Analysis of the Pressure-wall Interaction at the Release of a Stop Closure

by

Lan Chen

B.Eng. Aerodynamics
Beijing University of Aeronautics and Astronautics, 1995
M.S. Acoustics
Institute of Acoustics, Chinese Academy of Sciences, 1998

SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN AERONAUTICS AND ASTRONAUTICS
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MARCH 2011

Signature of Author: _____

Lan Chen
Department of Aeronautics and Astronautics
May 11, 2011

Certified by: _____

Wesley L. Harris
Professor of Aeronautics and Astronautics
and Associate Provost, Faculty Equity
Thesis Supervisor

Accepted by: _____

Eytan H. Modiano
Associate Professor of Aeronautics and Astronautics
Chair, Graduate Program Committee

Analysis of the Pressure-wall Interaction at the Release of a Stop Closure

by

Lan Chen

Submitted to the Department of Aeronautics and Astronautics
on May 11, 2011 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Aeronautics and Astronautics

Abstract

In producing a stop consonant, a soft tissue articulator, such as the lower lip, the tongue tip, or the tongue body, is raised to make an airtight closure. Stevens [1] pp 329-330 hypothesized that the interaction of the air pressure with the yielding soft-tissue wall would lead to a plateau-shaped release trajectory, and the duration of the plateau is progressively longer for bilabial, alveolar, and velar (Fig. 1–1). This thesis analyzes the pressure-wall interaction when a stop closure is released. Three flow models are implemented to derive the release trajectory: quasi-steady incompressible, unsteady incompressible, and unsteady compressible flow. Results from the models confirm Stevens' hypothesis. In the unsteady flow models, this thesis contributes a new method – deformable control volume analysis – to the pressure-wall interaction for small openings. This method may also be applied to quantify the unsteady effect during the closing and opening of the vocal folds and during the initial transient phase of a stop consonant, when the cross-sectional area is small. Indirect means of measuring an unknown parameter in the pressure-wall interaction analysis is discussed with the aid of a closure model which derives the condition of retaining a complete closure against air pressure buildup. In comparison with real speech data, an acoustic measure is defined for determining the duration of the frication noise of voiceless alveolar and velar stop consonants in syllable-initial positions. This newly defined measure is based on the time variation of the average FFT magnitude in the whole frequency range and the magnitude in a 50-Hz-wide frequency band containing the front cavity resonance for the signal in every 5 milliseconds (a moving 5-ms window). This measure is found applicable to 25 releases out of 32 releases from TIMIT database. The means of the collected durations are found closest to the estimated duration calculated with the unsteady compressible flow model.

Thesis Supervisor: Wesley L. Harris
Title: Professor of Aeronautics and Astronautics and Associate Provost, Faculty Equity

Acknowledgements

I would like to express my sincere thanks to my thesis supervisor Prof. Wesley L. Harris for his kindly support, practical help, and effective guidance during the course of completing this thesis. I would also like to express my sincere thanks to Prof. Kenneth, N. Stevens for his support and guidance.

I am also very grateful to all my friends and all the people who have supported me and helped me during this time. They are cherished forever in my heart.

At last, I would like to thank MIT Office of Graduate Education for partial financial support.

Contents

List of figures

List of tables

List of symbols

A_c	Cross-sectional area of the supraglottal constriction
A_{c_start}	Cross-sectional area of the supraglottal constriction at the time of release
A_g	Cross-sectional area of the glottal constriction
CV	Control volume
D	Length of the supraglottal constriction along the dimension perpendicular to the midsagittal plane
d	Derivative
F_{P_1}	Average pressure force acting on the upper mass
F_{kc}	Force generated by the spring connecting the two masses
f_{P_m}	Horizontal pressure force
f_k	Kinetic frictional force
f_s	Static frictional force between the contacting surfaces
h_c	Downward displacement of the tongue surface
k	Lumped spring constant or the stiffness of the soft tissue surface
k_c	Spring constant connecting the two masses
L	Inductance
L_c	Axial length of the supraglottal constriction
m	Lumped mass of the soft tissue surface
P_{c_rel}	Contact pressure at the time of release
P_e	Static pressure at the exit of the tube
P_i	Static pressure at the inlet of the tube
P_m	Intraoral pressure
P_s	Subglottal pressure
p_x	Static pressure at the location x
r	Lumped damping of the soft tissue surface
t	Time
U	Volume velocity going through the supraglottal constriction
U_g	Volume velocity going through the glottis
u_x	Flow velocity at the location x
V	Constant release velocity
V	Instantaneous downward velocity of the upper mass
y_1	Displacement of the upper mass
y_{eq1}	Equilibrium position of the upper mass
y_2	Displacement of the lower mass

y_{eq2}	Equilibrium position of the lower mass
ρ	Density of the air
μ_m	Maximum static frictional coefficient
μ_k	Kinetic frictional coefficient

Chapter 1

Introduction

1.1 Research background

Speech is the most convenient way of human communication. A person may not realize any difficulty in speaking his/her own native language, but when he/she starts to learn a foreign language, production problem occurs. Some patients with diseases such as Parkinson's and ALS (*Amyotrophic lateral sclerosis*), take even greater effort in producing an understandable speech sound. He/she needs to know how and be able to produce the correct sound so that others can understand the word being expressed.

Speech is also the most convenient means of man-machine communication. In this task, a computer is expected to recognize human speech accurately, which is called automatic speech recognition; and also to "speak" as naturally as a human speaker, which is realized by speech synthesis.

Human speech is powered by the pulmonary pressure in the lung. The air coming out of the lung interacts with various structures along the vocal tract (the air way from the throat to the lips), and generates the sound sources, which are then filtered by the cavity resonances determined by the shape of the vocal tract when the particular speech sound is produced. Both generating the sound sources and shaping the vocal tract require the coordination of multiple respiratory and articulatory structures.

As a special type of sound, speech is the subject studied in speech production, a discipline under acoustics. Acoustics studies the generation, propagation, and reception of sound. Similarly, speech production studies speech sound generation and propagation along the vocal tract of a human subject. The perception of speech sound is studied in speech perception, or psychoacoustics. Based on the physics of speech sound generation and propagation, mathematical models are developed as tools for quantifying the physiological activities involved in producing a speech sound.

Speech production models also provide a knowledge-based approach to identify features of the speech sound for speech recognition [2], and supply rules for speech synthesis [3].

Speech sounds are composed of two primary categories: vowels and consonants. In the production of a vowel, vocal fold vibration is present and the sound source is located at the laryngx. In producing a consonant, the vocal folds do not vibrate, or vibrate in a more moderate mode compared with the vibration in producing a vowel.

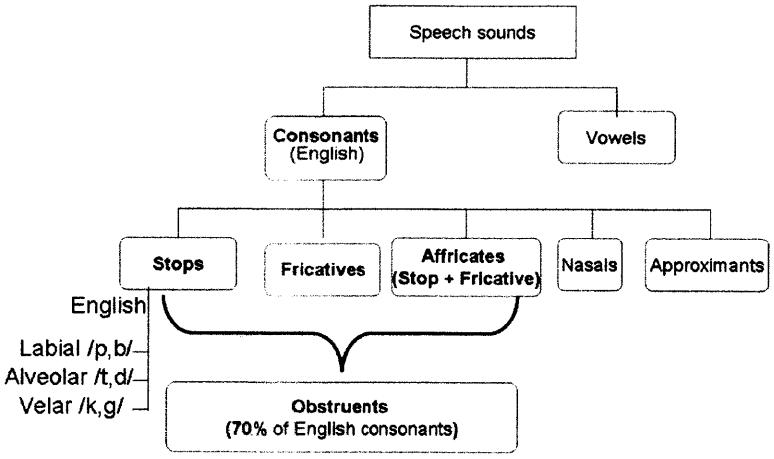


Fig. 1–1 Concept Map - Speech sounds are composed of vowels and consonants. English consonants can be sub-categorized into stops, fricatives, affricates, nasals, and approximants, according the manner of articulation. Among them, stops, fricatives and affricates are also called obstruents, and this name implies obstruction of air in their production. In English, 70 percent of consonants are obstruents.

Effective use of the vowel models started from Chiba and Kajiyama [4] in 1941 for the Japanese language, and the vowel models for the English language were first reported by Jakobson, Fant, and Halle [5] in 1952. Later, Fant's work in 1960 [6], "Acoustic Theory of Speech Production", established the theoretical framework widely and currently used.

Compared to the vowel models, slow advance has been made in the consonant models [7] pp29-30. One of the obstacles, considered by Fant, is the difficulty involved in modeling all the relevant factors in the acoustic production process.

According to the way of producing the sound (also called manner of articulation by phoneticians), English consonants can be further grouped into stop consonants, fricatives, affricates, nasals, and approximants, as shown in the concept map in Fig. 1–1. Among them, stop consonants, fricatives, and affricates are also called obstruent consonants. The name "obstruent" means that obstruction of air exists in their production. In English, 70 percent of consonants belong to obstruent consonants.

In this group with the largest number of consonants, stop consonants (/p, b, t, d, k, g/) and fricatives (/f, v, s, z, ʃ, ʒ, θ, ð/) comprise the majority, and the remnant of two affricates (/ʧ, ʤ/) can be taken as the combination of a stop consonant and a fricative.

The articulation and the acoustics of a stop consonant /t/ and a fricative /s/ are compared in Fig. 1-2. The midsagittal profile of the vocal tract in producing the closure interval of /t/ is shown in Fig. 1–2a, and the profile in producing /s/ is shown in Fig. 1–2b. A complete closure is made with the tongue tip for /t/ and a constriction is made for /s/, both at the alveolar ridge. Air pressure is being built up behind either the closure or the constriction.

The acoustic waveform of utterance /ata/ and /asa/ are shown in Fig. 1–2c and Fig. 1–2d respectively. The closure interval in the acoustic waveform of /t/ is indicated in Fig. 1–2c, corresponding to the articulation profile shown in Fig. 1–2a. No sound is supposed to be produced during this interval, and the signal recorded in the acoustic waveform belongs to the background noise. When the closure is released, a brief interval of turbulence noise is generated immediately from the air stream coming out from the released closure. This interval is labeled as "/t/" in the acoustic waveform in Fig. 1–2c. In producing the fricative /s/, an air stream goes through the constriction shown in Fig. 1–2b, and creates continuous turbulence noise. This interval is indicated by the label "/s/" in the acoustic waveform in Fig. 1–2d. Both acoustic waveforms demonstrate that turbulence noise is characteristic of obstruent consonants. A stop consonant distinguishes itself from a fricative at the onset of the turbulence noise, where an initial transient is generated first. The transient looks like an impulse in the waveform (Fig. 1–2c) at the beginning of the /t/ interval.

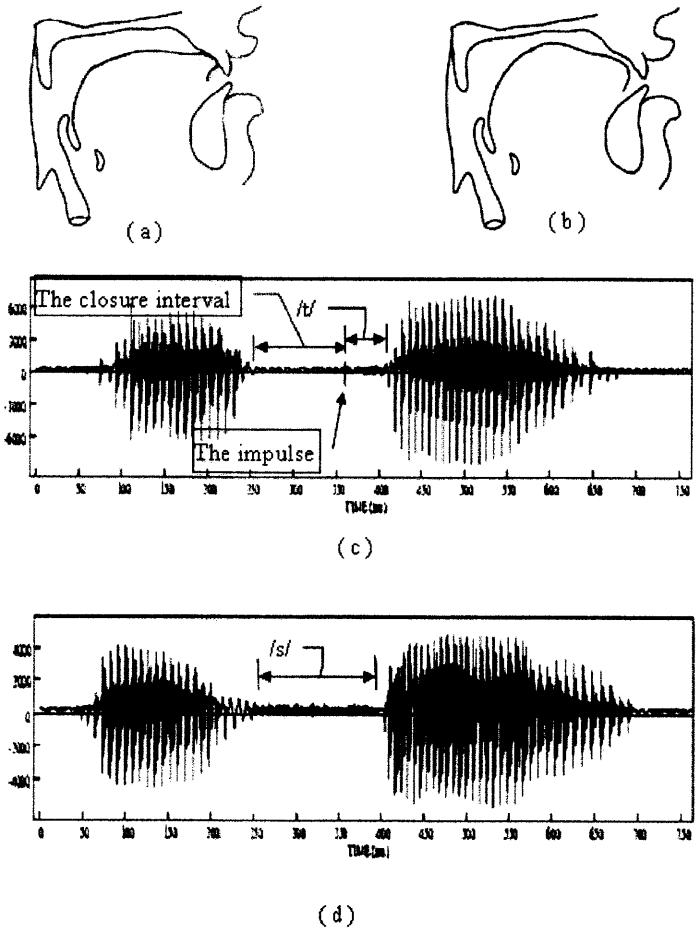


Fig. 1–2 (a) Midsagittal section of the vocal tract when a stop consonant /t/ is produced [1]. The tongue blade is making a complete closure at the alveolar ridge. (b) Midsagittal section of the vocal tract when a fricative /s/ is produced [1]. The tongue blade is making a constriction at the alveolar ridge. (c) Acoustic waveform of the utterance /ata/. The closure interval, the interval of the /t/ sound, and the impulse-like transient are labeled. (d) Acoustic waveform of the utterance /asa/. The interval belonging to the /s/ sound is labeled.

As the name "obstruent" suggests, obstruent consonants are produced with obstruction of air in the vocal tract. In the course of obstructing the air, a soft tissue articulator (it is often the primary articulator utilized in producing the consonant) such as the tongue blade, the tongue dorsum, or the lower lip, is always involved to make a

complete closure with the target surface in producing a stop consonant, or to create a constriction in the airway in producing a fricative. Air pressure is then built up in the vocal tract, and this increase in air pressure is necessary for generating the turbulence noise characteristic of these consonants.

Stevens discussed the role of the yielding surface of soft-tissue articulators in the articulatory models of stop and fricative consonants [1] pp 325, 328-329, 382 respectively. He created a model of the interaction of steady flow with the yielding wall in fricative production, which showed that the compliant surface makes it easier for the speaker to generate the turbulence noise with the maximum intensity at the supraglottal constriction[1] pp110-111. He also considered the pressure-wall interaction during the release of a stop closure, and his hypothesis serves as the motivation of this thesis, as discussed in the next section.

1.2 Motivation: Stevens' hypothesis

Stevens [1] pp 329-330 suggested that the pressure-wall interaction would lead to a plateau-shaped release trajectory shown in solid line in Fig.1–3a. He also illustrated schematically in the panels in Fig.1–3b the hypothesized sequences of events after a velar closure is released.

The mechanism underlying his hypothesis was that the pressurized air behind the closure may cause the closure to be released (in solid line) earlier than the time when it would be without the influence of the intraoral pressure buildup (in dashed line). After the release, owing to the decrease in intraoral pressure, the compliant tissue surface would bounce back, and a plateau in the release trajectory could be retained for a while when the tongue dorsum keeps moving downward. The slower this downward motion, the longer the plateau could be retained.

Hanson and Stevens [8] found perceptual evidence for this hypothesis: the stop consonants were perceptually more "acceptable" when they were synthesized with a plateau progressively longer for labial, alveolar, and velar in the release trajectories, and they had acoustic characteristics matching better to normal speech.

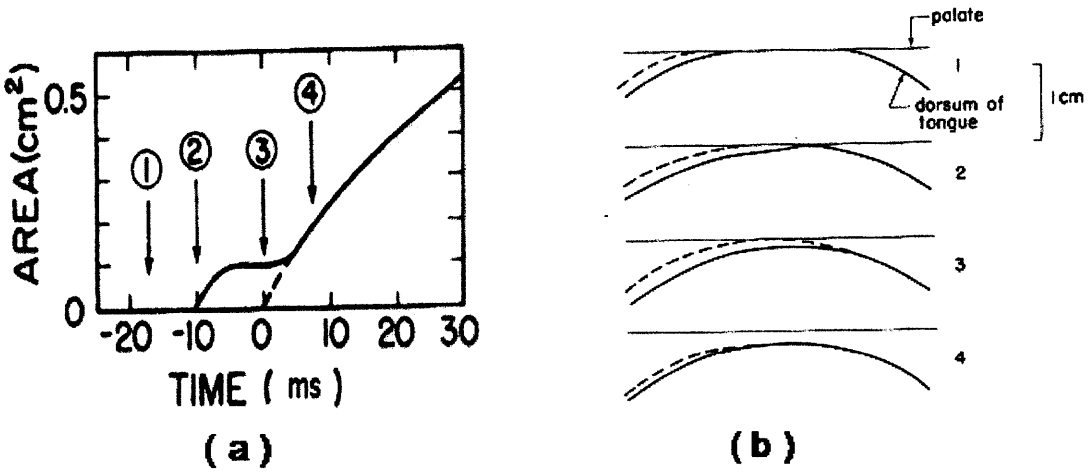


Fig.1–3 (a) The plateau-shaped release trajectory under the influence of the intraoral pressure (solid line) and the straight release trajectory without the influence of the intraoral pressure (dashed line). The labels with arrow indicate the time corresponding to the release sequences showed in the panels in (b)[1] pp329. (b) Hypothesized sequences before and after the release of a velar closure under the influence of the intraoral pressure (solid line) and without the influence of the intraoral pressure (dashed line). Owing to the intraoral pressure, the closure could be released before the time when the tongue dorsum leaves the palate without the influence of the intraoral pressure [1] pp329.

However, it is difficult to measure this plateau-shaped release trajectory directly because (1) the pressure-wall interaction may last for very short time, on the order of tens of milliseconds; (2) the dynamic range of the displacement of the yielding wall is so small that a good space resolution at the tissue boundary is required; (3) the systems capable of tracking articulatory movements with X-ray or EM-waves can not capture the release trajectories right at the closure, as the sensors are intentionally placed away from the location of the closure, in order to prevent interference with speech production[9].

This situation fits what Fant [7] pp29-30 commented about the slow advance in consonant models compared with vowel models. He considered that "the obstacles are our lack of reliable data on the details of the vocal tract and reliable physiological data

with respect to consonants, and also the difficulty involved in modeling all relevant factors in the acoustic production process".

1.3 Objective and thesis outline

We initially aim to test Stevens' hypothesis by developing a mathematical model of the pressure-wall interaction during the release of a stop closure. This effort would contribute to a more complete scientific pursuit of the physiology of consonant production.

Relevant knowledge in speech production is introduced in the next section, which includes speech aerodynamics, mechanical properties of the compliant tissue walls, turbulence noise source in speech production, the place of articulation, and the acoustics of stop consonants.

Previous studies on the pressure-wall interaction in speech production are reviewed in Chapter 2. Stevens' fricative model mentioned in Section 1.1 is introduced first, which studied the type of consonants closest to stop consonants. Another consonant model, McGowan's tongue tip trill model, is introduced next, which studied the interaction of an unsteady flow the yielding wall. At last, vocal fold models are introduced.

New models developed in this thesis for the pressure-wall interaction in stop consonant production are discussed in Chapter 3. The results calculated with these models are presented in Chapter 1. Comparison of the model results and relevant acoustic data are discussed in Chapter 5, and also an unknown parameter in the models is discussed.

1.4 Relevant knowledge in speech production: speech aerodynamics, compliant walls, turbulence noise source, the place of articulation, and the acoustic pattern of stop consonants

In this section, some concepts and knowledge in speech production theory are introduced. They are referred to in the following chapters.

1.4.1 Speech aerodynamics and compliant walls

Speech aerodynamics studies the pressure and flow along the vocal tract, the air way between the throat and the lips, which can be simplified as a tube with cross-sectional area variable both in time and along its axis (Fig. 1–4a). The configuration of the vocal tract is determined by the positioning of the articulators such as the tongue, the lips, and the jaw, which move continuously during speech production.

As the most active articulator, the tongue functions not only to modify the shape of the vocal tract, or the acoustic characteristics in the radiated sound, but also works as a valve for either inhibiting or stopping the flow of air in the mouth. Working together with the teeth, alveolar processes, and the palate, the tongue is a part of the noise generator in the vocal tract. [10]

The pressure and flow in the vocal tract has been studied with a circuit model shown in Fig. 1–4b. [11] Pressure P is represented as potential and volume velocity U is represented as current. Lumped elements in the circuit model represent aerodynamic properties along the vocal tract and relevant physiological activities such as forming a supraglottal constriction or full closure, adducting or abducting the glottis, adjusting the stiffness of the vocal tract walls, and active expanding or contracting the vocal tract.

Below the glottis (the space between the vocal folds which is shown as an elliptic circle located at the inferior end of the vocal tract in Fig. 1–4a), a constant pressure P_s is assumed during speech production. This pressure is represented as a voltage source with magnitude P_s in the circuit model shown in Fig. 1–4b. Z_c is the resistance of the supraglottal constriction and Z_g is the resistance of the glottis.

For constrictions with the dimensions encountered in speech production, an empirical pressure-flow relation $\Delta P = \frac{1}{2}\rho(\frac{U}{A})^2$ [1] pp30 can be applied. This relation comes from an average of the coefficient k_L (in $\Delta P = k_L \frac{1}{2}\rho(\frac{U}{A})^2$) measured by Heinz [12] and Jw. van den Berg et al. [13] Subsequently, resistance Z_c and Z_g are obtained from $Z=\frac{\Delta P}{U}=\frac{1}{2}\rho U$ at the particular location, and they are nonlinear elements in the circuit.

The movement of the jaw and some laryngeal structures can actively expand or contract the volume of the vocal tract. This volume change is represented as a current source U_a in the circuit.

