CROSS-SECTIONAL AREA AND LENGTH CHANGES WITH A SEMI-OCCLUDED VOCAL TRACT: THE ROLE OF THE EPILARYNX IN HUMAN PHONATION

by

Amanda C. Heller-Stark

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF SPEECH, LANGUAGE, AND HEARING SCIENCES

In Partial Fulfillment of the Requirements

For The Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2023

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Amanda C. Heller-Stark, titled Cross-Sectional Area and Length Changes with a Semi-Occluded Vocal-Tract: The Role of the Epilarynx in Human Phonation and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Robin A. Samlan — Date: Apr 24, 2023

Brad H. Story — Date: Apr 24, 2023

Kate Bunton (Apr 24, 2023 13:24 PDT)
Kate Bunton — Date: Apr 24, 2023

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

We hereby certify that we have read this dissertation prepared under our direction and recommend that it be accepted as fulfilling the dissertation requirement.

Robin A. Samlan
Dissertation Committee Co-Chair
Speech, Language, and Hearing Sciences
Date: Apr 24, 2023

Brad H. Story
Dissertation Committee Co-Chair
Speech, Language, and Hearing Sciences
Date: Apr 24, 2023

ACKNOWLEDGEMENTS

I am deeply grateful to those who contributed to the completion of this dissertation. First and foremost, I would like to express my sincere gratitude to my supervisors, Drs. Robin Samlan and Brad Story, for their support, encouragement, and guidance throughout this process. I would also like to thank my committee member, Dr. Kate Bunton, for her time and expertise that helped me refine my research. Drs. Ingo Titze and Lynn Maxfield for the financial support that allowed this project to be completed. Additionally, I would like to acknowledge the participants in my study, without whom this research would not have been possible.

DEDICATION

To Adeline and Hudson, the most important people in my life.

Table of Contents

LIST OF FIGURES

CHAPTER 1:

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF TABLES

ABSTRACT

The configuration of the vocal tract acoustically transforms the sound produced by the vibration of the vocal folds into the vowels and consonants listeners recognize as speech or song. Although the phonetic components of speech are well known to be produced by modulating the vocal tract with expansions and constrictions, the focus of this study was on understanding how specific types of vocalizations targeted to enhance voice quality generate refinements of the vocal tract shape through further selective widening (expansions/lengthening) and narrowing (constrictions/contractions).Positioned between the vocal folds (source) and vocal tract (filter), the epilarynx has been proposed to play a critical role in vowel identity and voice quality. Yet, the physiological changes or series of modifications that take place in this region during specific types of vocalization are not well understood. The purpose of this study was to provide a better understanding of vocal tract shapes across several voicing profiles, commonly used during clinical interventions for the treatment of voice disorders, by quantifying them using high-resolution magnetic resonance imaging (MRI).

CHAPTER 1: INTRODUCTION

1.1. The Vocal Tract

The vocal tract is the airway extending from the glottis (the space between the vocal folds) to the lips. Its acoustic characteristics shape the sounds we recognize as voice and speech. Anatomically the vocal tract can be divided into three main areas: the oral cavity, the pharynx, and the epilarynx. Biologically, the vocal tract plays a primary role in air exchange and swallowing. For sound production, these structures provide the ability for humans to vocalize and achieve phonemic contrasts for speech. Each portion or section of the vocal tract varies in its functionality, purpose, and role in human vocalization.

The oral cavity is formed collectively by the lips, teeth, tongue, jaw, hard palate and soft palate, as well as the surrounding soft tissue, whereas the pharynx is the airspace that extends from the soft palate to the larynx, and can be subdivided into regions called the nasopharynx, oropharynx, and laryngopharynx. Together, the oral cavity and pharynx comprise the vocal tract and as the anatomical structures move or are positioned within them, the vocal tract shape and length change to form vowels and consonants in isolation and connected speech.

The larynx, an intricate part of the vocal tract, provides the vibratory source for vocalization. Positioned between the pharynx and the trachea, the larynx comprises three unpaired cartilages, three paired cartilages, and several muscles (Hirano, 1974). The vocal folds themselves are a paired, five-layered structure that, when abducted, allow air to travel between the upper and lower airway and, when adducted, provide a seal to protect the lower airways from food, liquid, and foreign bodies. The space above the vocal folds has been called the vestibule, epilaryngeal tube, epilarynx, and, more recently, the laryngeal canal (Titze & Story, 2022).

For speech, each component of the vocal tract can be influenced, shaped, and modified by the speaker depending on a particular acoustic or phonetic target. Significant variability exists between humans regarding their vocal tract size, shape, and configuration. This variability, along with difficulty isolating individual structures, complicates our understanding of how vocal tract shape, and more exclusively the epilarynx, influences voice production and quality.

1.2. Epilarynx

The epilarynx, as it will be referred to throughout this document, is the airspace directly above the vocal folds and is defined by the anatomical boundaries of the laryngeal vestibule, ventricles, and ventricular fold glottis (Titze & Story, 2022). The range of terminology used to refer to the space above the vocal folds stems from distinct differences in its anatomical, physiological, and geometrical influence. The epilarynx has often been conceived to be tube-like in its shape (cf., Sundberg, 1974); however, this view doesn't encapsulate the complex geometry of how this region can be configured (Titze & Story, 2022). Moisik (2013) argues for using the term epilarynx because it "refers to both the boundaries of the physical structures and the space it encloses." Yet, a more recent argument has been made to name this space the laryngeal canal. The term laryngeal canal (Titze & Story, 2022) incorporates the complex geometry and the tube-like shape of this space. Additionally, the term "canal" can imply and link many acoustic similarities to the ear canal (Titze & Story, 2022).

The epilarynx is hypothesized to have multiple roles and functions within the larynx, but its role has been typically considered to be quite limited for speech production. Physiologically the epilaryngeal space follows closely to the function of the vocal folds as it becomes narrowly closed and sealed off to protect the airway during swallowing, throat clearing, coughing, and bearing down. This space is generally thought to be open during breathing and could be argued to also be “open” during vocalization and speech. Although, its exact geometric configuration (narrowed versus opened) in human voice and speech is unknown.

Developmentally the epilarynx is used in infancy to create turbulent sounds (squeals, screams, glottal stops) through constricting or narrowing the space for optimal power transfer (Bettany, 2004; Esling et al., 2004; Benner, 2009; Titze, 2002). In adult speech, more oral articulatory actions are required, and it has been suggested the epilarynx plays an influential role (Bettany, 2004; Esling et al., 2004; Benner, 2009). The developmental shift in epilaryngeal shape and configurations implies a fundamental role of this space in speech that adapts across the lifespan. An argument still exists for the role of the epilarynx in voice quality, vocal production, and how its configuration shapes its interaction with the vocal folds and supraglottic structures to change sound production.

One of the major obstacles to understanding the role of the epilarynx in speech has been observing the dynamic changes of this space in vivo. Few researchers have attempted to objectively quantify the epilaryngeal space as part of the sound production system (Bergan et al., 2004; Story, 2002, 2005, 2008; Story et al., 1996,1998, 2001; Titze & Story, 1997; Titze, 2002; Titze et al., 2003) or to study the differences in epilaryngeal configuration and the consequences on speech output (Samlan et al., 2009, 2011, 2013; Samlan & Kreiman, 2014; Story et al., 2001). Moreover, no consensus has been made regarding the exact anatomical boundaries of this region (space and structure), and therefore replication of findings has not been made.

1.3. Speech Sound Production

To understand how sound is produced, we must first look at the individual contributions of both the source (vocal folds) and the filter (vocal tract). The vocal folds have two main positions, abducted (opened/breathing) or adducted (closed/speaking or singing). When the vocal folds are abducted during rest breathing, air is flowing through an open glottis. The glottal flow, or flow of air through the glottis, of completely abducted vocal folds, represents a steady movement of air from the lungs. During voicing, vocal fold oscillation causes a rapid opening and closing of the glottis, and the constant stream of air is modulated into a series of flow pulses. Graphically, glottal airflow can be represented in a glottal flow waveform. Peak glottal flow occurs the highest when the glottis is large and decreases rapidly when the vocal folds are moving toward midline, closing the glottis. (Fant, 1986)

In contrast to glottal flow, the glottal area represents the temporal variation of the glottal opening area as the vocal folds rapidly move together medially and then apart laterally (Story, 2015). Glottal area can also be represented graphically similar to the glottal flow waveform, showing the area over time as the vocal folds are opening, maximally opened, closing, and maximally closed. Glottal area is greatest when the vocal folds are apart and smallest when the vocal folds are in close proximity to the glottal midline. Glottal area and flow waveforms share a similar appearance; however, these two waveforms are not directly proportional (Titze, 1984). The slight disproportion between glottal flow and area is due to the peak of glottal airflow being skewed to the right of peak glottal area (Fant, 1986). The skewing of glottal flow means that maximum airflow through the glottis lags slightly behind maximum glottal area. The skewing phenomenon occurs because of the acoustic and aerodynamic interaction of the source and filter (Titze, 1984; Fant, 1986). Cyclic variations of the glottal airflow produced by the oscillating vocal folds generates a complex acoustic wave consisting of a fundamental frequency and many harmonic components.

The vocal tract is an acoustic system whose resonances filter the sound transmitted through it. Resonance is a physical phenomenon where a system responds to an oscillatory input by reflecting and transmitting energy. This means that the vocal tract configuration will determine the acoustic output by selectively altering the frequency spectrum of the sound based on the morphometric configuration of the vocal tract (cf., Fant, 1960). As the sound wave travels through the vocal tract, the amplitude of harmonics in close proximity to the natural resonances of the vocal tract, called resonant frequencies (Fant, 1960), are enhanced whereas other harmonics are suppressed in amplitude. When viewed in an amplitude spectrum, each of these resonant peaks are referred to as a formant, which represents a specific band of harmonic frequencies that will be enhanced by the vocal tract. Formant frequencies are the primary acoustic characteristics that encode phonetic information and are also important in voice quality.

For a system to be self-oscillating or a self-sustained oscillatory system several requirements must be met. First, the system requires an energy source. Second, the system can grow in maximum excursion (amplitude) if energy input is greater than internal energy loss. Conversely, the system becomes damped/reduced if energy input is smaller than internal energy loss. Lastly, the system is steady in amplitude if energy input is equal to internal energy loss. The vocal folds are inherently self-sustaining and independent from the motor system in their ability to vibrate. The motor system is required to put the vocal folds in the position to vibrate.

The pressure below the vocal folds, subglottal pressure (P_s) is the driving force of energy below the vocal folds. The pressure above the vocal folds, supraglottal pressure (P_i) is the driving force of energy above the vocal folds, and the pressure inside the glottis or between the vocal folds is intraglottal pressure (P_g). (Figure 1.1).

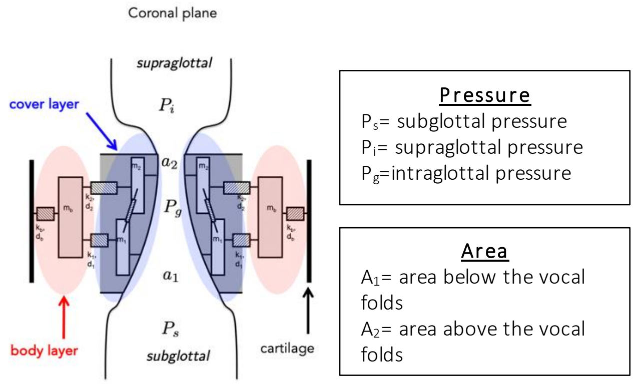


Figure 1.1: Model of the vocal folds that illustrate the mechanisms of vocal fold vibration (based on course slides adapted by B. Story from Story (2015, Ch. 3)).

1.3.1 Mechanism #1: Self-sustaining oscillation by convergent and divergent configurations

In the convergent profile, P_s is greater than P_i, making the P_g nearly equal to P_s. This pressure differential creates a lateral movement or glottal opening, making the area below the vocal folds larger than the area above the vocal folds. As the vocal folds are displaced laterally, the restoring force of the tissue, based on their elastic properties, opposes the lateral movement. As the restoring force opposes the lateral movement, P_g begins to decrease, and the glottic area gets wider. With the combination of these two forces (increased restoring force and decrease P_g), the tissue begins to change its direction of movement. In the divergent profile, P_s is less than P_i, which creates a medial movement where P_g is nearly equal to P_i, and a glottal closing is achieved. In the divergent profile, the area below the vocal folds is smaller than the area above the vocal folds. For the vocal folds to sustain their vibration over time, the P_g must decrease over the entire cycle as illustrated in both Figures 1.2 and 1.3.

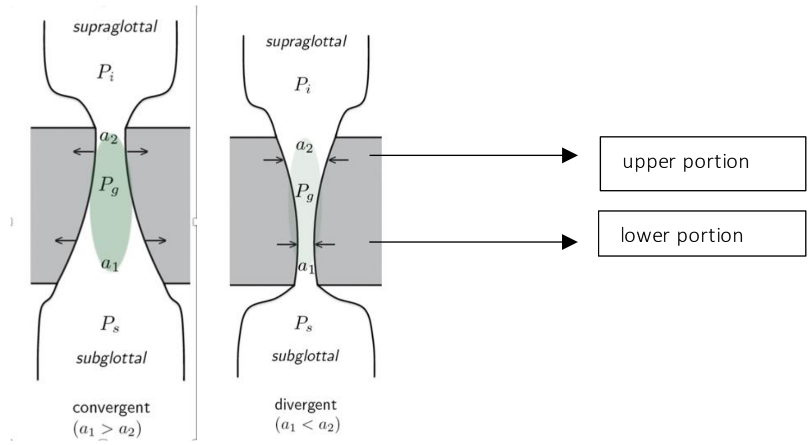


Figure 1.2: Illustration of the intraglottal pressure during the convergent and divergent phases of a vibratory cycle (based on course slides adapted by B. Story from Story (2015, Ch. 3)).

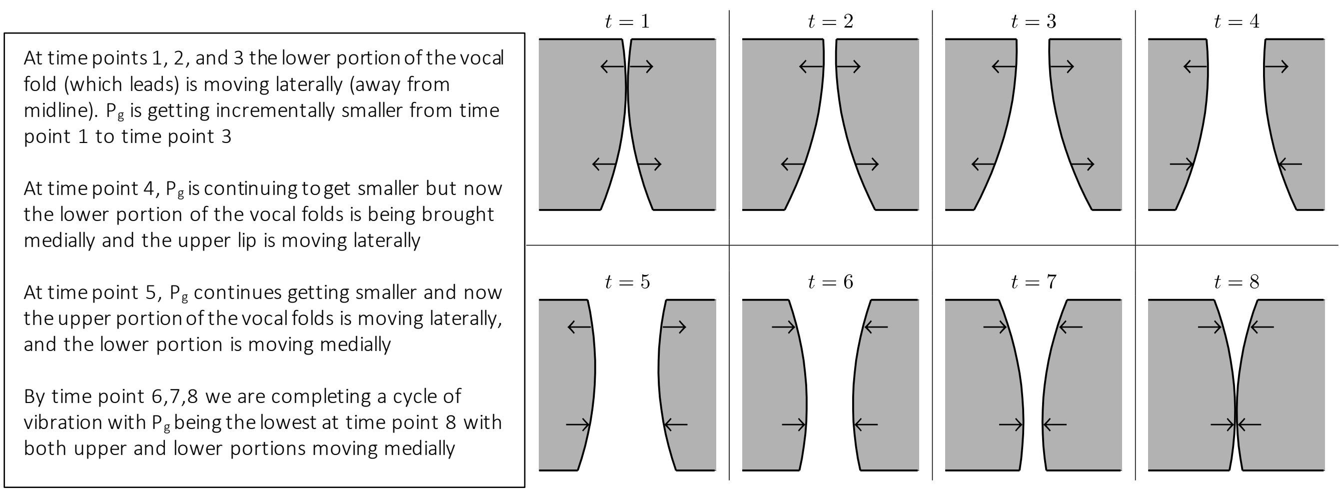


Figure 1.3: Illustration of the coronal plane of the vocal fold during one cycle of vibration (based on course slides adapted by B. Story from Story (2015, Ch. 3)).

1.3.2 Mechanism #2: Self-sustaining oscillation by the interaction of the vocal tract and larynx

The second mechanism by which the vocal folds can self-sustain vibration is due to the inertia of the air column above the vocal folds formed by the vocal tract. As illustrated in Figure 1.4, during the glottal opening phase the vocal folds move laterally and airflow through the glottis increases. This accelerates the air particles above (downstream) the glottis and increases the pressure at the entry to the vocal tract (P_i in Figure 1.2) which results in an increased outward force on the vocal folds. Eventually the restoring forces generated by tissue elasticity will overcome the intraglottal pressure and the vocal folds will begin to move medially toward the glottal midline. Simultaneously, the air in the vocal tract continues to move upward because of its inertia, and this results in a decrease in pressure above and within the glottis. Thus, the inertance of the vocal tract has the effect of increasing the intraglottal pressure during lateral vocal fold motion and decreasing it during medial motion, producing the driving forces needed to sustain oscillation.

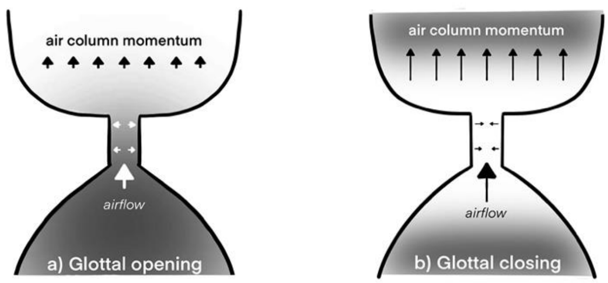


Figure 1.4: Illustration of the mechanism of vocal fold oscillation due to the vocal tract inertance (Adapted from Švec, J. G., Schutte, H. K., Chen, C. J., & Titze, I. R. (2021))

1.3.3 Characterization of the glottal flow waveform

Open quotient (Q_o) and skewing quotient (Q_s) are measures that can provide information about the glottal source and vocal tract. Q_o and Q_s are extracted from the glottal flow waveform and calculated over the period duration. The glottal flow waveform is a visual and numeric representation of the transglottal flow and the displacement of the flow. These measures tell us about the shape of the waveform, which can subsequently inform us about the quality of the sound generated.

The Q_o is a measure influenced directly by the glottal width or the area between the right and left vocal processes. The more adducted (closed) the vocal processes are, the lower the Q_o value will be. The more abducted (open) the glottal width is, the higher or greater the Q_o value.

Holmberg, Hillman, and Perkell (1988) reported normative values for the open quotient. These values are as follows:

0.4-0.7= normal voice quality
Less than 0.4=strained quality
Greater than 0.7=breathy quality

If the open quotient were to represents the abduction or adduction of the vocal folds, then a value that is less the normative range will be the more adducted (closed) glottis and vice versa. The skewing quotient tells us about the ratio of rise rate to fall rate in glottal airflow. In a level 1 interaction, the change in glottal airflow will directly result from the acoustic feedback from the vocal tract. An increased skewing quotient of the glottal airflow will increase the maximum flow declination rate (MFDR). MFDR is related, in part, to how fast the vocal folds are closing and has been perceptually linked to vocal intensity. The skewing quotient values typically range from 1.0 to 5.0 in normal voice quality (Bergen et al., 2004), where a value of 1.0 represents an equal ratio of rise to fall time. In the source-filter model, the Q_s is a function of the vocal tract inertance and has been theorized to be dictated by the epilarynx (Titze & Story, 1997). Inertia is a physical property defined as a body that tends to resist any change in motion. The inertance of the vocal tract can then be thought of as the vocal tract's ability to resist the change of airflow and subsequently dictate the acoustic output. Q_s will therefore increase or decrease depending on the shape of the vocal tract. If $Q_s=T_p/T_n$, or the ratio of the time of increase flow (positive slope) over the time of decreasing flow (negative slope), there would have to be an inverse relationship with the vocal tract size. So, if the vocal tract is getting larger, the skewing quotient will decrease, and as the vocal tract narrows, the skewing quotient will increase. To maximize the acoustic benefit perceptually, speakers increase their skewing quotient by nonlinear source-filter coupling (Titze,

2004) or by speeding up glottic closure (Granqvist et al., 2003). The skewing quotient would be expected to be smaller in people with a greater glottal area (increase in glottal area).

When a person produces higher pitches, the energy in the second harmonic can increase or decrease rapidly depending on how close it is to the resonance. This has a perceptual characteristic if the second harmonic passes through the first harmonic. In the example of changing your register (gliding from a low pitch to a high pitch), this can be perceived as a change in quality because you can gain or lose energy in the second harmonic. The closer the harmonics are to the nearest resonance, the more "boost" the harmonic will be given. If the harmonic goes past the resonance, then it will be diminished in its energy. This has been documented to be either reinforce or not reinforce the resonance.

