



Numerical study on the phonation mechanisms of sibilant [s]

呂，學叡

(Degree)

博士（工学）

(Date of Degree)

2022-09-25

(Date of Publication)

2023-09-01

(Resource Type)

doctoral thesis

(Report Number)

甲第8468号

(URL)

<https://hdl.handle.net/20.500.14094/0100477894>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



Doctor Thesis

Numerical study on the phonation mechanisms
of sibilant [s]

歯茎摩擦音[s]の発声メカニズムに関する数値的研究

July 2022

Graduate school of system Informatics

Kobe University

呂學叡 LU HSUEH JUI

Contents	I
Table captions	III
Figure captions	IV
Abstract	VIII
Acknowledgement	X
Abbreviation	XII
Chapter 1. General Introduction	1
Chapter 2. Numerical method	11
2.1 Introduction	12
2.2 Governing equation	13
2.3 Numerical scheme	14
2.4 Adaptively switched time stepping scheme	19
2.4.1 Methodology	19
2.4.2 Results and discussion	22
2.4.3 Conclusion	26

Chapter 3. Investigation of effects of incisor angle on production of sibilant /s/	43
3.1 Introduction	44
3.2 Vocal tract geometry	45
3.3 Computational conditions	47
3.4 Results and discussion	48
3.5 Conclusion	53
 Chapter 4. Investigation of effects of tongue articulation and velopharyngeal closure	 71
4.1 Introduction	72
4.2 Vocal tract geometry	72
4.3 Computational conditions	74
4.4 Results and discussion	75
4.5 Conclusion	80
 Chapter 5. Summary and future work	 105
 Reference	 109

Table captions

Table 2-1. Computational parameters of flow-induced noise around a cylinder.

Table 2-2. Comparison of results with CFL=20, 40, 80, and 160.

Table 2-3. Computational parameters of flow around a 3D flat plate.

Table 3-1. Overjet (horizontal overlap) and overbite (vertical overlap) for the 0° to 30° models.

Table 3-2. test of grid resolution.

Table 3-3. Computational parameters of the 0° model.

Table 4-1. Computational parameters of the moving process case.

Table 5-1. The parameters and conditions of all cases.

Figure captions

Figure 2-1. Compressible immersed boundary method configuration.

Figure 2-2. Grid distribution.

Figure 2-3. The Contours of instantaneous velocity field and Karman vortex shedding.

Figure 2-4. The Contours of instantaneous vorticity field and Karman vortex shedding.

Figure 2-5. Comparison of the contours of instantaneous pressure fluctuations.

Figure 2-6. Comparison of instantaneous pressure fluctuations with results in [55].

Figure 2-7. Drag coefficients C_D and lift coefficients C_L with respect to dimensionless time.

Figure 2-8. Comparison of time required to reach the quasi-steady state between constant and variable steps.

Figure 2-9. The cube distribution. (Cube division: $16 \times 16 \times 16$)

Figure 2-10. The computational condition.

Figure 2-11. Contours of fluctuation pressure for phase- and spanwise-averaged flow fields.

Figure 2-12. The mean velocity field around the plate.

Figure 2-13. The mean velocity profile from surface of the plate at the downstream.

Figure 3-1. (a) CT scan image of /s/. (b) Vocal tract geometry for the simulation.

Figure 3-2. (a) Vocal tract geometry for pronouncing /s/. (b) Top and (c) side views of the vocal tract.

Figure 3-3. Vocal tract geometry with the incisor angle increased from the original position (0°) to 30° : (a) original geometry with 0° ; (b) modified geometry with 30° ; and (c) modified geometry with the incisor angle ranging from 0° to 30° .

Figure 3-4. (a) Cube and (b) cell distributions.

Figure 3-5. The three-view diagram of overall computational domain of the current model.

Figure 3-6. Sound pressure level (SPL) spectrum of the base simulation (0° , solid black line) and measurements of the printed replica (dashed red line) and the average value of an actual subject for 18 recordings (blue circles). The error bars denote the range for 18 repeated recordings.

Figure 3-7. The experimental setting.

Figure 3-8. (a) Normalized instantaneous velocity magnitude and (b) root mean square (RMS) of the velocity fluctuations for the 0° model.

Figure 3-9. The predicted SPL spectra with incisor angles from 0° to 30° .

Figure 3-10. (a, c, e) Normalized instantaneous velocity magnitude and (b, d, f) root mean square (RMS) of the velocity fluctuations at $t = 0.015$ s for the (a, b) 10° , (c, d) 20° , and (e, f) 30° models.

Figure 3-11. Fast Fourier transforms of the velocity fluctuations at (a, c) 5 kHz and (b, d) 10 kHz for the (a, b) 0° and (c, d) 30° models.

Figure 3-12. Fast Fourier transforms of the velocity fluctuations at (a, c) 10 kHz and (b, d) band-pass filter field of 8-12 kHz for the (a, b) 0° and (c, d) 30° models.

Figure 3-13. Acoustic fields simulated with the monopole acoustic source at (a, c) Point 1 (49.1, -8.4, -7.7) and (b, d) Point 2 (58.5, -7.7, -4.5) for 0° and 30° models, respectively.

Figure 3-14. SPL spectra for the 0° and 30° models predicted for the monopole acoustic sources at Point 1 (0.049, -0.0084, -0.0077) and Point 2 (0.0585, -0.0077, -0.0045).

Figure 4-1. MRI scan images of the phonation process of Japanese word /usui/.

Figure 4-2. MRI images and vocal tract geometry of the (a) vowel /u/ to (c) the sibilant sound /s/ in the pronouncing process of /usui/.

Figure 4-3. The cross-section during the phonation process of (a) the tongue constriction and (b) the velopharyngeal valve.

Figure 4-4. The motion and the cross-sectional area of the velopharyngeal valve at (a) $t_1 = 0.0$ s, (b) $t_2 = 0.05$ s and (c) $t_3 = 0.10$ s.

Figure 4-5. The cross-sectional area of the tongue constriction and the velopharyngeal valve during the movement.

Figure 4-6. The modified geometries for velum speed investigating at $t_2 = 0.05$ s.

Figure 4-7. The difference of the cross-sectional area ratio of the different velum closing speed.

Figure 4-8. The movement of the velopharyngeal valve with different closing speed.

Figure 4-9. The grid distribution (every 16th gridline is shown for clarity).

Figure 4-10. The three-view diagram of overall computational domain.

Figure 4-11. (a, b, c) Instantaneous flow field and (d, e, f) pressure fluctuation along the sagittal plane at (a, d) $t = 0.01$ s, (b, e) $t = 0.05$ s and (c, f) $t = 0.10$ s.

Figure 4-12. The spectrogram of the propagating sound collected 80 mm from the lip outlet.

Figure 4-13. The spectrogram at the far-field of the (a) faster closing case (0.220 m/s), (b) base simulation (0.208 m/s) and (c) slower closing case (0.152 m/s).

Figure 4-14. The OASPL value of the different velum closing speed cases.

Figure 4-15. The OASPL of the different velum closing speed based on the tongue constriction.

Figure 4-16. The maximum velocity in the velopharyngeal channel for different closing speed of velopharyngeal valve.

Figure 4-17. The area ratio of the velopharyngeal channel to the oral channel.

Figure 4-18. The flow rate ratio at the velopharyngeal channel to the oral channel.

Figure 4-19. The flow rate at the velopharyngeal channel.

Figure 4-20. The flow rate at the nasal channel.

Figure 4-21. The far-field spectrogram of the (a) faster closing case (0.220 m/s), (b) base simulation (0.208 m/s) and (c) slower closing case (0.152 m/s) based on the velopharyngeal valve opening.

Figure 4-22. The OASPL value of the different closing speed cases based on the velopharyngeal valve opening ratio.

Figure 4-23. The relation of OASPLs and the velum closure speed.

Abstract

Speech disorders impact directly a person's social and occupational activity. The current study investigated the effect of the shape of the vocal tract on the sound generation mechanisms by the developed numerical framework. In the proposed Adaptively Switched Time Stepping scheme, the Newton linearization error can be eliminating, thus saving the computational resource at the same time keeping the accuracy for both fluid and acoustic calculations.

Firstly, to examine the sound changes by the inclination angles of the incisor, the simulations of the vocal tract geometry of sibilant /s/ sound at different incisor angles were conducted. Both of the airflow and the sound behavior of the /s/ sound were predicted to explore the cause of the sound changes. The results showed that increasing the incisor angle reduced the amplitude of the sound in the frequency range from 8 kHz to 12 kHz. Furthermore, the acoustic simulations conducted with different position of the monopole sound sources indicated that if the sound source position was located closer to the exit of the vocal tract, the source didn't couple strongly with the resonator, thereby the reduced the sound amplitude and generated a weaker sibilant /s/ sound.

In addition, to investigate the articulation process of velopharyngeal closure and tongue movement, the simulation of sibilant /s/ production with the realistically pronouncing process was conducted. The sound production process was extracted in the part of the Japanese word /usui/, including the tongue elevation and the velopharyngeal valve closure. The effect of tongue movement and velopharyngeal closure during the articulation process of the sibilant /s/ production was described by analyzing the spectrogram at the far-field. Moreover, the further simulations with the different closing speed of the velopharyngeal valve indicated that the size of the tongue

constriction mainly affects the sibilant /s/ sound even with the small velopharyngeal closure. Besides, the comparison of the far-field spectrogram based on the velopharyngeal valve opening indicated that the slower closing speed of the velopharynx will enlarge the sound amplitude during the co-articulation process, and this undesired sound will let the listener difficult to distinguish the word.

The current study using the implicit numerical simulation with the developed ASTS schemes investigated the effect of the incisor angle and the effect of the velopharyngeal closure on sibilant /s/ sound. These results contributed to the dental treatment for generating the sibilant fricatives. By further investigation, the present results and the current framework will enable to provide the underlying insights of the sibilant sound production prior to prosthetic surgery.

Acknowledgement

The present doctoral dissertation is the research work carried out at Kobe University, Kobe, Japan, during April' 2018 to September' 2022 under the supervision and guidance of Professor. Makoto Tsubokura.

I would like to extend my deep gratitude to all those who have offered me a lot of help and support in the process of my thesis writing during the four years at Kobe University.

First of all, my sincere thanks go to Professor Makoto Tsubokura and Professor ChungGang Li, my supervisors, who have offered me numerous valuable comments and suggestions. Without their painstaking teaching and insightful advice, the completion of this thesis would have been impossible.

Also, I would like to express my sincere gratitude to all the professors that greatly broadened my horizon and enriched my knowledge in the acoustic field. Specifically, I thank Professor Tsukasa Yoshinaga, Toyohashi University of Technology, and Professor Kazunori Nozaki, Osaka University Dental Hospital for kindly providing the vocal tract geometry with the CT scan images and the experimental results.

Besides, I thank Professor Akiyoshi Iida, the Toyohashi University of Technology for helpful discussion, and advice for investigation.

In addition, I thank Doctor Rahul Bale, Complex Phenomena Unified Simulation Research Team, RIKEN, for the technical support. Moreover, I appreciate the RIKEN Junior Research Associate Program supported my research financially.

Finally, I would like to express the most heartfelt gratitude to my family members and friends who have provided me with great help and support.

呂學叡 LU HSUEH JUI

Prof. Dr. Makoto Tsubokura Laboratory

Graduate School of System Informatics

Kobe University, Japan

July' 2022

Abbreviation

The following abbreviations are used in this dissertation.

ASTS	Adaptively Switched Time Stepping
CAA	Computational Aeroacoustics
CFL	Courant–Friedrichs–Lewy
CT	Computational Tomography
DAC	Direct Aeroacoustic Computations
FFT	Fast Fourier Transform
IBM	Immersed Boundary Method
IC	Interface Cell
IP	Image Point
LMRoe	Low-Mach-fix for Roe method
LUSGS	Lower–Upper Symmetric Gauss–Seidel
MRI	Magnetic Resonance Imaging
MUSCL	Monotone Upstream-centered Scheme for Conservation Laws
OASPL	Overall Sound Pressure Level
RMS	Root Mean Square

SLTS	Solution-Limited Time Stepping
SPL	Sound Pressure Level
C_D	Drag coefficient
C_L	Lift coefficient
N_{cube}	Total number of cubes
N_{cell}	Total number of cells
Pr	Prandtl number
Re	Reynolds number
St	Strouhal number
SL_{cell}	Size of the largest cell
SS_{cell}	Size of the smallest cell
x_i	Cartesian coordinates in i^{th} direction
u_i	Velocity components in i^{th} direction
F_i	Flux vectors in i^{th} direction
δ_{ij}	Kronecker delta
μA_{ij}	Stress term

e	Total energy
Φ_{IP}	Value of image point
Φ_{IC}	Value of interface cell
Φ_i	Value of surrounding points
Φ_{Wall}	Wall value
ω_i	Value calculated using the VanderMonde matrix
ε	A tiny value to prevent numerical divergence when IC is very close to the wall
D	Diameter (mm)
L	Length of the plate (mm)
s	Separation (mm)
b	Thick of the plate (mm)
t	Physical time (s)
τ	Artificial time (s)
U	Conservative vector (m/s)
c	Speed of sound (m/s)
P	Pressure (Pa)
$\bar{P}$	Time-averaged pressure (Pa)

P_{ref}	Reference pressure (Pa)
ρ	Density (kg/m ³)
ρ_0	Ambient fluid density (kg/m ³)
T	Temperature (K)
T_0	Ambient fluid temperature (K)
μ	Dynamic viscosity (N·s/m ²)
μ_0	Ambient fluid dynamic viscosity (N·s/m ²)
R	Gas constant (J/kg/K)
k	Thermal conductivity (W/mK)
γ	Heat capacity ratio

Chapter 1. General Introduction

Speech disorders impact directly a person's social and occupational activity. In Japan, the number of children who have articulation disorder is 556,000; the number of adults who have articulation disorder is 293,000; the number of people who have stuttering is 710,000 [1]. The articulation disorder can be caused by a variety of reasons, such as abnormal anatomy, abnormal neurophysiology, or mislearning. The current study put the emphasis on the morphology of the vocal tract and we investigate the effects of the incisor angle as well as the effect of the velopharyngeal closure on the sound generation mechanisms by the numerical simulation.

To investigate the effect of the vocal tract geometry, the sibilant /s/ sound was chosen because this sound is one of the most difficult consonants to pronounce. The human is the only mammal who can pronounce /s/ sound well, even though the gorilla cannot. Therefore, it is often reported in speech disorders. Besides, it is used in daily conversation frequently and can be found in many languages. In addition, it is an unvoicing sound which means it can be pronounced without considering the vibration of the vocal cord so that the current study can focus on the effect of the shape of the vocal tract by applying /s/ sound.

The sibilant fricative /s/ is generated by the airflow passing through a constriction formed between the alveolar ridge and the tongue without the voicing. When airflow is forced to go through the narrow constriction (sibilant groove), the turbulent jet flow is generated, and the velocity fluctuation downstream will be considered as the source of the sound [2] [3]. The fricative sounds are characterized as broadband noise in the frequency range above fricative's characteristic peaks, about 4–7 kHz for /s/ and 2–3 kHz for /ʃ/ [4].

It is known that the central incisor position and angle affect speech production, especially the fricative consonant /s/ [5–9]. Because the /s/ sound is generated by the turbulent jet flow leaving the constriction [10], inaccurate formation of the constriction results in speech production

difficulties for /s/. Runte et al. [5] constructed a maxillary denture, and the inclination angle of the central incisor was changed within the range from -30° to $+30^\circ$ to investigate the influence of the angle on the production of /s/. Meanwhile, the jaw movement during the phonation of /s/ was measured to observe the closest oral tract space with the horizontal and vertical overlap of the incisors [11]. Hamlet et al. [12] measured the tongue movement with a contact sensor and showed the effects of dental prostheses on the timing and duration of the constriction formation of sibilant fricatives. However, because the /s/ sound is generated by jet flow in the oral tract, the detailed mechanisms of the effects of dental prostheses on the observed sound changes are unclear.

The production mechanisms in the absence of articulatory dysfunction had been investigated by modeling the vocal tract geometry. Shadle [13] proposed a simplified vocal tract model for fricative consonants and investigated the effects of teeth-like obstacle positions on the generated sounds. In addition, numerical flow simulations were applied to a simplified model and the effects of geometrical differences on the turbulent jet flow and its sound sources were examined [14] [15].

To further investigate the airflow associated with the production of /s/ and its source characteristics in an actual oral tract, a vocal tract replica was constructed from computational tomography (CT) images, and the flow and sound generated in the replica were measured using a microphone and an anemometer [16]. A numerical flow simulation of a realistic vocal tract geometry revealed that the aeroacoustic sound source is located near the upper and lower incisor surfaces [17].

Recently, both turbulent flow for /s/ and the sound generated by the vocal tract geometry were predicted via numerical simulations using high-performance computing resources [18–21]. These simulations revealed that the sound source is located downstream of the upper and lower incisors and that the acoustic characteristics of /s/ are formed by the geometry downstream of the

constriction. However, the effects of geometrical differences resulting from the dental prosthesis, e.g., the incisor positions and angles, on the flow and sound generation are still unclear.

Therefore, in chapter 3, we conducted numerical flow simulations of the vocal tract geometry of /s/ at different incisor angles to examine the origin of sound changes using the inclination angles of the incisor reported by Runte et al. [5]. Both the airflow for /s/ and the sound from the vocal tract geometry were predicted using numerical simulations to explore the cause of the sound changes.

In addition to the incisor angle, the closure of the velopharyngeal valve also affects sound generation. Children who have cleft lip, cleft palate and will have insufficient velopharyngeal closure is 0.18%. Kummer [22] mentioned that the velopharyngeal dysfunction is a speech disorder where the velopharyngeal valve does not close consistently and completely during the production of oral sounds. The causes for the velopharyngeal dysfunction might be abnormal anatomy (velopharyngeal insufficiency), abnormal neurophysiology (velopharyngeal incompetence), or particular articulation errors (velopharyngeal mislearning).

The velopharyngeal valve, which consists of the velum (soft palate), the lateral and posterior pharyngeal walls, controls the amount of airflow going to the nasal cavity. During the normal nasal breathing or producing the nasal sounds, the soft palate rests, and the velopharyngeal valve opens the flow channel and the airflow goes through the nasal cavity, which results in an acoustic coupling between the oral and nasal cavities. On the contrary, during the production of the oral consonants, such as /s/ or /p/, the complete closure of the velopharyngeal valve is required so that the soft palate moves superior and posterior to close the velopharyngeal channel against the airflow going to the nasal cavity.

If the flow escapes through the nasal cavity during the sibilant sound production, the nasal emission occurs. In this situation, the speech distortion will be created by an acoustic mechanism where undesirable sound propagates in the nasal cavity. Nasal emission can occur on all pressure-sensitive phonemes or it can be phoneme-specific. Oren et al. [23] provided a brief review of the types of nasal emission, the terms used to describe it, and the potential causes and indicated that when there are characteristics of velopharyngeal insufficiency, the audibility of nasal emission can give the examiner clues as to the relative size of the opening. Kummer et al. [24] investigated the comparison of velopharyngeal gap size in patients with hypernasality, hypernasality, and nasal emission, or nasal turbulence (rustle) and indicated that the degree of the velopharyngeal insufficiency can be predicted to some extent based on perceptual assessment. If the presence of a nasal rustle suggested a small velopharyngeal opening, despite the severe distortion of speech. Kummer et al. [25] also predicted the velopharyngeal gap size based on an individual's speech. Although it was not an exact science, the results revealed that some predictions of gap size can be made on the basis of the speech assessment alone. If the patient has nasal rustle, which predicts a small gap; or moderate to severe hypernasality, which is associated with a larger opening.

The effects of the velopharyngeal opening and the mechanism of the nasal emission have been investigated by the measurements [26–28] and numerical simulations [29–32]. Beer et al. [26] compared the dynamic near-real-time magnetic resonance imaging with the videofluoroscopy technique and confirmed that real-time MRI correctly depicted the pattern of velopharyngeal closure in correspondence to videofluoroscopy in all cases, and may deliver additional information in selected cases. Bunton et al. [27] examine the relation of perceptual ratings of nasality by experienced listeners, measures of nasalance, and the size of the nasal port opening for three simulated English corner vowels, /i/, /u/, and /ɑ/, and noted the differences in nasalance and

perceptual ratings of nasality among the three vowels. Besides, they indicated that the listeners detected nasality for the high and low vowels simulated with nasal port areas of 0.01 and 0.15 cm^2 , respectively. Furthermore, Bunton [28] extended previously published modeling work examining the relationship between nasal port opening, measures of nasalance, and perceptual ratings of nasality. The findings indicate that perceptual ratings of nasality and nasalance increased for samples generated with nasal port areas up to and including 0.16 cm^2 but plateaued in samples generated with larger nasal port areas.

Sundström conducted a series of studies by using the large eddy simulation to investigate the effect of the velopharyngeal insufficiency. To determine the flow behavior and the sound generating mechanism in the vocal tract, the small and large velopharyngeal opening cases based on the specific patient were conducted [29]. They confirmed that the turbulence in the nasal cavity indeed affected the far-field acoustic behavior, and found that the generated sound was the turbulence interacting with the solid nasal wall. To examine the aeroacoustic phenomenon on the sibilant /s/ sound caused by the existing of the velopharyngeal opening, they conducted four cases by two patient-specific models with and without a small size of the velopharyngeal valve [30]. The results showed that since the aeroacoustic mechanisms in the nasal cavity were eliminated by removing the small opening of the velopharynx, the far-field spectrum changed significantly. In addition, to clarify the relationship between the flow characteristics in the nasal cavity and the size of the velopharyngeal opening, they investigated the flow characteristics and sound generation mechanism by changing the size of the velopharyngeal opening based on a subject who was distorting nasal emission [31]. They found that the velocity and the turbulence were inversely proportional to the size of the opening if the flow resistance across the velopharyngeal valve was smaller than that across the oral constriction. On the other hand, if the flow resistance was higher

across the velopharyngeal valve than the oral constriction, the flow characteristic was reduced at a higher rate. They indicated that the area ratio of the velopharyngeal valve to the oral constriction can be considered as a factor to determine the airflow characteristics during the production of sibilant /s/ sound. They also mentioned that the most severe nasal emission was produced when the size of the velopharyngeal opening was equal to the size of the oral constriction. To examine changes in the sound source mechanisms as the different sizes of the velopharyngeal opening during the production of the sibilant /s/ sound, the patient-specific models with modified velopharyngeal opening were conducted [32]. The results showed that the acoustic energy in the nasal cavity and the far-field acoustic behavior was directly related to the size of the velopharyngeal opening.

However, almost all studies of the flow simulation used a fixed vocal tract geometry which only provides the stationary result instead of the time-varying sound production in the actual human speech process. Therefore, in chapter 4, the numerical simulation of sibilant /s/ production with the realistically moving vocal tract was conducted to investigate the articulation process of velopharyngeal closure and tongue movement. The sibilant /s/ sound production process was simulated in the part of the Japanese word /usui/ which is extracted after the vowel production of /u/ to the sibilant sound /s/, including the tongue elevation and the velopharyngeal valve closure. By analyzing the spectrogram of the generated sound, the current study clarifies the effect of tongue movement and velopharyngeal closure during the articulation process of the sibilant /s/ production.

Additionally, the moving speed and timing of the speech organ is also an important factor in the sound production during the phonation process. Moll et al. [33] investigated the timing of the velar movements during a speech by taking high-speed cinefluorographic films of four normal

subjects speaking English sentences containing various combinations of nasal consonants, consonants, and vowels. They found that the velar-opening gesture for the nasal is initiated near the beginning of a primary articulatory movement toward the first vowel in the sequence in which a nasal consonant is preceded by one or two vowel sounds. Dalston et al. [34] investigated the reaction times among normal adult speakers by nasometric and phototransductive measurements. In the measurement, four female and four male subjects sustained production of the nasal consonant /m/ until a computer-driven imperative tone was presented at which time they were required to say /pi/ as quickly as possible. The results showed that there were 146 ms for initiation of velopharyngeal movement toward closure, 224 ms for voicing offset, 280 ms for onset of velopharyngeal closure, and 317 ms for the onset of plosive release. Warren et al. [35] measured the timing characteristics associated with velopharyngeal closure in 11 subjects by the pressure-flow technique. They indicated that the hypernasality seems to be associated with the actual time the velopharyngeal mechanism is open, rather than the volume of air escaping from the nasal chamber.

Hence, to investigate the effect of the closing speed of the velopharyngeal valve, in chapter 4, one faster and one slower closing cases were conducted by modifying the size of the velopharyngeal valve during the phonation process. By comparing the spectrum based on the velopharyngeal opening, the amplitude difference at the co-articulation process was clarified.

In this study, the numerical flow simulations of the vocal tract geometry were conducted to clarify the effect of the geometry on the sibilant /s/ sound. To explore the cause of the sound changes, both the airflow and the sound in the vocal tract geometry for /s/ sound were predicted using numerical simulations solving the three-dimensional compressible Navier–Stokes equations [36–38]. However, while applying computational aeroacoustics (CAA) remains challenging,

because the different order of scales in the acoustic and flow fields coexist in the same computational domain. As a result, highly accurate methods require very short time steps to capture the physical variables in the acoustic field. Therefore, the explicit method was adopted, generally, because of its low levels of numerical dispersion and dissipation. One of the most popular explicit techniques is the Runge–Kutta method. To achieve higher accuracy and more efficient computations, several optimized Runge–Kutta methods have been proposed. For example, a modified Runge–Kutta coefficient can achieve high-accuracy CAA by minimizing the dispersion and dissipation in the Fourier space over a large range of frequencies for linear operators [39–41]. However, even with optimized Runge–Kutta methods, the Courant–Friedrichs–Lewy (CFL) condition restricts the time step, potentially resulting in very long computation times. Hence, it might not be appropriate to use explicit methods for practical acoustics applications.

Relatively few DACs use implicit methods, because the larger numerical dissipation and dispersion result in lower accuracy. To adopt a larger time step while retaining sufficient accuracy in an implicit scheme, Lian et al. [42] developed the solution-limited time stepping (SLTS) approach to accelerate the speed of convergence to the steady state. This approach uses the value of the residual as the criterion for the stability of the numerical scheme, and thus for determining whether a larger CFL number should be adopted to accelerate the convergence speed. Their scheme was extended to transient state problems by introducing a dual time stepping technique [43]. However, this excellent idea cannot be applied to DAC problems because of the Newton linearization error [44] introduced by the artificial time term in the dual time stepping approach.

Hence, a framework with an implicit time scheme based on the solution-limited time stepping [42] and the lower–upper symmetric Gauss–Seidel (LUSGS) schemes has been developed for direct aeroacoustic computations problems, as described in chapter 2. In the proposed Adaptively

Switched Time Stepping (ASTS) scheme, the Newton linearization error can be eliminated and the acceptable accuracy can be achieved in both aerodynamic and aeroacoustic problems.

Besides, the immersed boundary method was adopted for the complex replica geometry of the realistic human vocal tract and the movement of the sound generating process. To reduce the grid generation time for the vocal tract geometry and to provide better load balancing and higher performance on parallelization, we applied the hierarchical structure grid [45] in the simulations. By further developing the proposed methodology, this simulation technology will enable us to predict the effects of speech organs on the production of sibilant fricatives for patients prior to prosthetic surgery.

Chapter 2. Numerical method

2.1 Introduction	12
2.2 Governing equation	13
2.3 Numerical scheme	14
2.4 Adaptively switched time stepping scheme	19
2.4.1 Methodology	19
2.4.2 Results and discussion	22
2.4.3 Conclusion	26

2.1 Introduction

The physical variables involved in a flow field are of very different scales to those of the aeroacoustic field within the flow. Thus, direct aeroacoustic computations (DACs) require extremely small time steps to capture the fluctuations in the aeroacoustic field and stabilize the numerical scheme. As a result, the computation time for DACs can be very long. This paper describes an implicit Adaptively Switched Time Stepping (ASTS) scheme that enables the use of larger time steps when solving DAC problems.

ASTS is based on the adaptive selection of the time stepping scheme according to the value of the residual. Early in the iteration process, an artificial time term is used to afford larger time steps while ensuring that the numerical scheme remains stable. Once the residual is less than some preset reference value, the artificial time term is removed, thus eliminating the Newton linearization error. Simulation results show that ASTS can reduce the computational cost of DACs by a factor of more than eight while maintaining the required level of accuracy. To be more specifically, ASTS scheme allows CFL numbers up to 40 to be adopted in DAC problems.

Additionally, to accelerate the time required to reach the quasi-steady state, an extremely large CFL number of up to 160 can be used to force the rapid development of the flow in the initial stages. After reaching the quasi-steady state, the time step gradually decreases to the target time step, which is better suited to DAC problems. This use of extremely large CFL numbers to quickly reach the quasi-steady state and larger CFL numbers for DAC problems means that the ASTS scheme is a promising tool for CAA studies.

2.2 Governing equation

The governing equations are the compressible Navier–Stokes equations:

$$\frac{\partial U}{\partial t} + \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} = 0, \quad (1)$$

where t is the time, x_i indicates three directions in Cartesian coordinate system ($i = 1, 2, 3$), and the conservative vector U is

$$U = (\rho \ \rho u_1 \ \rho u_2 \ \rho u_3 \ \rho e)^T, \quad (2)$$

where the flux vectors F_i are

$$F_i = \begin{pmatrix} \rho u_i \\ \rho u_i u_1 + P \delta_{i1} - \mu A_{i1} \\ \rho u_i u_2 + P \delta_{i2} - \mu A_{i2} \\ \rho u_i u_3 + P \delta_{i3} - \mu A_{i3} \\ (\rho e + P) u_i - \mu A_{ij} u_j - k \frac{\partial T}{\partial x_i} \end{pmatrix}, \quad \forall i = 1, 2, 3, \quad (3)$$

ρ is the density, u_i indicates the velocity components ($i = 1, 2, 3$), and δ_{ij} is the Kronecker delta.

The total energy e is calculated as

$$e = \frac{P}{\rho(\gamma - 1)} + \frac{1}{2} (u_1^2 + u_2^2 + u_3^2), \quad (4)$$

and μA_{ij} is the stress term, with

$$A_{ij} = \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} (\nabla \cdot u) \delta_{ij}. \quad (5)$$

The pressure P follows the ideal gas equation:

$$P = \rho RT . \quad (6)$$

The dynamic viscosity μ and thermal conductivity k are based on Sutherland's law:

$$\mu(T) = \mu_0 \left(\frac{T}{T_0} \right)^{\frac{3}{2}} \frac{T_0 + 110}{T + 110} , \quad (7)$$

$$k(T) = \frac{\mu(T) \gamma R}{(\gamma - 1) \text{Pr}} , \quad (8)$$

where $\rho_0 = 1.1842 \text{ kg/m}^3$, $\mu_0 = 1.85 \cdot 10^{-5} \text{ N}\cdot\text{s/m}^2$, $T_0 = 298.06 \text{ K}$, $\gamma = 1.4$, $R = 287 \text{ J/kg/K}$, and the Prandtl number (Pr) is 0.71.

