of these synthetic consonant sequences were administered. The results of the tests were used to verify the validity of each principle and the extent of its applicability.

Before diving into the specifics of the study, the next chapter will provide a review of past research conducted in the area of gestural overlap of stop-consonant sequences. In addition, Chapter 3 will provide a review of the mechanics and acoustical properties of stop consonants, the basics of gestural timing estimation, as well as an outline of the parameters and usage of HLsyn.

Chapter 2: Review of Literature

In the framework of articulatory phonology proposed by Browman and Goldstein (1992), gestures are the "basic units of phonological contrast". Utterances are modeled as organized patterns of overlapping and non-overlapping gestures. There also exists an abstract, theoretical framework containing general principles of how the gestures are organized in speech production; gestural relations are expressed in terms of dynamics and phasing. Research is still being conducted to develop the framework.

A different model of acoustic speech, represented in terms of segments described by sets of binary distinctive features, was presented by Stevens (Stevens, 2002). These features may be articulator-bound or articulator-free; the articulator-free features fall into one of three classes: vowel, glide, or consonant. In addition to the primary articulatory and acoustic correlates of a feature, enhancing actions might also exist. For example, the somewhat fronting of the tongue-body in the production of an alveolar stop is an enhancing gesture. Such gestures are thought to heighten perceptual contrast and remain intact in consonant sequences; this raises a question investigated in this study: Are defining gestures more susceptible to obliteration when they are supported by enhancing actions?

Questions concerning the linguistic effects on gestural coordination have also been posed; they have been the focus of several studies. Byrd's research on the articulatory timing of consonant sequences using electropalatography (EPG) found that consonants in word-onset positions exhibit less gestural overlap than those in coda positions (Byrd, 1996). In addition, more overlap was found in coronal-velar consonant sequences than velar-coronal sequences.

Chitoran, Goldstein, and Byrd's study on the gestural overlap of stop consonants in Georgian using the EMMA (Electromagnetic Midsagittal Articulometer) magnetometer system led to similar findings (Chitoran, Goldstein, and Byrd, 2002). Consonants in word-initial positions had less gestural overlapping than those in word-internal positions. Sequences with front-to-back order of place of articulation were found to have more gestural overlap than those with back-to-front order. Examples having front-to-back order of place of articulation are labial-coronal, coronal-dorsal, and labial-dorsal stop-stop sequences. Examples having back-to-front order of place of articulation are coronal-labial, dorsal-coronal, and dorsal-labial stop-stop sequences.

Zsiga's study on acoustical evidence for gestural overlap of VC#CV stop-consonant sequences also found indications of more overlap for sequences with front-to-back order of place of articulation than for those with back-to-front order (Zsiga, 1994). Formant transitions into the [d] closure for [d#p] and [d#k] sequences were compared to those of [d#t]. The reasoning behind the comparison was that if the second consonant did not influence the formant transitions into the closure of the first consonant, then the transitions into the first consonant should be similar for all three stop-stop sequences. The alveolar-alveolar stop sequence [d#t] was used as a control. The results of the study showed that formant movements into the [d] closure were more affected when followed by [k] than by [p] (Zsiga, 1994).

In addition to acoustic and articulatory data analysis, perception tests involving stop-stop sequences have been administered. Byrd's perception study involving synthesized [b#d] and [d#b] stop sequence stimuli found that as the overlap increased, correct identification of C1, the first consonant in a VCCV sequence, was more reduced for [d#b] than for [b#d] (Byrd, 1992). Similar findings were made from Surprenant and Goldstein's perception experiment involving natural speech [p#t] and [t#p] sequences in English. With considerable but similar amounts overlap in the sequences, C1 in [p#t] was correctly identified much more often than C1 in [t#p] (Surprenant and Goldstein, 1998). These studies suggest that perception recoverability may be a possible explanation for why there is less overlap in sequences with back-to-front order of place of articulation. However, in Repp's perception tests of carefully articulated stop-stop sequences, no strong suggestions of C2's co-articulatory influences on C1 were found; C2 is the second consonant in the VCCV sequence (Repp, 1983). He did find significant statistical evidence for a higher F2 values just prior to the C1 closure for [b#g] sequences than for [b#d] sequences.

These research efforts played crucial roles in fueling the motivation and guiding the direction of this study. First, stop consonants have been shown to be not discretely articulated. Many demands are placed on the lips, tongue, and glottis to articulate stop-stop sequences in relatively short durations, making such utterances interesting subjects in the study of articulatory coordination. Second, previous findings of word position and order of place of articulation influencing gestural overlap of stop-stop sequences indicated that linguistic effects must be considered and examined. Furthermore, previous studies' concentration solely on oral gestures prompted the inclusion of laryngeal gestures in this study. Finally, the possible link between the presence of enhancing gestures and defining gestures' susceptibility to obliteration was also examined.

Chapter 3: Theoretical Framework

This chapter provides a review of the mechanics and acoustical properties of stop consonants, the basics of gestural timing estimation, as well as an overview of the parameters and usage of HLsyn.

3.1 Mechanics and Acoustical Properties of Stop Consonants

The production of stop consonants consists of three general stages. First, the appropriate articulator makes a closure at some point along the vocal tract. For a labial stop consonant, the closure is formed at the lips; for an alveolar, the closure is formed with the tongue tip pressing against the inner ridge of the upper-front-teeth gums; for a velar stop, the closure is formed with the back of the tongue touching the soft palate (see Table 1). Pressure behind the closure is built up as the vocal tract remains closed in the second stage. The final stage involves the release of the closure (Stevens, 1998).

Table 1: Types of Stop Consonants

	Voiced	Voiceless
Labial	b	p
Alveolar	d	t
Velar	g	k

Voicing characteristics are an important differentiating factor among stop consonants. For a voiced stop consonant, vocal folds continue to vibrate for some time after the closure; whereas for a voiceless stop, the vibration stops upon closure (see Figure 1). The release of all stops is typically followed by a burst of transient noise and frication noise from the constriction; aspiration noise from the glottis is usually produced for voiceless stops (see Figure 1). Voicing starts quickly following the release burst for voiced stop consonants; voicing onset is delayed for voiceless stops (Stevens, 1998).

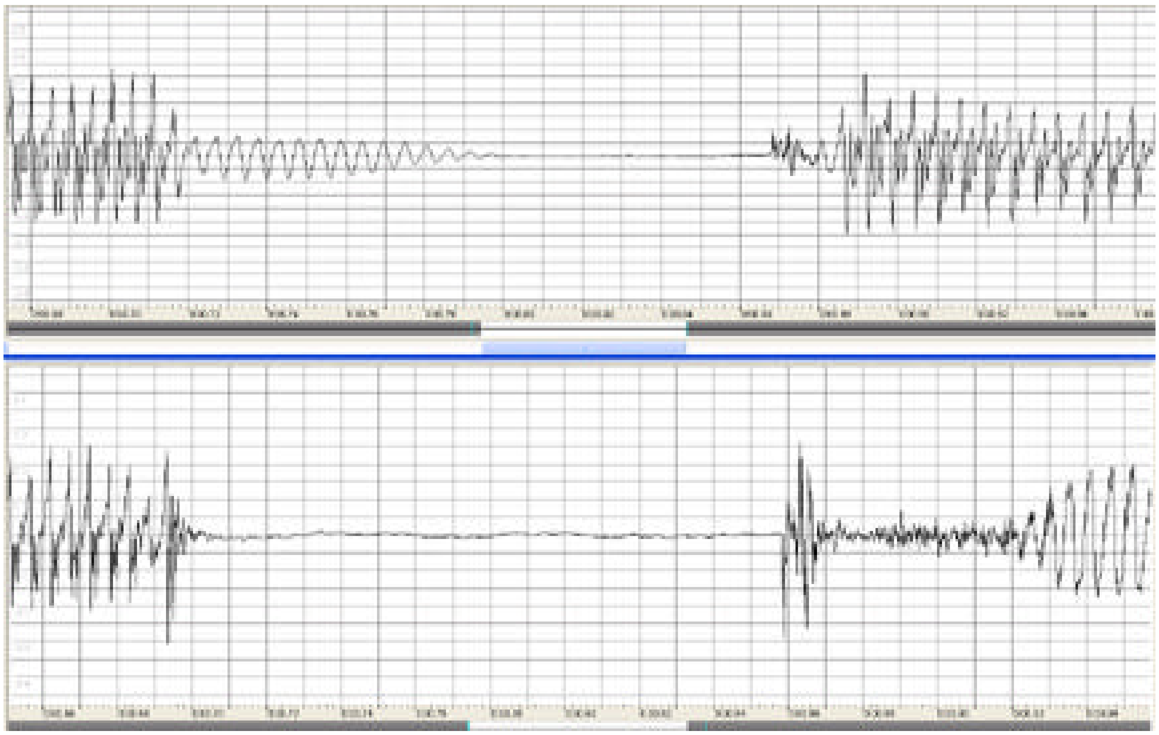

Figure 1. Waveforms for [ad#a] and [at#a]. The top panel displays the waveform for [ad#a]; the bottom panel displays the waveform for [at#a]. Notice the vocal fold vibrations present even after the d-closure; vibrations stop upon the t-closure. Also, the d-burst is much smaller than the t-burst.

The place of articulation is another differentiating factor among stop consonants. Evidence from formant data can be used to reveal information about the place of articulation as well as the movements of oral articulators. In general, F1 of stop consonants decreases as the constriction size decreases and increases as the constriction size increases. Movements of F2 and F3 differ among labial, alveolar, and velar stops, as well as whether the consonant is followed by a front or back vowel. In this study, the back vowel precedes and follows stop consonants. F2 of a back vowel going into a labial closure usually shows a small downward movement; the back body resonance is not affected by the front-body resonance change due to the lip constriction release. F3 also decreases, although not dramatically (see Figure 2). F2 of a back vowel going into an alveolar closure usually increases as the tongue tip forms a closure with the upper-front-teeth gums and the tongue body exhibits slight fronting (Stevens, 1998). F3 decreases, although not dramatically (see Figure 2). Finally, F2 increases during the transition into a velar closure from a back vowel, while F3 decreases. The two formants are close together just prior to the point of closure (Stevens, 1998).

Formant transitions of a stop consonant moving into a back vowel approximately mirror those of the back vowel moving into the same consonant (Stevens, 1998).

3.2 Basics of Gestural Timing Estimation

Gestural timing estimates are made based on acoustic data. This section first discusses gestural timing estimation in cases where there is no overlap, and then discusses acoustic manifestations of gestural overlap and how they can be used to make articulatory timing estimates.

In cases where there is no gestural overlap, the acoustic properties of consonants in V1C1#C2V2 sequences should not be different from the same C1 in VC#V sequences and the same C2 in V#CV sequences. The movements of the articulators for the production of C1 should be the same whether it is in V1C1#C2V2 or VC#V sequences;

similarly, the movements should be the same for C2 whether it is in V1C1#C2V2 or V#CV sequences.

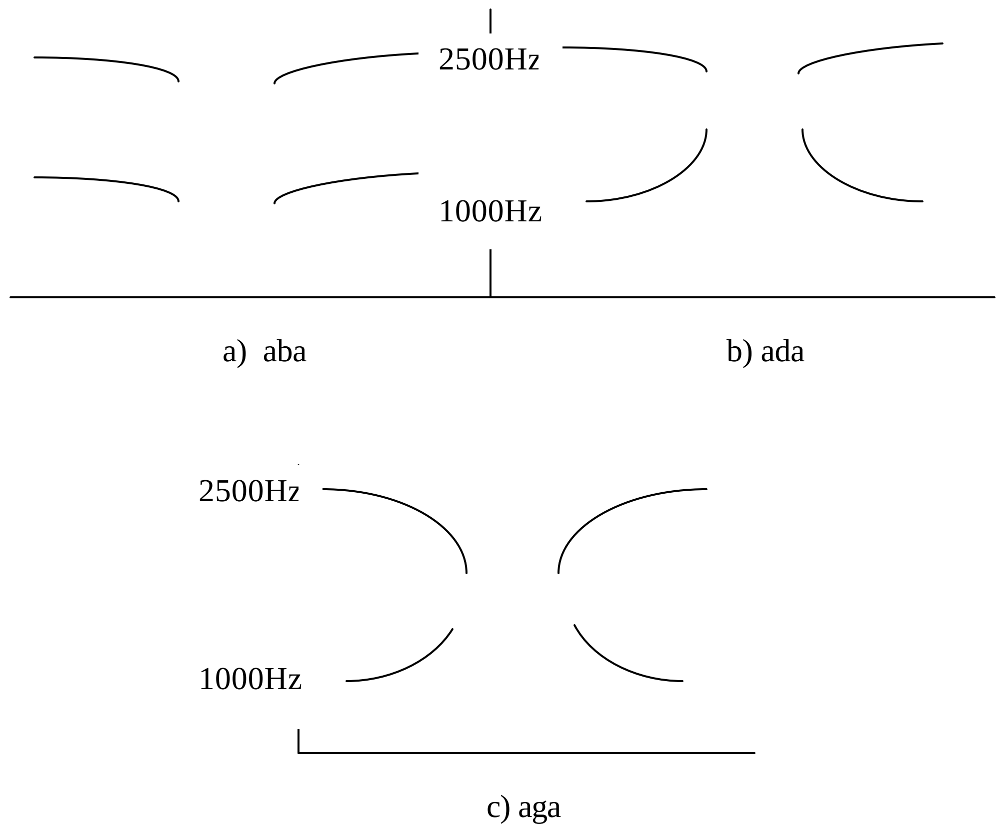

Figure 2. Formant Movements of Stop Consonants. a) F2 and F3 movements for [aba]. b) F2 and F3 movements for [ada]. c) F2 and F3 movements for [aga].

For V1C1#C2V2 sequences with no gestural overlap, the formation of the C1 closure by the appropriate oral articulator begins toward the end of V1 and should be made by the time pitch periods end. If C1 is voiced, the area of the glottis remains relatively unchanged. On the other hand, if C1 is voiceless, the area of the glottis should start to increase starting around the end of V1; the spreading of the glottis should reach its maximum around the time just prior to the C1 release. The C1 closure remains until the time of its release. The vocal folds remain lax for the voiced-stop burst and stiffen for the voiceless-stop burst. The release of the C1 closure generates a burst which should match the timing of the actual event as indicated by acoustic evidence. If both C1 and C2 are voiced, the area of the glottis and the stiffness of the vocal folds remain unchanged. If both consonants are voiceless, the area of the glottis remains large and the vocal folds remain stiff for C2. If C1 is voiced and C2 is voiceless, upon the release of C1, the glottis should begin to spread; the increase in the area of the glottis should reach a maximum around the time just prior to the C2 release. In addition, the vocal folds would stiffen for the voiceless C2 release. If C1 is voiceless and C2 is voiced, upon the release of C1, the stiffness of the vocal folds begins to decrease. The area of the glottis should also begin to decrease to a point where voicing is possible; the decrease should terminate around the time when the C2 closure is made or some time soon afterwards, since the glottis usually does not close immediately.

Since no gestural overlapping is assumed here, the formation of the C2 closure by the appropriate articulator(s) begins after the C1 release. The C2 closure remains until the time of its release. The release of the C2 closure generates a burst which should match the timing of the actual event as indicated by acoustical evidence. If C2 is voiced, the area of the glottis does not change and the onset of voicing begins fairly quickly. Also, the vocal folds remain lax for the voiced-stop release. If C2 is voiceless, the area of the glottis starts to decrease upon the release of the closure ending at a point where voicing is possible. The vocal folds stiffen for the voiceless-stop release. The glottis does not close immediately, since the onset of voicing is usually delayed for voiceless stops.

However, gestural overlap may very well exist for the production of stop-consonant sequences. In cases where there is overlap, the acoustic properties of the consonants in VC#CV sequences should be different from those of the same consonants in VCV sequences. The differences in the movements of the articulators between the production of stop consonants in sequence and those in singleton VCV sequences should manifest themselves acoustically and possibly reflect the influence exerted by the neighboring consonant. The following discusses such acoustic manifestations and how they may be used to infer gestural timing estimates.

3.2.1 Formant Movements

Formant movements reveal information about the movements of the oral articulators and the tongue body in the production of stop consonants. If the oral-articulator movements are the same for the production of stop consonants in sequence as for the production of the same consonants in singleton, the formant movements should also be the same. In addition, if the production of C2 does not overlap with the production of C1, formant movements for a particular consonant should all be the same no matter what the neighboring consonant is. Therefore, attention is paid to the differences in formant movements of stop-consonant sequences where C1 or C2 is the only variable. For example, if differences in formant movements into the C1 closure exist and is C2-dependent, then such evidence would suggest that the formation of the C2 closure began before the release of C1. The extent of the overlap may be inferred from the magnitude of the C1- or C2-dependent deviations in the formant movements.

3.2.2 Closure Durations

The closure duration of the stop consonants in sequence is another focal point in the estimation of gestural overlap. Evidence supporting no overlap in the production of stop-consonant sequences would be if the closure durations of C1 and C2 in sequence are roughly equal to the sum of singleton C1 and C2. However, if there is overlap, the closure durations would very likely be shortened. Therefore, the following measure will be taken of consonant sequences:

(closure duration VC1#C2V)/[(clos. dur. VC1#V)+(clos. dur. V#C2V)]   Eq. 1

This measure gives a ratio of how the closure duration for the consonant sequence compares with the combined closure durations of the consonants individually. Closure durations that are notably shorter than those of the combined C1 and C2 closure durations suggest overlap in the production of the consonants.

3.2.3 Absence of the C1 Burst

The absence of C1 bursts in stop-consonant sequences is another strong indicator of overlap. A likely cause of the absence of the C1 burst is the formation of the C2 closure before the C1 release.

3.2.4 Voicing Characteristics

Voicing characteristics reveal information about the movements of laryngeal articulators in the production of stop-consonant sequences, because it is determined by vocal-fold stiffness adjustments and glottal opening adjustments. If the voicing of C2 has no influence on the voicing of C1, voicing characteristics for a particular consonant in the C1 position should be the same no matter what consonant is in the C2 position. Similarly, if the voicing of C1 exerts no influence on that of C2, the voicing characteristics should reflect no difference among sequences where C1 is the only variable.

To measure the voicing characteristics of the stop consonants, focus is placed on the vowels preceding and following the consonants. The difference between the amplitude of the first harmonic (H1) and that of the second harmonic (H2) for vowels preceding and following the stop consonant(s) is positively correlated to the open quotient, the percent of the cycle in which the glottis is open (Hanson, 1997). The difference between H1 and the amplitude of the strongest harmonic in the first-formant range (A1) for vowels preceding and following the stop consonant(s), is positively correlated to the acoustic loss at the glottis. Such measurements may be made for pitch periods of V1 just prior to the C1 closure as well as for those just after the onset of V2 voicing to infer voicing characteristics of C1 and C2, respectively.

3.3 HLsyn

HLsyn plays a crucial role as an analysis and applications tool in this study. It is a commercial product of Sensimetrics Corporation. There are two parts to the synthesizer. The lower-level part consists of a formant synthesizer that carries out the actual sound generation. It is controlled by 40-some acoustically-oriented parameters. The higherlevel part consists of 13 articulatory-oriented parameters and a set of mapping equations which use these variables to calculate the lower-level parameters (Hanson and Stevens, 2002). When properly controlled, HLsyn has been shown to produce natural-sounding speech.

In this study, gestural estimates are directly transferred into manipulating HLsyn's articulatory-oriented parameters for the generation of the stop-consonant sequences. Table 2 lists the 13 articulatory parameters and how they are controlled.

Table 2: Manipulating HLsyn Parameters

Parameter	***Manipulation Technique***
al (cross-sec. area formed by lips)	Use gestural estimations for the lips.
ab (cross-sec. area formed by tongue blade)	Use gestural estimations for the tongue blade.
ag (area of glottal opening)	Use gestural estimations for the glottis opening.
an (cross-sec. area of velopharyngeal port)	Velopharyngeal port is closed for stop consonant production.
f0 (pitch)	Use pitch estimations.
F1- f4 (formants)	Use formant estimations; these parameters control tongue body gestures.
ps (subglottal pressure)	Assumed to be 8 cm H2O for males and 6.5cm H2O for females.
dc (compliance of vocal folds)	Set to negative to increase the phonation threshold for voiceless stops and set to positive to decrease the phonation threshold for voiced stops.
ue (active expansion of the vocal tract volume)	Positive for active expansion; negative for contraction.
ap (area of posterior glottal opening	Presence of opening normal for female speakers; set to 1mm^2 throughout

3.4 Analysis of Variance

In this study, the analysis of variance (ANOVA) is used to test for significance in the differences between groups of data. The null hypothesis tested by ANOVA is always the statement that the groups in question have the same mean. The resulting p-value from the analysis indicates the strength of the evidence against the null hypothesis (H_0). Table 3 contains an interpretation of p-values (Devore, 1995). Based on convention, alpha is set to 0.05. The p-value must be equal to or less than alpha in order to reject the null hypothesis.

Table 3: Interpretation of P-Values

P-Value	Interpretation
$P < 0.01$	very strong evidence against H_0
$0.01 \leq P < 0.05$	moderate evidence against H_0
$0.05 \leq P < 0.10$	suggestive evidence against H_0
$0.10 \leq P$	little or no real evidence against H_0

Chapter 4: Phase I

In phase one of this study, acoustic data such as formant movements, closure durations, release bursts, and spectrum shape at low frequencies, were collected and analyzed for a number of naturally-spoken stop-consonant sequences. Based on the analysis, gestural timing estimates were inferred. Finally, from the gestural timing estimates, an initial set of possible principles was derived.

4.1 Phase I Methodology

4.1.1 Natural Speech Recording

Four native speakers of American English, two male and two female, were asked to read the list of phrases (shown in Figure 3) at a normal pace. Two repetitions of each phrase were uttered. These utterances include individual English stop consonants in word-initial (V#CV) and word-final (VC#V) positions, as well as all possible combinations of VC#CV English stop-consonant sequences. The utterances were recorded on DAT tape in a sound-proof recording room. The recordings were digitized at a sampling rate of 11kHz.

4.1.2 Acoustic Data Collection

Spectrograms, pitch estimations and formant estimations were generated for each consonant sequence using the software tools xkl and lspecto. The following sections detail how formant movements, closure durations, stop consonant burst presence, and voicing characteristics were measured from the recorded speech data.

Say **saw pot** again.	Say **top ox** again.
Say **saw top** again.	Say **dot ox** again.
Say **saw cop** again.	Say **dock ox** again.
Say **saw bob** again.	Say **bob ox** again.
Say **saw dot** again.	Say **cod ox** again.
Say **saw got** again.	Say **dog ox** again.

Say **top pot** again.	Say **got pot** again.	Say **dock pot** again.
Say **top tot** again.	Say **got top** again.	Say **dock top** again.
Say **top cop** again.	Say **got cop** again.	Say **dock cop** again.
Say **top bob** again.	Say **got bob** again.	Say **dock bob** again.
Say **top dot** again.	Say **got dot** again.	Say **dock dot** again.
Say **top got** again.	Say **dot got** again.	Say **dock got** again.

Say **mob pot** again.	Say **cod pot** again.	Say **dog pot** again.
Say **mob top** again.	Say **cod top** again.	Say **dog top** again.
Say **mob cop** again.	Say **cod cop** again.	Say **dog cop** again.
Say **mob bob** again.	Say **cod bob** again.	Say **dog bob** again.
Say **mob dot** again.	Say **cod dot** again.	Say **dog dot** again.
Say **mob got** again.	Say **cod got** again.	Say **dog got** again.

Figure 3. Phrases Containing Stop-Consonant Sequences. Subjects are asked to read the list of phrases as naturally as they can.

4.1.2.1 Recording Formant Movements

F2 and F3 formant movements were recorded by measuring the values in the middle of the vowel preceding C1, as well as at 60ms, 40ms, 20ms, and immediately prior to the closure of C1. Measurements were also made at the onset of the vowel following C2, as well as at 20ms, 40ms, and 60ms after the vowel onset, and in the middle of the same vowel.

Each formant reading was taken as follows. Using xkl, a 6.4 Hamming window was placed over a pitch period at an appropriate time prior to the closure of C1 or after V2 onset. The first formant was measured at the frequency of the maximum low-frequency peak. The second formant was recorded at the frequency of the highest peak in magnitude after the first formant. The third formant was recorded at the frequency of the highest peak in magnitude after the second formant (see Figure 4).

4.1.2.2 Measuring Closure Duration

The closure duration for singleton consonants was measured starting at the end of the last pitch period prior to the stop closure and ending just before the release of the stop consonant burst. Figure 5 details how closure durations were measured for singleton consonants.

As for the closure durations for stop-consonant sequences, if both C1 and C2 release bursts were present, the closure duration of C1 was measured starting at the end of the last pitch period prior to the C1 closure and ending just before the C1 release burst; that of C2 was measured starting at the termination of the C1 burst and ending at the release of C2 (see Figure 5). If the C1 burst was absent, the closure duration was measured starting at the end of the last pitch period prior to the C1 closure and ending just before the C2 release burst.