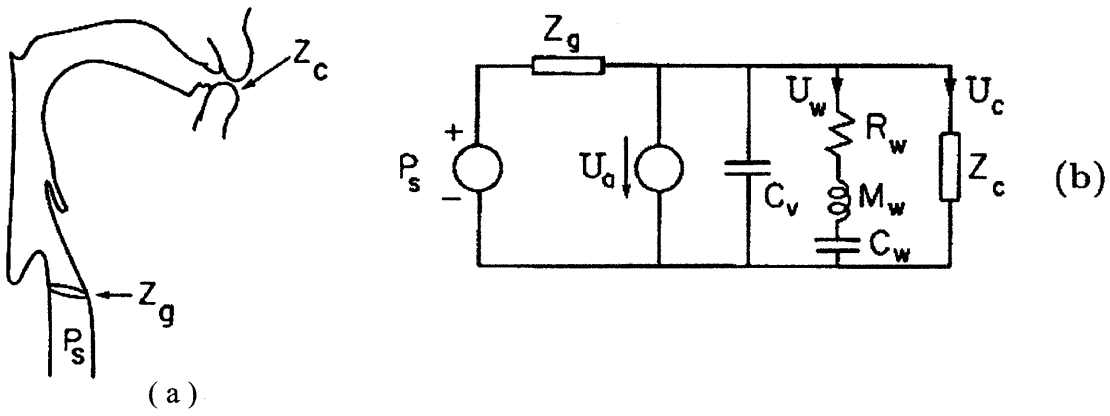


Fig. 1–4 (a) Midsagittal profile of the vocal tract when a supraglottal constriction is formed with the lips. Z_g and Z_c are the impedance at the glottal and supraglottal constriction.[1]
(b) Equivalent circuit model for the air flow in the vocal tract configuration in (a): the voltage source P_s represents the subglottal pressure; Current source U_a represents active expansion or contraction of the vocal tract volume; C_v is the compliance of the air in the vocal tract volume; and R_w, M_w and C_w are the resistance, mass, and compliance of the vocal tract walls respectively. [1]

A significant deviation from common duct flow is caused by the massive yielding walls of soft tissue surfaces along the vocal tract, with only a small portion of hard wall along the teeth and the hard palate. The yielding walls can be put into motion by pressure variations in the vocal tract, and the degree of mobility is described by the inverse of their mechanical impedance per unit area [1] pp26, which is defined as the ratio of the pressure acting on the wall over the resulting velocity. The impedance in the frequency range up to 100 to 200 Hz is the combination of a compliance C_s, a mass M_s, and a resistance R_s in series, $Z_s = \frac{1}{j\omega C_s} + j\omega M_s + R_s$.

Based on the impedance measured on cheek tissue conducted by Ishizaka, et al. [14], the ranges of the three elements (in per unit area) were estimated as

$C_s \sim 1.0\times10^{-5}$ to $3.0\times10^{-5}\ cm^3/dyne$

$M_s \sim 1.0$ to $2.0\ gm/cm^2$

$R_s \sim 800$ to $2000\ dyne\cdot s/cm^3$

Svirsky et al. [15] also measured the in-vivo compliance C_s of the tongue surface by tracking the displacement of a flesh point during the closure portion of voiced and voiceless stop consonants. The recorded displacements were averaged to derive the surface compliance, which showed the same order as the impedance estimated by Ishizaka et al. [14] Additionally, they found that the surface compliance was about 3 times larger for a voiced than for a voiceless stop consonant. The more compliant tongue surface in producing a voiced stop consonant had been considered as a means of maintaining voicing at the glottis [16]. In a pressurized tank model of making a stop closure later discussed in Appendix B, the difference in stiffness also results from the force balance required under difference levels of intraoral pressure buildup.

The impedance in per unit area introduced above is used to derive the elements C_W, M_W, and R_W in the circuit model shown in Fig. 1–4b. For a surface area S of the

walls between the glottis and the supraglottal constriction, multiplying C_s by the area S gives C_W, and dividing M_s and R_s by the area S respectively gives M_W and R_W.

The last element C_v in Fig. 1–4b represents the volume compliance of the air in the vocal tract. As a small quantity compared with the wall impedance, this element is often neglected.

1.4.2 Turbulence noise source and the place of articulation

The acoustic signal of obstruent consonants is characterized by the dominance of wideband noise, which is generated by the jet coming out from a narrowing called constriction in the vocal tract. This type of sound source is also called turbulence noise source because it is generated by the turbulence in the jet.

The turbulence noise source in obstruent production is located near to a constriction in the vocal tract. Once the pressure and flow along the vocal tract have been derived with the circuit model introduced above, the intensity of the turbulence noise sources can be estimated from empirical equations. Experiments with mechanical models[17] found that the sound power in the middle- and high- frequency range of the turbulence noise source generated at the constriction is proportional to the third power of the pressure buildup behind a constriction. For a given cross-sectional area A and a pressure drop ΔP_m across the constriction, the magnitude of the turbulence noise source is derived as: $p_s = K\Delta P_m^{\frac{3}{2}} A^{\frac{1}{2}}$ (K is a constant determined by the flow rate and the specific configuration of the constriction).

The turbulence noise source at the constriction is then filtered by the front cavity, which is the space between the constriction and the lips, and the spectrum of the radiated sound usually has a prominence at the frequency corresponding to the lowest natural frequency of the front cavity, also called the front-cavity resonance. As this resonant frequency is inversely proportional to the length of the front cavity, it gives information about the location of the constriction.

The location of the constriction is also called the place of articulation by phoneticians, as it is the place where an articulatory closure or constriction is made. The frequency of the front-cavity resonance has been found an important cue for identifying the place of articulation of obstruent consonants [18]-[20], [1].

1.4.3 The acoustic pattern of stop consonants

The acoustic events following the release of a stop consonant have been described by Fant [6]. The sequence starts with a brief pulse of volume velocity, which is called the initial transient. The transient is then followed by a burst of turbulence noise source called the frication noise, which locates at the constriction; and then a possible brief interval of turbulence noise source called the aspiration noise, which is generated at the glottis.

During the first 1-2 ms of the release, the air flow is accelerated by the pressure gradient across the released closure. This transient flow generates a sharp impulse at the released closure. The impulse contains the most accurate information of the place of articulation without any smearing from noise, and it could be a perceptually important cue to the place of articulation. However, with weak energy, this impulse was often found buried in the frication noise. [22]

Following the transient, the frication noise is generated by the air stream coming out of the released closure. The frication noise also contains the front-cavity resonance, thus providing information about the place of articulation.

Following the frication noise, the aspiration noise may be generated by the air stream at the glottal constriction, when the glottis is adducted for voicing. The aspiration noise does not contain salient information of the front cavity; because it is located at the glottis.

In English, a stop consonant may be produced at three different places of articulation: bilabial (at the lips, /p/ and /b/), alveolar (with the tongue tip and at the alveolar ridge, /t/ and /d/), and velar (with the tongue body and at the soft palate, /k/ and /g/).

Chapter 2

Literature review: previous studies on the pressure-wall interaction in speech production

At present, only two models have been found in the literature addressing the pressure-wall interaction in consonant production: McGowan's sprung-trap-door model[23] for tongue-tip trills and Stevens' 2-section model[1] pp 109-112 for the influence of steady flow on the yielding wall in the vicinity of a constriction in fricative production. In contrast, a large number of mathematical models have been developed for the vibration of vocal folds in the presence of air stream. [24]–[30]

The relation of these models to speech sounds is indicated in the concept map in Fig. 2–1. The model developed in this thesis is also included.

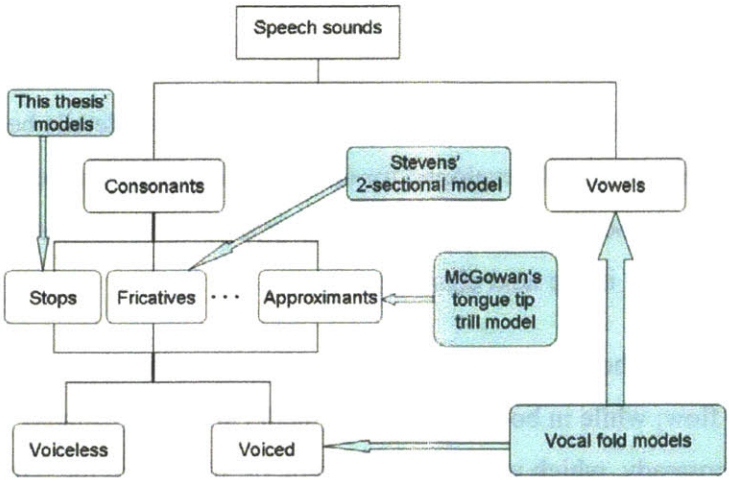


Fig. 2–1 Concept map: the pressure-wall interaction models in relation to speech sounds.

As Steven's model addresses fricatives, which is the type of consonants closest to the subject of this thesis, stop consonants, his model is discussed in more details. McGowan's tongue tip trill model focused on the role of the compliance of the vocal tract wall in sustaining the tongue-tip vibration. English does not have trill sounds. They are

special consonants in some languages such as Italian and Russian, and they can be categorized as approximants.

Both consonant models did not treat the collision of the primary articulator with the target plane, because collision is negligible in their cases; however, in stop consonant production, the motion of the primary articulator starts from a collided condition – the primary articulator in contact with the target plane, so the collision cannot be neglected.

The collision leads to a contact force between the primary articulator and the target plane, and in Section 5.3 a proper amount of contact pressure is demonstrated to be a necessary condition to retain a complete seal when air pressure is built up. The importance of the contact force in making a complete closure was also suggested by Löfqvist A. et al. [31] The contact force needs to be treated in a more complete stop production model.

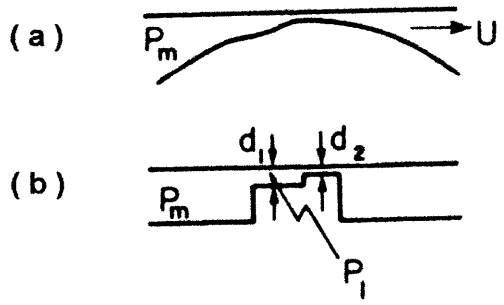
In contrast to the small number of consonant models addressing the pressure-wall interaction, a larger number of models have been developed for the pressure-wall interaction in vocal fold vibration, and some of them have also treated the collision.

The vocal fold models are not going to be covered in details because of the large amount. As the new model developed for stop consonants in this thesis is similar to the famous two-mass vocal fold model[25] in many aspects, more details about this vocal fold model is introduced in the next chapter.

As regards the type of flow in these models, Stevens' 2-sectional model deals with steady flow; while in both McGowan's model and the two-mass vocal fold model, the flow is unsteady, which is also characteristic of the flow in the release of a stop closure.

2.1 Steven's 2-section model for fricative production

Stevens [1] pp110-111 used a 2-section model (Fig. 2-2 Left (b)) to represent the yielding wall in the vicinity of a constriction in fricative consonant production. The tube with two sections of height d_1 and d_2 respectively, represents a typical tapered constriction with yielding wall shown in Fig. 2-2 Left (a). The section with height d_2 corresponds to the region with the minimum cross-sectional area.



When $P_m = P_{atm}$, and no air flow goes through the constriction, $d_2 = d_{20}$.

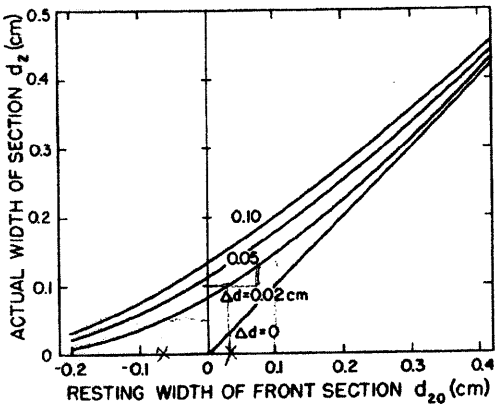


Fig. 2–2 Left: Diagram of Stevens' 2-section model. (a) A typical tapered constriction with yielding wall. P_m is the intraoral pressure and U is the volume velocity coming out of the constriction [1]. (b) The 2-section model used to represent the compliant constriction [1]. d_1 and d_2 are the height of the two sections respectively. Fig. 2-2 Right: The final height of the front section d_2 calculated with the 2-section model versus its initial resting height d_{20} [1]. d_{20} is the height of the front section when $P_m = P_{atm}$.

For such a compliant tube, the initial height of the front section is d_{20} without air flow, and the height becomes d_2 when a steady air flow goes through the tube after the intraoral pressure P_m is established upstream. The cross-sectional area of the section

with height d_2 determines the intensity of the turbulence noise, and it is also referred to as the size of the constriction.

The 2-section model has two parameters: the surface compliance per unit area C_1, and the difference Δd between the height of the front section (d_2) and that of the adjacent upstream section (d_1), indicating the degree of tapering. They are assumed to be constant in producing a fricative.

The model calculated the final height d_2 versus the initial height d_{20}, and the result is plotted in Fig. 2-2 Right for different degrees of tapering Δd. This result showed that the presence of the air flow enlarged the size of the fricative constriction and the amount of increase depends on the degree of tapering Δd, and the initial size of the constriction.

The model results also suggested that a yielding wall also makes it easier to form a supraglottal constriction whose cross-sectional area is optimal for generating turbulence noise with the maximum intensity at the constriction. Based on the pressure-flow relation and the empirical equation of the intensity of turbulence noise source introduced in the previous chapter, it has been found that when the cross-sectional area of the supraglottal constriction is $1/\sqrt{5}$ times the glottal constriction area, the maximum turbulence noise intensity can be achieved at the supraglottal constriction[1] pp110-111, for a constant subglottal pressure and fixed glottal constriction area.

As the glottal constriction area is estimated to be in the range of 0.1 ~ 0.3 cm^2 when a male speaker produces a voiceless fricative[1] pp110-111, the optimal supraglottal constriction area for the voiceless fricative is in the range of 0.04-0.13 cm^2. For a constriction width of 2 cm, the height is 0.02-0.065 cm in the vertical dimension. To achieve this small range of constriction area, precise control over the positioning of the primary articulator is required. However, from Fig. 2-2 Right, we know that this range of d_2 corresponds to a wider range of d_{20} up to 0.1 cm because of the pressure-wall

interaction, thus it becomes easier for the speaker to adjust the constriction size to be within the optimal range[1] pp111.

Additionally, the optimal range includes negative initial heights. A negative initial constriction area means that a complete closure is initially made because of a displacement intruding the roof of the constriction. The resultant contact pressure and the rich distribution of sensory receptors along the palate and the tongue could also contribute to the ease of control over both the vertical and axial positioning of the primary articulator.

Stevens' model suggested that an initial complete closure is allowed in producing a fricative; however, it also has a limitation which suggests that a negative initial height can always go above zero as long as there is any tapering in the shape of the constriction (Fig. 2-2b), and it is impossible to make a complete closure against air pressure buildup. As we know from the production of a stop consonant, when the initial height of a constriction is negative enough, the intraoral pressure will not be able to open it, and a complete closure can be retained before it is released.

2.2 McGowan's sprung trap door model for the tongue tip trills

Trill is a special type of speech sound presented only in some languages such as Italian and Russian. They are produced by sustained vibration of a soft tissue articulator. The model introduced here addressed the trill sounds produced with the tongue tip, but other articulators such as the lips and other parts of the tongue can produce trill sounds too.

McGowan [23] created a model of the tongue tip interacting with unsteady air flow in a sustained vibration, and quasi-steady flow was assumed (Fig. 2–3). The tongue-tip articulator was treated as a sprung trap door, moving under the combined action of the Bernoulli force in the air flow and the intraoral pressure P_c behind the "door".

The tongue tip model has a moment of inertia I, and it is hinged on the rest part of the tongue by springs, with a rest angle θ relative to the vertical direction. In deriving the moment of inertia I, the tongue tip is modeled as a rectangular beam with height h_T, thickness l_C and breadth (dimension into the paper) b_T. The parameter A_c represents the cross-sectional area of the constriction between the tip of the tongue and the palate, and U_c is the volume velocity coming out of this constriction.

The constriction area A_c and the intraoral pressure P_c are related in a nonlinear differential equation governing the motion of the "door":

$$\frac{d^2\theta}{dt^2}+\frac{r}{I}\frac{d\theta}{dt}+\frac{\kappa}{I}(1+\eta\theta^2)\theta=\frac{\tau}{I}.$$

In this equation, θ is the rest angle; $I=\rho_T l_C b_T(h_T/3)$; $A_c=b_T h_T[1-\cos(\theta)]$; κ and $\eta\kappa$ are linear and cubic spring constants; r is the resistive constant in the tongue tip; and ρ_T is the density of the tongue tip. τ is the torque exerted by the air pressure, derived as $\tau=\frac{1}{2}b_T h_T^{\,2}\left[P_c-K\frac{\rho_A}{2}\left(\frac{U_c}{A_c}\right)^2\right]$. ρ_A is the density of the air, and K is a parameter that

quantifies the reduction in static pressure surrounding the tongue-tip surface because of the Bernoulli effect.

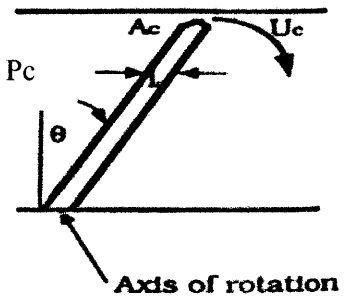


Fig. 2–3 McGowan's sprung trap door model for the tongue-tip trill. P_c is the intraoral pressure upstream of the constriction; θ is the rest angle; A_c represents the cross-sectional area of the constriction formed with the tongue tip and the palate; U_c is the volume velocity at the constriction; and L_c is the thickness of the tongue tip.

The intraoral pressure P_c is then related to the compliance and other lumped mechanical properties of the vocal tract wall behind the tongue tip by means of electrical circuit analogy. In order to solve these governing equations, McGowan had to estimate the order of magnitude of quite a few parameters with and without the aid of simultaneous intraoral pressure and volume flow data.

2.3 Vocal fold models

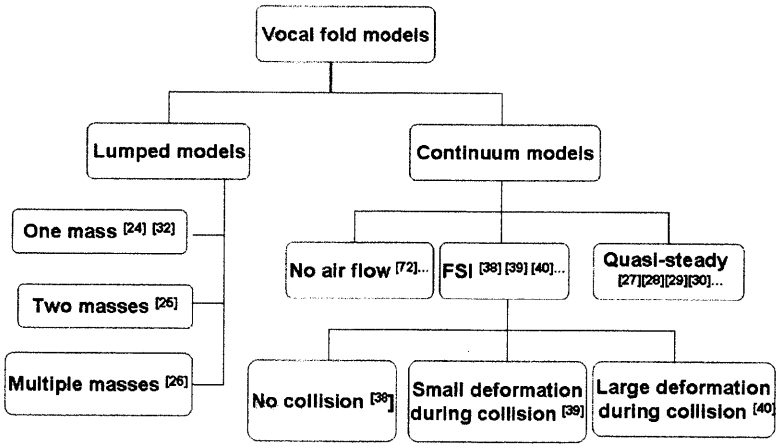


Fig. 2–4 Concept map: vocal fold models

The vocal folds have been simulated with lumped-element models, such as the one-mass model[24],[32], two-mass model[25] and multi- mass models[26], and also the continuum models[27], [28],[29],[30],[72] (Fig. 2–4).

Most vocal fold models treat the pressure-wall interaction under the quasi-steady assumption, which validates the application of the flow variables obtained on a static vocal fold model to the vocal folds in motion [33], [34], except for a short time before the vocal folds are collided and after they are separated apart.[35], [36],[37]

Some models [38],[39],[40] do not require the quasi-steady assumption to apply the empirical relations. They calculate the deformation of the vocal folds and the air flow between them simultaneously before coupling the solutions in both domains together. This process is also called FSI (Fluid-structure Interaction). Among these FSI vocal-fold models, Tao and Jiang [38] did not treat the collision; Luo et al.[39] treated small deformations caused by the collision; and an FSI continuum model capable of treating large deformations during the collision has recently been developed by Zhang et al. [40].

The famous two-mass model developed by Ishizaka and Flanagan [25] is a widely used lumped-element model. It can deal with the collision regardless of the degree of resultant deformation; however, it requires the quasi-steady assumption.

With similar method applied to the two-mass model, we first formulate a model devoted for the release of a stop closure in the next chapter. The stop consonant model differs in that the two masses representing the yielding surface of the primary articulator have an additional uniform rigid body motion. In vocal folds vibration, the rigid body motion of the folds is negligible. Moreover, at the time of release, an initial deformation due to the contact pressure at the time of release exists in the stop model.

Like the two mass model, quasi-steady flow is assumed in the first stop consonant model. The model is then further improved by relaxing the quasi-steady flow assumption, and two unsteady flow models are developed: unsteady incompressible and unsteady compressible flow.

Chapter 3

Analyses of the pressure-wall interaction during the release of a closure

A stop consonant is produced by creating a complete closure in the air way with the primary articulator, and then air pressure is built up behind the closure. During the closure interval, a yielding wall in the vicinity of the closure is essential to prevent the air from escaping, functioning as an O-ring seal used in tubing systems containing fluids.

Furthermore, a proper amount of contact pressure is required between the two surfaces in contact for effective sealing. Both this contact pressure and the intraoral pressure buildup during the closure interval act on the yielding wall as external forces. Therefore, certain amount of elastic energy has already been stored in the wall before the release starts.

After the closure is released, the contact pressure disappears and the air pressure in the vicinity of the released closure drops suddenly, the yielding wall would move in respond to this change in the external forces. The resultant motion would again change the cross-sectional area of the released closure, and also the surrounding air pressure.

This interaction of the air pressure and the yielding wall is analyzed in this chapter. Since releasing the closure is accompanied by a rapid airflow which produces a burst of sound composed of a brief initial transient and turbulence noise, with duration of a few tens of milliseconds, the pressure-wall interaction would affect the airflow, and eventually influence the generated sound.

In order to analyze the pressure-wall interaction, a solid model of the yielding wall and a flow model of the air flow going through the released closure are required, and then the two models are solved simultaneously.

The goal of the analysis is to calculate the evolution of the cross-sectional area $A_c(t)$ after the closure is released, also called the release trajectory. Stevens hypothesized a plateau-shaped release trajectory, as shown in Fig.1–3a, so we first hope to find out if a plateau really exists in the calculated release trajectory. If a plateau does exist, we also hope to know how long it lasts during each type of release, as this duration would determine the duration of the frication noise in the acoustic signal of the particular type of stop consonant.