2.3 Numerical scheme

The dual time stepping method is applied to solve the Navier–Stokes formulation in Eq. (1).

Take the 2-D governing equation as an example

$$\frac{\partial U_p}{\partial \tau} + \frac{\partial U}{\partial t} + \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} = 0 , \quad (9)$$

where U_p is the primitive form of $[P, u_1, u_2, T]$. τ and t are, respectively, the artificial and physical time, and U is the conservative form of $[\rho, \rho u_1, \rho u_2, \rho e]$.

The discretized form of Eq. (7) is

$$\begin{aligned} & \frac{\bar{U}_p^{k+1} - \bar{U}_p^k}{\Delta \tau} + \frac{3\bar{U}^{k+1} - 4\bar{U}^n + \bar{U}^{n-1}}{2\Delta t} \\ & + \frac{1}{\Delta x_1} (\bar{F}_{1(i+1/2,j)}^{k+1} - \bar{F}_{1(i-1/2,j)}^{k+1}) + \frac{1}{\Delta x_2} (\bar{F}_{2(i,j+1/2)}^{k+1} - \bar{F}_{2(i,j-1/2)}^{k+1}) = 0 \end{aligned} \quad (10)$$

The superscripts k and n indicate the iteration numbers of artificial time and the preceding step of real time, respectively. When the artificial time term $\partial U_p / \partial \tau$ converges to zero, the quantities of the $(k+1)th$ iteration in the artificial time term are approximately transferred to the quantities of the $(n+1)th$ time step in real time, and Eq. (10) becomes the original Navier–Stokes equation including the transient term.

Eq. (10) can be rearranged as

$$[M^{-1} \frac{I}{\Delta \tau} + \frac{3}{2\Delta t} + M^{-1}(\delta_{x_1} A_p^k + \delta_{x_2} B_p^k)] \Delta U_p = M^{-1} R^k. \quad (11)$$

where $M = \partial U / \partial U_p$, δ_{x_i} is the central-difference operator, $A_p = \partial F_1^k / \partial U_p$ is the flux Jacobian, and $R^k = -(3U^k - 4U^n + U^{n-1}) / (2\Delta t) - (\delta_{x_1} \bar{F}_1^k + \delta_{x_2} \bar{F}_2^k)$.

In the computation of R^n on the right-hand side of Eq. (11), the terms involving F_i in Eq. (3) can be divided into two parts. One is the inviscid term $F_{inviscid}$, expressed as

$$F_{inviscid} = \begin{pmatrix} \rho u_i \\ \rho u_i u_1 + P \delta_{i1} \\ \rho u_i u_2 + P \delta_{i2} \\ (\rho e + P) u_i \end{pmatrix}, \quad (12)$$

and the other is the viscous term $F_{viscous}$, expressed as

$$F_{viscous} = - \begin{pmatrix} 0 \\ \mu A_{i1} \\ \mu A_{i2} \\ \mu A_{ij} u_j + \lambda \frac{\partial T}{\partial x_i} \end{pmatrix}. \quad (13)$$

To obtain more accurate results at low Mach numbers, Eq. (12) is solved using the low-Mach-fix for Roe (LMRoe) method [49], whereby

$$F_{inviscid,i+1/2} = \frac{1}{2}(F_R + F_L) + F_d, \quad (14)$$

where F_d is the Roe upwind dissipation term given by

$$F_d = -\frac{1}{2} \left\{ |U| \begin{bmatrix} \Delta \rho \\ \Delta(\rho u) \\ \Delta(\rho v) \\ \Delta(\rho E) \end{bmatrix} + \delta U \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho H \end{bmatrix} + \delta p \begin{bmatrix} 0 \\ 1 \\ 0 \\ u \end{bmatrix} \right\} \quad (15)$$

and

$$\delta U = (c - |u|) \frac{\Delta p}{\rho c^2} + \frac{u}{c} f(M) \Delta u, \quad (16)$$

where $f(M) = (|u| + |v| + |w|) / c$, and

$$\delta p = \frac{u}{c} \Delta p + (c' - |u|) \rho \Delta u, \quad c' = f(M)c. \quad (17)$$

To calculate F_R and F_L in Eq. (14), the fifth-order monotone upstream-centered scheme for conservation laws (MUSCL) [50] with no limiter function is adopted. The fifth-order MUSCL without a limiter function can be written as

$$U_{i+1/2}^L = 1/60 \times (2U_{i-2} - 13U_{i-1} + 47U_i + 27U_{i+1} - 3U_{i+2}), \quad (18)$$

$$U_{i-1/2}^R = 1/60 \times (-3U_{i-2} + 27U_{i-1} + 47U_i - 13U_{i+1} + 2U_{i+2}). \quad (19)$$

In addition to the inviscid term, the derivatives of A_{ij} in the viscous term of Eq. (13) are computed using the second-order central difference.

To simulate the complex geometry, e.g., a realistic human oral cavity, the immersed boundary method with the hierarchical structure grid [45] was applied for the grid spacing. Using a hierarchical structure grid can shorten the working time required to build computational grids and simultaneously provide better load balancing and higher performance for parallel computations. As boundary conditions on the wall, a no-slip condition is imposed on the velocity and a zero-gradient is applied in the normal direction for the pressure and density using the immersed boundary method (IBM) for compressible flow [51]. The compressible immersed boundary method configuration is shown in Figure 2-1.

The IBM can be summarized in the following steps:

- (i) The interface cell (IC) is defined by calculating the distance between the wall and the center of the cell within a range of one cell.
- (ii) The image point (IP) is defined as the point at which the distance between the IP and IC is equal to the distance between the IC and the wall in the normal direction.
- (iii) The value of the IP can be interpolated using the surrounding points:

$$\Phi_{IP} = \omega_1 \Phi_1 + \omega_2 \Phi_2 + \omega_3 \Phi_3 + \omega_{IC} \Phi_{IC}, \quad (20)$$

where ω_1 , ω_2 , ω_3 and ω_{IC} are calculated using the VanderMonde matrix which corresponds to the bilinear interpolation scheme. Moreover, the value of IC can be interpolated using the IP and the wall value.

$$\Phi_{IC} = \frac{1}{2}(\Phi_{IP} + \Phi_{Wall}), \quad (21)$$

(iv) Thus, the IP and IC are coupled. By solving the coupled equation, Dirichlet or Neumann boundary conditions can be imposed.

$$\Phi_{IC} = (\omega_1\Phi_1 + \omega_2\Phi_2 + \omega_3\Phi_3 + \Phi_{assign})/(2 - \omega_{IC}), \text{ (Dirichlet condition)} \quad (22)$$

$$\Phi_{IC} = (\omega_1\Phi_1 + \omega_2\Phi_2 + \omega_3\Phi_3)/(1 - \omega_{IC} + \varepsilon), \text{ (Neumann condition)} \quad (23)$$

where ε is a tiny value to prevent numerical divergence when IC is very close to the wall (ω_{IC} is nearly 1).

In this study, Eq. (22) and Eq. (23) had been extended to three-dimensional scenario. Besides, since the no-slip and isothermal condition are adopted on the wall, Eq. (22) is used for velocity. On the other hand, Eq. (23) is applied for pressure.

The whole framework is based on a hierarchal structure system named CUBE. Using the hierarchical structure grid system can shorten the working time required to build computational grids and simultaneously provide better load balancing and higher performance for parallel computations. Absorbing and non-reflecting boundary conditions [52] are adopted at the outermost boundary to prevent reflections from affecting the results.

2.4 Adaptively switched time stepping scheme

2.4.1 Methodology

To set the artificial time step $\Delta\tau$ in Eq. (11) so as to accelerate the convergence speed, Lian et al. [42] proposed the brilliant solution-limited time stepping (SLTS) method. This was employed alongside the lower–upper symmetric Gauss–Seidel (LUSGS) implicit method to adaptively adjust the CFL number and determine $\Delta\tau$.

When adopting LUSGS to solve Eq. (11), the estimation value ΔQ_{est} , which is defined in Eq. (20), provides a criterion for determining whether the calculation is stable or not [42].

$$\Delta Q_{est} = -\Delta\tau M^{-1}[-(3U^k - 4U^n + U^{n-1}) / (2\Delta t) - (\delta_{x_1} \bar{F}_1^k + \delta_{x_2} \bar{F}_2^k)], \quad (24)$$

According to the numerical test in [42], if any component of ΔQ_{est} is less than the reference value ΔQ_{ref} defined in Eq. (25), the calculation is stable. Hence, the CFL number based on the artificial time step $\Delta\tau$ can be increased to accelerate the speed of convergence. In this situation, CFL is set to some value $CFL_{specified}$, which has a recommended value of 1000 [42].

$$\Delta Q_{ref} = \begin{pmatrix} \alpha_1 \times \max[0.5 \times \rho(u_1^2 + u_2^2), \Delta P_{sur}, P_{global} \times 10^{-9}] \\ \alpha_2 \times \max[(u_1^2 + u_2^2), \frac{\Delta P_{sur} \times c}{\gamma P}, V_{global} \times 10^{-9}] \\ \alpha_3 \times \max[(u_1^2 + u_2^2), \frac{\Delta P_{sur} \times c}{\gamma P}, V_{global} \times 10^{-9}] \\ \alpha_4 \times T \end{pmatrix}, \quad (25)$$

where ΔP_{sur} is the maximum difference between the pressure at surrounding points and P_{global} ,

V_{global} denotes some global value that ensures the reference values are always greater than 0, c is

the speed of sound, and γ is the heat capacity ratio. $[\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4]$ is the coefficient of the maximum allowable fractional change, which has a suggested value of $[0.1 \ 2 \ 2 \ 0.1]$ [42].

If any one component of ΔQ_{est} is greater than ΔQ_{ref} , the calculation is not stable, and the CFL number should be decreased to prevent the calculation from diverging. Eq. (26) is used to decrease the CFL number appropriately:

$$CFL_{allowable} = CFL_{specified} \times \alpha \times \Delta Q_{residual} / |\Delta Q_{est}|, \quad (26)$$

where a value of $\alpha = 0.2$ has been used in previous studies [42]. $CFL_{allowable}$ is then adopted as the artificial time step.

Under the SLTS method, the maximum allowable CFL number is automatically evaluated, and the maximum allowable artificial time step $\Delta \tau$ can be determined to accelerate the speed of convergence. Additionally, the value of the physical time step is not restricted by the CFL condition because of the dual time stepping. Overall, a larger physical time step and faster speed of convergence can be achieved, thus drastically reducing the computational resources required.

However, SLTS is not suitable for aeroacoustic simulations because the Newton linearization error of the term $\partial U_p / \partial \tau$ persists, even when the convergence criteria are satisfied. This kind of small error will destroy the accuracy of the acoustic field, because physical values in the acoustic field are usually 2–3 orders of magnitude smaller than those in the flow field.

A straightforward method of removing the Newton linearization error was suggested by Xu et al. [53]. They found that the term containing the artificial time term, $M^{-1}I / \Delta \tau$, can be set to zero in Eq. (11). This drives the Newton linearization error to zero, leaving

$$\left[\frac{3}{2\Delta t} + M^{-1}(\delta_{x_1} A_p^k + \delta_{x_2} B_p^k)\right]\Delta U_p = M^{-1}R^k. \quad (27)$$

In this situation, the artificial time step is effectively set to infinity in Eq. (11), significantly increasing the convergence speed. However, Eq. (27) cannot accommodate such a large physical time step because of the lack of the artificial time term. Therefore, to adopt a larger physical time step while still avoiding the Newton linearization error, we propose the Adaptively Switched Time Stepping (ASTS) scheme.

Inside the iteration process, when ΔQ_{est} is greater than ΔQ_{ref} , the numerical scheme is not yet stable, and so the whole framework is the same as SLTS, which means Eq. (27) is used to obtain the artificial time step for solving Eq. (11). Once ΔQ_{est} is less than ΔQ_{ref} , however, Eq. (28) can be applied to remove the Newton linearization error, because the numerical scheme is already stable.

The solution procedures of ASTS can be briefly summarized in the following two steps:

1. Calculate Eqs. (24) and (25) to obtain the values of ΔQ_{est} and ΔQ_{ref} .
2. If $\Delta Q_{est} > \Delta Q_{ref}$, use Eq. (26) to calculate $\Delta \tau$ and solve Eq. (9); if $\Delta Q_{est} < \Delta Q_{ref}$, solve Eq. (27) directly.

When performing the implicit time stepping scheme, the initial guess for the iteration is usually the result of the current time step, which means we set $U^{k+1,k=0} = U^n$ for R^k in Eq. (11). After N_k iterations, $U^{k+1,k=N_k}$ approaches U^{n+1} , i.e., the iteration process has converged, and we move to the next physical time step. For ASTS, Eq. (11) is first applied until $\Delta Q_{est} < \Delta Q_{ref}$, whereupon Eq. (27) is used. Hence, the initial guess for Eq. (27) is actually the SLTS result after several iterations,

which means the initial guess for Eq. (27) in ASTS is already an approximation of U^{n+1} . This is why Eq. (27) can use a larger time step while retaining sufficient accuracy.

2.4.2 Results and discussion

The proposed ASTS scheme was validated by simulating and capturing the tonal noise generated by a flow passing a two-dimensional circular cylinder at $Re = 200$ with a free stream velocity of Mach 0.2. The resulting aerodynamic properties, such as the drag coefficient, lift coefficient, and Strouhal number, were then compared with those reported by Linnick and Fasel [54], whereas the properties of the acoustic field, such as the pressure fluctuations, were compared with those presented by Seo and Mittal [55].

The computational parameters are listed in Table 2-1 and the grid distribution is shown in Figure 2-2. The finest mesh around the cylinder and its wake is around $0.015D$, where D is the diameter, while far from the cylinder, a much coarser resolution of around $0.48D$ is adopted. According to [55], the wavelength in the far field is around $25D$, where a sufficient resolution of around 50 points per wave can be ensured in the present simulations. Additionally, an absorbing, non-reflecting boundary is adopted at the outermost cubes.

Using the proposed ASTS scheme, simulations were conducted with four different CFL numbers (20, 40, 80, and 160, based on sound speed plus the inlet speed). The number of inner iterations was set to 100 to ensure a fair evaluation of the computation time. For explicit methods, the CFL number should generally be less than 1 to ensure that the numerical scheme is stable. For implicit methods, the CFL number is typically less than 10 to guarantee accurate results. The

proposed scheme can use larger time steps that imply a CFL number above 20. The simulation results are summarized in Table 2-2.

For the case of CFL=20, Figure 2-3 shows the instantaneous velocity field; Figure 2-4 shows the instantaneous vorticity field. Karman vortex shedding in the wake region can be clearly observed. Additionally, the drag coefficient, lift coefficient, and Strouhal number are in good agreement with the results presented in [54]. Based on the results with CFL=20, it is clear that ASTS not only allows a larger CFL number, but also obtains accurate results in terms of the aerodynamic properties.

As the CFL number increases, the discrepancies with respect to the results in [54] become larger. When CFL=40, the drag coefficient and the Strouhal number are slightly overestimated, but the error is within 3%. Once the CFL number is greater than 80, the discrepancies in both the lift coefficient and the Strouhal number are greater than 10%. Overall, when the CFL number is less than 40, ASTS can still obtain accurate results.

Figure 2-5 shows the instantaneous pressure fluctuation, which is given by

$$P' = \frac{P - \bar{P}}{\rho_0 c^2}, \quad (28)$$

where $\bar{P}$ is the time-averaged pressure field, ρ_0 is the ambient density, and c is the speed of sound. For CFL=20 and CFL=40, the acoustic field presented in the figure compares well with the corresponding field reported by [55]. The propagation of pressure fluctuations from the cylinder to the far field can be observed, and the vortex sound source behind the cylinder can be identified. For the case of CFL=80, the enhanced numerical dissipation caused by the large time step degrades

the shape of the wave, especially near the vortex sound source, which changes the wavelength of the propagation.

The instantaneous pressure fluctuations along the line $x=0$ above the cylinder are compared in Figure 2-6. Note the extremely small scale of the pressure fluctuations, which are $\sim 0.01\%$ of the dominant pressure, $\rho_0 c^2$. A highly accurate scheme is required to capture such small-scale phenomena. For the case of CFL=20 and CFL=40, the distribution is highly consistent with that reported in [54]. However, for CFL=80, the amplitude is underestimated while the wavelength is overestimated.

From the above results, it can be concluded that ASTS not only enhances the computational efficiency by enabling a larger time step, but also obtains accurate results. It has been shown that ASTS can be applied to acoustic problems. As well as increasing the time step to accelerate the time required to reach the quasi-steady state, ASTS provides a robust method for adjusting both the calculation time and the accuracy.

Figure 2-7 shows the history of the drag and lift coefficients with respect to (dimensionless) time (D/U) for four sphere cases. Although CFL=80 and CFL=160 produce inaccurate results as the time steps are excessively large, they reach the quasi-steady state much faster than when CFL=20. Therefore, an idea is proposed here. In the early stages, the development of the flow field can be accelerated by adopting an excessively large time step. Once the flow field reaches the quasi-steady state, the time step is gradually decreased to an appropriate value for ensuring accurate results.

An example is provided in Figure 2-8, which shows the history of the drag and lift coefficients with CFL=20 and with a variable time step. For CFL=20, the time step is constant, and

approximately 1200 iterations are required to reach the quasi-steady state. In the case with variable time steps, CFL=160 is adopted in the early stages, and then the time step is gradually decreased with a decay coefficient of 0.9 until CFL is less than 20. This decay process takes only another 20 steps ($160 \times 0.9^{20} = 19.45$). Finally, another 200 steps are required until the flow is fully developed. In total, fewer than 300 iterations are required to reach the quasi-steady state.

Compared with CFL=20 and a constant time step, this represents a reduction of over 75% in terms of the computation time, although the level of accuracy remains unchanged. Compared with the traditional LUSGS implicit scheme, where the CFL number is typically less than 10 for transient state problems, the computation time of the proposed ASTS method is up to eight times faster (four times faster with the variable time steps and twice as fast with larger time steps). This is because the calculation of Eq. (24), which is also needed for the traditional LUSGS, and Eq. (25) can be almost neglected.

To convince the current framework is reliable on three-dimensional acoustic computation, the simulation of the flow around the flat plate was conducted to validate the result in Ref [46] by Yokoyama et al.. The parameter of the simulation is shown in Table 2-3. The cube distribution and the computational condition are shown in Figures 2-9 and 2-10. According to Yokoyama et al. [46], while the inlet velocity matches the parameter of the flat plate geometry, the sound will be resonated, and the resonant frequency of the half-wavelength mode was found to be predicted by the following empirical formula:

$$f_{res} = 0.5 \left(\frac{c}{L} \right) / \left(1 + \alpha \left(\frac{s}{L} \right)^\beta \right), \alpha = 0.7, \beta = 0.84. \quad (29)$$

where c is the sound speed, L is the length of the plate, and s is the separation.

As the contours of the fluctuation pressure for phase- and spanwise-averaged flow fields shown in Figure 2-11, the expansion and compression waves radiate around the flat plate were observed, both the frequency of vortex shedding and the wavelength of the radiating waves were in a good agreement with [46]. The mean velocity field around the plate is shown in Figure 2-12. In Figure 2-13, the comparison of the mean velocity profile of the boundary layer from the upper surface also showed good accuracy. These results indicate that the current framework can adequately capture the acoustic and flow fields in three-dimensional acoustics simulation.

2.4.3 Conclusion

A new implicit time scheme based on the SLTS and LUSGS schemes has been developed for DAC problems. In the proposed ASTS scheme, the residual term is used as an indicator for adaptively determining whether the governing equation includes an artificial time step or not, thus eliminating the Newton linearization error. Simulation results show that when the CFL number is less than 40, acceptable accuracy can be achieved in both aerodynamic and aeroacoustic problems. Additionally, larger time steps reduce the calculation time. Therefore, in the early stages, larger time steps are used to quickly develop the flow, and then the time step is gradually decreased to ensure accurate results. This approach significantly reduces the computation time. According to our results, the computation time can be reduced by a factor of eight compared with the traditional LUSGS. Therefore, the proposed ASTS scheme has the advantages of high efficiency, high accuracy, and widespread applicability for both fluid and acoustic calculations.

Table 2-1. Computational parameters of flow-induced noise around a cylinder.

Time step (s)	3.26×10^{-8} ; 6.52×10^{-8} ; 1.30×10^{-7} ; 2.61×10^{-7}
CFL number	20; 40; 80; 160
Total number of cubes N_{cube}	16,186
Cell number in a cube	16×16 (2-dimensional)
Total number of cells N_{cell}	4,143,616
Size of the largest cell SL_{cell} (mm)	0.0216
Size of the smallest cell SS_{cell} (mm)	6.75×10^{-4}
Computational domain $X_1 \times X_2$ (mm ³)	11.0592×11.0592
Inlet velocity U (m/s)	69
Dimensionless pressure (Pa) $\rho_0 c^2$	1.4182×10^5
Cylinder diameter D (mm)	0.045
Reynolds number	200

Table 2-2. Comparison of results with CFL=20, 40, 80, and 160.

	Drag	Lift	Strouhal
	coefficient C_D	coefficient C_L	number St
CFL=20	1.40 (2.2%)	0.69 (1.4%)	0.201 (1.0%)
CFL=40	1.41 (2.9%)	0.71 (1.4%)	0.193 (1.3%)
CFL=80	1.46 (6.6%)	0.81 (15.7%)	0.170 (14.6%)
CFL=160	1.50 (9.5%)	0.87 (24.3%)	0.118 (40.7%)
Ref. [54]	1.37	0.7	0.199

Table 2-3. Computational parameters of flow around a 3D flat plate.

Thick of plate b	2 mm
Length of plate	30 mm
Inlet velocity	44 m/s
Re num.	5,800
Mach num.	0.13
Min grid size	0.05 mm
Cell num.	82,000,000
dt	10^{-6}

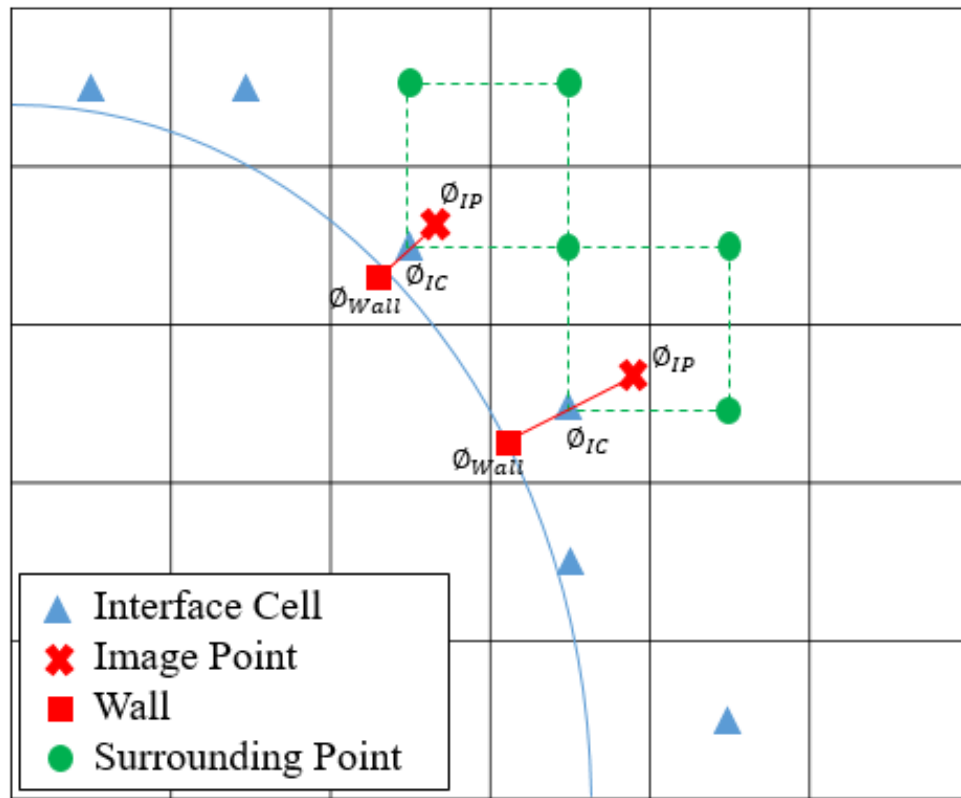
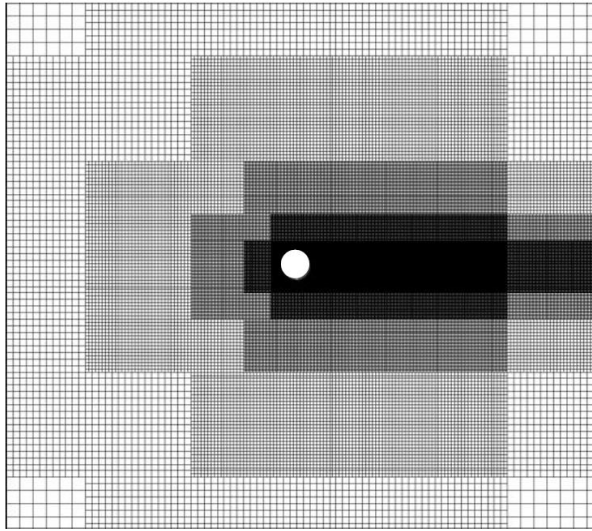
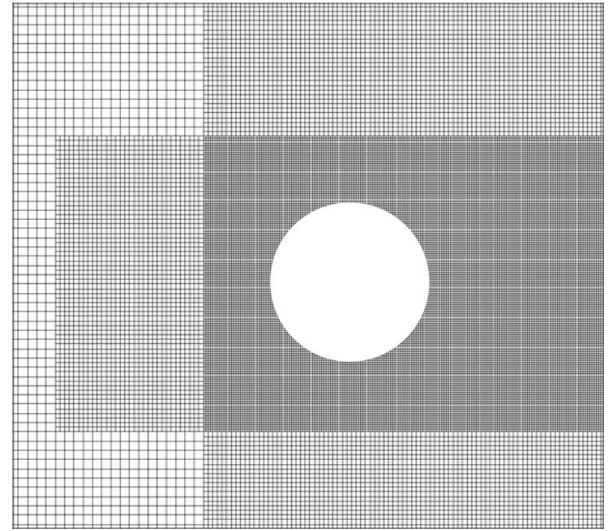


Figure 2-1. Compressible immersed boundary method configuration.



(a) Grid distribution around the cylinder



(b) Magnified view of the grid distribution

Figure 2-2. Grid distribution.

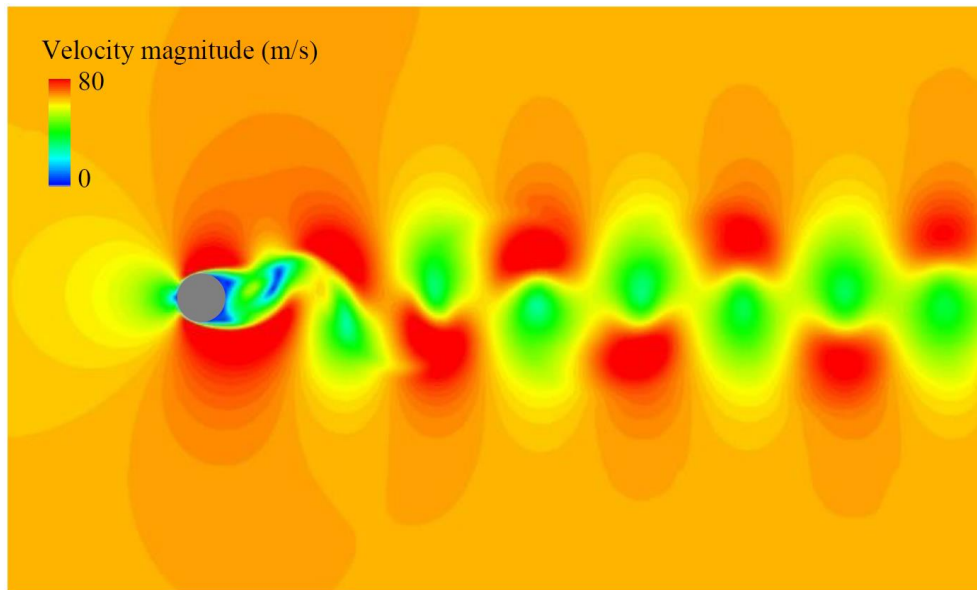


Figure 2-3. The Contours of instantaneous velocity field and Karman vortex shedding.

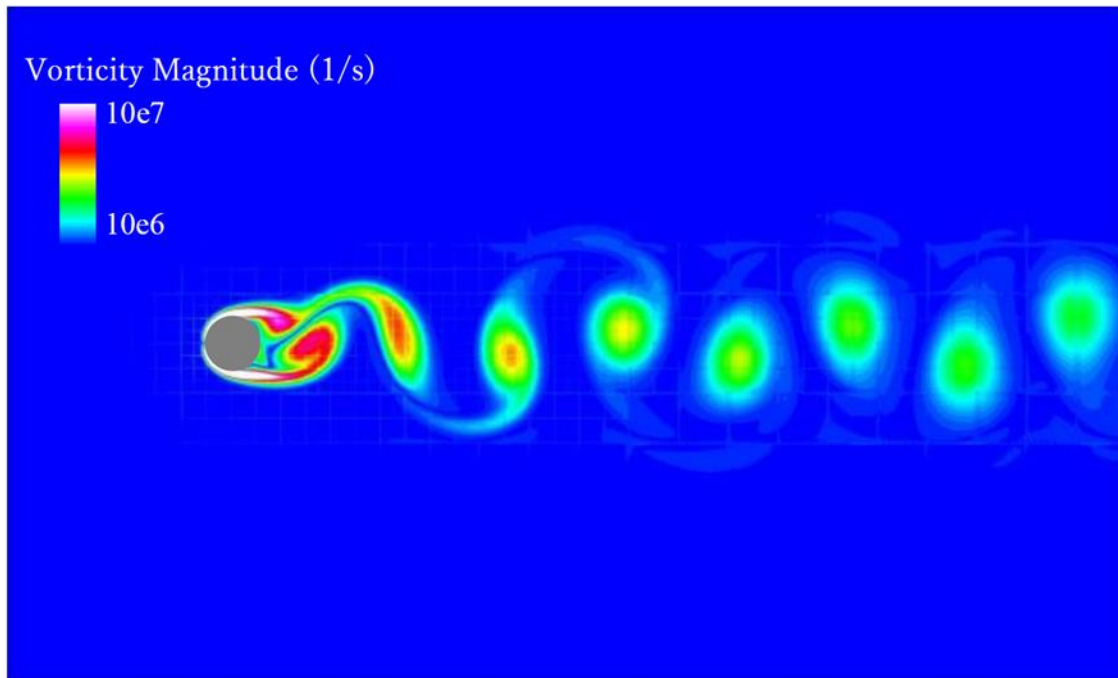


Figure 2-4. The Contours of instantaneous vorticity field and Karman vortex shedding.