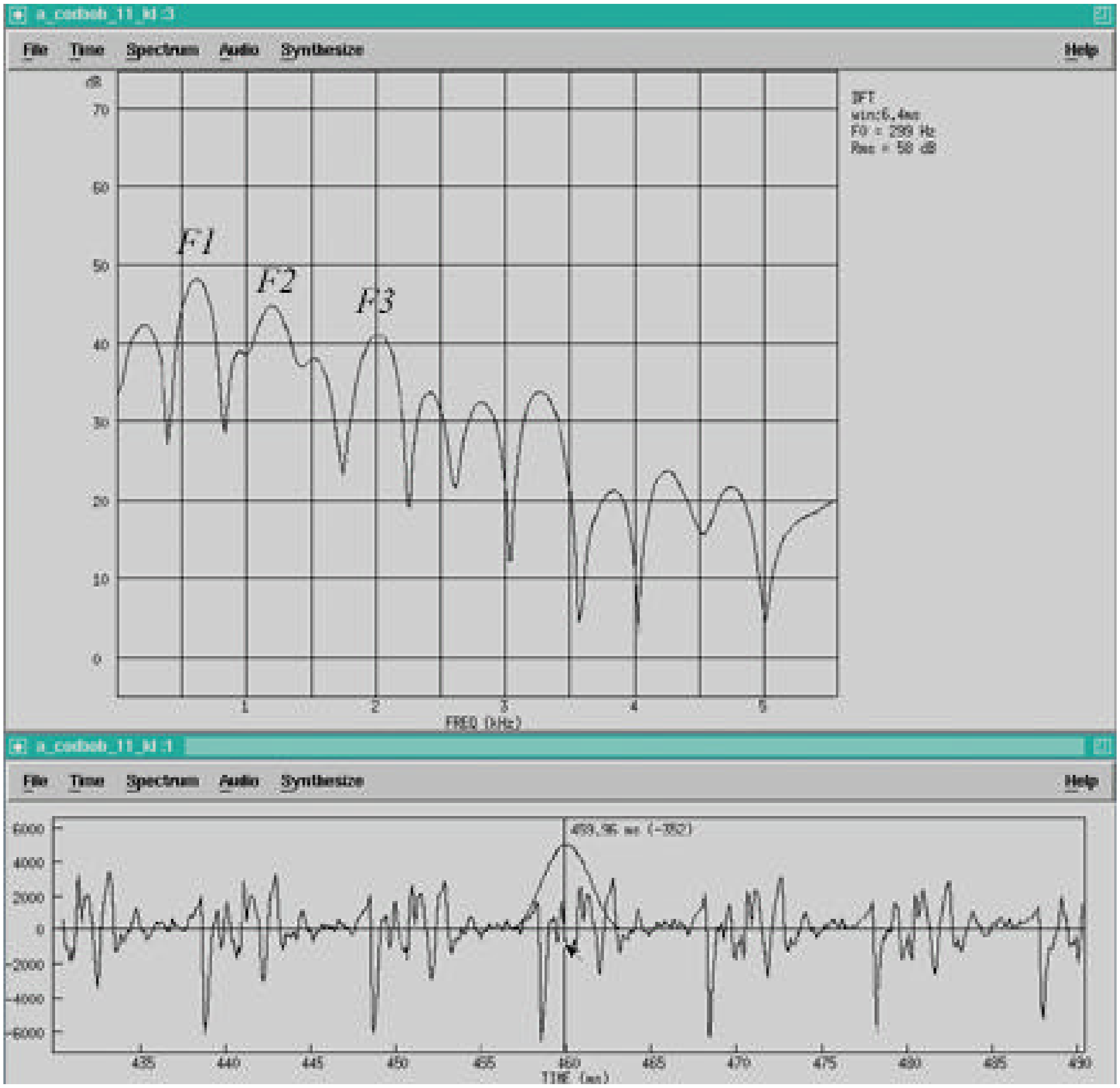

Figure 4. Formant Measurements. A 6.4 Hamming window was placed over a pitch period. The first formant was measured at the frequency of the highest peak in magnitude. The second formant was recorded at the frequency of the highest peak in magnitude after the first formant. The third formant was recorded at the frequency of the highest peak in magnitude after the second formant (see Figure 1).

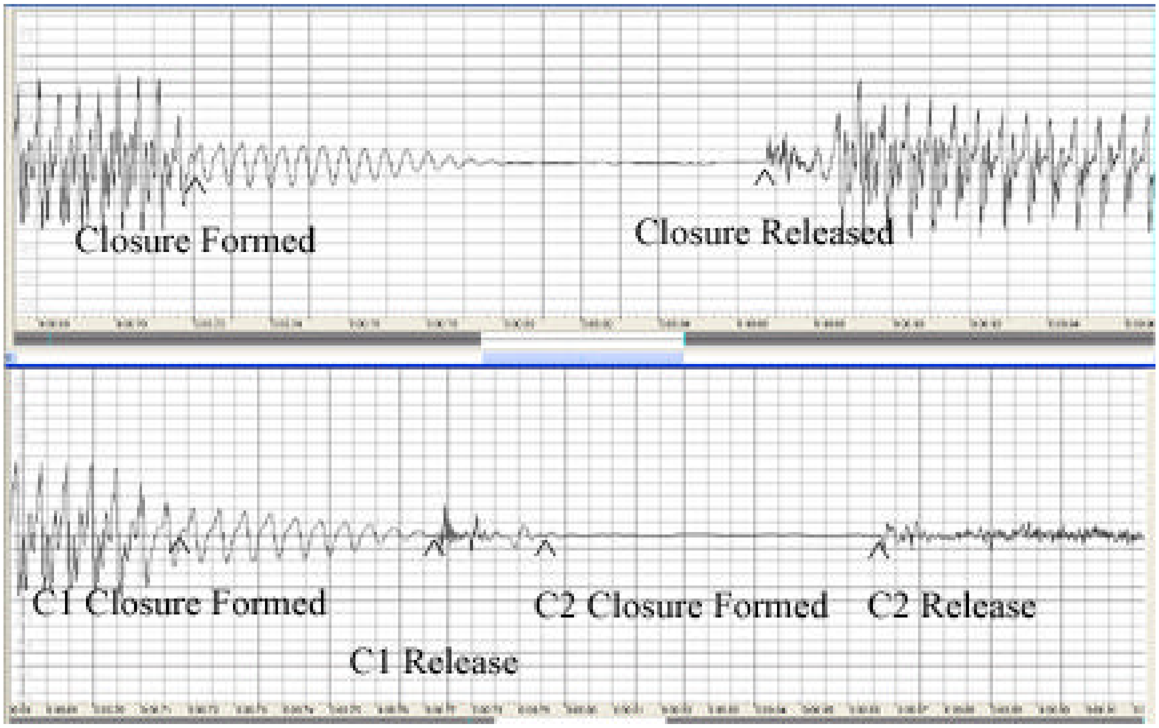

Figure 5. Closure Duration Measurements. The top panel shows how closure duration was measured for VCV sequences. The bottom panel shows how closure durations were measured for VCCV sequences, for cases where C1-bursts were present.

4.1.2.3 Presence and Length of Stop Burst

Acoustic evidence for the stop-consonant release in the form of bursts was examined for each individual consonant and consonant sequence; the time of the release burst was recorded (see Figure 5). The length of the release burst was measured starting at the onset of the release-burst and terminating at the end of the noise burst.

4.1.2.4 Measuring H1, H2 and A1

H1 (first harmonic), H2 (second harmonic), and A1 (strongest harmonic in the first-formant range) were measured by applying a Hamming window over the last four complete pitch periods of V1 prior to the closure of C1 (see Figure 6). The same process was repeated for the first four complete pitch periods of V2 following the C2 release. The length of the Hamming window varied from 30ms to 45ms depending on the pitch of the speaker. In addition, the frequency of each harmonic was recorded.

H1 and H2 were also adjusted to account for the F1 boost on the vocal-tract transfer function (Hanson, 1997). F1 readings were taken from the same pitch periods where the harmonics were measured. The final F1 value was the average of the readings taken at each of the last four complete pitch periods of V1 or at each of the first four complete pitch periods of V2; a 6.4ms Hamming window was used to measure F1. The method for measuring F1 was the same as that used for measuring F2 and F3 detailed in 4.1.2.1. The quantity, $20 \log_{10} [F1^2 / (F1^2 - f^2)]$, was subtracted from both H1 and H2; f is the frequency at which the harmonic is located (Hanson, 1997). From here on, H1* and H2* represent the adjusted values of H1 and H2.

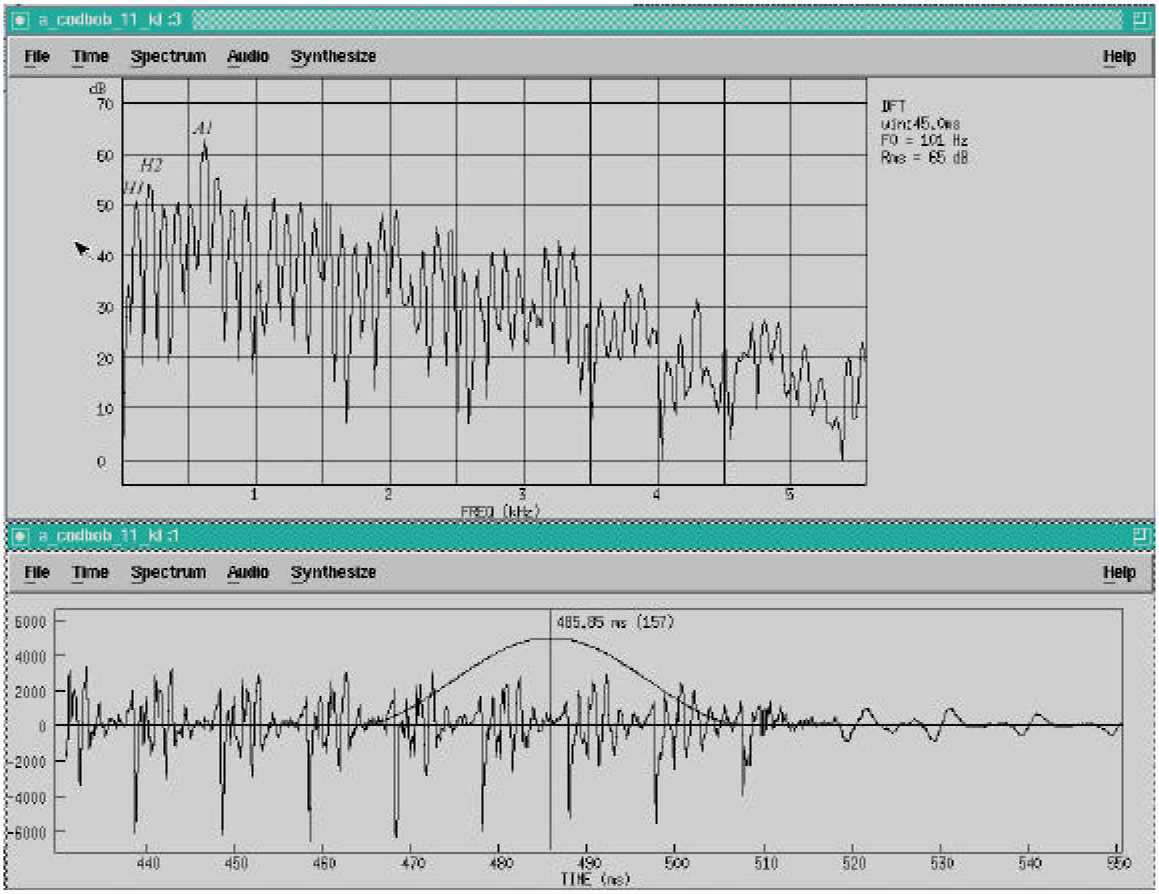

Figure 6: Spectrum Shape at Low Frequencies. H1 (first harmonic), H2 (second harmonic), and A1 (strongest harmonic in the first-formant range) were measured by applying a Hamming window over the last four complete pitch periods of V1 prior to the closure of C1.

4.2 Phase I Results

Figure 7 details the average F2 movement of V1 into the C1 closure, measured from the middle of the vowel to just before the closure, with respect to the order of place of articulation. Averages were calculated from detailed data gathered on the formant movements from all four speakers (see Appendix C though F).

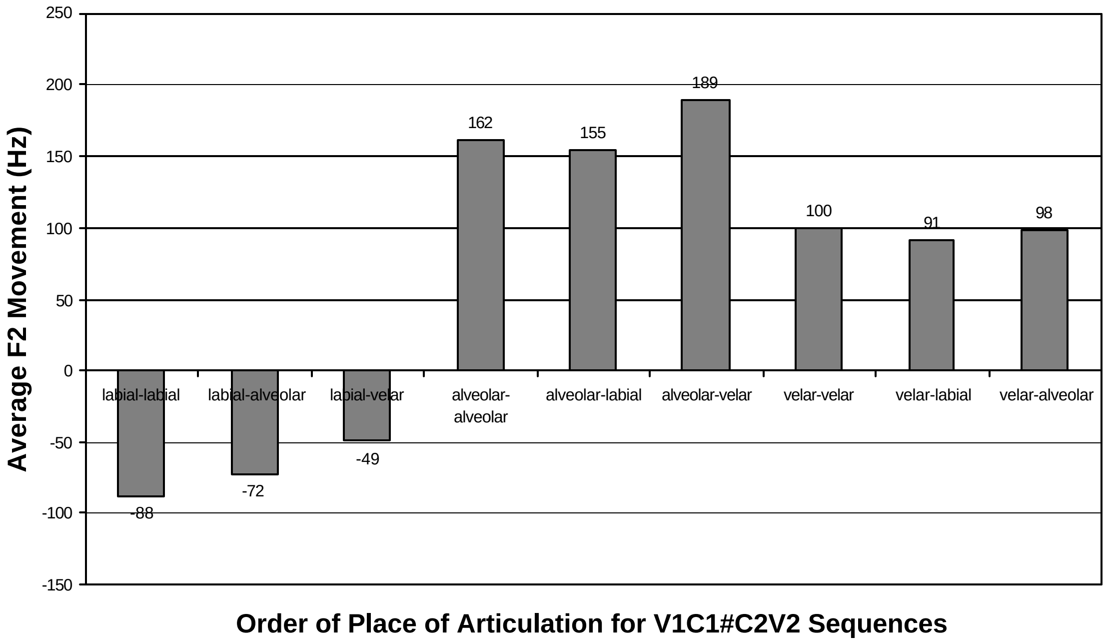

Figure 7. F2 Movement of V1 into C1 vs. Order of Place of Articulation. A plot of the average F2 movement of V1 into the C1 closure, measured from the middle of the vowel to just prior to the closure, with respect to the order of place of articulation. Averages were taken from both voiced and voiceless sequences with respect to order of place of articulation.

The average F2 movements into the labial-C1 closure seem affected by C2. When the labial was followed by an alveolar rather than another labial, the average decrease of F2 into the closure was smaller by 16Hz (see Figure 7). When the labial was followed by a velar, the average decrease of F2 was smaller than that of a labial-labial sequence by 39Hz.

As for sequences with an alveolar stop in the C1 position, the average F2 increase into an alveolar-labial sequence was smaller than that into an alveolar-alveolar sequence by 7Hz. On the other hand, the average F2 increase into an alveolar-velar sequence was larger than that into an alveolar-alveolar sequence by 27Hz.

Finally, for sequences with a velar stop in the C1 position, the average F2 increase into a velar-labial sequence was smaller than that into a velar-velar sequence by 9Hz. The average F2 increase into a velar-alveolar sequence was smaller than that into a velar-velar sequence by 2Hz.

Figure 8 details the average F2 movement of V2 following the C2 release, measured from the onset of V2 voicing to the middle of the vowel, with respect to the order of place of articulation. Averages were calculated from detailed data gathered on the formant movements from all four speakers (see Appendix C through F).

The average increase of F2 from C2 to V2 in alveolar-labial sequences was larger than that for labial-labial sequences by 8Hz (see Figure 8). The average F2 increase for velar-labial sequences was larger than that of labial-labial sequences by 34Hz.

The average F2 decrease into V2 for labial-alveolar sequences was larger than that of alveolar-alveolar sequences by 54Hz. The average F2 decrease into V2 for velar-alveolar sequences was larger than that of alveolar-alveolar sequences by 27Hz.

As for sequences with a velar stop in the C2 position, when a labial is in the C1 position instead of a velar, the average F2 decrease was larger by 19Hz. When an alveolar was in the C1 position instead of a velar, the average F2 decrease was larger by 72Hz.

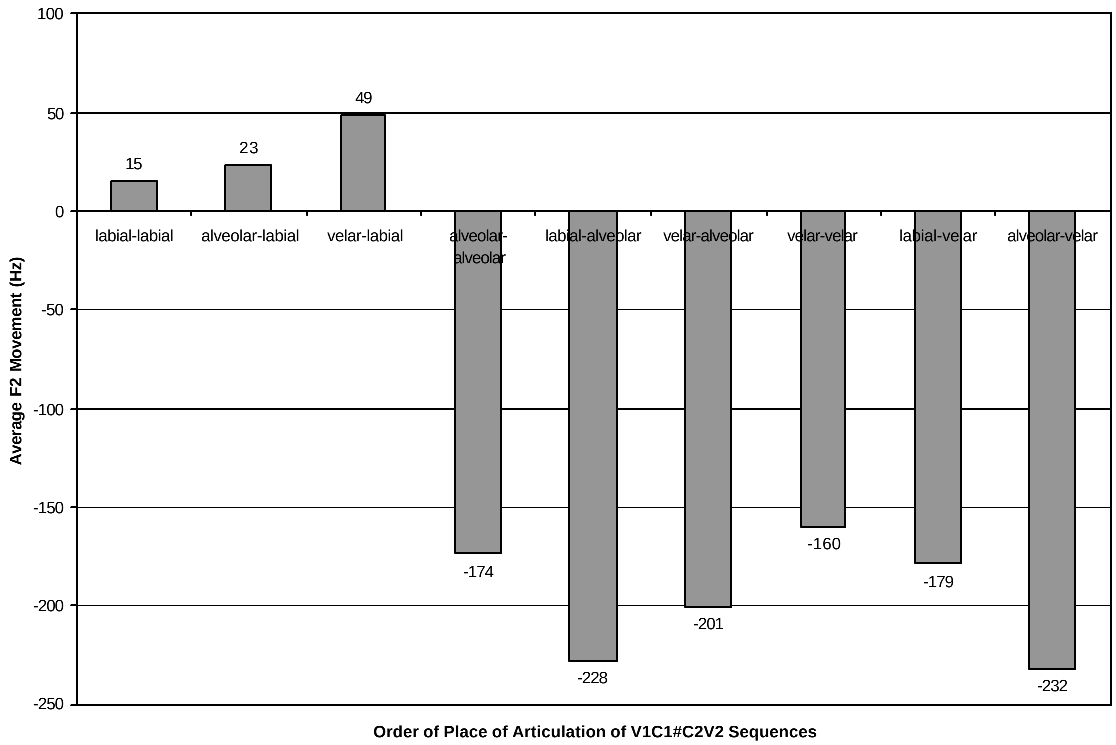

Figure 8. F2 Movement from C2 to V2 vs. Order of Place of Articulation. A plot of the average F2 movement of V2 following the C2 release, measured from the onset of V2 voicing to the middle of the vowel, with respect to the order of place of articulation.

Figure 9 is a plot of the average ratio of the consonant-sequence closure duration to the sum of the closure durations of singleton C1 and C2 with respect to the order of place of articulation. Averages were taken of all four speakers.

Sequences with a velar stop in the C1 position had closure durations roughly equal to the sum of the closure durations of singleton C1 and C2; their ratios were around 1 (see Figure 9). Alveolar-alveolar sequences had a closure duration ratio of .93. As for the rest of the sequences, the closure durations ratios ranged from .85 to .88.

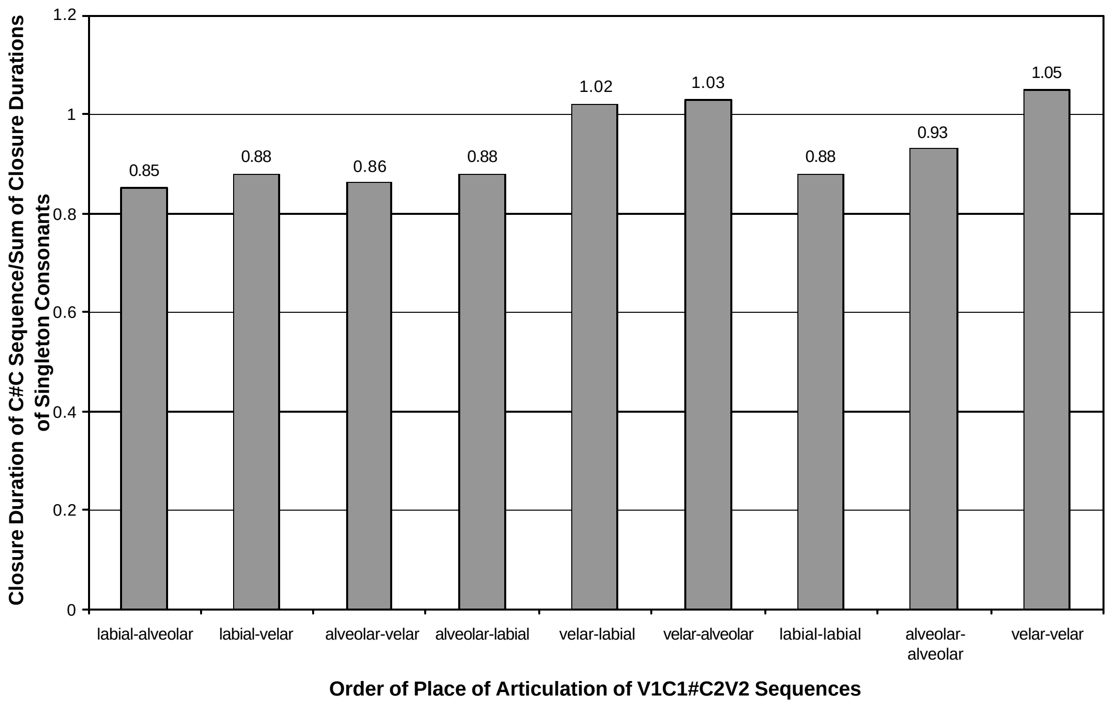

Figure 9. Closure Duration Ratio vs. Order of Place of Articulation. A plot of the average ratio of the consonant-sequence closure duration to the sum of the closure durations of singleton C1 and C2 with respect to the order of place of articulation

Figure 10 is a plot of the average ratio of the consonant-sequence closure duration to the sum of the closure durations of singleton C1 and C2 with respect to the order of voicing. Averages were taken of all four speakers.

The closure duration of sequences with voiced stop consonants in the C1 position were roughly equal to the sum of singleton C1 and C2 closure durations. On the other hand, voiceless-voiceless and voiceless-voiced sequences had lower closure duration ratios at .83 and .85, respectively.

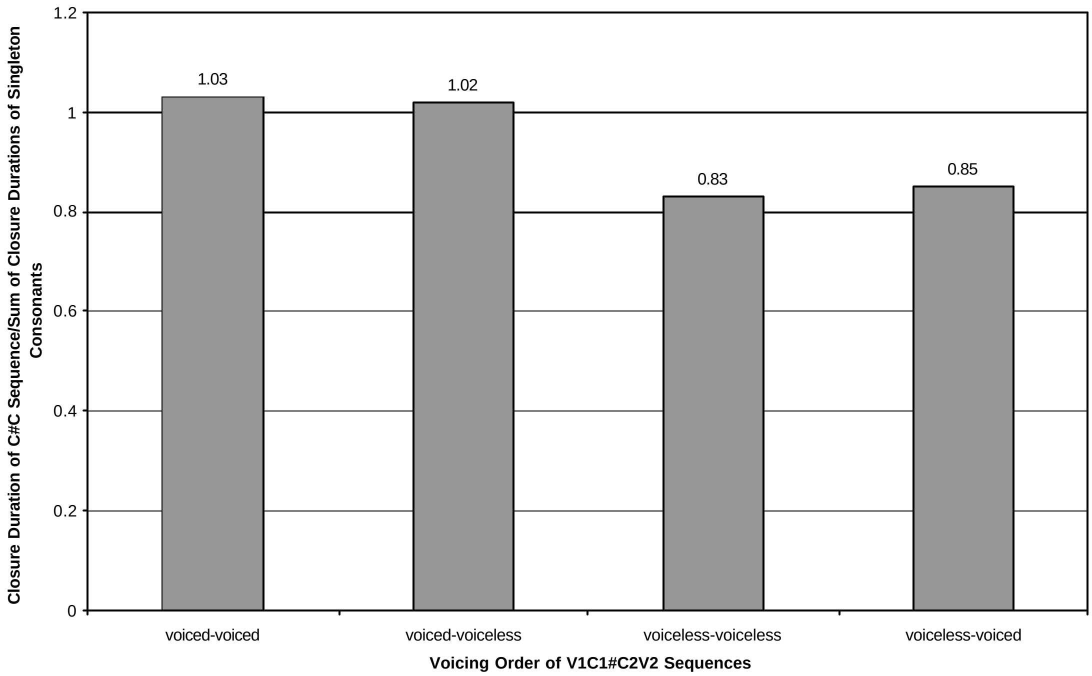

Figure 10. Closure Duration Ratio vs. Voicing Characteristics. A plot of the average ratio of the consonant-sequence closure duration to the sum of the closure durations of singleton C1 and C2 with respect to the order of voicing.