The H1-H2 measure can be calculated to quantify the difference between the amplitude of the fundamental frequency and $2F_0$. So, when there is a higher value of H1-H2, this can indicate either a prolonged opening and closing time or that the glottal flow does not have any abrupt change that would normally be seen during vocal fold collision. This number represents the relative strength of the relationship of harmonics one and two. Samlan and Story, 2011 documented that the modeled H1-H2 of the glottal area increased with greater adduction of the vocal processes, but the H1-H2 of the glottal flow did not. H1-H2 was estimated from the spectrum of the acoustic pressure signal in this computation. Due to the presence of the first resonance (F1), to adequately quantify H1-H2, a correction for the first resonance must be made. By making this correction, they were able to represent the temporal (time) and spectral characteristics of the glottal flow waveform. Despite being a computational study, this work elucidated that different vocal fold shapes, nodal point/pivot point, and vocal tract configurations play a dramatic role in influencing H1-H2. Even though H1-H2 had been theorized to represent the degree of breathiness acoustically, these author's concluded that, in fact, given the variability of the different vocal configurations, this metric was not a good indicator of breathy voice quality.

In summary, the source-filter theory is based on two primary levels of interaction that influence acoustic airway pressure, or the sound we hear. The interaction occurs at two different levels. Level 1, the acoustic airway pressure is influenced by the glottal airflow at the source. Level 2, the acoustic airway pressure is influenced by the vibration pattern of the vocal folds (described above). The interaction between level 1 and level 2 becomes enhanced when the harmonic frequencies and amplitudes (source) are near the resonance's frequencies and amplitudes (vocal tract). In a level 1 interaction, transglottal pressure and glottal airflow affect the acoustic vocal tract pressures with little to no influence on the vocal fold vibration. Conversely, in a level 2 interaction, intraglottal pressure becomes the driving force for the modes of vibration of the vocal folds (Palaparthi et al., 2019).

Because of the *mostly* nonlinear interaction of the source-filter theory, it can be assumed that changes from downstream (subglottal) and upstream (supraglottal) will impact vocal fold vibration. In the linear theory, the vocal fold vibration is considered independent of the vocal tract. However, the linear theory of vibration fails to encapsulate the interactive nature of the entire system. In a computational study by Titze and Story (1997), the vocal tract was found to facilitate optimal vocal fold vibration by matching the high impedance of the vocal tract to the relatively low impedance of the glottis. The impedance matching of the source filter suggests that, by narrowing or constricting certain portions of the vocal tract (i.e., the epilarynx), an improvement in voice quality or potentially vocal effort may be achieved.

1.4. Vocal Tract Configuration and Voice Quality

The acoustic resonance characteristics of speech sounds are produced by the shape of the entire vocal tract, extending from glottis to lips. Although particular regions of the vocal tract are often discussed as independent systems (e.g., oral cavity, velopharynx, etc.), the shape of the airspace at any given instant of time is formed collectively by the relative positioning of the articulators, thus configuring the entire vocal tract. For example, if it were thought that the difference in tongue and lip position were the only influencing maneuvers separating a front from a back vowel, then the relative positioning of the velum or area of the pharynx would be irrelevant. In the original linear theory (Fant, 1960), the vocal folds (vibrating source) and vocal tract (filter) were assumed to be independent of one another and therefore combined to generate an output. Recently, it has been accepted that a nonlinear relationship exists between the source and filter, and vowel identity should therefore be thought of as the interaction between the two (Titze, 2008).

A speaker can quickly and dynamically change the shape of their vocal tract to produce a targeted output. Vowels, and vowel-like sounds, are differentiated based primarily on the first two formants (Peterson & Barney, 1952; Hillenbrand et al., 1995), with some influence from the third formant (Story et al., 2016). Principal component analyses of vocal tract area functions based on MRI measurements (Story & Titze, 1998; Story 2005) showed that the first two principal components map in a 1:1 manner to the first two formants (F_1/F_2). Story (2005) systematically manipulated the cross sectional-areas of a "neutral" vocal tract to achieve targeted F_1 and F_2 formant frequencies using a computational modeling framework. This work found that the /i/ vowel (acoustically low F_1 and high F_2) was produced when an expanded pharynx accompanied a constricted oral cavity. An expanded oral cavity and constricted pharynx were required to produce an /ɑ/ vowel (high F_1, low F_2). /u/ vowels (low F_1 and F_2) are created with constriction near the lips and velum but with an expanded oral and lower pharynx (Table 1). These findings confirm that while the target f_o may change, to retain the integrity of the vowel or vowel-like sound, the speaker must manipulate large portions of their vocal tract area and possibly length to position the first two formant frequencies (Story, 2006).

Table 1: Summary of vocal tract shape and vowels

Vowel	F_1/F_2	Oral Cavity	Pharynx
/ɑ/	Low/high	Expanded	Constricted
/i/	High/low	Narrowed	Expanded
/u/	Low/low	Expanded	Expanded

Timbre, quality of sound, and the overall identity of the voice are dependent on both the speaker's intention and the listener's perception. It is known that spectral slope influences voice quality. Source spectral slope for four segments of the spectrum appear to have unique contributions to overall voice quality: the differences between the first and second harmonics (H1–H2), second and fourth harmonics (H2–H4), overall spectral slope, and high-frequency noise (Kreiman et al., 2014). Story (2016) commented that sound quality of voice is also based on the location or the clustering of the acoustic resonance F3, F4, and F5. In the long-term average spectrum (LTAS), the region of F3-F5 has been called the "singer's formant cluster," given the broad spectral energy resulting from two or more resonances of the vocal tract. (Story, 2016). Sundberg (1974) offered that the singer's formant resulted from the epilarynx functioning as its own resonator. If the epilarynx were not connected to the vocal tract and were open to free space, its resonance frequency would be 2500-3000 Hz, depending on its length (Sundberg, 1974). How the singer’s formant cluster is achieved functionally has been argued by researchers but has been suggested to be created by narrowing and lengthening of the epilarynx (Story et al., 1996, 2001, 2005, 2008; Echternach et al., 2011) or through a higher or lowered laryngeal position (Wang, 1986). While several modeling studies have shown that the singer's formant cluster is achieved by narrowing the epilarynx, Detweiler (1994) did not confirm this finding through MRI, stroboscopy, or acoustic analysis.

It could be argued that just like for singers, certain speakers also demonstrate a "speaker's formant cluster." The idea of a "speaker's formant cluster" has been hypothesized, but limited supporting evidence has been provided. In theory, given that the overall quality of the voice is influenced by the location of the formant frequencies, then it becomes plausible that identifying clustering of speaking formants may help define and understanding how voice quality is achieved, manipulated, or deviates from "normal". Krieman et al. (2014) suggested that the identification of acoustic parameters that account for or predict the perception of voice quality must mean that speakers can control or manipulate their system physiologically to achieve such targets.

1.5. Semi-Occluded Vocal Tract (SOVT)

Semi-occlusion of the vocal tract (SOVT) refers to an articulatory action that *nearly* closes the airway at some location between the glottis and lips. For example, a /z/ sound requires that the tongue tip nearly occludes the vocal tract by positioning it almost in contact with alveolar ridge; whereas an /m/ is a complete oral occlusion with impedance of the nasal system. Semi-occlusion may also refer to temporarily lengthening the vocal tract through internal means (e.g., lip trill, tongue trill) or with assistance from a tube introduced at the lips (i.e., artificial lengthening). The concept of using a semi-occluded vocal tract to accomplish derived changes in vocal tract function is unique, and it has been shown that the semi-occlusion or occlusion during voicing tasks can reduce the cross-sectional area of a portion of the vocal tract other than just the location of the semi-oclusion (Story et al., 2000; Titze, 2006b). Reduction in the cross-sectional area is theorized to maximize the nonlinear interaction between the source (vocal folds) and filter (vocal tract), with an overarching goal of encouraging impedance matching, reducing energy loss within the systems, and consequently minimizing vocal fold contact/collision.

Vocal tract semi-occlusions can balance aerodynamic and acoustic energy through three different means: changing the shape of the vocal tract, changing vocal tract length, and narrowing/constricting the epilarynx. (Titze and Verdolini-Abbott, 2012; Rosenberg, 2014). One of the main ways SOVT productions improve the aerodynamic energy within the system is by increasing vocal tract inertance. Increasing vocal tract inertance is achieved by impedance matching, or the coupling of the glottal source with the vocal tract. In this interactive system, by reducing the phonation threshold pressure and increasing the skewing of the glottal flow waveform, vocal tract impedance can be increased and affect the glottal function through acoustic-aerodynamic interactions and mechano-acoustic interactions.

Several pivotal findings on inertive reactance of the vocal tract have come from computational modeling. Overall, it has been found that all SOVT productions amplify the inertive reactance of the vocal tract predominantly in the fundamental frequency range (Titze, 2004), which is most commonly used in conversational speech. (Story et al., 2000; Titze & Laukkanen, 2007). The increase of inertive reactance in the 100-200 Hz F_0 range alters the interaction and subsequent functioning of the vocal tract and vocal folds by reducing phonation threshold pressure. (Story et al., 200l; Conroy et al., 2014; Horacek et al., 2019b; Kang, Scholp et al., 2019, Tangeney et al., 2019).

The physical and physiological effects of using a semi-occluded vocal tract to improve voicing efficiency and, subsequently, voice quality have been of interest to voice scientists and therapists. Several researchers have proposed that using semi-occluded sounds allows a talker to automatically align the vocal tract so that the benefits of semi-occlusions are achieved without the need for explicit instruction on "how to produce the sound." This idea of requiring no direction or additional maneuvering to achieve improved vocal efficiency and quality seems attractive.

Relying solely on a semi-occluded vocal tract, however, poses some challenges and questions. First, what is the goal of the semi-occluded voicing tasks/production? Second, what is the voice user's knowledge base of sound production, and can they recognize suboptimal voicing patterns? In addition to translating these approaches to disorders populations, these challenges and questions have been a primary focus of recent literature in voice science. It is not really understood how the principal mechanism of semi-occluded vocal tract tasks translates to or how the benefits of these productions are maintained over time.

1.6. Use of Semi-Occluded Vocal Tract in Voice Therapy

The goal of most voice therapy is to improve voice quality and function. Voice therapy is often the first line of treatment for people with voice complaints. One common approach to voice therapy is "resonant voice," in which patients are asked to produce sound that is "easy" and buzzy in the facial tissues (Titze & Abbot, 2012), or "forward focused." The resonant productions are taught across speaking activities beginning with sustained sounds and progressing to conversation. A resonant voice is considered to balance both degree of laryngeal adduction (breathy and pressed) and optimal interaction of the vocal folds and vocal tract (Titze, 2001). Computational modeling has shown that resonant voice can be achieved acoustically by narrowing the epilarynx to maintain an inertive vocal tract (Titze, 2001). Clinically this can be targeted by having speakers create more "space in the pharynx," lowering the larynx and reducing paralaryngeal tension and hyperfunction. While clinical anecdotes have supported theories of how a resonant voice is achieved, the exact physiological or anatomical changes responsible for creating a resonant voice are still being identified.

One approach to teaching resonance is to contrast a sound produced in a typical manner (i.e., say “hmm”) with specific directions for how to produce that sound (i.e., say “hmm,” noticing the vibrations in your lips, and allowing the buzzy feeling to increase). The typical production (i.e., say hmm) can be considered “untargeted,” and the production with specific directions “targeted,” in that the patient is targeting a feeling of vibration in their lips rather than just saying “hmm.” SLPs rely on their subjective impression of the perceptual characteristics of the voice quality to determine if a resonant voice is achieved. Enhancing kinesthetic feedback can also improve the patient’s ability to perceive the target (Verdolini-Abbott, 2000). Most voice therapy treatments use a combination of approaches to individualize treatments based on the patient's specific voice quality and complaints and the type of feedback most salient to the particular patient. An individualized approach is important in the clinical setting to ensure patient-centered care but has complicated our scientific understanding of what independent or interactive effect occurs during voice therapy.

“Semi-occluded vocal tract exercises (SOVTEs)” has been coined as a term referring to a class of vocalization techniques that use a semi-occluded vocal tract as “therapeutic exercises” targeted at improving voicing (e.g., voicing efficiency, voice quality, voice function, etc.) (Stemple et al., 1994; Verdolini-Abbott, 2000; Laukkanen et al., 2008; Rosenberg, 2014; Dargin et al., 2015; Meerschman et al, 2019; Stark et al., 2023). Use of the term “exercise” in this context indicates a repeated activity with the goal of altering a pattern, rather than an activity designed to improve muscle strength or efficiency.

Resonant voice, as described previously, is one of many therapy techniques that uses an SOVT. Semi-occlusions can also involve phonating while the vocal tract is partially occluded using articulators or external aids such as straws or tubes. Examples of SOVT productions include trilling (lips or tongue), voiced fricative or nasal consonants (e.g., /z/, /v/, /m/, /n/, /ŋ/), tube phonation, water-resisted phonation, and Y-buzz. Variations in tube size, length, size of the oral aperture, depth of water, targeted pitch, and volume all introduce variations of the semi-occlusion. Additionally, several formalized therapy approaches (Lessac-Madsen Resonant Voice Therapy, Vocal Function Exercises, and Accent Method) use semi-occluded techniques to elicit improved voice quality. Within the last decade, a proliferation of research articles on the scientific findings underscoring the utility of SOVTs has emerged (Andrade et al., 2016; Andrade et al., 2014; Bele, 2005; Bonette et al., 2020; Calvache et al., 2020; Conroy et al., 2014; Costa et al., 2011; Croake et al., 2017; Dargin & Searl, 2015; Dargin et al., 2016; Frisancho et al., 2020; Guzman et al., 2013a, 2013b, 2015, 2016, 2017a, 2017b; Kapsner-Smith et al., 2015; Laukkanen et al., 2008, 2012; Meerschman et al., 2019). These scientific findings, combined with clinical anecdotes, outcomes SOVT clinical studies have all documented changes across several domains of voicing. These changes have challenged researchers to seek to understand the complex interaction of SOVTs and improvements in vocalization.

1.7. Quantification of changes using an SOVT

Quantification of changes occurring with an SOVT has been documented through acoustic, aerodynamic, biomechanical, perceptual, quality of life, and morphometric measures (Table 2). Researchers have obtained measurements in these domains across a multitude of tasks (sustained phonation through connected speech), with most measurements taken immediately before and after an SOVT production (Andrade et al., 2016; Andrade et al., 2014; Bele, 2005; Bonette et al., 2020; Calvache et al., 2020; Conroy et al., 2014; Costa et al., 2011; Croake et al., 2017; Dargin & Searl, 2015; Dargin et al., 2016; Frisancho et al., 2020; Guzman et al., 2013a, 2013b, 2015, 2016, 2017a, 2017b; Laukkanen et al., 2008, 2012). Other studies have focused on outcome measurements following a course of formal SOVTE therapies (Kapsner-Smith et al., 2015; Meerschman et al., 2019). Changes in acoustic characteristics have been shown with SOVT productions in; F_0, sound pressure level (SPL), long-term average spectrum (*LTAS*), cepstral peak prominence (CPP), and several perturbation measures. Aerodynamic measures of air pressure (phonation threshold pressure, PTP), airflow (phonation threshold flow), and maximum flow declination rate (MFDR) have also shown positive changes. Minimal work has been done to explore in vivo changes that occur during SOVT productions, specifically with morphometric measurements (physical dimensions of the vocal tract).

Table 2: Domain-specific outcomes of an SOVT

Acoustic	**Aerodynamic**
Fundamental Frequency (F_0) Sound Pressure Level (SPL) Long-term Average Spectrum (LTAS) Cepstral Peak Prominence (CPP) Perturbation Measures AVQI	Air Pressure -Oral Pressure -Subglottal Pressure -Transglottal Pressure -Intraoral Pressure -Phonation Threshold Pressure Air Flow Airflow Phonation Threshold Flow Maximum Flow Declination Rate (MFDR)

Biomechanical	Perceptual	Quality of Life
Glottal Contact Quotient/Open Quotient Maximum Area Declination Rate (MADR)	GRBAS CAPE-V Other	Voice-Related Quality of Life Voice Handicap Index Other

Table 3: Summary of CT imaging studies of human vocal tract

CT Study	**Major Findings**			
	Subjects (Healthy/Disordered) (Trained/Untrained)	**Tasks**	**Analysis Technique**	**Product/Results**
Guzman et al., 2012	Healthy/Trained/Adult (N=1)	a/ before and after straw	Sagittal	Lowered laryngeal position, hypopharynx increased, larger inlet of pharynx and epilaryngeal tube.
Guzman et al., 2017	Disordered/Untrained/Adults (N=10)	Before , during, and after straw (stirring and drinking) on /a/	Sagittal and Transversal	Increased volume of vocal tract with straw- specifically in pharynx. More changes seen in stirring straw

Table 4: Summary of MRI imaging studies of human vocal tract

	Major Findings			
	Subjects	Tasks	Analysis Technique	Product/Results
Baer et al., 1991	Healthy/Trained/Adult (N=2)	9 vowels	Axial	Cross-sectional areas
Sulter et al.., 1992	Healthy/Trained/Adult (N=1)	[i, a, ε]	Axial; centerline	Resonance frequencies highly correlated
Story et al., 1996	Healthy/Trained/Adult (N=1)	12 vowels, 3 nasals, 3 plosives	Sagittal	Area functions provided.
Story et al., 1998	Healthy/Untrained/Adult (N=1)	22 vowels & consonants	Axial	Anatomical sex difference in size, but similar shape
Fitch et al., 1999	Healthy/Untrained/Toddler to Adult (N=129)	Static vocal tract	Sagittal	Sex difference in vocal tract length in adults not childhood most account for in pharynx
Clement et al., 2007	Healthy/Trained (N=1)	French vowels (/i/, /a/, /u/)	Sagittal	MRI and formant matching based on area functions for each vowel.
Yamasaki et al., 2011	Healthy/Disordered (N=10 nodules; N=10 normal)	Static vocal tract	Sagittal and Axial	Morphometric differences with nodules having greater laryngeal vestibule constriction; narrowed distance of true vocal folds; reduced AP distance in larynx.
Laukkanen et al., 2012	Healthy/Trained (N=1)	/a/ with and without straw	Sagittal and Axial	Oral cavity, pharynx and epilaryngeal increase during straw and stayed larger after- wider pharynx plays more of a role than widening of epilarynx.
Venturna et al., 2013	Healthy/ Trained/ Adult (N=3)	Spoken and Sung [i,a,u] low, comfortable and high pitch.	Sagittal	Singing voice increase oral cavity volume; pharyngeal area increased in acting voice.
Moisik et al., 2015	Healthy/Trained (N=1)	Creaky voice, glottal stop, aryepiglottal stop	Sagittal and Axial	Epilaryngeal constriction plays a role in suppressing vocal fold vibration.

Yamasaki et al., 2017	Healthy/Disordered/Adult (N=20)	Before and after straw in water at rest and phonation of /æ/	Sagittal	Positional differences at rest and during phonation between healthy and disordered. Vocal tract adjustments made during straw in water.
Leppavuori et al., 202	Healthy/Trained (N=4)	Singing vowels	Sagittal	Vocal modes in complete vocal technique found differences in vocal tract shapes with each vocal mode.

Morphometric changes of the physical dimensions of the vocal tract with semi-occluded productions have been explored computationally (Bickley &Stevens, 1986; Dollinger & Montequin, 2006; Mainka et al., 2015; Moisik, 2013; Samlan & Story, 2011; Samlan et al., 2013; Samlan & Kreiman, 2014; Story et al., 1996, 1998, 2000, 2001, 2002, 2005, 2008; Titze, 2002; Titze, 2006; Titze & Laukkanen, 2007; Titze & Palaparthi, 2016; Tom et al., 2001) and in vivo during computer tomography (CT) and magnetic resonance imaging (MRI) (Guzman, Castro, et al., 2013; Guzman, Laukkanen, et al., 2013; Guzman, Miranda, et al., 2017; Laukkanen et al., 2012, Yamaskai et al., 2011 & 2017). In one preliminary study of a trained vocal user, findings from CT and MRI studies have confirmed laryngeal depression (lowering of the larynx), narrowing of the velopharyngeal port, expansion of pharyngeal space, and increase in epilaryngeal area during and after semi-occlusion (Guzman et al, 2013). Although these are promising findings, a larger sample size, and observation of morphometric differences in untrained speakers are needed. The exact causative mechanism by which these changes occur has not been confirmed. From the data collected on SOVT productions, theories have supported that the lower PTP, increased inertive reactance, and increase in oral, pharyngeal, and intraglottal pressure are facilitated by passive stretching of the vocal tract and relaxation of pharyngeal and extrinsic laryngeal musculature. (References provided in Table 3 and 4).

1.8. MRI in Speech Science

Rokkaku et al., (1986) were the first group of researchers to systematically use still image MRI to evaluate speech production for Japanese vowel production. Shortly thereafter Baer et al., (1987) used similar methodology in MRI but explored English vowels. In the late 1980s image acquisition time took over 25 seconds for a single 5mm slice, leaving researchers restricted to single sagittal slices to extract vocal tract shapes from. With technological advances in MRI imaging, Baer et al., (1991) were able to combine images from all three planes to create the first 3D reconstruction of the vocal tract. Narayanan and Alwan (1995) used MRI to image the vocal tract during the production of /s, sh, f, θ, z, yog/ and found significant interspeaker differences in across all consonant context. In 1996, Story et al (1996), provided both qualitative and quantitative information about vocal tract configurations of 12 vowels, 3 nasal, and 3 plosive sounds. Over the next several decades several researchers investigated vocal tract configurations using volumetric MR images in several sustained voicing tasks (Table 4). Most recently MRI imaging has turned to real-time/dynamic MRI (dMRI) image acquisition of the vocal tract to capture the temporal changes of the speech structures during speech. dMRI can now capture the complex geometry of the vocal tract across a frame sampling rate more reflective of speech/voice frequency (Lingala, et al, 2017; Lim, et al., 2019).