3.1 Solid model

A lumped-element solid model is developed to represent the viscoelastic properties of the soft tissue surface of the primary articulator (the lower lip, the tongue tip, or the tongue body in English). Like the two-mass model of vocal folds, this solid model of the soft-tissue articulator is composed of two masses, three springs, and two dampers, as shown in Fig. 3–1a. The upper mass represents the part of the surface in contact with a rigid target plane when a closure is formed, and the lower mass represents the part of the surface exposed to the intraoral pressure.

Two masses are applied to represent the soft tissue surface in the vicinity of the closure, as the part of the surface right at the closure and that upstream of the closure do not move in phase because the pressure forces acting on them are different. Consequently, different amount of potential energy is stored in them at the time of release.

The same per-unit-area value of the mass (m), spring constant (k), and damping coefficient (r) are assigned to the elements belonging to the two parts of surface, as the properties of the soft tissue surface can be assumed uniform in the vicinity of the closure. These quantities are determined according to the experiments done by Ishizaka et al.[14] on the relaxed cheek tissue (refer to Appendix A for choosing the measured value of relaxed versus tense cheek tissue). The spring constant k_c connecting the two masses is determined as $1.5k$ according to the convention used in the two-mass model of the vocal folds [25].

The target plane is assumed to be rigid and it is represented by a straight dashed line in the 2-D model shown in Fig. 3–1. This line represents the palate for alveolar and velar stop consonants (refer to Section 1.4.3, and Fig. 1–1); and for bilabial stop consonants, it represents a virtual plane of symmetry between the two lips.

During the release, the base of the two-mass system moves downward with a constant velocity V, as shown in Fig. 3–1. The subsequent displacement of the upper mass $y_1(t)$, with t representing the time, would lead to a change of $y_1(t)D$ in the cross-sectional area of the supraglottal constriction (D is the length of the constriction along the dimension perpendicular to the midsagittal plane.). The real-time cross-sectional area of the released closure is thus

$$A_c(t) = Vt - y_1(t)D. \qquad (1)$$

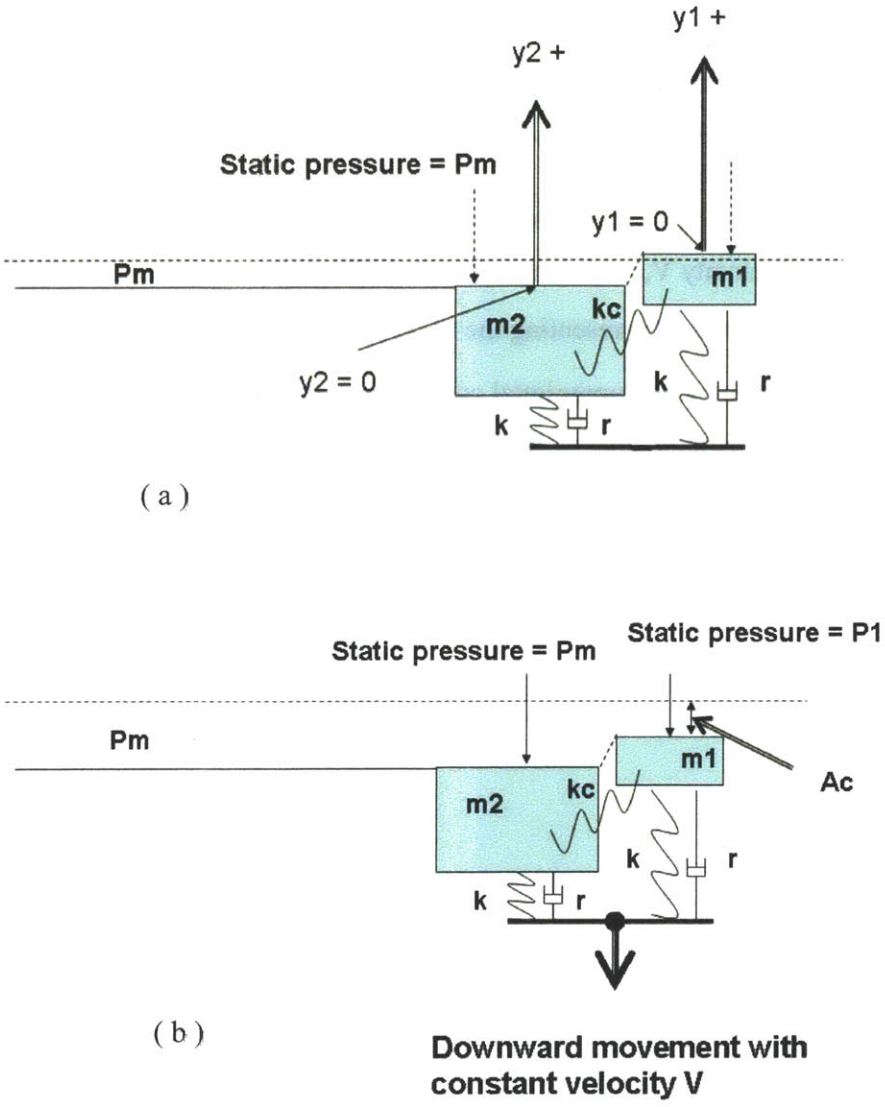


Fig. 3–1 Lumped elements representing the soft tissue surface in the vicinity of the closure. The rigid target plane is represented by a straight dashed line above the lumped elements. The upper-mass represents the part of surface in contact with the target plane and the lower mass represents the part of surface right behind the closure. (a) The configuration of the lumped elements during the closure interval. The force received on the lower mass is the intraoral pressure P_m and the force received on the upper mass is the contact pressure. (b) The configuration of the lumped elements when the closure is released. The force on the lower mass is the intraoral pressure P_m and the force on the upper mass is P_1.

The governing equations of the motion of the two masses are formulated in the y_1- and the y_2- coordinate respectively, as shown in Fig. 3–1a. Both coordinates move downward with a constant velocity V, so they are inertial coordinates. In each coordinates, the position of the mass at the time of release is set as the origin.

According to Newton's second law, the governing equation of the upper mass m_1 is formulated as:

$$m\frac{d^2y_1}{dt^2} + r\frac{dy_1}{dt} + k\left(y_1 - y_{eq1}\right) = F_{kc} + F_{P1} \tag{2}$$

$$F_{kc} = -k_c\left(\left(y_1 - y_{eq1}\right) - \left(y_2 - y_{eq2}\right)\right) \tag{3}$$

$$y_{eq1} = \frac{P_{c\_rel}}{k} \tag{4}$$

$$y_{eq2} = \frac{P_s}{k} \tag{5}$$

The initial conditions are $t = 0, y_1 = 0, \frac{dy_1}{dt} = 0$.

In Equation (2), F_{kc} is the force generated by the spring connecting the two masses, and F_{P1} is the average pressure force acting on the upper surface of mass m_1. As the lumped elements are considered for unit area, the force F_{P1} equals the static pressure P_1. y_{eq1} and y_{eq2} in Equation (3) are the equilibrium positions of the two masses respectively.

P_{c_rel} in Equation (4) is the contact pressure at the time of release. If the two contacting surfaces are separated in the absence of air pressure, the contact pressure is zero and no deformation retains. However, because of the higher intraoral pressure behind a stop closure, the primary articulator could be pushed apart from the target plane, with some deformation retained in the yielding wall.

Although the value of P_{c_rel} is unknown, its range can be estimated from related measurements. Matsumura, et al.[41] found that the maximum lingual-palatal contact stress during alveolar consonant production was in the range of 5 to 6 kPa, which is over 5 times larger than the average subglottal pressure P_s. The values obtained by McGlone et al.[42] were lower, with an average of about 3 kPa for syllable initial /t/. Therefore, in the models, this parameter is varied from $0.5P_s$ to $5P_s$, and the corresponding release trajectories are calculated. Indirect means of measuring this parameter is also discussed in Section 5.3.

The governing equation of the lower mass m_2 is:

$$m\frac{d^2y_2}{dt^2}+r\frac{dy_2}{dt}+k\left(y_2-y_{eq2}\right)=-F_{kc}-P_m \tag{6}$$

The initial conditions are $t=0, y_2=0, \frac{dy_2}{dt}=0$. In this equation, P_m is the intraoral pressure, which is also the static pressure acting on the lower mass m_2.

In the two governing equations discussed above, the average static pressure acting on the two masses, F_{P_1} and P_m, are unknown flow variables. In order to derive them, flow models are developed.

3.2 Flow models

A schematic configuration of the vocal tract is shown in Fig. 3–2. The subglottal pressure P_s and the cross-sectional area of the glottal constriction A_g are assumed constant. The cross-sectional area of the supraglottal constriction $A_c(t)$ is under the influence of pressure-wall interaction, and the panels illustrating the release events hypothesized by Stevens in Fig.1–3b is also attached below the supraglottal constriction.

As the glottal constriction has a constant cross-sectional area A_g, the empirical pressure-flow relation $\Delta p = \frac{1}{2}\rho(\frac{U}{A})^2$ for constrictions with the dimensions encountered in speech production[1]pp30 can be directly applied to the volume flow U_g going through the glottis. The intraoral pressure P_m can be thus derived as:

$$P_m(t,y) = P_s - \frac{1}{2}\rho\left(\frac{U_g}{A_g}\right)^2 \tag{7}$$

The flow going through the space between the upper mass m_1 and the rigid target surface at the supraglottal constriction is in the region of pressure-wall interaction. This space is indicated by a dashed circle in Fig. 2–4, and it is simplified as a short uniform tube whose cross-sectional area changes with time.

As the length of this tube is short compared with both the wavelengths of the acoustic wave in the air and the elastic wave propagating through the compliant wall, the lower boundary of the tube moves uniformly; hence it is lumped as the upper surface of a single mass. Consequently, the flow is simplified as going through a uniform tube with length L_c and with time-varying cross-sectional area $A_c(t)$, as shown in Fig. 3–3.

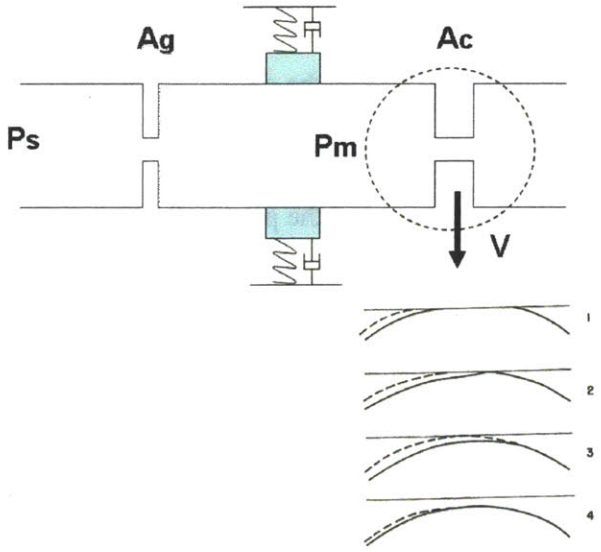


Fig. 3–2 Simplified configuration of the vocal tract for demonstrating the pressure-flow along the vocal tract. P_s is the subglottal pressure, A_g is the cross-sectional area of the glottal constriction, P_m is the intraoral pressure, and A_c is the time-varying cross-sectional area of the supraglottal constriction at which the pressure-wall interaction occurs. The lower boundary of the supraglottal constriction moves downward with a constant velocity V. The release events [1] pp329 hypothesized by Stevens are also attached for illustrating the big picture along the whole vocal tract.

In order to calculate the flow variables along the vocal tract, three flow models are established: 1) quasi-steady incompressible flow model, 2) unsteady incompressible flow model, and 3) unsteady compressible flow model. In all the pressure-wall interaction models discussed in the previous chapter, the flow was assumed incompressible. A compressible flow model is developed in this thesis out of some considerations in developing a fully unsteady flow.

Viscosity is included in the quasi-steady flow model, by using the empirical pressure-flow relation; however, it is not treated in the unsteady flow models. Elaborate

treatment of the viscosity in vocal fold vibration can be found in Deverge et al. [43] and Vilain, et al.[37]

3.2.1 Flow model 1: quasi-steady incompressible flow

In the two-mass model of vocal folds, the pressure drop ΔP across the supraglottal constriction is related to the volume velocity U with an equation derived from electrical circuit analogy: ΔP=RU+L$\frac{dU}{dt}$. In the first term on the left side, R represents the resistance, which is determined under the quasi-steady assumption. The empirical pressure-flow relation for static constrictions $\Delta p = \frac{1}{2}\rho(\frac{U}{A_c})^2$ [1] pp30 is applied to the supraglottal constriction with cross-sectional area A_c at time t. In the second term, L represents the inductance calculated as $L = \frac{\rho L_c}{A_c}$, in which L_c is the length of the tube.

With this pressure-flow equation applied to the flow through the supraglottal constriction, the intraoral pressure P_m is derived as:

$$P_m = \frac{\rho}{2}\left(\frac{U}{Ac}\right)^2 + L\frac{dU}{dt} \tag{8}$$

Substitute (9) in (7), and we have

$$P_s - \left(\frac{\rho}{2}\left(\frac{U}{Ac}\right)^2 + L\frac{dU}{dt}\right) = \frac{\rho}{2}\left(\frac{U}{A_g}\right)^2 \tag{9}$$

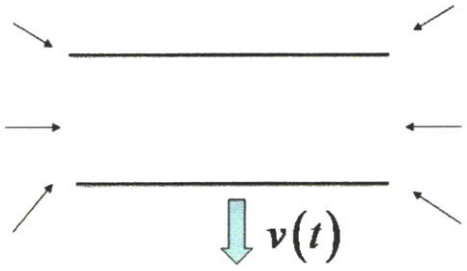
From this equation, the volume velocity U is derived first, which is then substituted in Equation (8) to derive P_m. The initial condition for solving Equation (9) is $t = 0,\ U = 0$.

As the cross-sectional area at the time of release $t = 0$ can not be zero, an initial area A_{c_start} has to be assigned.

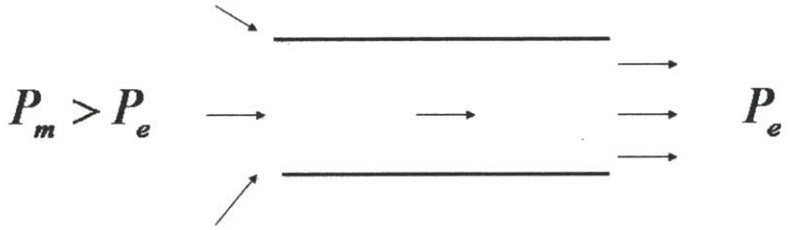
In order to derive F_{P_1} in Equation (2), the Bernoulli equation for unsteady flow is applied to the uniform tube, and the static pressure is found to vary linearly along the length of tube. The same result is obtained using the unsteady term (the second term on the right side) in the electrical circuit analogy. As the static pressure at the outlet P_e equals the atmospheric pressure, and the static pressure at the inlet is close to the intraoral pressure P_m, the average static pressure acting on the mass m_1 $F_{P1} = \frac{P_m}{2}$.

3.2.2 Flow model 2: unsteady incompressible flow and deformable control volume analysis

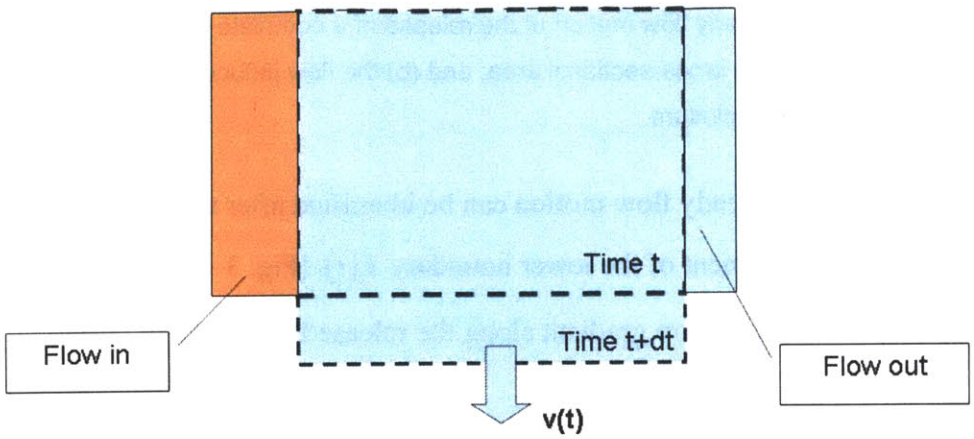
At the release of a complete closure, the air in the uniform tube above the upper mass m_1 is accelerated from the state of rest, and the surrounding flow field is intrinsically unsteady.



(a)



(b)



$$\frac{d}{dt}(system)=\frac{d}{dt}(CV)+(Flow-out)-(Flow-in)$$

Fig. 3–4 Illustration of a system property transporting within a deformable control volume and the modified Reynolds Transport Theorem. The blue color marks the system occupying the control volume at time t; the orange color marks the extra substance flowing into the control volume from the environment during the time interval Δt.

Symmetry is a salient feature of this type of unsteady flow caused by a moving boundary, and it helps to simplify the analysis (refer to Fig. 3–5). First, the velocity in the middle of the tube can be set as zero. Second, only the velocity and pressure distribution along half of the tube needs to be derived.

A deformable control volume of length $(L_c/2-x)$ is then used to derive the velocity and pressure distribution along the tube in the first type of unsteady flow (Fig. 3–5). The left boundary of the control volume is at a location x inside the tube, with the origin $x=0$ sitting in the middle of the tube; the right boundary of the control volume is at the exit of the tube. The control volume deforms as the cross-sectional area of the tube $A_c(t)$ changes with time.

Applying the conservation of mass to this control volume, we have

$$0=\rho A_c(L_c/2-x)\big|_{t+\Delta t}-\rho A_c(L_c/2-x)\big|_t+\rho u_x A_c\Delta t-\rho u_e A_c\Delta t$$

Fig. 3–3 Two types of unsteady flow motion at the release of a complete closure: (a) the flow caused by the change in the cross-sectional area; and (b) the flow induced by the pressure gradient along the released closure.

Two types of unsteady flow motion can be identified after the release: (1) the air flow caused by the movement of the lower boundary $v(t)$ (Fig. 3–3a); and (2) the starting flow due to the pressure gradient along the released closure (Fig. 3–3b).

The second type of unsteady motion can be well treated by the Bernoulli equation for unsteady flow, leading to the same unsteady term $L\frac{dU}{dt}$ in the first model. However, in the first type of unsteady flow, the boundary moves in the direction perpendicular to the mean flow, which is not a common flow problem with one-dimensional analysis.

In order to treat the first type of unsteady flow, a deformable control volume analysis is deployed to solve the moving-boundary problem. The name, "deformable" control volume, suggests that it is a type of control volume which can deform, in contrast to the "fixed" control volume commonly used in fluid dynamics.

A deformable control volume is illustrated in Fig. 3–4 with a rectangle of dashed line. The lower boundary of the control volume moves at the speed of $v(t)$, which leads to an increase in the volume at time $t+\Delta t$. The system occupying the control volume at time t would occupy the area filled with the blue color in the diagram. The control volume at time $t+\Delta t$ also contains extra substance marked with the orange color, which flows in from the environment.

The Reynolds Transport Theorem modified for the system in such a deformable control volume is

$$\frac{d(system)}{dt}=\frac{d(CV)}{dt}+(Flow-out)-(Flow-in)$$

This analysis has also been discussed by White [44] pp137.

For incompressible flow, the density ρ is constant. u_x is the flow velocity at the location x inside the tube, and u_e is the flow velocity at the exit. Positive directions of the velocities are indicated in Fig. 3–5. Let $\Delta t \to 0$, and we have

$$u_x = u_e - \frac{\left(\frac{L_c}{2} - x\right)}{A_c}\frac{dA_c}{dt} \tag{10}$$

As the velocity at $x = 0$ is zero because of symmetry, we have

$$u_e = \frac{L_c}{2A_c}\frac{dA_c}{dt} \tag{11}$$

This equation suggests that when the tube is extended, i.e. $\frac{dA_c}{dt} > 0$, the flow velocity at the exit is positive (in the positive direction indicated in the diagram) and the air is absorbed into the tube; while when the tube is contracted, i.e. $\frac{dA_c}{dt} < 0$, the flow velocity at the exit is negative (opposite to the positive direction indicated in the diagram) and the air is squeezed out of the tube.

Substitute (11) in (10), and we have the velocity distribution along the tube

$$u_x = \frac{x}{A_c}\frac{dA_c}{dt} \tag{12}$$

Next, the momentum equation is applied to the same control volume in order to derive the pressure distribution p_x.

$$\left(-p_x A_c + P_e A_c\right)\Delta t = \int_x^{L_c/2} \rho u_x A_c dx\Big|_{t+\Delta t} - \int_x^{L_c/2} \rho u_x A_c dx\Big|_t + \rho u_x |u_x| A_c \Delta t - \rho u_e |u_e| A_c \Delta t$$

Let $\Delta t \to 0$, and we have

$$\frac{(-p_x + P_e)A_c}{\rho} = \frac{d}{dt}\int_x^{L_c/2} u_x A_c dx + \left(u_x|u_x| - u_e|u_e|\right)A_c \tag{13}$$

The static pressure at the exit equals the atmospheric pressure P_e. Substitute (11) and (12) in (13), and we have

$$\frac{(p_x - P_e)}{\rho} = -\frac{1}{2A_c}\frac{d^2A_c}{dt^2}\left(\frac{L_c^2}{4} - x^2\right) - \frac{x^2}{A_c^2}\frac{dA_c}{dt}\left|\frac{dA_c}{dt}\right| + \frac{L_c^2}{4A_c^2}\frac{dA_c}{dt}\left|\frac{dA_c}{dt}\right| \tag{14}$$

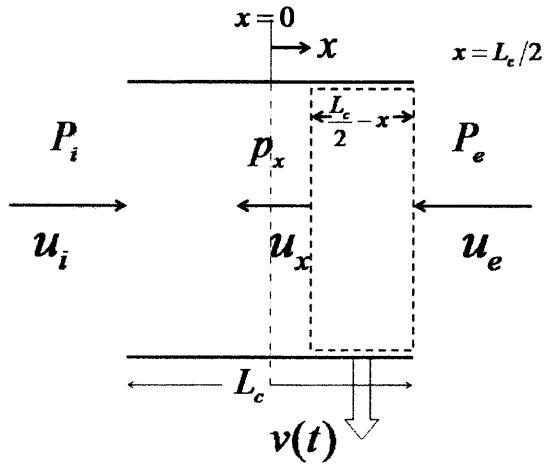


Fig. 3–5 A control volume (in dashed line) used for deriving the governing equations in the unsteady incompressible flow model in a uniform tube with length L_c and time-varying cross-sectional area $A_c(t)$ (not shown in the figure). The flow velocity at the inlet and the exit of the tube are u_i and u_e respectively, and the static pressure at the inlet is P_i, and P_e at the exit. The control volume shown in the diagram occupies a part of the tube with length $\frac{L_c}{2} - x$. The left boundary of the control volume is at the location x, whose origin is in the middle of the tube, and the right boundary of the control volume is at the exit of the tube.