Figure 11 details the ratio of sequences with absent C1 bursts to the total number of sequences with respect to the order of place of articulation. A stop burst was considered absent when there was no acoustical evidence to support its presence.

Labial-alveolar, labial-velar, alveolar-velar, and velar-labial sequences all had C1-burst absence ratios of .25 or less (see Figure). Alveolar-labial and velar-alveolar sequences had at least a .5 C1-burst absence ratio. For sequences where C1 and C2 had the same place of articulation, the absence ratio was at least .63.

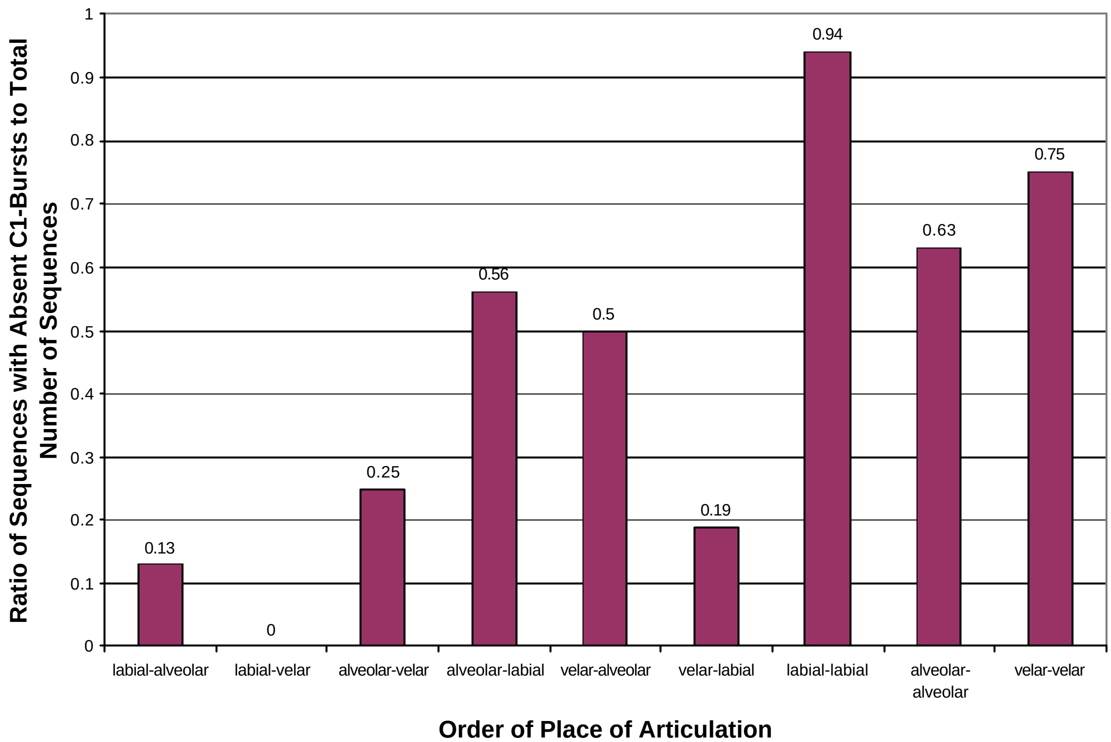

Figure 11. C1 Burst Absence vs. Order of Place of Articulation. A plot of the ratio of sequences with absent C1 bursts to the total number of sequences with respect to the order of place of articulation.

Figure 12 is a plot of the ratio of sequences with absent C1 bursts to the total number of sequences with respect to the stop consonant in the C1 position. Data indicated that a ‘g’ in the C1 position had the highest absence ratio at .71. A ‘d’ in the C1 position had the next highest absence ratio at .5, followed closely by ‘t’ with an absence ratio of .46. A ‘b’ in the C1 position had a release-burst absence ratio of .33, while the ratio for ‘p’ was .38. A C1 ‘k’ had the lowest absence ratio at .25. On average, the C1 burst absence ratios for voiced stops were larger than those for voiceless stops.

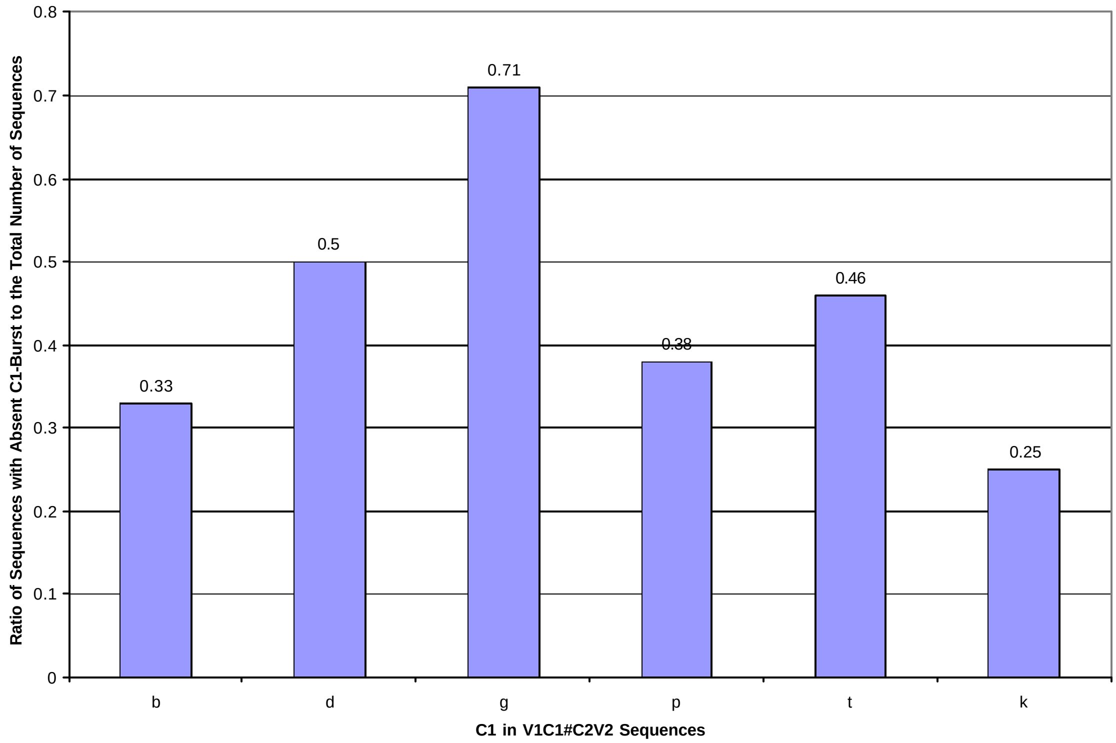

Figure 12. C1 Burst Absence vs. Stop Consonant in C1 Position. A plot of the ratio of sequences with absent C1 bursts to the total number of sequences with respect to the stop consonant in the C1 position.

Figure 13 is a plot of the average H1*-H2* and H1*-A1 values based on data taken from the last four complete pitch periods prior to the C1 closure of V1C1#C2V2 sequences, with respect to the order of voicing. The averages consisted of data from all four speakers.

On average, when a voiced C1 was followed by a voiceless C2 instead of a voiced C2, H1*-H2* was larger by .9dB and H1*-A1 was larger by .86dB. When a voiceless C1 was followed by a voiced C2 instead of a voiceless C2, the average H1*-H2* was smaller by .82dB and the average H1*-A1 was smaller by .08dB.

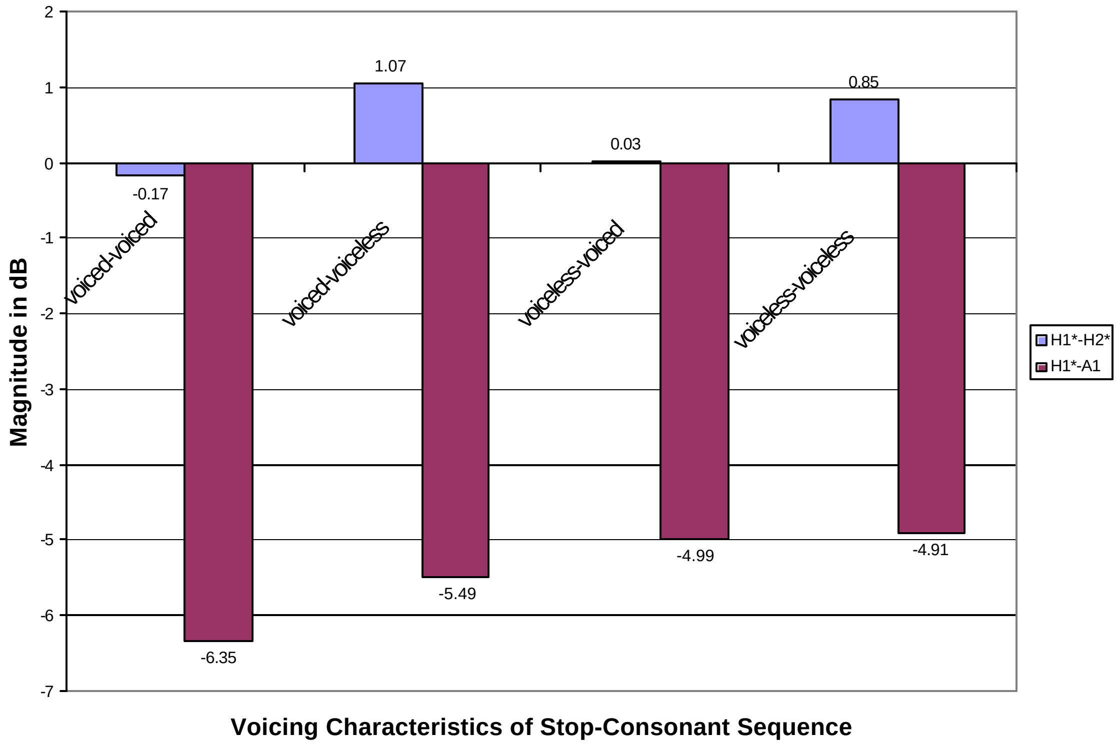

Figure 13. H1*-H2* and H1*-A1 of C1 vs. Voicing Characteristics. A plot of the average H1*-H2* and H1*-A1 values based on data taken from the last four complete pitch periods prior to the C1 closure of V1C1#C2V2 sequences, with respect to the order of voicing.

Figure 14 is a plot of the average H1*-H2* and H1*-A1 values based on data taken from the first four complete pitch periods at the onset of V2 voicing in V1C1#C2V2 sequences, with respect to the order of voicing. The averages consisted of data taken from all four speakers.

On average, when a voiced C2 was preceded by a voiceless C1 rather than a voiced C1, H1*-H2* was larger by .35dB and H1*-A1 was larger by 2.37dB. When a voiceless C2 was preceded by a voiced C1 instead of a voiceless C1, the average H1*-H2* was smaller by .04dB and the average H1*-A1 was smaller by 1.58dB.

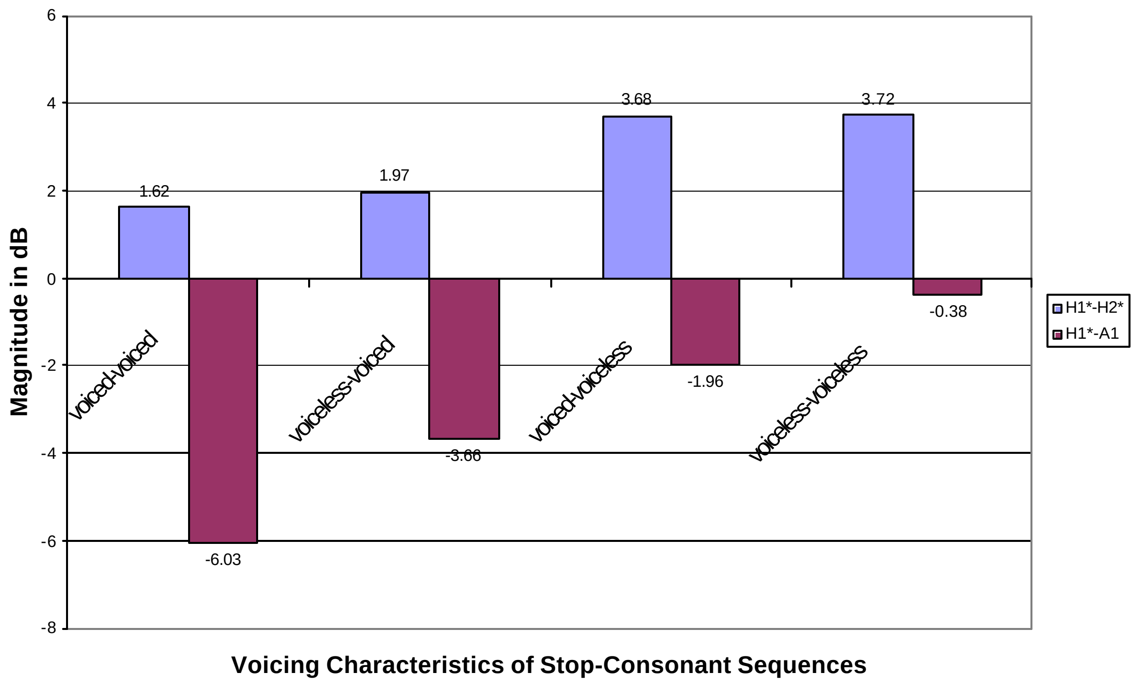

Figure 14. H1*-H2* and H1*-A1 of C2 vs. Voicing Characteristics. A plot of the average H1*-H2* and H1*-A1 values based on data taken from the first four complete pitch periods at the onset of V2 voicing in V1C1#C2V2 sequences, with respect to the order of voicing.

4.3 Acoustic Data Analysis

4.3.1 Trends in F2 Movements into the C1 Closure in V1C1#C2V2 Sequences

The average F2 movements into the C1 closure for V1C1#C2V2 sequences seemed to deviate depending on the place of articulation for C2. For example, the average decreasing movements of F2 into the C1 closures of labial-alveolar and labial-velar sequences were smaller than that of labial-labial sequences. The differences seemed to indicate a possible C2 influence on the F2 movements into C1; since F2 generally increases going into alveolar and velar closures. As for alveolar-labial and velar-labial sequences, the average increases in F2 movements were smaller than those for alveolar-alveolar and velar-velar sequences, respectively. Again, the deviations seemed to indicate a possible influence exerted by the labial, since F2 movements usually decrease going into labial closures.

Compared to the average increasing F2 movement into alveolar-alveolar sequences, that of alveolar-velar sequences was higher. The average F2 movement into velar-alveolar sequences was slightly lower than that into velar-velar sequences. Such deviations agree with the observation that F2 values terminate at a higher frequency going into velar stops than going into alveolar stops. However, the observation was made based only on the data collected in this study.

Furthermore, deviations in F2 movements in sequences with front-to-back order of place of articulation were greater in magnitude, ranging from 16Hz to 39Hz, than in sequences with back-to-front order where deviations ranged from 2Hz to 9Hz.

Overall, the data seemed to reflect C2-dependent deviations in the average F2 movements into the C1 closure, with sequences with front-to-back order of place of articulation exhibiting deviations that were larger in magnitude than those of sequences with back-to-front order. These trends suggest the existence of C2-dependent movements of the tongue body and possibly other oral articulators prior to the C1 closure. However, such trends were found to be statistically insignificant using analysis of variance (ANOVA) with alpha of 0.05 (see Table 4). The resulting p-value must be smaller than alpha in order to reject the null hypothesis.

Table 4. Results ANOVA Tests Examining Differences in F2 Movements into the C1 Closure Between Groups with Different Orders of Place of Articulation

Group 1 (12 sequences per group)	Group 2 (12 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
labial-labial	labial-alveolar	0.49	Cannot Reject
labial-labial	labial-velar	0.22	Cannot Reject
alveolar-alveolar	alveolar-labial	0.88	Cannot Reject
alveolar-alveolar	alveolar-velar	0.50	Cannot Reject
velar-velar	velar-labial	0.81	Cannot Reject
velar-velar	velar-alveolar	0.92	Cannot Reject

4.3.2 Trends in F2 Movements of C2 Transitioning into V2

Special attention was paid to the possible influence of C1 on C2 in the analysis of F2 movements of C2 transitioning into V2. Although the average F2 movements differed for sequences with varying places of C1 articulation, no trends were observed. For example, the average F2 movement of labials transitioning into V2 was actually higher when alveolar and velar stops were in the C1 position instead of labials. The average decreases in F2 movements for labial-alveolar and labial-velar sequences transitioning into V2 were greater than for alveolar-alveolar and velar-velar sequences, respectively. Clearly, the deviations were not C1-dependent and no trends were discovered.

The F2-movement differences between groups of varying orders of place of articulation were further tested using analysis of variance (ANOVA) with alpha of 0.05. These differences were found to be statistically insignificant (see Table 5).

Table 5: Results ANOVA Tests Examining Differences in F2 Movements into V2 Between Groups with Different Orders of Place of Articulation

Group 1 (12 sequences per group)	Group 2 (12 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
labial-labial	alveolar-labial	0.72	Cannot Reject
labial-labial	velar-labial	0.11	Cannot Reject
alveolar-alveolar	labial-alveolar	0.25	Cannot Reject
alveolar-alveolar	velar-alveolar	0.54	Cannot Reject
velar-velar	labial-velar	0.68	Cannot Reject
velar-velar	alveolar-velar	0.09	Cannot Reject

Trends in F2 movements seemed to indicate a possible influence of C2 on C1 but not the other way around. Sequences with front-to-back order of place of articulation were more affected by C2 than sequences with back-to-front order, a finding which agrees with results of past studies (Byrd, 1996; Chitoran, Goldstein, and Byrd, 2002; Zsiga, 1994). However, these trends were found to be statistically insignificant.

4.3.3 Closure Duration and Order of Place of Articulation

On average, sequences with front-to-back order of place of articulation had smaller closure duration ratios than those of sequences with back-to-front order. The average ratio for sequences with front-to-back order was .86 compared to .98 for sequences with back-to-front order. The difference between the two groups was tested using analysis of variance (ANOVA) with alpha of 0.05. The resulting p-value of 0.028 indicated that the difference between the two groups was statistically significant (see Table 6). This finding seemed to indicate that there was more overlap for sequences with front-to-back order of place of articulation. Perhaps articulators started to form the C2 closure prior to the C1 release, maybe even before the C1 closure, therefore shortening the closure duration.

Table 6: ANOVA Test Result Examining Differences in Closure Duration Ratios Between Sequences with Front-to-Back and Back-to-Front Orders of Place of Articulation

Group 1 (48 sequences)	Group 2 (48 sequences)	p-Value	H_0: Means from Group 1 and Group 2 are the same
Front-to-Back	Back-to-Front	0.028	Rejected

4.3.4 Closure Duration and Voicing

On average, sequences beginning with a voiceless stop consonant had smaller closure duration ratios than those of sequences with a voiced C1. The average ratio for sequences with voiceless C1 was .84 compared to 1.02 for sequences with voiced C1. In order to test the significance of the difference, analysis of variance (ANOVA) was performed with an alpha of 0.05. The resulting p-value of $6.35*10^{-5}$ was much smaller than alpha; therefore, there was very strong evidence rejecting the null hypothesis (see Table 7). This finding seemed to indicate that there was more overlap in sequences with voiceless stops in the C1 position.

Table 7: ANOVA Test Result Examining Differences in Closure Duration Ratios Between Sequences with Voiced C1 and Sequences with Voiceless C1

Group 1 (72 sequences)	Group 2 (72 sequences)	p-Value	H_0: Means from Group 1 and Group 2 are the same
Voiced C1	Voiceless C1	$6.35*10^{-5}$	Rejected

4.3.5 Absence of C1 Burst

C1-burst absence rates were very high for sequences where C1 and C2 had the same place of articulation, with ratios ranging from .68 to .94. Sequences with back-to-front order had absence ratios ranging from .19 to .56. Surprisingly, sequences with front-to-back order of place of articulation had the lowest C1-burst absence rates, ranging from 0 to .25. This finding was unexpected because sequences with front-to-back order of place of articulation were thought to have more gestural overlap than sequences with back-to-front order. The low absence ratios may suggest that the overlapping of gestures in such sequences did not extend so far as to obliterate the C1 burst.

One trend emerged from comparing C1-burst absence ratios among the different stop consonants. On average, the absence ratios were larger for voiced stops in the C1 position than for voiceless stops. The average absence ratio for sequences with voiced C1 was 0.51, compared to 0.36 for sequences with voiceless C1. In the previous section, greater overlap for sequences with voiceless C1 was suggested because of their smaller closure duration ratios compared to sequences with voiced C1. This finding seemed to suggest that although such sequences may have greater overlap, the extent of the overlap did not go so far as to obliterate the C1 burst.

Unfortunately, the differences in C1-burst absence ratios were not statistically analyzed because of the small sample size; there were only three voiced and three voiceless stop consonants.

Finally, no conclusive evidence was found indicating a definite link between the presence of enhancing gestures and defining gestures' susceptibility to obliteration. For example, the absence ratio of t-bursts is slightly larger, at .46, than that of p-bursts, at 0.38. The higher rate of absence for the t-burst, a defining gesture, is perhaps explained by the presence of the fronting of the tongue body, an enhancing gesture. However, significance of the ratio differences is unclear; statistical analysis was not performed because of the small sample size.

4.3.6 Order of Voicing

The average values of both H1*-H2* and H1*-A1, measured from the last four complete pitch periods prior to the C1 closure, were larger when a voiced C1 was followed by a voiceless C2 instead of a voiced C2. Both measurements were larger by 0.9dB and correlated to a larger open-quotient and bigger acoustic loss at the glottis. However, the differences between the voiced-voiced and voiced-voiceless groups proved to be statistically insignificant using analysis of variance (ANOVA) with alpha of 0.05 (see Table 8).

The average values of both H1*-H2* and H1*-A1, measured from the last four complete pitch periods prior to the C1 closure, were smaller when a voiceless C1 was followed by a voiced C2 instead of a voiceless C2. H1*-H2* was smaller by 0.8dB and H1*-A1 was smaller by 0.08dB; the differences correlated to a smaller open-quotient and slightly less acoustic loss at the glottis, respectively, for voiceless-voiced sequences compared to voiceless-voiceless sequences. However, the differences between the voiceless-voiced and voiceless-voiceless groups proved to be statistically insignificant using analysis of variance (ANOVA) with alpha of 0.05 (see Table 9). The high p-values indicated that the voicing of C1 did not seem to be affected by that of C2.

Table 8: ANOVA Test Results Examining Differences in Voicing Characteristics of C1 Between Voiced-Voiced and Voiced-Voiceless Sequences

Group 1 (36 sequences per group)	Group 2 (36 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* prior to C1 for **voiced**-voiced	H1*-H2* prior to C1 for **voiced**-voiceless	0.43	Cannot Reject
H1*-A1 prior to C1 for **voiced**-voiced	H1*-A1 prior to C1 for **voiced**-voiceless	0.58	Cannot Reject

Table 9: ANOVA Test Results Examining Differences in Voicing Characteristics of C1 Between Voiceless-Voiced and Voiceless-Voiceless Sequences

Group 1 (36 sequences per group)	Group 2 (36 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* prior to C1 for **voiceless**-voiced	H1*-H2* prior to C1 for **voiceless**-voiceless	0.66	Cannot Reject
H1*-A1 prior to C1 for **voiceless**-voiced	H1*-A1 prior to C1 for **voiceless**-voiceless	0.96	Cannot Reject

The average values of both H1*-H2* and H1*-A1, measured from the first four complete pitch periods at the onset of V2 voicing, were larger when a voiced C2 is preceded by a voiceless C1 instead of a voiced C1; the average H1*-H2* value was larger by .35dB and H1*-A1 was larger by 2.37dB. Such differences indicated a larger open-quotient and bigger acoustic loss at the glottis, respectively, for C2 in voiceless-voiced sequences than in voiced-voiced sequences. However, the differences between the voiceless-voiced and voiced-voiced groups proved to be statistically insignificant using analysis of variance (ANOVA) with alpha of 0.05 (see Table 10).