The complexity of the human vocal tract and shape configurations for different sounds, voice qualities, and intensity, etc., and the physical or anatomic variations between people have been significant barriers in MRI imaging for speech. Accurately representing and getting high-resolution imaging of the anterior neck (and the structures contained internally) has been a long-standing issue in the voice literature. A crucial part of any MRI is the MRI radio frequency coils that communicate to send and receive pulse signals. Four different types of coils can be used in

MRI image acquisition; volume, shim, gradient and surface coils. Surface coils, coils that are placed directly on a region/area of interest, allow signals to be received by the surrounding tissue. Moreover, MRI coils and the coils' placement play an invaluable role in MRI image acquisition. Coils are designed to provide high signal sensitivity. Each MRI scanner has numerous built-in coils. For the purposes of this dissertation, the Siemens MAGNETOM Prisma 3T scanner has built-in/default coils in the head and posterior neck. Given our goal of imaging the anterior neck and the known issues with imaging this region, the University of Utah Center for Advanced Imaging Research (UCAIR) developed a close-fitting anterior neck coil designed initially to optimize our imaging. The unique 7-channel neck coil developed by Dr. Rock Hadley and the UCAIR team has been studied extensively in carotid artery, subglottis stenosis, and our pilot work. (Hadley et al., 2005; Beck et al., 2020).

1.9. Specific Aims and Hypotheses

The goal of this study was to describe where and how different SOVT tasks, with and without a target of enhancing vibratory sensations, alter the cross-sectional area(s) and length of the vocal tract in six vocally healthy individuals.

Specific Aims:

1. To compare methodology for measuring the vocal tract cross-sectional areas by measuring a previously analyzed and published imaging set (Story et al., 2008).
   a. Goal: Comparison of area functions of /u/, /i/, and /a/ from published work in Story et al., 2008.
2. To determine if cross-sectional area and vocal tract length change when participants are instructed to target a sensation of vibration at the location of three different lip semi-

occlusions. Area functions (which included cross-sectional area and length) obtained for the targeted tasks were compared to those measured for the untargeted tasks.

a. Hypotheses:

   i. Targeting lip vibrations for (/m/ and /u/) will decrease cross-sectional area measurements of the epilarynx compared to typical, untargeted productions.

   ii. Targeting lip vibrations for /u/ compared to targeting /u/ with a straw will result in the following changes.

      1. Increased vocal tract length due to expansion of oral cavity and lower position of the larynx across all participants.
      2. Increased area in the oral cavity and pharynx.
      3. Decreased area of the epilarynx.

CHAPTER 2: Comparison of Methodology (Aim 1)

Presented in this chapter is a description of the methodology for measuring the vocal tract cross-sectional areas from MR-based image sets. First, the image processing software and measurement techniques are explained and demonstrated using selected image sets collected for this study. As a comparison procedure, the method was applied to three image sets for vowels /u/, /i/, and /a/ collected by Story (2008). The resulting area functions were then compared to the area functions reported in Story (2008).

2.1. Method

2.1.1 Image Management

The image sets for this study each consisted of 60 slices where each slice was 3 mm thick, 151 mm wide, and 220 mm long with an in-slice pixel resolution of 512x352. The image sets were stored in DICOM format and then transferred to Analyze format. The image sets used for comparison with Story (2008) contained slices that were 5 mm thick, 240 mm wide, and 240 mm long with an in-slice resolution of 256x256 pixels.

Software: Mimics Innovation Suite (MIS)

To concatenate all slices, raw DICOM images were inserted into ImageJ. Once a continuous stack of images for each task was created and saved in Analyze format, the images were transferred into Mimics Innovation Suite (MIS) [Materialise Technologielaan 15 3001 Leuven Beligum]. MIS is a commercially available 3D segmentation software that automatically converts images into three planes (axial, transverse, and sagittal) sections. An example is shown in Figure 2.1 using the image set collected for a participant that will be described in a later section producing an /u/ vowel with a straw inserted between the lips (the straw configuration does not appear in the image) The images in Figure 2.1a and 2.1b are inverted due to the order in which MIS read the concatenated stack of images produced by ImageJ; the inversion did not affect the analysis. The axial slice (Figure 2.1c) is the typical orientation where the bottom of the image is anatomically posterior and the top anterior. A "mask" was created for each stack of images using an automated algorithm using a self-selected pixel density (Figure 2.1e). After masking has been applied, MIS automatically creates a "rough" 3D model (Figure 2.1d).

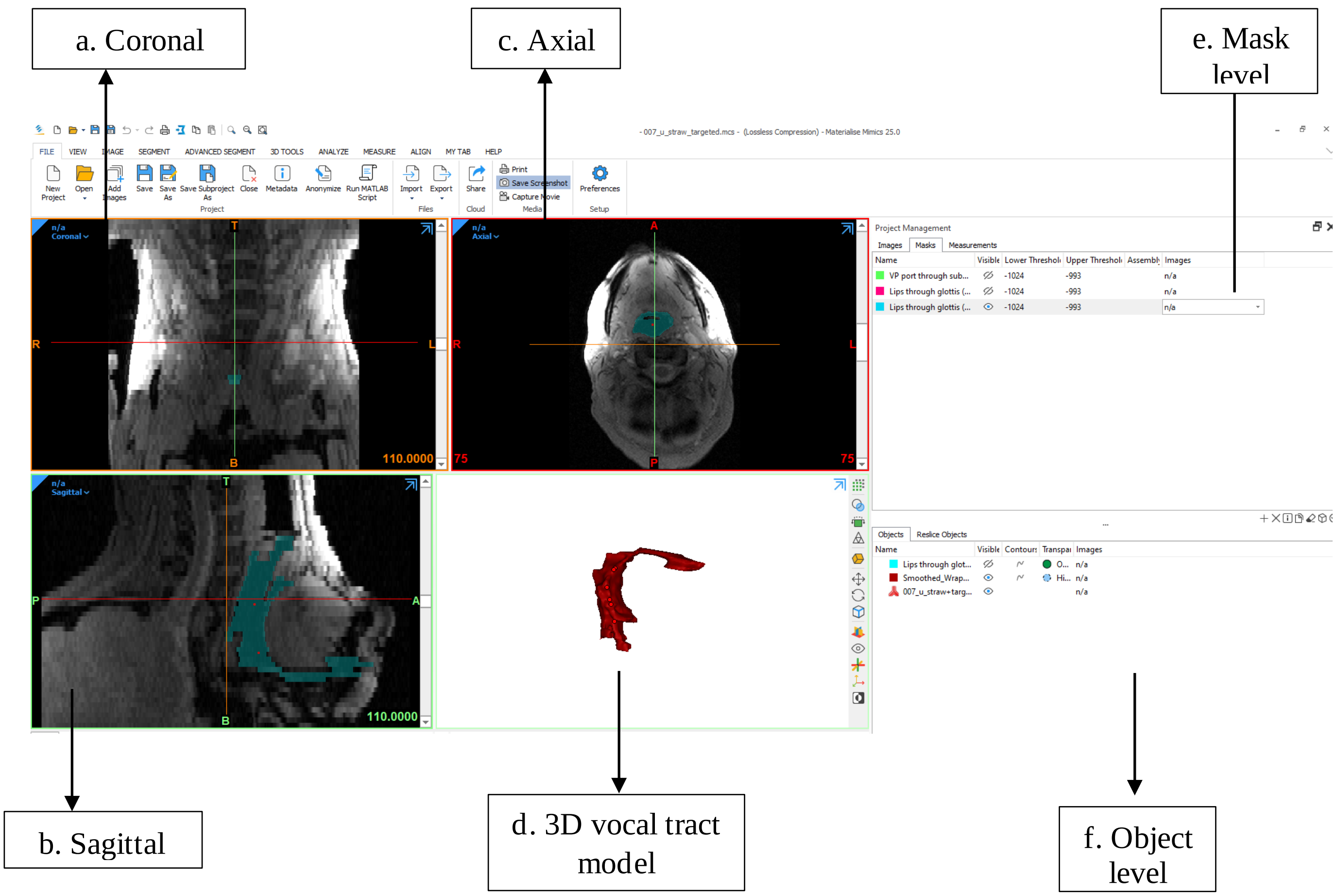


Figure 2.1: Mimics Innovation Suite view following the import of image stack and segmentation. a) coronal plane corresponding to the axial slice. b) sagittal plane corresponding to axial slice. c) axial plane 3mm slice. d) 3D vocal tract model reconstructed from axial segmentation. e) three segmentation "masks." f) 3D objects created from the segmentation masks.

The top text box in the upper right is the "mask" level. The Mimics airway analysis module allows for several mask levels. For this study, three masks were applied. The first was a mask from the velopharyngeal port through the subglottis (Figure 2.1e green box and magnified in Figure 2.2). The mask was then paired down to only include the lips through the glottis (Figure 2.1e and magnified in Figure 2.2 pink box). The final mask is teal and represents the lips through the glottis.

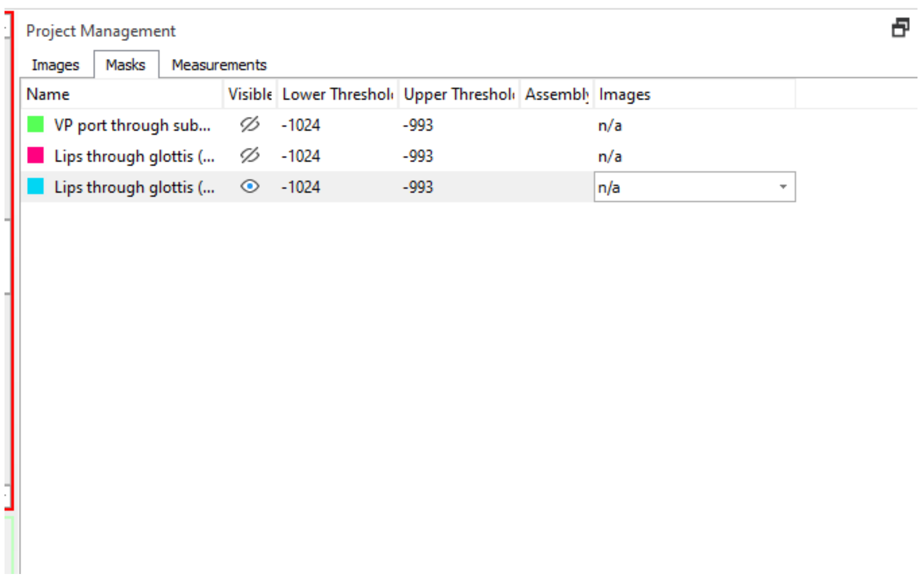
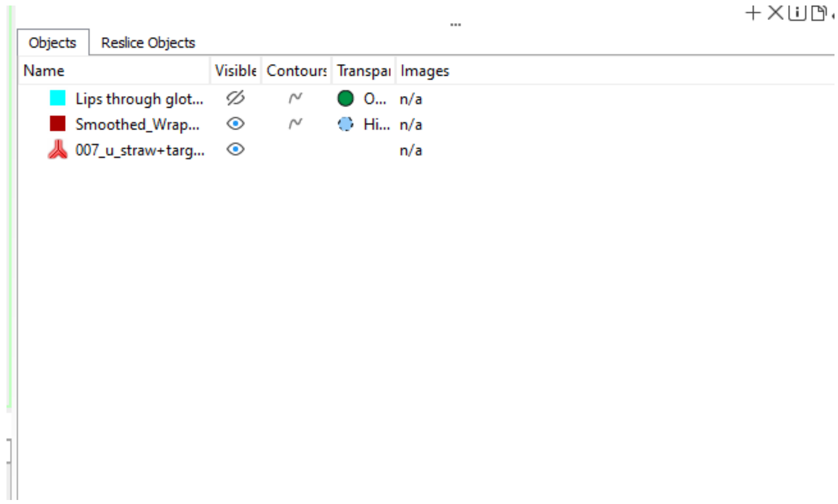


Figure 2.2: Zoomed-in view of the “project management” window showing three unique masks created from the image stack.

Once the mask was “executed,” a 3D model (referred to as an “object” in Mimics) is generated. The object was then additionally smoothed and wrapped (Figure 2.3 “Smoothed_Wrap red box). The process of smoothing and wrapping was automated and eliminated the unsmoothed edges from the mask level. This step allowed for a more accurate visual and anatomical representation of the vocal tract. Finally, after the object (3D image) was smoothed and wrapped, a centerline was fit from the lips through the glottis (Figure 2.4 "07_u_straw +target").



Object window to 3D vocal tract with centerline

Figure 2.3: Zoomed-in view of the "objects" window where the completed 3D model is created.

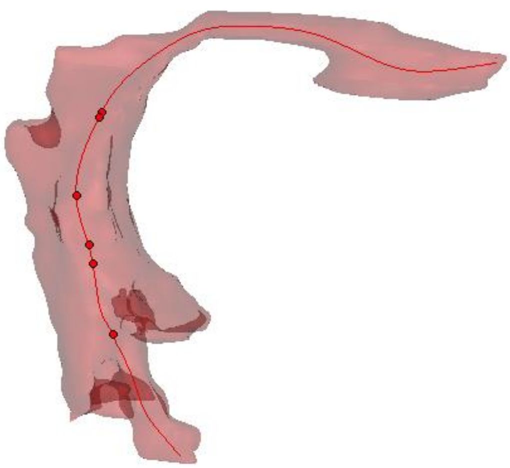


Figure 2.4: 3D vocal tract reconstruction with red centerline fitted (for participant 07 producing an /u/ with a straw at the lips). The red curved line indicates the automated centerline selection, whereas the red dots are selected points on the centerline.

2.1.2. Automatic and Manual Segmentation

The MIS software allows for the automatic segmentation of air from surrounding tissue in each image and then reconstructs the 3D airspace extending from the laryngeal region to the lips. The initial automated threshold boundaries were inflated (red box in the left panel of Figure 2.5) and included pixels outside the anatomical structures (purple around the head and neck image). Therefore, the mask was cropped (yellow box in the middle panel of Figure 2.5), and additional manual cropping and segmentation of the regions of interest were applied (black box in the right panel of Figure 2.5). This process of manual segmentation was completed on each of the 60 slices.

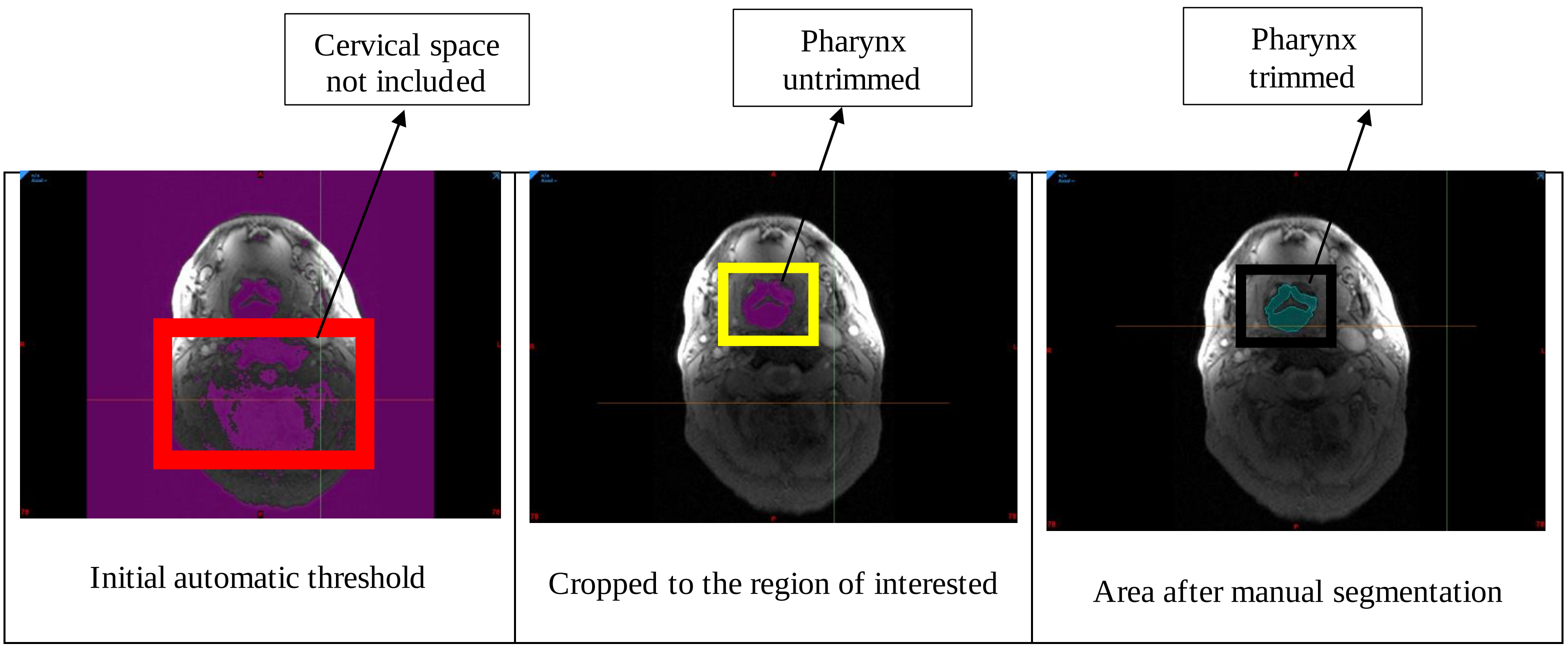


Figure 2.5: Image progression from automatic to manual area segmentation.

The segmentation for each image set was deemed complete by consensus between Drs. Brad Story and Robin Samlan, and the author on appropriately selected air, regions of interest, and anatomical boundaries.

2.2. Length and Area Boundaries

For this project, the air space within the pyriform sinus cavity was not included in the measurements used for the vocal tract. Boundaries of the airspace began at the lips and continued through the oral cavity. The oral cavity only included the air space; therefore, the teeth, and hard and soft palate were excluded from the measurement. The velopharyngeal port was not included, and consequently, no airspace from below the sinus cavity through the uvula was included. The airspace from the uvula extending down through the oropharynx was segmented and included in the area. Next, the space from the oropharynx through the pharynx and down to the aryepiglottic folds was measured. The aryepiglottic tissue, epiglottis, hyo-epiglottic ligament, and thryo-epiglottic ligament was not measured (Figure 2.6).

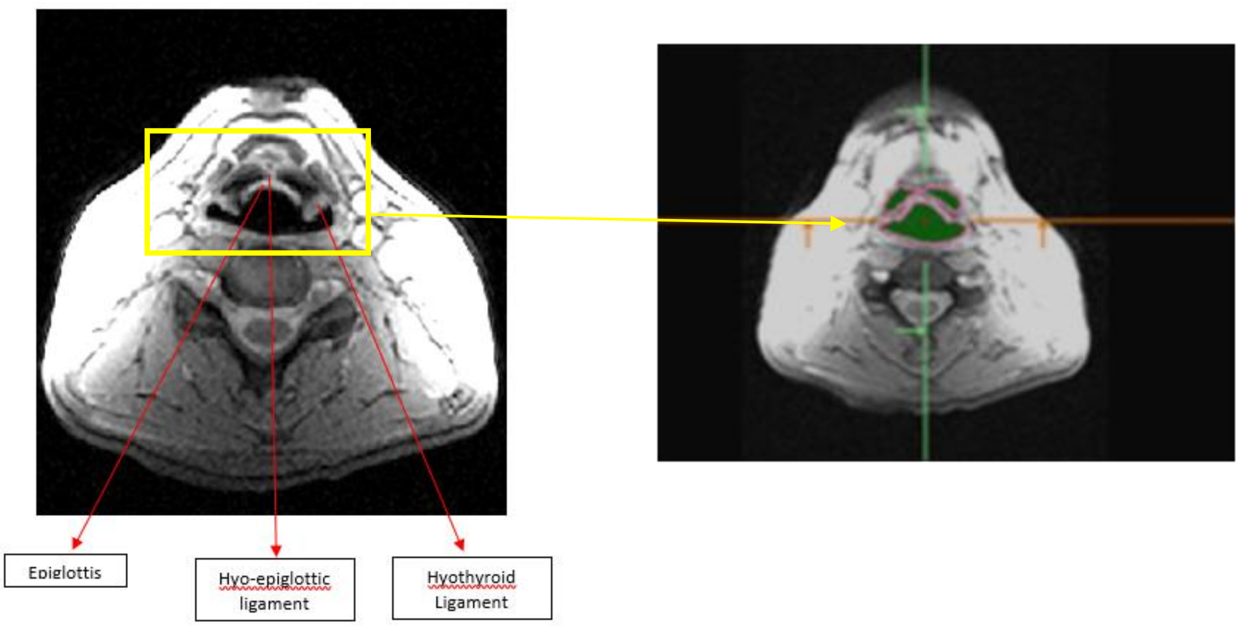


Figure 2.6: a) Structures not included in the measurement indicated with red arrows. The black region outlined in the yellow box is the airspace area b) segmented area, which was seen in the panel on the right as black is now shadowed in green and outlined in pink in the panel on the right.