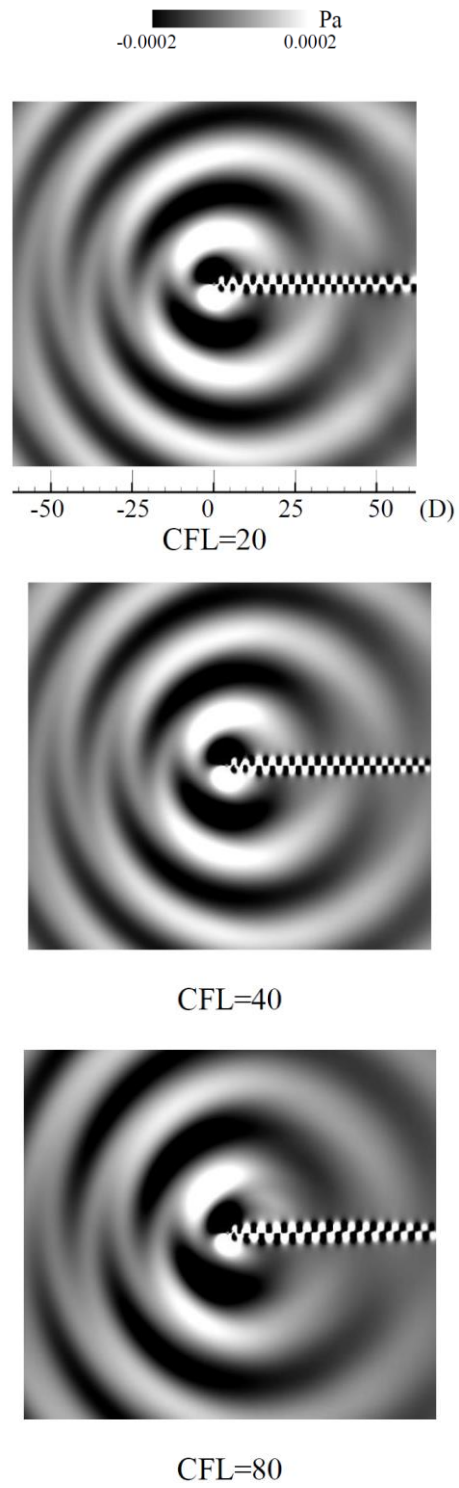


Figure 2-5. Comparison of the contours of instantaneous pressure fluctuations.

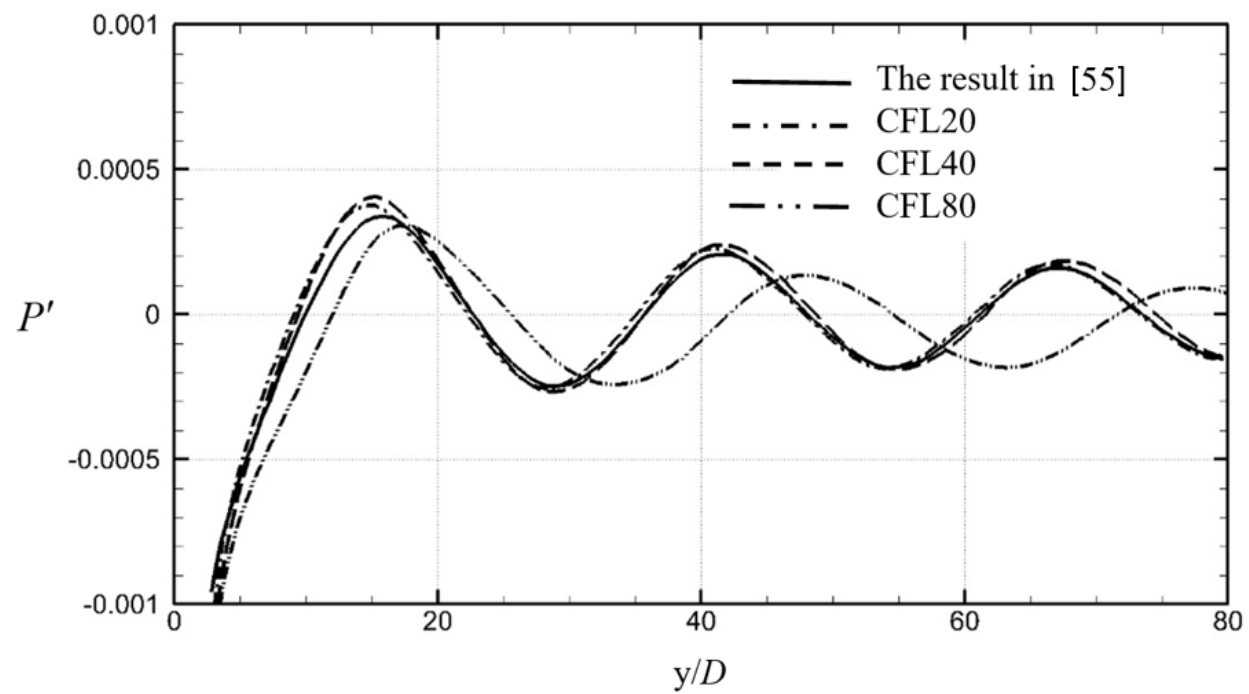


Figure 2-6. Comparison of instantaneous pressure fluctuations with results in [55].

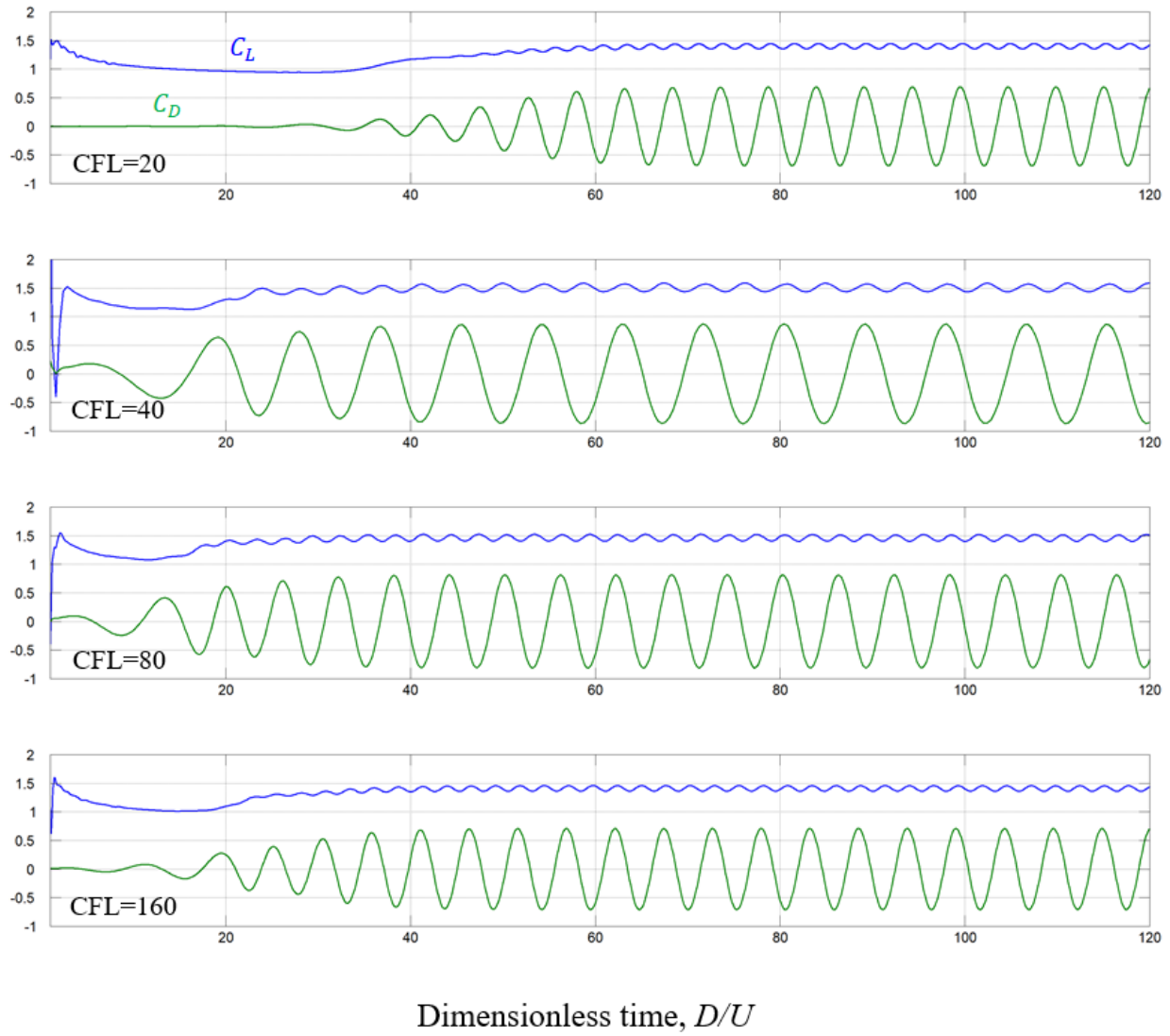


Figure 2-7. Drag coefficients C_D and lift coefficients C_L with respect to dimensionless time.

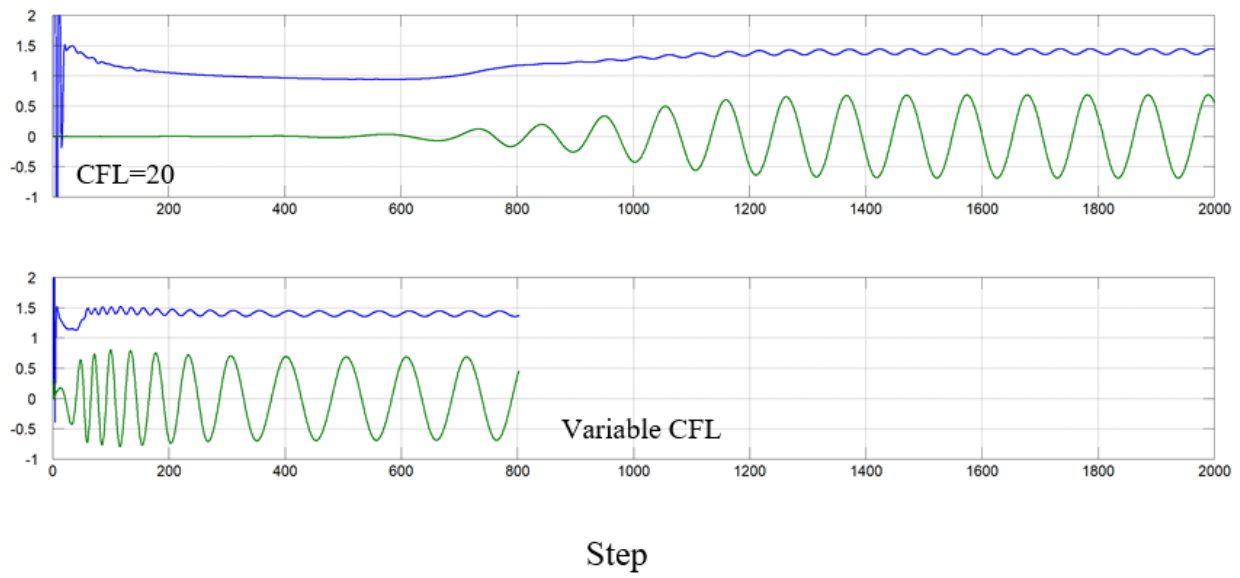


Figure 2-8. Comparison of time required to reach the quasi-steady state between constant and variable steps.

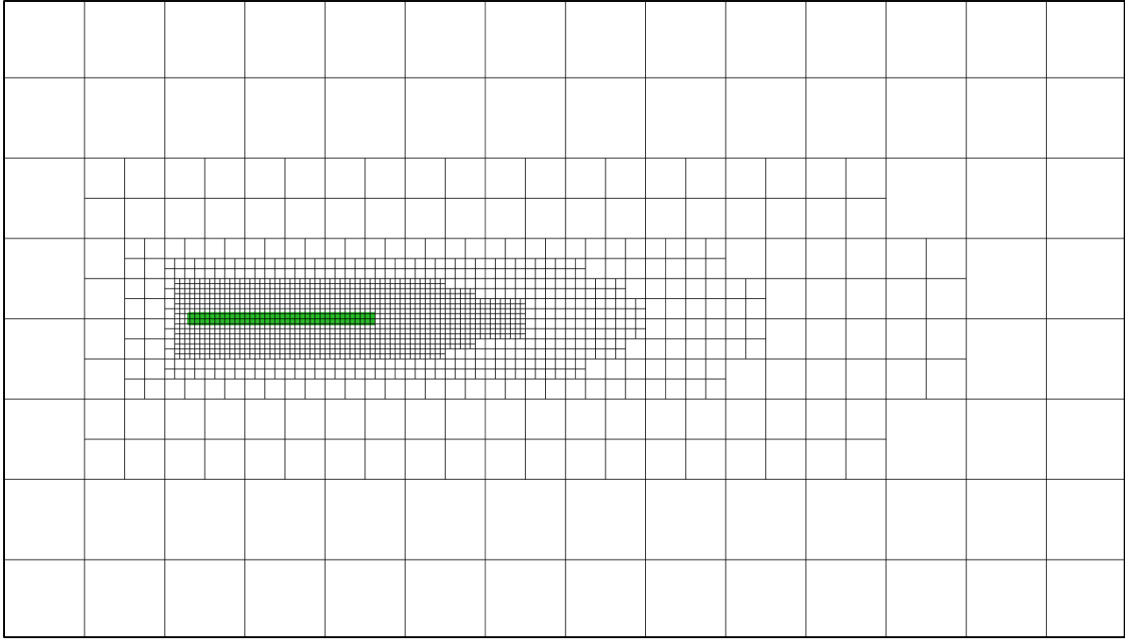


Figure 2-9. The cube distribution. (Cube division: $16 \times 16 \times 16$)

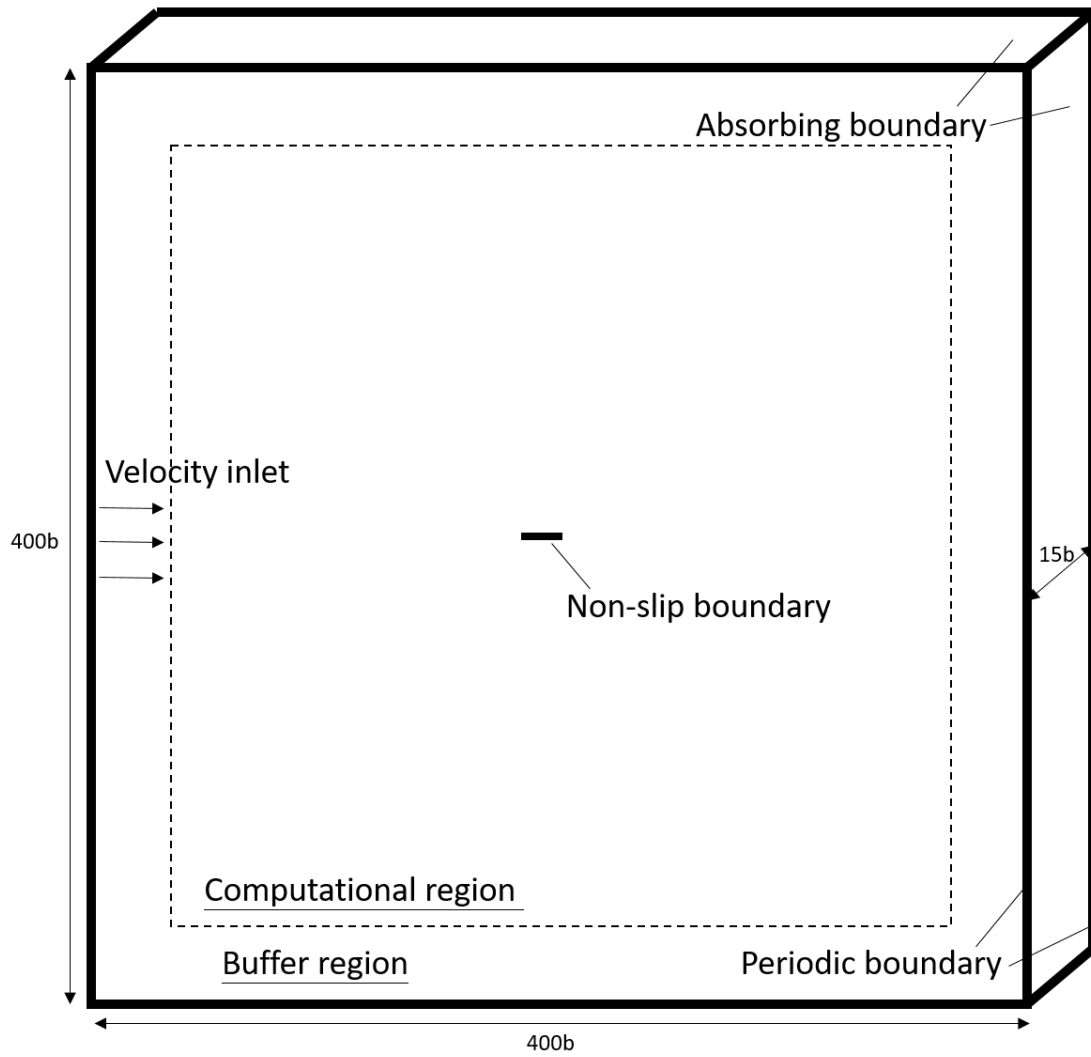
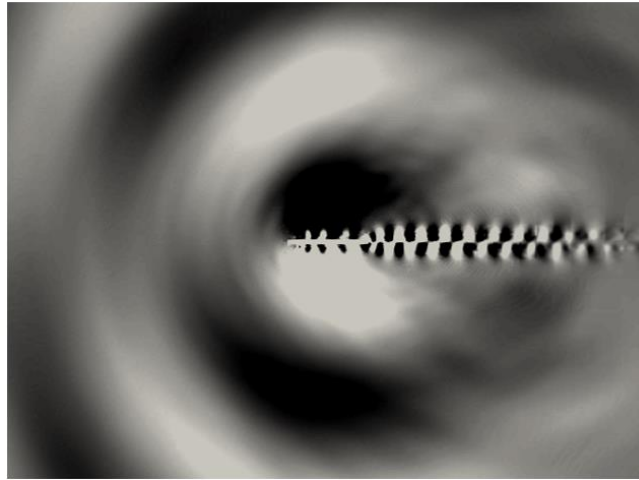
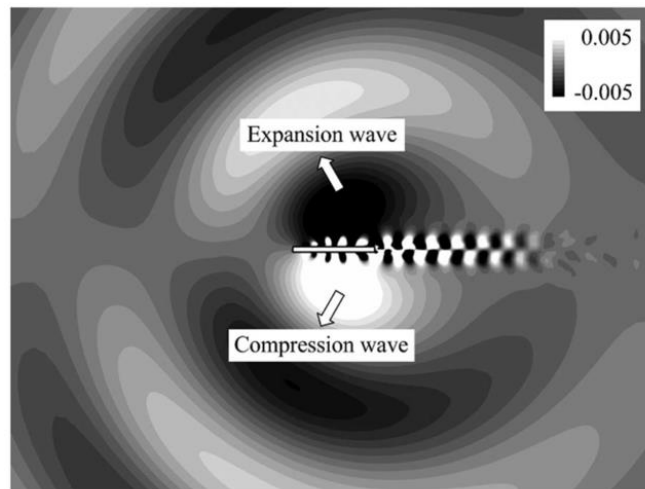


Figure 2-10. The computational condition.



(a) Present result.



(b) Benchmark result [46].

Figure 2-11. Contours of fluctuation pressure for phase- and spanwise-averaged flow fields.

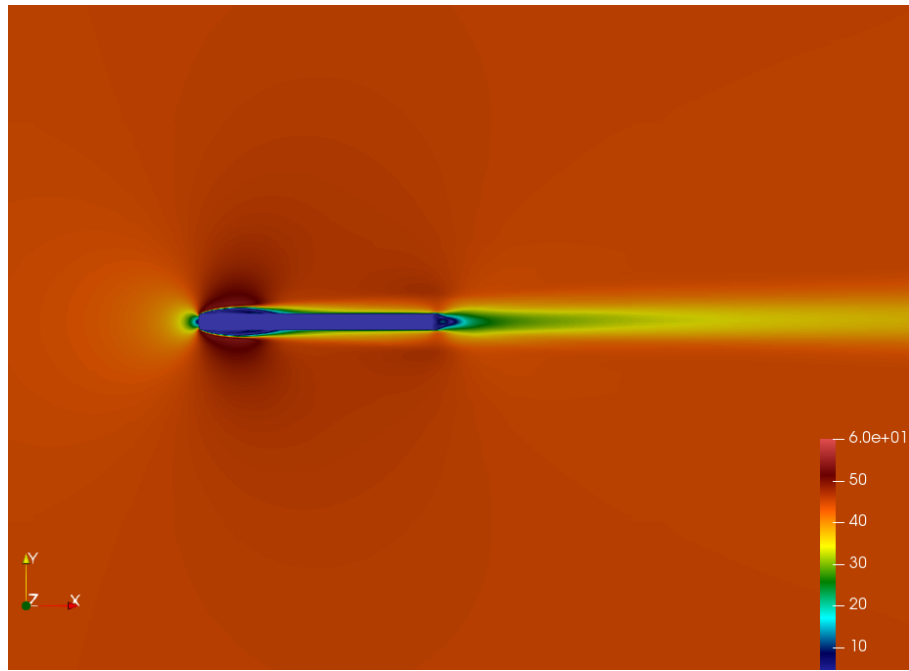


Figure 2-12. The mean velocity field around the plate.

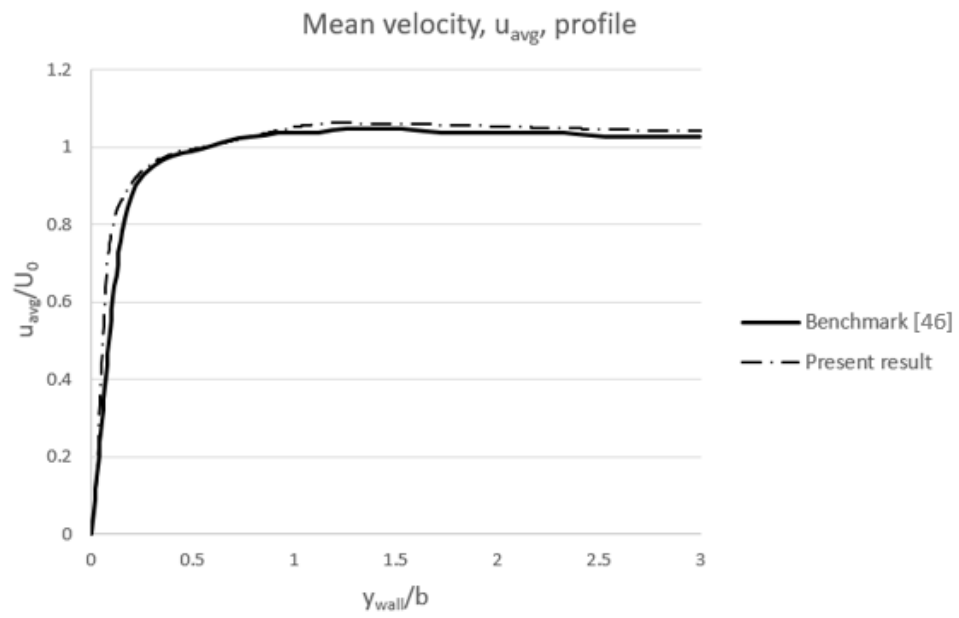


Figure 2-13. The mean velocity profile from surface of the plate at the downstream.

Chapter 3. Investigation of effects of incisor angle on production of sibilant /s/

3.1 Introduction	44
3.2 Vocal tract geometry	45
3.3 Computational conditions	47
3.4 Results and discussion	48
3.5 Conclusion	53

3.1 Introduction

In this chapter, we investigate the effects of the inclination angle of the central incisor on the speech production of the fricative consonant /s/ using an implicit compressible flow solver. The airflow and sound during the pronunciation of /s/ were simulated while the angle of the incisor in the vocal tract was increased up to 30° reported by Runte et al. [5]. Consequently, increasing the incisor angle was found to decrease the amplitudes of the sound in the frequency range from 8 kHz to 12 kHz and affected the flow configuration.

Furthermore, instead of a constant velocity inlet from the throat, the simulations with acoustic monopole sources were conducted for the 0° and 30° cases to determine the relationship between the positions of the velocity fluctuations and the far-field SPL spectra. Performing a Fourier transform on the velocity fluctuation, we found that the large magnitudes at 10 kHz moved toward the lip outlet when increasing the incisor angle. Moreover, separate acoustic simulations showed that the shift in the sound source position toward the lips decreased the far-field sound amplitudes. These results indicate that the change in the flow configurations and the aeroacoustic sound sources caused a decrease in the far-field sound amplitudes of /s/.

These findings clarify the effects of geometrical differences on flow and sound generation. To be more specifically, the raising incisor angle will affect the flow configuration in the vocal tract and reduce the sound amplitude of the sibilant /s/ sound. These results provide the underlying insights necessary to design dental prostheses for the production of sibilant fricatives.

3.2 Vocal tract geometry

To simulate the phenomena involved in the pronunciation of /s/, the geometry of a vocal tract replica was constructed from CT images [16], as shown in Figure 3-1. The subject was a 32-year-old Japanese male who self-reported no speech disorders with a normal dentition of angle class I, without inter-teeth spacing. The effect of inter-teeth spacing is out of the scope of this paper and thus it will be future work. In addition, the sidewalls of inclined incisors were filled with a straight wall to prevent the formation of small gaps that lead to the instability of flow simulation.

The CT images were taken while the subject sustained the pronunciation of /s/ for 9.6 s without vowel context, and the image resolution was $0.1 \times 0.1 \times 0.1 \text{ mm}^3$. The surfaces of the vocal tract geometry were extracted based on brightness values using the software itk-SNAP [56]. Since the resolution of the CT scan is 0.1 mm and the tips of the incisors were slightly smoothed through the vocal tract extraction from the CT scan, we have confirmed that the extracted vocal tract geometry reproduced the subject's pronunciation of /s/ up to 14 kHz by constructing an oral replica [16]. The ethics committee of the graduate school of Osaka University certified this study (H26-E39).

Figure 3-2(a) shows the vocal tract geometry, including the throat, tongue, hard palate, incisors, and lips. The x_1 is defined as the anterior-posterior direction; the x_2 is defined as the inferior–superior direction; x_3 is defined as the transverse direction. The initial inclination angle of the incisor to the maxillary plane was 108° for this subject. For the inclined cases, the variation from $+10^\circ$ to $+30^\circ$ was resulted in 118° to 138° based on this original incisor angle, and the region of the inclination incisor is marked in red color ($-10 \text{ mm} < x_3 < 10 \text{ mm}$). The geometries from the top and side views are shown in Figure 3-2(b) and 3-2(c), respectively.

We confirmed that the exclusion of the upstream vocal tract geometry on the simulation of /s/ was negligible for the main acoustic characteristics of /s/ [18]. According to Runte [5], by comparing the sound of /s/ generated by the human subject with the subject's vocal tract replica made of plaster using a 3D printer, the frequency characteristics of /s/ were produced up to 16 kHz with the maximum discrepancy of 8 dB. It indicates that the solid wall condition is valid to investigate the sound mechanisms up to 16 kHz.

To investigate the effect of the inclination angle of the upper incisor, the incisor position ($-17.7 \text{ mm} < x_1 < -12.2 \text{ mm}$) was raised from its original position (Figure 3-3(a), 0°) up to 30° (Figure 3-3(b)). Figure 3-3(c) shows the mid-sagittal plane in the modified geometry ($x_3 = 0$) with the incisor angle increased in 10° increments from 0° to 30° , the length of the incisor is kept as the same as the original geometry.

Table 3-1 shows the overjet (horizontal overlap) and overbite (vertical overlap) of the central incisor angle in each model. As shown in Figure 3-3(a), the overjet (horizontal overlap) and overbite (vertical overlap) of the incisor angle in the original model are 2.3 mm and 0.3 mm, respectively. When the inclination angle increases, the x_1 distance between the upper and lower incisors becomes longer; therefore, the overjet increases from 2.3 mm to 2.5 mm. Conversely, the x_2 distance becomes wider; therefore, the overbite decreases from 0.3 mm to -1.7 mm. These values are within the range of clinical measurements [12]. The sidewalls of each tooth were modified to a smooth appearance to prevent the formation of small gaps between the teeth that lead to the instability of flow simulation.

3.3 Computational conditions

After testing the space resolution shown in Table 3-2, the minimum grid size was set as 0.05 mm around the upper incisors to keep the accuracy of the immersed boundary around the turbulent region. The cube and cell distributions are shown in Figure 3-4, while the cube division is $16 \times 16 \times 16$ cells. To simulate the sound waves propagating through the lip outlet, the far-field region was set outside of the vocal tract model. The total grid numbers of the 0–30° cases were 8.7×10^7 , 7.2×10^7 , 7.8×10^7 , and 6.7×10^7 , respectively.

The three-view diagram of overall computational domain with the boundary condition and tracking point for the current model are shown in Figure 3-5. The inlet was set to a uniform velocity condition to simulate the pronunciation of /s/. The uniform inlet velocity was set to 1.5 m/s, which resulted in a physiological flow rate of 430 cm³/s [17]. The Reynolds number was 5632, based on the maximum velocity ($\overline{|u|}_{\max} = 50.8$ m/s) inside the oral cavity in the original geometry.