Table 10: ANOVA Test Results Examining Differences in Voicing Characteristics of C2 between Voiced-Voiced and Voiceless-Voiced Sequences

Group 1 (36 sequences per group)	Group 2 (36 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* at V2 onset for voiced-**voiced**	H1*-H2* at V2 onset for voiceless-**voiced**	0.76	Cannot Reject
H1*-A1 at V2 onset for voiced-**voiced**	H1*-A1 at V2 onset for voiceless-**voiced**	0.15	Cannot Reject

The average values of both H1*-H2* and H1*-A1, measured from the first four complete pitch periods at the onset of V2 voicing, were smaller when a voiceless C2 is preceded by a voiced C1 instead of a voiceless C1; the average H1*-H2* value was smaller by .04dB and H1*-A1 was smaller by 1.58dB. Such differences indicated a smaller open-quotient and smaller acoustic loss at the glottis, respectively, for C2 in voiced-voiceless sequences than in voiceless-voiceless sequences. However, the differences between the two groups proved to be statistically insignificant using analysis of variance (ANOVA) with alpha of

0.05 (see Table 11).

Table 11: ANOVA Test Results Examining Differences in Voicing Characteristics of C2 between Voiced-Voiceless and Voiceless-Voiceless Sequences

Group 1 (36 sequences per group)	Group 2 (36 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* at V2 onset for voiced-**voiceless**	H1*-H2* at V2 onset for voiceless-**voiceless**	0.96	Cannot Reject
H1*-A1 at V2 onset for voiced-**voiceless**	H1*-A1 at V2 onset for voiceless-**voiceless**	0.23	Cannot Reject

Relatively small p-values of 0.15 and 0.23 resulted from ANOVA tests examining differences in H1*-A1 data between voiced-voiced sequences and voiceless-voiced sequences, as well as between voiced-voiceless and voiceless-voiceless sequences, respectively. This prompted a closer look at the voicing characteristics data.

One important detail to mention is that H1*-H2* and H1*-A1 values differ significantly between male and female speakers. More specifically, female speakers usually have larger open quotient and bigger acoustical loss at the glottis than male speakers (Hanson, 1999). Data collected in this study also reflect such differences (see Appendix G). In order to take a closer look at the voicing characteristics data, the next two graphs reflect data separated by gender.

Figure 15 is a plot of the average H1*-H2* and H1*-A1 values based on data taken from the first four complete pitch periods at the onset of V2 voicing in V1C1#C2V2 sequences, with respect to the order of voicing. Averages consisted of data taken from female speakers.

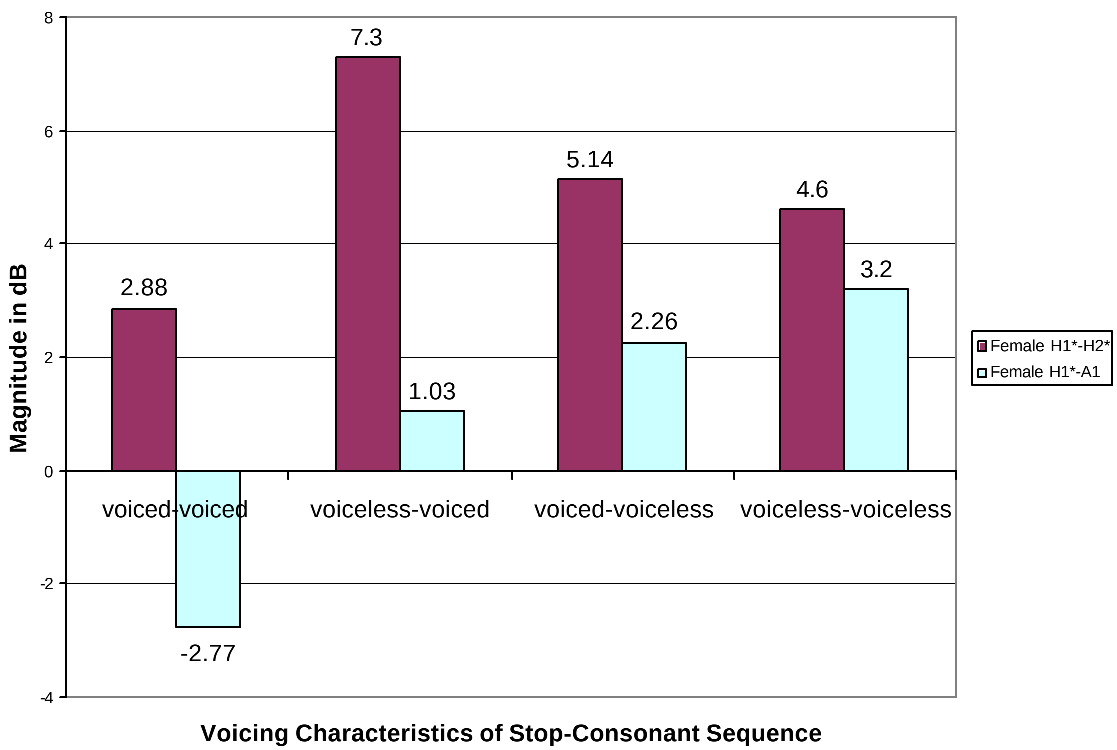

Figure 15. H1*-H2* and H1*-A1 of C2 vs. Voicing Characteristics of Sequences by Female Speakers. A plot of the average H1*-H2* and H1*-A1 values based on data taken from the first four complete pitch periods at the onset of V2 voicing in V1C1#C2V2 sequences, with respect to the order of voicing. Averages consisted of data taken from female speakers.

On average, when a voiced C2 was preceded by a voiceless C1 rather than a voiced C1, H1*-H2* was larger by 4.42dB and H1*-A1 was larger by3.8dB for female speakers. When a voiceless C2 was preceded by a voiced C1 instead of a voiceless C1, the average H1*-H2* was actually larger by .54dB and the average H1*-A1 was smaller by .96dB for female speakers.

These differences between the groups of sequences were tested using analysis of variance with alpha of 0.05 (see Table 12).

Table 12: ANOVA Test Results Examining Differences in Voicing Characteristics of C2 between Voiced-Voiced and Voiceless-Voiced Sequences Uttered by Female Speakers

Group 1 (18 sequences per group)	Group 2 (18 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* at V2 onset for voiced-**voiced**	H1*-H2* at V2 onset for voiceless-**voiced**	0.0047	Rejected
H1*-A1 at V2 onset for voiced-**voiced**	H1*-A1 at V2 onset for voiceless-**voiced**	0.018	Rejected

The analysis of variance indicated strong evidence for the rejection of the null hypothesis; differences between H1*-H2* and H1*-A1 values, measured at the first four complete pitch periods at V2 onset, between voiced-voiced and voiceless-voiced sequences were statistically significant. This finding indicated that for female speakers, the voicing characteristics of the voiced C2 were affected by those of the voiceless C1. More specifically, the voiced C2 was less "voiced" when it was preceded by a voiceless C1 than by a voiced C1.

Table 13 details the ANOVA test results examining differences in voicing characteristics of C2 between voiced-voiceless and voiceless-voiceless sequences uttered by female speakers. No statistical significance to the differences was found.

Table 13: ANOVA Test Results Examining Differences in Voicing Characteristics of C2 between Voiced-Voiceless and Voiceless-Voiceless Sequences Uttered by Female Speakers

Group 1 (18 sequences per group)	Group 2 (18 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* at V2 onset for voiced-**voiceless**	H1*-H2* at V2 onset for voiceless-**voiceless**	0.85	Cannot Reject
H1*-A1 at V2 onset for voiced-**voiceless**	H1*-A1 at V2 onset for voiceless-**voiceless**	0.5	Cannot Reject

Figure 16 is a plot of the average H1*-H2* and H1*-A1 values based on data taken from the first four complete pitch periods at the onset of V2 voicing in V1C1#C2V2 sequences, with respect to the order of voicing. Averages consisted of data taken from male speakers.

On average, when a voiced C2 was preceded by a voiceless C1 rather than a voiced C1, H1*-H2* was actually smaller by 1.97dB and H1*-A1 was larger by .95dB for male speakers. When a voiceless C2 was preceded by a voiced C1 instead of a voiceless C1, the average H1*-H2* was smaller by .22dB and the average H1*-A1 was smaller by 2.3dB for male speakers.

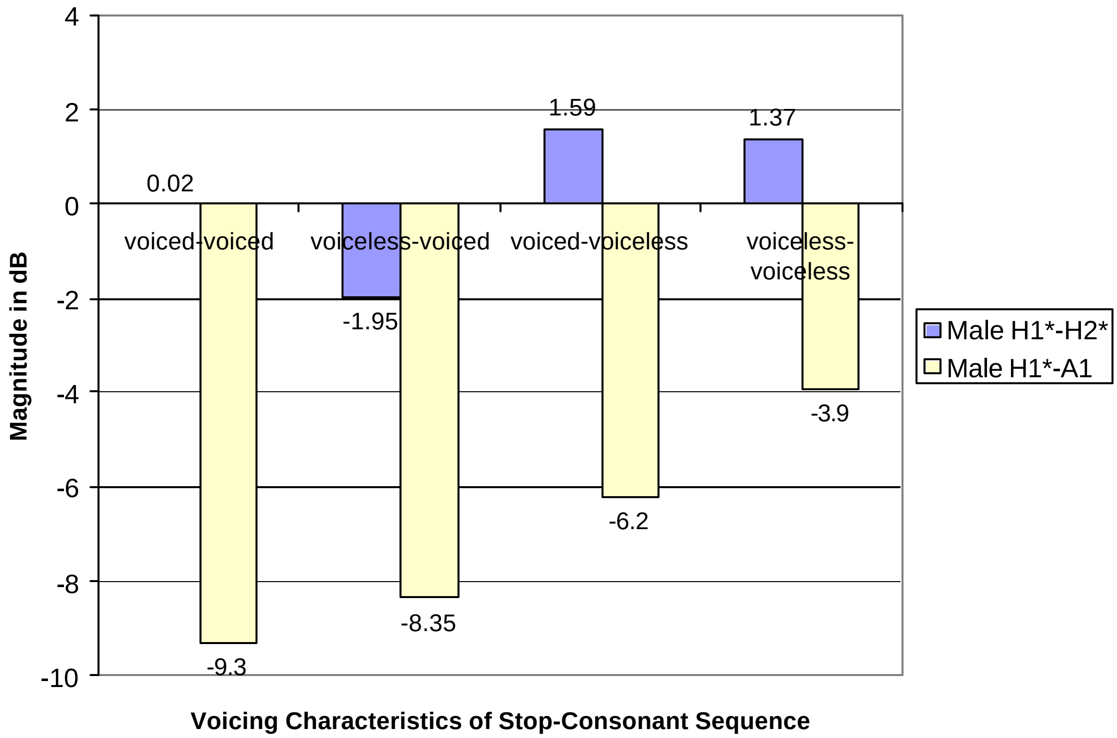

Figure 16. H1*-H2* and H1*-A1 of C2 vs. Voicing Characteristics of Sequences by Male Speakers. A plot of the average H1*-H2* and H1*-A1 values based on data taken from the first four complete pitch periods at the onset of V2 voicing in V1C1#C2V2 sequences, with respect to the order of voicing. Averages consisted of data taken from male speakers.

These differences between the groups of sequences were tested using analysis of variance with alpha of 0.05 (see Tables 14 and 15).

The large p-values indicated that the differences in H1*-H2* and H1*-A1 data, measured from the first four complete pitch periods of V2, between voiced-voiced and voiceless-voiced sequences were not statistically significant.

Table 14: ANOVA Test Results Examining Differences in Voicing Characteristics of C2 between Voiced-Voiced and Voiceless-Voiced Sequences Uttered by Male Speakers

Group 1 (18 sequences per group)	Group 2 (18 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* at V2 onset for voiced-**voiced**	H1*-H2* at V2 onset for voiceless-**voiced**	0.22	Cannot Reject
H1*-A1 at V2 onset for voiced-**voiced**	H1*-A1 at V2 onset for voiceless-**voiced**	0.66	Cannot Reject

Table 15 details the ANOVA test results examining differences in voicing characteristics of C2 between voiced-voiceless and voiceless-voiceless sequences uttered by male speakers. No statistical significance to the differences was found.

Table 15: ANOVA Test Results Examining Differences in Voicing Characteristics of C2 between Voiced-Voiceless and Voiceless-Voiceless Sequences Uttered by Male Speakers

Group 1 (18 sequences per group)	Group 2 (18 sequences per group)	p-Value	H_0: Means from Group 1 and Group 2 are the same
H1*-H2* at V2 onset for voiced-**voiceless**	H1*-H2* at V2 onset for voiceless-**voiceless**	0.89	Cannot Reject
H1*-A1 at V2 onset for voiced-**voiceless**	H1*-A1 at V2 onset for voiceless-**voiceless**	0.08	Cannot Reject

The analysis of voicing characteristics of stop-consonant sequences indicated no evidence for the influence of C1 by C2. Although there was evidence indicating that the voicing of

C1 seems to affect the voicing of C2, such effects did not always occur (see Appendix G). For example, the only statistically significant data suggesting C1's influence on the voicing of C2 were shown in voiced-C2 sequences uttered by female speakers. A closer look at the H1*-H2* and H1*-A1 sorted by speaker reveal that the extent of C1's influence varies greatly from individual to individual (see Appendix G).

4.4 General Gestural Timing Estimates

This section discusses what the results of acoustic data analysis mean in terms of gestural timing estimates.

The analysis of acoustic data disclosed several interesting findings. First, F2 data revealed C2-dependent deviations in the movements of F2 into the C1 closure. Such deviations were larger in magnitude for sequences with front-to-back order of place of articulation than for sequences with back-to-front order. However, these trends were found to be statistically insignificant. Nevertheless, the trends agree with similar findings in past studies examining formant movements in stop-consonant sequences (Zsiga, 1994).

How do these trends translate to general gestural timing estimates? The existence of C2-dependent deviations in F2 movements into the C1 closure suggests the presence of C2-dependent movements of the tongue body and of possibly other oral-articulators into the C1-closure. Greater deviations in sequences with front-to-back order of place of articulation suggest more overlapping than in sequences with back-to-front order.

Data also indicated smaller closure duration ratios for sequences with front-to-back order of place of articulation than those with back-to-front order. This suggests more overlap in sequences with front-to-back order of place of articulation than in sequences with back-to-front order. Very possibly, the formation of the C2 closure by the appropriate articulators begins prior to the release of C1, perhaps even before the C1 closure. However, the absence rates of C1-bursts were lower for sequences with front-to-back order of place of articulation than for those with back-to-front order. The low absence rates of such sequences suggest that the formation of the C2 closure does not occur early enough to obliterate the C1 release.

In addition, data indicated smaller closure duration ratios for voiceless-C1 sequences than for voiced-C1 sequences. However, the C1-burst absence rate was lower for voiceless-C1 sequences than for voiced-C1 sequences. It is unclear why closure duration ratios and C1-burst absence rates differ between voiceless-C1 and voiced-C1 sequences.

Finally, voicing data indicated possible influence of C1's voicing on the voicing of C2, although such effects varied from individual to individual. This finding suggests that the actions of the laryngeal articulators, such as glottal opening adjustments and vocal-fold stiffness adjustments, for the production of C1 may overlap with those for the production of C2. For example, if C1 is voiceless, the area of the glottis may still be a bit spread and the vocal folds a bit stiff going into a voiced C2.

4.5 Derivation of Possible Principles

Based on the analysis of acoustic data and the inferred gestural timing estimates, two principles were derived.

Principle 1: There exist C2-dependent movements of the tongue body and of possibly other oral-articulators into the C1-closure. The formation of the C2 closure by the appropriate articulators is likely to start to occur before the release of C1, perhaps even prior to the C1 closure, for stop-consonant sequences. The extent of such overlapping is greater in sequences with front-to-back order of place of articulation than in back-to-front order, however, overlapping does not extend so far as to obliterate the C1 release burst.

Principle 2: The actions of the laryngeal articulators for C1 production, such as glottal opening adjustments and vocal-fold stiffness adjustments, may overlap with those for C2 production.

The derived principles, along with the inferred gestural estimates, were verified and refined in the next phase of this study.

Chapter 5: Phase II

In phase two of this study, both the general gestural estimates and the derived principles were verified and refined through quasi-articulatory synthesis using HLsyn. Consonant sequences were first generated using sets of specific gestural timing estimates inferred from acoustic data. The synthetic utterances were acoustically and perceptually compared to the actual utterances in order to verify and refine the articulatory timing estimates; the resulting gestural timing estimates were used to cross-check and refine the general estimates inferred in phase one. Next, in order to verify the revised principles, synthetic sequences with varying degrees of overlap were generated. Perception tests composed of these synthetic consonant sequences were administered. The results of the tests were used to verify the validity of each principle and the extent of its applicability.

5.1 Phase II Methodology

5.1.1 Verification and Refinement of Gestural Timing Estimates

The first step in this phase of the study was to verify and refine the general gestural timing estimates made from acoustical analysis. The process involved making specific gestural timing estimates for naturally-spoken stop-consonant sequences and transferring these estimates to HLsyn to generate synthetic counterparts to the spoken utterances. The synthetic sequences were generated to be as acoustically identical to the actual sequences as possible. Through the process, the general gestural timing estimates were refined so that they were in agreement with the specific gestural estimates.

Gestural timing estimates were made for 12 singleton stop consonants (VCV) and 12 VC#CV sequences spoken by a male speaker, using the concepts detailed in Section 3.2. Sequences with voiced and voiceless alveolar stops in the C1 position were chosen; such sequences include both back-to-front and front-to-back orders of articulation.

The gestural timing estimates, along with the corresponding formant data, were transferred directly to HLsyn to generate synthetic sequences (refer to Section 3.3 for HLsyn manipulation). Appropriate adjustments were made to make the synthetic sequences acoustically and perceptually similar to the spoken utterances (see disk attachment of HLsyn files).

The resulting gestural timing estimates were used to cross-check the general gestural timing estimates inferred from phase one. The process found no discrepancies significant enough to change the general gestural timing estimates made in phase one. Therefore, the derived principles were also not modified.

5.1.2 Verification and Refinement of Derived Principles

The validity of the derived principles and the extent of their applicability were verified through perception testing. The tests were composed of synthetic sequences with varying degrees of overlap. Four sets of such sequences were generated; each set was aimed at the verification and refinement of an aspect of a particular principle. The following describes those sets of synthesized sequences.

The first set of synthetic sequences consisted of 15 d#d sequences with the F2 value just prior to the C1 closure varying in 50Hz steps from 1000Hz to 1700Hz. F2 at middle of V1 was kept constant at 1200Hz for all sequences (see Appendix for articulatory parameters of the sequences). The control sequence in this group had a F2 value of 1400Hz just prior to the C1 closure.

This set of d#d-sequences was aimed at modeling the potential deviations in the F2 movements into the C1 closure. The delta of the F2 movement into a d-closure was gradually decreased to that of a b-closure in one extreme, and gradually increased to that of a g-closure in the other extreme. The perception accuracies of these sequences should reveal a possible window of allowable deviation that did not negatively affect the perception of the sequence. Since the F2 movements reflect the oral gestures, the results were used to verify and refine Principle 1. For example, if any slight deviation in the F2 movement, even for sequences with front-to-back order of place of articulation, significantly decreased perception accuracy, then the validity of Principle 1 would be questioned. In addition, differences in perception accuracies of sequences with front-to-back F2 movement deviations (toward that of 'g') and sequences with back-to-front F2 movement deviations (toward that of 'b') would further substantiate or question Principle 1's statement of greater overlap for front-to-back sequences.

The second set consisted of d#b, d#p, d#g, d#k, t#b, t#p, t#g, t#k sequences with C1 bursts and their counterparts without C1 bursts. Sequences with absent C1 bursts were synthesized by forming the C2 closure before the release of C1 closure. The perception accuracies of these sequences should provide information on the importance of the C1-burst in the perception of the sequences. The results were used to verify and refine Principle 1's statement about the non-obliteration of the C1-burst in the gestural overlap of stop consonant sequences.

The third set consisted of a series of 11 d#p sequences with changing voicing characteristics of 'p' to eventually sound like 'b' (see Table 16). The first sequence in the series was a normal d#p sequence. With each successive sequence, the voicing characteristics of 'p' were modified to be acoustically closer and closer to that of a 'b'. The last sequence of the series had the voicing characteristics of a normal d#b sequence. The voice bar following the voiced-C1 closure was not a changing factor. This set modeled the potential effects of a voiced C1 on the voicing characteristics of a voiceless C2. More specifically, perception accuracies of these sequences should answer the question: When preceded by a 'd', how "voiced" can the 'p' be and still be accurately perceived?

The last set consisted of a series of 11 t#b sequences with changing voicing characteristics of 'b' to eventually sound like 'p' (see Table 17). The first sequence in the series was a normal t#b sequence. With each successive sequence, the voicing characteristics of 'b' were modified to be acoustically closer and closer to that of a 'p'. The last sequence of the series had the voicing characteristics of a normal t#p sequence. This set modeled the potential effects of a voiceless C1 on the voicing characteristics of a voiced C2. More specifically, perception accuracies of these sequences should answer the question: When preceded by a ‘t’, how “voiceless” can the ‘b’ be and still be accurately perceived?

Table 16: D#P Sequences with Changing HLsyn Voicing Parameters Around the P-Release

	HLsyn Voicing Parameters Around the ‘p’-Release
d#p sequence 1 (normal d#p sequence)	ag=50, ue=-100, dc=-70
d#p sequence 2	ag=45.4, ue=-90, dc= -56
d#p sequence 3	ag=40.8, ue=-80, dc=-42
d#p sequence 4	ag=36.2, ue=-70, dc=-28
d#p sequence 5	ag=31.6, ue=-60, dc=-14
d#p sequence 6	ag=27, ue=-50, dc=0
d#p sequence 7	ag=22.4, ue=-40, dc=14
d#p sequence 8	ag=17.8, ue=-30, dc=28
d#p sequence 9	ag=13.2, ue=-20, dc=42
d#p sequence 10	ag=8.6, ue=-10, dc=56
d#p sequence 11 (acoustically similar to d#b)	ag=4, ue=0, dc=70

The perception accuracies of the last two sets of sequences should provide a possible window of allowable laryngeal gestural overlap, which would be used to refine Principle 2. In addition, the validity of Principle 2 would also be verified. For example, a non-existent window of allowable overlap would cast doubt on whether overlap of laryngeal gestures exists.

5.1.3 Perception Tests

Two perception tests were administered. The first test was composed of the four sets of synthesized sequences with varying degrees of overlap; listeners were asked to indicate the consonant-sequence they heard. The second test was composed of spoken sequences and their synthetic counterparts; listeners were asked to rate how natural the synthetic sequences were when compared to the spoken ones. Ten subjects participated in the series of two perception tests. The next two paragraphs detail the composition and administration of each perception test.

Table 17: T#B Sequences with Changing HLsyn Voicing Parameters Around the B-Release

	HLsyn Voicing Parameters Around the ‘b’-Release
t#b (normal t#b sequence)	ag=4, ue=0, dc=70
t#b	ag=7, ue=-10, dc=56
t#b	ag=10, ue=-20, dc=42
t#b	ag=13, ue=-30, dc=28
t#b	ag=16, ue=-40, dc=14
t#b	ag=19, ue=-50, dc=0
t#b	ag=22, ue=-60, dc=-14
t#b	ag=25, ue=-70, dc=-28
t#b	ag= 28, ue=-80, dc=-42
t#b	ag= 30, ue=-90, dc=-56
t#b (acoustically similar to t#b)	ag=33, ue=-100, dc=-70

The first perception test was composed of the four sets of synthesized sequences with varying degrees of overlap detailed in section 5.1.2. The 52 uniquely synthesized sequences were each repeated 5 times for a total of 260 sequences. Each sequence was composed of two repeats of a single synthesized VC#CV utterance. There was a 500ms pause in between repeats. Each synthesized sequence was separated by a 2s pause.

The sequences were given non-descriptive names and randomly mixed in the perception test play-list. Subjects were asked to follow the instructions detailed in Appendix A.

After taking the first perception test, a second test was administered. The second perception test was composed of spoken d#b, d#p, d#g, d#k, t#b, t#p, t#g, and t#k sequences and their synthetic counterparts. Each sequence was repeated twice, with a 500ms in between. Each spoken sequence was presented first, followed by a 2-second pause; next, its synthetic counterpart was presented. The listener was asked to rate how natural each synthetic sequence sounded compared to its natural counterpart (see Appendix B). The naturally produced VC#CV sequences were extracted from the spoken utterances. The goal of this test was to have an idea of the quality of the synthesized sequences.

5.2 Perception Test Results

Figure 17 contains the accuracy ratios of the d#d sequences as the F2 value just prior to the C1 closure changes from 1000Hz to 1700Hz. The F2 value at mid-V1 was constant at 1200Hz for all sequences. The control sequence had a F2 of 1400Hz just prior to the C1 closure.

Data indicated that when F2 was at 1400Hz or higher, the sequences were accurately perceived about 90% of the time. However, as the F2 value prior to the C1 closure got smaller than 1400Hz, perception became less accurate. Subjects began to hear b#d and d#b sequences. When F2 fell below 1200Hz, perception accuracy was about 40%, with about equal numbers d#b and b#d sequences perceived.