The average static pressure acting on the upper mass m_1 is

$F_{P_1} = \frac{2}{L_c/2}\int_0^{L_c/2}(p_x - P_e)dx$. From (14), we can integrate p_x and obtain

$$F_{P_1}^{(1)} = -\frac{\rho L_c^{\,2}}{6A_c}\frac{d^2A_c}{dt^2} + \frac{\rho L_c^{\,2}}{3A_c^{\,2}}\frac{dA_c}{dt}\left|\frac{dA_c}{dt}\right| \tag{15}$$

The label "(1)" on the upper right corner indicates that this pressure force results from the first type of unsteady flow.

This type of unsteady motion does not directly change the intraoral pressure left to the tube; instead, it creates a volume velocity $\frac{L_c}{2}\frac{dA_c}{dt}$, which is counted as a current source in the electrical circuit analogy. This volume velocity, or current source, requires a flow velocity of $\frac{L_c}{2A_g}\frac{dA_c}{dt}$ to be created at the glottal constriction upstream, thus inducing a decrease in the intraoral pressure:

$$P_m^{(1)} = P_s - \frac{1}{2}\rho\left(\frac{L_c}{2A_g}\frac{dA_c}{dt}\right)^2 \tag{16}$$

Now, both flow variables required in the solid model are derived for the first type of unsteady flow.

In analyzing the second type of unsteady motion, a regular fixed control volume analysis is applied to the tube with a constant cross-sectional area $A_c(t)$ at time t. The velocity is the same throughout the length of the tube because viscosity is neglected, so the conservation of mass is satisfied automatically. Applying the momentum equation to a control volume of length x from the exit of the tube, we have

$$p_x - P_e = \rho x\frac{du_e}{dt} \tag{17}$$

The static pressure at $x = L_c$ is close to the intraoral pressure P_m, so we have

$$P_m^{(2)} = \rho L_c \frac{du_e}{dt} \tag{18}$$

P_m is the gauge pressure, which equals the absolute pressure minus the atmospheric pressure.

The average static pressure acting on the upper mass m_1 is $F_{P_1} = \frac{1}{L_c}\int_0^{L_c}(p_x - P_e)dx$. Substitute Equation (17) in the integral, and we have the pressure force resulting from the second type of unsteady motion:

$$F_{P_1}^{(2)} = \frac{P_m}{2} \tag{19}$$

The volume velocity $u_e^{(2)} A_c$ also creates a flow velocity of $\frac{u_e A_c}{A_g}$ at the glottal constriction, thus inducing a decrease in the intraoral pressure:

$$P_m^{(2)} = P_s - \frac{1}{2}\rho\left(\frac{u_e A_c}{A_g}\right)^2 \tag{20}$$

Now, the governing equation of the upper mass is formulated as:

$$\mathrm{m}\frac{\mathrm{d}^2 \mathrm{y}_1}{\mathrm{dt}^2} + \mathrm{r}\frac{\mathrm{dy}_1}{\mathrm{dt}} + \mathrm{k}\left(\mathrm{y}_1 - \mathrm{y}_{\mathrm{eq1}}\right) = \mathrm{F}_{\mathrm{kc}} - \mathrm{F}_{\mathrm{P}_1}, \tag{2}$$

$$F_{kc} = -k_c\left(\left(y_1 - y_{eq1}\right) - \left(y_2 - y_{eq2}\right)\right), \tag{3}$$

$$F_{P_1} = F_{P_1}^{(1)} + F_{P_1}^{(2)} = -\frac{L_c^2}{6A_c}\frac{d^2 A_c}{dt^2} - \frac{L_c^2}{3A_c^2}\frac{dA_c}{dt}\left|\frac{dA_c}{dt}\right| + \frac{P_m}{2} \tag{21}$$

P_m is synthesized from Equation (16) and (20):

$$P_m = P_s - \frac{1}{2}\rho\left(\frac{L_c}{2A_g}\frac{dA_c}{dt} + \frac{u_e A_c}{A_g}\right)^2 \tag{22}$$

This equation also couples the two types of unsteady motion together.

The cross-sectional area of the supraglottal constriction is related to the displacement of the upper mass.

$$\frac{d^2 A_c}{dt^2} = -D\frac{d^2 y_1}{dt^2} \tag{23}$$

$$\frac{dA_c}{dt} = V - D\frac{dy_1}{dt} \tag{24}$$

$$\mathrm{A_c(t) = A_{c\_start} + Vt - y_1(t)\,D} \tag{1}$$

A third governing equation of the flow velocity at the exit of the tube in the second type of unsteady motion is added

$$\frac{du_e}{dt} = \frac{P_m}{\rho L_c} \tag{25}$$

The initial condition is t=0; $u_e = 0$.

Solve the coupled ordinary differential equations (2), (6), and (25) simultaneously, and then the evolution of the cross-sectional area $A_c(t)$ after the release can be derived.

3.2.3 Flow model 3: unsteady compressible flow

When the closure is just released, the rate of change in the volume of the air between the upper mass m_1 and the target plane is large. If the compressibility of the air is considered, the density of the air could vary in time and also along the tube, thus changing the pressure force acting on the upper mass.

Using the deformable control volume analysis and the procedure introduced in the previous section, the density, velocity, and pressure distribution along the tube are derived for the first type of unsteady motion. The deformable control volume analysis for the second type of unsteady motion gives the same result as in the unsteady incompressible flow model.

A deformable control volume of length x is defined in Fig. 3–6, which deforms as the cross-sectional area of the tube $A_c(t)$ changes with time. The left boundary of this control volume is the location x inside the tube, with the origin of x at the exit of the tube. This location of origin is different from the one defined in the incompressible flow model, as here the flow field in the tube is approximated as perturbations about a reference point at the exit of the tube.

Apply the conservation of mass to this control volume, and we have

$$0 = \int \rho_x A_c dx\big|_{t+\Delta t} - \int \rho_x A_c dx\big|_t + \rho_x u_x A_c \Delta t - \rho_e u_e A_c \Delta t$$

u_x and u_e are the flow velocity at the location x and at the exit of the tube respectively, and ρ_x and ρ_e are the density of the air at the location x and at the exit of the tube respectively. Let $\Delta t \to 0$, and we have

$$\frac{d}{dt}\int_0^x \rho_x A_c dx + \left(\rho_x u_x - \rho_e u_e\right) A_c = 0 \tag{26}$$

As the velocity at $x = \frac{L_c}{2}$ is zero because of symmetry, we have

$$u_e = \frac{1}{\rho_e A_c}\frac{d}{dt}\int_0^{L_c/2} \rho_x A_c dx \tag{27}$$

Apply the momentum equation to the same control volume, and we have

$$(-p_x A_c + P_e A_c)\Delta t = \int_0^x \rho_x u_x A_c dx\big|_{t+\Delta t} - \int_0^x \rho u_x A_c dx\big|_t + \rho u_x |u_x| A_c \Delta t - \rho u_e |u_e| A_c \Delta t$$

Let $\Delta t \to 0$, and we have

$$(P_e - p_x) A_c = \frac{d}{dt}\int_0^x \rho_x u_x A_c dx + (\rho_x u_x |u_x| - \rho_e u_e |u_e|) A_c \qquad (28)$$

The static pressure at the exit equals the atmospheric pressure P_e, and p_x is the static pressure at the location x.

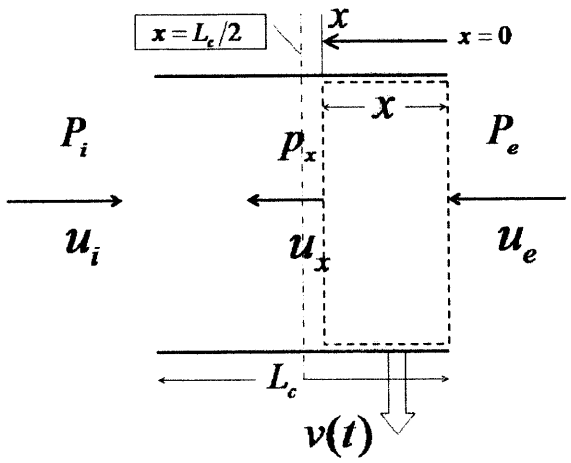


Fig. 3–6 A control volume (in dashed line) defined in a uniform tube with length L_c and time-varying cross-sectional area. The control volume has length x, with the left boundary at the location x and the right boundary at the exit of the tube. The flow velocity at the inlet and the exit of the tube are u_i and u_e respectively; the static pressure at the inlet is P_i; and the static pressure at the exit is P_e. The static pressure and the flow velocity at the location x are p_x and u_x respectively.

Apply the energy equation to the same control volume, and we have

$$(P_e A_c u_e \Delta t - p_x A_c u_x \Delta t) - F_{p_x}\left(\frac{1}{D}\frac{dA_c}{dt}\right)\Delta t = \int_0^x \rho_x \frac{u_x^2}{2} A_c dx\big|_{t+\Delta t} - \int_0^x \rho_x \frac{u_x^2}{2} A_c dx\big|_t + A_c \Delta t\left(\rho_x \frac{u_x^2}{2} - \rho_e \frac{u_e^2}{2}\right)$$

In this equation, $F_{p_x} = \int_0^x p_x dx$, which represents the total pressure force that the air in the control volume acts on the upper mass, and the term $F_{p_x}\left(\frac{1}{D}\frac{dA_c}{dt}\right)\Delta t$ represents the work that is done to the upper mass during the time interval Δt by the air in the control volume. Let $\Delta t \to 0$, and we have

$$\left(P_e u_e - p_x u_x\right) A_c - F_{p_x}\left(\frac{1}{D}\frac{dA_c}{dt}\right) = \frac{d}{dt}\int_0^x \rho_x \frac{u_x^2}{2} A_c dx + A_c\left(\rho_x \frac{u_x^2}{2} - \rho_e \frac{u_e^2}{2}\right) \tag{29}$$

As the tube is very short, with length of about 1 cm, it is assumed that the air in the tube goes through an isothermal process, and we have

$$\frac{p_x}{\rho_x} = \frac{P_e}{\rho_e} \tag{30}$$

Up to now, we have acquired four equations for three unknown variables p_x, ρ_x, and u_x, which are functions of both the time t and the space x. An additional unknown variable u_e, varies only with time t.

In order to derive $F_{p1} = \frac{2}{L_c}\int_0^{L_c/2}\left(p_x - P_e\right)dx$, some approximations are made. As the tube is short, x is a small variable in this problem; therefore, a perturbation series can be used to approximate the unknown variables. The density of the air $\rho(x,t)$ is first approximated as $\rho(x,t) = a_0(t) + a_1(t)x + a_2(t)x^2 + a_3(t)x^3 + \cdots$, then $p(x,t)$ can be derived directly from Equation (30).

Now we consider a tube of length $L_c = 1cm$, then we have x $\leq 0.5cm$, so we can further simplify the density of the air as $\rho(x,t) = a_0(t) + a_1(t)x$ by neglecting all the terms of order equal to and higher than x^2. A perturbation analysis is then implemented to derive $a_0(t)$ and $a_1(t)$ from the governing equations (26) – (30) (refer to Appendix A):

$$\frac{d}{dt}\left(\frac{d}{dt}\left(A_c\left(a_0\cdot\frac{L_c}{2}+a_1\cdot\frac{L_c^{\,2}}{8}\right)\right)\right)-A_c\left(\left|\frac{\rho_e u_e}{a_0}\right|\left(\frac{a_0}{A_c}\frac{dA_c}{dt}+\frac{da_0}{dt}\right)+\rho_e u_e\left|\frac{\rho_e u_e a_1}{a_0^{\,2}}+\frac{1}{A_c}\frac{dA_c}{dt}+\frac{1}{a_0}\frac{da_0}{dt}\right|\right)$$
$$=-\frac{P_e}{\rho_e}a_1 A_c. \tag{31}$$

The density of the atmosphere is taken as the density of the air at the exit of the tube, i.e. $x=0, \rho=\rho_e$, so we have $a_0=\rho_e$. Equation (31) then becomes a second-order ODE of the unknown variable $a_1(t)$. At $t=0$, as the density is uniform at any location in the tube, we have $a_1=0$; the velocity is also zero at any location. From the conservation of mass, we have $\frac{d}{dt}(\rho A_c)=0$, which gives $\frac{da_1}{dt}=-\frac{\rho_e}{A_c}\frac{dA_c}{dt}$ at $t=0$.

This ODE of $a_1(t)$ can be solved simultaneously with the governing equations of the two masses, which are also coupled to the second type of unsteady motion via the equations listed below.

$$F_{P_1}=-\frac{P_e L_c}{4\rho_e}a_1-\frac{P_m}{2} \tag{32}$$

$$P_m=P_s-\frac{1}{2}\rho\left(\left(u_e^{(1)}+u_e^{(2)}\right)\frac{A_c}{A_g}\right)^2 \tag{33}$$

$$\frac{du_e^{(2)}}{dt}=\frac{P_m}{\rho L_c} \tag{34}$$

$$u_e^{(1)}=\frac{1}{\rho_e A_c}\frac{d}{dt}\left(A_c\left(\rho_e\cdot\frac{L_c}{2}+a_1\cdot\frac{L_c^{\,2}}{8}\right)\right) \tag{35}$$

All the primary equations discussed in the three models are summarized in Table 3-1.

able 3-1 Equation set

$$m\frac{d^2y_1}{dt^2}+r\frac{dy_1}{dt}+k\left(y_1-y_{eq1}\right)=F_{kc}+F_{P1}\;;\; m\frac{d^2y_2}{dt^2}+r\frac{dy_2}{dt}+k\left(y_2-y_{eq2}\right)=-F_{kc}-P_m\;;\; F_{kc}=-k_c\left(\left(y_1-y_{eq1}\right)-\left(y_2-y_{eq2}\right)\right)$$

$$y_{eq1}=\frac{P_{c\_rel}}{k}\;;\; y_{eq2}=\frac{P_s}{k}$$

Flow model 1 (quasi-steady incompressible)	Flow model 2 (unsteady incompressible)	Flow model 3 (unsteady compressible)
$P_m=\frac{\rho}{2}\left(\frac{U}{Ac}\right)^2+L\frac{dU}{dt}$	$P_m=P_s-\frac{1}{2}\rho\left(\frac{L_c}{2A_g}\frac{dA_c}{dt}+\frac{u_eA_c}{A_g}\right)^2$	$P_m=P_s-\frac{1}{2}\rho\left(\left(u_e^{(1)}+u_e^{(2)}\right)\frac{A_c}{A_g}\right)^2$
$F_{P1}=-\frac{P_m}{2}$	$F_{P_1}=\frac{\rho L_c^{\,2}}{6A_c}\frac{d^2A_c}{dt^2}-\frac{\rho L_c^{\,2}}{3A_c^{\,2}}\frac{dA_c}{dt}\left\lvert\frac{dA_c}{dt}\right\rvert-\frac{P_m}{2}$	$F_{P_1}=-\frac{P_eL_c}{4\rho_e}a_1-\frac{P_m}{2}$
		$\frac{d}{dt}\left(\frac{d}{dt}\left(A_c\left(\rho_e\cdot\frac{L_c}{2}+a_1\cdot\frac{L_c^{\,2}}{8}\right)\right)\right)-$ $A_c\left(\lvert u_e\rvert\left(\frac{\rho_e}{A_c}\frac{dA_c}{dt}+\frac{d\rho_e}{dt}\right)+\rho_eu_e\left\lvert\frac{u_ea_1}{\rho_e}+\frac{1}{A_c}\frac{dA}{dt}\right.\right.$ $=-\frac{P_e}{\rho_e}a_1A_c.$
$\frac{dU}{dt}=\frac{1}{L}\left(P_s-\frac{\rho}{2}\left(\frac{U}{Ac}\right)^2-\frac{\rho}{2}\left(\frac{U}{A_g}\right)^2\right)$	$\frac{du_e^{(2)}}{dt}=\frac{P_m}{\rho L_c}$	$\frac{du_e^{(2)}}{dt}=\frac{P_m}{\rho L_c}$
	$u_e^{(1)}=\frac{L_c}{2A_c}\frac{dA_c}{dt}$	$u_e^{(1)}=\frac{1}{\rho_eA_c}\frac{d}{dt}\left(A_c\left(\rho_e\cdot\frac{L_c}{2}+a_1\cdot\frac{L_c^{\,2}}{8}\right)\right)$

Chapter 4

Results

The primary equations in the solid model and the three flow models summarized in Table 3-1, are solved using a multi-step solver in MATLAB, and the release trajectory – the evolution of the cross-sectional area of the released closure $A_c(t)$ – is calculated.

The three flow models are applied to two case studies and the results are compared. In one case, the starting cross-sectional area of the released closure A_{c_start} is varied (Section 4.1); in the other case, the length of the released closure L_c is varied (Section 4.2).

The typical release trajectory of voiceless English stop consonants /p, /t/, and /k/ (belonging to three places of articulation respectively) are calculated according to the typical release velocity V of the consonant. The contact pressure at the time of release P_{c_rel} is also varied. These release trajectories are a function of V and P_{c_rel}, i.e. $A_c(t) = f(V, P_{c\_rel})$.

4.1 The starting cross-sectional area of the released closure A_{c_start}

The starting cross-sectional area of the released closure A_{c_start} appears in the initial conditions explicitly only in Model 3 (unsteady compressible); in the other two models, it affects the solution by prescribing the initial cross-sectional area of the uniform tube in which air flow is initiated.

In order to investigate the influence of the starting cross-sectional area, the release trajectory of a single case is calculated with the three flow models. In this case, a closure is released at a downward release velocity $V = 50\ cm^2/s$, with the subglottal pressure $P_s = 8\ cmH_2O$, the glottal constriction area $A_g = 0.2\,\text{cm}^2$, the length of the

released closure L_c=1cm, and the contact pressure at the time of release $P_{c\_rel} = 3P_s$. The starting cross-sectional area A_{c_start} is decreased from 0.1 cm^2 to 0.0000001 cm^2.

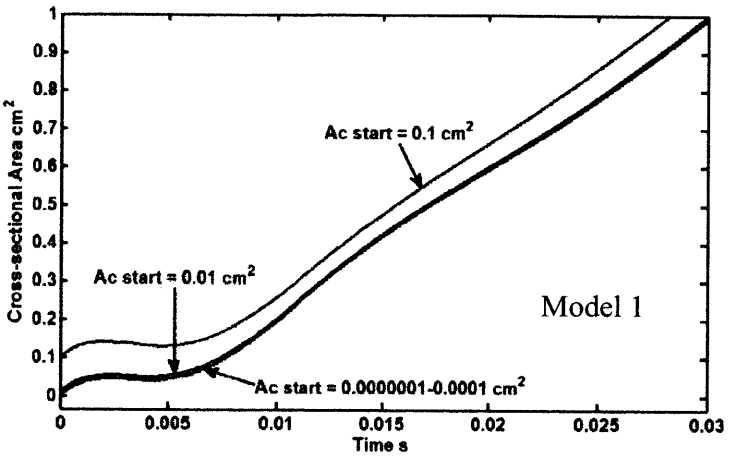


Fig. 4–1 The release trajectories of A_{c_start}=0.0000001 cm^2 to 0.1 cm^2 calculated with Model 1 (quasi-steady incompressible flow). The release trajectories change little when $A_c \leq 0.01cm^2$.

Shortly after the release starts, a plateau appears in the release trajectories calculated with all the flow models (Fig. 4–1 to 4–3). Only when $A_{c\_start} < 0.001cm^2$, the release trajectory calculated with Model 2 (unsteady incompressible flow) deviate from the plateau pattern.

The release trajectories calculated with Flow model 1 (quasi-steady incompressible flow) show little difference when the starting cross-sectional area A_{c_start} is smaller than and equal to 0.01 cm^2 (Fig. 4–1). However, the release trajectories calculated with Flow model 2 (unsteady incompressible flow) (Fig. 4–2) and 3 (unsteady compressible flow) (Fig. 4–3) continue to change when $A_{c\_start} < 0.01cm^2$, but the pattern of variation is different.

For Model 2 (unsteady incompressible flow), when $0.00001cm^2 < A_{c\_start} < 0.001cm^2$, the plateau gradually becomes a peak, and the magnitude of the peak

increases when A_{c_start} is further decreased. When $A_{c\_start} \geq 0.00001cm^2$, the system becomes so stiff that computation stops at a very short time after the release. Matlab indicates that it is "unable to meet integration tolerance without reducing the step size below the smallest value allowed." In Fig. 4–2, it is shown as a straight line on the vertical axis.

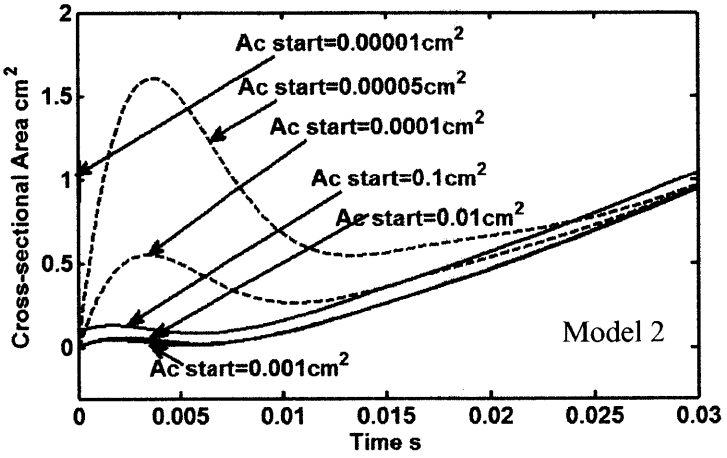


Fig. 4–2 The release trajectories of A_{c_start}=0.00001 cm^2 to 0.1 cm^2 calculated with Model 2 (unsteady incompressible flow). The plateau becomes a peak when $0.00001cm^2 < A_{c\_start} < 0.001cm^2$, and a straight line when $A_{c\_start} \geq 0.00001cm^2$.