To keep the flow in the computational domain from being polluted by reflecting pressure waves, an absorbing boundary condition was used as the outlet condition. The absorbing boundary condition used in the current study is based on JB Freund [57] and extend by Li [52] which is adjusted for the low flow speed simulation. The time step was set to 10^{-6} s, so that the CFL number was 6.8, which fulfilled the condition for the ASTS method. The physical time of the performed simulations was 0.015 s and required parallel computing with 1152 cores on 32 nodes for 30 h. Table 3-3 summarizes the computational parameters for the simulations.

Fast Fourier transform (FFT) using the Hann window was applied to the waveforms sampled 100 mm from the lip outlet ($x_l = 100$ mm) to analyze the far-field sound spectrum. The FFT sampling frequency was 50 kHz with 256 points averaged five times. The sound pressure level

(SPL) was calculated based on the reference pressure $P_{\text{ref}} = 20 \times 10^{-6}$ Pa. In addition to the sound spectrum, the magnitudes of the velocity fluctuations at each frequency were calculated via FFT on each grid to identify the highest contribution position for the potential sound source [58].

3.4 Results and discussion

To ascertain the accuracy of the simulation in the present framework, the result was compared with experimental results and measurement of the actual subject. The acoustic pressure was collected 100 mm from the lip outlet ($x_1 = 100$ mm), and the frequency spectra of the sound was calculated via FFT, as shown in Figure 3-6. We have confirmed that the velocity magnitude around the tracking point is small enough to be neglected, which is less than 0.05 m/s. In the experiment, the vocal tract replica was constructed using a 3D printer (Objet30Pro, Stratasys, USA; accuracy: ± 0.1 mm) and a constant flow was input to the model using a compressor (YC-4RS, Yaezaki, Tokyo, Japan), as shown in Figure 3-7. The sound generated by the model was measured using a microphone (type 4939, Bruel & Kjaer, Nærum, Denmark) at a distance of $x_1 = 100$ mm in an anechoic chamber (with a volume of 8.1 m^3). The measurement result was recorded with an actual subject for 18 times of utterances and plotted on Figure 3-6 in blue dot with error bar. The subject sustained /s/ for 3 seconds without vowel context, and before the signal was calculated via FFT, some portions of signals after the onset and before the offset of /s/ are removed. Therefore, inclusion of the effect of the onset and offset was avoided.

The sibilant sound is characterized as a broadband noise above 4 kHz. At a frequency of 3 kHz, SPL rapidly increased, and the first characteristic peak reached around 5 kHz. This result shows

good agreement with the measurements for the actual subject and oral replica, and the same characteristics of sibilant /s/ sound was observed by Runte [5]. Accordingly, this result suggests that the present framework of the direct aeroacoustic computation can provide aeroacoustic predictions with reasonable accuracies.

Figure 3-8 shows the normalized instantaneous velocity magnitude $|\tilde{u}|/|\tilde{u}|_{max}$ and the root mean square (RMS) value of the velocity fluctuation $|\tilde{u}|_{rms}/|\tilde{u}|_{max}$ on the mid-sagittal plane, where $|\tilde{u}|_{max} = 50.8$ m/s. The highest velocity occurs at the narrow channel between the tongue and the hard palate (the sibilant groove). Downstream of the sibilant groove, the flow became turbulent at the region between the teeth and the lower lip ($-15 \text{ mm} < x_l < -11 \text{ mm}$) as a result of the jet flow leaving the sibilant groove. Accordingly, the velocity RMS is relatively high in the area downstream of the jet flow. According to Lighthill's analogy [56], the aeroacoustic sound source is primarily produced by the time variation in the space derivatives of the Reynolds stress tensor. Therefore, the sound source likely appeared in the region of high RMS values at the chamber between the teeth and the lower lip.

To investigate the effects of the incisor angle, the incisor angle was varied from 0° to 30° ; the far-field SPL spectra for these angles are shown in Figure 3-9. According to Runte [5], when the incisors were moved, a significant decrease in the upper boundary frequency was observed for the /s/ sounds. In contrast, in the lower boundary frequency of the noise band were observed increasing while the incisor moving labially and decreasing while the incisor moving palatally. In this study, amplitudes from 8 kHz to 12 kHz were decreased by the increase of incisor angle. This means that the upper boundary frequency of noise was decreased and the noise band range became smaller with the inclined incisors. While the noise band range of 0° was approximately 4 kHz to 12 kHz, the noise band range of 30° was 4 kHz to 8 kHz. This result is consistent with that found by Runte

[5]. Besides, according to Snow [59], general audible frequency range for male and female speech are up to 7 kHz, and 9 kHz, respectively, and the characteristic peak of the sibilant sound at around 4 kHz was observed for all teeth angles. Therefore, the sounds generated by all the cases can be characterized and recognized as sibilant fricative /s/. However, the decreasing amplitudes in the frequency range 8 kHz to 12 kHz might affect the recognition of the sound.

Figure 3-10 shows the normalized instantaneous velocity magnitude $|u|/\overline{|u|}_{max}$ and root mean square (RMS) value of the velocity fluctuations $|u|_{rms}/\overline{|u|}_{max}$ from the 10° to 30° models on the mid-sagittal plane ($x_3 = 0$). As shown in Figure 3-10(a), 3-10(c), and 3-10(e), the instantaneous velocities of all cases are accelerated at the narrow channel between the tongue and the hard palate (the sibilant groove) ($-25 \text{ mm} < x_1 < -16 \text{ mm}$). Downstream of the sibilant groove, the flow became turbulent at the region between the teeth and the lower lip ($-15 \text{ mm} < x_1 < -11 \text{ mm}$) as a result of the jet flow leaving the sibilant groove. However, because the exit of the sibilant groove became wider in superior direction with the inclined incisors, the reduction of the occlusion made the mainstream flow faster and the flow reached more distant positions. Hence, the high RMS region shown in Figure 3-10(b), 3-10(d), and 3-10(f) moved from the cavity between the lower incisor and the lip to the tip of the lower lip with increasing incisor angle. Meanwhile, the turbulence intensity was not affected by the raised incisor angle. According to Lighthill's analogy [56], the aeroacoustic sound source is primarily produced by the time variation in the space derivatives of the Reynolds stress tensor, which means that the sound source likely appeared in the region of high RMS values. Therefore, the different flow configurations caused by the raised incisor angle might be considered as the reason for the difference in the acoustic field.

To identify the cause of the different SPL values around 10 kHz between 0° and 30° in Figure 3-9, the positions of the potential sound sources for the 0° and 30° models were calculated. The

magnitudes of the velocity fluctuations at specific frequencies (5 kHz and 10 kHz) were calculated via FFT on each grid. The magnitudes of the velocity fluctuations inside the vocal tract for the 0° model along the x_1 - x_2 -plane are shown in Figure 3-11(a) at 5 kHz ($x_3 = 1.1$ mm) and Figure 3-11(b) at 10 kHz ($x_3 = -7.7$ mm), respectively. At 5 kHz, the maximum value is located behind the incisor, which is the exit of the sibilant groove. As seen in the flow configurations (Figure 3-8(a) and Figure 3-10(e)), the jet flow emerged at this position. Conversely, the maximum value at 10 kHz appeared at the cavity between the teeth and the lower lip, which is at the position (49.1, -8.4, -7.7). This corresponds to the exit of the jet flow being the gap between the teeth.

The magnitudes of the velocity fluctuations for the 30° model at 5 kHz ($x_3 = 1.1$ mm) and 10 kHz ($x_3 = -4.5$ mm) are shown in Figure 3-11(c) and 3-11(d). The maximum value for 5 kHz appeared at the exit of the sibilant groove, which is the same as the position for the 0° model at 5 kHz. These results indicate that the velocity fluctuations downstream of the constriction formed the characteristic peak at 5 kHz for both the 0° and 30° models. Besides, it also can explain the increasing amplitude around 4-6 kHz in Figure 3-7. Based on Figure 3-8(a), 3-10(e), 3-10(c), and 3-10(e), the maximum velocity which is increasing with the raising incisor angle is at the sibilant groove and it will affect the amplitude around 4-6 kHz. Conversely, at 10 kHz, the jet flow of the 30° model passed along the surface of the incisor and the maximum value of the velocity fluctuations appeared above the lower lip, which was at the position (58.5, -7.7, -4.5).

To include whole range of the reduced frequency in Figure 3-9 from 8-12 kHz, besides of the single frequency (10 kHz), the band-pass filter field of 8-12 kHz for 0° and 30° models are shown in Figure 3-12. For the 0° model, the high value region also occurs between the teeth and lower lip. As for the 30° models, the position of maximum value is slightly different but it still locates out of the oral cavity. These results show consistency with Figure 3-11.

To determine the relationship between the positions of the velocity fluctuations, i.e., the assumed aeroacoustic sound sources, and the far-field SPL spectra, instead of the base simulation with constant velocity inlet from the throat, the acoustic simulations with acoustic monopole sources were conducted for the 0° and 30° cases. In the previous acoustic studies, monopole to quadrupole sound sources were used to emulate the sound generated by the turbulent flow [60] [61]. Therefore, for the simplicity, the monopole sources composed of white noise were applied in the current study at Point 1 (49.1, -8.4, -7.7) and Point 2 (58.5, -7.7, -4.5), which corresponded to the positions of maximum velocity fluctuations at 10 kHz for both models, as shown in Figure 3-13.

A comparison of the far-field sound spectra in Figure 3-14 shows an obvious amplitude difference between the two sound source positions in the frequency range of 4–8 kHz. When the acoustic source is located at Point 1, SPL shows the characteristics of the sibilant sound for both the 0° and 30° models. Conversely, SPL generated by the sound source at Point 2 consisted of a broad noise with amplitudes lower than those generated at Point 1. This is because the acoustic source at Point 1 was located inside the modeled vocal tract geometry and the sound wave was resonated by the front oral cavity. The distance from the glosso-palatal constriction to anterior lip surface was 1.39 cm, which was close to the quarter of wavelength around 6 kHz. Therefore, SPL around 6 kHz was larger and the sound generated by Point 1 still had the characteristics of the sibilant sound in the far field. In contrast, because the acoustic source at Point 2 was located almost entirely outside of the vocal tract, the source did not couple strongly with the resonator, and no significant sound could be caught in the entire frequency range at the far-field. For this reason, the far-field SPL around 10 kHz was smaller in the 30° case. Consequently, the slight change of the

geometry affected less to the turbulent intensity, but changed the flow configuration and shifted the position of the potential sound source, thereby influenced the far-field SPLs of /s/.

3.5 Conclusion

To clarify the effect of the inclination angle of the incisor on the generating of the sibilant /s/ sound, numerical flow simulations of a vocal tract geometry with different incisor angles was conducted in this chapter. According to the far-field SPL spectrum, increasing the incisor angle had no influence on the characteristic peak of the sibilant sound at 4 kHz, but reduced the amplitude of the sound in the frequency range from 8 kHz to 12 kHz. In the flow field, the flow configuration inside of the vocal tract had been changed by the increasing incisor angle, which also shift the position of the potential sound source.

By setting the monopole source at the different sound source positions of the 0° and 30° cases, we found that if the sound source position was located closer to the exit of the vocal tract, the source didn't couple strongly with the resonator, thereby at far-field no significant frequency could be observed. Consequently, the raising angle of the incisor affected less to the turbulent intensity, but changed the flow configuration and shifted the position of the potential sound source, thereby reducing the amplitude of the sibilant /s/ sound.

Table 3-1. Overjet (horizontal overlap) and overbite (vertical overlap) for the 0° to 30° models.

	0°	10°	20°	30°
Overjet (mm)	2.3	2.36	2.43	2.5
Overbite (mm)	0.3	-0.4	-1.05	-1.7

Table 3-2. test of grid resolution.

Minimum grid size	0.1 mm	0.05 mm	0.025 mm
Total grid number	26 million	87 million	399 million
Error in SPL	9.6 %	4.5 %	4.2 %

Table 3-3. Computational parameters of the 0° model.

Size of the largest cell (mm)	1.6
Size of the smallest cell (mm)	0.05
Computational domain (mm ³)	409.6 × 409.6 × 409.6
Inlet velocity (m/s)	1.5
Re	5,632
Time step (s)	10 ⁻⁶
Physical time (s)	0.015

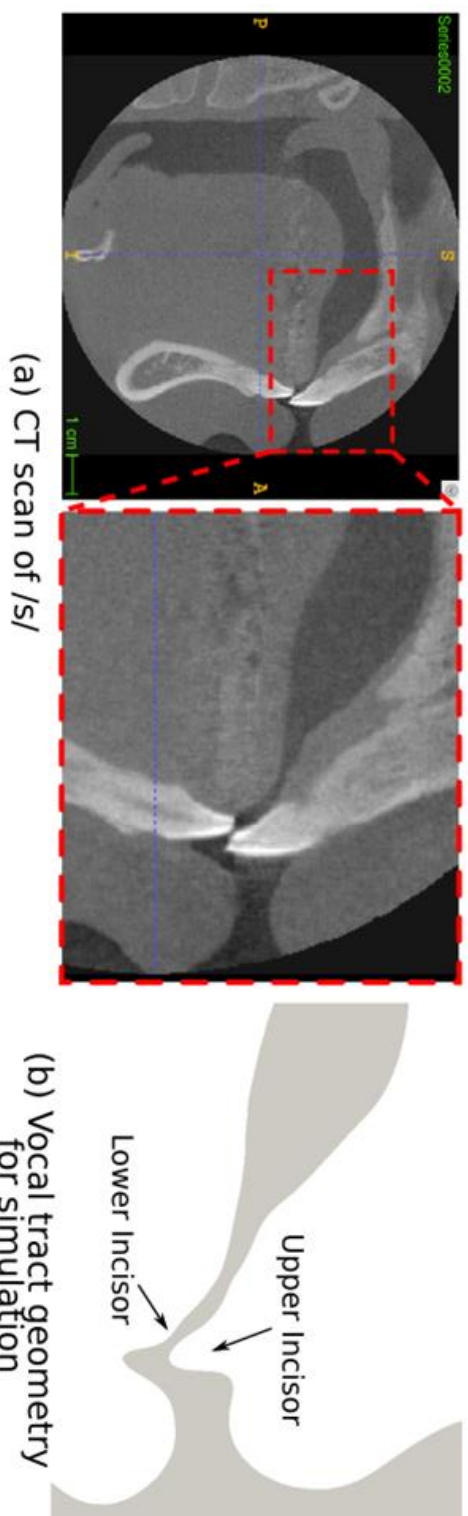


Figure 3-1. (a) CT scan image of /s/. (b) Vocal tract geometry for the simulation.

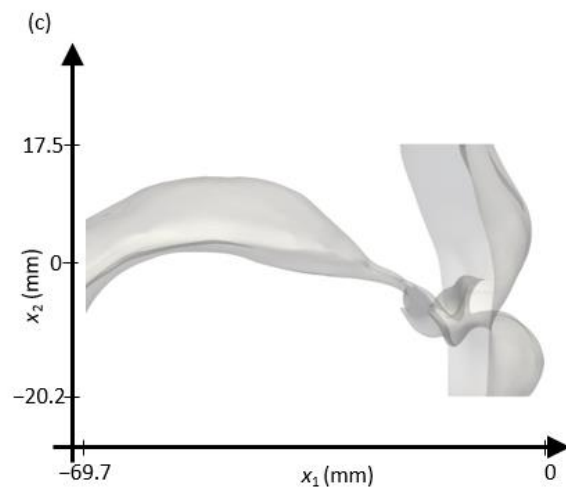
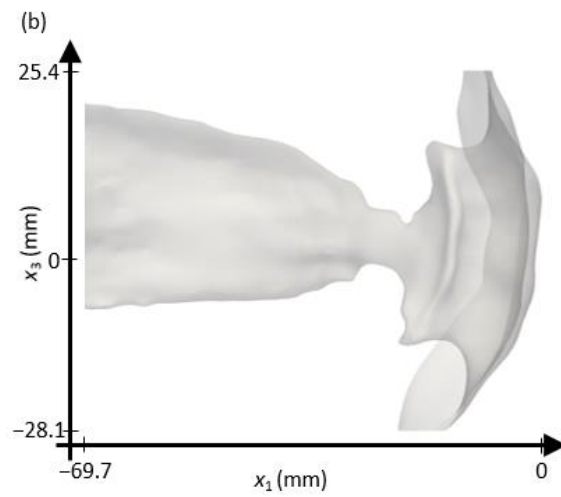
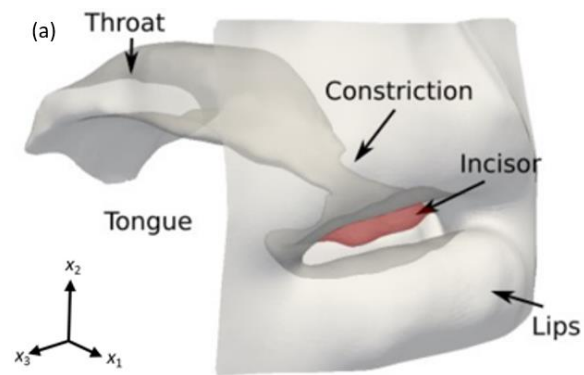
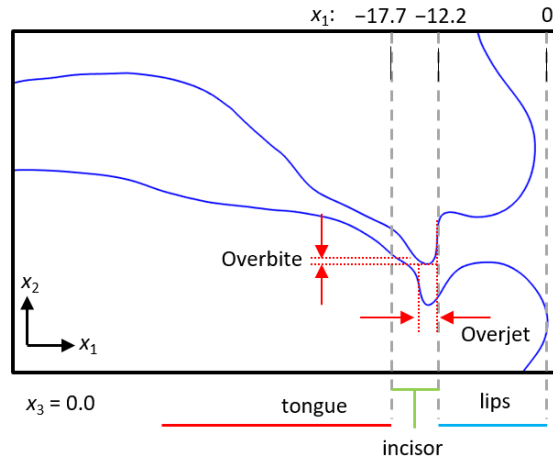
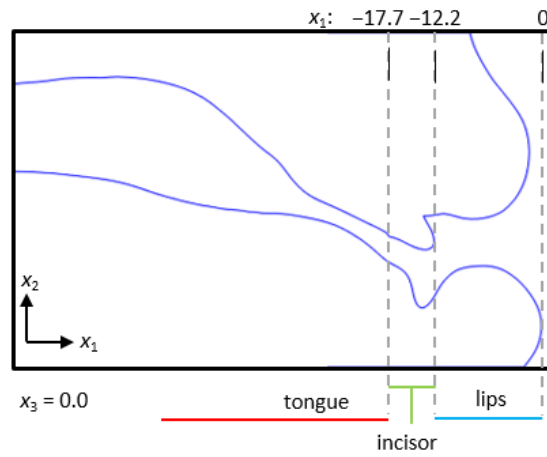


Figure 3-2. (a) Vocal tract geometry for pronouncing /s/.

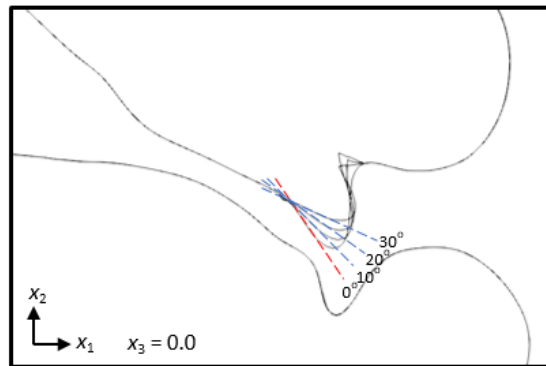
(b) Top and (c) side views of the vocal tract.



(a) Original geometry with a 0° incisor



(b) Modified geometry with a 30° incisor



(c) Modified geometry with the incisor angle ranging from 0° to 30°

Figure 3-3. Vocal tract geometry with the incisor angle increased from the original position (0°) to 30°: (a) original geometry with 0°; (b) modified geometry with 30°; and (c) modified geometry with the incisor angle ranging from 0° to 30°.

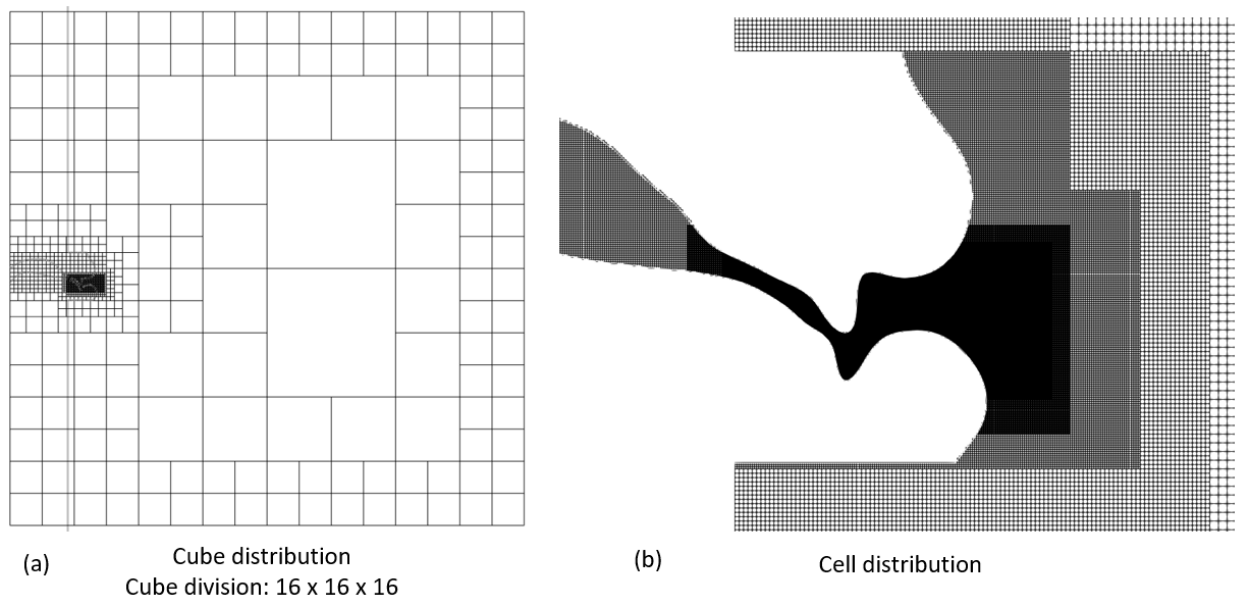


Figure 3-4. (a) Cube and (b) cell distributions.

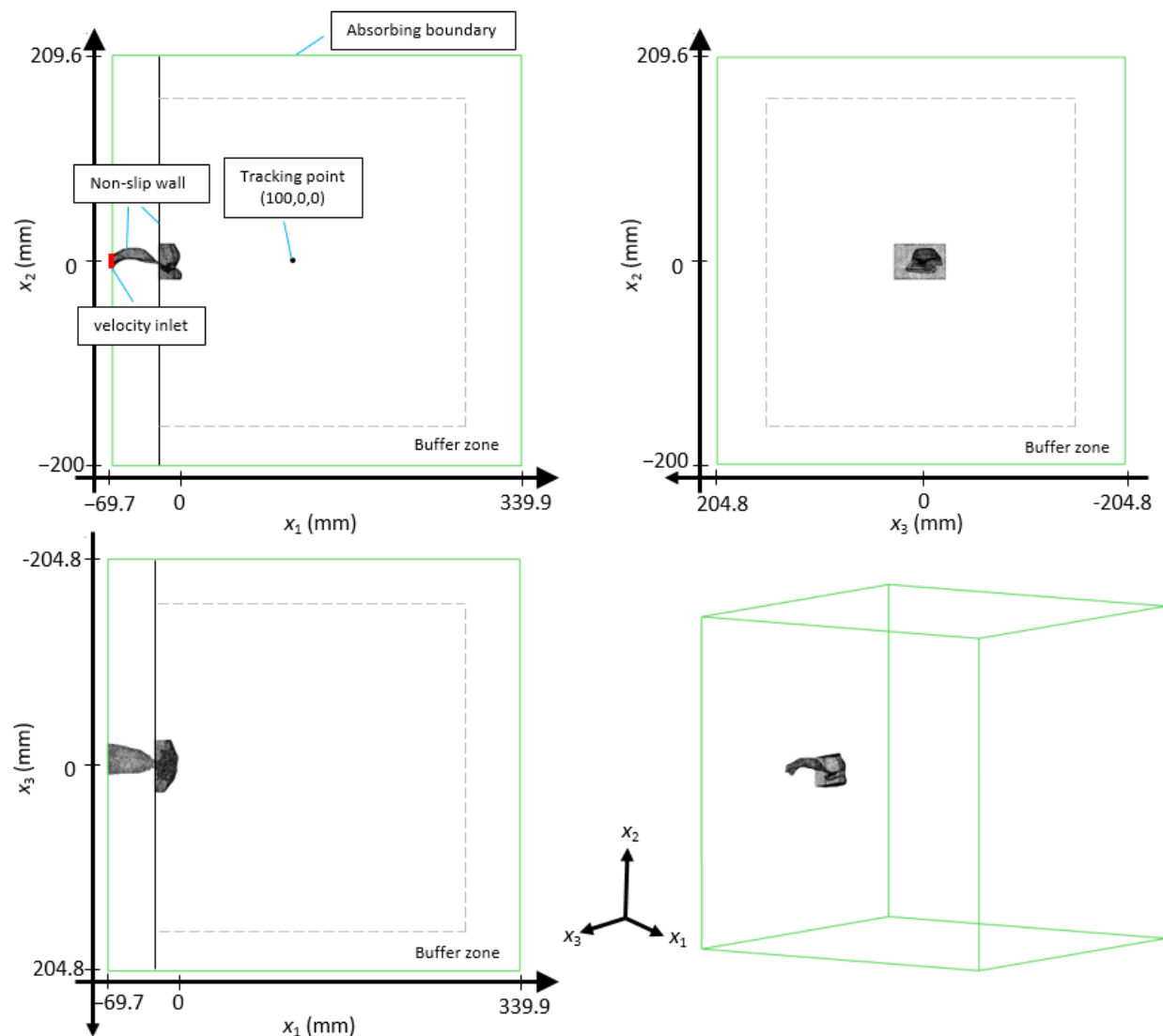


Figure 3-5. The three-view diagram of overall computational domain of the current model.

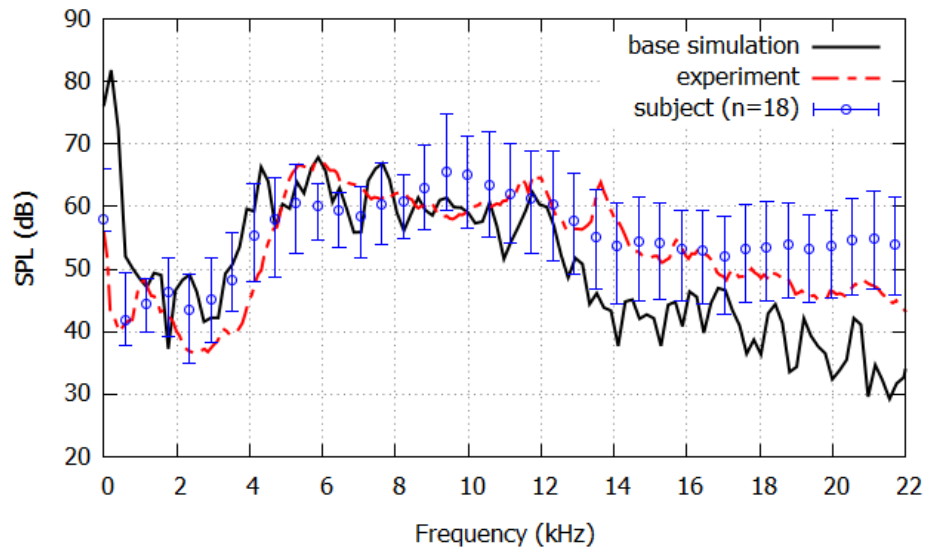


Figure 3-6. Sound pressure level (SPL) spectrum of the base simulation (0°, solid black line) and measurements of the printed replica (dashed red line) and the average value of an actual subject for 18 recordings (blue circles). The error bars denote the range for 18 repeated recordings.

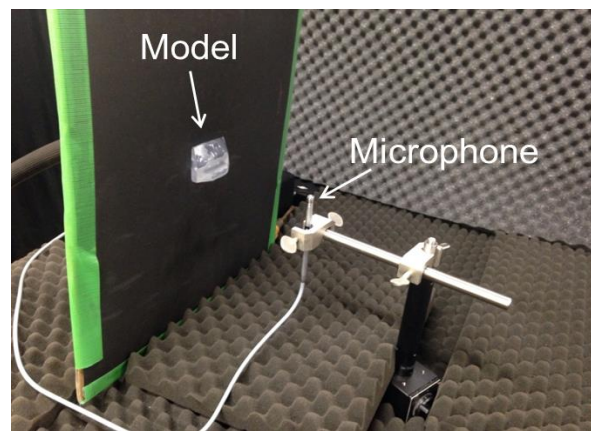
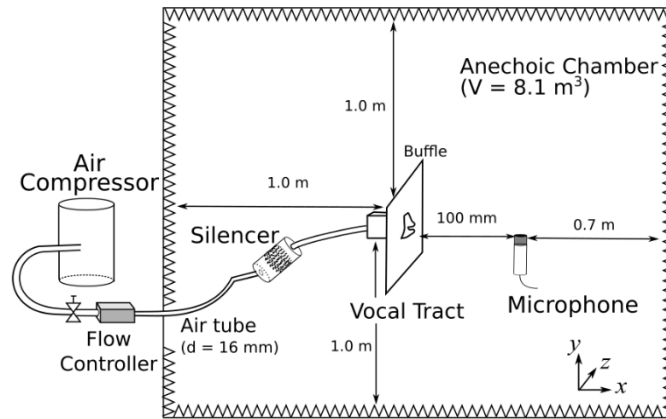


Figure 3-7. The experimental setting.