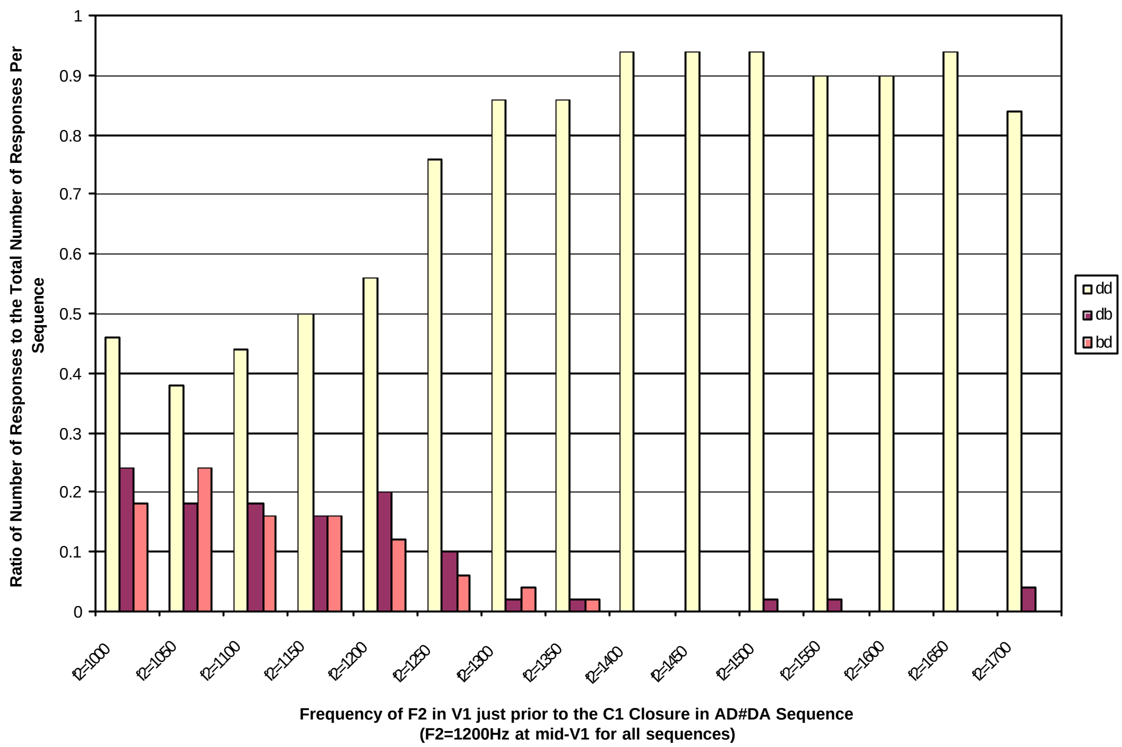

Figure 17. Perception Accuracy vs. Final F2 Prior to C1 Closure. A plot of the accuracy ratios of the d#d sequences as the F2 value just prior to the C1 closure changes from 1000Hz to 1700Hz.

Figure 18 details the effects of the absent C1 release bursts on the perception of V1C1#C2V2 sequences. Each V1C1#C2V2 sequence had a synthesized version where the C1 burst was present and a version where the C1 burst was absent. The perception accuracies of the sequences were plotted on the graph.

Data indicated that perception accuracy was lowered when C1 bursts were absent. Perception accuracy ratios for d-sequences with C1 bursts ranged from 0.82 to 0.92, while those for d-sequences without C1 bursts ranged from 0.62 to 0.84. Perception accuracy ratios for t-sequences with C1 bursts ranged from 0.8 to 0.86, whereas those for t-sequences without bursts ranged from 0.42 to 0.68.

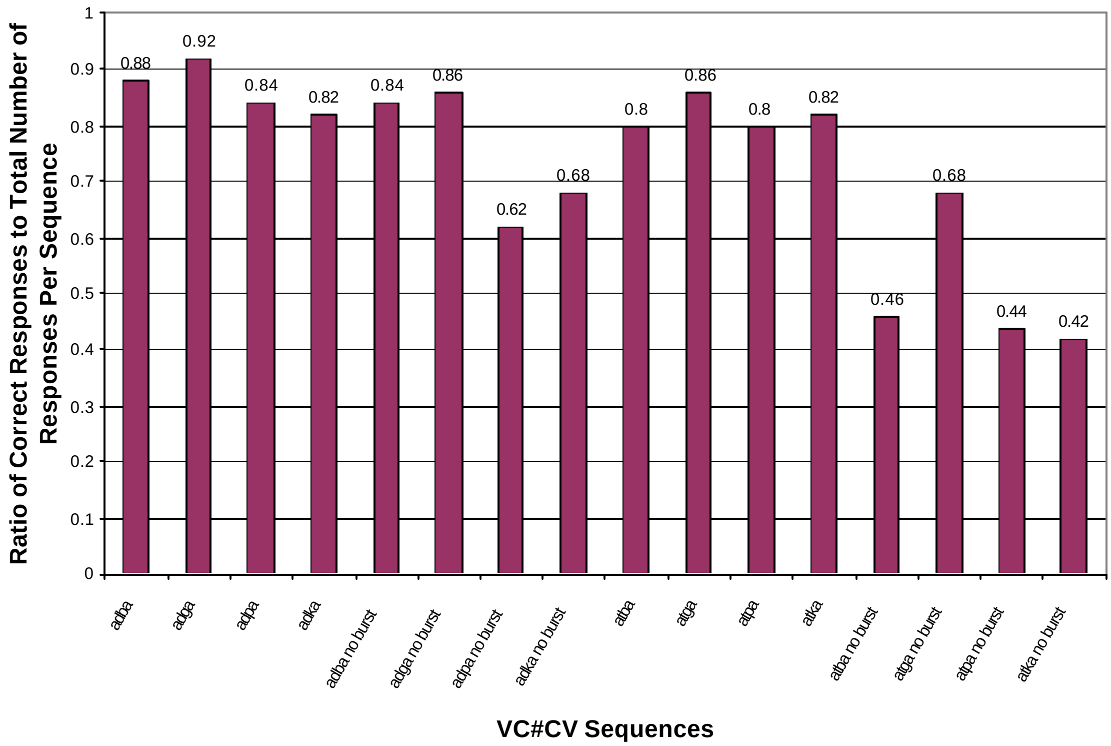

Figure 18. Perception Accuracy vs. Presence of C1-Burst. A Plot of the perception accuracy ratios of V1C1#C2V2 sequences with and without C1 bursts.

Figure 19 is a plot of the perception accuracy ratios of the series of d#p sequences, as the ‘p’ in each successive sequence gets acoustically closer and closer to a ‘b’. Perception accuracy ratios for the first several sequences were not significantly affected by the changes in voicing parameters, with ratios centered at 0.75. However, starting with the eighth sequence (ag=17.8, ue=-30, dc=28) in the series, the perception accuracy ratio dropped below 0.6; the number of d#b sequences perceived increased significantly. The perception accuracy ratios continued to drop and the ratio of perceived d#b sequences continued to increase as the last few sequences in the series got acoustically closer and closer to a d#b sequence. The last d#p sequence in the series only had a perception accuracy ratio of 0.02 and a d#b perception ratio of 0.86.

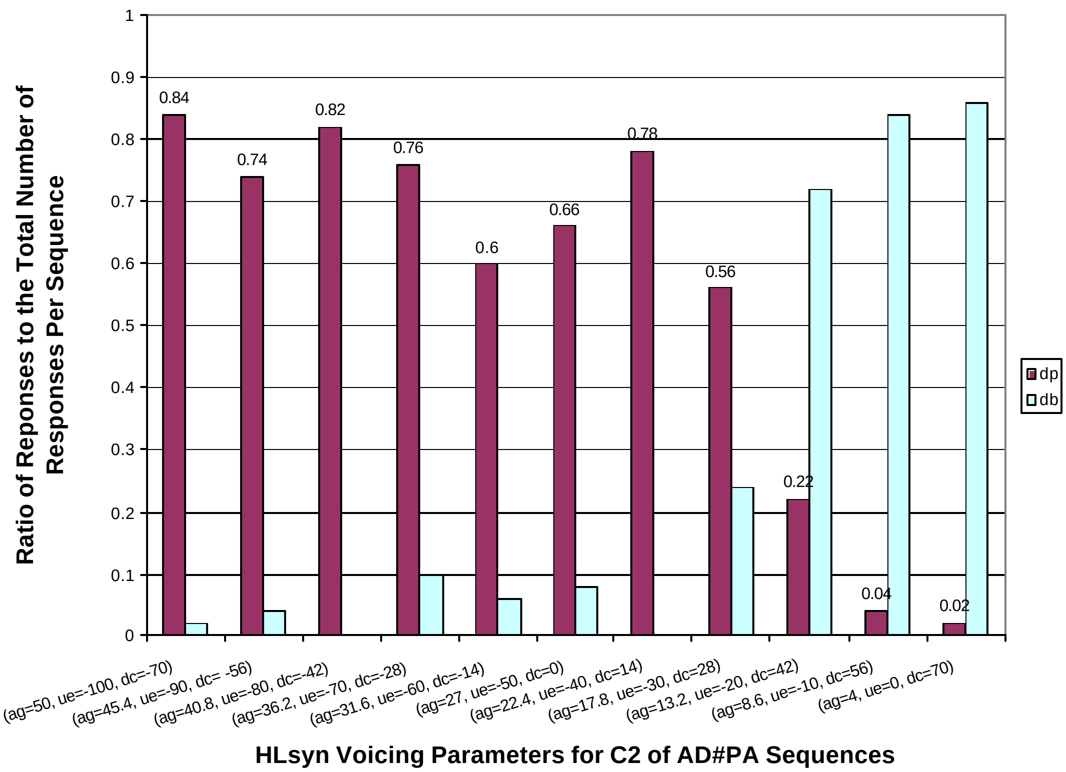

Figure 19. Perception Accuracy vs. Changing HLsyn Voicing Parameters of D#P. A plot of the perception accuracy ratios of the series of d#p sequences, as the 'p' in each successive sequence gets acoustically closer and closer to a 'b'.

Figure 20 is a plot of the perception accuracy ratios of the series of t#b sequences, as the 'b' in each successive sequence gets acoustically closer and closer to a 'p'. The first t#b sequence in the series had a perception accuracy ratio of 0.8. As the glottis spread and the vocal folds became less slack, the perception accuracy ratio fell to 0.66 for the second t#b ratio. As the voicing parameters got closer to those for the voiceless labial, perception accuracy ratios continued to decrease as the ratio of perceived t#p sequences increased. The last t#b sequence in the series had a perception accuracy ratio of 0.1 and a t#p perception ratio of 0.65.

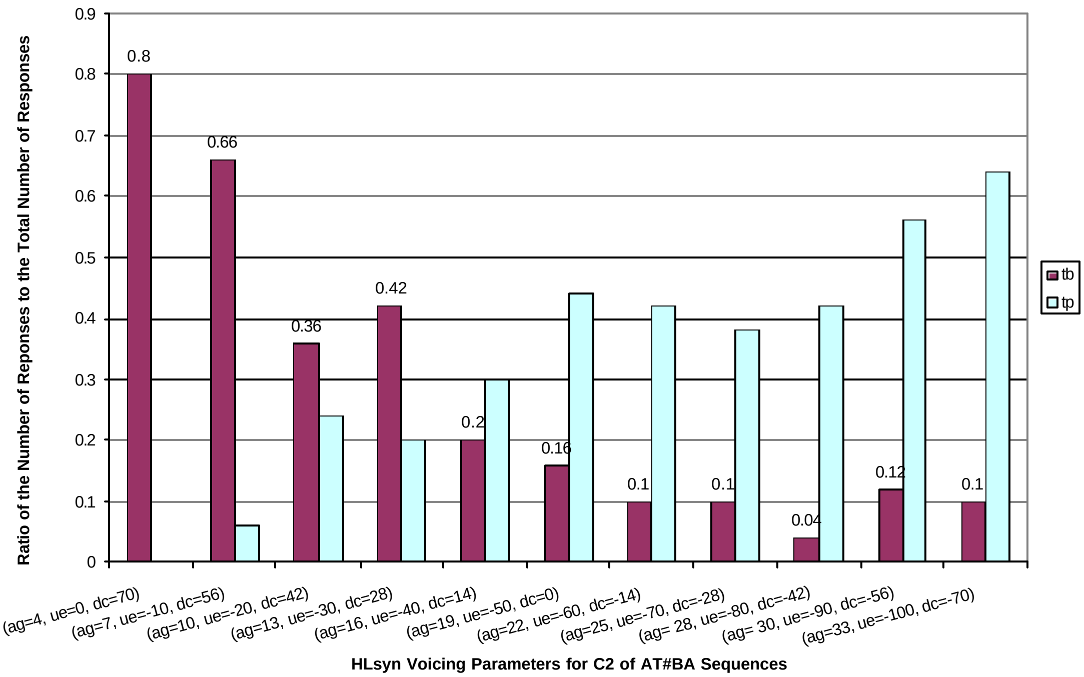

Figure 20. Perception Accuracy vs. Changing HLsyn Voicing Parameters of T#B. A plot of the perception accuracy ratios of the series of t#b sequences, as the 'b' in each successive sequence gets acoustically closer and closer to a 'p'

The results of the second perception test are contained in Table 18. Subjects rated the naturalness of the synthesized sequences compared to their spoken counterparts. The scale ranged from -5 to 5, with -5 being "extremely less natural", 0 being "just as natural", and 5 being "extremely more natural".

The overall average rating for the synthesized sequences was around 0, meaning that the listeners thought that the sequences were just as natural-sounding as the spoken sequences. The standard deviation of the ratings was around 2. Listeners seemed to think either that all the synthesized sequences were more natural-sounding than the ones extracted from natural speech or that they were all less natural-sounding than those extracted from natural speech.

Table 18: Naturalness Ratings on the Synthesized Sequences When Compared to Their Spoken Counterparts (Scale of -5 to 5)

	db	dk	dd	dg	dp	tb	tk	tg	tp
Listener 1	-1	-1	-3	-3	-2	-2	-1	-1	-3
Listener 2	4	1	2	4	3	-1	3	2	1
Listener 3	-3	-4	-1	-2	-3	-3	-2	-2	-4
Listener 4	3	-2	-3	0	-2	-3	0	2	-2
Listener 5	3	2	2	3	1	1	2	2	1
Listener 6	-2	2	3	2	1	1	2	2	1
Listener 7	3	0	3	0	-1	0	2	2	-1
Listener 8	2	0	3	1	3	2	0	0	0
Listener 9	-3	-1	0	-1	-3	-3	-1	-1	-3
Listener 10	0	-1	-1	0	-1	-1	0	0	-1
Average	0.6	-0.4	0.5	0.4	-0.4	-0.9	0.5	0.6	-1.1
Overall Avg. Rating	-0.02								
Standard Dev.	2.077								

5.3 Discussion of Perception Test Results

The perception accuracies of the set of d#d sequences with varying F2 values indicated that the F2 movements into the d-closure were sensitive to deviations toward that into a labial-closure. As the final F2 value started to get smaller than 1400Hz, the perception accuracy also started to fall. However, when the final F2 value was increased from 1400Hz all the way up to 1700Hz, the perception accuracies remained relatively unchanged. When the final F2 values were less than 1400Hz, significant numbers of b#d and d#b sequences were perceived. However, no sequences containing a velar stop(s) were perceived. Perhaps the manipulation of F3 transitions may increase the number of velars perceived, since F2 and F3 usually converge prior to velar closures. It cannot be explained why d#b sequences were perceived, since the only variable was the final F2 value prior to the C1 closure.

In addition, the results seemed to indicate an allowable range of values for the final F2 to take, even if its movements resembled that into a velar closure. However, this allowable range of values did not extend to F2 movements resembling those going into a labial-closure; such movements have very small or negative deltas. Therefore, results seemed to indicate a higher sensitivity to C2-dependent deviations in the F2 movements reflecting overlap in back-to-front sequences than that for front-to-back sequences.

These findings agree with Principle 1 derived in the previous phase. The results indicate a much bigger window of allowable C2-dependent deviation in F2 movements into the C1 closure for sequences with front-to-back order of place of articulation than for those with back-to-front order. Furthermore, the existence of such windows supports the validity of Principle 1's assertion of the presence of C2-dependent movements of the tongue body and possibly of other oral articulators while going into the C1 closure. In addition, the results agree with Principle 1's suggestions that the formation of the C2 closure may begin prior to the C1 release, perhaps even before the C1-closure, and not negatively affect perception accuracy.

Compared to VC#CV sequences with C1 bursts present, sequences without C1 bursts had lower perception accuracies. The differences between the two groups were found to be statistically significant using an analysis of variance (ANOVA) with alpha of 0.05 (see Table 19). This suggests the important role C1 bursts play in the perception of V1C1#C2V2 sequences.

Table 19: ANOVA Test Result Examining the Differences in Perception Accuracies between Sequences with C1 Bursts and without C1 Bursts

Group 1	Group 2	p-Value	H_0: Means from Group 1 and Group 2 are the same
Sequences with C1 bursts (8)	Sequences without C1 bursts (8)	0.004	Rejected

In addition, absent t-bursts were found to decrease perception accuracies more than absent d-bursts. The differences between the perception accuracies of sequences with absent t-bursts and those with absent d-bursts were found to be statistically significant using an analysis of variance with alpha of 0.05 (see Table 20). This finding suggests that t-bursts may play a bigger role in the perception of t-sequences than d-bursts in the perception of d-sequences.

The importance of the C1-burst in perception suggests a reason for the non-obliteration of the C1 bursts in sequences that may have greater gestural overlap. The findings do not conflict with Principle 1's statement that the extent of overlap usually does not go so far as to obliterate C1 bursts.

Table 20: ANOVA Test Result Examining the Differences in Perception Accuracies between Sequences with Absent D-Bursts and Sequences with Absent T-Bursts

Group 1	Group 2	p-Value	H_0: Means from Group 1 and Group 2 are the same
D-sequences without d-bursts (4)	T-sequences without t-bursts (4)	0.026	Rejected

Perception test results from d#p and t#b sequences with varying voicing parameters suggest that a window of allowable C1-dependent deviations for the voicing of C2 exist. For instance, the perception accuracies of d#p sequences were not significantly affected until voicing parameters ag decreased from 50 to 13.8, ue increased from -100 to -30, and dc increased from -70 to 28. The perception accuracies of t#b sequences were not significantly affected until voicing parameters ag increased from 4 to 10, ue decreased from 0 to -20, and dc decreased from 70 to 42. The existence of such windows supports the validity of the existence of laryngeal gestural overlap, as stated in Principle 2.

In addition, application of the principle in HLsyn revealed a gray area which seemed to separate voiced and voiceless consonants. In this gray area, ag ranged from 10 to 13, ue ranged from -20 to -30, and dc ranged from 20 to 40. However, it is important to note that these parameters were modified dependently as a set of variables in the generation of d#p and t#p sequences for the perception test.

Finally, participants of the perception tests generally found the synthesized sequences to be just as natural as those extracted from actual speech. This finding lessens the possibility that the results of the first perception test were skewed because of the unnaturalness of synthesized sequences.

Chapter 6: Conclusion and Recommendation

Through acoustic data analysis, quasi-articulatory synthesis verification, and perception testing, two possible principles governing the coordination of oral and laryngeal articulators in the production of English stop-consonant sequences emerge from this study.

First, V1C1#C2V2 stop-consonant sequences with front-to-back order of place of articulation have more overlap of articulators than those with back-to-front order. More specifically, there exist C2-dependent movements of the tongue body and possibly other oral articulators while going into the C1 closure. In addition, the formation of the C2 closure by the appropriate articulators is likely to begin prior to the release of C1, perhaps even before the C1 closure, for such sequences. The extent of the overlapping usually does not go beyond the obliteration of the C1 release burst. This finding agrees with results from previous studies demonstrating more gestural overlap in sequences with front-to-back order or place of articulation (Byrd, 1996; Chitoran, Goldstein, and Byrd, 2002; Zsiga, 1994).

Second, gestural overlap of laryngeal articulators, such as glottal opening adjustments and vocal-fold stiffness adjustments, usually affects the voicing of C2 in V1C1#C2V2 sequences. The existence of the overlap varies from individual to individual. There may be a window of allowable overlap of laryngeal gestures that does not significantly sacrifice speech perception accuracy.

Future extensions from this study may involve closer examinations of possibly greater overlap in sequences with voiceless C1s than in those with voiced C1s. Acoustic analysis in this study has found such sequences to have shorter closure duration ratios than sequences with voiced C1s; such differences are statistically significant.

In addition, fundamental frequency, voice-bar duration, and voicing onset readings may be measured and analyzed for the spoken stop-consonant sequences to further substantiate voicing characteristics data for the inference of laryngeal gestural timing estimates and the derivation of possible principles.

Singleton VCV sequences need also be acoustically analyzed. Acoustic data of VCV sequences may be compared to those of VCCV sequences to reveal and investigate possible evidence suggesting gestural overlap in the production of stop-consonant sequences.

Finally, perception tests involving sequences with varying degrees of overlap not limited to alveolar stops in the C1 position may be administered. This would give more data demonstrating the effects of overlap in sequences with various orders of place of articulation and various voicing orders on perception accuracies.

Bibliography

Browman, C. and Goldstein, L. (1992) Articulatory phonology: an overview. *Phonetica*, **49**, 155-180.

Byrd, D. (1996) Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, **24**, 209-244.

Byrd, D. (1992) Perception of assimilation in consonant clusters: a gestural model. *Phonetica*, **49**, 1-24.

Chitoran, I., Goldstein, L. and D. Byrd (2002) Gestural overlap and recoverability: Articulatory evidence from Georgian. C. Gussenhoven, T. Rietveld, and N. Warner (eds.) Papers in Laboratory Phonology 7, Cambridge University Press.

Devore, J. L. *Probability and Statistics for Engineering and the Sciences*. 4th ed. Wadsworth Publishing, 1995.

Hanson, H. M. (1997). "Glottal characteristics of female speakers: Acoustic correlates," J*ournal of the Acoustical Society of America*, **101**, 466-481.

Hanson, H.M. and Stevens, K.N. (2002) A quasi-articulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer. In Press.

Repp, B.H. (1983) Bidirectional contrast effects in the perception of VC-CV sequences. *Perception and Psychophysics*, **33(2)**, 147-155.

Stevens, K.N. (1998) *Acoustic Phonetics*. MIT Press, Cambridge, MA.

Stevens, K.N. (2002) Towards a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America*, **111(4)**, 1872-1891.

Surprenant, A.M. and Goldstein, L. (1998) The perception of speech gestures. *Journal of the Acoustical Society of America*,**104(1)**, 518-529.

Zsiga, E.C. (1994) Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, **22**, 121-140.

Zsiga, E. C. (2000) Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics*, **28**, 69-102.

Appendix A: Perception Test 1 Instruction Sheet

You will hear a series of vowel-consonant-consonant-vowel sequences (e.g. "adba"). Each sequence is repeated twice. Please indicate the 2 consonants you hear in each sequence to the best of your ability.

Possible answers are: db, dd, dg, dp, dt, dk, tp, tt, tk, td, tg, tb. Sequences may be repeated. Please type one response per line.

To begin, load the playlist named "test1.m3u" in winamp and play it. Please use headphones to listen to the sound sequences. You may pause the test if needed. Please note that sequences with the same number indicator do not necessary mean that they are identical.

After you finish taking the test, please save this file and email it to szhao@mit.edu and I will email you a short summary of what the tests are for. Thanks.

Appendix B: Perception Test 2 Instruction Sheet

You will hear a series of vowel-consonant-consonant-vowel sequences (i.e. "adba"). Each sequence is repeated twice. Sequences with names ending in "synthesized" are synthesized. Each synthesized sequence follows one that is spoken by a speaker. Please use headphones to listen to the sound sequences. Please indicate on a scale of -5 to 5 how natural the synthesized version is compared to the actual.

5 = extremely more natural
4 = much more natural
3 = definitely more natural
2 = fairly more natural
1 = a bit more natural
0 = just as natural
-1 = a bit less natural
-2 = fairly less natural
-3 = definitely less natural
-4 = much less natural
-5 = extremely less natural

To begin, load the playlist named "test2.m3u" in Winamp and play it. You may pause the test if needed. Please type only one response per line.

After you finish taking the test, please save this file and email it to szhao@mit.edu and I will email you a summary of what the tests are for. Thanks.