Unlike Model 2 (unsteady incompressible flow), Model 3 (unsteady compressible flow) continues to derive a plateau whose magnitude decreases with decreasing A_{c_start} until the solution is collapsed (A_c becomes zero and the computation stops.) at $t \approx 2$ milliseconds when $A_{c\_start} = 0.00005cm^2$. When A_{c_start} is further reduced, the same pattern remains. In general, the release trajectories are closer to each other when $A_{c\_start} \leq 0.0001cm^2$.

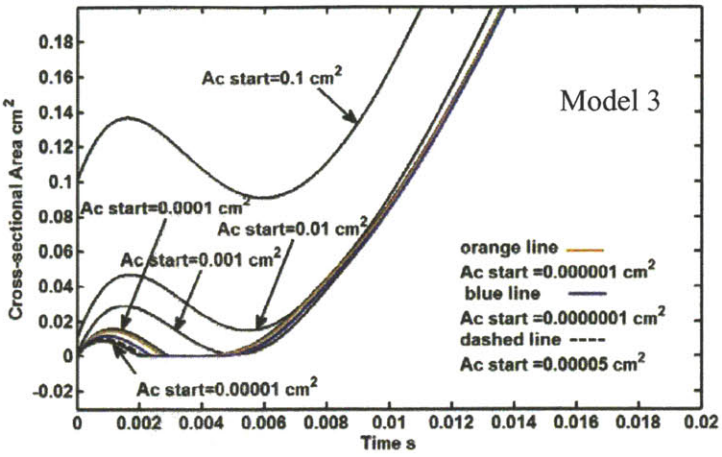


Fig. 4–3 The release trajectories of A_{c_start} =0.0000001 cm^2 to 0.1 cm^2 calculated with Model 3 (unsteady compressible flow). The magnitude of the plateau decreases with decreasing A_{c_start} until a zero cross-sectional area is reached at around 2 millisecond when $A_{c\_start} = 0.00005cm^2$.

4.2 The length of the released closure L_c

The influence of the length of the released closure L_c is evaluated next. The case studied also has the downward release velocity $V = 50\ cm^2/s$, the subglottal pressure $P_s = 8\ cmH_2O$, the glottal constriction area $A_g = 0.2\,\text{cm}^2$, and the contact pressure at the time of release $P_{c\_rel} = 3P_s$.

First, the three flow models are used to calculate the release trajectories for L_c = 1 cm, and A_{c_start} =0.001 cm^2 (Fig. 4–4). The release trajectories calculated with Model 2 (unsteady incompressible flow) and Model 3 (unsteady compressible flow) are closer to each other than to the trajectory calculated with Model 1 (quasi-steady incompressible flow), and they both have a plateau longer than the one calculated with Model 1.

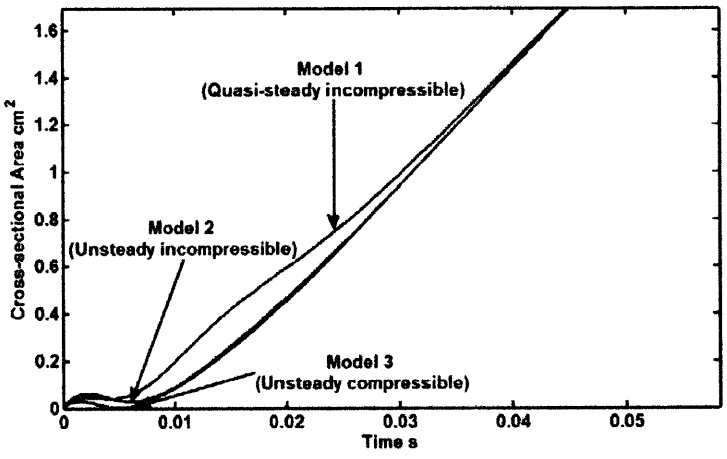


Fig. 4–4 The release trajectories of length $L_c = 1cm$ and A_{c_start} =0.001 cm^2 calculated with the three models

Next, the length of the released closure L_c is varied in each model ($L_c = 1, 0.5,$ and $0.1cm$), and the starting cross-sectional area $A_{c\_start} = 0.001cm^2$. The release trajectories are found insensitive to the length in all models, although the degree of insensitiveness

differs in each model. Fig. 4–5 shows that the release trajectories from model 1 (quasi-steady incompressible flow) overlap. In Fig. 4–6, the release trajectories calculated with model 2 (unsteady incompressible flow) only deviate a little bit from each other during the time $t = 3 - 13$ milliseconds. The results from model 3 (unsteady compressible flow) in Fig. 4–7 deviate even less but they are less overlapped compared with the trajectories from model 1 in Fig. 4–5 .

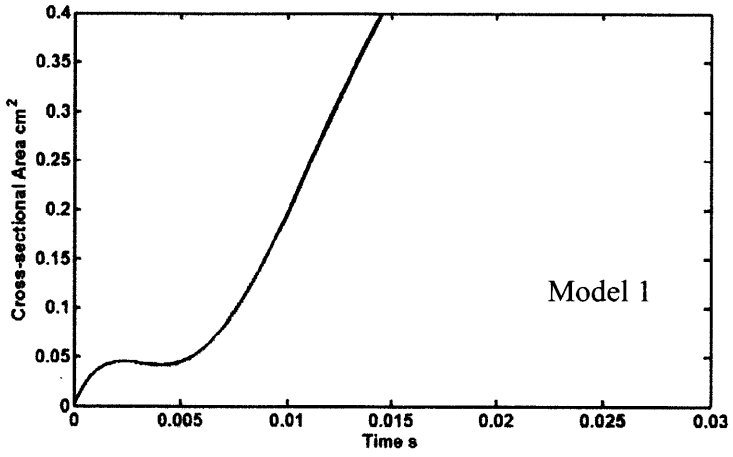


Fig. 4–5 The release trajectories of a closure with length $L_c = 0.1, 0.5$ and $1cm$ and $A_{c\_start} = 0.001cm^2$ calculated with Model 1 (quasi-steady incompressible flow).

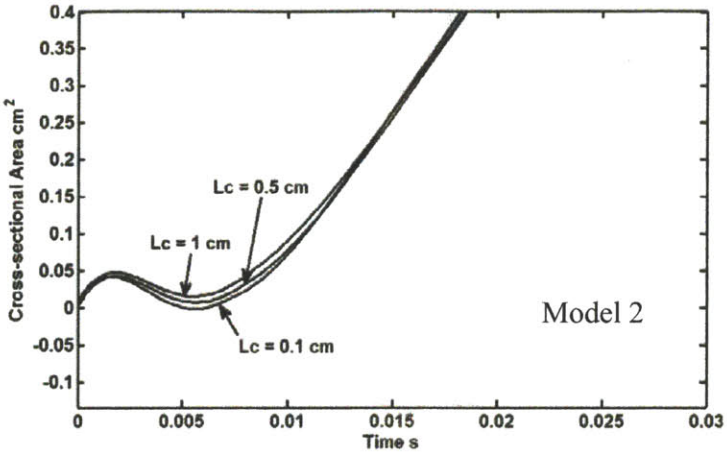


Fig. 4–6 The release trajectories of a closure with length $L_c = 0.1, 0.5$ and $1cm$ and $A_{c\_start} = 0.001cm^2$ calculated with Model 2 (unsteady incompressible flow).

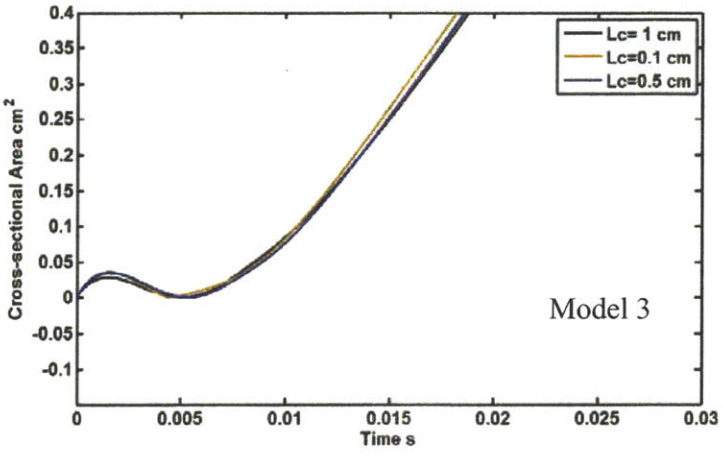


Fig. 4–7 The release trajectories of a closure with length $L_c = 0.1, 0.5$ and $1cm$ and $A_{c\_start} = 0.001cm^2$ calculated with Model 3 (unsteady compressible flow)

However, when the starting cross-sectional area A_{c_start} is further reduced, the dependence of the release trajectory on the length diverges in each model. The release trajectories calculated with model 1 (quasi-steady incompressible flow) remain insensitive to the difference in the length; the trajectories calculated with mode 2

(unsteady incompressible flow) and 3 (unsteady compressible flow) show larger difference for differentiated lengths (Fig. 4–8 and 23).

For model 2 (unsteady incompressible flow), when the starting cross-sectional area A_{c_start} is smaller than a critical value determining the change of the release pattern from a plateau to a peak (discussed in the previous section.), the dependence of the length becomes significant. Fig. 4–8 demonstrates the release trajectories for two lengths $L_c = 1, \text{and} 0.5cm$, with a common $A_{c\_start} = 0.00005cm^2$. A_{c_start} is chosen for both lengths to fall into the peak-pattern. The calculated release trajectories deviate from each other with the maximum difference of around 1.4 cm^2, and the longer closure releases with a peak of greater magnitude.

The release trajectories calculated with model 3 (unsteady compressible flow) also show greater dependence on the length when $A_{c\_start} = 0.0001cm^2$ (Fig. 4–9). However, different from model 2 (unsteady incompressible flow), the cross-sectional area is smaller during the plateau region in the release trajectory of longer closure.

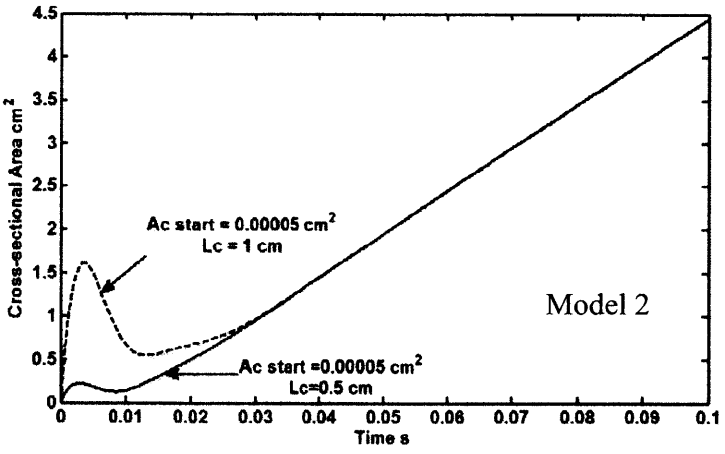


Fig. 4–8 The release trajectories of a closure with $L_c = 1$ and $0.5cm$ and $A_c = 0.00005cm^2$, calculated with Model 2 (unsteady incompressible flow)

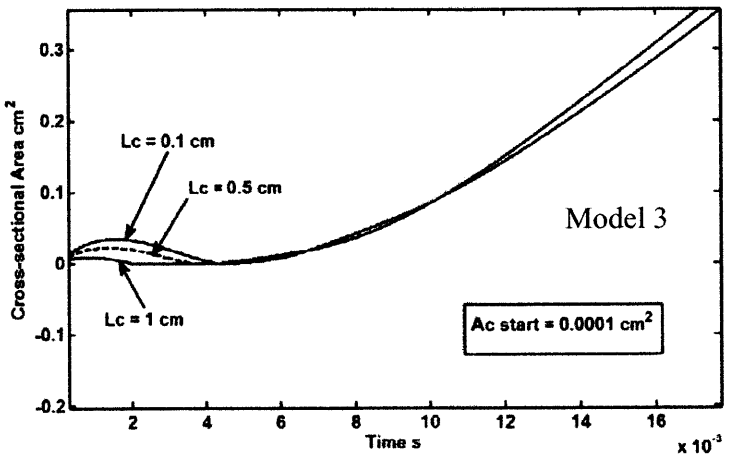


Fig. 4–9 The release trajectories of a closure $L_c = 1, 0.5,$ and $0.1cm$ and $A_c = 0.0001cm^2$ calculated with Model 3 (unsteady compressible flow).

4.3 The release trajectories of /p/, /t/, and /k/: the release velocity V, contact pressure at release P_{c_rel}, and collapse of the released closure

The typical release trajectories of English voiceless stop consonants /p/, /t/, and /k/ are calculated with Model 1 (quasi-steady incompressible flow) and 3 (unsteady compressible flow). Model 2 (unsteady incompressible flow) is not applied, as it derives releases trajectories with unrealistically large cross-sectional area (the peak pattern) when the starting cross-sectional area is small.

A /p/-release is represented by the typical release velocity of the lower lip of $100\,cm^2/s$. In a typical /t/-release, the tongue tip moves downward at the speed of $50\,cm^2/s$, and the tongue body moves downward at $25\,cm^2/s$ in a /k/-release (Stevens, 1998).

The contact pressure at the time of release P_{c_rel} is unknown, but the range of its magnitude must be smaller than 5-6 kPa, the maximum lingual-palate contact pressure measured in alveolar consonant production (Matsumura, et al.1994). We are going to show in Section 5.3 that P_{c_rel} is dependent on the intraoral pressure during the closure interval, or the subglottal pressure P_s. Therefore, P_{c_rel} is varied from $0.5\,P_s$ to $5\,P_s$, and $P_s = 8cmH_2O$.

The /p/-release trajectories calculated with Model 1 (quasi-steady incompressible flow) are shown in Fig. 4–10; the /t/-release trajectories in Fig. 4–11; and the /k/-release trajectories in Fig. 4–12. The starting cross-sectional area A_{c_start} is $0.0001\,cm^2$. They show that for larger contact pressure at the time of release (P_{c_rel}=$5\,P_s$ in the /p/-releases, $P_{c\_rel} \geq 3\,P_s$ in the /t/-releases, and $P_{c\_rel} \geq 2\,P_s$ in the /k/-releases,), a plateau exists in all places of articulation.

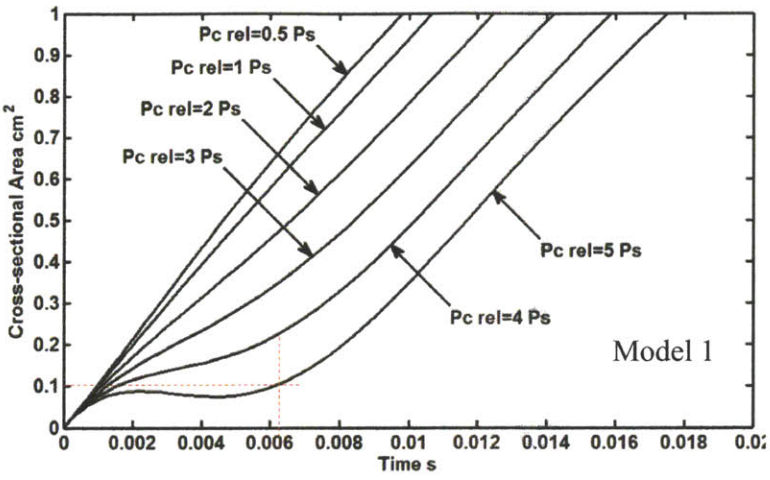


Fig. 4–10 The /p/-release trajectories calculated with Model 1 (quasi-steady incompressible flow)

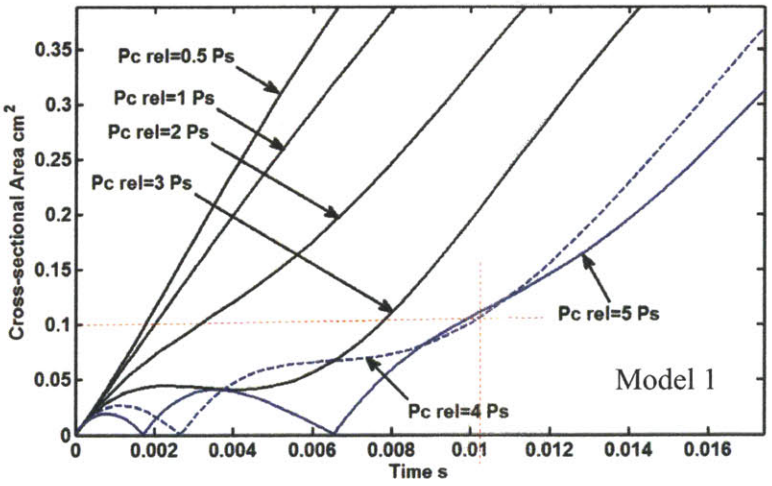


Fig. 4–11 The /t/-release trajectories calculated with Model 1 (quasi-steady incompressible flow)

For slower releases, collapse of the released closure occurs at a smaller P_{c_rel}. No collapse is found in the /p/-releases. In the /t/-releases, it occurs when $P_{c\_rel} \geq 4\,P_s$; and in the /k/-release with the smallest release velocity, it occurs when $P_{c\_rel} \geq 3\,P_s$. Among the

releases with the same downward velocity, collapse occurs at an earlier time for greater P_{c_rel}.

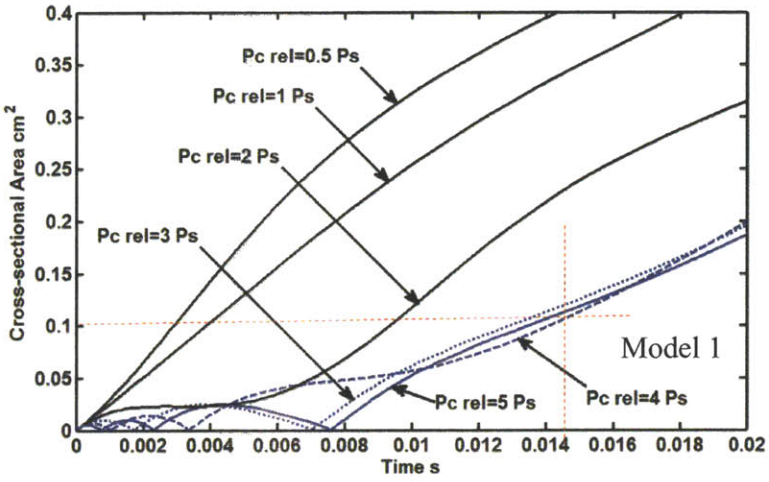


Fig. 4–12 The /k/-release trajectories calculated with Model 1 (quasi-steady incompressible)

The /p/-release trajectories calculated with Model 3 (unsteady compressible) are shown in Fig. 4–13, the /t/-release trajectories in Fig. 4–14, and the /k/-release trajectories in Fig. 4–15. The starting cross-sectional area A_{c_start} is 0.0001 cm^2 for /p/-, /t/-releases, and the /k/-releases with smaller $P_{c\_rel} = 0.5,\ \text{and}\ 1P_s$. For the /k/-releases with larger P_{c_rel}, A_{c_start} is reduced to 0.000001 cm^2 so that a smooth release trajectory can be calculated.

The release trajectories calculated with Model 3 (unsteady compressible) for each place of articulation show a longer duration of plateau than those calculated with Model 1 (quasi-steady incompressible). The durations of the frication noise estimated from the released trajectories calculated with both models are labeled in each figure and they are also listed in Table 4-1 for each place of articulation.