The epilarynx was defined as the space visible at the laryngeal surface of the epiglottis and extending down to the surface of the vocal folds. The glottis was defined as the 3mm section (one slice thickness) where the vocal folds were visible bilaterally. In the original work by Story et al. (1996), the pyriform sinuses are not included in the main vocal tract area measurements but instead treated as separate airspaces that branch off the vocal tract. The current project followed these recommendations. The subglottis boundaries were those directly below the vocal fold section through the membranous trachea, but the subglottal area was not included in the area functions.

2.3. Cross-Sectional Area and Length

Following the segmentation of each slice, a completed 3D vocal tract was created (Figure 2.7). The next step was to fit a centerline through the 3D model from the glottis to the lips. Given the complex geometry of the vocal tract, the initial automatically-computed centerline required manual editing to ensure accuracy and shape. Along each centerline that passes through the corresponding 3D reconstruction, several red dots (Fig. 2.7) indicate the automatic selection of landmarks central to the model reconstruction. Each red dot corresponds to a data point (i.e., area value) perpendicular to the centerline. Finally, the cross-sectional areas were extracted and exported to text files.

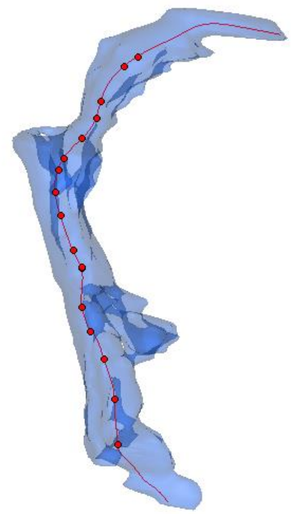


Figure 2.7: 3D vocal tract reconstruction with the centerline for production of an /m/. The red line represents the centerline from lips to glottis. Each red dot indicates the anchor points that the automatic algorithm used to place the centerline, but then was manual manipulated to appropriately align with the geometry of the vocal tract.

In its final form, the vocal tract data consists of a set of cross-sectional areas measured from oblique sections perpendicular to each (x,y,z) coordinate point along the centerline. The 3D coordinates were then used to calculate the Euclidean distance between each point along the centerline such that the set of cross-sectional areas could be represented as an “area function.” That is, the vocal tract configuration for each condition produced by a participant was ultimately characterized by cross-sectional area as a function of the distance from the glottis, indicated mathematically as “A(x)” where A is cross-sectional area and “x” is distance from the glottis. For ease of graphical presentation and for eventual computational modeling, each area function was filtered with a 4-point FIR filter and then resampled to consist of 44 cross-sections; the x-y coordinates along the centerline were also resampled to 44 points to coincide with the area function. The vocal tract length was determined as the distance from the glottis to the lip termination (Figure 2.7).

2.4. Anatomical Boundaries in 3D Model

The oral cavity, pharynx, and epilaryngeal boundaries were defined from the visual 3D reconstruction. Anatomical boundaries were not determined from the individual DICOM images but instead based on the 3D reconstructions or the area function (i.e., distance from glottis). As a general definition, the oral cavity was assumed to extend from the lips to the uvula. The pharynx was the region from below uvula to below the laryngeal surface of the epiglottis (Figure 2.8). Finally, the epilarynx was the region above the vocal folds through the laryngeal surface of the epiglottis. (Figure 2.8).

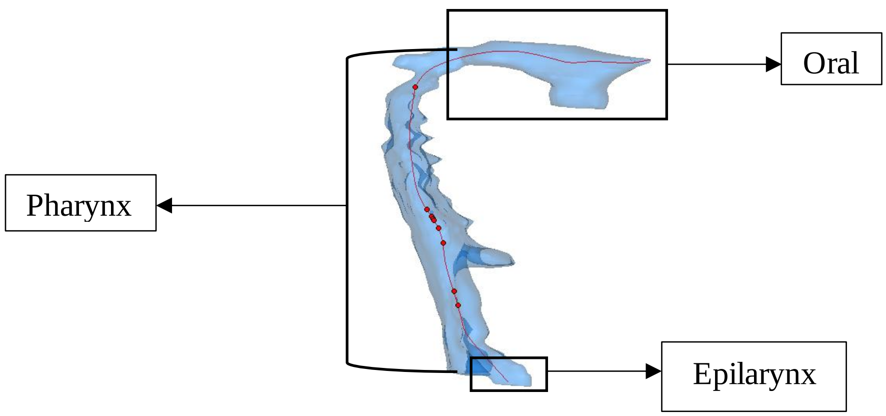


Figure 2.8: 3D reconstruction of vocal tract with gross anatomical regions outlined.

2.5. Results Objective 2

Comparison of MIS Methodology

To ensure the accuracy of the semi-automated segmentation process of MR-based image sets obtained from the participants in this study, MR images from the adult male participant in Story (2008) were segmented based on the method described in the previous sections. The image sets were provided by Dr. Brad Story provided for a /u/, /a/, and /i/ vowels for the male talker in Story (2008).

The results will be presented below with the side-by-side comparison of the 3D reconstructions from each segmentation and reconstruction methodology. The following images will show each vowel's MIS view for the centerline placement. From the centerline, cross-sectional areas were extracted, and the area functions were plotted. The MIS area function was then compared and overlaid on the area function published by Story(2008), represented in blue (Figures 2.10, 2.13, and 2.16), and the MIS 2023 in orange for each vowel.

/u/ vowel

Figure 2.9 shows the 3D reconstructions for the /u/ vowel from Story (2008). The panel on the left is the 3D reconstruction from the original 2008 manuscript, whereas the right panel shows the reconstruction from Mimics Innovation Suite (MIS). The reconstructions are similar, with minimal visual deviations.

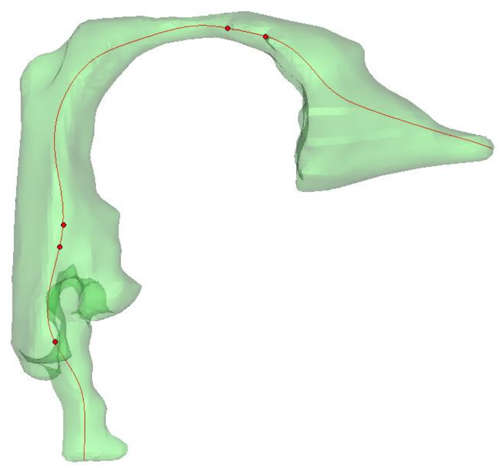
Story 2008 , /u/	MIS 2023_Story /u/ 2008
	

Figure 2.9: 3D model comparison from Story (2008) /u/ vowel reconstruction compared to MIS 2023 3D model reconstruction with the centerline.

Figure 2.10 is the view from MIS for the /u/ vowel. The upper right (coronal plane), upper left axial plane), and lower left (midsagittal) quadrants all represent the 3D reconstruction from one 4mm axial slice. The lower right quadrant is the completed 3D reconstruction from the /u/ vowel, with the centerline placed from the glottis to the lips.

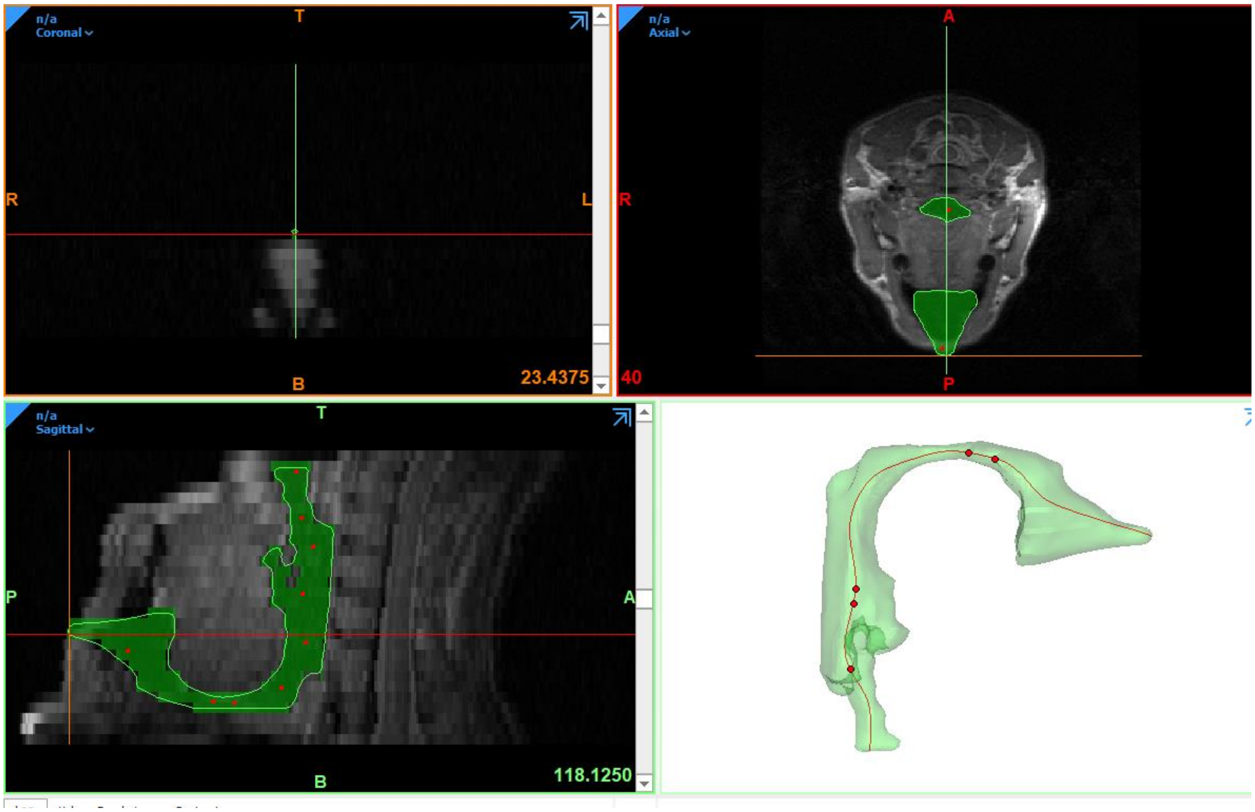


Figure 2.10: Mimics view of the /u/ vowel reconstruction. The upper left quadrant is the coronal plane. The bottom left quadrant is the mid-sagittal plane. The upper right quadrant is the axial plane. The bottom left quadrant is the 3D reconstruction with the red centerline.

Figure 2.11 shows the smooth area functions overlaid for both segmentation methodologies. The epilarynx length was nearly the same, with a slight expansion in the MIS Heller-Stark 2023 area from approximately 2 cm to 5 cm from the glottis. In the 2023 Heller-Stark MIS area function, a subtle shift of expanded regions leftward by ~1 cm can be observed at 4 cm and 15 cm from the glottis. Another slight variation is seen at the lip termination, where a slight shortening of the vocal tract was seen in the Heller-Stark 2023 MIS area function (17.5 cm overall length in orange and 18.2 cm in blue). The deviations across the two area functions are likely due to slight differences in the centerline placement algorithm, segmentation thresholds, exclusion of the teeth from the airspace and the way in which the vocal tract was terminated at the lips.

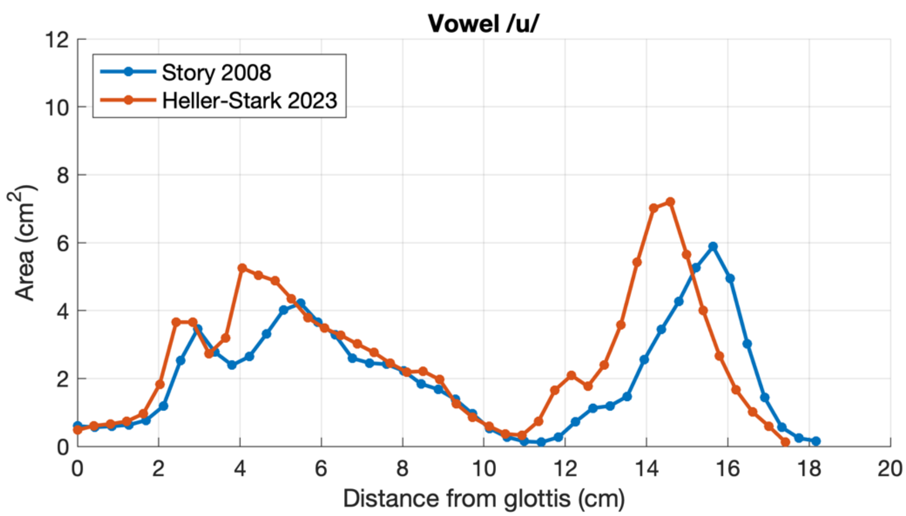


Figure 2.11: Smoothed area function overlaid of original Story 2008 and the MIS for /u/ vowel segmentation comparison.

/a/ vowel

Figure 2.12 shows the 3D reconstructions for the /a/ vowel from Story (2008). The panel on the left is the 3D reconstruction from the original 2008 manuscript, whereas the right panel shows the reconstruction from Mimics Innovation Suite (MIS). The reconstructions are similar, with minimal visual deviations.

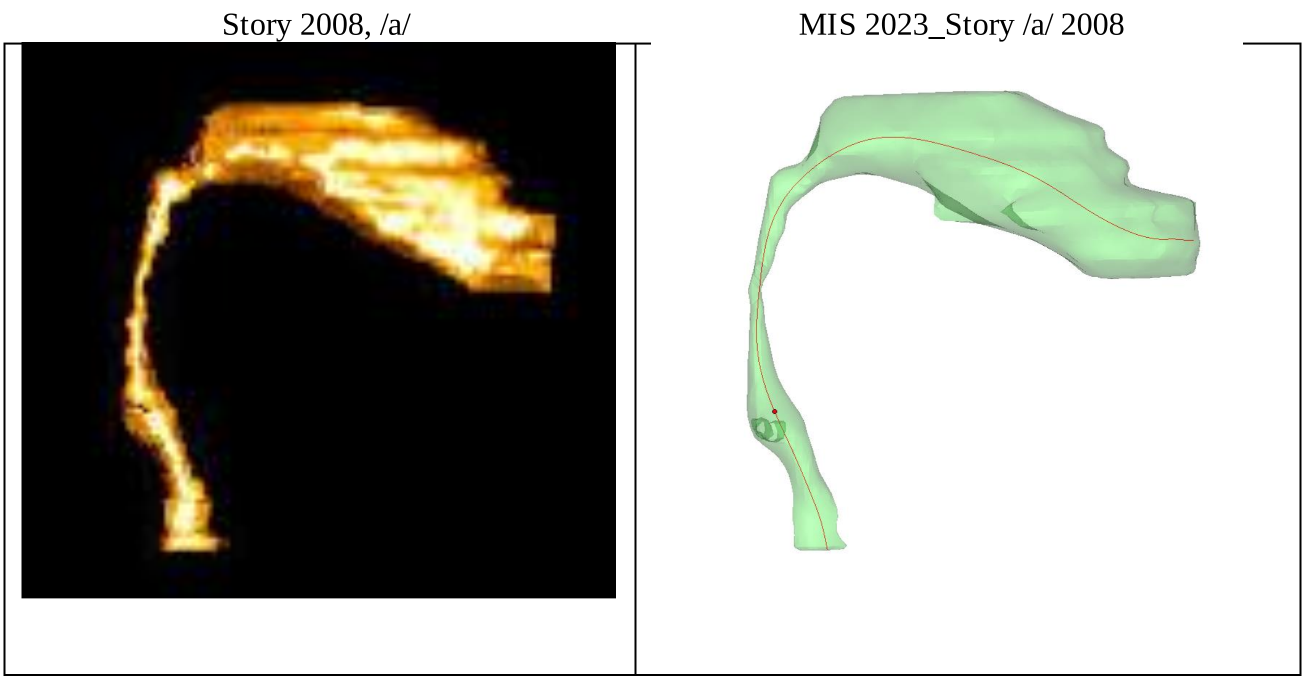


Figure 2.12: 3D model comparison from Story (2008) /a/ vowel reconstruction compared to MIS 2023 3D model reconstruction with the centerline.

Figure 2.13 is the view from MIS. The upper right (coronal plane), upper left axial plane), and lower left (midsagittal) quadrants all represent the 3D reconstruction from one 4mm axial slice. The lower right quadrant is the completed 3D reconstruction from the /a/ vowel, with the centerline placed from the glottis to the lips.

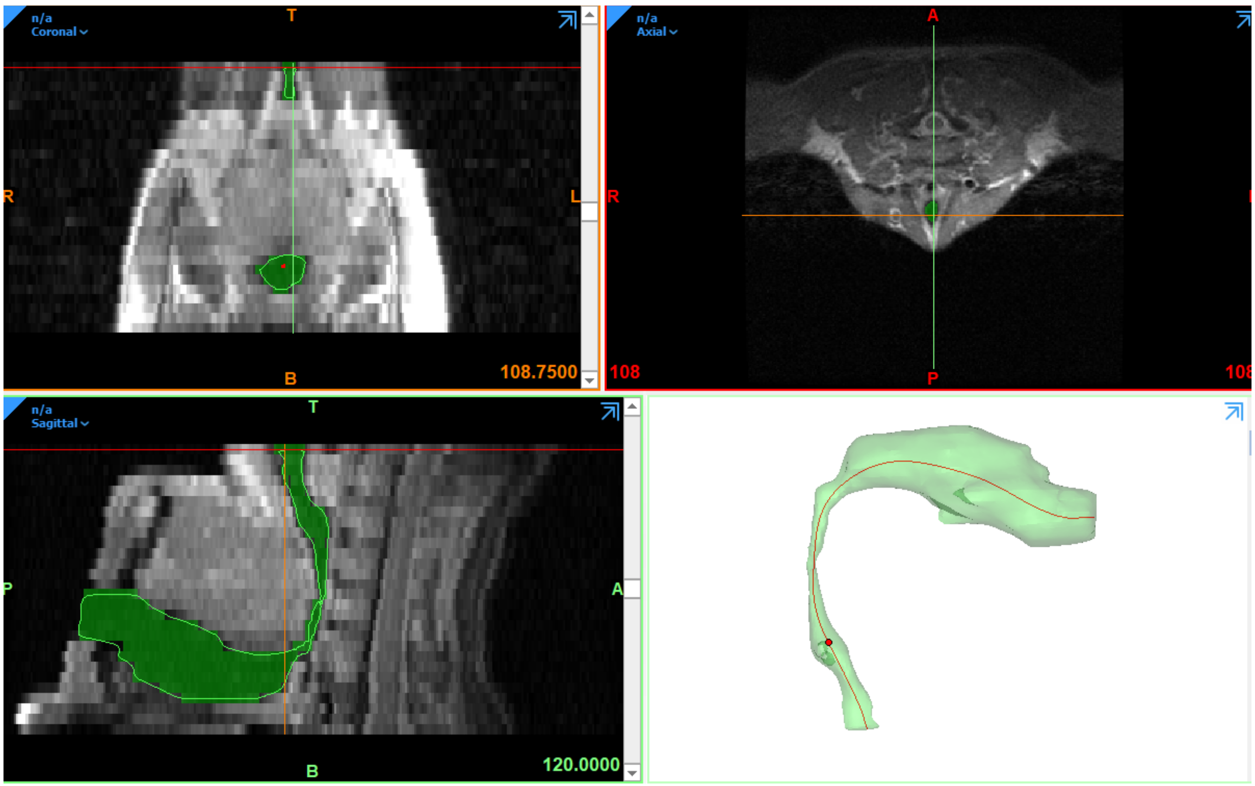


Figure 2.13: Mimics view of the /a/ vowel reconstruction. The upper left quadrant is the coronal plane. The bottom left quadrant is the mid-sagittal plane. The upper right quadrant is the axial plane. The bottom left quadrant is the 3D reconstruction with the red centerline.

Figure 2.14 shows the smooth area functions overlaid for both segmentation methodologies. The epilarynx length was identical, with a slight area reduction in the MIS Heller-Stark 2023 area function, from approximately 2 cm to 5 cm from the glottis. In the 2023 Heller-Stark MIS, a minimal expansion was seen at 2.8 cm, with a mostly 1:1 comparison until approximately 12 cm from the glottis. An approximately 2.5 cm^3 expansion was seen from 12 cm to 15 cm from the glottis. This slightly inflated area in the oral cavity from the 2023 Heller-Stark reconstruction was, again, due to how the teeth were handled in the MIS reconstructions. Another slight variation is seen at the lip termination. The Story (2008) methodology employed a closed, fixed boundary for the lip termination. The same technique was not used in the current work.