Appendix C: Speaker 1's Acoustic Data

Speaker1	codtop	codpot	codgot	coddot	codcop	codbob
f2_midVC1	1174,1586(427)	1152,1586(445)	1152(444)	1174,1565(446)	1174(402)	1174(431)
f2_60msPC1	1195,1565(447)	1152,1586(464)	1174(463)	1174,1565(465)	1195(422)	1195,1565(450)
f2_40msPC1	1174,1521(467)	1174,1565(482)	1174,1500(482)	1152,1543(485)	1195(442)	1174,1565(470)
f2_20msPC1	1108,1521(487)	1130,1500(500)	1521(502)	1521(505)	1500(462)	1152,1543(490)
f2_PC1	1108,1521(509)	1152,1586(518)	1521(521)	1521(525)	1521(482)	1543(509)
f3_midVC1	2412,2782	2412,2847	2021,2412	2456	2391	2021,2434
f3_60msPC1	2412,2803	2412,2782	1999,2391	2434	2456	2021,2434
f3_40msPC1	2391,2803	2412,2803,3260	1999,2391	2391	2391	2021,2434
f3_20msPC1	2391,2760	2391,2738,3260	1999,2391	2412	2391	1999,2412
f3_PC1	2391,3064	2391,3151	1978,2391	2391	2369	1999,2391
f2_OnC2V2	1304(672)	1152(694)	1543(671)	1521(663)	1217,1543(663)	1130(645)
f2_20msOV2	1174(691)	1152(713)	1478(693)	1478(684)	1195(684)	1130(666)
f2_40msOV2	1174(712)	1152(733)	1217(714)	1195(705)	1174(704)	1087(687)
f2_60msOV2	1174(733)	1152(754)	1174(735)	1174(728)	1174(725)	1043
f3_OnC2V2	2391,2803	2325,2869	1999,2412	2412	2021,2434,2803	2021,2456
f3_20msOV2	2391,2760	2369,2869	2021,2412	2412	2021,2391,2847	2021,2412
f3_40msOV2	2369,2803	2369,2825	2043,2434	2391	2021,2412,2782	1999,2391
f3_60msOV2	2391,2803	2369,2825	2043,2434	2412	2021,2369,2803	1999,2369
dur_closure	81.2ms	109.2ms.	36.4ms,59.9ms	98.2	48.0ms,38.3ms	105.3ms.
closure_2_V	153.1(517,669)	164.4(527,690)	136.0(532,668)	124.8(534,659)	168.1(492,660)	121.8(519,633)
release_2_V	71.5	54.9	35.6	27.1	74.6	17.3
burst	d(0)t(598)	d(0)p(636)	d(568),g(633)	d(0),d(632)	d(541),c(586)	d(0),b(624.3)
H1-H2,V1	43.1-45.3(451)	52.9-49.3(475)	49.8-47.5(449)	46.4-46.6(458)	45.4-48.4(428)	46.4-49.3(464)
V2	40.8-38.8(727)	40.3-42.2(738)	38.3-39.2(751)	37.9-44.4(721)	36.6-41.5(717)	37.1-44.2(724)
H1-A1,V1	43.1-52.8	52.9-58.2	49.8-54.9	46.4-53.7	45.4-56.4	46.4-57.6
V2	40.8-50.9	40.3-52.3	38.3-49.1	37.9-54.0	36.6-50.7	37.1-54.1

Speaker1	docktop	dockpot	dockgot	dockdot	dockcop	dockbob
f2_midVC1						
f2_60msPC1	1195(405)	1152(382)	1195(364)	1195(370)	1195(420)	1174(424)
f2_40msPC1	1174(427)	1130(403)	1174(385)	1174(391)	1174(441)	1174(446)
f2_20msPC1	1174(448)	1130(422)	1174(407)	1152(412)	1174(461)	1174(466)
f2_PC1	1195(468)	1152(441)	1195(429)	1174(433)	1195(481)	1174(488)
f3_midVC1						
f3_60msPC1	1956,2304	1934,2260,2695	2347	1956,2347	1934,2347	1978,2369
f3_40msPC1	1978,2369	1978,2304,2695	1978,2369	1956,2369	1999,2391	1978,2347
f3_20msPC1	1978,2412	1934,2369,2695	1978,2412	1956,2347	1999,2391	1999,2347
f3_PC1	1956,2238	1999,2673	1978,2456	1978,2325	2021,2412	1999,2369
f2_OnC2V2	1369(689)	1152(648)	1500(599)	1456(619)	1326(702)	956(712)
f2_20msOV2	1195(709)	1152(667)	1413(619)	1239(641)	1217(722)	978(733)
f2_40msOV2	1195(729)	1152(687)	1217(641)	1195(662)	1195(742)	1087(754)
f2_60msOV2	1174(750)	1174(708)	1195(663)	1174(684)	1174(764)	1021(776)
f3_OnC2V2	1934,2347	2260	1956,2347	2347	1978,2347,2782	2304
f3_20msOV2	1956,2347	2304	1999,2369	2391	1978,2347,2760	2325
f3_40msOV2	1956,2325	2282	1978,2391	2347	2021,2760	2369
f3_60msOV2	2304	1978,2412	2347	2021,2412	2347	
dur_closure	129.4ms	63,64.6	137.1	76.5,65.6	143.7	67,97.3
closure_2_V	205.2(478,680)	193(450,644)	161(436,596)	169.6(440,608)	208.5(490,699)	211.3(498,708)
release_2_V	top(76.9)	pot(43.2)	got(23.9)	dot(21.1)	cop(77.9)	bob(23.1)
burst	k(0)t(602.6)	k(500)p(595.4)	k(0)g(571.9)	k(515.4)d(587.8)	k(620.56)	k(560)b(681)
H1-H2,V1	37.3-41.6(408)	40.8-43.9(405)	36.1-43.5(379)	39.1-43.0(384)	42.7-41.0(435)	42.7-41.6(430)
V2	39.4-42.3(726)	34.0-40.7(701)	39.3-42.8(669)	37.7-42.4(699)	35.3-38.7(759)	34.3-40.2(793)
H1-A1,V1	37.3-50.6	40.8-55.9	36.1-52.2	39.1-50.8	42.7-51.6	42.7-52.7
V2	39.4-47.5	34.0-48.2	39.3-50.6	37.7-49.3	35.3-46.5	34.3-47.5

Speaker 1	dogtop	dogpot	doggot	dogdot	dogcop	dogbob
f2_midVC1	1130(381)	1108(375)	1174(372)	1152(394)	1152(363)	1130(365)
f2_60msPC1	1021(402)	1087(396)	1174(392)	1152(414)	1130(383)	1108(386)
f2_40msPC1	1043(423)	1065(416)	1174(414)	1152(435)	1108(403)	1108(406)
f2_20msPC1	1130(444)	1130(437)	1174(436)	1174(456)	1174(425)	1152(427)
f2_PC1	1152(467)	1195(457)	1326(456)	1260(476)	1239(445)	1195(447)
f3_midVC1	2325	2304	2369	2412	2347	1978,2369,2803
f3_60msPC1	2325	2304	2412	2391	2347	1978,2369,2847
f3_40msPC1	2325	2325	2391	2391	2347	1978,2369,2847
f3_20msPC1	2325	2325	2412	2412	2369	1978,2347,2847
f3_PC1	2238	1978,2391	2477	2499	2412	1999,2456,2869
2869						
f2_OnC2V2	1195(663)	1087(673)	1478(611)	1456(626)	1195(664)	956,1217(609)
f2_20msOV2	1174(682)	1000(693)	1260(633)	1217(648)	1174(683)	1000(629)
f2_40msOV2	1152(703)	1065(714)	1195(656)	1195(669)	1152(703)	1000(650)
f2_60msOV2	1152(724)	1174(736)	1174(679)	1152(691)	1152(724)	1087(672)
f3_OnC2V2	2391	2325	1978,2391	2369	2021,2412,2803	2217,2499
f3_20msOV2	2347	2325	1999,2412	2412	2021,2803	2304
f3_40msOV2	2391	2325	2021,2434	2369	1999,2391,2825	2347
f3_60msOV2	2391	2325	2021,2456	2369	2021,2434,2847	2391
dur_closure	117.6	138	109.5	109.6	122.4ms	139.6ms
closure_2_V	183.7(476,659)	195.1(467,662)	144.9(466,608)	139.8(467,605)	199.2(454,652)	148(457,604)
release_2_V	69.4	57.2	35.9	30.3	76.3	11.3
burst	g(0)t(589.6)	g(0)t(604)	g(0)g(573)	g(0)d(592.9)	g(0)c(584)	d(0)b(594)
H1-H2,V1	39.9-43.3(397)	40.5-46.3(379)	38.7-40.9(387)	40.5-44.0(407)	42.4-45.8(378)	42.0-43.7(380)
V2	31.9-40.1(717)	34.1-42.0(719)	33.7-39.7(673)	37.7-40.5(707)	38.8-42.8(707)	34.4-40.5(688)
H1-A1,V1	39.9-54.4	40.5-57.8	38.7-52.6	40.5-53.6	42.4-54.6	42.0-55.6
V2	31.9-48.4	34.1-50.5	33.7-47.8	37.7-47.9	38.8-51.1	34.4-49.3

Speaker 1	gottop	gotpot	gotdot	gotcop	gotbob	dotgot
f2_midVC1					1217b(388)	1195(491)
f2_60msPC1	1217(377)	1239(391)	1195(353)	1195(404)	1195(408)	1174(522)
f2_40msPC1	1217(396)	1195(413)	1217(373)	1195(425)	1174(429)	1152(542)
f2_20msPC1	1217(407)	1195(434)	1217(393)	1217(446)	1195(449)	1174(562)
f2_PC1	1217(429)	1174(454)	1521(415)	1369(468)	1217(470)	1239(583)
f3_midVC1					2021,2412	2456
f3_60msPC1	1978,2434	2021,2456	2456	2021,2434	2021,2456	2456
f3_40msPC1	1978,2456	1999,2477	2456	2021,2456	1999,2477	2456
f3_20msPC1	1978,2434	1999,2477	2434	2021,2456	1999,2477	2456
f3_PC1	1978	1999,2456	2456	2043,2477	1978,2521	2456
f2_OnC2V2	1239(620)	1174(664)	1478(549)	1260(651)	1174(623)	1434(769)
f2_20msOV2	1195(639)	1152(682)	1391(569)	1195(671)	1174(643)	1239(790)
f2_40msOV2	1174(659)	1174(701)	1217(590)	1195(691)	1174(665)	1217(812)
f2_60msOV2	1174(680)	1195(722)	1195(612)	1174(713)	1174(687)	1195(834)
f3_OnC2V2	2434	2043,2499	2456	2043,2477	1999,2477	2043,2456
f3_20msOV2	2456	2043,2477	2434	2021,2499	2021,2456	2043,2477
f3_40msOV2	2499	2021,2456	2434	2043,2521	2043,2456	2043,2477
f3_60msOV2	2456	1956,2456	2456	2043,	2043,2477	2065,2477
dur_closure	17.5,69.4	114.6	108.5	115.7	127.0	25.9,93.6
closure_2_V	170.3(446,617)	168.2(491,660)	125.2(420,545)	172.4((475,647)	127.0(492,627)	167.6(598,766)
release_2_V	72.8	53.5	16.7	56.3	0	33.7
burst	t(463.9)t(543.4)	t(0)p(606.9)	t(0)d(528.6)	t(0)k(591.5)	t(0)d(619.1)	t(624.6)g(732.9)
H1-H2,V1	35.5-41.3(381)	38.6-41.6(408)	33.4-38.3(346)	37.5-40.1(388)	38.7-41.5(413)	37.5-42.7(505)
V2	34.4-40.4(673)	41.2-43.4(717)	37.5-39.3(607)	36.1-38.8(695)	34.1-41.4(703)	36.5-40.0(839)
H1-A1,V1	35.5-47.1	38.6-51.5	33.4-48.4	37.5-48.2	38.7-50.6	37.5-51.6
V2	34.4-49.9	41.2-49.5	37.5-50.0	36.1-51.1	34.1-48.8	36.5-47.5

Speaker 1	mobtop	mobpot	mobgot	mobdot	mobcop	mobbob
f2_midVC1	934,1108(381)	1152(390)		1130(429)	1152(327)	
f2_60msPC1	1108(401)	1152(413)	1130(411)	1130(450)	1152(347)	1065(384)
f2_40msPC1	956,1087(421)	1130(432)	1130(432)	1152(470)	1152(367)	1021(403)
f2_20msPC1	913,1152(442)	956(451)	1130(453)	1152(491)	1152(387)	978(424)
f2_PC1	978(461)	913(470)	1174(475)	1152(511)	1152(407)	934(444)
f3_midVC1	1934,2391,2760	2434		2434,2847	2412,2869	
f3_60msPC1	1956,2391,2869	2412	2434	2434,2825	2412,2869	2391,2803
f3_40msPC1	1956,2391,2869	2391	2434	2412,2803	2412,2803	2369,2847
f3_20msPC1	1956,2412,2890	2347	2434,	2369,2803	2412,2825	2369,2825
f3_PC1	1956,2825	2217	2260,2890	1978,2499,2847	2304,2869	2217,2803
f2_OnC2V2	1174(654)	1152(662)	1521(648)	1521(695)	1239(605)	978(608)
f2_20msOV2	1217(673)	913,1174(680)	1239(670)	1217(716)	1217(624)	1021(628)
f2_40msOV2	1195(693)	934,1130(701)	1195(693)	1195(737)	1195(643)	1087(650)
f2_60msOV2	1174(713)	1152(721)	1195(715)	1174(760)	1174(664)	1108(673)
f3_OnC2V2	2434,2912	2369	2369,2825	2434	2021,2391,2847	2325
f3_20msOV2	2456,2869	2369	2391,2869	2391,2847	2021,2434,2847	2391,2869
f3_40msOV2	2412,2803	2369	2412,2869	2434,2847	2021,2391,2825	2412,2847
f3_60msOV2	2391,2825	2347	2456,2803	2412	2021,2391,2760	2412,2847
dur_closure	58.2,46.5	113.1	59.9,57.8	86.2,36.7	48.7,46.7	135.1
closure_2_V	179.6(472,651)	179.8(479,659)	158.9(486,644)	169.6(522,691)	174.7(427,602)	152.0(454,606)
release_2_V	64.0	67.0	35.3	35.3	63.7	17.2
burst	b(529),t(587)	b(0),p(590)	b(545),g(608)	b(608),d(656)	b(476),c(538)	b(0),b(589)
H1-H2,V1	43.3-45.8(406)	44.4-47.2(418)	39.3-43.4(406)	40.6-44.1(435)	43.7-46.7(362)	44.5-45.4(390)
V2	40.9-39.1(718)	38.3-42.5(706)	39.9-37.2(720)	34.4-41.1(764)	39.5-38.7(659)	35.4-41.5(711)
H1-A1,V1	43.3-51.1	44.4-53.9	39.3-50.5	40.6-48.9	43.7-53.1	44.5-50.9
V2	40.9-50.4	38.3-51.7	39.9-46.1	34.4-48.1	39.5-49.1	35.4-47.9

Speaker 1	toptot	toppot	topgot	topdot	topcop	topbob
			last/first			
f2_midVC1						
f2_60msPC1	1217(404)	1152(447)	1217(398)	1174(364)	1326(339)	1174(411)
f2_40msPC1	1174(424)	1152(467)	1217(419)	1195(384)	1195(358)	1174(430)
f2_20msPC1	1130(442)	1108(485)	1195,1586(439)	1174(404)	1174(378)	1130(449)
f2_PC1	1174(462)	1174(505)	978,1239,1565(458)	1174(425)	1195(398)	1065(469)
f3_midVC1						
f3_60msPC1	2434	2021,2412	2043,2477	2434	2043,2456	2434
f3_40msPC1	2456	2043,2412	2043,2456	2434	2021,2456	2434
f3_20msPC1	2434	2021,2412	2043,2434	2434	2043,2456	2456
f3_PC1	2434	2043,2434	2065,2456	2456	2043,2477	2434
f2_OnC2V2	1195(626)	1152(680)	1565(638)	1304(557)	1326(590)	1174(625)
f2_20msOV2	1174(646)	1152(699)	1456(657)	1195(577)	1260(610)	1174(644)
f2_40msOV2	1174(665)	1195(718)	1282(678)	1217(597)	1195(630)	1174(665)
f2_60msOV2	1043(686)	1043(739)	1239(700)	1217(618)	1195(650)	1174(686)
f3_OnC2V2	2412	2021,2412	2065,2434	2021,2434	2021,2434	2065,2477
f3_20msOV2	2412	2021,2434	2021,2434	2434	1999,2456	2043,2477
f3_40msOV2	2412	1999,2434	1978,2434	2456	2043,2412	2021,2434
f3_60msOV2	2369	2391	1978,2434	2434	2043,2456	2021,2412
dur_closure	74.4	95.4	71.4	73.4,24.2	59.6,45.3	109.2
closure_2_V	144.2(480,623)	151.9(523,677)	152.5(471,624)	124.5(431,553)	182.8(406,588)	122.9(499,621)
release_2_V	69.7	58.1	74.1	21.3	68.3	14
burst	p(0)t(553.3)	p(0)p(619)	p(0)g(537)	p(503)d(532.3)	p(464.6)c(519.3)	p(0)b(607)
H1-H2,V1	45.6-48.1(437)	49.1-48.7(472)	43.3-45.8(414)	43.5-45.2(389)	44.3-46.9(374)	45.2-48.1(445)
V2	44.8-46.8(670)	46.9-47.3(723)	37.7-43.2(682)	37.1-44.4(622)	44.4-45.1(627)	39.7-43.1(700)
H1-A1,V1	45.6-55.3	49.1-54.7	43.3-50.4	43.5-49.4	44.3-51.5	45.2-52.2
V2	44.8-48.9	46.9-50.9	37.7-50.7	37.1-50.4	44.4-45.2	39.7-48.6

Appendix D: Speaker 2's Acoustic Data

Speaker 2	codtop	codpot	codgot	coddot	codcop	codbob
v1f1	717	717	717	739	739	717
f2_60msPC1	1565(487.0)	1434(587.9)	1521(629.3)	1521(601.5)	1500(602.2)	1456(621.8)
f2_40msPC1	1608(505.1)	1478(604.6)	1586(653.3)	1608(624.9)	1565(625.4)	1500(639.9)
f2_20msPC1	1652(525.0)	1543(625.7)	1652(671.8)	1673(643.5)	1630(643.3)	1543(662.5)
f2_PC1	1630(543.4)	1521(647.1)	1608(689.7)	1652(660.7)	1652(666.3)	1521(685.9)
f3_60msPC1	2456	2434	2391	2369	2238	2282
f3_40msPC1	2543	2521	2434	2347	2304	2391
f3_20msPC1	2586	2608	2586	2717	2477	2456
f3_PC1	2651	2651	2630	2651	2651	2564
v2f2	717	717	717	717	739	739
f2_OnC2V2	1478(733.6)	1239(877.7)	1717(858.7)	1760(831.0)	1695(864.2)	1347(847.4)
f2_20msOV2	1434(751.3)	1369(894.5)	1652(876.5)	1673(848.9)	1478(886.4)	1347(865.1)
f2_40msOV2	1434(770.3)	1413(911.7)	1565(894.6)	1565(867.0)	1434(904.8)	1369(885.0)
f2_60msOV2	1391(788.4)	1456(928.6)	1521(912.5)	1521(885.6)	1434(923.5)	1304(903.0)
f3_OnC2V2	2369	2304	2304	2651	2282	2282
f3_20msOV2	2325	2325	2304	2369	2217	2152
f3_40msOV2	2347	2347	2282	2282	2217	2108
f3_60msOV2	2325	2434	2304	2282	2195	2086
dur_closure	118.1	119.4	83.4,36.4 138.0		79.0,36.3 135.3	
closure_2_V	175.9	219.1	151.4	157.5	186.5	135.3
release_2_V	57.5	99.7	19.7	19.4	67.6	0
burst	d() t(671.7-729.2)	d() p(774.3-874.0)	d(780.5-792.5) g(828.9-848.6)	d() d(807.2-826.6)	d(752.7-756.4) k(792.7-860.2)	d() b(827.6)
V1: h1	152 Hz 50.1	174 Hz 58.1	174 Hz 57.1	174 Hz 56.3	174 Hz 54.8	174Hz, 55.8
h2	326 Hz 47.5	369 Hz 53.9	326 Hz 51.3	326 Hz 51.5	348 Hz 53.0	348Hz, 50.1
a1	630 Hz 54.5	739 Hz 60.1	674 Hz 57.3	674 Hz 58.2	695 Hz 58.1	695Hz, 55.2
V2: h1	152 Hz 52.6	174 Hz 57.8	174 Hz 56.6	152 Hz 54.7	174 Hz 53.9	152Hz, 54.0
h2	326 Hz 50.5	348 Hz 53.0	326 Hz 52.0	326 Hz 51.2	326 Hz 53.5	326Hz, 50.3
a1	630 Hz 54.8	717 Hz 57.6	674 Hz 59.2	630 Hz 56.3	826 Hz 58.7	804Hz, 57.2

Speaker 2	docktop	dockpot	dockgot	dockdot	dockcop	dockbob
v1f1	761	739	739	739	739	739
f2_60msPC1	1434(440.5)	1434(478.3)	1413(439.4)	1456(468.1)	1347(593.6)	1478(506.7)
f2_40msPC1	1434(463.5)	1413(500.9)	1369(462.9)	1478(486.8)	1413(615.4)	1456(529.0)
f2_20msPC1	1456(480.1)	1434(522.4)	1413(480.9)	1500(505.3)	1478(638.1)	1478(550.8)
f2_PC1	1456(502.1)	1456(542.4)	1500(499.3)	1565(523.3)	1500(662.6)	1391(570.9)
f3_60msPC1	2282	2260	2217	2238	2217	2238
f3_40msPC1	2282	2260	2173	2282	2238	2238
f3_20msPC1	2282	2282	2086	2369	2217	2195
f3_PC1	2369	2173	2065	2434	2282	2434
v2f2	739	717	761	739	695	717
f2_OnC2V2	1478(747.2)	1391(861.6)	1673(733.7)	1673(811.8)	1391(977.5)	1195(795.6)
f2_20msOV2	1434(763.9)	1434(878.3)	1608(752.1)	1608(830.1)	1326(997.8)	1195(812.5)
f2_40msOV2	1434(781.6)	1456(895.3)	1543(768.4)	1521(849.7)	1347(1013.2)	1195(830.2)
f2_60msOV2	1413(799.7)	1521(912.9)	1521(786.6)	1456(868.2)	1413(1033.8)	1195(847.3)
f3_OnC2V2	2260	2282	2369	2543	2130	2325
f3_20msOV2	2282	2238	2238	2456	2130	2304
f3_40msOV2	2260	2282	2282	2325	2173	2304
f3_60msOV2	2282	2369	2304	2260	2238	2325
dur_closure	85.4,59.4 154.8,36.7	182.7	90.1,139.2 167.7	82.4,95.2		
closure_2_V	224.0	294.8	209.0	260.7	280.4	206.9
release_2_V	72.6	91.3	26.3	11.0	112.6	5.5
burst	k(599.5-606.2) p(665.6-738.2)	k(713.1-725.2) p(761.8-853.2)	k() g(703.8-730.1)	k(637.7-658.2) d(797.4-808.4)	k() k(856.7-969.4)	k(662.9-686.8) b(781.9-787.4)
V1: h1	174 Hz 53.3	174 Hz 53.1	174 Hz 53.1	152 Hz 49.7	174 Hz 54.0	174Hz,52.9
h2	348 Hz 49.5	348 Hz 47.1	326 Hz 50.2	326 Hz 45.8	348 Hz 49.1	348Hz,48.1
a1	695 Hz 53.8	717 Hz 49.5	674 Hz 57.7	804 Hz 52.6	717 Hz 55.5	695Hz,54.0
V2: h1	174 Hz 57.1	174 Hz 58.0	174 Hz 55.3	152 Hz 48.4	196 Hz 56.0	174Hz,54.6
h2	348 Hz 53.0	348 Hz 55.4	348 Hz 52.6	326 Hz 46.8	391 Hz 53.0	348Hz,49.9
a1	695 Hz 56.9	695 Hz 60.9	695 Hz 58.3	782 Hz 57.4	587 Hz 58.7	695Hz,57.0