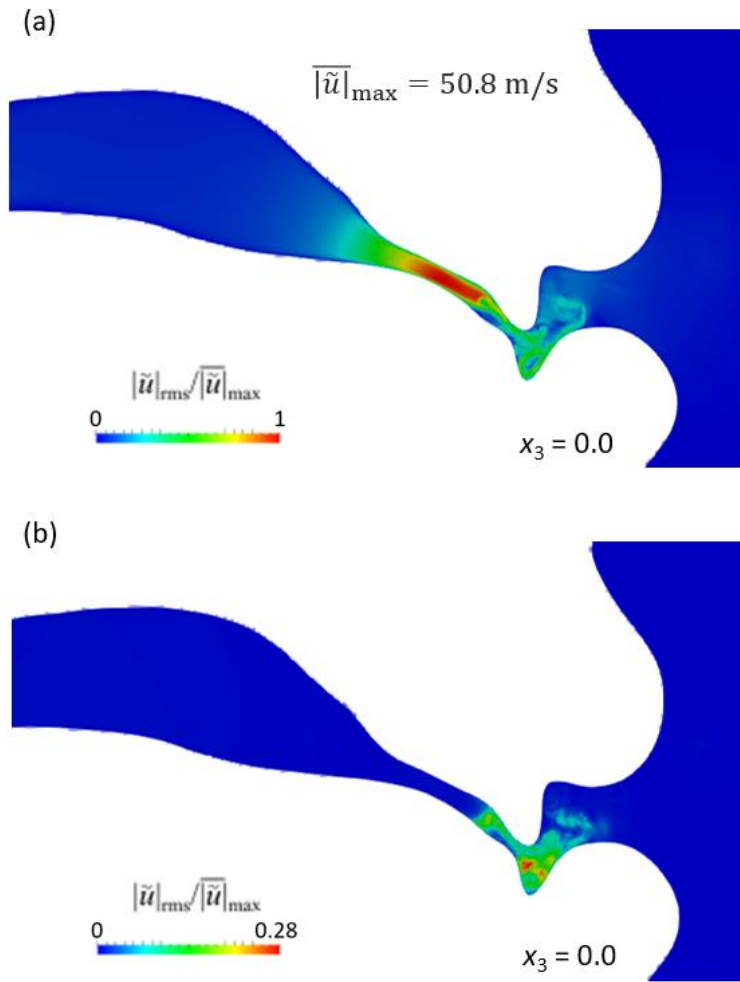


Figure 3-8. (a) Normalized instantaneous velocity magnitude and (b) root mean square (RMS) of the velocity fluctuations for the 0° model.

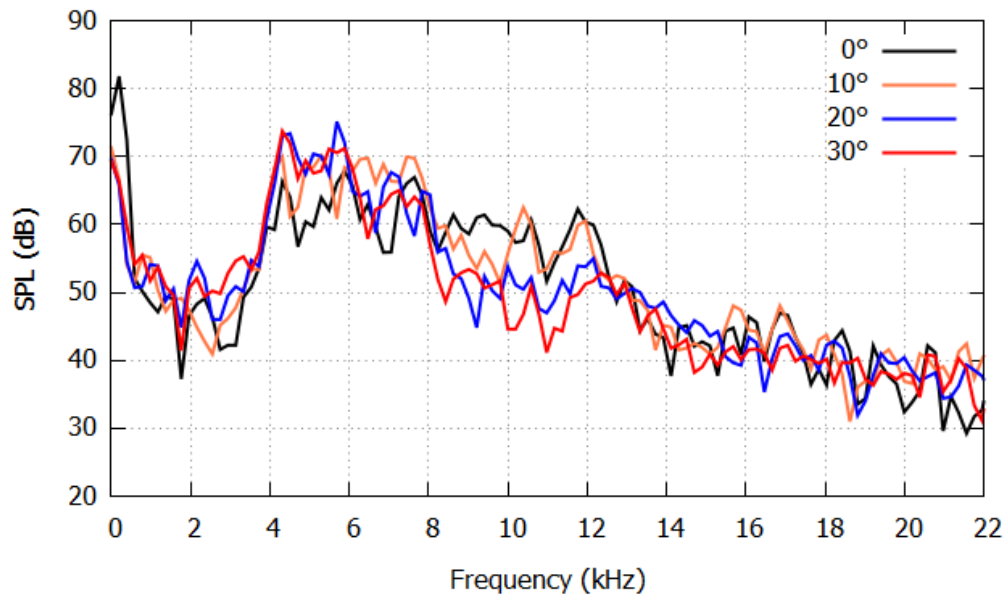


Figure 3-9. The predicted SPL spectra with incisor angles from 0° to 30°.

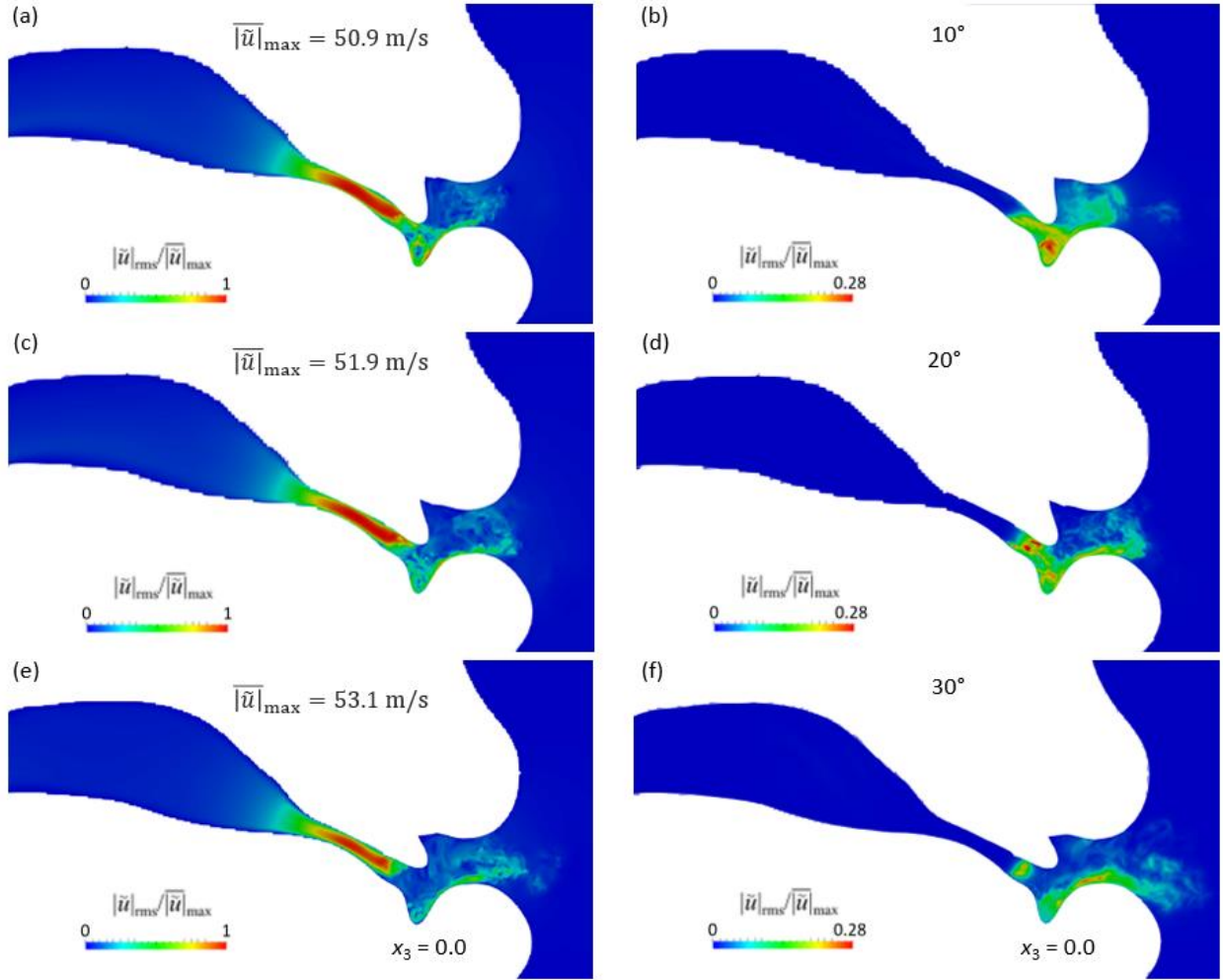


Figure 3-10. (a, c, e) Normalized instantaneous velocity magnitude and (b, d, f) root mean square (RMS) of the velocity fluctuations at $t = 0.015 \text{ s}$ for the (a, b) 10° , (c, d) 20° , and (e, f) 30° models.

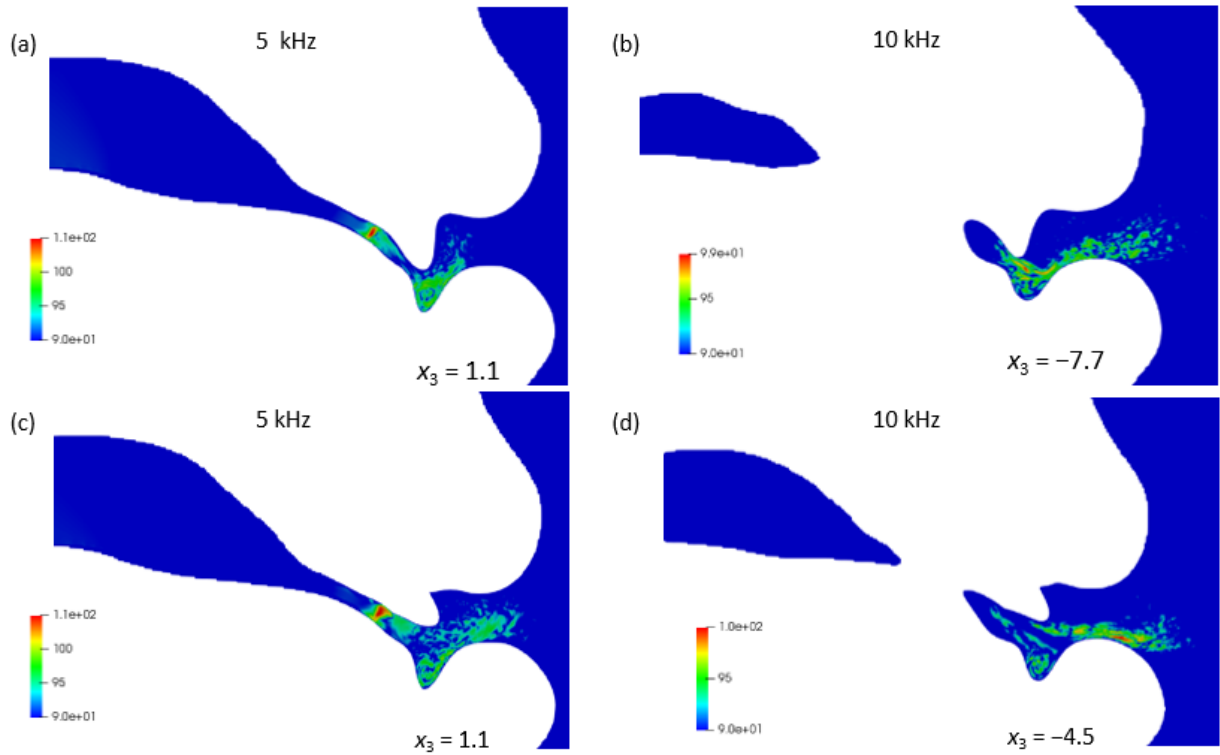


Figure 3-11. Fast Fourier transforms of the velocity fluctuations at (a, c) 5 kHz and (b, d) 10 kHz for the (a, b) 0° and (c, d) 30° models.

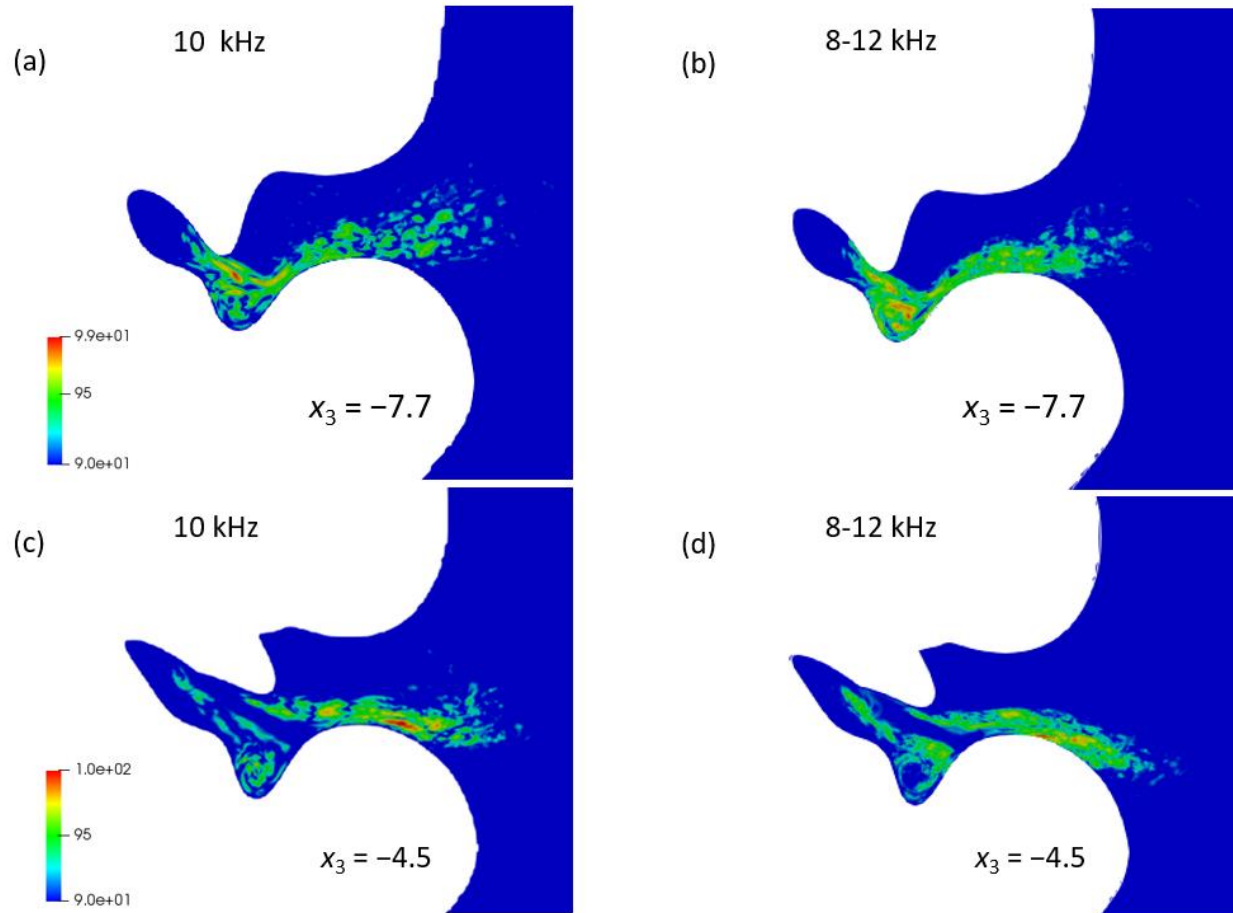


Figure 3-12. Fast Fourier transforms of the velocity fluctuations at (a, c) 10 kHz and (b, d) band-pass filter field of 8-12 kHz for the (a, b) 0° and (c, d) 30° models.

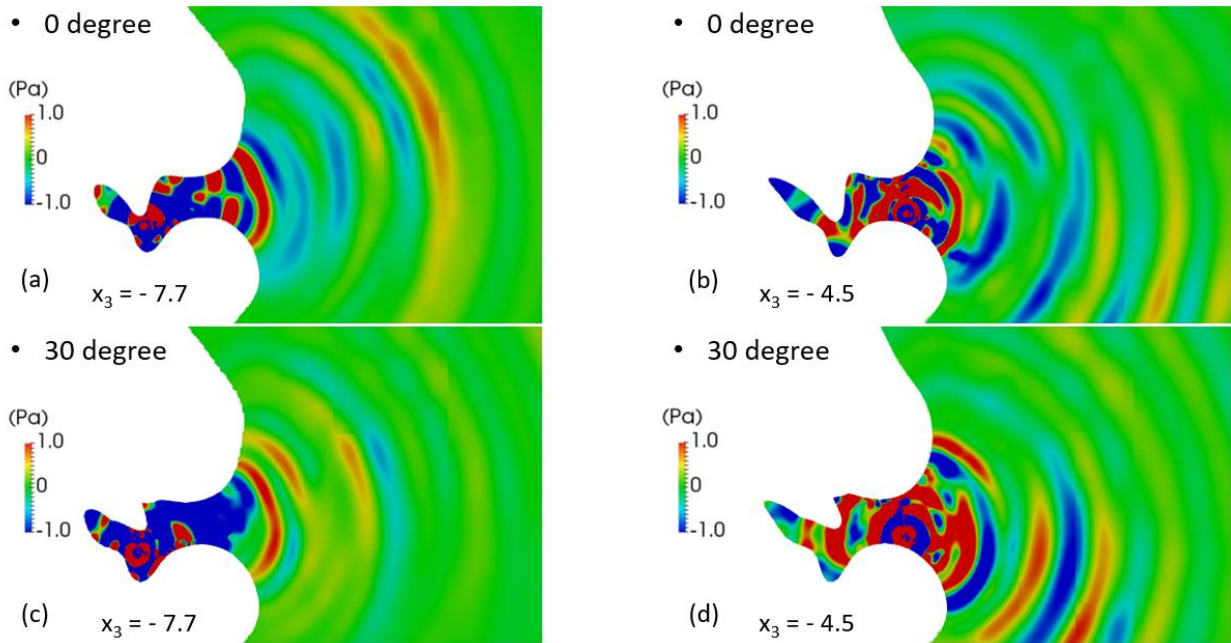


Figure 3-13. Acoustic fields simulated with the monopole acoustic source at (a, c) Point 1 (49.1, -8.4, -7.7) and (b, d) Point 2 (58.5, -7.7, -4.5) for 0° and 30° models, respectively.

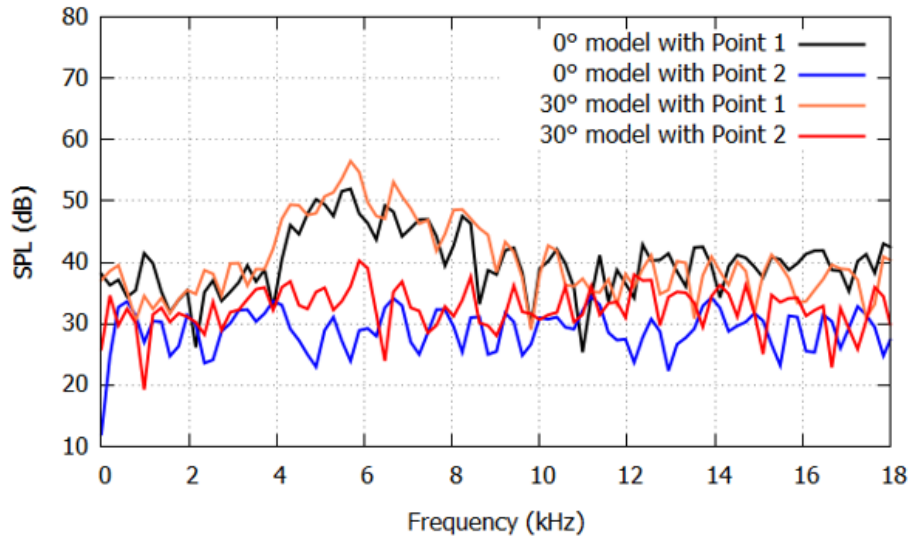


Figure 3-14. SPL spectra for the 0° and 30° models predicted for the monopole acoustic sources at Point 1 (0.049, -0.0084, -0.0077) and Point 2 (0.0585, -0.0077, -0.0045).

Chapter 4. Investigation of effects of tongue articulation and velopharyngeal closure

4.1 Introduction	72
4.2 Vocal tract geometry	72
4.3 Computational conditions	74
4.4 Results and discussion	75
4.5 Conclusion	80

4.1 Introduction

After investigation of the effect on the sustained sibilant /s/ sound with the stationary geometry, the human phonation process of the sibilant /s/ sound were aimed. In this chapter, the numerical simulation of sibilant /s/ production with the realistically moving vocal tract was conducted to investigate the articulation process of velopharyngeal closure and tongue movement.

The sibilant /s/ sound production process was simulated in the part of the Japanese word /usui/ which is extracted after the vowel production of /u/ to the sibilant sound /s/, including the tongue elevation and the velopharyngeal valve closure. Additionally, to investigate the effect of the velum closure speed during the sibilant sound, the simulations with a faster and a slower closing speed of the velopharyngeal valve were conducted. The velopharyngeal valve motion was accelerated and decelerated from the original speed during the phonation process.

By analyzing the spectrogram of the predicted sound, the current study clarifies the effects of tongue movement and velopharyngeal closure speed during the articulation process. Specifically, the phonation process with slower closing speed of the velum enlarges the sound amplitude at co-articulation process, which might confuse the listener.

4.2 Vocal tract model

To simulate the pronouncing process, the geometry of a vocal tract was extracted from four-dimensional computed tomography (4D-CT) images of a subject pronouncing the Japanese word /usui/ as shown in Figure 4-1. The subject was a 42-year-old Japanese male who self-reported no

speech disorders with a normal dentition of angle class I. The CT scan was conducted by a 320-row Area Detector CT (20 fps; 320 slices of 512×512 pixels). The subject pronounced the Japanese word /usui/ (“thin” in English) for 1.6 s while the subject was sitting down. The vocal tract geometry was collected from the scans of the phonation process and extracted from the images based on the brightness values, as shown in Figure 4-2, including the oral cavity, nasal cavity, tongue constriction and velopharyngeal valve. During the simulation, the geometries at each time step were interpolated from these three geometries (Figure 4-2 (d, e, f)) to mimic the realistic motion of the vocal tract.

The geometries used in the current study were the duration after the vowel /u/ (Figure 4-2(a, d)) to the sibilant sound /s/ (Figure 4-2(c, f)), which includes the movement of tongue elevation and the closure of the velopharyngeal valve. In Figure 4-3, the cross-section during the phonation process of the tongue constriction and the velopharyngeal valve are shown at $t_1 = 0.0$ s, $t_2 = 0.05$ s, and $t_3 = 0.10$ s. Besides, the motion and the cross-sectional area of the velopharyngeal valve during the process is shown in Figure 4-4. The figure of the cross-sectional area of the tongue constriction and the velopharyngeal valve during the movement are plotted in Figure 4-5. The preliminary simulation with the stationary geometry was from $t = -0.005$ s to $t = 0$ s to avoid the effects of unstable state on the results of the far-field sound. After $t = 0.0$ s, the cross-sectional areas of the tongue and velum constrictions are reduced from 121 mm^2 and 55.6 mm^2 to 12.9 mm^2 and 0.01 mm^2 , respectively.

In addition, to investigate the effect of the velopharyngeal closure speed, the geometry was changed by modifying the cross-sectional area of the velopharyngeal valve at $t_2 = 0.05$ s, as shown in Figure 4-6. While the original closing speed of velum was 0.208 m/s , a faster velopharyngeal closure with the closing speed of velum equals to 0.220 m/s and a slower velopharyngeal closure

with the closing speed of velum equals to 0.152 m/s were conducted. As shown in Figure 4-7, a faster closing case (red line) and a slower closing case (gray line) were plotted. The initial and the end of the velopharyngeal valve size of all cases are set to be the same, however, the velopharyngeal channel was totally closed at $t = 0.05$ s in the faster case, while the channel was still opening 50% at $t = 0.05$ s in the slower case. The closure speeds of the velum are 0.220 m/s, 0.208 m/s and 0.152 m/s for three cases. In Figure 4-8, the cross-section at $t_1 = 0.0$ s, $t_2 = 0.05$ s, and $t_3 = 0.10$ s during the phonation process of the velopharyngeal valve are shown.

4.3 Computational condition

After testing the space and temporal resolution in the preliminary simulation, the time step was set as 2×10^{-6} and the minimum grid size was chosen to 0.05 mm near the constriction to keep the accuracy of the immersed boundary around the turbulent region. The total grid number was around 1.82×10^8 , and the grid distribution on the mid-sagittal plane distribution (every 16th gridline is shown for clarity) is shown in Figure 4-9. Although the physiologically flow rate varies during the real phonation process, the constant flow rate was set in this simulation to focus on examining the effect of the movement of the speech organs. The uniform inlet velocity was set to 1.5 m/s on the pharynx inlet (the red line shown in Figure 4-9), which resulted in a physiological flowrate of 450 cm³/s. Since the sibilant /s/ sound will be generated by the geometry of the vocal tract [62], therefore, there is no voice source set in this simulation. The detailed parameters were shown in Table 4-1.

To keep the flow in the computational domain from being polluted by reflecting pressure waves, an absorbing boundary condition was used as the outlet condition. The three-view diagram of the overall computational domain is shown in Figure 4-10. Fast Fourier transform (FFT) using the Hann window was applied to the pressure waveforms sampled at 80 mm from the lip outlet to analyze the far-field sound spectrogram. The FFT sampling frequency was 50 kHz with 256 points. The overlap ratio was 0.9 for the spectrogram calculation.

4.4 Results and discussion

The instantaneous flow field and the pressure fluctuation along the sagittal plane from /u/ to /s/ are shown in Figure 4-11 for the CT scans with the original velum speed of 0.208 m/s. At $t = 0.01$ s (Figure 4-11(a) and 4-11(d)), the tongue was at the lower position so that the size of the tongue constriction was not small enough to produce the sound of /s/. In addition, the velopharyngeal valve was opened which let the airflow go to both nasal and oral cavities. Specifically, there was 35% airflow go through the velopharyngeal valve and 65% airflow goes to the oral cavity. Therefore, the velocity of the mainstream at the oral cavity was only around 10 m/s, and there was no sound propagation from both nasal and oral cavities. At $t = 0.05$ s (Figure 4-11(b) and 4-11(e)), since the velopharyngeal valve kept closing the channel to nasal cavity, the higher flow resistance decreased the flow rate from 35% to 13% in the nasal cavity and increased the flow rate from 65% to 87% in the oral cavity. Moreover, the elevated tongue position caused the reduction of the cross-sectional area of the oral cavity so that the velocity of the mainstream flow in the oral cavity was accelerated to around 20 m/s and the periodic pressure wave was

observed near the lips. Besides, because of the narrow velopharyngeal valve, the velocity became turbulent at the nasal cavity which resulted in the pressure wave at the exit of the nose in Figure 4-11(e). At $t = 0.10$ s (Figure 4-11(c) and 4-11(f)), the tip of the tongue moved posterior direction and elevated towards the hard palate. The velocity of the mainstream in the oral cavity reached around 25 m/s and the amplitude of the sound was increased at the far-field. At this time, since the velopharyngeal valve was completely closed, there was no nasal sound emission. Consequently, the generation of /s/ sound mainly depended on the movement of tongue constriction, and the velopharyngeal valve also involved the far-field acoustics behavior.

The spectrogram of the propagating sound collected at 80 mm from the lip outlet, is shown in Figure 4-12. For $t = 0.01$ s to $t = 0.03$ s, since the cross-sectional area of the oral tract was decreased with the tongue elevation and velopharynx closure, the amplitude peak at 6 kHz gradually appeared after $t = 0.02$ s. The amplitude of the broadband noise around 5–9 kHz was observed with relatively small amplitudes since the velopharyngeal valve was still open and some of the airflows kept going through the nasal cavity. For $t = 0.03$ s to $t = 0.07$ s, because of the closing velopharyngeal valve, the broadband noise was amplified in the extended frequency range of 3–11 kHz. After $t = 0.07$ s, when the size of the velopharyngeal valve became smaller than that of the tongue constriction, the airflow tended to go to the oral cavity rather than the nasal cavity. Since the tongue constriction at this moment was narrow enough for the production of the sibilant sound, the typical sibilant sound was generated with the characteristics peak around 4–6k Hz, and the broadband noise was observed around 4–12k Hz. This spectrogram is similar to the sound pronounced by the subject [62]. Furthermore, the relation between the size of the velopharyngeal valve and the sound amplitude observed at the far-field was consistent with the previous studies [24] [25] [31] [62].

To investigate the effect of the closing speed of the velopharyngeal valve, simulations with closing speeds of 0.220 m/s, 0.208 m/s and 0.152 m/s were conducted with the cross-sectional area ratio shown in Figure 4-7. As the spectrograms with the different closing speeds shown in Figure 4-13, the characteristic peak of the sibilant /s/ sound around 4–6 kHz was observed in all cases after $t = 0.09$ s. In the period of 0.01 s to 0.09 s, the frequency range of the broadband noise of /s/ was extended from 6 kHz to 3–7 kHz during the phonation process due to the tongue elevation in all cases.

The overall sound pressure level (OASPL) values from 1.5–20 kHz during the phonation process are shown in Figure 4-14. The OASPL of the 0.220 m/s was larger than the other cases around $t = 0.01$ s, because of the higher flow rate through the oral cavity caused by the higher flow resistance in the nasal cavity. However, the OASPL of 0.152 m/s became larger than that of 0.220 m/s after $t = 0.04$ s. The amplitude of 0.220 m/s at this moment became the lowest since the velopharyngeal valve was totally closed. On the contrary, because the channel to the nasal cavity was still 50% opened for 0.152 m/s, the audible nasal emission made the OASPL higher at $t = 0.05$ s. Besides, the OASPL value based on the cross-sectional area of tongue constriction is shown in Figure 4-15. The OASPL amplitude tendencies in all cases were relatively flat with the larger tongue constriction and were increasing rapidly while the tongue constriction was less than 25 mm². It indicated that rather than the closing speed of the velopharyngeal valve, the size of the tongue constriction contributes more to the amplitude of the sibilant /s/ sound. Figure 4-16 shows the maximum velocity in the velopharyngeal channel. Since the timing of complete closure was earlier in the case with velum speed of 0.220 m/s, the velocity reduced to 0 around $t = 0.05$ s. On the other hand, because of the slower closing speed of the velopharyngeal valve, the velocity reached the highest value later in the 0.152 m/s velum speed case.

The area ratio of the velopharyngeal opening to the oral constriction is shown in Figure 4-17; the flow rate ratio of the velopharyngeal channel to the vocal tract is shown in Figure 4-18. According to Sundström [31], the most severe nasal emission was produced while the size of the velopharyngeal opening was equal to the size of the oral constriction. This phenomenon also can be observed in the current simulation. The flow rate at the velopharyngeal channel and the nasal channel of the slower closing case were highest in all cases at the $t = 0.05$ s, which potentially results in more severe nasal emission., as shown in Figure 4-19 and Figure 4-20.