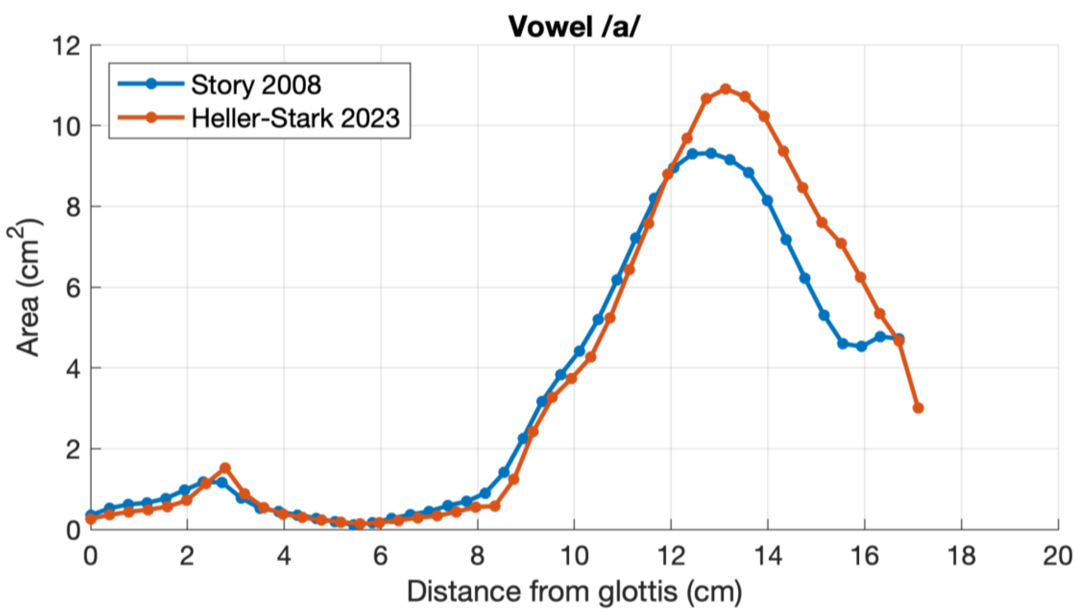


Figure 2.14: Smoothed area function overlaid of original Story (2008) and the MIS for /a/ vowel segmentation comparison.

/i/ vowel

Figure 2.15 shows the 3D reconstructions for the /i/ vowel from Story (2008). The panel on the left is the 3D reconstruction from the original 2008 manuscript, whereas the right panel shows the reconstruction from Mimics Innovation Suite (MIS). The reconstructions are similar, with minimal visual deviations.

Story 2008, /i/ | MIS 2023_Story /i/ 2008

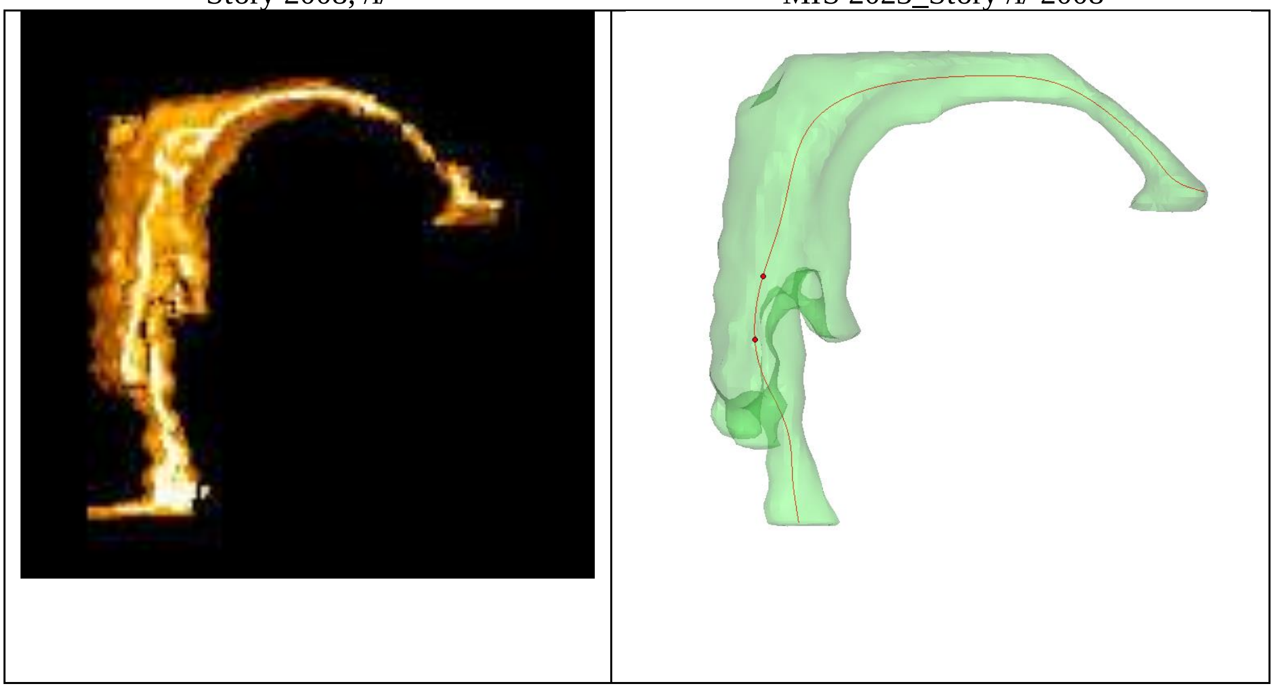


Figure 2.15: 3D model comparison from Story (2008) /i/ vowel reconstruction compared to MIS 2023 3D model reconstruction with the centerline.

Figure 2.16 is the view from MIS for the /i/ vowel. The upper right (coronal plane), upper left axial plane), and lower left (midsagittal) quadrants all represent the 3D reconstruction from one 4mm axial slice. The lower right quadrant is the completed 3D reconstruction from the /i/ vowel, with the centerline placed from the glottis to the lips.

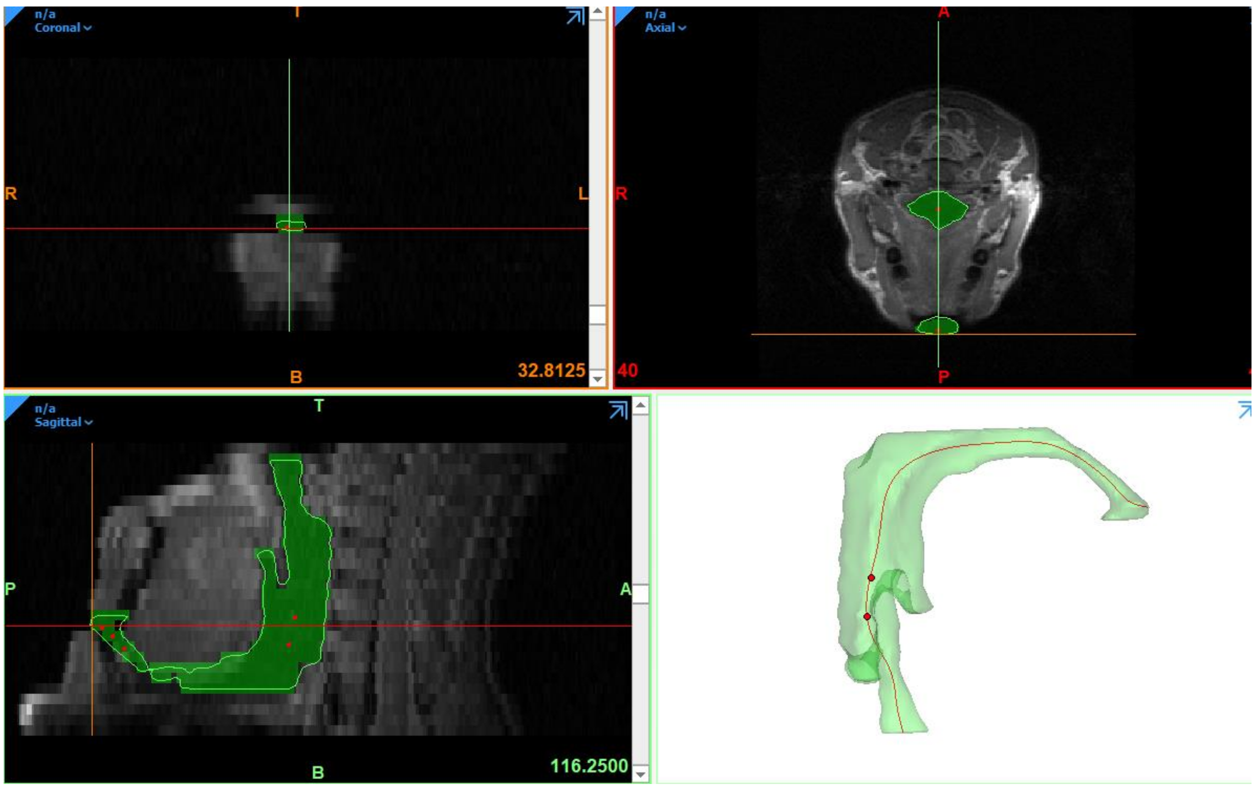


Figure 2.16: Mimics view of the /i/ vowel reconstruction. The upper left quadrant is the coronal plane. The bottom left quadrant is the mid-sagittal plane. The upper right quadrant is the axial plane. The bottom left quadrant is the 3D reconstruction with the red centerline.

Figure 2.17 shows the smooth area functions overlaid for both segmentation methodologies. The epilarynx length and area are identical. Almost identical areas are seen from 0 cm to 14.2 cm from the glottis. Again, a slight expansion of the oral cavity at approximately 15 cm from the glottis is seen. This variation in oral cavity area and lip termination is due to the segmentation variations between Story (2008) and the current study.

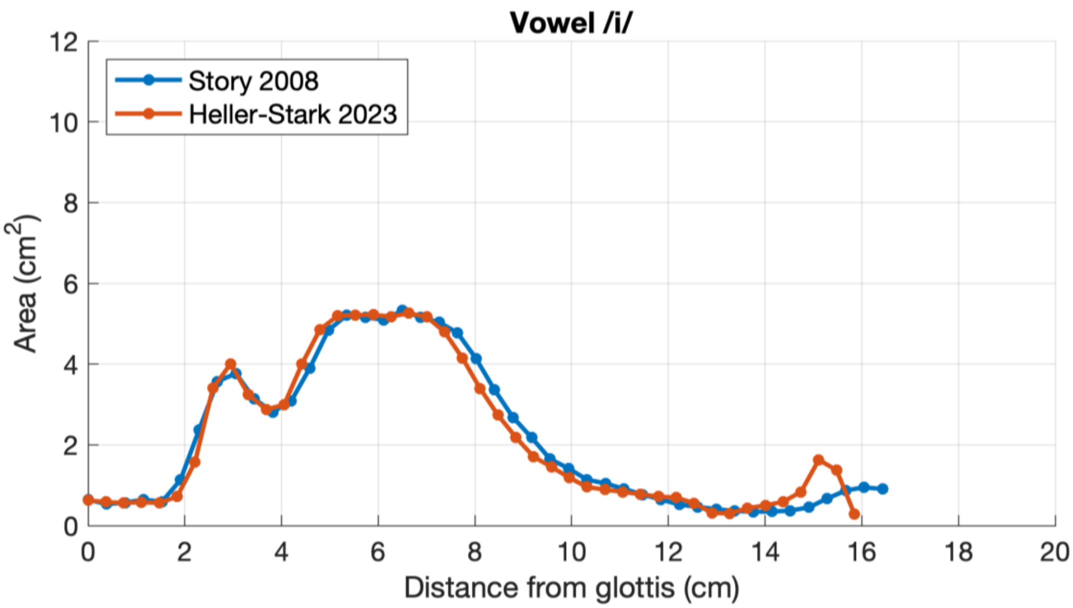


Figure 2.17: Smoothed area function overlaid of original Story 2008 and the MIS for /i/ vowel segmentation comparison.

2.6. Discussion Objective 1

The current methodology for segmentation and obtaining the resulting area functions were compared to Story (2008). Almost identical areas were seen across all three vowels from the epilarynx to the starting boundaries of the oral cavity. There were slight variations and discrepancies in regions within the oral cavity. These deviations resulted from methodological differences in how the teeth segmentation and lip termination boundaries were set. The Story (2008) method for teeth casting and setting a closed boundary at the lips were more accurate.

Although different, a close approximation of the area was still obtained with the new segmentation modality.

CHAPTER 3: Methods for Aim 2

3.1. Participants

Seven vocally healthy individuals were recruited from the University of Utah clinical trials posting or through word of mouth for participation in the study. One male and one female from each of our three age groups (young=18-40; middle=41-64; aging (65 years and older) were recruited. Eligibility for the study was determined based on the following inclusion criteria: self-reported normal voice functioning as determined by the Voice Handicap Index (VHI; Jacobson et al., 1997), no history of a voice problem lasting more than four weeks, and voice quality in the normal range based on auditory-perceptual screen by the author. Exclusion criteria included claustrophobia, physical deficits that prevented participants from being in an MRI scanner, currently pregnant or nursing, or any formal voice training.

Participant 05 underwent the initial session only to be excluded from the MRI session due to a stapes implant. All participants scored within normal limits on three voice screening tools (VHI, laryngoscopy, and AVQI) (Table 5). Participant 01's data were not analyzed due to an error in the MRI data acquisition.

Table 5: Demographic information for each participant

Participant	Age	Sex	VHI	AVQI	Stroboscopic findings
01	26	F	10	2.7	No pertinent deviations
02	67	M	5	3.34	Mild bowing (judged by laryngologist to be within normal limits for age)
03	49	F	7	4.02	No pertinent deviations
04	35	M	1	3.24	Capillary ectasia (judged by laryngologist to not impact vibratory features)
06	60	M	7	4.19	No pertinent deviations
07	72	F	2	2.61	No pertinent deviations

An Institutional Review Board (IRB) application was submitted to the University of Utah and the University of Arizona. IRB approval from both institutions was obtained (IRB_00149122 University of Utah; Study_00000824 University of Arizona). Volunteers participated in two different sessions across one to two weeks. Once identified, participants completed an MRI screening questionnaire via email (Appendix A) or over the phone.

Recruitment and screening telephone call

Participants were recruited via word-of-mouth and emails sent by study investigators and research assistants from the University of Utah student body and the community. Participants were contacted initially by telephone. The initial phone call served to determine eligibility for MRI and self-report of no previous history of voice problems, voice therapy, or formal voice training.

All eligible participants were evaluated at the Voice, Airway, Swallowing Translational (VAST) Research Lab run by Dr. Julie Barkmeier-Kraemer at the University of Utah. Informed consent was obtained at the beginning of session one. Participants were all informed of the study's aims, and no details regarding the nature of the study were withheld.

<u>3.2. Session One (Voice Evaluation and MRI training)</u>

After providing informed consent, participants underwent a thorough voice evaluation following standard procedures (Patel et al., 2018). The evaluation was used to determine whether the participant met the requirement of normal voice and laryngeal functioning. Participants also completed the following: MRI screening questionnaire (Appendix B), Voice Handicap Index (VHI) (Appendix C), laryngostroboscopy (Appendix D), brief perceptual voice assessment (Appendix E and F), and acoustic recording.

All acoustic recordings were conducted in a sound booth with a AKG head-mounted condenser microphone (model C520) placed 45 degrees to the side of the mouth at a distance of 4 cm. A battery of voicing tasks was recorded in both the upright and supine positions. All participants were within normal limits for all three domains of the voice evaluation. Participants 02 and 04 presented with irregularities on laryngoscopic examination and the consulting physician confirmed these irregularities to be normal anatomical variants that were consistent with aging or determined not to impact the vibratory features of the vocal folds.

The VHI is a psychometrically robust quality of life questionnaire used frequently in voice clinics. A cut-off score of less than 19 is typically used to indicate no voice concerns. (Jacobson et al., 1997).

Mock MRI training session

Following the voice evaluation, the participants were familiarized with the testing procedures and requirements for the MRI scanner session (Appendix G). The participants were verbally instructed on how to perform each and for how long they would perform each task. Finally, they had several opportunities to practice the tasks with the author. During the training session, it was described to the participants that targeted productions were intended to use a vibratory focus at the point of the semi-occlusion, and untargeted productions were to be produced "as if talking to a friend" with no particular emphasis on vibration. Participants then practiced targeted and untargeted productions of /z/ and /v/. They were asked to switch between their targeted production which had a vibratory focus and untargeted production that did not have a specific cue to feel vibration, to feel the contrast between what a vibratory-focused production was versus no vibratory focus. Participants were asked to describe where and how they felt differences in the sound productions. When the clinician, AS (the author), did not perceive an increase in vibratory focus during the targeted tasks, participants were given cueing to increase the feeling or sensation of vibratory energy at the teeth (/z/) or lips (/v/).

Participants were then walked to the MRI suite and familiarized with the radiology space. All procedural aspects of the MRI, including the head mount, neck coil, and MRI equipment, were discussed. Following completion of the radiology documentation, the participants were placed into a mock MRI tube. The participants were supine in the mock tube, with headphones playing MRI scanner noise. Participants were instructed to perform the practice voicing tasks they had previously learned. A sound level meter was used to measure the relative loudness during each production to ensure vocal intensity did not exceed 70 dB SPL. Verbal feedback was provided on the targeted intensity level (i.e., comfortable loudness) based on the SPL and accurate vowel and consonant production. No additional cueing about resonance was provided during the mock tube trials. The mock tube trial was approximately 15 minutes, depending on the amount of prompting needed.

3.3. Session Two (MRI)

Participants spent approximately one hour in the MRI scanner. They were refamiliarized with what was expected of them in the MRI scanner and informed of the sounds they would be producing during the scanner. Participants were then fitted in one of three sizes of neck coils designed by Dr. Rock Hadley at the Utah Center for Advanced Imaging Research (UCAIR) (Beck et al., 2020). The UCAIR proprietary seven-channel neck coil was designed to accommodate the human neck's complex geometry and varying shapes and sizes. The neck coil provides high signal sensitivity (i.e., improved signal-to-noise ratio (SNR)) (Beck et al., 2017). The neck coil was placed around the front of the participant's neck and was worn for the entirety of the MRI scanning. The participant's head was placed into a head mount with a strap across the forehead. Headphones were placed bilaterally into the ears, and then pads were placed on either side of the head. Once the radiology technician had the participant appropriately in the scanner, the author instructed the participant on the placement of the straw for the /u/ with straw productions. Specifically, they were asked to limit their hand movement as they placed the straw into the mouth over the head coil. They were asked to hold the straw between their lips and place their hands back down by their side. Additionally, they were instructed to breathe around the straw or through their nose to limit additional respiratory movement. The participants were informed there was a bidirectional microphone so they could speak to the study team and hear the study team and the directions. Testing protocol was followed as detailed in Appendix G.

Figure 3.1: SIMENS T3 MRI bore with participant.

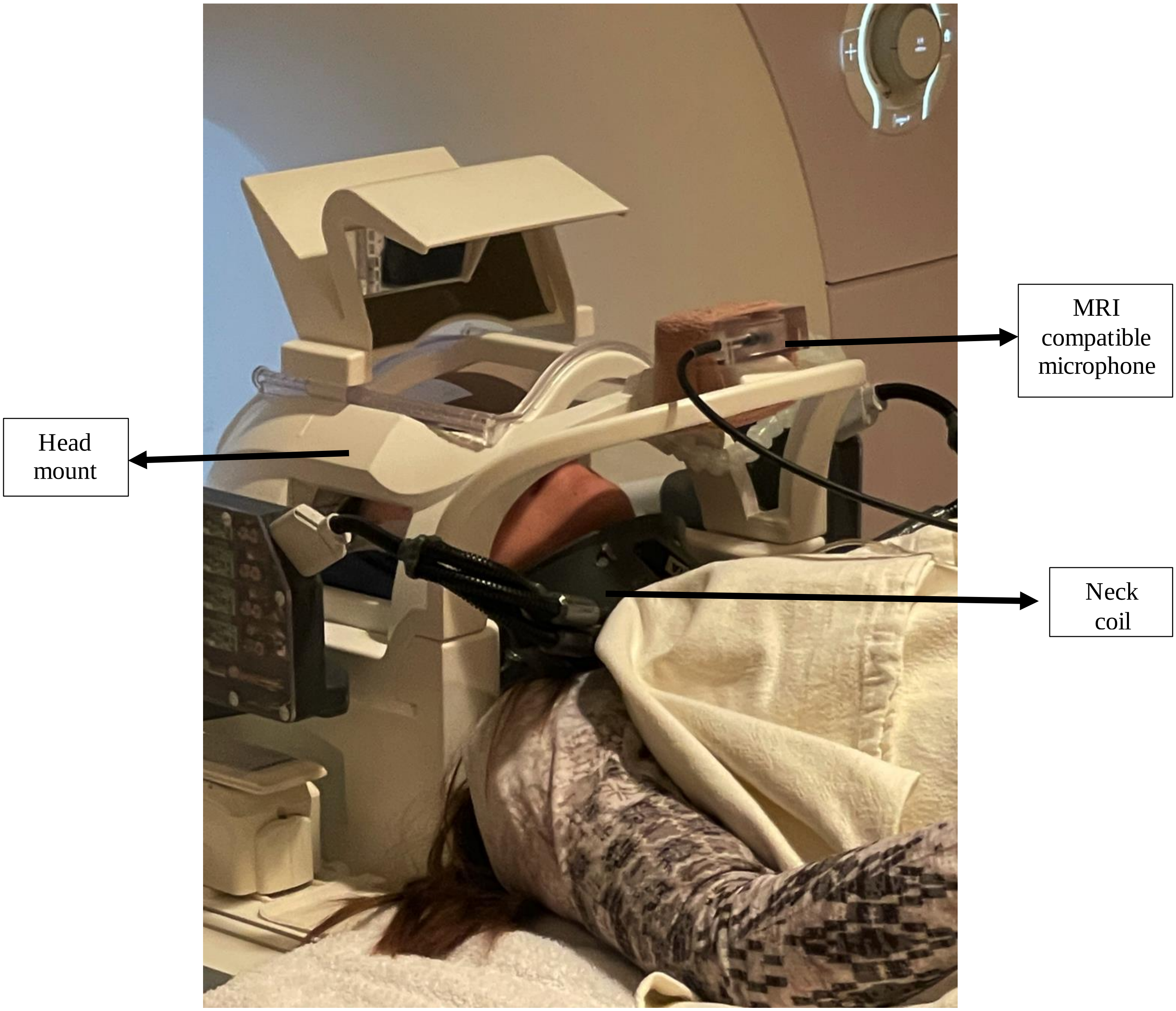


Figure 3.2: MRI scanner session 2 with participant fitted with head mount, MRI microphone and 7-channel neck coil.