Speaker 2	dogtop	dogpot	doggot	dogdot	dogcop	dogbob
v1f1	695	717	695	695	695	674
f2_60msPC1	1239(514.3)	1391(458.8)	1217(428.4)	1347(492.6)	1239(504.9)	1347(458.5)
f2_40msPC1	1195(537.9)	1304(481.8)	1217(453.9)	1217(514.0)	1217(528.9)	1326(470.5)
f2_20msPC1	1195(555.6)	1282(504.8)	1239(473.0)	1195(535.9)	1217(546.8)	1326(488.5)
f2_PC1	1195(578.8)	1282(523.2)	1195(491.1)	1195(557.1)	1217(564.1)	1217(506.2)
f3_60msPC1	2282	2304	2021	2282	2304	2238
f3_40msPC1	2347	2347	2260	2391	2391	2282
f3_20msPC1	2391	2369	2412	2456	2434	2456
f3_PC1	2477	2434	2543	2586	2499	2499
v2f2	717	717	717	739	695	739
f2_OnC2V2	1434(828.4)	1326(756.2)	1717(670.9)	1673(746.1)	1500(770.8)	1260(723.4)
f2_20msOV2	1282(844.3)	1260(773.1)	1695(688.8)	1586(762.8)	1456(787.6)	1369(742.1)
f2_40msOV2	1326(862.8)	1326(790.5)	1565(706.1)	1565(780.1)	1478(805.3)	1434(761.2)
f2_60msOV2	1239(880.0)	1391(808.5)	1565(725.9)	1500(797.5)	1456(823.1)	1347(779.3)
f3_OnC2V2	2282	2325	2630	2543	2304	2304
f3_20msOV2	2195	2304	2630	2586	2304	2282
f3_40msOV2	2282	2325	2347	2304	2260	2108
f3_60msOV2	2282	2347	2456	2260	2304	2282
dur_closure	143.3	103.1,31.0	138.8	146.3	127.0	104.7,69.8
closure_2_V	232.2	210.9	155.7	168.0	189.9	194.8
release_2_V	88.8	49.2	16.9	21.6	62.8	16.0
burst	g()	g(639.3-666.7)	g(check) g()	g()	g(629.8-634.1)	
	t(734.7-823.5)	p(697.7-746.9)	g(651.5-668.4)	d(716.3-737.9)	c(698.8-731.7)	b(704.0-720.0)
V1: h1	174 Hz 54.7	174 Hz 53.0	152 Hz 53.2	174 Hz 56.9	174 Hz 54.2	174Hz 55.3
h2	348 Hz 49.4	348 Hz 49.3	326 Hz 50.1	369 Hz 51.1	326 Hz 49.7	326Hz 50.6
a1	674 Hz 56.9	674 Hz 57.3	630 Hz 55.7	717 Hz 57.7	674 Hz 56.6	674Hz 56.5
V2: h1	174 Hz 55.6	174 Hz 56.9	152 Hz 55.6	174 Hz 57.4	174 Hz 58.4	152Hz 55.2
h2	326 Hz 50.4	326 Hz 52.0	326 Hz 52.5	348 Hz 52.8	326 Hz 51.7	326Hz 52.1
a1	674 Hz 56.5	674 Hz 58.1	652 Hz 56.6	674 Hz 55.1	674 Hz 57.1	804Hz 58.3

Speaker 2	gotpot	gotdot	gotcop	gotbob	dotgot
v1f1	717	717	739	739	739
				1500(tb549.2)	
f2_60msPC1	1521(469.6)	1478(502.2)	1478(520.4)	1521(565.4)	1478(598.2)
f2_40msPC1	1543(486.3)	1521(524.9)	1521(542.1)	1565(584.7)	1543(615.3)
f2_20msPC1	1608(506.6)	1543(545.9)	1586(562.1)	1586(601.5)	1630(630.9)
f2_PC1	1630(526.3)	1608(564.3)	1695(588.5)	1565(616.7)	1695(658.4)
				2477(tb)	
f3_60msPC1	2456	2325	2434	2586	2456
f3_40msPC1	2521	2456	2499	2651	2586
f3_20msPC1	2564	2412	2586	2695	2651
f3_PC1	2673	2325	2586	2760	2630
v2f2	695	717	717	695	739
f2_OnC2V2	1434(792.0)	1630(732.5)	1478(803.1)	1347(754.7)	1760(833.5)
f2_20msOV2	1456(814.4)	1586(750.2)	1369(824.0)	1369(771.8)	1695(851.4)
f2_40msOV2	1500(831.2)	1565(772.9)	1326(845.0)	1347(793.7)	1586(870.6)
f2_60msOV2	1543(848.2)	1521(790.3)	1282(865.7)	1347(811.0)	1565(889.7)
f3_OnC2V2	2325	2564	2412	2369	2391
f3_20msOV2	2456	2608	2412	2304	2304
f3_40msOV2	2543	2260	2456	2369	2369
f3_60msOV2	2564	2282	2434	2369	2499
dur_closure	114.4	36.4,51.7 39.3,38.5	50.3,28.8 93.9, 35.6		
closure_2_V	238.2	155.4	196.7	117.8	158.7
release_2_V	123.8	20.5	82.6	26.2	28.2
burst	t()	t(609.9-656.7)	t(637.5-673.9)	t(680.1-692.6)	t(763.5-766.1)
	p(664.4-788.2)	d(708.4-728.9)	k(712.3-794.9)	b(721.5-747.6)	g(801.9-809.1)
V1: h1	174 Hz 56.8	174 Hz 56.5	196 Hz 55.0	174 Hz 55.9	152 Hz 49.1
h2	369 Hz 51.4	348 Hz 51.9	369 Hz 51.0	348 Hz 50.4	304 Hz 50.1
a1	739 Hz 58.6	695 Hz 57.4	739 Hz 58.2	717 Hz 55.8	782 Hz 52.8
V2: h1	174 Hz 59.3	174 Hz 55.8	196 Hz 58.4	174 Hz 59.6	152 Hz 54.3
h2	348 Hz 53.6	326 Hz 50.8	369 Hz 51.8	348 Hz 53.9	326 Hz 49.6
a1	695 Hz 59.1	652 Hz 55.3	761 Hz 58.4	695 Hz 55.5	804 Hz 52.7

Speaker 2	mobtop	mobpot	mobgot	mobdot	mobcop	mobbob
v1f1	695	695	717	717	717	695
f2_60msPC1	1217(452.2)	1304(502.3)	1347(438.2)	1326(434.9)	1369(471.3)	1326(490.9)
f2_40msPC1	1282(475.4)	1304(526.2)	1347(457.2)	1304(453.7)	1413(494.8)	1304(514.9)
f2_20msPC1	1239(497.9)	1217(548.8)	1260(475.7)	1239(477.2)	1434(513.0)	1260(537.9)
f2_PC1	1152(520.7)	1130(570.8)	1130(494.5)	1152(500.2)	1152(535.5)	1174(561.6)
f3_60msPC1	2608	2760	2499	2521	2499	2651
f3_40msPC1	2630	2738	2521	2456	2543	2651
f3_20msPC1	2608	2738	2543	2499	2543	2673
f3_PC1	2630	2717	2564	2564	2564	2586
v2f2		717	739	761	717	739
f2_OnC2V2	1478(737.7)	1391(763.9)	1673(723.8)	1717(695.2)	1434(778.6)	1347(766.9)
f2_20msOV2	1456(756.4)	1456(785.2)	1586(742.9)	1630(714.3)	1326(795.5)	1391(784.0)
f2_40msOV2	1413(773.2)	1478(802.4)	1543(761.7)	1521(731.9)	1217(812.9)	1369(801.9)
f2_60msOV2	1391(791.2)	1500(820.4)	1500(781.1)	1521(751.2)	1195(824.9)	1347(821.0)
f3_OnC2V2	2304	2347	2456	2695	2195	2304
f3_20msOV2	2347	2412	2369	2673	2347	2304
f3_40msOV2	2325	2412	2325	2217	2391	2304
f3_60msOV2	2347	2434	2347	2238	2347	2325
dur_closure	76.8,37.4 121.9	73.6,100.3	80.2,58.5 64.6,62.6	172.6		
closure_2_V	195.5	171.2	206.1	173.4	226.0	188.0
release_2_V	68.0	49.6	27.7	15.9	93.3	15.5
burst	b(610.6-623.8) t(661.2-729.2)	b() p(705.3-754.7)	b(582.1-586.7) g(687.0-714.7)	b(593.7-612.6) d(671.0-686.9)	b(613.6-619.1) c(681.7-775.0)	b() b(747.3-762.7)
V1: h1	174 Hz 57.0	174 Hz 57.2	152 Hz 51.8	174 Hz 55.0	174 Hz55.4	174Hz56.3
h2	348 Hz 50.3	348 Hz 50.8	326 Hz 47.9	326 Hz 48.3	348 Hz49.1	348Hz47.9
a1	695 Hz 54.6	695 Hz 54.5	804 Hz 51.8	674 Hz 50.7	674 Hz53.4	674Hz53.8
V2: h1	174 Hz 54.7	174 Hz 56.3	152 Hz 53.0	152 Hz 51.0	174 Hz57.0	174Hz56.9
h2	326 Hz 50.3	326 Hz 51.2	326 Hz 52.1	326 Hz 48.7	348 Hz52.5	326Hz51.7
a1	674 Hz 58.3	674 Hz 56.8	804 Hz 58.3	804 Hz 56.6	674 Hz58.4	652Hz56.9

Speaker 2	toptot	toppot	topgot	topdot	topcop	topbob
v1f1	717	695	717	717	717	739
f2_60msPC1	1369(530.4)	1413(588.5)	1434(476.5)	1434(520.0)	1500(474.3)	1434(479.6)
f2_40msPC1	1304(551.4)	1347(603.2)	1434(493.4)	1434(541.6)	1478(486.6)	1369(501.2)
f2_20msPC1	1260(571.7)	1304(619.2)	1434(515.4)	1456(558.9)	1478(503.9)	1369(522.8)
f2_PC1	1174(591.6)	1174(640.6)	1239(536.2)	1239(580.1)	1304(521.3)	1260(544.3)
f3_60msPC1	2325	2477	2304	2369	2304	2369
f3_40msPC1	2325	2456	2325	2412	2304	2564
f3_20msPC1	2325	2456	2391	2391	2347	2412
f3_PC1	2369	2434	2434	2477	2325	2456
v2f2	695	674	717	717	695	739
f2_OnC2V2	1521(804.8)	1391(1391)	1673(745.6)	1717(784.8)	1434(769.7)	1413(717.2)
f2_20msOV2	1478(825.8)	1413(875.3)	1630(763.0)	1630(800.7)	1434(787.6)	1282(737.2)
f2_40msOV2	1500(842.5)	1456(892.4)	1565(780.3)	1456(817.8)	1391(804.1)	1282(754.6)
f2_60msOV2	1521(859.9)	1500(909.9)	1521(797.7)	1434(835.2)	1347(820.3)	1326(772.0)
f3_OnC2V2	2564	2434	2477	2738	2434	2391
f3_20msOV2	2369	2412	2369	2695	2456	2304
f3_40msOV2	2347	2391	2282	2325	2499	2260
f3_60msOV2	2347	2412	2304	2238	2499	2369
dur_closure	77.7,33.7	117.5	73.3,92.5	77.5,75.2	59.0,77.5	144.6
closure_2_V	198.6	199.5	200.4	178.8	228.4	156.4
release_2_V	71.9	82.1	25.4	16.4	83.3	11.7
burst	p(675.3-690.6) t(724.3-796.2)	p() p(774.2-856.2)	p(614.8-624.0) g(716.5-741.9)	p(674.4-684.1) d(759.3-775.8)	p(596.7-605.3) k(-766.1)	p() b(701.4-713.0)
V1: h1	196 Hz 52.8	196 Hz 53.5	174 Hz 56.1	174 Hz 54.0	174 Hz 51.8	196Hz 55.4
h2	391 Hz 46.5	369 Hz 47.9	369 Hz 50.3	369 Hz 51.4	326 Hz 50.3	369Hz 51.8
a1	782 Hz 54.1	739 Hz 56.4	717 Hz 57.0	739 Hz 56.6	674 Hz 54.1	739Hz 57.6
V2: h1	174 Hz 57.8	174 Hz 58.8	174 Hz 56.3	174 Hz 54.8	174 Hz 57.2	174Hz 58.7
h2	348 Hz 50.2	348 Hz 52.8	348 Hz 53.0	348 Hz 51.2	348 Hz 52.4	348Hz 54.5
a1	717 Hz 52.5	695 Hz 56.0	695 Hz 59.1	674 Hz 53.6	717 Hz 57.5	695Hz 55.8

Appendix E: Speaker 3's Acoustic Data

Speaker 3	codtop	codpot	codgot	coddot	codcop	codbob
f2_60msPC1	1282(348.0)	1260(430.5)	1282(342.0)	1304(322.6)	1260(380.3)	1260(333.4)
f2_40msPC1	1413(368.0)	1413(448.3)	1391(361.4)	1413(343.9)	1347(399.4)	1326(350.5)
f2_20msPC1	1608(387.0)	1565(466.5)	1565(381.4)	1521(363.7)	1434(419.3)	1456(369.4)
f2_PC1	1586(406.4)	1695(483.3)	1695(401.3)	1608(382.8)	1543(438.8)	1695(390.0)
f3_60msPC1	2260,2760	2238	2173,2738	2238	2238	2238
f3_40msPC1	2260,2673	2304	2195,2738	2217	2195	2195
f3_20msPC1	2564	2195,2760	2260,2738	2173	2195	2195
f3_PC1	2673	2630	2564	2173	2195	2304
f2_OnC2V2	1217(645.1)	1130(742.7)	1673(612.7)	1717(560.7)	1478(660.9)	1130(585.2)
f2_20msOV2	1174(668.3)	1152(765.7)	1369(632.2)	1478(581.6)	1174(683.4)	1087(600.2)
f2_40msOV2	1152(696.0)	1217(791.4)	1304(652.1)	1282(606.7)	1152(710.4)	1130(623.7)
f2_60msOV2	1152(710.3)	1260(804.5)	1217(670.3)	1217(633.2)	1152(725.3)	1130(637.7)
f3_OnC2V2	2260	2456	2152	2608	2325	2238
f3_20msOV2	2369	2499	2173	2673	2260	2260
f3_40msOV2	2434	2412	2173	2304	2282	2347
f3_60msOV2	2521	2391	2195	2347	2304	2391
dur_closure	156.7	65.6,95.1	145.4,28.5	158.8	44.2,52.7	70.9,85.3
closure_2_V	232.0	173.9	207.3	158.8	213.6	177.4
release_2_V	75.4	251.6	57.3	13.2	96.7	12.3
burst	d(0) t(564.8-640.2)	d(554.9-567.1) p(662.2-741.0)	d(547.9-552.4) g(580.8-589.7)	d(546.0-554.3) d(0)	d(487.6-508.0) c(560.6-657.3)	d(466.9-475.8) b(561.1-563.6)
H1-H2,V1	53.1-51.2(339.4)	46.7-55.1(415.9)	53.0-51.3(345.6)	53.3-51.6(317.8)	53.8-52.6(382.4)	48.9-55.9(346.5)
V2	47.3-46.4(686.0)	46.7-46.8(796.4)	46.1-47.0(689.0)	47.1-47.9(624.2)	48.7-44.5(715.7)	48.7-45.1(658.1)
H1-A1,V1	53.1-56.8	46.7-61.5	53.0-56.8	53.3-58.1	53.8-57.8	48.9-61.4
V2	47.3-50.9	46.7-51.9	46.1-51.4	47.1-54.1	48.7-50.6	48.7-51.5

Speaker 3	docktop	dockpot	dockgot	dockdot	dockcop	dockbob
f2_60msPC1	1195(346.9)	1174(337.6)	1217(303.4)	1217(328.3)	1174(344.3)	1195(316.6)
f2_40msPC1	1239(368.1)	1217(359.6)	1239(328.1)	1260(349.8)	1217(366.1)	1260(336.8)
f2_20msPC1	1260(388.0)	1391(380.5)	1304(349.7)	1304(371.0)	1326(387.0)	1347(356.0)
f2_PC1	1391(409.5)	1456(401.0)	1456(373.1)	1391(392.5)	1456(407.0)	1391(375.3)
f3_60msPC1	2173	2217	2217	2238	2195	2195
f3_40msPC1	2130	2195	2152	2195	2152	2173
f3_20msPC1	2043	2195	2043	2152	2130	2130
f3_PC1	2152	2086	1934	2043	2065	2108
f2_OnC2V2	1413(676.5)	1152(692.2)	1608(677.8)	1652(601.9)	1347(722.0)	1130(564.2)
f2_20msOV2	1195(701.4)	1152(715.6)	1369(700.8)	1413(626.1)	1174(746.1)	1130(586.3)
f2_40msOV2	1195(715.7)	1174(742.2)	1239(726.8)	1282(640.7)	1130(774.1)	1152(612.0)
f2_60msOV2	1174(732.5)	1239(756.8)	1239(739.6)	1239(656.8)	1152(789.3)	1130(640.2)
f3_OnC2V2	2173	2325	2173	2564	2173	2238
f3_20msOV2	2217	2325	2043	2217	2217	2217
f3_40msOV2	2260	2304	2195	2195	2282	2282
f3_60msOV2	2304	756.8	2238	2173	2325	2325
dur_closure	74.4,110.5	59.8,85.6	38.6, 171.2	187.3	55.6,135.7	64.9,82.7
closure_2_V	253.6	272.8	294.0	199.1	313.1	177.9
release_2_V	75.5	41.3	30.0	10.6	85.4	7.5
burst	k(493.2-498.8) t(609.2-684.7)	k(475.9-562.0) p(647.6-688.9)	k(417.7-461.0) g(632.2-662.3)	k(0) d(584.5-596.3)	k(461.6-497.6) k(633.3-718.7)	k(446.5-467.8) b(550.5-558.0)
H1-H2,V1	43.1-52.6(341.1)	52.5-42.0(340.7)	53.6-45.0(320.3)	55.3-49.4(333.1)	53.1-48.1(334.6)	44.9-51.6(318.8)
V2	45.9-43.2(707.1)	44.8-42.7(747.3)	44.3-40.5(745.0)	45.5-42.2(644.2)	45.6-45.5(765.4)	47.1-45.3(631.8)
H1-A1,V1	43.1-58.6	52.5-51.2	53.6-56.5	55.3-59.1	53.1-57.2	44.9-57.1
V2	45.9-47.3	44.8-50.1	44.3-46.3	45.5-55.3	45.6-53.5	47.1-54.5

Speaker 3 dogtop dogpot doggot dogdot dogcop dogbob

f2_60msPC1 1152(380.9) 1195(375.8) 1152(359.8) 1152(410.9)1174(400.3)1174(509.4)
f2_40msPC1 1152(404.1) 1282(395.1) 1239(380.8) 1195(431.8)1217(421.6)1217(529.6)
f2_20msPC1 1217(424.0) 1434(419.3) 1326(402.7) 1239(453.9)1260(441.2)1391(549.6)
f2_PC1 1304(444.1) 1586(436.5) 1456(424.6) 1304(475.5)1282(460.2)1521(569.2)

f3_60msPC1 2195 2152,2803 2195 2152 2152 2434
f3_40msPC1 2217 2217,2803 2195 2130 2195 2434
f3_20msPC1 2217 2260,2782 2173 2173 2152 2369
f3_PC1 2217 2217,2760 2152 2238 2130 2195

f2_OnC2V2 1500(683.6) 1174(683.5) 1695(594.9) 1630(643.4)1347(689.3)1108(799.2)
f2_20msOV2 1174(707.6) 1152(707.5) 1413(616.6) 1456(665.1)1239(711.3)1130(825.7)
f2_40msOV2 1174(721.9) 1174(720.1) 1195(642.0) 1304(689.1)1174(737.9)1152(854.3)
f2_60msOV2 1174(737.5) 1195(733.7) 1174(669.5) 1260(715.6)1174(752.1)1130(869.1)

f3_OnC2V2 2391 2543 2130,2630 2412 2152 2260
f3_20msOV2 2304 2412 2130,2630 2065,2412 2195 2325
f3_40msOV2 2347 2304,2673 2195,2673 2195,2717 2217 2304
f3_60msOV2 2304 2325,2717 2238,2673 2238,2847 2217 2304

dur_closure 149.0 64.1,86.5 141.8 153.9 165.8 107.3,92.7
closure_2_V 225.8 247.8 161.7 164.2 216.1 206.3

release_2_V 76.4 169.2 19.8 13.7 50.2 8.0?

burst g() g(497.0-512.1) g(0) g(0) g(0) g(676.6-690.2)
t(602.8-679.2) p(598.5-681.3) g(572.3-585.9) d(627.5-635.6) c(636.1-686.3)b(775.0-783.0?)

H1-H2,V1
55.1-49.8(372.8)46.2-56.8(360.2)55.9-49.3(349.6)56.9-52.4(402.5)57.2-52.2(398.9)48.4-57.0(462.2)

V2
45.9-39.0(728.5)46.5-47.6(725.4)45.1-45.8(647.6)46.1-47.5(707.5)46.5-45.3(743.7) 43.3-48.2(874.0)

H1-A1,V1 55.1-58.9 46.2-65.3 55.9-61.4 56.9-61.3 57.2-65.2 48.4-62.3
V2 45.9-48.0 46.5-56.9 45.1-51.7 46.1-56.1 46.5-54.5 43.3-45.9

Speaker 3		gottop		gotpot		gotdot		gotcop	gotbob		dotgot
f2_60msPC1		1304(314.6)		1347(340.4)		1326(351.9)		1347(371.6)1347(383.0)1304(329.8)			
f2_40msPC1		1434(335.4)		1434(359.1)		1391(372.1)		1413(391.6)1434(403.8)1391(351.3)			
f2_20msPC1		1500(353.6)		1543(378.0)		1456(392.5)		1500(412.6)1456(426.3)1521(374.9)			
f2_PC1		1760(377.6)		1630(401.0)		1521(416.8)		1652(433.8)1565(452.2)1673(401.9)			
f3_60msPC1		2238		2238,2717	2869		2217		2217		2238
f3_40msPC1		2543		2195,2651	2847		2260		2195,2717	2217	
f3_20msPC1		2564		2586		2195,2738	2195		2847		2238
f3_PC1		2695		2738		2543		2217		2564	2564
f2_OnC2V2		1195(627.3)		1174(612.5)		1630(616.6)		1543(660.8)1239(613.5)1456(574.1)			
f2_20msOV2		1174(651.7)		1152(636.6)		1304(641.0)		1174(687.8) 1174(638.6)1239(599.6)			
f2_40msOV2		1152(665.8)		1195(649.5)		1195(656.2)		1152(702.8)1152(667.2)1195(614.7)			
f2_60msOV2		1174(680.0)		1282(679.5)		1217(691.9)		1174(718.6)1152(683.7)1195(630.2)			
f3_OnC2V2		2369,2782	2325		2630		2325		2217		2173
f3_20msOV2		2760		2304		2586		2217	2260		2152
f3_40msOV2		2412,2782	2304		2195,2630	2238		2347	2195		
f3_60msOV2		2412		2325		2217		2260	2369		2217
dur_closure		146.9		80.6,48.5	31.2,21.9	77.4,35.7	32.1,19.6	86.8,27.3			
closure_2_V		237.3		200.3		177.7		196.6	144.2		148.2
release_2_V		90.1		59.2		10.8		72.6	76.8		17.7
burst		t()		t(490.2-501.6)		t(463.6-578.5)		t(527.7-536.7)		t(493.8-508.0)	
	t(500.9-508.3)										
		t(534.0-624.4)		p(550.1-609.9)		d(601.1-610.1)		k(573.5-646.1)		b(527.5-604.4)	
	g(535.6-553.3)										
H1-H2,V1	44.8-53.0(317.3)47.3-55.8(327.9)43.0-50.8(354.6)42.8-50.1(375.1)43.3-51.5(367.1)								52.6-49.5(332.9)		
V2		45.7-43.1(668.0)46.6-48.1(655.1)41.8-43.6(665.4)45.1-43.1(708.8)41.9-43.1(708.8)								45.5-41.8(620.8)	
H1-A1,V1	44.8-54.7	47.3-58.9	43.0-50.4	42.8-50.3	43.3-53.5	52.6-53.0					
V2		45.7-42.5	46.6-51.4	41.8-44.6	45.1-48.2	41.9-44.9	45.5-44.9				

Speaker 3	mobtop	mobpot	mobgot	mobdot	mobcop	mobbob
f2_60msPC1	1130(352.8)	1152(343.1)	1108(358.9)	1108(406.8)	1108(365.5)	1152(396.0)
f2_40msPC1	1130(375.8)	1152(361.3)	1108(379.4)	1108(427.4)	1108(387.6)	1152(417.5)
f2_20msPC1	1152(397.6)	1130(379.3)	1108(400.1)	1108(447.8)	1152(408.6)	1130(438.5)
f2_PC1	1065(421.9)	1065(397.8)	1130(424.6)	1087(463.9)	1108(426.9)	1087(459.2)
f3_60msPC1	2391	2325	2543	2412	2369	2477
f3_40msPC1	2391	2347	2477	2456	2369	2369
f3_20msPC1	2369	2325	2391	2412	2347	2434
f3_PC1	2347	2325	2325	2391	2369	2282
f2_OnC2V2	1456(654.0)	1152(598.5)	1760(607.3)	1673(659.9)	1152(686.6)	1043(634.6)
f2_20msOV2	1174(677.7)	1152(622.5)	1434(628.9)	1456(681.8)	1130(710.8)	1087(654.2)
f2_40msOV2	1152(690.7)	1152(636.0)	1282(652.8)	1260(707.4)	1152(724.6)	1087(681.3)
f2_60msOV2	1130(720.1)	1195(650.0)	1239(678.3)	1217(722.1)	1130(740.3)	1108(608.7)
f3_OnC2V2	2260	2564	2152	2586	2304	2217
f3_20msOV2	2325	2499	2173	2695	2282	2282
f3_40msOV2	2347	2434	2304	2369	2347	2369
f3_60msOV2	2347	2391	2238	2369	2347	2434
dur_closure	159.9	122.8	53.1,91.8	132.7,38.1	71.1,41.3	159.7
closure_2_V	225.9	198.3	181.2	188.9	246.6	159.7
release_2_V	62.1	77.4	23.4	9.4	102.9	0
burst	b(0)	b(0)	b(477.2-483.2)	b(600.6-605.1)	b(508.0-534.5)	0
	t(589.1-651.2)	p(519.9-597.5)	g(575.1-598.5)	d(643.2-650.2)	k(575.8-678.7)	0
H1-H2,V1	56.9-50.5(367.6)	52.1-55.8(346.5)	54.6-51.6(391.3)	55.3-54.3(388.7)	56.2-49.9(357.3)	56.2-49.2(388.8)
V2	46.6-47.0(696.4)	46.2-48.2(641.7)	47.9-45.2(672.0)	48.6-44.6(728.6)	42.9-45.7(717.1)	46.5-44.1(713.0)
H1-A1,V1	56.9-57.3	52.1-60.0	54.6-56.2	55.3-60.1	56.2-59.1	56.2-49.5
V2	46.6-52.7	46.2-51.6	47.9-49.1	48.6-53.0	42.9-51.4	46.5-50.8