To clarify the relationship between the sound amplitude and velum closing speed, the spectrograms of the different closing speeds are shown in Figure 4-21 based on the velopharyngeal valve opening. The horizontal axis is plotted with the velum opening ratio, and the complete closure (0% opening) was around $t = 0.05$ s for the case with 0.220 m/s, $t = 0.09$ s for the original case, and $t = 0.10$ s for the case with 0.152 m/s. As seen in the figure, at 80% valve opening ratio, peaks were mainly observed around 6 kHz in all cases. While the valve opening ratio reached 20%, the frequency range of the broadband noise was increased to 3–10 kHz in all cases. Although the amplitudes of the three cases were different, the tendency of the broadband noise indicated that the sound is strongly related to the opening of the velopharyngeal valve. Besides, the amplitudes of 0.152 m/s in the frequency range 3 to 7 kHz were larger than those of 0.208 m/s and 0.220 m/s.

To identify the effect of the velum closure speed, the OASPL based on the velopharyngeal valve opening ratio was shown in Figure 4-22. As shown in the figure, the amplitude was amplified for 0.152 m/s in the opening ratio from 75% to 5%. The mean OASPL of the 0.220 m/s, 0.208 m/s, and 0.152 m/s are 61.8 dB, 62.1 dB, and 64.2 dB, respectively, during the phonation process ($t = 0$ to 0.1 s). At the velum opening of 55%, the OASPL of 0.152 m/s was 5 dB louder than that of

0.220 m/s because the area ratio of velopharyngeal to oral channel was larger in slower closing case.

Moreover, the relation between the mean OASPLs and the velum closure speed was plotted in Figure 4-23. The relation was in the inverse proportion with the slope equaled to $-35.3 \text{ dB}\cdot\text{s/m}$. This indicated that the sound generated before the /s/ will be enlarged while the slower closing speed of the velopharyngeal valve. This amplified sound might be related to area ratio of velopharyngeal to oral constriction mentioned by Sundström [31]. Consequently, the slower velum closure speed will result in the larger sound amplitude during the co-articulation and might disturb the acoustic characteristics of pronounce sounds. Although we performed the simulation only for this subject, further simulation with additional patients with hypernasality may help the diagnosis of the velopharyngeal insufficiency.

4.5 Conclusion

To investigate the effect of the velopharyngeal closure speed and the tongue constriction formation on the sibilant sound production, a numerical simulation of the phonation process after the vowel /u/ to the sibilant sound /s/ was conducted in this chapter. The broadband noise generation during the co-articulation process was observed and consistent with the measurement result. To clarify the effect of velopharyngeal closure speed, the simulations with different velum closing speed were conducted. The spectrograms based on the velopharyngeal valve opening showed that the frequency range of the broadband noise of /s/ was extended from 6 kHz to 3–7 kHz during the co-articulation process in all cases, and this indicated that the sound amplitude is

strongly related to the velopharyngeal opening. In addition, the sound amplitude was increased in the case with the slow velum closure during the co-articulation process, indicating the possible phonation of indistinguishable sound.

Table 4-1. Computational parameters of the moving process case.

Size of the largest cell (mm)	1.6
Size of the smallest cell (mm)	0.05
Computational domain (mm ³)	$409.6 \times 204.8 \times 204.8$
Total number of cells N_{cell}	182,394,880
Inlet velocity (m/s)	1.5
Time step (s)	2×10^{-6}
Physical time (s)	0.12

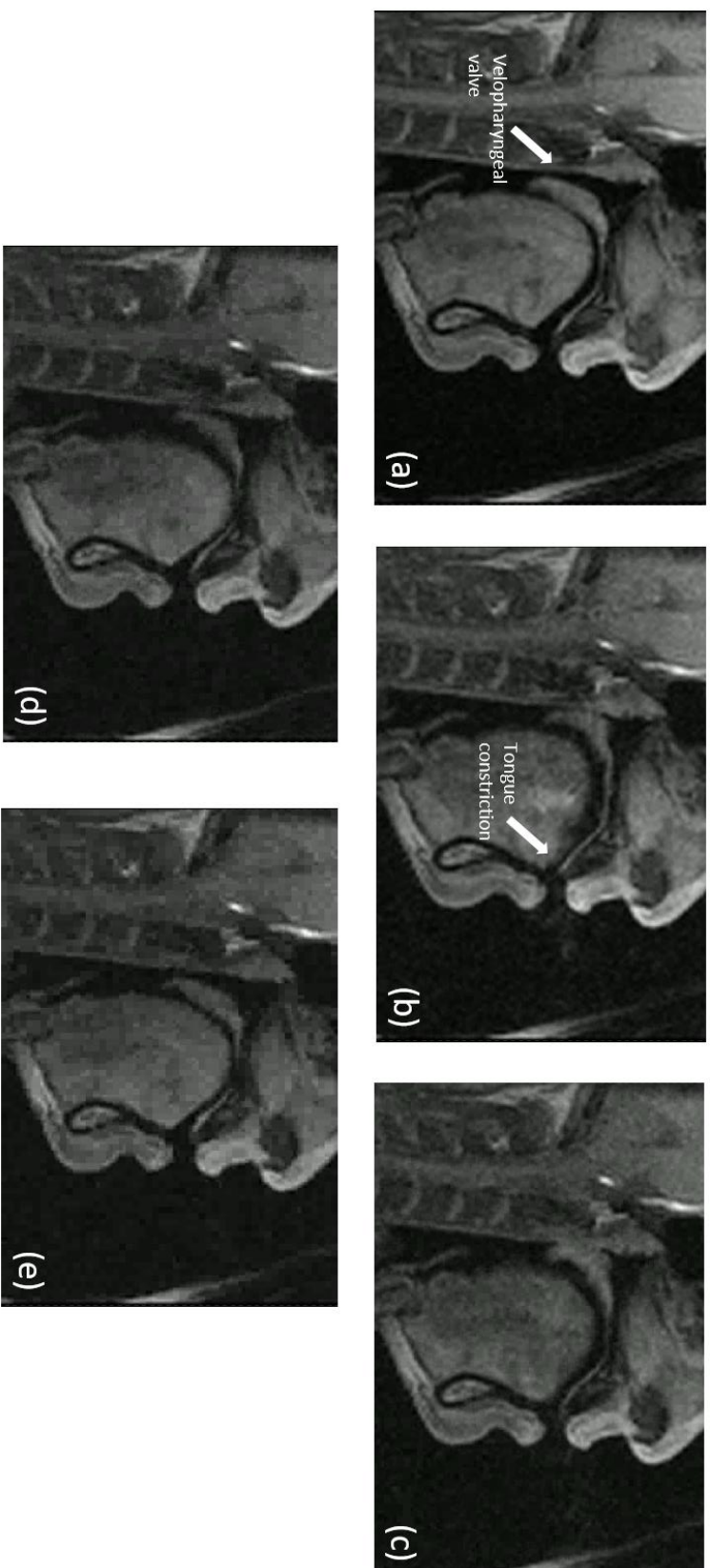


Figure 4-1. MRI scan images of the phonation process of Japanese word /usui/.

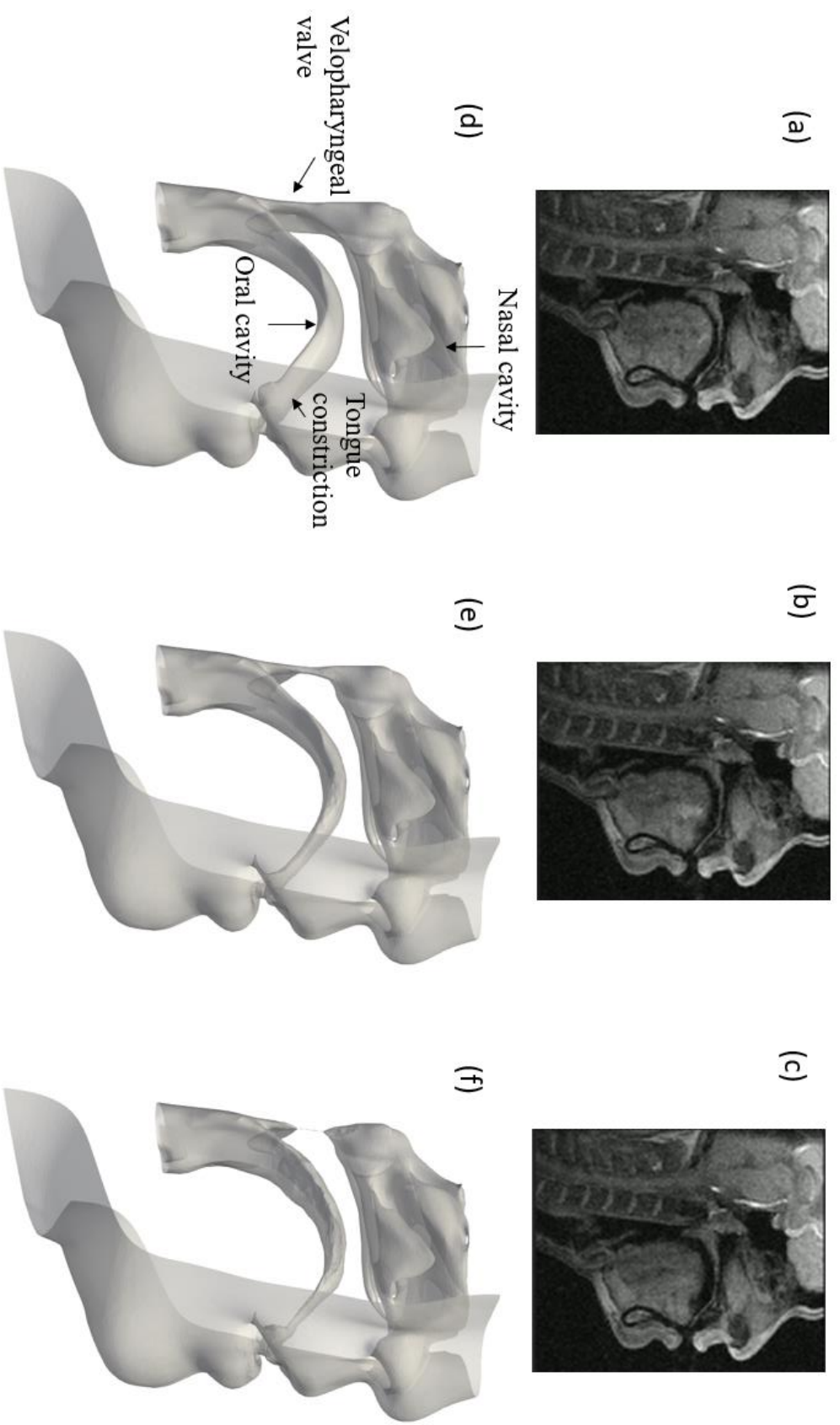


Figure 4-2. MRI images and vocal tract geometry of the (a) vowel /u/ to (c) the sibilant sound /s/ in the pronouncing process of /usu/.

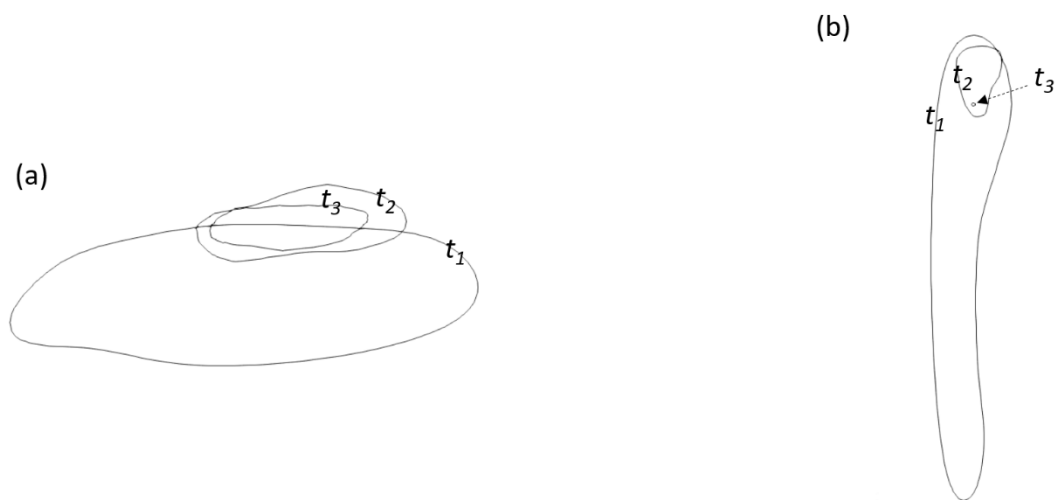


Figure 4-3. The cross-section during the phonation process of (a) the tongue constriction and (b) the velopharyngeal valve.

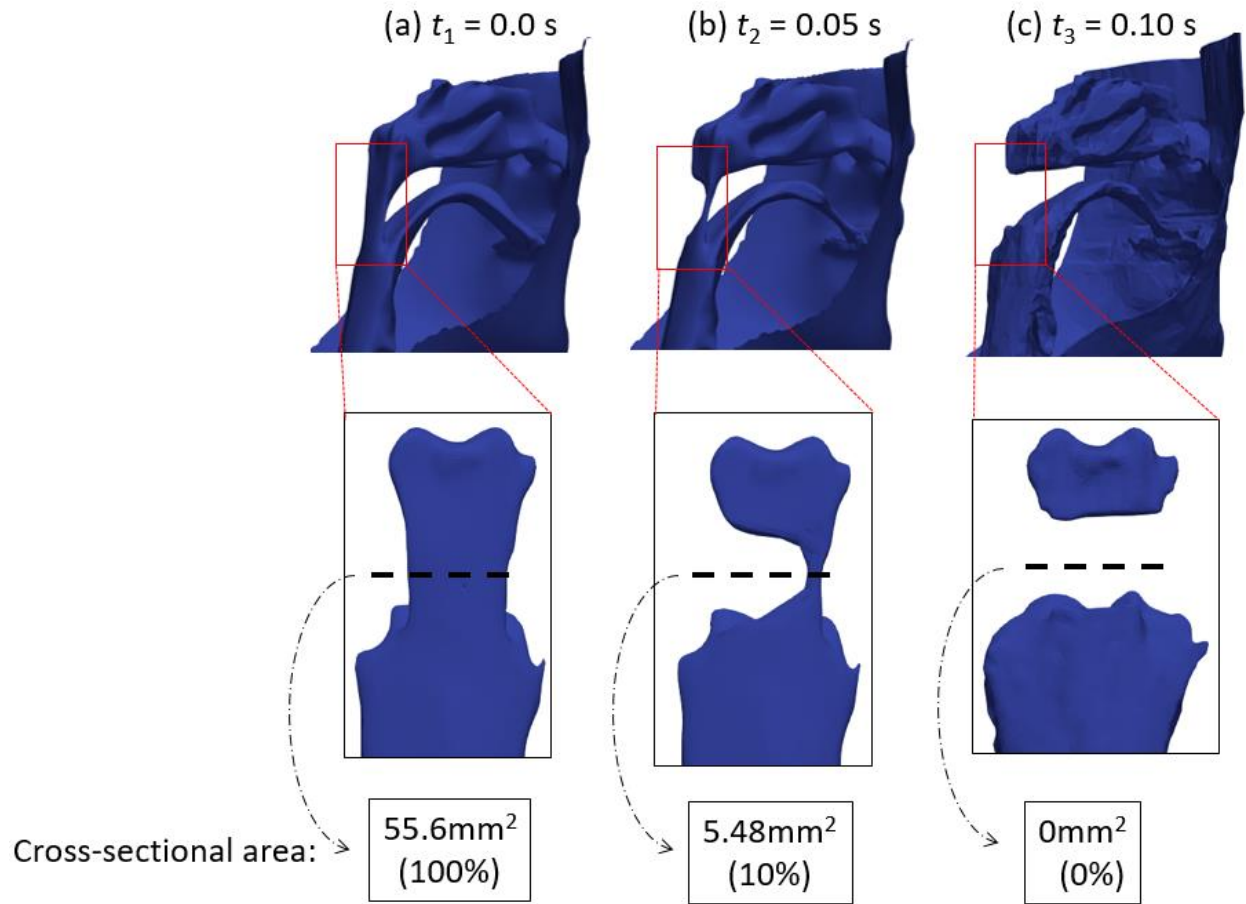


Figure 4-4. The motion and the cross-sectional area of the velopharyngeal valve at (a) $t_1 = 0.0 \text{ s}$, (b) $t_2 = 0.05 \text{ s}$ and (c) $t_3 = 0.10 \text{ s}$.

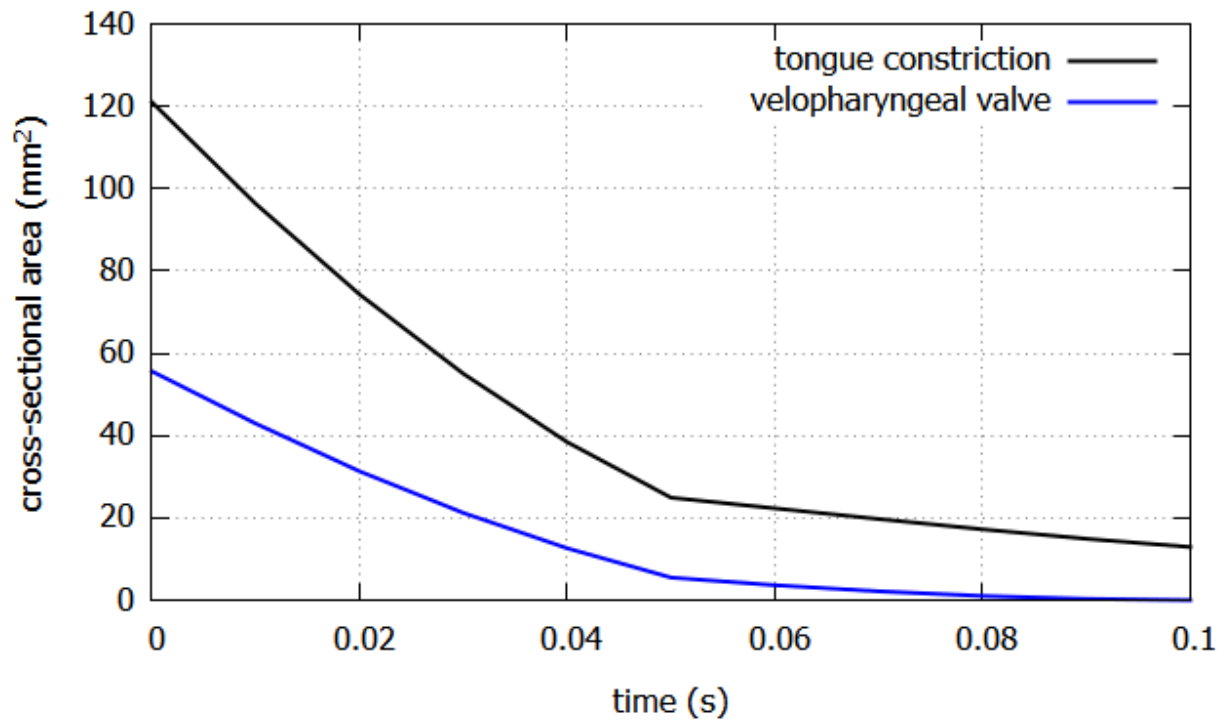


Figure 4-5. The cross-sectional area of the tongue constriction and the velopharyngeal valve during the movement.

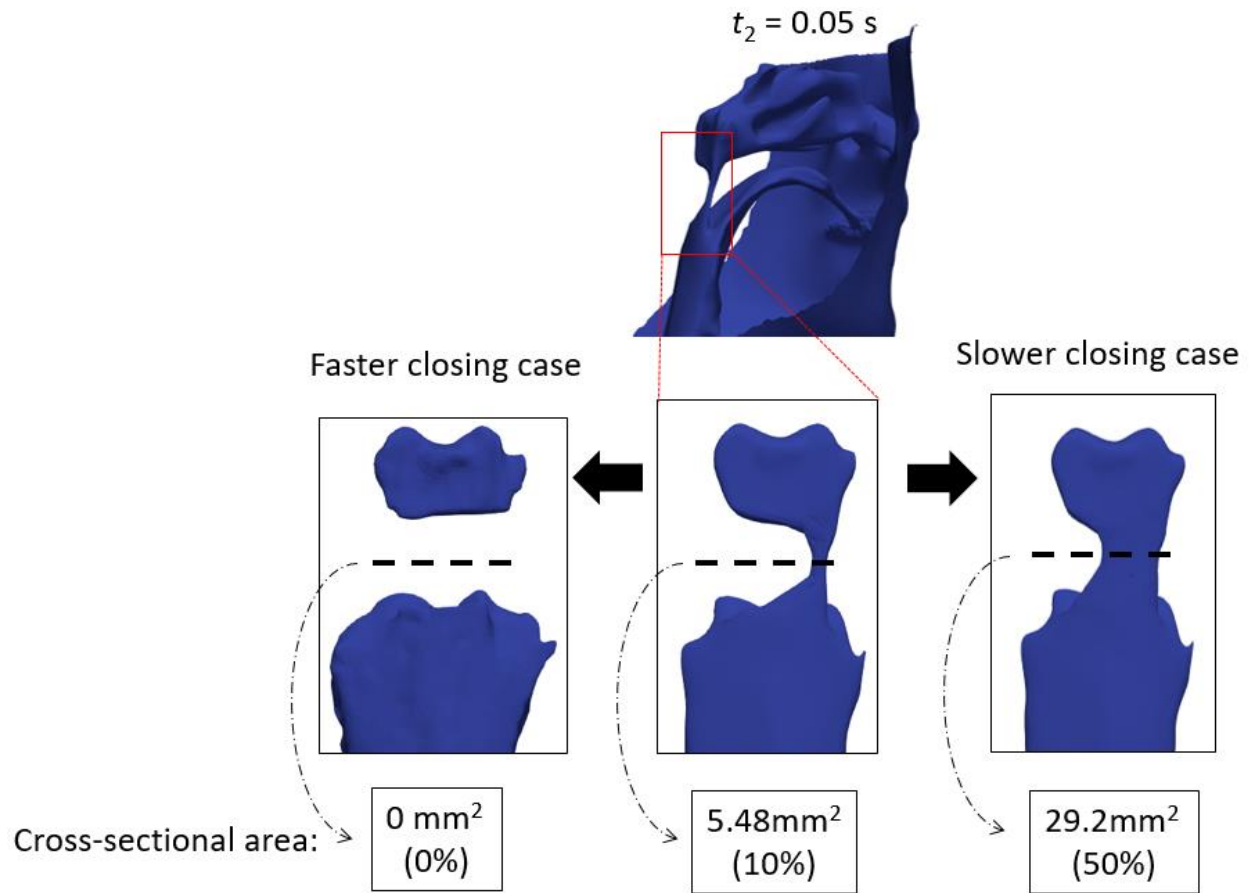


Figure 4-6. The modified geometries for velum speed investigating at $t_2 = 0.05$ s.

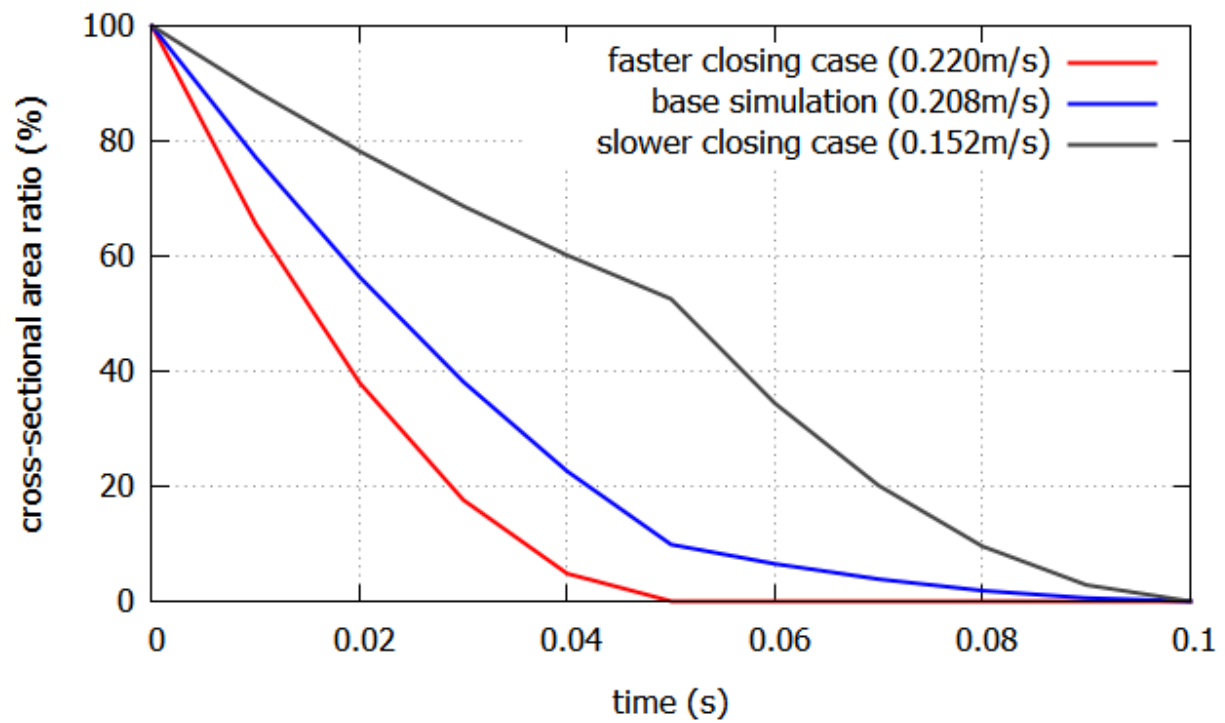
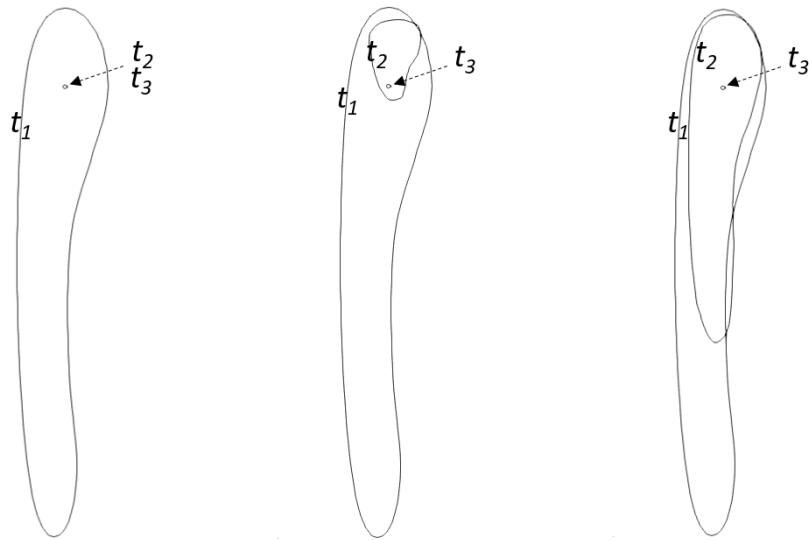


Figure 4-7. The difference of the cross-sectional area ratio of the different velum closing speed.



(a) Faster closing speed. (b) Original speed. (c) Slower closing speed.

Figure 4-8. The movement of the velopharyngeal valve with different closing speed.

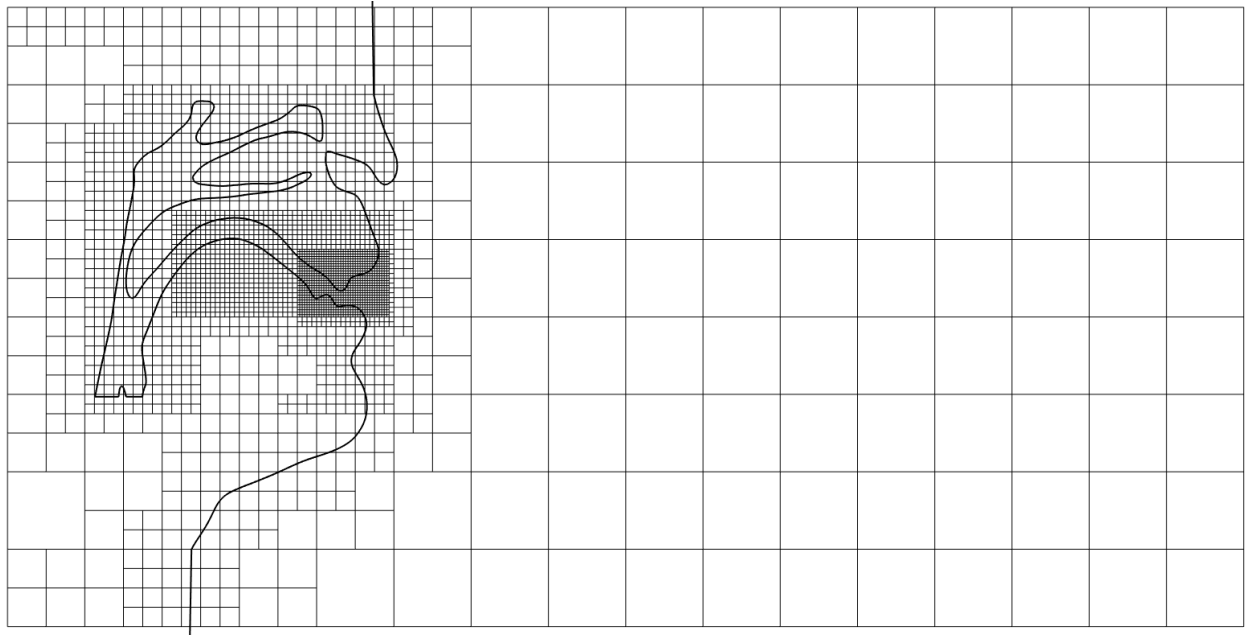


Figure 4-9. The grid distribution (every 16th gridline is shown for clarity).