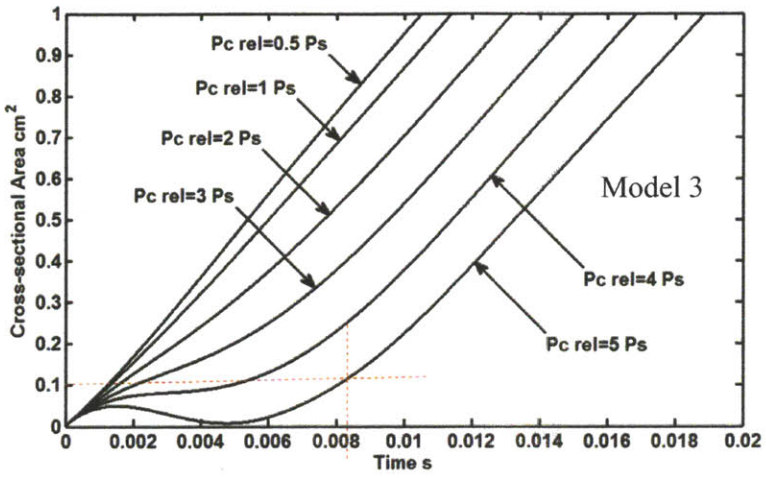


Fig. 4–13 The /p/-release trajectories calculated with Model 3 (unsteady compressible)

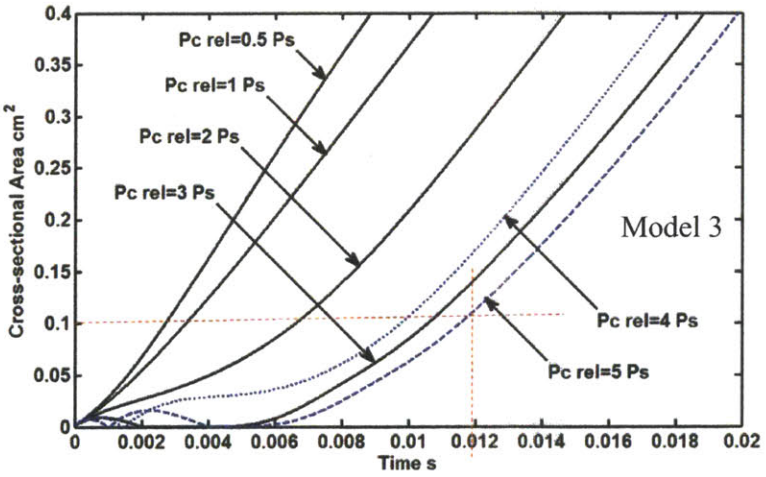


Fig. 4–14 The /t/-release trajectories calculated with Model 3 (unsteady compressible)

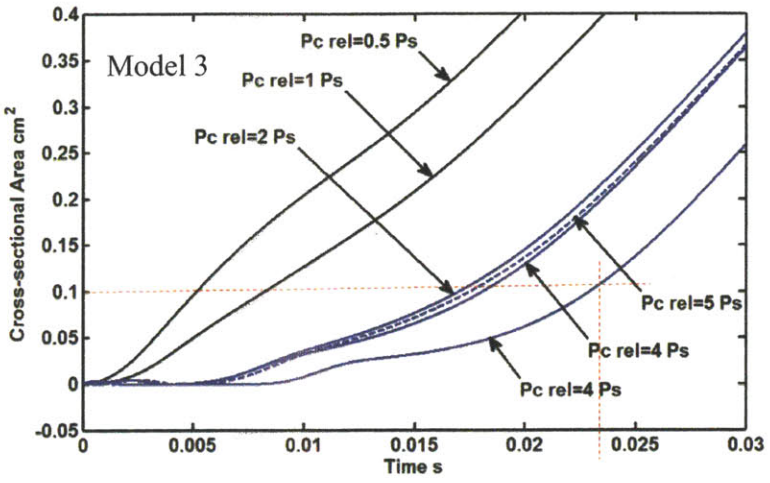


Fig. 4–15 The /k/-release trajectories calculated with Model 3 (unsteady compressible)

Table 4-1 Duration of the plateau of voiceless stop consonants calculated with Flow model 1 (quasi-steady incompressible) and 3 (unsteady compressible)

Place of articulation	Duration of the frication noise estimated from the release trajectory calculated with Model 1 (quasi-steady incompressible)	Duration of the frication noise estimated from the release trajectory calculated with Model 3 (unsteady compressible)
/p/	6.3	8.3
/t/	10	11.9
/k/	14.6	23

Chapter 5

Discussion

Results from the physical-analytical models that we have obtained for the pressure-wall interaction during the release of a stop closure, demonstrate that a plateau-like lingering does show up after the release, as Stevens portrayed, and the duration of this "plateau" gets longer with slower release movement of the primary articulator.

In developing the flow model for the pressure-wall interaction, we noted the limitation of the quasi-steady approximation in the two-mass vocal fold model, and improved it with an unsteady incompressible flow model, and then went further with an unsteady compressible flow model.

The release trajectories calculated with the quasi-steady incompressible and unsteady compressible flow models are then compared with the duration of the frication noise that is measured from the acoustic data of real speech.

In the analysis, the contact pressure at the time of release is an unknown variable, and we can only roughly estimate its range from the literature. Real-time measurement techniques for the lingua-palatal contact pressure during speech production are reviewed, and a static model is then discussed for the condition of retaining a stop closure. From this model, an indirect means of measuring the contact pressure at the time of release is proposed. A related physiological variable, the tongue strength, is also discussed in the same section.

5.1 Comparison of the flow models: quasi-steady versus unsteady; incompressible versus compressible

In a pressure-wall interaction model, a fundamental flow problem is to relate the pressure drop to the flow velocity or volume velocity through a constriction with time-

varying cross-sectional area. This pressure-flow relation in an unsteady flow is different from that in a steady or quasi-steady flow.

In Stevens' fricative model, steady flow was assumed because it is the equilibrium position of the yielding wall that needs to be examined, so the pressure-flow relation for steady flow $\Delta p = \frac{1}{2}\rho(\frac{U}{A})^2$ was applied.

In both McGowan's trill model and the two-mass model of vocal folds, the flow is intrinsically unsteady, and both models used this equation, ΔP=RU+L$\frac{dU}{dt}$, which originates from electrical circuit analogy, with R representing the resistance and L representing the inductance.

In determining the resistance R in the first term, as flow resistance is most conveniently measured on static constrictions[45],[13], the measured flow resistance can be applied to a constriction whose cross-sectional area changes with time, if quasi-steady flow is assumed[33].

The second term takes the form of an unsteady term, and it was applied to address the issue that "in the time-varying condition of the cords, the inertance of the air masses involved should be taken in account"[25] pp1240. The parameter "L" in the unsteady term equals $\frac{\rho L_c}{A_c}$, with the same form as an acoustic mass, indicating that the second term represents an inertance.

However, the origin of the second term $L\frac{dU}{dt}$ lacks an explanation in the view of fluid dynamics. This term may represent a linearized unsteady flow, as it has the same form of an acoustic mass. The acoustic wave is a type of linearized unsteady flow with zero mean flow, and the fluctuations in the flow variables are so small that the governing equations can be linearized.

We also note that $L\frac{dU}{dt}=\frac{\rho L_c}{A_c}\frac{d(A_c u)}{dt}=\frac{\rho L_c}{A_c}\left(\frac{dA_c}{dt}u+A_c\frac{du}{dt}\right)$, so it can capture the changes induced by the variations in both the flow velocity and the cross-sectional area. Massey [22] also applied this term in the analysis of the transient source during the initial 1-2 ms after the release of a stop closure; in this case, $\frac{dA_c}{dt}$ is a significant driver for the unsteady flow.

Although the term $L\frac{dU}{dt}$ does not accurately represent the unsteady flow effect, this error is disguised by the fact that the quasi-steady approximation is valid in the near field expect for a short time before and after the constriction is closed [35],[36],[37]. Zhang, et al.[33] also verified that the quasi-steady approximation is valid in the far-field.

Since the unsteady flow effect is more significant in the near field during a short time when the cross-sectional area is small, its influence is mostly in the high-frequency range. For the voicing sound generated by vocal fold vibration, the frequency range of interest mostly locates in the fundamental frequency, so the unsteady effect may not be important in most cases. However, for stop consonants, the frequency range up to 5 kHz, is perceptually important (refer to Section 5.2), and this is why we are considering the release of a stop closure in such a small time scale – tens of milliseconds after the release. The unsteady flow effect should be considered for the evolution of the cross-sectional area after the release.

Flow model 2 and 3 are derived in the view of fluid mechanics, and by means of deformable control volume analysis. They address the unsteady flow effect more completely than Flow model 1, which is established on the quasi-steady assumption and an unsteady term from electrical circuit analogy discussed above.

As discussed above, the unsteady term in Flow model 1 (quasi-steady incompressible) can be separated into two terms, $\frac{\rho L_c u}{A_c}\frac{dA_c}{dt}+\rho L_c\frac{du}{dt}$. The first term

$\frac{\rho L_c u}{A_c}\frac{dA_c}{dt}$ is related to the first type of unsteady flow motion illustrated in Fig. 3–3, and the second term addresses the second type of unsteady flow. The term for the second type of unsteady flow has the same form in all the three models, but the first type of unsteady flow is treated differently in each model.

The term most similar to $\frac{\rho L_c u}{A_c}\frac{dA_c}{dt}$ in Flow model 2 (unsteady incompressible) is $\frac{\rho}{A_c}\int u\frac{dA_c}{dt}dx$, which comes from $\frac{\rho}{A_c}\frac{d}{dt}\int uA_c dx = \frac{\rho}{A_c}\int\left(\frac{du}{dt}A_c + u\frac{dA_c}{dt}\right)dx$. In Flow model 3 (unsteady compressible), the most similar term is $\frac{1}{A_c}\int \rho u\frac{dA_c}{dt}dx$, which comes out of $\frac{1}{A_c}\frac{d}{dt}\int \rho u A_c dx = \frac{1}{A_c}\int\left(\frac{d\rho}{dt}uA_c + \rho A_c\frac{du}{dt} + \rho u\frac{dA_c}{dt}\right)dx$. The extra terms in Model 2 and 3 show that the first type of unsteady flow motion is only partially treated in Model 1.

Next, when compressibility is considered, the first type of unsteady flow motion provides a stabilizing force for the variation of the cross-sectional area of the tube. When the cross-sectional area A_c increases, the air density in the tube tends to decrease, and the static pressure drops consequently. The lowered static pressure would hold the moving boundary so that the cross-sectional area A_c would not increase as rapidly as before. This effect is more significant when A_c is small, because of the greater rate of density change.

This stabilizing force may also provide the second bouncing-back mechanism for the upper mass (at the beginning in Chapter 1, we only noticed the potential energy initially stored in the yielding wall as a bouncing-back mechanism), and this mechanism only exists in Flow model 3 (unsteady compressible).

Flow model 2 (unsteady incompressible) gives a complete representation of the first type of unsteady motion; however, it induces a destabilizing force as an incompressible model, especially for small cross-sectional area. This destabilizing force

accounts for the peak pattern replacing the plateau when the starting cross-sectional area $A_{c\_start} \leq 0.001 cm^2$ (Fig. 4-2).

In Fig. 4–5 of Section 4.1, the release trajectories calculated from Flow model 1 (quasi-steady incompressible) demonstrate the least variation in the duration of the plateau when the length of the released closure L_c is changed. This insensitivity may result from an incomplete treatment of the first type of unsteady flow motion.

The release trajectories calculated with the three flow models for the same L_c and A_{c_start} show the same pattern in Fig. 4-4. However, when the starting cross-sectional area is further reduced, the release trajectories calculated from flow model 2 (unsteady incompressible) (Fig. 4-8) and flow model 3 (unsteady compressible) (Fig. 4-9) show larger variations for closures with different length. This greater sensitivity to L_c for small cross-sectional area is believed to be an indication of the significance of the first type of unsteady motion for small cross-sectional area.

Among the release trajectories calculated with Flow model 1 (quasi-steady incompressible) and 3 (unsteady compressible) for the /p/-, /t/-, and /k/- releases with varied contact pressure at the time of release P_{c_rel} (Section 4.3), the plateau starts to appear at a larger P_{c_rel} in the trajectories calculated with Flow model 1, compared with those calculated with Flow model 3. This trend can be explained by the additional bouncing-back mechanism provided in the unsteady compressible flow model. A plateau may exist even though the potential energy initially stored in the upper mass is not larger enough to counteract the downward movement of the base.

5.2 The acoustic effect of the pressure-wall interaction in a syllable-initial voiceless stop consonant

The pressure-wall interaction makes it possible for a supraglottal constriction to retain a small cross-sectional area for a while – the plateau in the release trajectory [1]. Acoustically this plateau would lead to longer frication noise in the release bust, according to the articulation-to-acoustics theory about stop consonant production and Stevens' hypothesis [1].

In this section, we define an acoustic measure for the duration of the frication noise (in order to separate it from the remainder of the release burst), and apply this measure to some syllable-initial voiceless stop consonants /t/ and /k/ contained in sentences in TIMIT database[46]. Bilabial releases are not included because their spectra lack prominences.

This acoustic measure is illustrated in Fig. 5–1, in which the waveform of a /t/-release is plotted in blue line, the average FFT magnitude for every 5 milliseconds in black line, and the magnitude of a frequency band (bandwidth of 50 Hz) containing the front cavity resonance for every 5 milliseconds in red line. The duration of the frication noise is defined as the time between the start of the release and the time when the average FFT magnitude of the resonant band (in red line) starts to drop rapidly; at the same time, the ratio of the average FFT magnitude of the resonant frequency band is at least twice of the average FFT magnitude in the whole frequency range (in black line). The FFT window effect is then corrected by adding the window length of 5 milliseconds to the duration just defined. For the /t/-release shown in Fig. 5–1, the duration of the frication noise is measured as 23 ms.

Two samples of the /k/-releases that the newly defined acoustic measure is applicable are also shown in Fig. 5–2 and Fig. 5–3. The duration of the frication noise is measured as 45ms and 38 ms respectively.

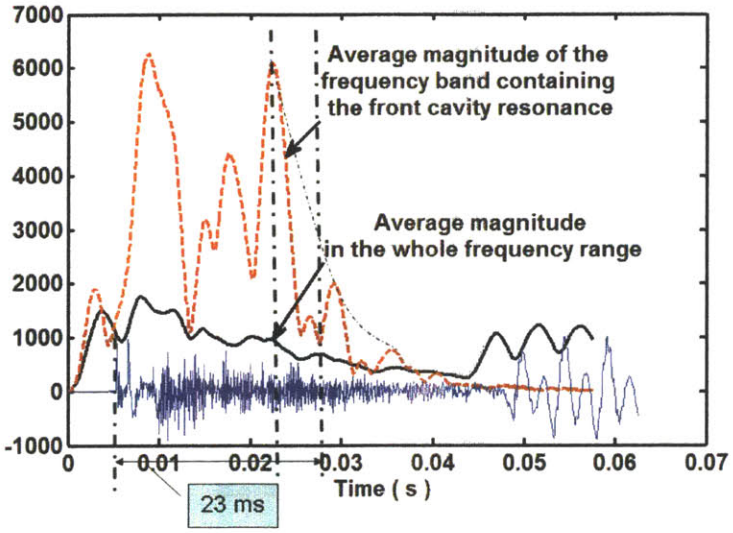


Fig. 5–1 The waveform of a /t/-release (in blue line) from the sentence "Don't ask me **to** carry an oily rag like that." (TIMIT\train\dr2\faem0\sa2). The average FFT magnitude Av in the whole frequency range (in black line), and the magnitude Ap in a frequency band containing the front cavity resonance (in red line). The duration of the frication noise is indicated as 23 ms, which is measured from the start of the release to the time when Ap (in red line) starts to drop rapidly; at the same time, Ap is at least twice of Av, and then corrected by adding the window length of 5 milliseconds.

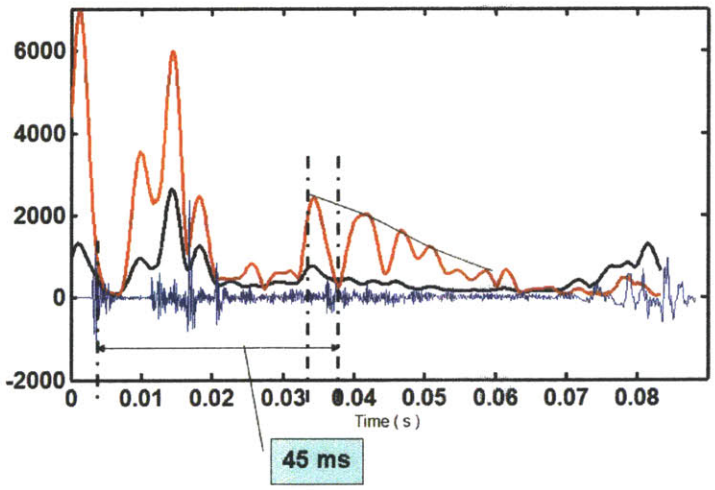


Fig. 5–2 The waveform of a /k/-release (in blue line) from the sentence "Don't ask me to **carry** an oily rag like that." (TIMIT\train\dr2\fcaj0\sa2). The duration of the frication noise is 45 ms.

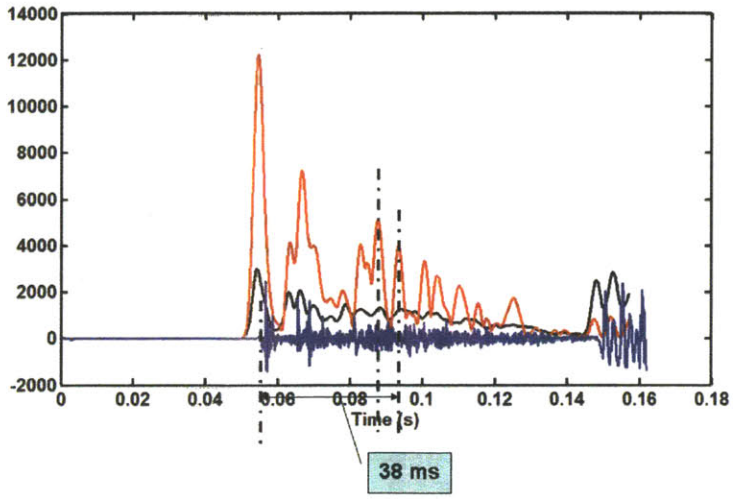


Fig. 5–3 The waveform of a /k/-release (in blue line) from the sentence "Don't ask me to **carry** an oily rag like that." (TIMIT\train\dr2\ faem0\sa2). The duration of the frication noise is 38 ms.

This newly defined acoustic measure can not be applied to all stop releases. In some stops consonants, the average FFT magnitude of the resonant frequency band increases after an initial drop, or remains large in the entire duration of the release burst. An example of such a /k/ release is shown in Fig. 5–4. Hanson and Stevens [47] also discussed this type of stop release. They found in the acoustic data collected from eight subjects that some subjects followed the phases described by Fant [6] (refer to Section 1.4.3), but other subjects produced a mixed frication and aspiration noise during the third phase.

Since the glottal constriction area A_g is assumed constant in the analysis, the calculated duration is not applicable to the entire duration of the stop consonant. In the later part of the release signal, the glottal constriction area A_g has to be reduced in order to prepare for the vibration of vocal folds.

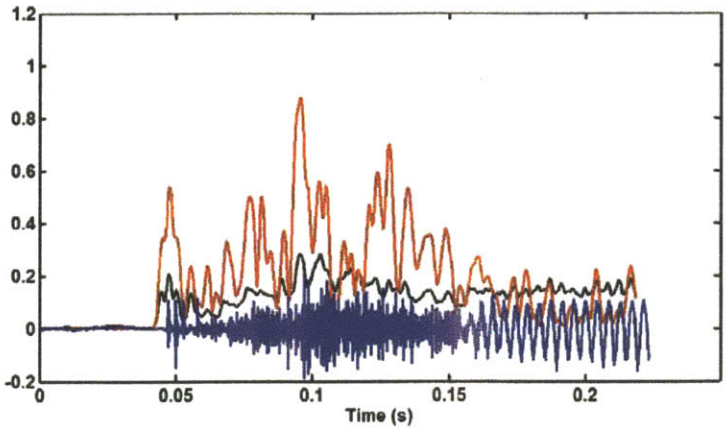


Fig. 5–4 An examples of a /k/ release with dominant frication noise throughout the release burst.

We have looked at 32 voiceless stop consonants in syllable-initial position, and found that the newly defined measure can be applied to 25 releases. The mean of the frication noise for the /t/-releases is 19.4 ms with std of 10.2 ms, and the mean for the /k/-releases is 25.9 ms with std of 15.2 ms. The means of both types of release are closer to the duration of the frication noise estimated from the calculated release trajectories with Flow model 3 (unsteady compressible flow) (11.5 ms for a /t/-release, and 23 ms for a /k/ release). From Flow model 1 (quasi-steady incompressible), the duration of the frication noise is estimated to be 10 ms for both /t/- and /k/- releases.

Stevens also estimated from acoustic data the duration of the plateau for labial and velar releases. "This duration can be as short as 5 milliseconds for a labial release, and can be 20 milliseconds or longer for a velar release. These durations are apparent in the length of the burst that occurs at the consonant release"[48]. The mean of the /k/-releases we have measured is also close to this estimation.

The durations measured in the acoustic signal (the mean), calculated with Flow model 1 (quasi-steady incompressible flow) and 3 (unsteady compressible flow), and estimated by Stevens [48] are listed in Table 5-1 respectively. The durations calculated with Flow model 3 is found closest to those measured in the acoustic signal.

Table 5-1 Comparison of the durations measured in the acoustic signal, calculated with two flow models models, and estimated by Stevens [48]

Place of articulation	The mean of the duration measured in the acoustic signal (millisecond)	Duration of the frication noise estimated by Stevens (2001) (millisecond)	Duration calculated with Model 1 (quasi-steady incompressible) (millisecond)	Duration calculated with Model 3 (unsteady-compressible) (millisecond)
/p/		5	6.3	8.3
/t/	**19.4**		10	**11.9**
/k/	**25.9**	20	14.6	**23**

In general, the pressure-wall interaction during the release of a stop closure causes a longer duration of the frication noise in the acoustic domain. This extended duration may enhance the perception of the place of articulation of a voiceless stop consonant in syllable-initial positions.

For a stop consonant in syllable-initial positions, the release burst [18]-[21] and formant transitions in the vowel [49]–[54] have been recognized as containing salient acoustic cues to the place of articulation.

To listeners, the release burst could be more important for voiceless aspirated stops than their voiced cognates, as the other cue, formant transitions, may not exist in the vowel[55], [56]. The waveform and spectrogram of an utterance /uhtA/ containing a voiceless stop consonant /t/ is shown in Fig. 5–5a, and an utterance /uhdA/ containing its voiced cognate /d/ is shown in Fig. 5–5b. Traceable formant movements are visible after the onset of voicing in /uhdA/, but not in /uhtA/. The formants after the onset of voicing in /uhtA/ are in straight lines and have no observable movement. Hence, for voiceless aspirated stop consonants, when formant transition cues are not present, listeners' perception of place relies on the release burst only. In support of this claim, the

perceptual effect of the amplitude of the release burst relative to vowel in classifying the place of articulation, was also found greater on voiceless stops than on voiced ones [57].