Speaker 3	toptot	toppot	topgot	topdot	topcop	topbob
f2_60msPC1	1260(384.9)	1260(347.4)	1260(311.9)	1260(377.6)	1260(335.5)	1239(306.3)
f2_40msPC1	1217(407.1)	1239(366.1)	1260(332.3)	1260(396.8)	1260(355.7)	1239(325.0)
f2_20msPC1	1260(429.4)	1217(384.8)	1239(355.0)	1260(415.5)	1217(376.8)	1217(343.5)
f2_PC1	1260(451.2)	1195(404.6)	1217(376.9)	1239(439.3)	1195(396.7)	1217(361.8)
f3_60msPC1	2282	2152	2282	2152	2217	2260
f3_40msPC1	2325	2195	2108,2760	2195	2282	2260
f3_20msPC1	2304	2282	2217	2217	2304	2282
f3_PC1	2217	2238	2238	2238	2282	2217
f2_OnC2V2	1521(655.2)	1152(656.7)	1630(573.1)	1521(634.1)	1456(652.4)	1108(618.0)
f2_20msOV2	1195(677.7)	1152(680.7)	1391(595.7)	1521(649.9)	1195(673.9)	1130(646.4)
f2_40msOV2	1174(689.9)	1217(708.3)	1260(621.6)	1282(666.8)	1174(699.8)	1130(661.5)
f2_60msOV2	1217(716.8)	1239(723.3)	1217(635.1)	1239(679.7)	1174(713.7)	1152(676.8)
f3_OnC2V2	2717	2499	2130,2717	2391	2173	2304
f3_20msOV2	2217	2434	1978	2456	2217	2325
f3_40msOV2	2217	2369	1999	2477	2325	2369
f3_60msOV2	2238	2325	1956	2173	2347	2369
dur_closure	85.9,53.9	185.9	82.2,63.3	177.4,20.0	50.9,104.7	227.6
closure_2_V	194.6	240.9	183.3	193.0	241.3	239.8
release_2_V	48.8	55.1	26.5	10.1	74.5	12.4
burst	p(542.9-547.9)	p(0)	p(464.1-475.3)	p(554.1-596.9)	p(458.4-469.1)	p(0)
	t(601.8-650.6)	p(597.7-652.8)	g(538.6-565.1)	d(617.1-627.1)	c(574.5-649.1)	b(594.9-607.4)
H1-H2,V1	53.5-45.8(388.9)	44.5-52.8(350.9)	51.9-49.3(330.3)	47.3-54.8(371.2)	52.1-49.9(337.8)	44.8-52.3(318.4)
V2	44.7-45.3(722.6)	46.7-47.0(699.3)	45.5-45.3(625.4)	46.7-46.3(698.6)	45.3-47.5(703.8)	40.8-47.6(683.4)
H1-A1,V1	53.5-45.8	44.5-51.7	51.9-52.6	47.3-54.3	52.1-52.3	44.8-55.6
V2	44.7-47.5	46.7-48.2	45.5-52.5	46.7-48.7	45.3-46.8	40.8-47.3

Appendix F: Speaker 4's Acoustic Data

Speaker 4	codtop	codpot	codgot	coddot	codcop	codbob
f2_60msPC1	1456(723.0)	1456(659.4)	1456(685.6)	1478(651.7)	1413(707.6)	1456(717.5)
f2_40msPC1	1543(744.6)	1478(680.2)	1586(707.4)	1565(672.0)	1521(729.6)	1500(738.9)
f2_20msPC1	1630(765.8)	1521(702.7)	1739(728.5)	1652(692.7)	1652(750.8)	1652(760.2)
f2_PC1	1695(786.6)	1630(722.7)	1847(750.4)	1739(713.8)	1782(772.0)	1739(781.0)
f3_60msPC1	2412	2499	2456	2564	2499	2499
f3_40msPC1	2477	2651	2521	2543	2412	2477
f3_20msPC1	2543	2651	2499	2564	2456	2477
f3_PC1	2651	2608	2695	2630	2369	2586
f2_OnC2V2	1586(981.4)	1239(950.1)	1760(913.9)	1695(887.6)	1500(969.8)	1174(969.2)
f2_20msOV2	1391(1002.5)	1239(972.0)	1630(936.6)	1608(905.2)	1478(987.0)	1195(991.7)
f2_40msOV2	1413(1022.5)	1282(993.8)	1586(954.5)	1521(924.3)	1413(1002.9)	1217(1009.1)
f2_60msOV2	1391(1044.0)	1369(1011.9)	1543(972.3)	1456(943.2)	1304(1022.8)	1239(1026.7)
f3_OnC2V2	2108	2347	2477	2499	2412	2238
f3_20msOV2	2434	2434	2217	2477	2369	2304
f3_40msOV2	2347	2434	2173	2391	2369	2304
f3_60msOV2	2412	2521	2369	2434	2412	2369
dur_closure	127.5	121.7	65.2,43.2	148.6	115.4	94.2,64.9
closure_2_V	188.6	221.1	151.8	165.1	187.7	173.3
release_2_V	61.2	99.3	37.2	16.2	72.8	10.5
burst	d() t(920.1-981.2)	d() p(846.9-946.3)	d(822.1-828.3) g(871.5-879.6)	d(0) d(867.4-883.9)	d() k(892.4-964.7)	d(880.4-884.1) b(-959.5)
H1-H2,V1	57.8-52.8(716.9)	59.5-56.2(643.6)	56.6-55.1(663.2)	57.3-53.8(639.3)	56.5-54.8(677.5)	59.5-56.3(697.9)
V2	56.6-47.2(1042.5)	53.9-53.8(1028.9)	54.0-54.0(991.7)	52.9-53.1(979.8)	52.1-50.1(1015.1)	55.1-54.2(1059.4)
H1-A1,V1	57.8-53.6	59.5-54.6	56.6-56.6	57.3-54.5	56.5-55.9	59.5-56.4
V2	56.6-48.7	53.9-55.7	54.0-55.8	52.9-52.1	52.1-49.1	55.1-56.6

Speaker 4	docktop	dockpot	dockgot	dockdot	dockcop	dockbob
f2_60msPC1	1217(584.4)	1239(656.4)	1391(670.3)	1260(628.8)	1495(658.4)	1304(742.9)
f2_40msPC1	1239(608.0)	1282(679.1)	1369(691.3)	1239(651.5)	1413(677.6)	1217(765.8)
f2_20msP C1	1347(630.5)	1413(701.6)	1434(712.7)	1434(673.0)	1413(697.6)	1391(788.4)
f2_PC1	1391(652.6)	1391(723.2)	1456(738.1)	1456(694.0)	1521(708.3)	1413(810.2)
f3_60msPC1	2477	2477	2543	2477	2695	2564
f3_40msPC1	2521	2434	2499	2456	2564	2477
f3_20msPC1	2456	2499	2499	2325	2695	2369
f3_PC1	2347	2412	2477	2412	2630	2195
f2_OnC2V2	1543(912.1)	1282(969.4)	1565(994.3)	1543(941.5)	1434(962.3)	1152(1071.5)
f2_20msOV2	1326(929.2)	1260(987.4)	1478(1017.1)	1478(962.2)	1391(980.6)	1195(1092.5)
f2_40msOV2	1217(946.8)	1260(1010.1)	1434(1040.0)	1413(979.1)	1391(998.5)	1217(1111.8)
f2_60msOV2	1260(967.4)	1326(1028.6)	1326(1058.1)	1347(997.3)	1369(1016.5)	1217(1132.4)
f3_OnC2V2	2325	2369	2043	2499	2391	2260
f3_20msOV2	2434	2391	2217	2586	2456	2173
f3_40msOV2	2521	2456	2238	2130	2456	2369
f3_60msOV2	2456	2456	2434	2238	2434	2260
dur_closure	80.1,50.4	86.4,61.5	74.2,80.4	88.6,60.8	73.9,94.0	92.8,109.3
closure_2_V	240.6	231.5	240.9	230.7	245.4	244.4
release_2_V	58.1	46.4	118.5	37.0	61.7	18.5
burst	k(747.3-799.4) t(849.8-907.8)	k(821.1-858.5) p(920.0-966.2)	k(819.9-868.5) g(949.0-986.6)	k(795.0-839.4) d(900.2-937.1)	k(787.5-803.5) k(897.5-959.1)	k(916.8-940.5) b(1049.8-1068.3)
H1-H2,V1	57.7-55.3(576.8)	57.1-55.3(640.2)	55.2-52.4(649.6)	55.8-53.8(618.4)	58.6-53.5(626.9)	59.1-57.4(740.8)
V2	52.7-54.3(976.1)	49.3-53.4(1055.5)	54.3-52.8(1074.7)	54.1-56.8(1027.7)	51.3-54.2(1018.2)	50.6-2.9(1142.0)
H1-A1,V1	57.7-55.0	57.1-54.6	55.2-52.1	55.8-51.1	58.6-53.3	59.1-56.1
V2	52.7-50.1	49.3-53.6	54.3-50.9	54.1-52.1	51.3-51.3	50.6-47.9

Speaker 4	dogtop	dogpot	doggot	dogdot	dogcop	dogbob
f2_60msPC1	1217(612.8)	1195(642.2)	1195(728.9)	1217(609.8)	1260(653.7)	1174(566.7)
f2_40msPC1	1217(634.9)	1217(663.7)	1217(750.4)	1239(631.3)	1326(675.9)	1217(588.9)
f2_20msPC1	1282(656.5)	1282(685.2)	1260(773.2)	1391(653.0)	1413(698.5)	1239(610.8)
f2_PC1	1369(677.8)	1391(706.6)	1282(794.9)	1413(673.5)	1413(719.6)	1217(633.1)
f3_60msPC1	2434	2586	2564	2477	2434	2586
f3_40msPC1	2434	2608	2521	2477	2543	2564
f3_20msPC1	2412	2651	2304	2412	2586	2521
f3_PC1	2391	2456	2347	2217	2608	2630
f2_OnC2V2	1586(916.4)	1304(954.6)	1586(944.3)	1673(858.8)	1521(929.6)	1195(806.8)
f2_20msOV2	1478(937.0)	1217(975.7)	1565(962.8)	1565(881.8)	1478(952.2)	1239(828.5)
f2_40msOV2	1456(959.3)	1326(992.3)	1543(985.3)	1434(900.2)	1391(974.4)	1282(852.1)
f2_60msOV2	1413(976.9)	1369(1008.4)	1521(1002.7)	1347(918.0)	1326(992.4)	1195(869.8)
f3_OnC2V2	2456	2564	2499	2586	2456	2586
f3_20msOV2	2434	2477	2347	2608	2195	2586
f3_40msOV2	2456	2477	2456	2217	2260	2586
f3_60msOV2	2499	2499	2434	2325	2304	2608
dur_closure	138.1,42.7 52.9,80.3	103.3	44.0,77.3 128.7		145.1	
closure_2_V	221.7	231.1	133.0	172.6	198.6	162.1
release_2_V	95.4	84.2	30.2	21.0	69.9	17.1
burst	g(747.6-773.8) t(816.5-911.9)	g(772.8-786.4) p(950.8-950.9)	g() g(906.4-936.2)	g(726.0-756.3) d(833.6-854.5)	g() k(855.5-925.4)	g() b(785.8-802.9)
H1-H2,V1	57.4-54.9(594.3)	58.3-54.9(619.9)	58.0-56.6(685.4)	57.8-56.3(570.6)	60.2-55.9(629.5)	58.7-56.9(560.3)
V2	55.5-54.4(961.7)	56.5-49.6(1026.2)	56.6-54.4(1029.0)	51.0-54.3(944.0)	54.6-55.2(988.0)	56.3-55.7(886.0)
H1-A1,V1	57.4-57.4	58.3-55.1	58.0-56.7	57.8-56.2	60.2-52.8	58.7-53.8
V2	55.5-54.7	56.5-53.3	56.6-52.4	51.0-54.2	54.6-55.7	56.3-54.8

Speaker 4	gottop	gotpot	gotdot	gotcop	gotbob	dotgot
v1f1	695	695	674	695	674	674
f2_60msPC1	1369(631.8)	1391(582.8)	1369(675.1)	1413(576.1)	1304(605.1)	1369(699.5)
f2_40msPC1	1347(653.1)	1369(603.8)	1326(696.5)	1413(596.4)	1304(627.0)	1413(720.0)
f2_20msPC1	1326(674.4)	1434(624.6)	1347(719.4)	1434(616.2)	1391(649.2)	1434(739.4)
f2_PC1	1369(694.8)	1391(644.4)	1413(738.0)	1434(636.6)	1434(670.0)	1500(756.7)
f3_60msPC1	2434	2521	2304	2477	2456	2543
f3_40msPC1	2521	2564	2521	2543	2521	2543
f3_20msPC1	2543	2543	2521	2521	2564	2521
f3_PC1	2521	2564	2499	2477	2564	2499
v2f1	674	674	674	717	674	695
f2_OnC2V2	1478(933.9)	1174(877.2)	1739(982.4)	1456(883.6)	1195(911.0)	1565(966.8)
f2_20msOV2	1521(956.0)	1217(898.9)	1717(1000.3)	1391(900.2)	1195(928.3)	1456(988.8)
f2_40msOV2	1456(975.4)	1239(921.4)	1608(1024.5)	1369(917.7)	1195(945.6)	1347(1011.3)
f2_60msOV2	1282(993.9)	1304(943.6)	1695(1047.1)	1304(935.1)	1195(962.7)	1478(1034.2)
f3_OnC2V2	2564	2499	2586	2434	2499	2543
f3_20msOV2	2477	2543	2369	2434	2543	
f3_40msOV2	2412	2586	2195	2412	2608	2195
f3_60msOV2	2521	2543	2456	2434	2586	2369
dur_closure	157.4	134.4	142.3,58.9 162.3		210.2	174.7
closure_2_V	232.8	224.3	229.9	232.1	226.1	200.6
release_2_V	75.3	90.0	18.4	69.8	15.9	25.8
burst	t()	t()	t(885.8-896.2)	t()	t()	t()
	t(858.5-933.9)	p(783.5-873.4)	d(955.1-973.4)	c(805.0-874.8)	b(885.8-901.7)	g(936.3-962.2)
V1: h1	196 Hz 56.8	196 Hz 58.0	174 Hz 59.0	196 Hz56.6	174 Hz 56.2	196Hz,57.8
h2	369 Hz 52.9	369 Hz 52.6	369 Hz 54.8	391 Hz51.2	348 Hz 53.5	391Hz,51.7
V2: h1	174 Hz 60.0	174 Hz 59.1	174 Hz 58.8	174 Hz54.8	174 Hz 57.3	174Hz,59.1
h2	348 Hz 55.9	348 Hz 54.3	348 Hz 51.9	348 Hz54.1	348 Hz 55.2	348Hz,53.0
V1: a1	739 Hz 51.8	739 Hz 49.8	717 Hz 51.8	587 Hz52.6	717 Hz 52.0	761Hz,52.4
V2: a1	695 Hz 53.8	717 Hz 52.6	674 Hz 49.1	674 Hz53.0	695 Hz 53.5	695Hz,48.0

Speaker 4	mobtop	mobpot	mobgot	mobdot	mobcop	mobbob
f2_60msPC1	1239(630.6)	1195(751.7)	1195(651.3)	1260(793.2)	1174(693.6)	1195(721.6)
f2_40msPC1	1260(652.9)	1217(773.2)	1195(673.8)	1239(815.2)	1174(715.5)	1195(744.4)
f2_20msPC1	1282(675.2)	1239(795.7)	1217(696.0)	1217(837.7)	1239(737.6)	1217(766.2)
f2_PC1	1239(697.5)	1152(816.7)	1174(718.2)	1152(860.4)	1217(759.1)	1217(787.5)
f3_60msPC1	2630	2717	2630	2651	2630	2825
f3_40msPC1	2630	2782	2608	2586	2673	2869
f3_20msPC1	2630	2782	2803	2586	2608	2912
f3_PC1	2630	2782	2760	2543	2760	2912
f2_OnC2V2	1782(902.9)	1239(1109.1)	1500(902.8)	1760(1048.4)	1500(980.2)	1217(958.7)
f2_20msOV2	1369(923.0)	1195(1130.6)	1456(921.3)	1543(1071.4)	1413(1001.1)	1239(975.3)
f2_40msOV2	1434(940.6)	1174(1156.5)	1391(938.6)	1456(1088.7)	1369(1018.1)	1282(992.6)
f2_60msOV2	1500(957.6)	1195(1177.8)	1304(956.8)	1369(1106.8)	1217(1036.4)	1217(1010.2)
f3_OnC2V2	2477	2391	2543	2521	2369	2434
f3_20msOV2	2412	2630	2499	2456	2434	2499
f3_40msOV2	2456	2521	2304	2499	2434	2499
f3_60msOV2	2434	2499	2173	2477	2456	2564
dur_closure	71.8,56.3	67.2,131.3	84.5,45.8	70.7,73.2	70.5,71.7	145.1
closure_2_V	196.8	282.6	168.8	175.9	211.7	154.6
release_2_V	56.4	71.0	30.2	22.9	58.4	9.3
burst	b(771.6-783.9)	b(889.1-902.0)	b(815.2-823.5)	b(938.3-947.3)	b(837.6-848.4)	b(0)
	t(840.2-896.6)	p(1033.3-1104.4)	g(869.3-899.5)	d(1020.5-1043.5)	k(920.1-978.9)	b(943.7-953.3)
H1-H2,V1	60.0-53.3(599.4)	59.3-52.3(734.9)	60.5-55.3(625.6)	59.3-55.5(768.1)	60.5-53.1(663.2)	57.4-52.6(690.0)
V2	57.8-53.9(957.8)	57.8-54.1(1188.9)	50.5-52.8(996.0)	52.1-54.4(1144.5)	55.3-51.4(1042.5)	56.0-53.0(1042.5)
H1-A1,V1	60.0-54.0	59.3-52.9	60.5-52.8	59.3-55.2	60.5-53.8	57.4-52.0
V2	57.8-53.1	57.8-53.3	50.5-49.8	52.1-51.2	55.3-51.0	56.0-51.3

Speaker 4	toptot	toppot	topgot	topdot	topcop	topbob
v1f1	652	674	695	652	674	674
f2_60msPC1	1260(689.4)	1304(744.9)	1195(645.8)	1239(688.0)	1260(675.2)	1304(724.3)
f2_40msPC1	1304(709.5)	1282(765.2)	1239(666.1)	1239(707.6)	1260(695.9)	1304(744.8)
f2_20msPC1	1260(729.7)	1282(785.7)	1217(687.7)	1195(727.0)	1282(715.8)	1369(766.2)
f2_PC1	1152(749.1)	1195(806.6)	1369(708.2)	1152(745.1)	1195(739.8)	1217(786.5)
f3_60msPC1	2630	2586	2608	2543	2477	2456
f3_40msPC1	2630	2586	2586	2521	2499	2477
f3_20msPC1	2608	2586	2673	2412	2564	2543
f3_PC1	2543	2564	2630	2543	2521	2586
v2f1	674	674	674	674	674	674
f2_OnC2V2	1673(1017.8)	1195(1103.4)	1521(967.1)	1543(1005.3)	1391(1018.5)	1174(1003.1)
f2_20msOV2	1239(1039.0)	1195(1123.8)	1500(987.1)	1543(1027.5)	1347(1035.9)	1217(1025.1)
f2_40msOV2	1282(1061.6)	1174(1144.5)	1456(1008.6)	1369(1049.0)	1304(1057.3)	1239(1047.5)
f2_60msOV2	1326(1083.9)	1456(1166.2)	1326(1026.0)	1260(1070.1)	1217(1074.5)	1217(1069.9)
f3_OnC2V2	2608	2543	2412	2304	2543	2456
f3_20msOV2	2369	2499	2412	2477	2543	2499
f3_40msOV2	2477	2543	2391	2412	2369	2456
f3_60msOV2	2456	2564	2477	2543	2499	2499
dur_closure	105.7,49.8 166.3	111.8,77.2	114.6,79.6 111.4,78.6	191.7		
closure_2_V	257.8	286.6	246.1	242.9	270.9	207.5
release_2_V	94.1	120.2	35.6	36.4	73.6	15.8
burst	p(866.6-874.8)	p()	p(828.2-849.7)	p(868.9-881.4)	p(855.5-862.9)	p()
	t(924.6-1018.7)	p(979.3-1099.7)	g(926.9-962.4)	d(961.0-997.2)	k(941.5-1015.0)	b(983.8-999.6)
V1: h1	196Hz,60.9	196Hz,57.3	196 Hz 59.6	196 Hz59.0	196 Hz59.2	196Hz,60.2
h2	391Hz,53.0	391Hz,53.0	369 Hz 53.9	413 Hz52.7	391 Hz52.1	391Hz,52.8
V2: h1	174Hz,62.0	196Hz,58.3	174 Hz 57.6	196 Hz57.2	174 Hz60.9	174Hz,59.2
h2	348Hz,54.5	369Hz,52.7	348 Hz 53.0	283 Hz51.1	348 Hz55.0	348Hz,54.3
V1: a1	587Hz,55.9	587Hz,54.5	761 Hz 52.5	609 Hz57.8	587 Hz53.8	587Hz,53.7
V2: a1	17Hz,52.1	761Hz,52.9	717 Hz 48.9	674 Hz51.5	717 Hz51.7	717Hz,49.1

Appendix G: Average Voicing Characteristics Across Individuals

Table 1: Order of Voicing vs. H1*-H2* for V1 in V1C#CV2 Sequences

	voiced-voiced	**voiced**-voiceless	**voiceless**-voiced	**voiceless**-voiceless
avg.	-0.17	1.07	0.027	0.845
male avg.	-6.37	-4.35	-7.04	-5.77
female avg.	6.03	6.5	7.09	7.45
m1	-9.24	-4.97	-8.86	-8.42
m2	-3.5	-3.72	-5.22	-3.11
f1	5.33	6.79	7.51	8.99
f2	6.73	6.2	6.67	5.9

Table 2: Order of Voicing vs. H1*- A1 for V1 in V1C#CV2 Sequences

	voiced-voiced	**voiced**-voiceless	**voiceless**-voiced	**voiceless**-voiceless
avg.	-6.35	-5.49	-4.99	-4.91
male avg.	-12.03	-10.67	-11.53	-11
female avg.	-0.67	-0.3	1.55	1.175
m1	-12.22	-9.31	-11.56	-11.39
m2	-11.84	-12.02	-11.5	-10.6
f1	1.41	2.6	5.1	4.95
f2	-2.74	-3.21	-2	-2.6

Table 3: Order of Voicing vs. H1*-H2* for V2 in V1C#CV2 Sequences

	voiced-**voiced**	voiceless-**voiced**	voiced-**voiceless**	voiceless-**voiceless**
avg.	1.62	1.97	3.68	3.72
male avg.	0.02	-1.95	1.585	1.37
female avg.	2.88	7.3	5.14	4.6
m1	2.07	-1.49	1.69	0.17
m2	-2.03	-2.4	1.48	2.57
f1	1.5	5.84	5.39	5.6
f2	4.26	8.75	4.88	3.6

Table 4: Order of Voicing vs. H1*- A1 for V1 in V1C#CV2 Sequences

	voiced-**voiced**	voiceless-**voiced**	voiced-**voiceless**	voiceless-**voiceless**
avg.	-6.03	-3.66	-1.96	-0.38
male avg.	-9.3	-8.35	-6.2	-3.9
female avg.	-2.77	1.03	2.26	3.2
m1	-4.09	-5.88	-4.8	-2.59
m2	-14.5	-10.82	-7.55	-5.25
f1	-1.79	5.74	4.23	6.4
f2	-3.75	-3.69	0.28	-0.07