Table 6: Summary of MRI tasks with scanning duration. Task order was randomized at the time of the MRI scan.

Task	Trials
/a/	6 repetitions (15 seconds each)
/i/	6 repetitions (15 seconds each)
/ɑe/	6 repetitions (15 seconds each)
/u/	6 repetitions (15 seconds each)
/m/ untargeted	6 repetitions (15 seconds each)
/m/ targeted	6 repetitions (15 seconds each)
/u/ targeted	6 repetitions (15 seconds each)
/u/ with straw untargeted	6 repetitions (15 seconds each)
/u/ with straw targeted	6 repetitions (15 seconds each)
Total:	42 repetitions (~810 seconds of vocalization= 13.5 minutes of voicing)

3.4. MRI Acquisition

MRI images were acquired using a Siemens MAGNETOM Prisma 3T scanner. MRI data were acquired with a susceptibility-weighted gradient-echo (GRE) sequence (field-of-view 22 cm, matrix 64x64, repetition time TR=7.7sec, echo time TE=3.04ms, slice thickness 3mm, FoV phase 68.8%). A phase resolution of 90% was used all additionally sequencing features can be found in Appendix I. All nine tasks were randomized for order of production for each participant (Table 5). Ten slices (3mm each) were acquired during each repetition time (15 seconds). Due to the necessity of scanning the entire vocal tract during phonation, participants were asked to repeat each task six times for a total of 50-60 slices from the oral cavity to the subglottis depending on the person's vocal tract length. Distortions caused by variations in magnetic susceptibility were removed with a Siemens built in distortion correction (2D) sequence using an elliptical filter. An MRI-compatible microphone was placed for each participant so that a real-time audio sample of each person could be acquired.

Figure 3.3: Head mount with MRI compatible microphone

CHAPTER 4: RESULTS

Objective 2: To determine if cross-sectional area and vocal tract length changed during targeted versus untargeted tasks.

The goal of objective two was to determine if and where a vibratory-focused cueing (targeted production) during semi-occluded vocal tract productions /u/, /u/ with straw, and /m/ influenced cross-sectional areas and length of the vocal tract.

4.1. Missing Data

Participant 01’s images were not measured or included in the study due to initial image acquisition errors. Table 7 shows the missing data for individual participants due to image acquisition errors during the MRI scanner session. Participant 02's /u/ straw targeted condition had ten missing slices from the uvula to the epiglottis. Participant 03's /m/ targeted condition had ten missing slices from the pharynx to the epilarynx. Participant 04 had ten missing slices in the /u/ straw targeted condition.

Table 7: Missing raw DICOM images. Gray shaded cells represent complete data sets in the DICOM images. Cells shaded blue are conditions where 10 slices were missing at the location indicated.

Participant	**/u/ straw untargeted**	**/u/ straw targeted**	**/u/ untargeted**	**/u/ targeted**	**/m/ untargeted**	**/m/ targeted**
02		Uvula to epiglottis				
03						Pharynx to epilarynx
04		Oral cavity				
05						
07						

4.2. Format of Results

The results of the MRI analysis are presented below in a standard format for each participant's data. These images were created using the processes detailed in the methodology. In brief, they represent three comparisons (no vibratory focus to vibratory focus from /u/ through a straw, no vibratory focus to vibratory focus from sustained /u/, and no vibratory focus to vibratory focus for sustained /m/). Additionally, the images are presented for each participant using a set of three figures, descriptions of key features of the data, and a table of cross-sectional area functions. The first figure presented for each participant is a 3D model reconstruction from the MIS.

The second figure for each case is a pseudo-mid-sagittal plot in which the equivalent diameter of each cross-sectional area is plotted on a perpendicular projection of the (x,y) coordinates of a specific centerline (cf., Story, Titze, and Hoffman, 2001, Figs 1-2 for an explanation). The x-axis represents the anteroposterior plane, with the corresponding inferior-superior plane area represented on the y-axis, and the inner and outer profiles connect the equivalent diameter projections. The purpose of the pseudo-midsagittal plots is to provide an anatomically intuitive representation of the cross-sectional variation of each vocal tract shape.

The third figure is the extracted smoothed area function, A(x). In these figures, the x-axis is the distance from the glottis (0 cm) to the lip termination, and the y-axis is the cross-sectional area (in cm^2). The pseudo-midsagittal plots and corresponding area functions are coded so that blue lines or images always represent the condition of "no vibratory focus production," and the red lines or figures represent the "vibratory focus production." All raw area functions for each segment in the smoothed area function figure are provided in a table at the end of each participant's data. The raw area function data based on each smoothed area function plot will be presented at the end of each participant's data. The raw area functions represent the numeric value in cm^2 of the corresponding section number. In these tables, the rows are the individual segments where segment 1 represents the glottis, and section 44 is the lip termination. Section 45 is the section length. The columns are the raw areas for /u/ untargeted, /u/ with targeted, /u/ with straw untargeted, and /u/ with straw targeted. The epilaryngeal area will be referred to as A_{epi}, the pharyngeal area as A_{ph}, and the oral cavity area as A_o.

Table 8: Variables of Objective Two

<table>
<tr><th colspan="2">Lip Occlusions</th><th>Targeting</th></tr>
<tr><td rowspan="4">Partially narrowed</td><td rowspan="2">/u/</td><td>No vibratory focus</td></tr>
<tr><td>Vibratory focus</td></tr>
<tr><td rowspan="2">/u/ with straw</td><td>No vibratory focus</td></tr>
<tr><td>Vibratory focus</td></tr>
<tr><td rowspan="2">Complete</td><td rowspan="2">/m/</td><td>No vibratory focus</td></tr>
<tr><td>Vibratory focus</td></tr>
</table>

4.3. Participant Data

4.3.1. Participant 02

Participant 02 /u/ straw

Figure 4.1 shows the 3D reconstructions for Participant 02 during the production of an /u/ with a straw placed at the lips. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the right panel shows the upper portion of the vocal tract shifted in the anterior direction relative to the lower portion, clearly a physiologically impossible configuration. This resulted from missing data and a slight positional change in the participant during the MR scan. The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analysis were carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

02 u straw untargeted | 02 u straw targeted

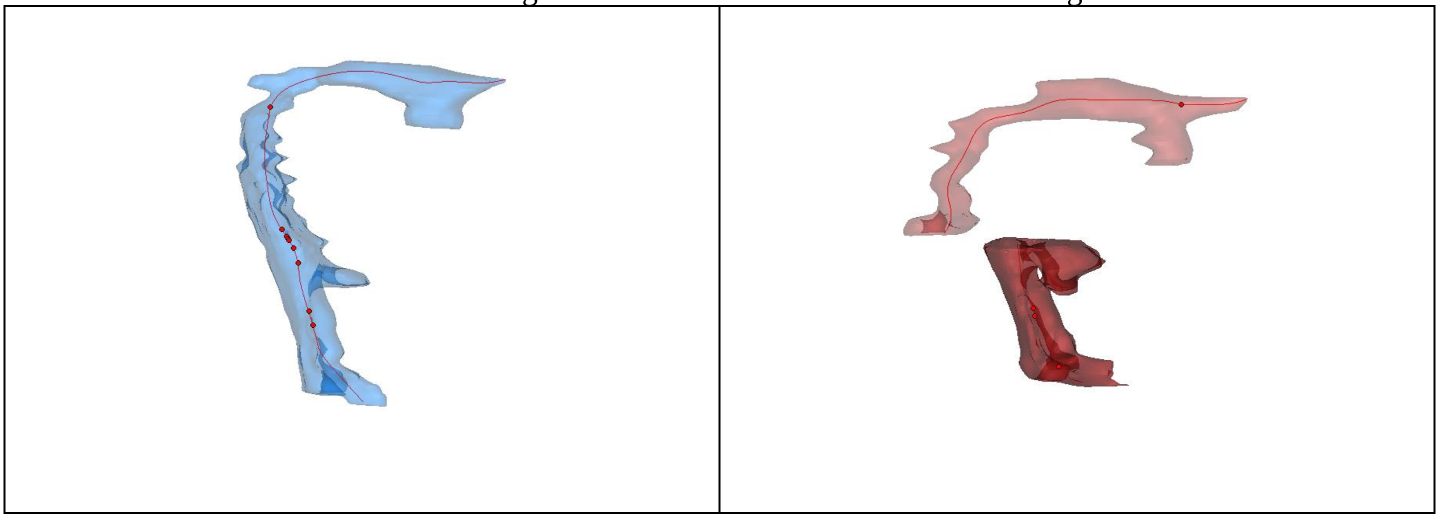


Figure 4.1: 3D model reconstructions for Participant 02 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). 10 slices are missing from the DICOM images from the uvula to the epiglottis and appear as a gap in the 3D model for the /u/ straw targeted condition.

In Figure 4.1, due to the missing data in the /u/ straw targeted condition, the pseudo-midsagittal plot will appear different than in the previous figure (Figure 4.2). In figure 4.2 the untargeted /u/ straw condition in the left panel, there is evidence of only a short epilaryngeal section which quickly opens into the lower pharynx. The vocal tract is constricted at about 10 cm from the glottis and then widens in the oral cavity before constricting around the straw at the lips. In the targeted condition, there is almost no epilaryngeal section that opens into a wider pharynx from approximately 2 cm to 4 cm from the glottis. Given the missing data from the uvula to the epiglottis, an interpolation of these sections was completed. The interpolation resulted in a narrowed portion from approximately 4 cm to 8 cm. This constriction was not consistent with the untargeted condition and was likely the result of the interpolation and not the actual anatomy.

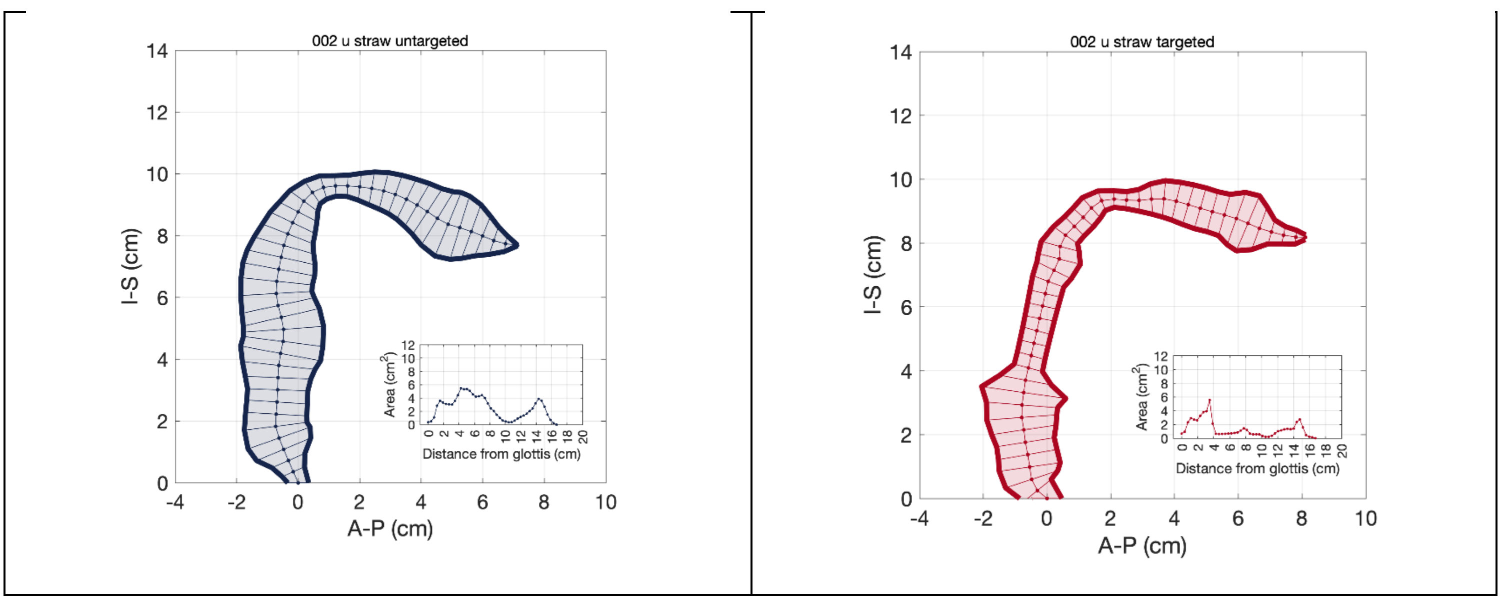


Figure 4.2: Pseudo mid-sagittal plots for Participant 02 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). /u/ targeted represents an interpolation of the vocal tract from the uvula to the epiglottis, given the missing data from the raw DICOM.

Area comparison of /u/ with straw (Participant 02)

Figure 4.3 graphically represents the area function of the /u/ straw tasks for participant 02 with and without a targeted condition. Above, the green areas overlaid on the smoothed area function represent approximate boundaries of the epilarynx, pharynx, and oral cavity.

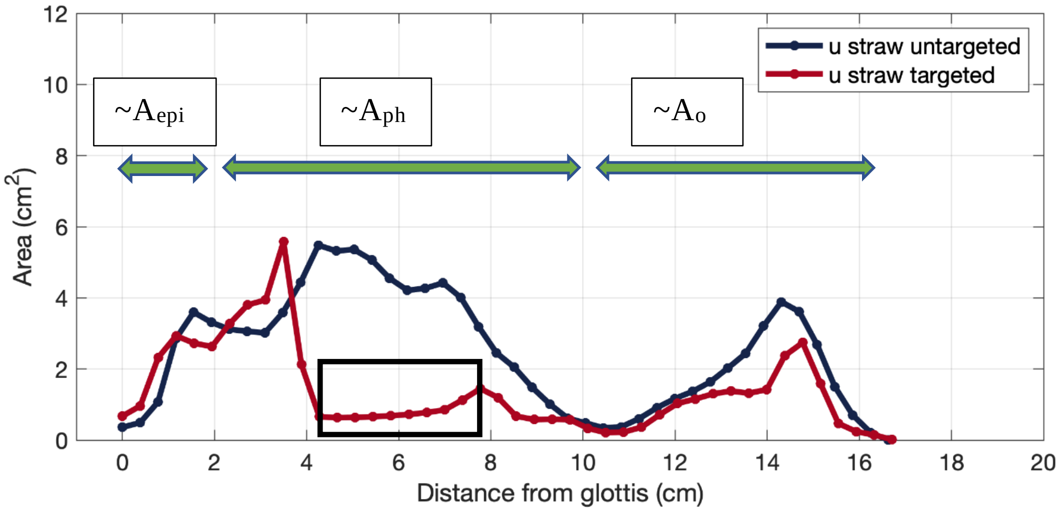


Figure 4.3: Smoothed area function for Participant 02 in the /u/ straw untargeted and targeted conditions. Green arrows represent an approximate region of the A_{epi}, A_{ph}, and A_{o}.

Epilarynx

A target of vibratory focus on an /u/ with the straw resulted in a relatively short but similar length epilarynx (~0 to .3 cm from glottis) with a larger area than the untargeted production (Figure 3.4 and 3.5). Overall, the epilaryngeal area (A_{epi}) of the targeted production was approximately two times that of the untargeted condition. However, the areas were considerably small in both cases, 0.96 cm^2 and 0.49 cm^2, respectively.

Pharynx

Determining pharyngeal area (A_{ph}) changes with targeting was difficult for the /u/ straw condition produced by participant 02 because ten slices were missing (Fig 4.3 black box) from the MRI data positioned approximately 4.2 cm from the glottis to about 8 cm from the glottis. To estimate the missing data, an interpolated version was created in the pseudo-mid-sagittal plots (Figure 4.2), which created a narrowing from 4.2 cm to about 7 cm from the glottis. It is unlikely that the narrowing observed in the pseudo-mid-sagittal plot was the trend for this participant, given that this narrowing was not seen in any other of 02's conditions. Instead, the expansion seen at 3.8 cm from the glottis is likely a more accurate representation of the pharynx. The area of the pharynx would likely be larger in the targeted condition if the area at 3.8 cm from the glottis were used to make inferences about the overall A_{ph}.

Oral Cavity

In both the untargeted and targeted /u/ with a straw, the velar constriction extends from about 10-11 cm from the glottis. Beyond this point, toward the lips, the cross-sectional area increases into an expansion for both conditions, but the untargeted case achieves a greater peak value than the untargeted case.

<u>Participant 02 /u/</u>

Figure 4.4 shows the 3D reconstructions for Participant 02 during the production of an /u/ with a straw placed at the lips. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus").

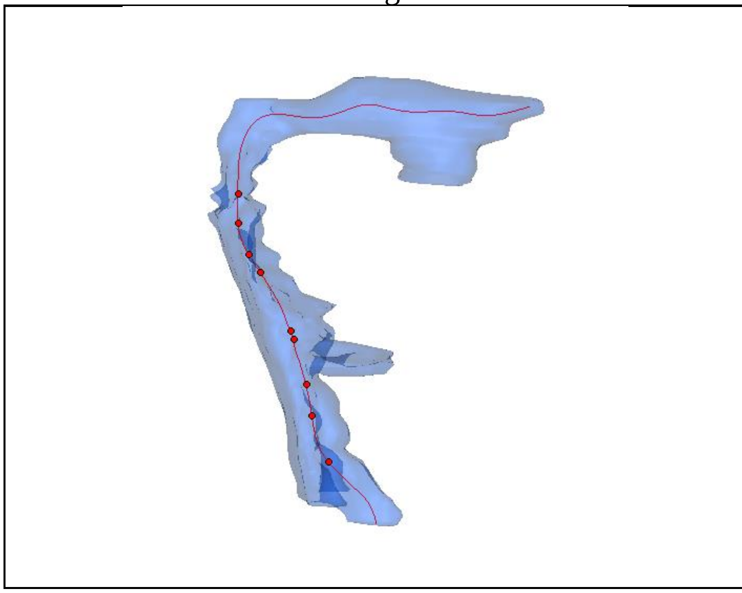
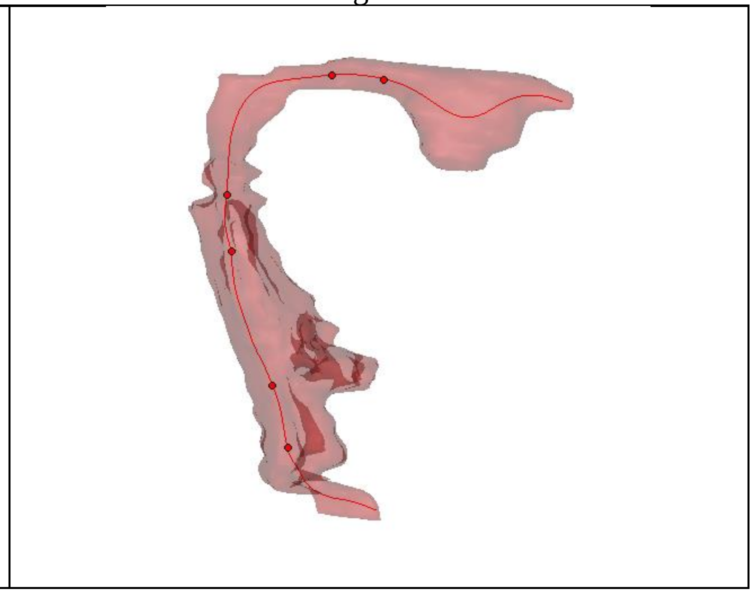
02 u targeted	02 u untargeted
	

Figure 4.4: 3D model reconstructions for Participant 02 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right).

Figure 4.5 is the pseudo-midsagittal plot for the /u/ straw conditions. The left panel is the plot for the untargeted condition, and the right panel is the targeted condition. A longer, more uniformed epilarynx is seen in the targeted condition. However, the vocal tract shape appears more condensed in the targeted condition due to the relative angle of the epilarynx. In the targeted condition, the vocal tract is constricted at about 10 cm from the glottis and then widens in the lower portion of the oral cavity before constricting at the lips. A larger and more open lip configuration is seen in the untargeted condition.