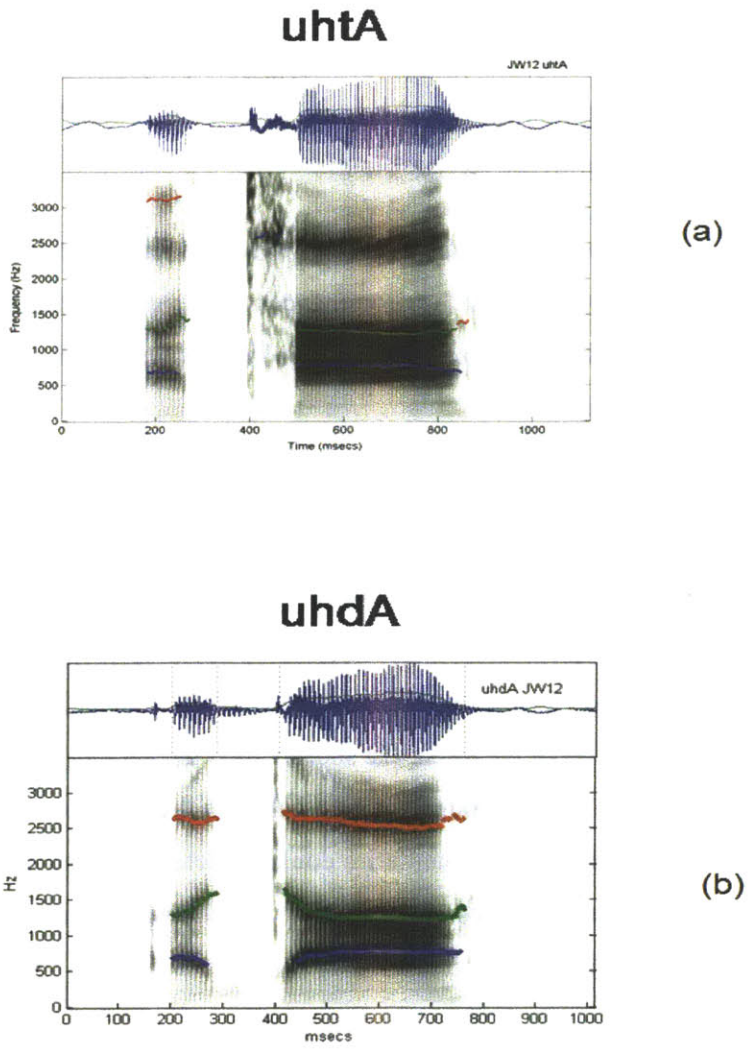


Fig. 5–5 (a) Waveform and spectrogram of utterance /uhtA/; (b) Waveform and spectrogram of utterance /uhdA/. Formants are tracked automatically with a code written by Mark Tiede, and are shown in the spectrogram in blue, green, and read thick lines. The onset of release starts around 400 ms in both /t/ and /d/. In (a), the formants are straight without any apparent movements after the voicing onset following /t/; while formant movements are observable right after the release of /d/.

As a short interval of explosive noise at the onset of a stop consonant, the release burst contains a vertical line in the spectrogram, whose spectral content [18],[19],[58] and amplitude relative to the vowel[57] have been found to carry cues for the place of articulation. Moreover, temporal cues could also exist in the release burst, as the auditory processing is both in the time and frequency domain. As correlates to the place of articulation with both temporal and spectral features, Kewley-Port [20] proposed the time-varying spectral play of the release burst.

It has been well understood that the place of articulation is reflected in the spectral content of the release burst by the prominences resulting from the filtering of the front cavity. However, temporal cues in the release burst have not been well studied in relation to the place of articulation.

The perceptual sense of both the amplitude and the duration of the release burst could be related to its physical duration T. First, the perception of the amplitude of a burst of noise has been found to depend on its duration[59]. Second, the JND (Just Noticeable Difference) for changes in duration is approximately proportional to the square root of T [60] when the duration T of a burst of noise is less than about 100 ms. Therefore, the shorter is the duration of the burst, the shorter JND the ear can sense. Listeners are thus expected to be able to discriminate at least between the duration of the burst of /p/ and that of /t, k/. The categorical boundaries could be determined by playing to listeners synthesized CV (consonant-vowel) syllables with varied duration of the release bursts whose amplitude is properly controlled.

The model results and tentative experimental results from real speech data show that the duration of the frication noise in the release burst is different for syllable-initial voiceless stop consonants with different places of articulation. Based on this fact and other observations introduced above, we would suggest investigating whether the duration of the release burst is a possible temporal cue or not for the place of articulation.

5.3 The contact pressure at the time of release

The contact pressure at the time of release is an unknown parameter in the model analysis. For alveolar and velar stop consonants, this contact pressure is between the tongue and hard palate, also called the lingua-palatal contact pressure (LPCP). This parameter is hard to be measured accurately in real time.

LPCP has been considered as an important parameter for evaluating the dynamic properties of the tongue during palatal consonant production [41]. It has been studied for various reasons: speech physiology, sensory-motor feedback system, and articulation [61]. Measuring the magnitude of LPCP during speech production has been attempted by several groups and with different methods [42], [41], [62].

McGlone et al.[42] used resistance strain gauges to measure the contact pressure at three locations (left, center, and right) along the palate during the production of alveolar consonants by 10 young adults. The frequency response of the pressure recording system is linear up to 180 Hz, which means the time resolution is more than 5 ms. This dynamic response is slow for real-time LPCP during the release of a stop closure.

Matsumura, et al.[41] used five piezoelectric strain gauges to measure the LPCP and contact pattern for an adult male speaker in producing /t/, /d/, and /n/. The dynamic response of their force sensor is adequate: in 0.28 ms, the output changes from 90% level to 10% level when a load of 5-gram weight is released. The measured temporal pattern of the LPCP in /t/ is very different from the pattern in /d/ and /n/. This approach could be the most promising method for measuring the contact pressure at the time of release.

Tiede, et al.[62] applied capacitive sheets for pressure sensing, and measured the LPCP differences in English lingual obstruent production for one subject. However, in this latest study, real-time measurement was found not reliable because of pervasive noise and the lack of a means of calibration over the full range of expected pressure limits.

In order to better understand the contact pressure at the time of release, statics is employed to analyze the force balance when a complete closure is retained against intraoral pressure buildup.

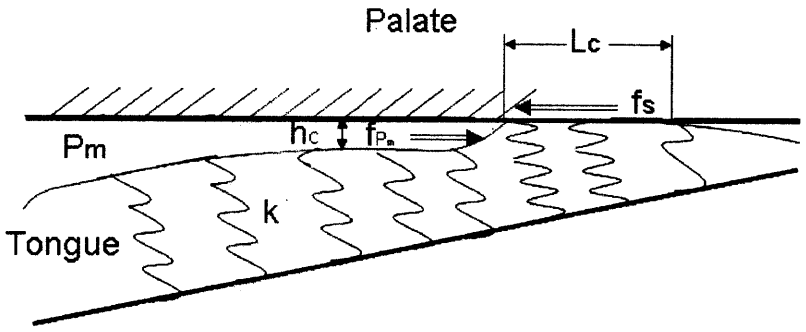


Fig. 5–6. A portion of the tongue making a complete closure of length L_C. The soft tissue surface of the tongue is represented as a layer of springs with stiffness *k* per unit area. Part of the tongue surface is in contact with the palate and part of it is exposed to the intraoral pressure P_m. The intraoral pressure presses the tongue surface downward of height h_c and also applies a horizontal force of f_{P_m} on the part of the tongue in contact with the palate. This force is balanced by the frictional force f_s between the surfaces in contact. Shear forces inside the tongue tissue are neglected.

Fig. 5-6 shows a portion of the tongue making a complete closure of axial length L_C with the palate. The dimension perpendicular to the midsagittal plane is considered in unit length. The yielding wall of tongue surface is represented as a layer of springs with stiffness constant per unit area of *k*. Air pressure is built up behind the closure, with magnitude P_m. This intraoral pressure causes a downward displacement of $h_c = P_s/k$ on the tongue surface upstream of the closure, and also applies a horizontal force of f_{P_m} on the part of the tongue in contact with the palate.

To satisfy the horizontal force balance,

$$f_{P_m} = P_m h_c = f_s \leq f_{sm}.$$

f_s is the static frictional force between the tongue surface and the palate, which is smaller than or equal to the maximum static frictional force f_{sm}, for closures of fixed location. The maximum frictional force f_{sm} is determined from the equation $f_{sm} = \mu_m P_c L_c$, in which P_C is the contact pressure; μ_m is the maximum static frictional coefficient; and the contact area equals L_C.

In case the location of closure keeps moving during the closure interval[31], the static frictional force would be replaced by the kinetic frictional force $f_k = \mu_k P_c L_c$, in which f_k is the kinetic frictional force and μ_k is the kinetic frictional coefficient.

For a closure with fixed location, we have

$$\mu_m P_C L_C \geq P_m h_C .$$

As $h_C = \frac{P_m}{k}$, we also have

$$P_C \geq \frac{P_m{}^2}{\mu_m k L_C}$$

This inequality shows that in order to make an air-tight closure, the contact pressure needs to be larger than or equal to the quantity on the right hand side. This quantity is the minimum contact pressure required to retain a complete closure, and also is the contact pressure at the time of release in the analysis discussed in Chapter 3.

As the minimum contact pressure required to retain a complete closure $\frac{P_m{}^2}{\mu_m k L_c}$ is proportional to the square of the intraoral pressure, a greater contact pressure is required to seal the closure for voiceless stop consonants than for voiced ones, and this has been reported by McGlone et al.[42] and Matsumura et al.[41].

From the minimum contact pressure required to retain a complete closure, we can also calculate a negative displacement of the primary articulator, $\frac{P_m^2}{\mu_m L_c k^2}$. In Chapter 2, we discussed that Stevens' 2-section model of fricative production does not give the lower limit of an initial positioning of the primary articulator in making a constriction. This lower limit missing in the fricative model is exactly the negative displacement just derived.

In causal speech, stop consonants have been observed to be occasionally produced as fricatives, and vice verse[63], [64], [65]. We can now explain this phenomenon: when the contact pressure is lower than the minimum amount required for retaining an air-tight closure, a fricative consonant is produced instead of a stop consonant. On the other hand, when the contact pressure resulted from an initial negative positioning is larger than the minimum amount of retaining a complete closure, or when the initial negative positioning has passed the lower limit of a fricative constriction, a stop consonant is produced in stead of a fricative.

The form $\frac{P_m^2}{\mu_m k L_c}$ also suggests that the contact pressure at the time of release could be measured indirectly by measuring the intraoral pressure P_m, the maximum static frictional coefficient μ_m, the stiffness of the wall k, and the length of the constriction L_c. These parameters are not dynamic, so they may be acquired more easily than the real-time LPCP.

The form $\frac{P_m^2}{\mu_m k L_c}$ also indicates that the contact pressure at the time of release is a function of the length of the constriction. However, this correlation was not implemented in the analysis in Chapter 3.

At last, a relevant physiological variable of the tongue is discussed, the tongue strength. Tongue strength is evaluated by measuring the maximum pressure that a person can produce in an air-filled bulb which is pressed against the hard palate. It can be

measured with an instrument called IOPI, the Iowa Oral Performance Instrument[66]. The measured tongue strength is the maximum lingua-palatal contact pressure a person can produce, but not during speech production.

Tongue strength is an important parameter for evaluating the chewing and swallowing disorders. Adequate lingual pressure is essential for bolus clearance. Potter and Short[67] measured the tongue strength in 150 children and adolescents at 3-16 of age. Their tongue strength was in the range from 20kPa to 90kPa, and increases with age. The measured tongue strength values are all far above the 5-6kPa maximum LPCP during adults' alveolar consonant production introduced above.

Tongue weakness is presumed to contribute to reduced articulatory precision and speech intelligibility, but studies of dysarthric speakers did not consistently find a correlation between tongue weakness and perceptible speech deficits[68]. Solomon, et al. [68] measured the tongue strength of 16 Parkinson's patients, and found a mean of 48.1kPa, compared with the mean of 55.5kPa on a control group of 16 neurologically normal adults.

Parkinson's patients have reduced tongue strength compared with the normal group; however, they still have the ability to produce the adequate maximum lingua-palatal contact pressure required for consonant production, which is only 12.5% of the mean tongue strength they have. This is a plausible explanation for the lack of correlation between tongue weakness and perceptible speech deficits.

Although the tongue strength in Parkinson's have not been found to be weakened to the extent that alveolar consonants would be degraded, significant positive correlations between severity of dysarthria and tongue strength in ALS (*Amyotrophic lateral sclerosis*) have been reported[69], [70](measured force in Newton, not pressure in Pa). The tongue strength of an ALS patient with severe dysarthria may be lower than the maximum LPCP required in consonant production.

Chapter 6

Conclusions and future work

The pressure-wall interaction at the release of a stop closure is analyzed with a lumped-element solid model and three flow models (quasi-steady incompressible, unsteady incompressible, and unsteady compressible). Analyses with the three models are summarized in Table 6-1.

Table 6-1 Comparison of the models

Models	Assumptions	Results	Conclusions
Flow model 1 (quasi-steady incompressible)	1. ΔP=RU+L$\frac{dU}{dt}$ 2. P_{c_rel}	1.Insensitive to A_{c_start} when $A_{c\_start} \le 0.02cm^2$ (for one case study) 2.Insensitive to L_c 3.The predicted duration of frication noise is shorter than the measured value on real speech	Good for larger cross-sectional area.
Flow model 2 (unsteady incompressible)	1. $\rho = const.$ 2. P_{c_rel}	1.The plateau becomes a peak when $A_{c\_start} \le 0.001cm^2$ (for one case study) 2.Sensitive to L_c for small A_{c_start}	Good for larger cross-sectional area
Flow model 3 (unsteady compressible)	1. $\rho = \rho(t,x)$ 2. P_{c_rel}	1.Release trajectories are close to each other when $A_{c\_start} < 0.0001cm^2$ (for one case study) 2.Sensitive to L_c for small A_{c_start} 3.The predicted duration of frication noise is closest to the measured value on real speech among all estimates.	Also good for small cross-sectional area

The calculated release trajectories confirm Stevens' hypothesis that a plateau would emerge right after the release, and the duration of the plateau would be progressively longer for labial, alveolar, and velar (Section 4.3).

An acoustic measure is defined to estimate the duration of frication noise in the release bursts of real speech. The means of the measured durations of the frication noise in 25 releases from TIMIT database agree better with the durations of the frication noise estimated from the release trajectories calculated with the unsteady compressible flow model, compared with those calculated with the quasi-steady incompressible flow model (Section 5.2).

Perception test is suggested to find out whether the duration of the frication noise could be an enhancing perceptual cue or not for identifying the place of articulation of voiceless stop consonants in syllable-initial positions.

In developing the unsteady flow models, a new analysis method – deformable control volume analysis – is introduced. It lays the foundation for analyzing the physical processes in the initial phases of a stop consonant (both the transient and the frication noise). Current speech production theory has defined these phases according to their salient acoustic feature, but lacked proper means to quantify them. For example, a previous analysis of the transient had to use an incomplete theoretical model[22]. Hence, analysis of the transient phase is recommended as the most immediate future work related to stop consonant models.

The deformable control volume analysis also contributes a new method to treating the unsteady effect during the closing and opening of the vocal folds [71], [36], [37]. Implementing this analysis in the two-mass model would dismiss the quasi-steady assumption.

At last, because of the instrumental difficulty in measuring the real-time LPCP, an indirect means of measuring the contact pressure at the time of release, an unknown parameter in the analysis, is suggest: by measuring the intraoral pressure P_m, the

maximum static frictional coefficient μ_m, the stiffness of the wall *k*, and the length of the constriction L_c, the contact pressure at the time of release could be acquired as $\frac{P_m^{\,2}}{\mu_m k L_c}$.

Bibliography

[1] Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: The MIT Press.

[2] Stevens, K. N. (2002) "Toward a model for lexical access based on acoustic landmarks and distinctive features" J. Acoust. Soc. Am., 111(4). 1872-1891.

[3] Hanson, H. and K. N. Stevens (2002). "A quasiarticulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using *HLsyn*" J. Acoust. Soc. Am. Volume 112, Issue 3, pp1158-1182.

[4] Chiba, T. and M. Kajiyama (1958) The vowel: its nature and structure, Phonetic Society of Japan

[5] Jakobson, R., Gunnar, F., and Morris, H. (1952) "Preliminaries to Speech Analysis." Technical Report no.13. Cambridge: MIT Acoustics Laboratory.

[6] Fant, G. (1960). Acoustic theory of speech production. The Hague: Mouton.

[7] Fant, G. (2004). "The relations between area functions and the acoustic signal." In *Speech acoustics and phonetics: Selected writings.* Dordrecht: Kluwer. 29-60. (Original work published 1980)

[8] Hanson, H. M., and K. N. Stevens (2000). "Modeling stop-consonant releases for synthesis (A)" J. Acoust. Soc. Am. 107, 2907.

[9] Westbury, J. R. (1994). *X-ray microbeam speech production database -user's handbook.* (University of Wisconsin, Madison, WI).

[10] Zemlin, W. R. (1988). *Speech and Hearing Science*. Englewood Cliffs, NJ: Prentice-Hall.

[11] Rothenberg, M. (1968). "The breath-stream dynamics of simple-released-plosive production." Biblioteca Phonetica No. 6, S. Karger AG, Basel

[12] Heinz, J. M. (1956). "Fricative consonants". Acoustic Laboratory Quarterly Report. October-December. MIT, Cambridge, MA. 5-7.

[13] Jw. van den Berg, J., J. T. Zantema, and P. Doornenbal, Jr. (1957) "On the Air Resistance and the Bernoulli Effect of the Human Larynx" J. Acoust. Soc. Am. 29(5), pp. 626-631

[14] Ishizaka, K., French, J. C., and Flanagan, J. L. (1975). "Direct determination of vocal tract wall impedan," IEEE Transactionon Acoustics, Speech, and Signal Processing ASSP-23, 370-373.

[15] Svirsky, M. A., K. N. Stevens, M. L. Matthies, J. Manzella, J. S. Perkell, and R. Wilhelms-Tricarico (1997). "Tongue surface displacement during obstruent stop consonants". J. Acoust. Soc. Am. 102, 562-571.

[16] Westbury, J. R. (1983). "Enlargement of the supraglottal cavity and its relation to stop consonant voicing," J. Acoust. Soc. Am. 73, 1322–1336.

[17] Shadle, C. (1985) The acoustics of fricative consonants. RLE Technical Report 506, MIT, Cambridge, MA.

[18] Stevens, K. N. and Blumstein, S. E. (1978). " Invariant cues for place of articulation in stop consonants," J. Acoust. Soc. Am. 64(5), 1358-1368.

[19] Blumstein, S. E. and Stevens, K. N. (1979). "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants," J. Acoust. Soc. Am. 66(4), 1001-1016.

[20] Kewley-Port, D. (1983). "Time-varying features as correlates of place of articulation in stop consonants." J. Acoust. Soc. Am., 73, 322-335.

[21] Lahiri, A., Gewirth, L., and Blumstein, S. E. (1984). "A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study," J. Acoust. Soc. Am. 76(2), 391-404.

[22] Massey, N. S. (1994). Transients at stop-consonant release, S. M. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

[23] McGowan, R (1992). "Tongue-tip trills and vocal-tract wall compliance," J. Acoust. Soc. Am. 91(5), 2903-2910.

[24] Flanagan, J. L. and Landgraf, L. L. (1968). "Self-oscillating source for vocal tract synthesizers," IEEE Trans. Audio Electroacoust. AU-16, 57-64.

[25] Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. 51, 1233-1268.

[26] Story, B. H. and I. R. Titze (1995) "Voice simulation with a body-cover model of the vocal folds". J. Acoust. Soc. Am. 97(2), 1249-1260

[27] Titze, I. R. (1973), "The Human Vocal Cords: A Mathematical Model. Part I," Phonetica 28, 129-170.

[28] Titze, I. R. (1974), "The Human Vocal Cords: A Mathematical Model. Part II," Phonefica 29, 1-21.

[29] Berry, D. A., and Titze, I. R. (1996). "Normal modes in a continuum model of vocal fold tissues," J. Acoust. Soc. Am. 100(5), 3345–3354.

[30] Alipour, F., Berry, D. A., and Titze, I. R. (2000). "A finite-element model of vocal-fold vibration," J. Acoust. Soc. Am. 108(6), 3003–3012.

[31] Löfqvist, A. and Gracco, V. L. (2002). "Control of oral closure in lingual stop consonant production," J. Acoust. Soc. Am. 111(6), 2811-2827.

[32] Zanartu, M., Mongeau, L., and Wodicka, G. R. (2007). "Influence of acoustic loadings on an effective single mass model of the vocal tract," J. Acoust. Soc. Amer. 121(2), 1119-1129.

[33] Zhang, Z., L. Mongeau, and S. H. Frankel (2002). "Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes", J. Acoust. Soc. Amer. 112(4), 1652-1663.

[34] Scherer, R. C., S. Torkaman, B. R. Kucinschi, and A. A. Afjeh (2010). "Intraglottal pressure in a three-dimensional model with a non-rectangular glottal shape". J. Acoust. Soc. Am., 128(2), 828-838.

[35] Mongeau, L., N. Franchek, C. H. Coker, and R. A. Kubli (1997). "Characteristics of a pulsating jet through a small modulated orifice, with application to voice production" J. Acoust. Soc. Am. 102(2), 1121-1133.

[36] Pelorson, X. (2001). "On the meaning and accuracy of the pressure-flow technique to determine constriction areas within the vocal tract". Speech Communication 35

[37] Vilain, C. E., X. Pelorson, C. Fraysse, M. Deverge, A. Hirschberg, J. Willems (2004). "Experimental validation of a quasi-steady theory for the flow through the glottis". Journal of Sound and Vibration. 276, 475-490.

[38] Tao, C. and J. J. Jiang (2006) "Anterior-posterior biphonation in a finite element model of vocal fold vibration" J. Acoust. Soc. Am. 120 (3), 1570-1577

[39] Luo, H., R. Mittal, and S. A. Bielamowicz (2009) "Analysis of flow-structure interaction in the larynx during phonation using an immersed-boundary method" J. Acoust. Soc. Am. 126 (2), 816-824

[40] Zhang, X., Q. Xue and R. Mittal (2010). "A coupled sharp-interface immersed-boundary-finite-element method for flow-structure interaction with application to human phonation" J Biomech Eng. 132(11):111003 (12pages).

[41] Matsumura, M., Kimura, K., Yoshino, K., Tachimura, T., and Wada, T. (1994). "Measurement of palatolingual contact pressure during consonant productions using strain gauge transducer mounted palatal plate," in *Proceedings of ICSL94* (Yokohama), Vol. 2, pp. 655-658.