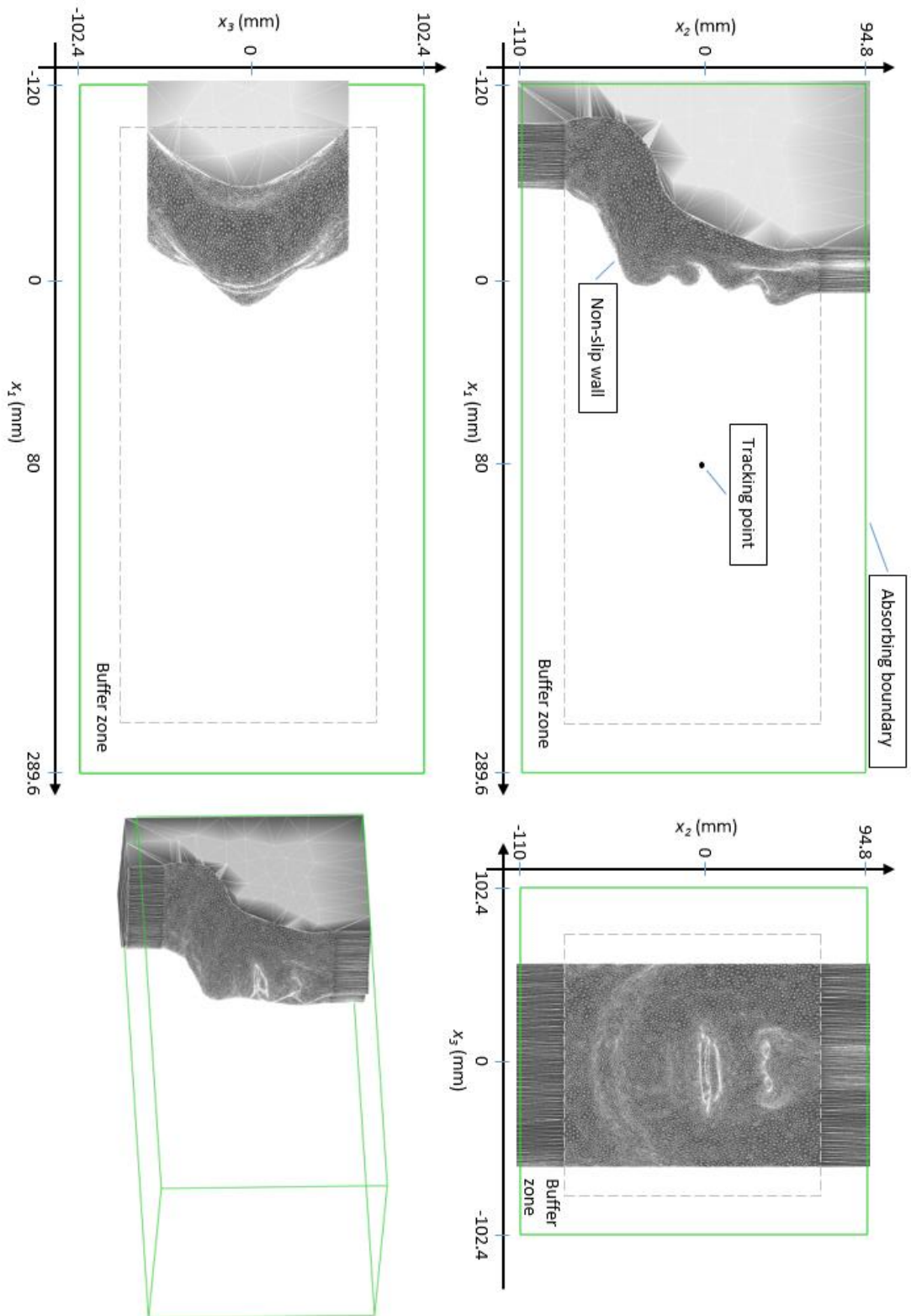


Figure 4-10. The three-view diagram of overall computational domain.

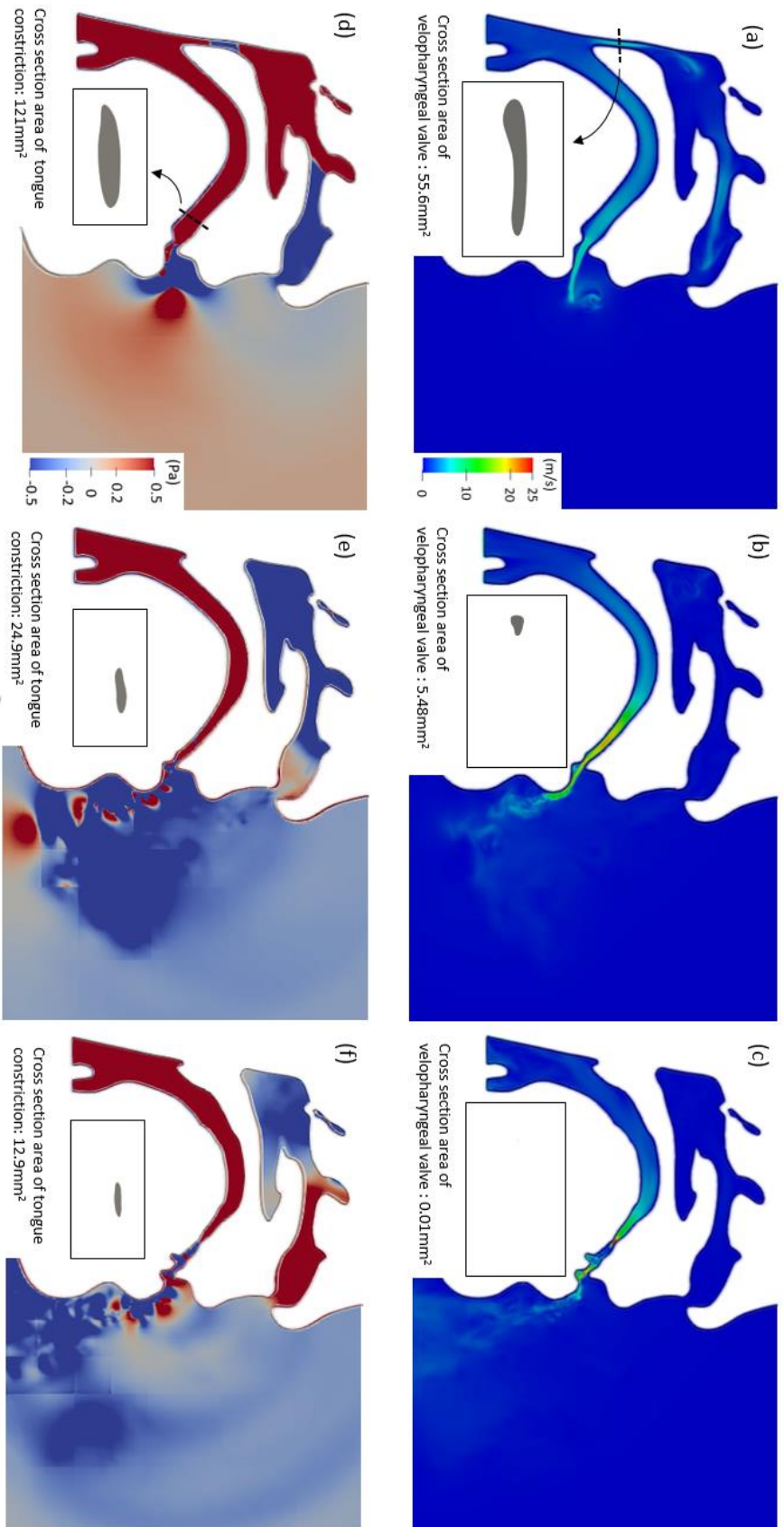


Figure 4-11. (a, b, c) Instantaneous flow field and (d, e, f) pressure fluctuation along the sagittal plane at (a, d) $t = 0.01$ s, (b, e) $t = 0.05$ s and (c, f) $t = 0.10$ s.

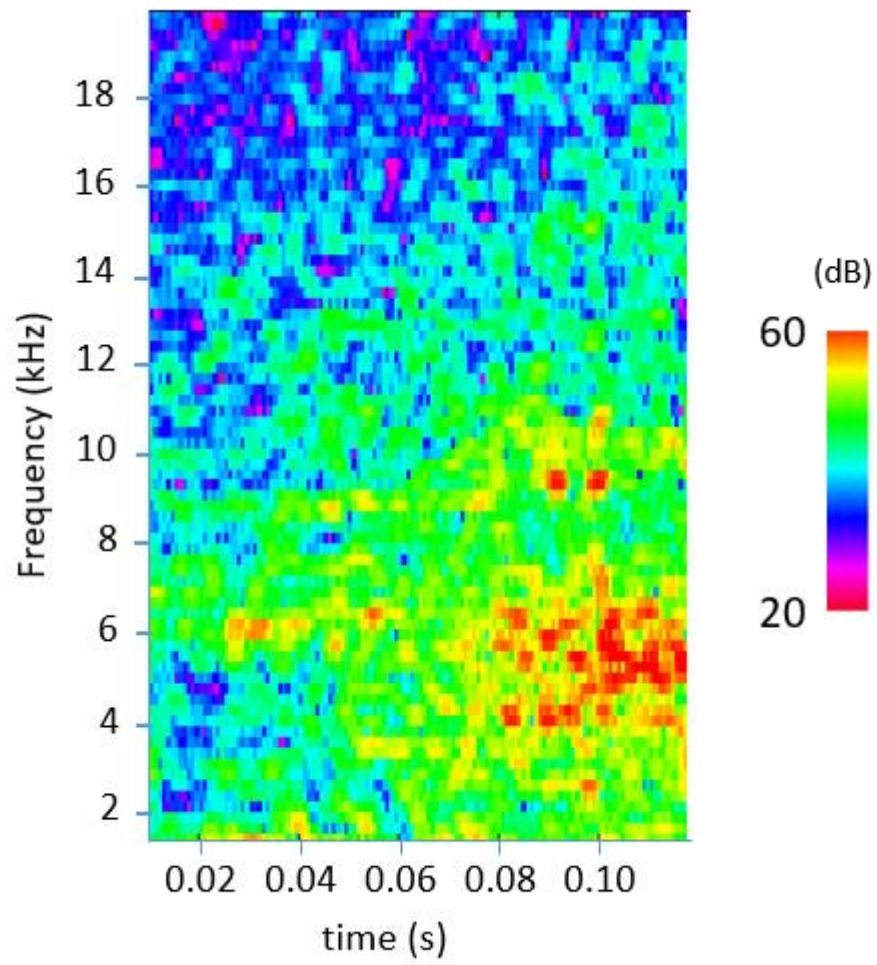


Figure 4-12. The spectrogram of the propagating sound collected 80 mm from the lip outlet.

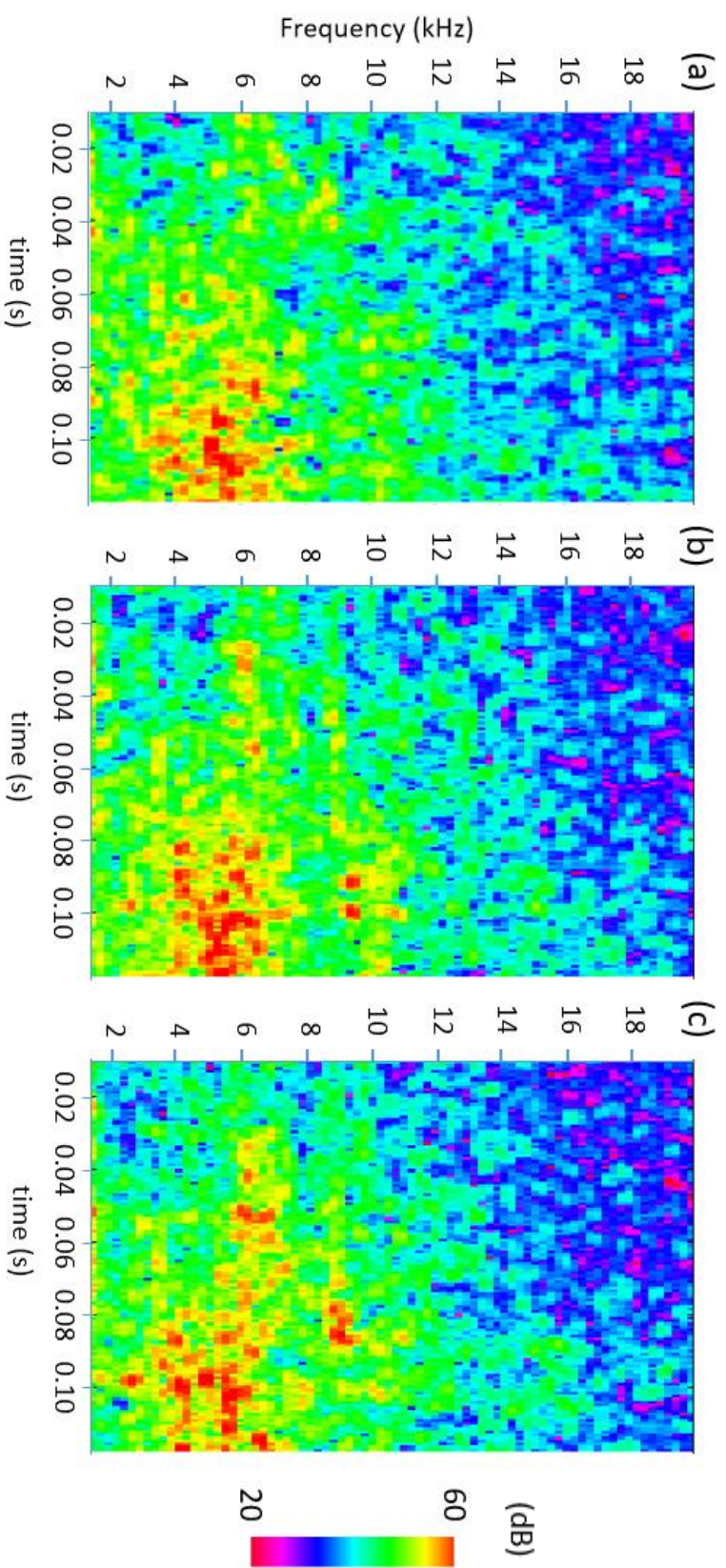


Figure 4-13. The spectrogram at the far-field of the (a) faster closing case (0.220 m/s), (b) base simulation (0.208 m/s) and (c) slower closing case (0.152 m/s).

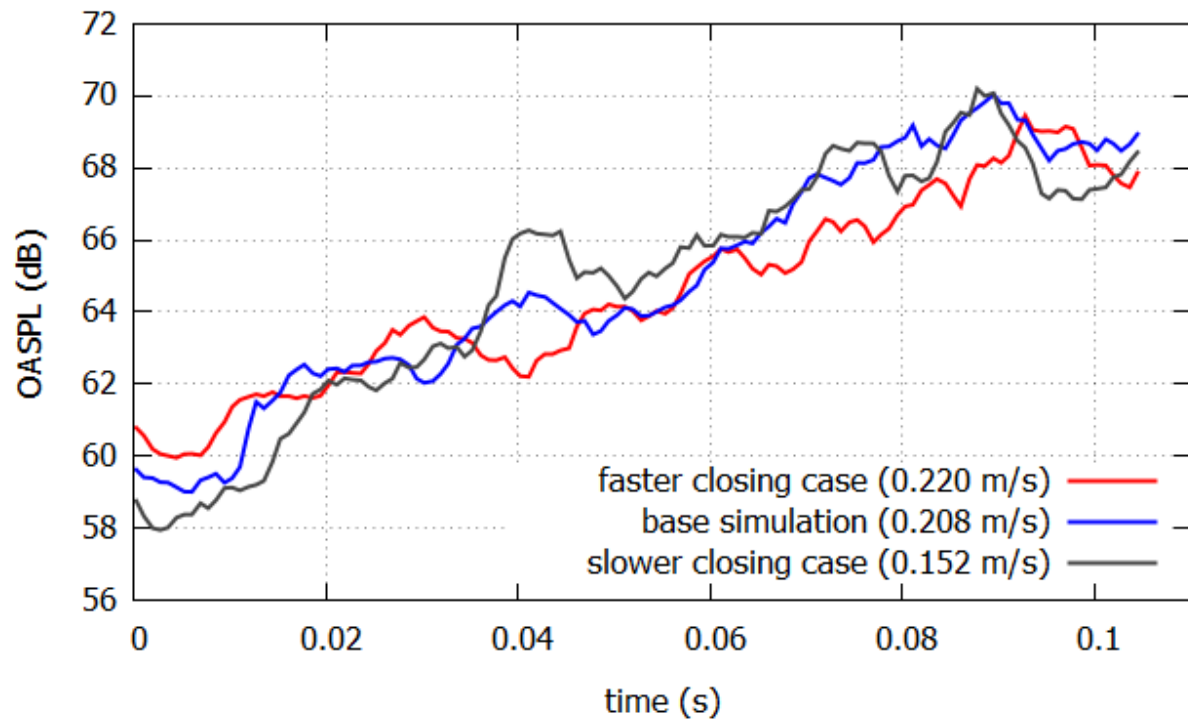


Figure 4-14. The OASPL value of the different velum closing speed cases.

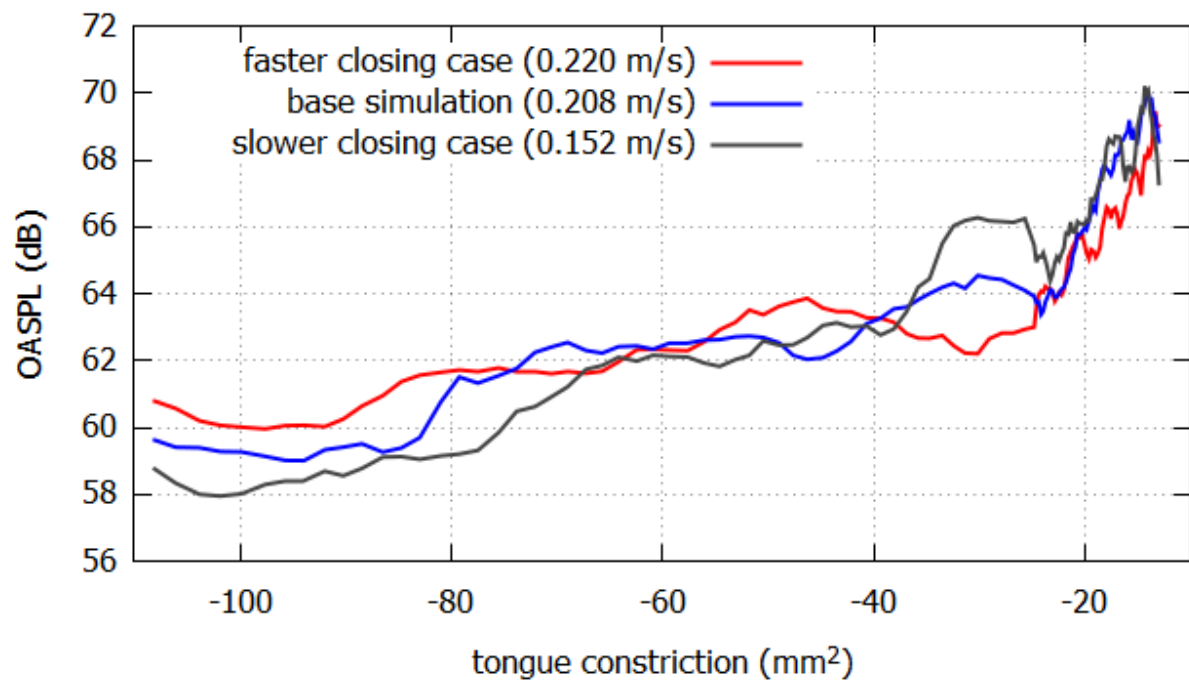


Figure 4-15. The OASPL of the different velum closing speed based on the tongue constriction.

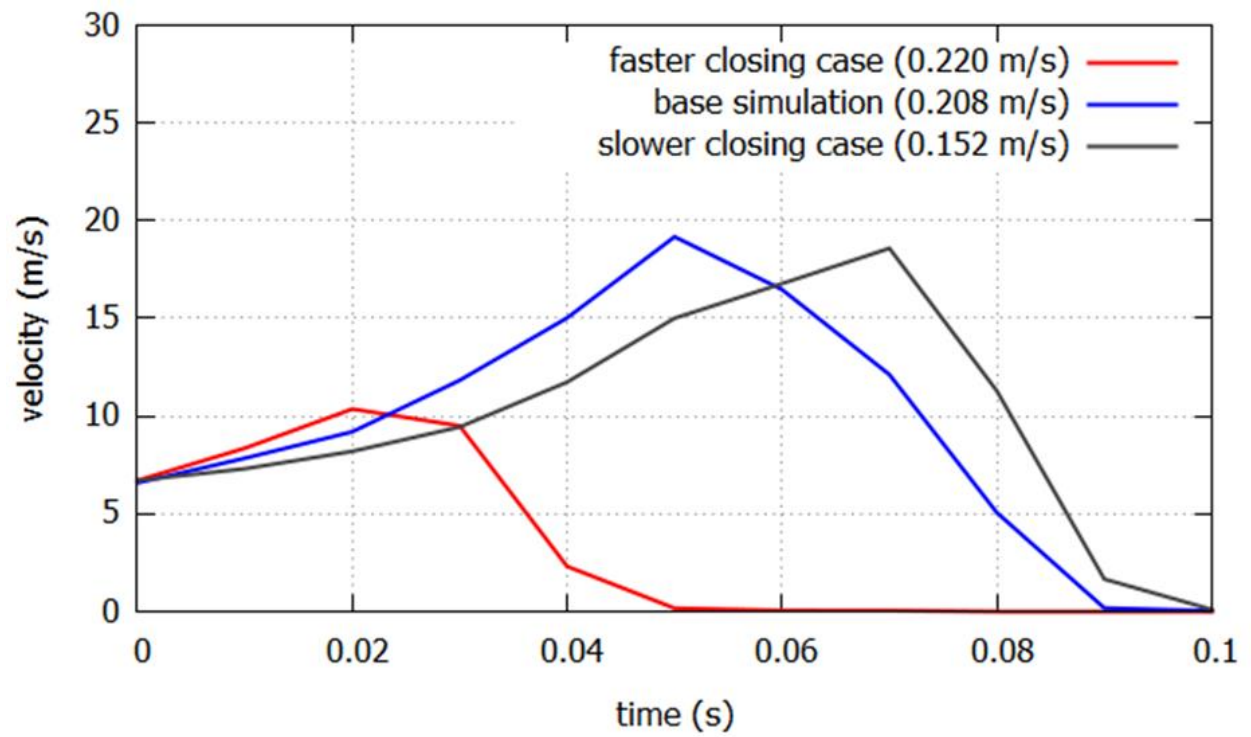


Figure 4-16. The maximum velocity in the velopharyngeal channel for different closing speed of velopharyngeal valve.

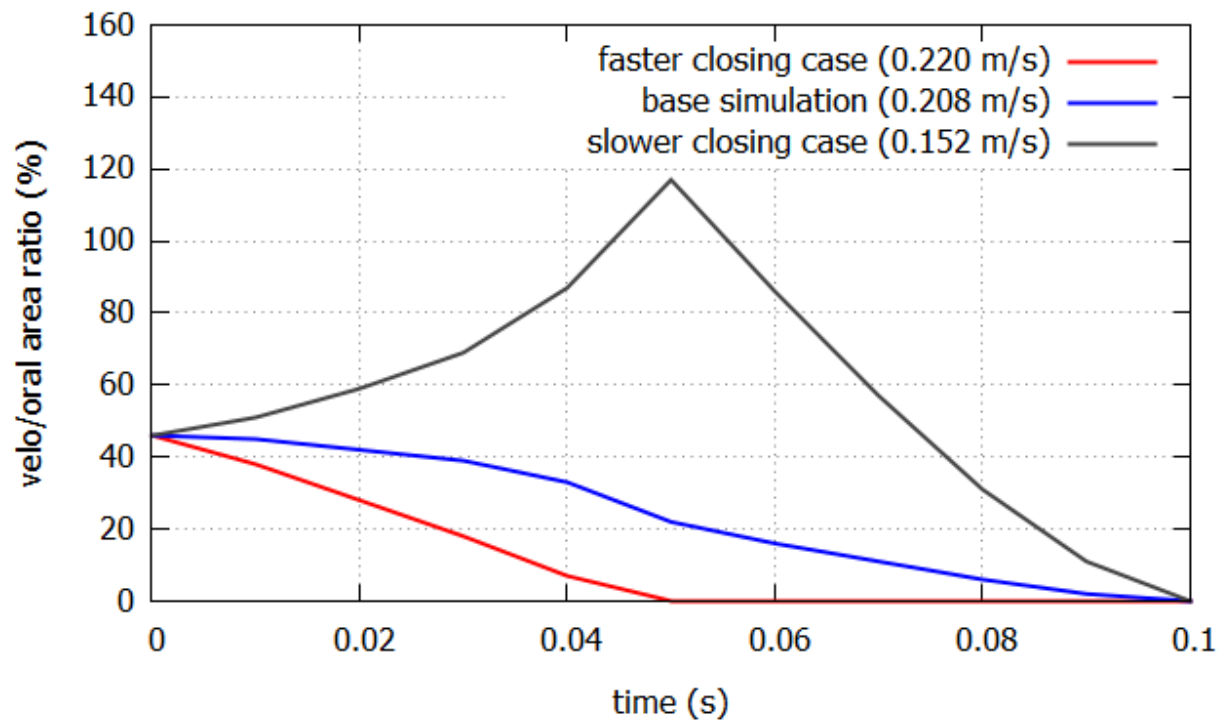


Figure 4-17. The area ratio of the velopharyngeal channel to the oral channel.

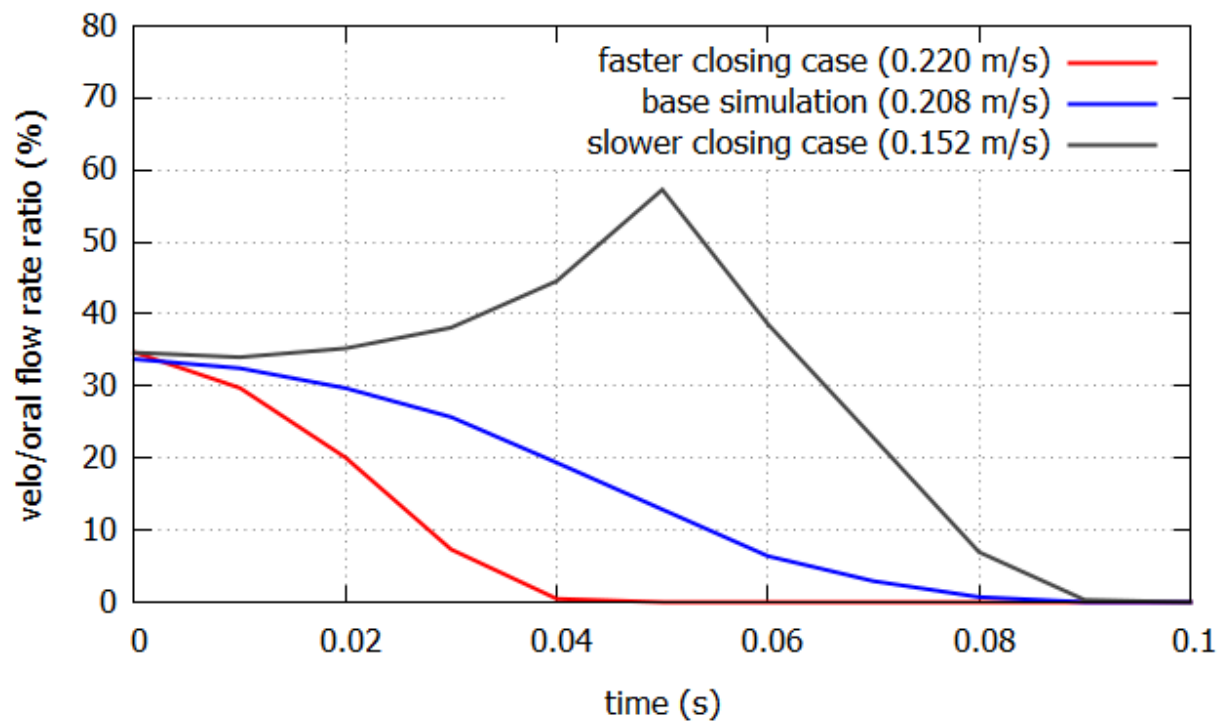


Figure 4-18. The flow rate ratio of the velopharyngeal channel to the oral channel.

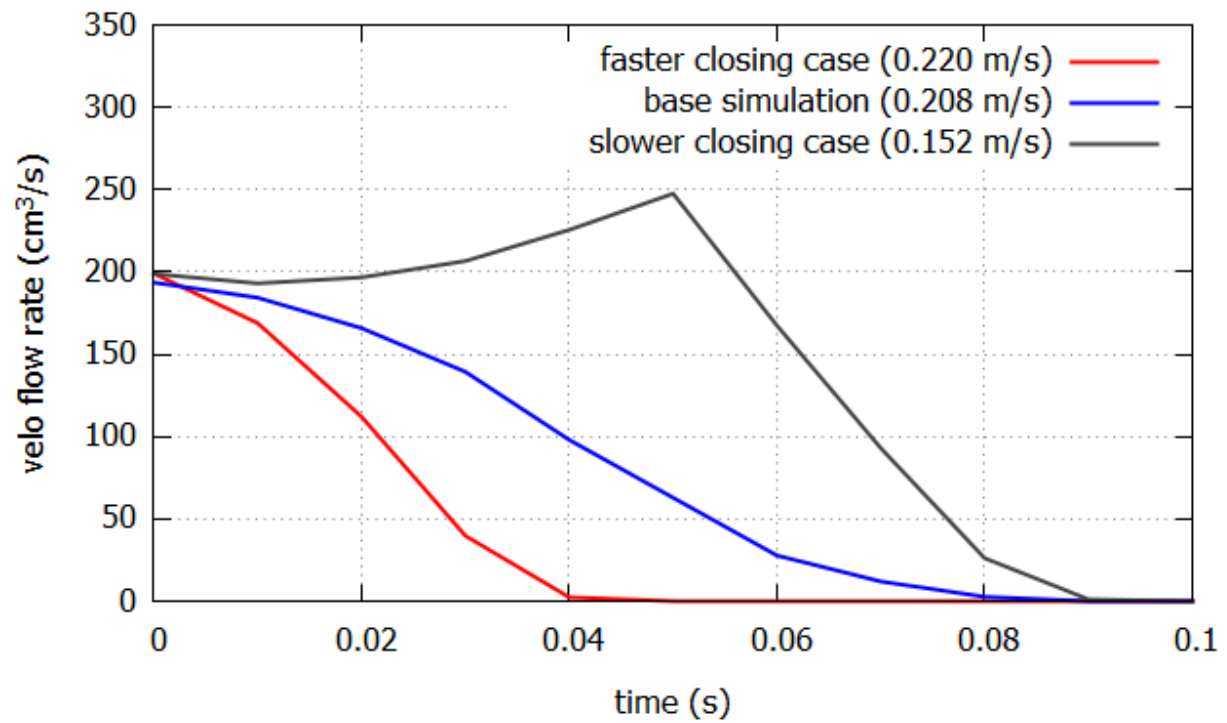


Figure 4-19. The flow rate at the velopharyngeal channel.