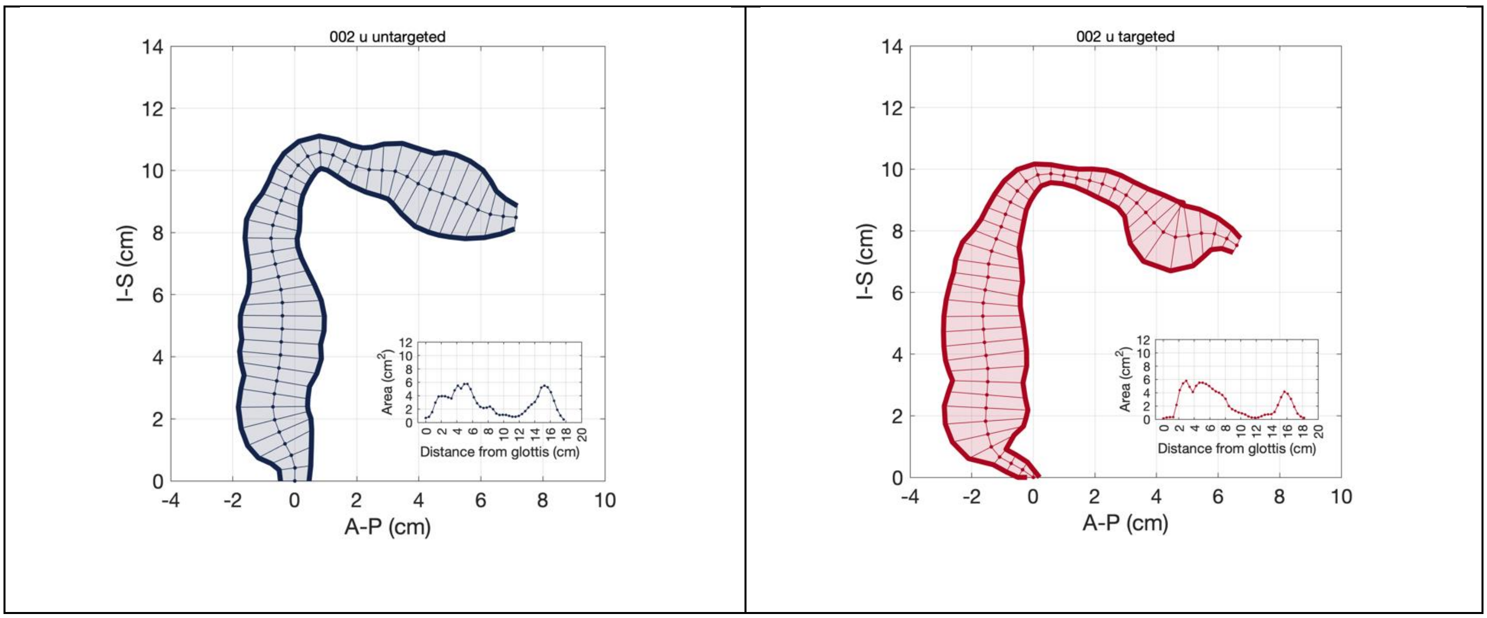


Figure 4.5: Pseudo mid-sagittal plots for Participant 02 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right).

<u>Area comparison of /u/ (Participant 02)</u>

Figure 4.6 shows targeting at the lips for the /u/ resulted in a distinct, nearly uniform, epilaryngeal tube with an overall smaller area than observed in the untargeted production (see the region 0 to 1.8 cm from the glottis, segment 1-4 in Fig. 4.6 black box). Targeting the production also led to a larger pharyngeal and smaller oral cavity area (A_o).

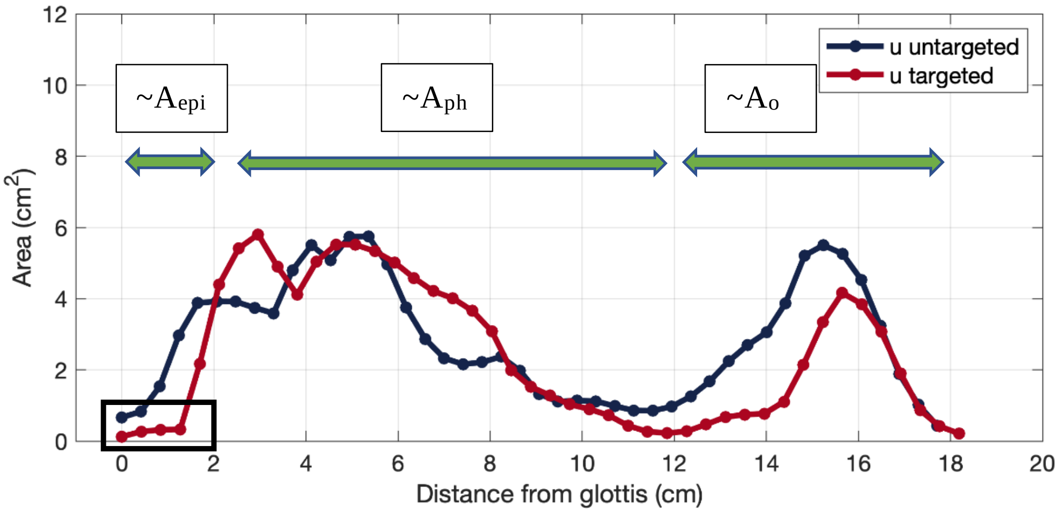


Figure 4.6: Smoothed area function for Participant 02 in the /u/ untargeted and targeted conditions. Black box at approximately 0 cm represents epilaryngeal tube shape for the /u/ targeted condition. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Area comparison of /u/ with and without straw (Participant 02)

Figure 4.7 is the area function for each /u/ condition with and without straw overlaid for comparison. Lengthening the vocal tract artificially (straw) with a vibratory target on the lips created the most uniform epilaryngeal shape (tube), the greatest increase in A_{ph}, and decreased A_o. The straw condition with targeting also resulted in the shortest vocal tract. Table 9 presents all the raw area values for Figure 4.7.

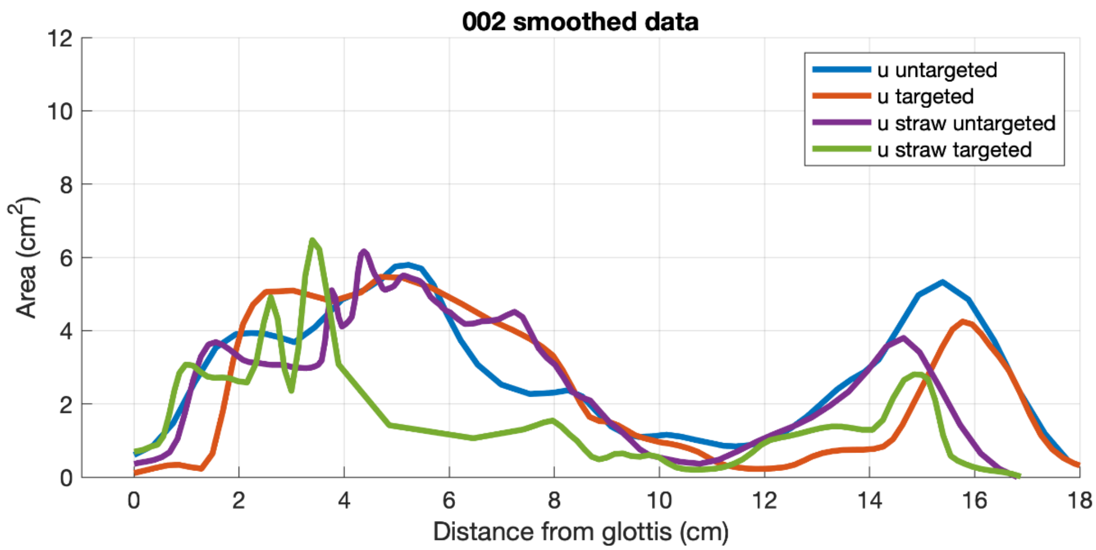


Figure 4.7: Combined area function smoothed for all /u/ tasks with and without straw participant 02.

Table 9: 02 area function raw data in cm^2

Section number (02)	/u/ untargeted	/u/ targeted	/u/ straw untargeted	/u/ straw targeted
1 glottis	0.67	0.13	0.36	0.68
2	0.83	0.27	0.49	0.96
3	1.54	0.32	1.07	2.32
4	2.96	0.33	2.85	2.93
5	3.88	2.17	3.6	2.73
6	3.93	4.4	3.31	2.63
7	3.92	5.41	3.12	3.28
8	3.75	5.8	3.06	3.81
9	3.59	4.89	3.02	3.94
10	4.8	4.11	3.58	5.58
11	5.5	5.04	4.43	2.13
12	5.08	5.52	5.48	0.66
13	5.74	5.51	5.32	0.63
14	5.75	5.33	5.36	0.64
15	4.97	5.02	5.07	0.66
16	3.75	4.58	4.55	0.69
17	2.87	4.21	4.21	0.72
18	2.33	4.01	4.27	0.78
19	2.16	3.66	4.42	0.85
20	2.22	3.09	4.01	1.13
21	2.37	1.99	3.19	1.45
22	1.97	1.53	2.45	1.19
23	1.32	1.28	2.05	0.67
24	1.11	1.03	1.49	0.58
25	1.14	0.9	1.01	0.59
26	1.12	0.72	0.62	0.58
27	0.98	0.43	0.48	0.33
28	0.86	0.27	0.35	0.21
29	0.85	0.23	0.37	0.23
30	0.97	0.27	0.6	0.36
31	1.26	0.47	0.91	0.71
32	1.68	0.68	1.17	1.03
33	2.24	0.74	1.37	1.16
34	2.69	0.77	1.64	1.3
35	3.05	1.11	2.03	1.38
36	3.88	2.14	2.44	1.32
37	5.21	3.34	3.21	1.42
38	5.5	4.16	3.89	2.37
39	5.26	3.85	3.61	2.74
40	4.53	3.07	2.69	1.59
41	3.23	1.9	1.5	0.47
42	1.89	0.87	0.71	0.24
43	1.02	0.42	0.22	0.15
44 lip termination	0.44	0.21	0	0.02
45 section length	0.41	0.42	0.39	0.39

Participant 02 /m/

Figure 4.8 shows the 3D reconstructions for Participant 02 during /m/ productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus").

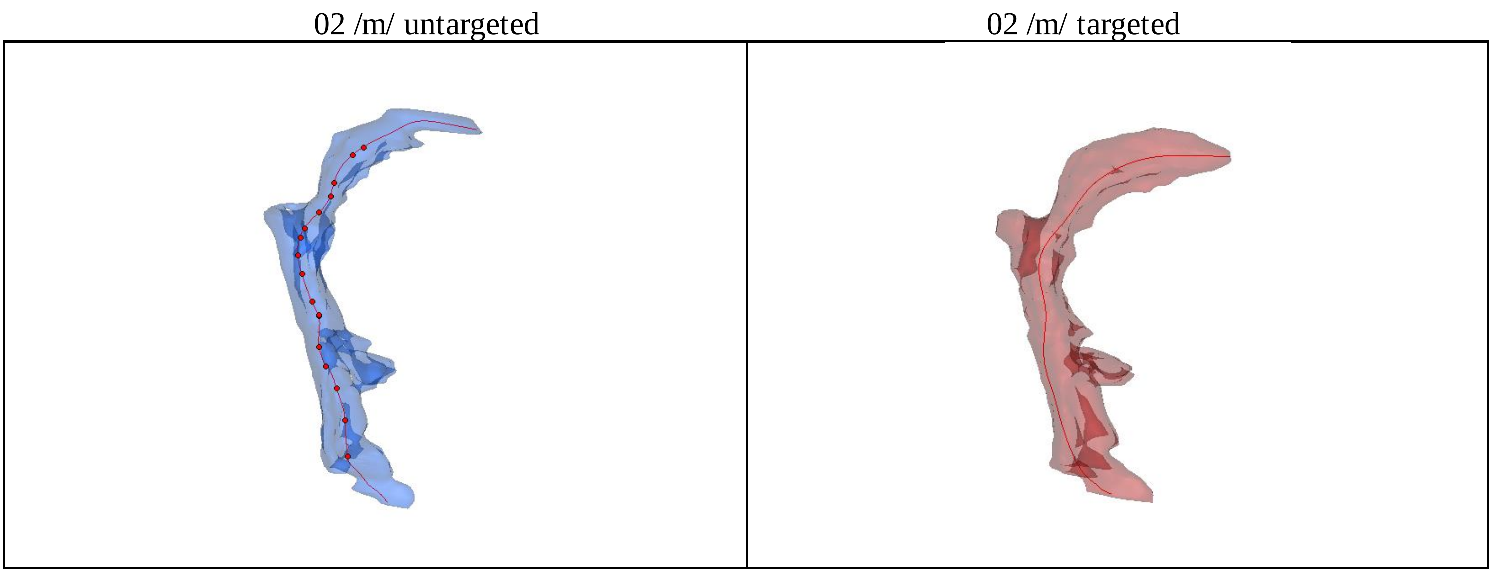


Figure 4.8: 3D model reconstructions for Participant 02 during production of an /m/ in the untargeted (left) and targeted conditions (right).

Figure 4.9 is the pseudo mid-sagittal plots for the /m/ in the untargeted (left panel) and targeted (right panel). Similar vocal tract configuration is seen visually in both conditions, with general expansions seen across all regions in the targeted condition. Complete lip occlusion is not seen in either condition.

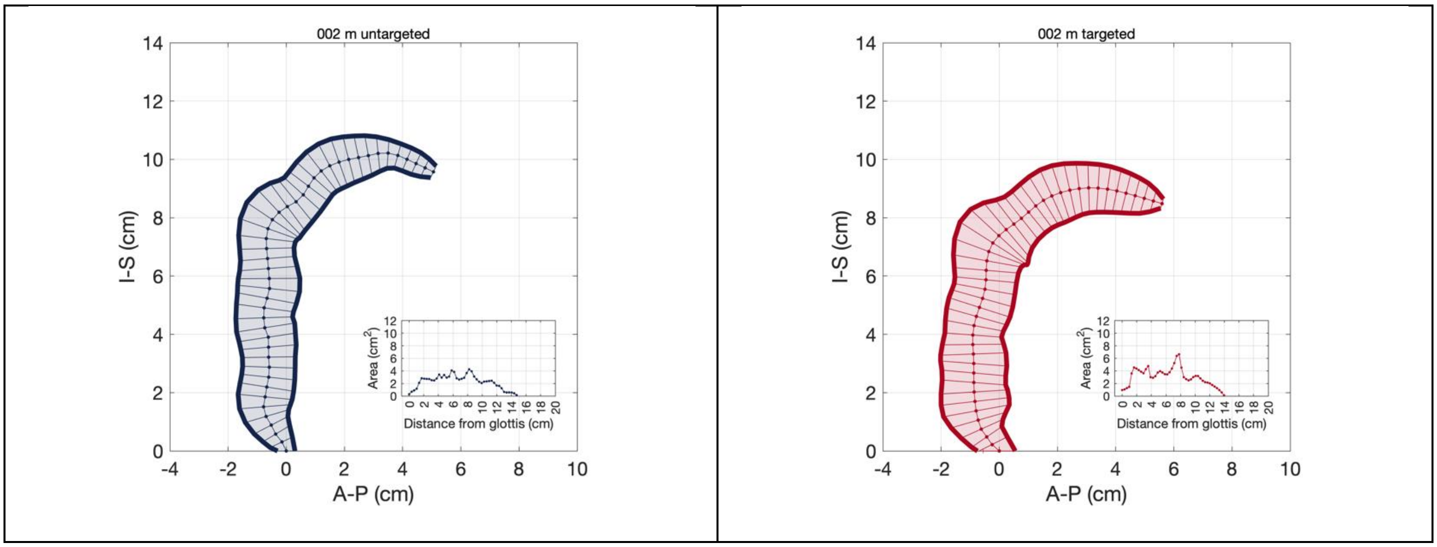


Figure 4.9: Pseudo mid-sagittal plots for Participant 02 during production of an /m/ in the untargeted (left) and targeted conditions (right).

Area comparison for /m/ (Participant 02)

Figure 4.10 is the smoothed area function for the /m/ conditions. Targeting vibrations in the oral cavity during /m/ resulted in a larger A_{epi} with a relatively comparable length and geometry to the untargeted production. Targeting the production also led to a larger pharynx with a slightly enlarged A_o from 9 to 11 cm from the glottis (Figure 4.9 black box). In the targeted condition, an overall shortening of the vocal tract length was noted (14 cm versus 14.8 cm). Informal listening confirmed that more hypernasality and visually more tongue retraction were observed for the targeted condition. Table 10 presents all the raw area values for Figure 4.10.

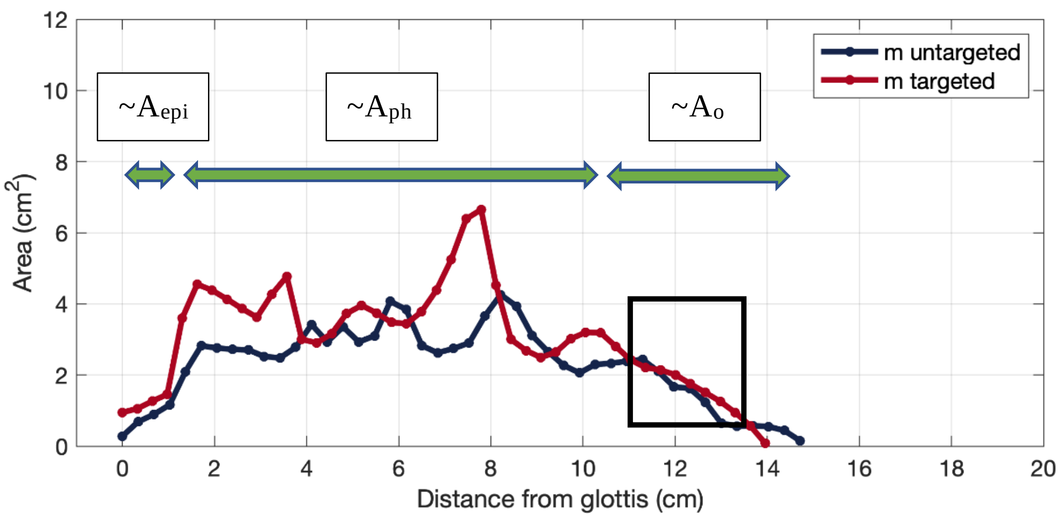


Figure 4.10: Smoothed area function for Participant 02 in the /m/ untargeted and targeted conditions. Black box at approximately 9 cm to 12 cm represents larger oral cavity in the targeted condition. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Table 10: Participant 02 area function raw data for /m/ in cm^2

Section number (02)	/m/ untargeted	/m/ targeted
1 glottis	0.28	0.95
2	0.69	1.05
3	0.89	1.26
4	1.16	1.46
5	2.1	3.6
6	2.82	4.55
7	2.76	4.38
8	2.72	4.13
9	2.7	3.87
10	2.52	3.62
11	2.48	4.27
12	2.78	4.77
13	3.42	3
14	2.93	2.91
15	3.35	3.16
16	2.93	3.73
17	3.1	3.96
18	4.07	3.74
19	3.84	3.48
20	2.83	3.44
21	2.62	3.77
22	2.74	4.38
23	2.9	5.25
24	3.66	6.39
25	4.26	6.65
26	3.94	4.53
27	3.11	3
28	2.66	2.69
29	2.27	2.49
30	2.06	2.65
31	2.3	3.02
32	2.33	3.2
33	2.39	3.19
34	2.44	2.81
35	2.1	2.42
36	1.67	2.21
37	1.62	2.14
38	1.23	2.01
39	0.65	1.75
40	0.56	1.51
41	0.57	1.26
42	0.55	0.95
43	0.44	0.58
44 lip termination	0.15	0.08
45 section length	0.34	0.32

Participant 02 Area Summary with Vibratory Focus

For participant 02, in two out of the three cases, a larger epilarynx (/m/ and /u/ with straw) and smaller A_o (/u/ with straw and /u/) were noted in targeted conditions compared to untargeted. A larger A_{ph} was noted across all targeted conditions (/u/ with straw, /u/, and /m/) compared to the untargeted condition. Vocal tract length did not change as a function of targeting.

4.3.2. Participant 03

Participant 03 /u/ straw

Figure 4.11 shows the 3D reconstructions for Participant 03 during the production of an /u/ with a straw placed at the lips. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left panel shows a gap from the uvula to the upper portion of the pharynx. This resulted from no measurable air present in the DICOM images to segment (Figure 4.12). The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analysis was carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

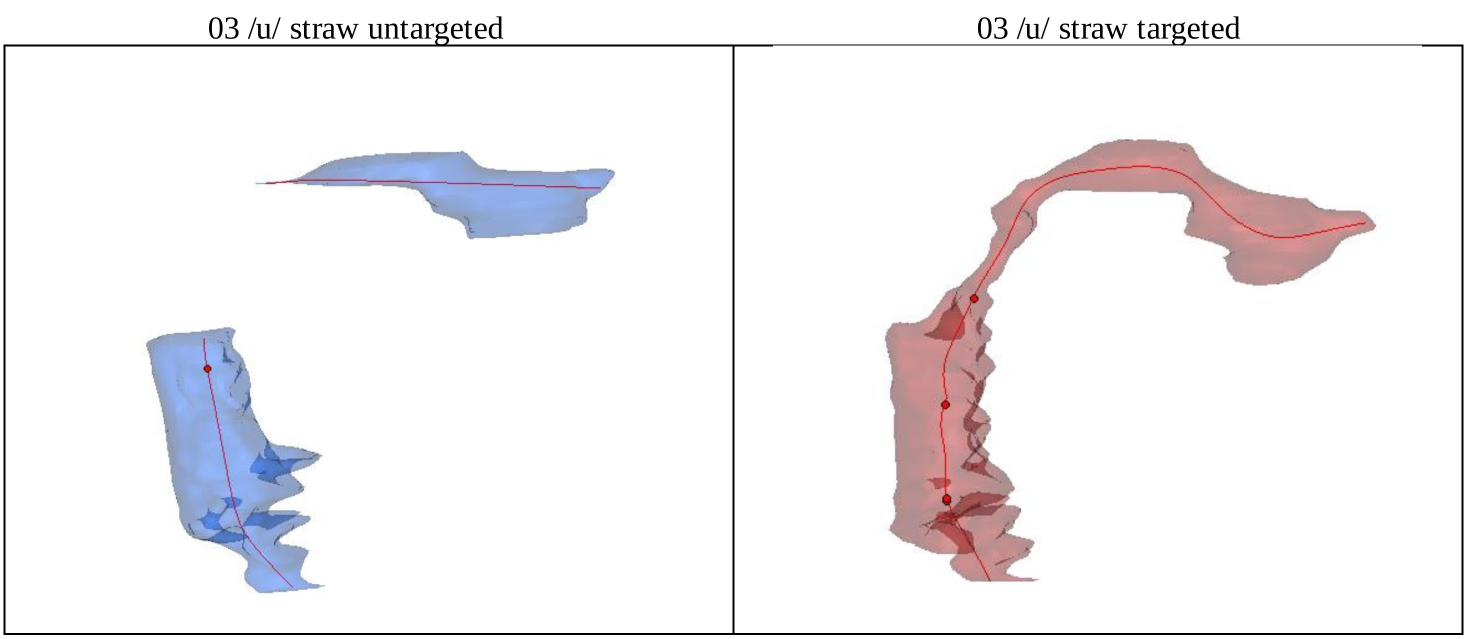


Figure 4.11: 3D model reconstructions for Participant 03 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). Gap in 3D model for /u/ straw untargeted condition due to no measurable air in the oropharynx.