[42] McGlone, R., Proffit, W., and Christensen, R. (1967). "Lingual pressure associated with alveolar consonants," J. Speech Hear. Res.10, 606-615.

[43] Deverge, M., X. Pelorson, C. Vilain, P.-Y. Lagrée, F. Chentouf, J. Willems, and A. Hirschberg (2003) "Influence of collision on the flow through in-vitro rigid models of the vocal folds." J. Acoust. Soc. Am. 114(6), 3354-3362

[44] White, F. (1999) Fluid Mechanics (Forth edition). Mcgraw-Hill.

[45] Wegel, R. (1930) "Theory of vibration of the larynx," Bell Syst. Tech. J. 9, 207–227

[46] Fisher, W. M., G. R. Doddington, and K. M. Goudie-Marshall (1986). "The DARPA speech recognition research database: specifications and status". *Proceedings of DARPA Workshop on Speech Recognition*. pp. 93–99.

[47] Hanson, H. and K. N. Stevens (2006). "The nature of aspiration in stop consonants in English (A)" J. Acoust. Soc. Am. Volume 119, Issue 5, 3393-3394

[48] Stevens, K. N. (2001) "The properties of the vocal tract walls help to shape several linguistic distinctions in language" Travaux du Cercle Linguistique de Copenhague XXXI, *Nina Grønnum & Jørgen Rischel (eds.)*: To honour Eli Fischer-Jørgensen. 2001. pp285-297

[49] Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M. and Gerstman, L. J. (1952). " Some experiments on the perception of synthetic speech sounds," J. Acoust. Soc. Am. 24(6), 597-605.

[50] Delattre, P., Liberman, A., and Cooper, F. (1955). "Acoustic loci and transitional cues for consonants," J. Acoust. Soc. Am. 27, 769-773.

[51] Dorman, M. F., Studdert-Kennedy, M., and Raphael, L.J. (1977). "Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues," Perception and Psychophysics 22, 109-122.

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

[53] Krull, D. (1987). "Second formant locus patterns as a measure of consonant-vowel coarticulation," Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm, PERILUS V, 43-61.

[54] Susman, H. M., McCaffrey, H. A., and Matthews, S. A. (1991). "An investigation of locus equation as a source of relational invariance for stop place categorization," J. Acoust. Soc. Am. 90, 1309–1325.

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

[56] Löfqvist, A. (1999). "Interarticulator phasing, locus equations, and degree of coarticulation," J. Acoust. Soc. Am. 106(4 Pt1), 2022-2030.

[57] Ohde, R. N. and Stevens, K. N. (1983). "Effect of burst amplitude on the perception of stop consonant place of articulation," J. Acoust. Soc. Am. 74(3), 706-714.

[58] Blumstein, S. E., and K. N. Stevens (1980) "Perceptual invariance and onset spectra for stop consonants in different vowel environments." J. Acoust. Soc. Am.67(2), 648-62.

[59] Zwislocki, J. J. (1969). "Temporal summation of loudness: An analysis," J. Acoust. Soc. Amer. 46(2), 431-441.

[60] Delgutte, B. (1980). "Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers," J. Acoust. Soc. Amer. 68, 843-857.

[61] Searl, J. P. (2003). "Comparison of transducers and intraoral placement options for measuring lingua-palatal contact pressure during speech." Journal of Speech, Language, and Hearing Research, 46, 1444-1456.

[62] Tiede, M., Perkell, J., Zandiour, M., Matthies, M., and Stockmann, (2003) "A new approach to pressure sensitive palatography using a capacitive sensing device" Proc. of 15th International Congress of Phonetic Sciences, 3149-3152.

[63] Lavoie, L. M. (2001) Consonant strength: Phonological patterns and phonetic manifestations. NewYork: Garland

[64] Degenshein, R. (2004) "Dholuo interdentals: fricatives or affricates? evidence from domain-initial strengthening." From Sound to Sense. June 11-June 13, 2004, MIT.

[65] Jones, M. J. and K. McDougall (2009) " The acoustic character of fricated /t/ in Australian English: A comparison with /s/ and /ʃ/ " J. International Phonetic Association. 39(3), 265-289

[66] Solomon, N. P., D. A. Robin, S. I. Mitchinson, D. J. VanDaele, and E. S. Luschei (1996). "Sense of effort and the effects of fatigue in the tongue and hand." *Journal of Speech and Hearing Research, 39,* 114-125.

[67] Potter, N. L. and R. Short (2009). "Maximal tongue strength in typically developing children and adolescents." Dysphagia 24, 391-397

[68] Solomon, N. P., D. A. Robin, and E. S. Luschei (2000). "Strength, endurance, and stability of the tongue and hand in Parkinson disease" Journal of Speech, Language, and Hearing Research, 43, 256-267.

[69] DePaul, R. and B. R. Brooks (1993). "Multiple orofacial indices in amyotrophic lateral sclerosis." Journal of Speech and Hearing Research, 36, 1158–1167

[70] Langmore, S. E. and M. E. Lehman (1994). "Physiologicdeficits in the orofacial system underlying dysarthria in amyotrophic lateral sclerosis." Journal of Speech and Hearing Research, 37, 28–37.

[71] McGowan, R (1993). "The quasi-steady approximation in speech production." J. Acoust. Soc. Am. 94(5), 3011-3013.

[72] Gunter, H. E. (2003). "A mechanical model of vocal-fold collision with high spatial and temporal resolution," J. Acoust. Soc. Am. 103, 994-1000.

[73] Warren, DW, Dalston, RM, Morr, KE, Hairfield, WM and Smith, LR (1989) "The speech regulating system -temporal and aerodynamic responses to velopharyngeal inadequacy". Journal of Speech and Hearing Research Vol.32 566-575.

[74] Kim, J., Zajac, D. J., Warren, D. W., Mayo, R., Essick, G. K. "The response to sudden change in vocal tract resistance during stop consonant production." Journal of Speech, Language,and Hearing Research 40, 848-857.

Appendices

A. Derivations of the perturbation approximation in Flow model 3

If we consider a tube of length 1 cm, then the maximum value of x is 0.5cm, and we further simplify $\rho(x,t) = a_0(t) + a_1(t)x$ by neglecting all terms of order equal to and higher than x^2.

Substitute $\rho(x,t)$ in Equation (27), and we have

$$u_e = \frac{1}{\rho_e A_c}\frac{d}{dt}\int_0^{L_c/2}(a_0 + a_1 x)A_c dx = \frac{1}{\rho_e A_c}\frac{d}{dt}\left(A_c\left(a_0 \cdot \frac{L_c}{2} + a_1 \cdot \frac{L_c^{\,2}}{8}\right)\right) \tag{36}$$

Form Equation (26) and Equation (27), we have

$$\rho u = \rho_e u_e - \frac{1}{A_c}\frac{d}{dt}\int_0^x \rho A_c dx = \rho_e u_e - \frac{1}{A_c}\left(\frac{dA_c}{dt}\int_0^x \rho dx + A_c\int_0^x \frac{d\rho}{dt}dx\right) \tag{37}$$

Substitute $\rho(x,t) = a_0(t) + a_1(t)x$ in this equation, and we have

$$u = \frac{\rho_e u_e - \left(\frac{a_0}{A_c}\frac{dA_c}{dt} + \frac{da_0}{dt}\right)x - \left(\frac{a_1}{2A_c}\frac{dA_c}{dt} + \frac{1}{2}\frac{da_1}{dt}\right)x^2}{a_0 + a_1 x} \tag{38}$$

By neglecting all terms of order equal to and higher, u is approximated as

$$u \simeq \frac{\rho_e u_e}{a_0} - \left(\frac{\rho_e u_e a_1}{a_0^{\,2}} + \frac{1}{A_c}\frac{dA_c}{dt} + \frac{1}{a_0}\frac{da_0}{dt}\right)x \tag{39}$$

$$|u| \simeq \left|\frac{\rho_e u_e}{a_0}\right| - \left|\frac{\rho_e u_e a_1}{a_0^{\,2}} + \frac{1}{A_c}\frac{dA_c}{dt} + \frac{1}{a_0}\frac{da_0}{dt}\right|x \tag{40}$$

$$\rho u = \rho_e u_e - \left(\frac{a_0}{A_c} \frac{dA_c}{dt} + \frac{da_0}{dt} \right) x \tag{41}$$

$$\rho u|u| = \rho_e u_e \left| \frac{\rho_e u_e}{a_0} \right| - \left(\left| \frac{\rho_e u_e}{a_0} \right| \left(\frac{a_0}{A_c} \frac{dA_c}{dt} + \frac{da_0}{dt} \right) + \rho_e u_e \left| \frac{\rho_e u_e a_1}{a_0^2} + \frac{1}{A_c} \frac{dA_c}{dt} + \frac{1}{a_0} \frac{da_0}{dt} \right| \right) x \tag{42}$$

$$\begin{aligned} \frac{d}{dt} \int_0^x \rho u A_c dx &= \frac{d}{dt} \left(A_c \int_0^x \left(\rho_e u_e - \frac{1}{A_c} \frac{d}{dt} \int_0^x (a_0 + a_1 x) A_c dx \right) dx \right) \\ &= \frac{d}{dt} \left(A_c \int_0^x \left(\rho_e u_e - \frac{1}{A_c} \frac{d}{dt} \left(A_c \left(a_0 x + \frac{a_1}{2} x^2 \right) \right) \right) dx \right) \\ &= \frac{d}{dt} \left(A_c \left(\rho_e u_e x - \frac{1}{A_c} \frac{d}{dt} \left(A_c \left(a_0 \frac{x^2}{2} + \frac{a_1}{6} x^3 \right) \right) \right) \right) \end{aligned} \tag{43}$$

Neglect all terms of order equal to and higher than x^2, and then $\frac{d}{dt} \int_0^x \rho u A_c dx$ is approximated as $x \frac{d}{dt} (A_c \rho_e u_e)$.

$$(P_e - p_x) A_c = \frac{d}{dt} \int_0^x \rho_x u_x A_c dx + \left(\rho_x u_x |u_x| - \rho_e u_e |u_e| \right) A_c \tag{44}$$

From Equation (30), we have $p = \frac{P_e}{\rho_e} \rho \simeq \frac{P_e}{\rho_e} (a_0 + a_1 x)$

Now the right-hand side of Equation (28) is approximated as

$$\begin{aligned} & x \frac{d}{dt} (A_c \rho_e u_e) - \rho_e u_e |u_e| + \\ & A_c \left(\rho_e u_e \left| \frac{\rho_e u_e}{a_0} \right| - \left(\left| \frac{\rho_e u_e}{a_0} \right| \left(\frac{a_0}{A_c} \frac{dA_c}{dt} + \frac{da_0}{dt} \right) + \rho_e u_e \left| \frac{\rho_e u_e a_1}{a_0^2} + \frac{1}{A_c} \frac{dA_c}{dt} + \frac{1}{a_0} \frac{da_0}{dt} \right| \right) x \right) \end{aligned}$$

and the left-hand side is approximated as $\left(P_e - \frac{P_e}{\rho_e}(a_0 + a_1 x)\right) A_c$. For this equation to hold, the coefficients of the x^0-term requires that

$$A_c\left(\rho_e u_e \left|\frac{\rho_e u_e}{a_0}\right|\right) - \rho_e u_e |u_e| A_c = \left(P_e - \frac{P_e}{\rho_e} a_0\right) A_c, \tag{45}$$

and the coefficients of the x^1-term requires that

$$\frac{d}{dt}\left(\frac{d}{dt}\left(A_c\left(a_0 \cdot \frac{L_c}{2} + a_1 \cdot \frac{L_c^2}{8}\right)\right)\right) - A_c\left(\left|\frac{\rho_e u_e}{a_0}\right|\left(\frac{a_0}{A_c}\frac{dA_c}{dt} + \frac{da_0}{dt}\right) + \rho_e u_e \left|\frac{\rho_e u_e a_1}{a_0^2} + \frac{1}{A_c}\frac{dA_c}{dt} + \frac{1}{a_0}\frac{da_0}{dt}\right|\right)$$
$$= -\frac{P_e}{\rho_e} a_1 A_c. \tag{46}$$

B. The stiffness of the yielding wall

Note that although the effect of muscle recruitment is known to increase the Young's modulus of the tongue, the tense cheek tissue measured by Ishizaka has a smaller stiffness, 3.33×10^3 dyne/cm, compared with the relaxed cheek tissue whose stiffness is 8.45×10^3 dyne/cm. The mechanical properties corresponding to the larger stiffness value is used in the analyses in Chapter 3, and the reason for this choice is provided here.

Intraoral pressure is an important parameter in stop consonant production, and the respiratory and articulatory structures have been found constrained to act together toward the common goal of retaining an adequate level of pressure for consonants.[73]. Moreover, receptors sensitive to the changes in the aerodynamic environment have been found in human lungs, trachea, larynx, nasopharynx and oral cavity [74].

When a complete closure is made against increased intraoral pressure during the closure interval of a stop consonant, the soft-tissue articulator involved can be also considered as a part of the wall of a pressurized tank, which has a circular radius of R and thickness of d, and is filled with air which has the intraoral pressure P_m, as shown in Fig. B-1a. Such a pressurized tank could be a model of a bilabial closure shown in Fig. B-1b. The tissue near to the closure is compressed in making a complete closure, but further away from the closure, the tissue is under tension owing to the expanding effect of the intraoral pressure.

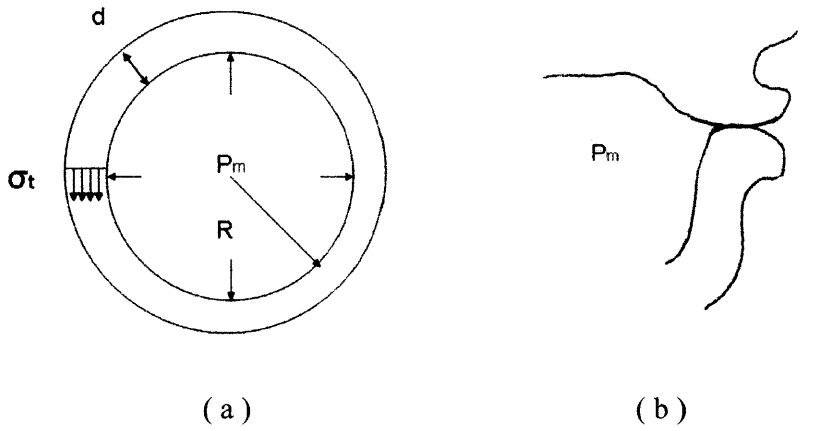


Fig. B-1 (a) Pressure Pm inside a circular tank with radius R and thickness d, and with tension σ_t inside the wall. (b) The tissue wall during the closure interval of a bilabial stop consonant /p/ can be taken as part of the wall of the tank shown in (a).

The tension in the soft-tissue articulator can be calculated from the force balance of an element in the wall. It can be shown that the circumferential stress σ_t in the wall is related to the pressure P_m, radius R, and thickness d as:

$$\sigma_t = \frac{P_m R}{d} \tag{47}$$

The circumferential tensile force in the wall is then

$$F = \sigma_t d = P_m R \tag{48}$$

This equation suggests that higher intraoral pressure would demand larger circumferential tension in the wall. As a tensed wall also exhibits larger surface stiffness, the stiffness of the wall would be found higher when the intraoral pressure is built up during the closure. This analysis agrees with what was found for the in vivo compliance C_s of the tongue surface by tracking the displacement of a flesh point during the closure portion of voiced and voiceless stop consonants.[5] The surface compliance was about 3 times larger in producing a voiceless stop consonant.

The pressurized tank model also suggests that the tissue properties of the yielding wall could be adjusted according to the level of the intraoral pressure, which is considered as equal to the subglottal pressure. Therefore, the stiffness of the wall might be a function of the subglottal pressure.

C. Matlab algorithms

The transcript file for the analysis with flow model 1 (quasi-steady incompressible):

```
clear all
rho=1.14e-3;
global Ac_start
Ac_start=0.000001;
Ps=7840;
k=84500;
y0=Ps/k;
global V
V=50; %cm^2/s
D=2;

[T,Y]=ode15s('two_mass_new_b',[0,100e-3],[0,0,0,0,0]')
Ac=(V*T+Ac_start-Y(:,1)*D);

figure(1)
plot(T,Ac,'k')
xlabel('Time s')
ylabel('Cross-sectional Area cm^2')


function [Ydot]=two_mass_new_b(t,Y)
Ag=0.2; % Length scale is in cm.
D=2;
Ps=7840; %10cmH2O in dyne/cm^2
Pc_min=3*Ps;
rho=1.14e-3;
Lc=1;
m=2.1;
r=800;
k=84500;
rm=r/m;
km=k/m;
kc_m=1.5*km;
y_c=Pc_min/k;
y0=Ps/k;
global V
y1eq=y_c;
y2eq=y0;
global Ac_start

Ac=(V*t+Ac_start-Y(1,:)*D);
L=rho*Lc/Ac;
F_kc=-kc_m*((Y(1,:)-y1eq)-(Y(3,:)-y2eq));
```

```
Pm=Ps-rho/2*(Y(5,:)/Ag)^2;

Ydot=[Y(2,:);...
-rm*Y(2,:)-km*(Y(1,:)-y1eq)+F_kc-Pm/2;...
Y(4,:);...
-rm*Y(4,:)-km*(Y(3,:)-y2eq)-F_kc-Pm;...
(Pm-rho/2*(Y(5,:)/Ac)^2)/L];
```

The transcript file for the analysis with flow model 2 (unsteady incompressible):

```
clear all
rho=1.14e-3;% density of the air in human vocal tract
global Ac_start
Ac_start=0.00004;
V=50; %cm^2/s
D=2;

[T,Y]=ode15s('final_b',[0,100e-3],[0,0,0,0,0]);

Ac=(V*T+Ac_start-Y(:,1)*D);

figure(2)
plot(T,Ac,'k')
xlabel('Time s')
ylabel('Cross-sectional Area cm^2')


function [Ydot]=final_b(t,Y)
V=50; % the release velocity


rho=1.14e-3;
Lc=0.5;% length of the constriction
Ag=0.2; % Length scale is in cm
D=2;
Ps=7840; %cmH2O to dyne/cm^
Pc_min=3*Ps;
m=2.1;
r=800;
k=84500;
kc=1.5*k;
y_c=Pc_min/k;
y0=Ps/k;
y1eq=y_c;
y2eq=y0;
global Ac_start
Ac=(V*t+Ac_start-Y(1,:)*D);
```

```
Add_y1=rho*D*Lc^2/Ac/6;
M_y1=m+Add_y1;
F_kc=-kc*((Y(1,:)-y1eq)-(Y(3,:)-y2eq));
dAc_dt=V-Y(2,:)*D;
ue1=Lc/2/Ac*dAc_dt;
ue2=Y(5,:);
Pm=Ps-rho/2*((ue1+ue2)*Ac/Ag)^2;
F_air_right=(-rho/3*Lc^2/Ac^2*dAc_dt*abs(dAc_dt)-Pm/2);
d2y_dt2=(-r*Y(2,:)-k*(Y(1,:)-y1eq)+F_kc+F_air_right)/M_y1;
d2Ac_dt2=-D*d2y_dt2;
due2_dt=Pm/rho/Lc; %ue

Ydot=[Y(2,:);...
d2y_dt2;...
Y(4,:);...
(-r*Y(4,:)-k*(Y(3,:)-y2eq)-F_kc-Pm)/m;...
due2_dt];
```

The transcript file for the analysis with flow model 3 (unsteady compressible):

```
clear all
rho=1.14e-3;
global V
V=50; %cm^2/s
D=2;
global Ac_start
Ac_start=0.0001;

[T,Y]=ode15s('compressible_b',[0,100e-3],[0,0,0,0,0,-rho*V/Ac_start,0]);
Ac=(V*T-Y(:,1)*D+Ac_start);

figure(3)
plot(T,Ac,'k')
xlabel('Time s')
ylabel('Cross-sectional Area cm^2')


function [Ydot]=compressible_b(t,Y)
y1=Y(1,:); dy1_dt=Y(2,:);
y2=Y(3,:); dy2_dt=Y(4,:);
a1=Y(5,:); da1_dt=Y(6,:);
ue2=Y(7,:);
global V
rho_e=1.14e-3;
P_e=1e6;
Lc=1;% length of the constriction
p1_air=P_e/rho_e*a1*Lc/4;
```

```
F_air_right=-p1_air;
Ag=0.2; %cm^2
D=2;
Ps=7840; % dyne/cm^2
Pc_min=5*Ps;
m=2.1;
r=800;
k=84500;
kc=1.5*k;
y_c=Pc_min/k;
y0=Ps/k;
y1eq=y_c;
y2eq=y0;
global Ac_start
Ac=(V*t+Ac_start-Y(1,:)*D);
F_kc=-kc*((Y(1,:)-y1eq)-(Y(3,:)-y2eq));
dAc_dt=V-Y(2,:)*D;
d2y_dt2=(-r*Y(2,:)-k*(Y(1,:)-y1eq)+F_kc+F_air_right)/m;
d2Ac_dt2=-D*d2y_dt2;
ue1=1/(rho_e*Ac)*(dAc_dt*(rho_e*Lc/2+a1*Lc^2/8)+Ac*Y(6,:)*Lc^2/8);
d2a1_dt2=(-P_e/rho_e*Y(5,:)*Ac-d2Ac_dt2*(rho_e*Lc/2+Y(5,:)*Lc^2/8)...
-(abs(ue1)*rho_e*dAc_dt+rho_e*ue1*Ac*abs(ue1*a1/rho_e+dAc_dt/Ac)))/(Ac*Lc^2/8);
Pm=Ps-rho_e/2*((ue1+ue2)*Ac/Ag)^2; %ue
d2y_dt2=(-r*Y(2,:)-k*(Y(1,:)-y1eq)+F_kc+F_air_right-Pm/2)/m;
due2_dt=Pm/rho_e/Lc;%ue

Ydot=[Y(2,:);...
d2y_dt2;...
Y(4,:);...
(-r*Y(4,:)-k*(Y(3,:)-y2eq)-F_kc-Pm)/m;...
Y(6,:);...
d2a1_dt2;...
due2_dt];
```