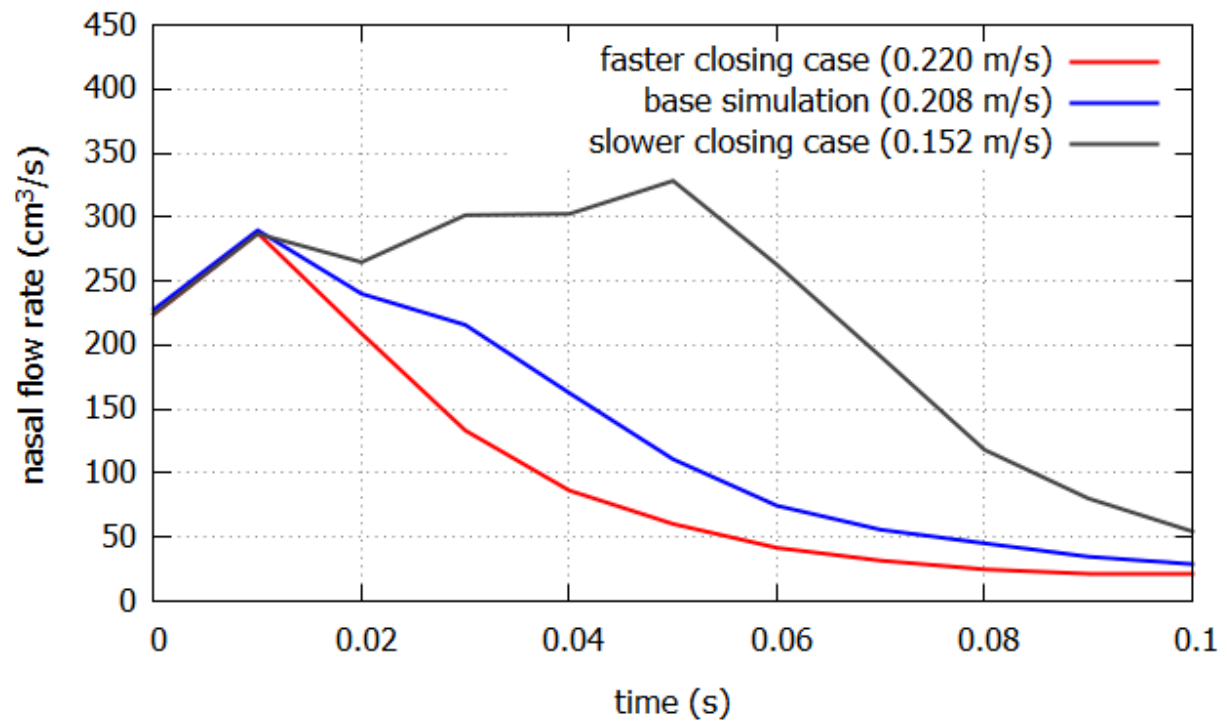


Figure 4-20. The flow rate at the nasal channel.

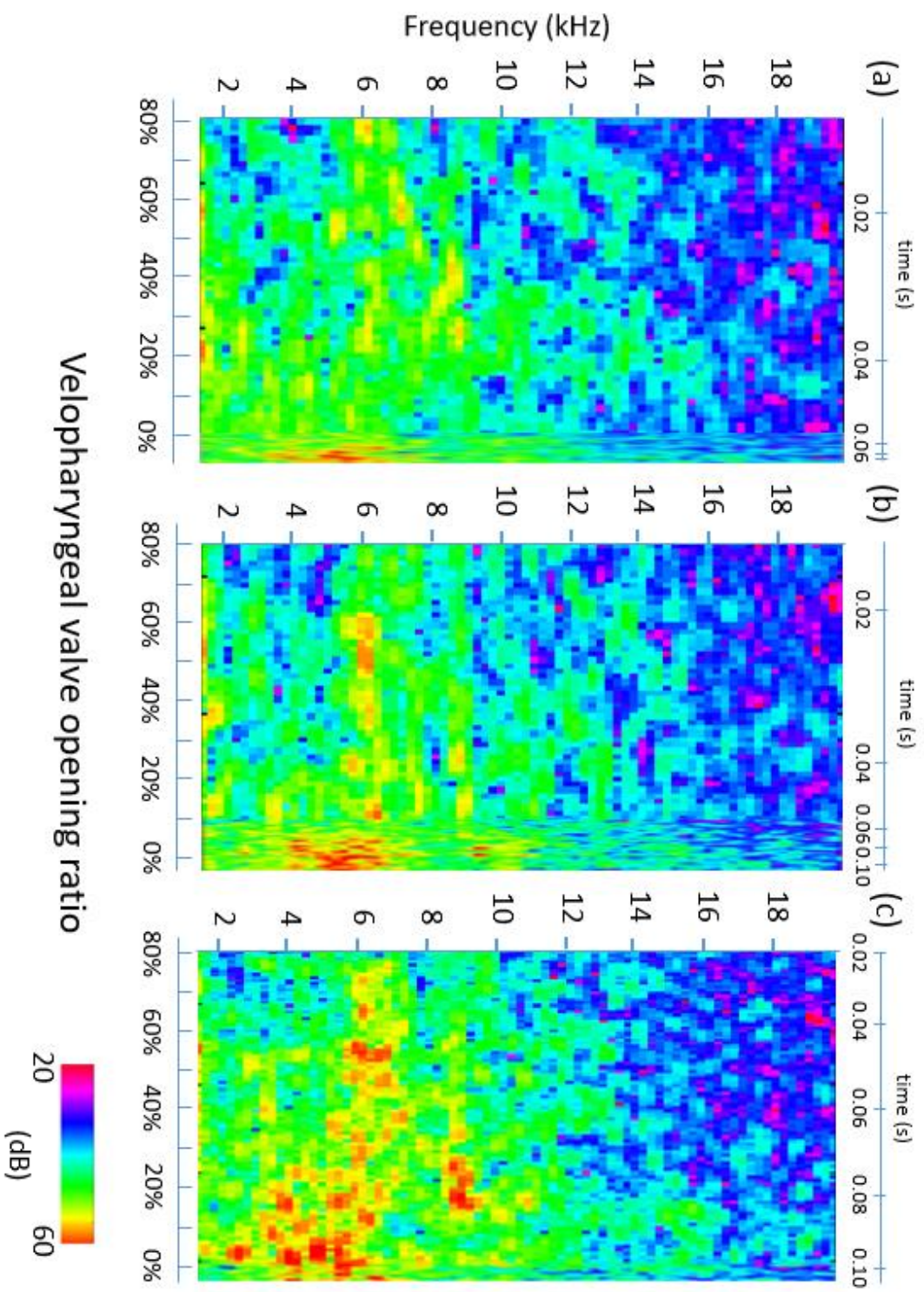


Figure 4-21. The far-field spectrogram of the (a) faster closing case (0.220 m/s), (b) base simulation (0.208 m/s) and (c) slower closing case (0.152 m/s) based on the velopharyngeal valve opening.

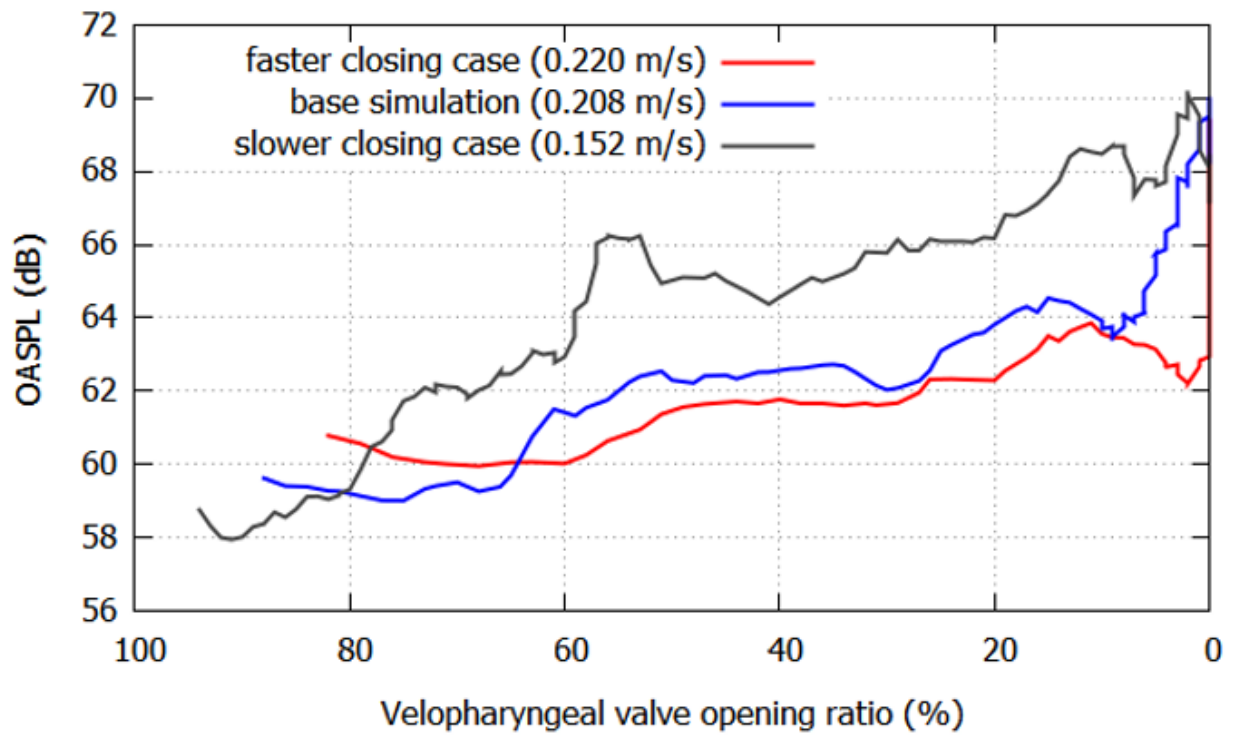


Figure 4-22. The OASPL value of the different closing speed cases based on the velopharyngeal valve opening ratio.

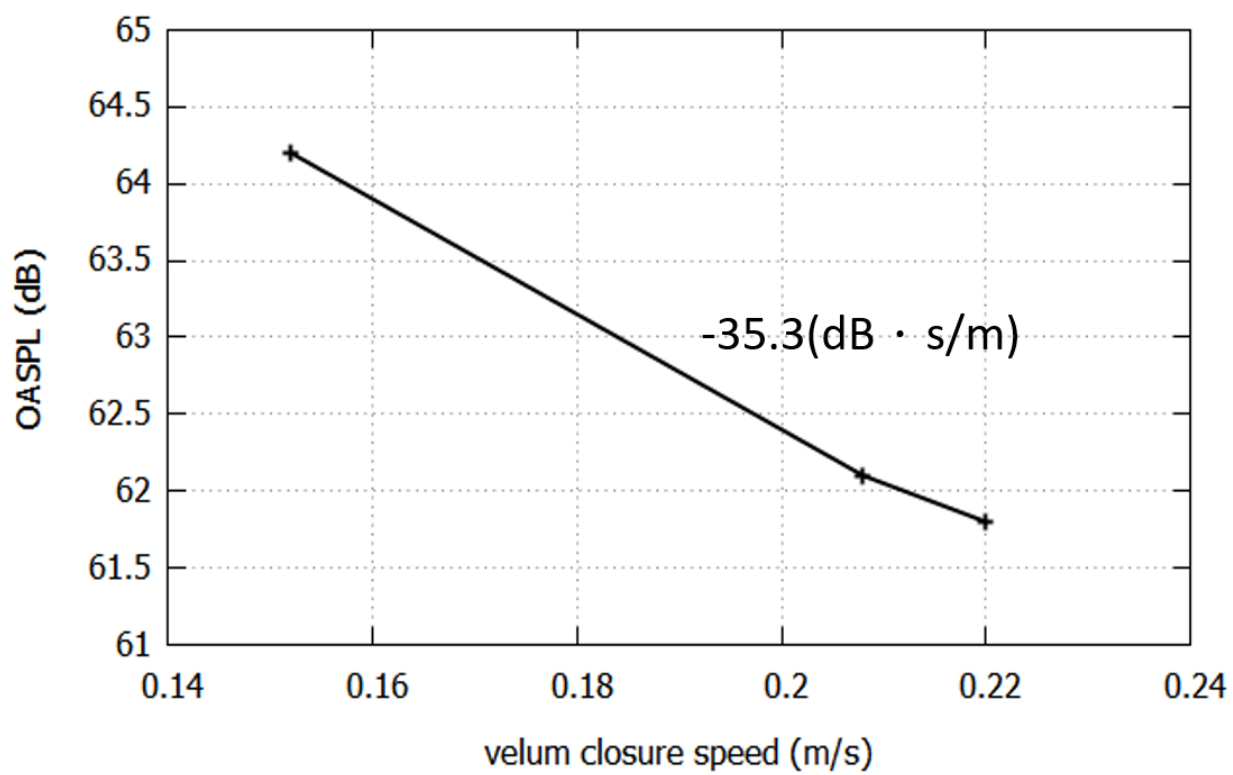


Figure 4-23. The relation of OASPLs and the velum closure speed.

Chapter 5. Summary and future work

To investigate the effect of the speech organ on the sibilant /s/ sound, the current study conducted the simulation of the realistic vocal tract model with the developed numerical framework. First, the simulations of the vocal tract geometry of sibilant /s/ sound were conducted to clarify the sound changes by the inclination angles of the incisor. Then, the simulation of sibilant /s/ production with the realistically pronouncing process was conducted to explore the effect of velopharyngeal closure and tongue movement. The parameters and conditions of all cases are shown in Table 5-1.

In chapter 3, the numerical flow simulations of the vocal tract geometry of /s/ at different incisor angles were conducted to examine the origin of sound changes. On the basis of the far-field SPL spectrum, increasing the incisor angle from 0° to 30° had no influence on the characteristic peak of the sibilant sound at 4 kHz. However, increasing the incisor angle reduced the amplitude of the sound in the frequency range from 8 kHz to 12 kHz. In the flow field, the turbulence intensity kept the same level and the maximum velocity occurred at the sibilant groove in all cases, while the high RMS value region moved from the cavity between the teeth and the lower lip to the tip of the lower lip when the inclination angle increased.

By conducting acoustic simulations with a monopole source at the potential sound source positions of the 0° and 30° cases, we found that the acoustic source position affects the resonance of the sound wave and influences the far-field SPL spectrum. Specifically, if the sound source position was located closer to the exit of the vocal tract, i.e., the lips, the source didn't couple strongly with the resonator, and no significant frequency could be caught at far-field. Because the flow channel downstream of the sibilant groove became wider when the incisor angle was increasing from 0° to 30°, the large velocity fluctuation region was shifted and the amplitude of the far-field sound around 10 kHz was reduced. Consequently, the slight change of the geometry

affected less to the turbulent intensity, but changed the flow configuration and shifted the position of the potential sound source, thereby influenced the performance on the acoustics field.

In chapter 4, the numerical simulation of sibilant /s/ production with the realistically moving vocal tract was conducted to investigate the articulation process of velopharyngeal closure and tongue movement. The sibilant /s/ sound production process was simulated in the part of the Japanese word /usui/ which is extracted after the vowel production of /u/ to the sibilant sound /s/, including the tongue elevation and the velopharyngeal valve closure. By analyzing the spectrogram of the generated sound, the current study provided a clear description of the sibilant /s/ articulation process which contributes to the clinical treatment for the sibilant sound production. The further simulations with the different closing speed of the velopharyngeal valve indicated that the sound amplitude is strongly related to the velopharyngeal opening due to the similar frequency range of the broadband noise during the co-articulation process in all cases. Besides, the comparison of the far-field spectrogram showed that the louder sound will be generated by the slower closing speed of the velopharyngeal valve and might make the co-articulation between two sounds ambiguous which will let the listener difficult to understand the word.

The current study clarified the effect of the incisor angle and the effect of the velopharyngeal closure on sibilant /s/ sound by the implicit numerical simulation with the developed ASTS schemes. These results provided the underlying insights necessary to the dental treatment for generating the sibilant fricatives. By further investigation, these findings and the present framework will enable us to quickly predict the effect of the vocal tract for more subjects on the production of sibilant fricatives to build the database for patients prior to prosthetic surgery.

Table 5-1. The parameters and conditions of all cases.

Chapter 3					Chapter 4		
Case	0° model	10° model	20° model	30° model	Faster closing case	Base simulation	Slower closing case
Boundary condition	Stationary boundary condition				Moving boundary condition		
Closing speed of velum (m/s)	-	-	-	-	0.220	0.208	0.152
Computational domain (mm³)	409.6 x 409.6 x 409.6	409.6 x 409.6 x 409.6	409.6 x 409.6 x 409.6	409.6 x 409.6 x 409.6	409.6 x 204.8 x 204.8	409.6 x 204.8 x 204.8	409.6 x 204.8 x 204.8
Total number of cells N_{cell}	8.7×10^7	7.2×10^7	7.8×10^7	6.7×10^7	1.82×10^8	1.82×10^8	1.82×10^8
Size of the largest cell (mm)	1.6	1.6	1.6	1.6	1.6	1.6	1.6
Size of the smallest cell (mm)	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Time step (s)	1×10^{-6}	1×10^{-6}	1×10^{-6}	1×10^{-6}	2×10^{-6}	2×10^{-6}	2×10^{-6}
Inlet velocity (m/s)	1.5	1.5	1.5	1.5	1.5	1.5	1.5

Reference

1. 荻安誠, 外山稔, & 松平登志正. (2016). コミュニケーション障害の疫学: 音声言語・聴覚障害の有病率と障害児者数の推定. 京都学園大学健康医療学部紀要, (1), 1-12.
2. Shadle, C. H. (1990). Articulatory-acoustic relationships in fricative consonants. In *Speech production and speech modelling* (pp. 187-209). Springer, Dordrecht.
3. Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, 108(3), 1252-1263.
4. Badin, P. (1991). Fricative consonants: acoustic and X-ray measurements. *Journal of phonetics*, 19(3-4), 397-408.
5. Runte, C., Lawerino, M., Dirksen, D., Bollmann, F., Lamprecht-Dinnesen, A., & Seifert, E. (2001). The influence of maxillary central incisor position in complete dentures on/s/sound production. *The journal of prosthetic dentistry*, 85(5), 485-495.
6. Lee, A. S., Whitehill, T. L., Ciocca, V., & Samman, N. (2002). Acoustic and perceptual analysis of the sibilant sound/s/before and after orthognathic surgery. *Journal of Oral and Maxillofacial surgery*, 60(4), 364-372.
7. Liu, R., Hu, A., Niu, L., Wang, F., Lu, Y., & Zhou, Q. (2019). Association of incisal overlaps with/s/sound and mandibular speech movement characteristics. *American Journal of Orthodontics and Dentofacial Orthopedics*, 155(6), 851-859.
8. Fonteyne, E., Van Doorne, L., Becue, L., Matthys, C., Bronckhorst, E., & De Bruyn, H. (2019). Speech evaluation during maxillary mini-dental implant overdenture treatment: A prospective study. *Journal of oral rehabilitation*, 46(12), 1151-1160.

9. Hu, S., Wan, J., Duan, L., & Chen, J. (2021). Influence of pontic design on speech with an anterior fixed dental prosthesis: A clinical study and finite element analysis. *The Journal of Prosthetic Dentistry*, 126(2), 204-e1.
10. Stevens, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants: Static considerations. *The journal of the acoustical society of America*, 50(4B), 1180-1192.
11. Burnett, C. A., & Clifford, T. J. (1993). Closest speaking space during the production of sibilant sounds and its value in establishing the vertical dimension of occlusion. *Journal of dental research*, 72(6), 964-967.
12. Hamlet, S. L., Cullison, B. L., & Stone, M. L. (1979). Physiological control of sibilant duration: Insights afforded by speech compensation to dental prostheses. *The Journal of the Acoustical Society of America*, 65(5), 1276-1285.
13. C. H. Shadle. (1985). *The acoustics of fricative consonants*, Ph.D. Thesis, Massachusetts Institute of Technology.
14. Van Hirtum, A., Grandchamp, X., Pelorson, X., Nozaki, K., & Shimojo, S. (2010). Les and" in Vitro" Experimental Validation of Flow around a Teeth-Shaped Obstacle. *International Journal of Applied Mechanics*, 2(02), 265-279.
15. Cisonni, J., Nozaki, K., Van Hirtum, A., Grandchamp, X., & Wada, S. (2013). Numerical simulation of the influence of the orifice aperture on the flow around a teeth-shaped obstacle. *Fluid Dynamics Research*, 45(2), 025505.
16. Nozaki, K., Yoshinaga, T., & Wada, S. (2014). Sibilant/s/simulator based on computed tomography images and dental casts. *Journal of dental research*, 93(2), 207-211.

17. Nozaki, K. (2010). Numerical simulation of sibilant [s] using the real geometry of a human vocal tract. In *High Performance Computing on Vector Systems 2010* (pp. 137-148). Springer, Berlin, Heidelberg.
18. Yoshinaga, T., Nozaki, K., & Wada, S. (2018). Experimental and numerical investigation of the sound generation mechanisms of sibilant fricatives using a simplified vocal tract model. *Physics of Fluids*, 30(3), 035104.
19. Pont, A., Guasch, O., Baiges, J., Codina, R., & Van Hirtum, A. (2019). Computational aeroacoustics to identify sound sources in the generation of sibilant/s. *International journal for numerical methods in biomedical engineering*, 35(1), e3153.
20. Yoshinaga, T., Nozaki, K., & Wada, S. (2019). Aeroacoustic analysis on individual characteristics in sibilant fricative production. *The Journal of the Acoustical Society of America*, 146(2), 1239-1251.
21. Yoshinaga, T., Nozaki, K., & Iida, A. (2020). Hysteresis of aeroacoustic sound generation in the articulation of [s]. *Physics of Fluids*, 32(10), 105114.
22. Kummer, A. W. (2011, May). Types and causes of velopharyngeal dysfunction. In *Seminars in speech and language* (Vol. 32, No. 02, pp. 150-158). © Thieme Medical Publishers.
23. Oren, L., Kummer, A., & Boyce, S. (2020). Understanding nasal emission during speech production: a review of types, terminology, and causality. *The Cleft Palate-Craniofacial Journal*, 57(1), 123-126.
24. Kummer, A. W., Curtis, C., Wiggs, M., Lee, L., & Strife, J. L. (1992). Comparison of velopharyngeal gap size in patients with hypernasality, hypernasality and nasal emission,

or nasal turbulence (rustle) as the primary speech characteristic. The Cleft palate-craniofacial journal, 29(2), 152-156.

25. Kummer, A. W., Briggs, M., & Lee, L. (2003). The relationship between the characteristics of speech and velopharyngeal gap size. The Cleft palate-craniofacial journal, 40(6), 590-596.
26. Beer, A. J., Hellerhoff, P., Zimmermann, A., Mady, K., Sader, R., Rummeny, E. J., & Hannig, C. (2004). Dynamic near-real-time magnetic resonance imaging for analyzing the velopharyngeal closure in comparison with videofluoroscopy. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, 20(5), 791-797.
27. Bunton, K., & Story, B. H. (2012). The relation of nasality and nasalance to nasal port area based on a computational model. The Cleft palate-craniofacial journal, 49(6), 741-749.
28. Bunton, K. (2015). Effects of nasal port area on perception of nasality and measures of nasalance based on computational modeling.
29. Sundström, E., & Oren, L. (2019). Pharyngeal flow simulations during sibilant sound in a patient-specific model with velopharyngeal insufficiency. The Journal of the Acoustical Society of America, 145(5), 3137-3145.
30. Sundström, E., & Oren, L. (2019). Sound production mechanisms of audible nasal emission during the sibilant/s. The Journal of the Acoustical Society of America, 146(6), 4199-4210.
31. Sundström, E., Boyce, S., & Oren, L. (2020). Effects of velopharyngeal openings on flow characteristics of nasal emission. Biomechanics and modeling in mechanobiology, 1-13.

32. Sundström, E., & Oren, L. (2021). Change in aeroacoustic sound mechanism during sibilant sound with different velopharyngeal opening sizes. *Medical & Biological Engineering & Computing*, 59(4), 937-945.
33. Moll, K. L., & Daniloff, R. G. (1971). Investigation of the timing of velar movements during speech. *The Journal of the Acoustical Society of America*, 50(2B), 678-684.
34. Dalston, R. M., & Seaver, E. J. (1990). Nasometric and phototransductive measurements of reaction times among normal adult speakers. *Cleft Palate Journal*, 27(1), 61-67.
35. Warren, D. W., Dalston, R. M., & Mayo, R. (1993). Hypernasality in the presence of “adequate” velopharyngeal closure. *The Cleft palate-craniofacial journal*, 30(2), 150-154.
36. Li, C. G., & Tsubokura, M. (2017). An implicit turbulence model for low-Mach Roe scheme using truncated Navier–Stokes equations. *Journal of Computational Physics*, 345, 462-474.
37. Li, C. G. (2020). A compressible solver for the laminar–turbulent transition in natural convection with high temperature differences using implicit large eddy simulation. *International Communications in Heat and Mass Transfer*, 117, 104721.
38. Li, C. G., Tsubokura, M., Fu, W. S., Jansson, N., & Wang, W. H. (2015). Compressible direct numerical simulation with a hybrid boundary condition of transitional phenomena in natural convection. *International Journal of Heat and Mass Transfer*, 90, 654-664.
39. Hu, F. Q., Hussaini, M. Y., & Manthey, J. L. (1996). Low-dissipation and low-dispersion Runge–Kutta schemes for computational acoustics. *Journal of computational physics*, 124(1), 177-191.

40. Bogey, C., & Bailly, C. (2004). A family of low dispersive and low dissipative explicit schemes for flow and noise computations. *Journal of Computational physics*, 194(1), 194-214.
41. Berland, J., Bogey, C., & Bailly, C. (2006). Low-dissipation and low-dispersion fourth-order Runge–Kutta algorithm. *Computers & Fluids*, 35(10), 1459-1463.
42. Lian, C., Xia, G., & Merkle, C. L. (2009). Solution-limited time stepping to enhance reliability in CFD applications. *Journal of Computational Physics*, 228(13), 4836-4857.
43. Lian, C., Xia, G., & Merkle, C. L. (2010). Impact of source terms on reliability of CFD algorithms. *Computers & Fluids*, 39(10), 1909-1922.
44. Xu, X., Lee, J. S., & Pletcher, R. H. (2005). A compressible finite volume formulation for large eddy simulation of turbulent pipe flows at low Mach number in Cartesian coordinates. *Journal of Computational Physics*, 203(1), 22-48.
45. Nakahashi, K. (2003). Building-cube method for flow problems with broadband characteristic length. In *Computational fluid dynamics 2002* (pp. 77-81). Springer, Berlin, Heidelberg.
46. Yokoyama, H., Kitamiya, K., & Iida, A. (2013). Flows around a cascade of flat plates with acoustic resonance. *Physics of Fluids*, 25(10), 106104.
47. Martín, M. P., & Candler, G. V. (2006). A parallel implicit method for the direct numerical simulation of wall-bounded compressible turbulence. *Journal of Computational Physics*, 215(1), 153-171.
48. Candler, G. V., Wright, M. J., & McDonald, J. D. (1994). Data-parallel lower-upper relaxation method for reacting flows. *AIAA journal*, 32(12), 2380-2386.

49. Rieper, F. (2011). A low-Mach number fix for Roe's approximate Riemann solver. *Journal of Computational Physics*, 230(13), 5263-5287.
50. Kim, K. H., & Kim, C. (2005). Accurate, efficient and monotonic numerical methods for multi-dimensional compressible flows: Part II: Multi-dimensional limiting process. *Journal of computational physics*, 208(2), 570-615.
51. Li, C. G., Tsubokura, M., & Bale, R. (2016). Framework for simulation of natural convection in practical applications. *International Communications in Heat and Mass Transfer*, 75, 52-58.
52. Li, C. G., Tsubokura, M., Fu, W. S., Jansson, N., & Wang, W. H. (2015). Compressible direct numerical simulation with a hybrid boundary condition of transitional phenomena in natural convection. *International Journal of Heat and Mass Transfer*, 90, 654-664.
53. Xu, X., Lee, J. S., & Pletcher, R. H. (2005). A compressible finite volume formulation for large eddy simulation of turbulent pipe flows at low Mach number in Cartesian coordinates. *Journal of Computational Physics*, 203(1), 22-48.
54. Linnick, M. N., & Fasel, H. F. (2005). A high-order immersed interface method for simulating unsteady incompressible flows on irregular domains. *Journal of Computational Physics*, 204(1), 157-192.
55. Seo, J. H., & Mittal, R. (2011). A high-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries. *Journal of computational physics*, 230(4), 1000-1019.
56. Lighthill, M. J. (1952). On sound generated aerodynamically I. General theory. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 211(1107), 564-587.

57. Freund, J. B. (1997). Proposed inflow/outflow boundary condition for direct computation of aerodynamic sound. *AIAA journal*, 35(4), 740-742.
58. Weiss, J. M., & Smith, W. A. (1995). Preconditioning applied to variable and constant density flows. *AIAA journal*, 33(11), 2050-2057.
59. Snow, W. B. (1931). Audible frequency ranges of music, speech and noise. *The Bell System Technical Journal*, 10(4), 616-627.
60. Yoshinaga, T., Van Hirtum, A., Nozaki, K., & Wada, S. (2020). Acoustic modeling of fricative/s/for an oral tract with rectangular cross-sections. *Journal of Sound and Vibration*, 476, 115337.
61. Pont, A., Guasch, O., & Arnela, M. (2020). Finite element generation of sibilants/s/and/z/using random distributions of Kirchhoff vortices. *International journal for numerical methods in biomedical engineering*, 36(2), e3302.
62. Yoshinaga, T., Nozaki, K., & Wada, S. (2019). A simplified vocal tract model for articulation of [s]: The effect of tongue tip elevation on [s]. *PloS one*, 14(10), e0223382.

神戸大学博士論文

「歯茎摩擦音[s]の発声メカニズムに関する数値的研究」，全132頁。

提出日 2022年07月12日

本博士論文が神戸大学機関リポジトリ**Kernel**にて掲載される場合、掲載登録日（公開日）はリポジトリの該当ページ上に掲載されます。

© 呂學叡 LU HSUEH JUI

本論文の内容の一部あるいは全部を無断で複製・転載・翻訳することを禁じます。

Doctor Thesis, Kobe University

“Numerical study on the phonation mechanisms of sibilant [s]”, 132 pages.

Submitted on 07/12/2022.

The date of publication is printed in cover of repository version published in Kobe University Repository Kernel.

© 呂學叡 LU HSUEH JUI

All Right Reserved, 2022