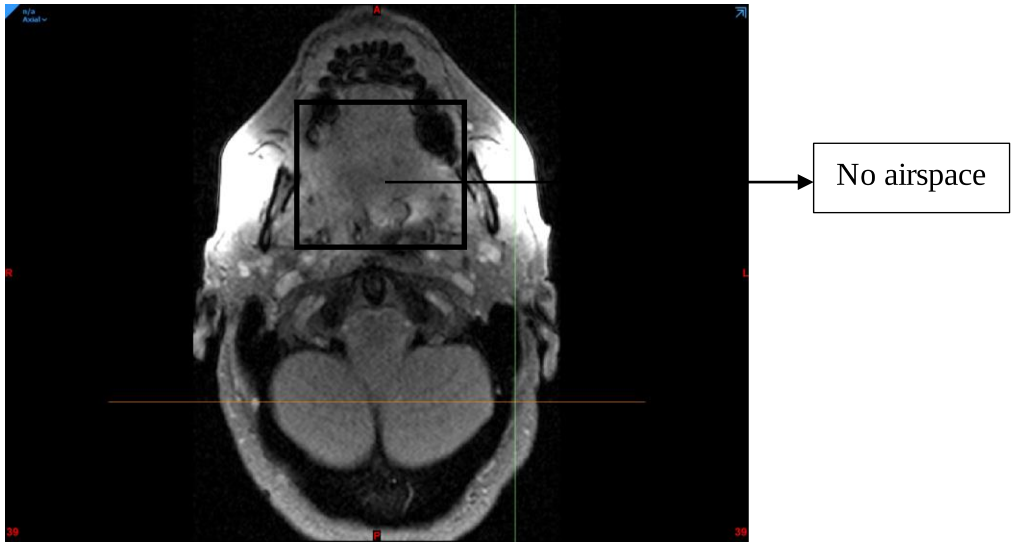


Figure 4.12: DICOM image during segmentation with no measurable area in oropharynx for 03 /u/ straw, untargeted. Black box represents the oral cavity with no air space to segment.

In Figure 4.11, due to the lack of measurable air in the/u/ straw untargeted condition, the pseudo-midsagittal plot (Fig. 4.13) will appear different than in Figure 4.11. In the untargeted /u/ straw condition in the left panel, there is evidence of only a short epilaryngeal section that opens into the lower pharynx. The vocal tract in the untargeted condition appears constricted at about 5.5 cm to about 9.5 cm from the glottis and then widens in the oral cavity before constricting around the straw at the lips. In the targeted condition, there is a more defined epilaryngeal section that opens into a wider pharynx from approximately 1 cm to 4.5 cm from the glottis. The /u/ straw targeted condition had a similar shape of the lower pharynx but had a greater expansion of the entire pharyngeal region. Additionally, the shape of the oral cavity shows greater expansion from the uvula to the soft palate and a significant expansion in the floor of the mouth with greater lip rounding.

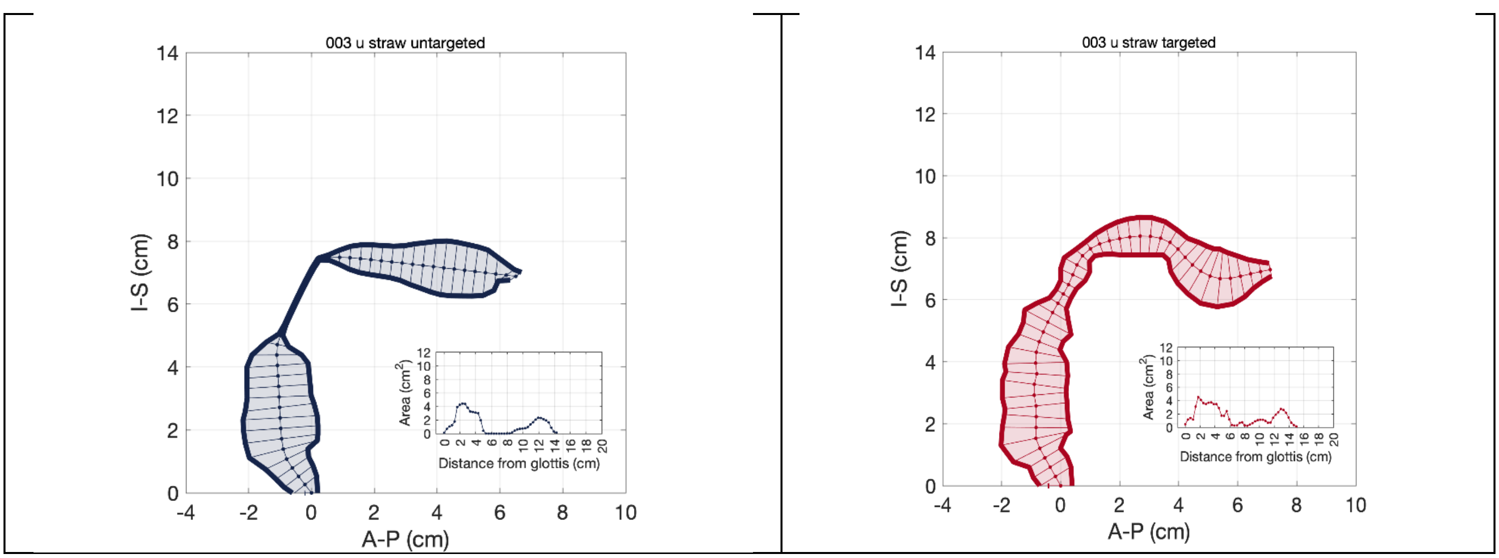


Figure 4.13: Pseudo mid-sagittal plots for Participant 03 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). /u/ straw untargeted shows zero area from approximately 5 cm to 8 cm due to no measurable air in the oropharynx.

Area comparison for /u/ straw (Participant 03)

Figure 4.14 is the area function for the /u/ straw conditions with and without straw for participant 03. Targeting the /u/ during the straw condition resulted in a larger A_{epi} compared to the untargeted production with a similar epilaryngeal shape during both conditions. Targeting led to a slightly different cross-sectional shape of the pharynx beginning around 2.2 cm from the glottis, with a general trend to being somewhat larger in area. A zero value is seen from approximately 5 to 8.2 cm from the glottis (Fig. 4.14 black box). It would be impossible to have a zero area at any anatomical level. So, while not a true zero, there was no measurable air in the raw DICOM; therefore, on the area function, a zero value is given. (Fig. 4.14). This participant appeared to narrow the oropharynx and increase the pharynx. Targeting also resulted in a larger oral cavity. The vocal tract length was longer (15 cm) in the targeted condition compared to 14.2 cm in the untargeted condition due to greater lip rounding.

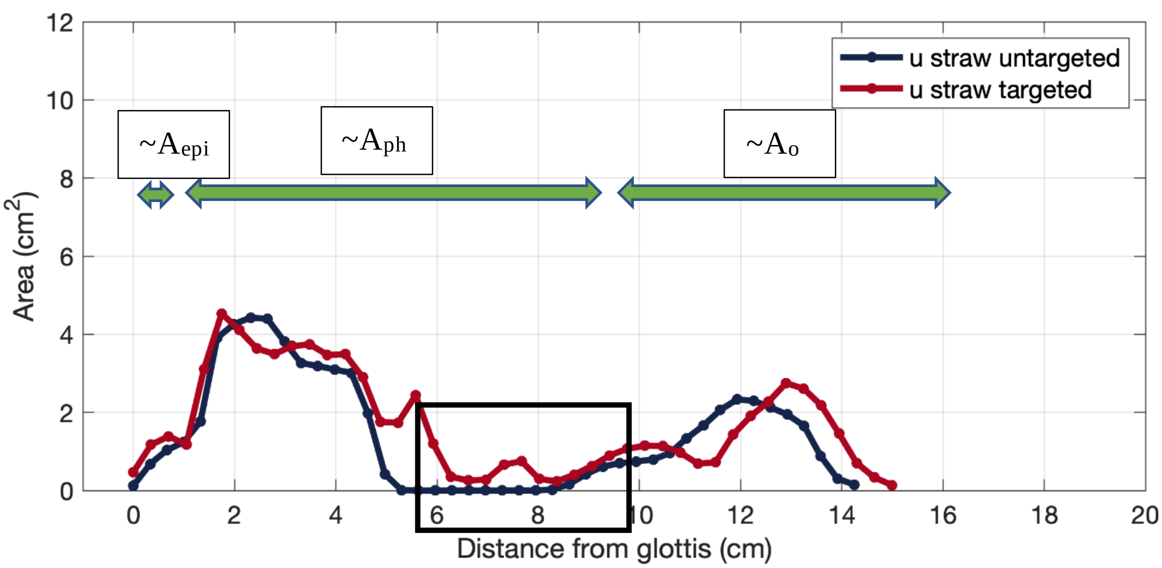


Figure 4.14: Smoothed area function for Participant 03 in the /u/ straw untargeted and targeted conditions. Black box at approximately 5 cm to 9 cm represents the region with no measurable air and, therefore, a zero area. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Participant 03 /u/ vowel

Figure 4.15 shows the 3D reconstructions for Participant 03 during production of an /u/ vowel. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). A comparable vocal tract shape was seen between Figures 4.15 and 4.16 There was a significant increase in the A_o seen for the /u/ targeted condition. This difference suggests that the interpretation of targeting for this participant was possibly a lowered jaw to accommodate the air in the floor of the mouth. The /u/ untargeted condition appears to have greater lip protrusion and a more similar lip shape that would be expected with the /u/ vowel.

03 /u/ untargeted 03 /u/ targeted

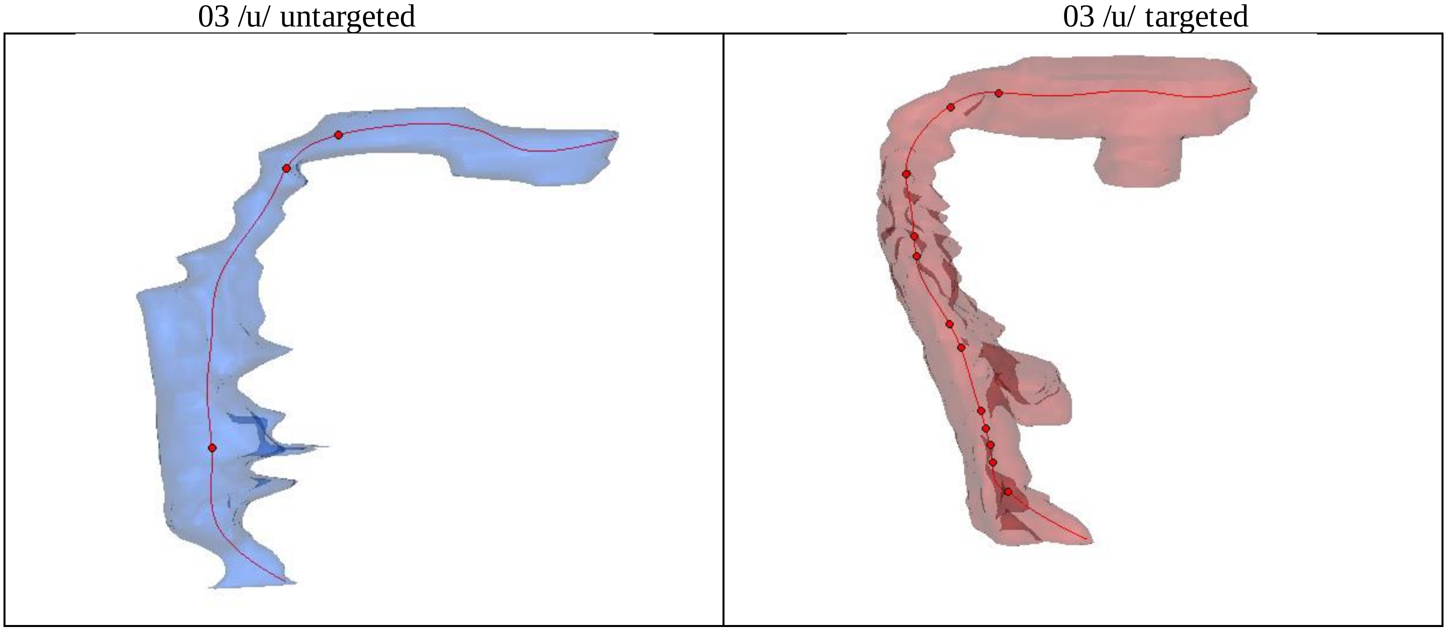


Figure 4.15: 3D model reconstructions for Participant 03 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right).

In Figure 4.16, is the pseudo midsagittal plots for the /u/ untargeted and targeted productions. The vocal tract length is substantially longer in the targeted condition, with a more distinct epilaryngeal tube. A larger oral cavity, and pharyngeal area is also appreciated in the /u/ targeted condition.

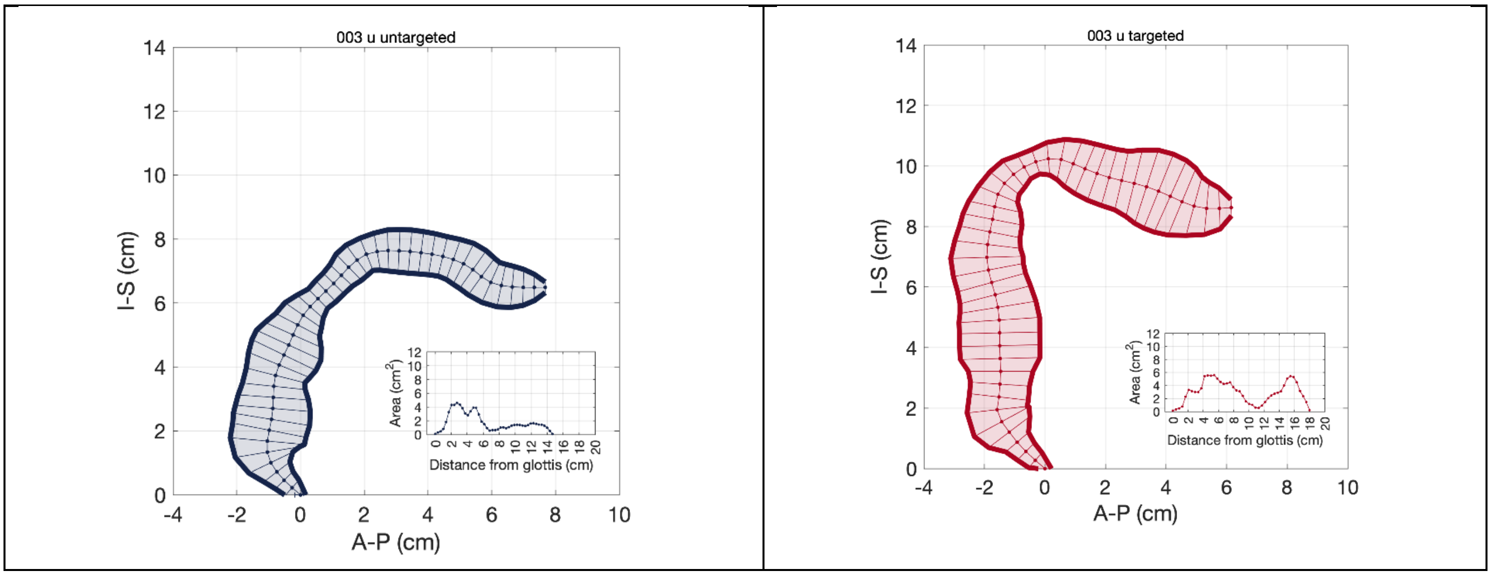


Figure 4.16: Pseudo mid-sagittal plots for Participant 03 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right).

Area comparison /u/ (Participant 03)

Figure 4.17 represents the cross-sectional areas for the /u/ conditions with and without targeting. Targeting /u/ resulted in an identical A_{epi} and slightly longer epilarynx compared to the untargeted condition and larger pharyngeal and oral areas. The vocal tract was 4.5 cm longer in the targeted condition.

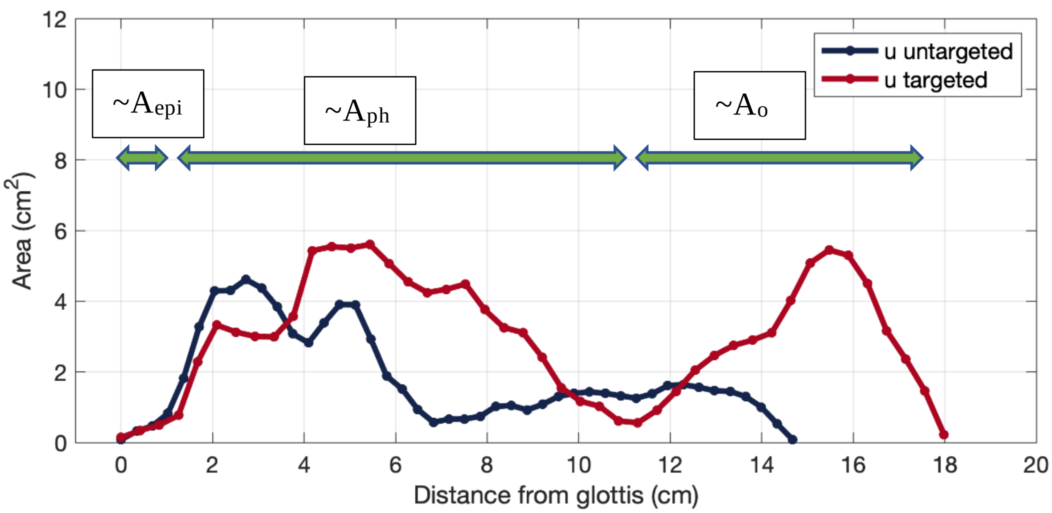


Figure 4.17: Smoothed area function for Participant 03 in the /u/ untargeted and targeted conditions. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

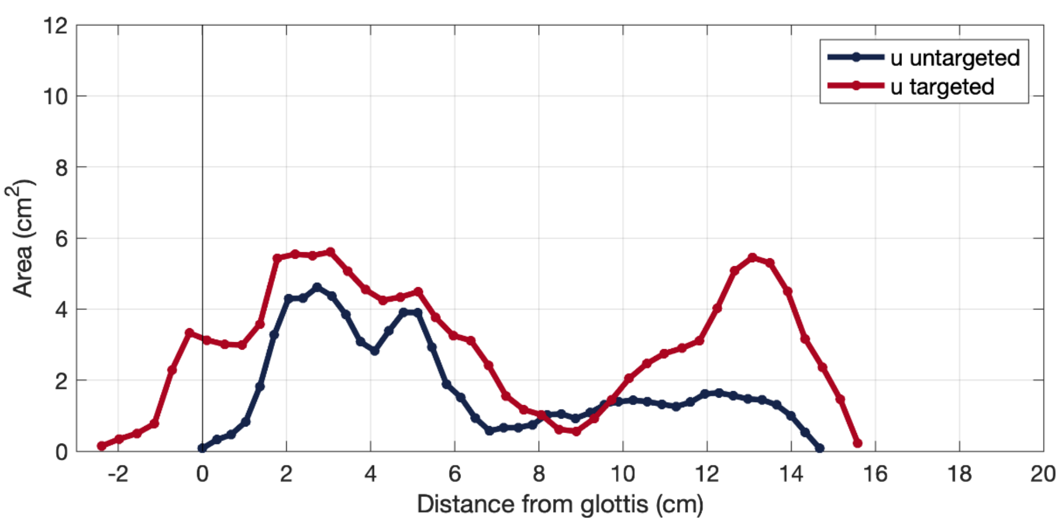


Figure 4.18: Smoothed area function aligned to the pharynx (shifted 3cm towards the glottis) for Participant 03 in the /u/ untargeted and targeted conditions.

Due to the vocal tract length discrepancy between the two conditions, the area function for the untargeted condition was shifted leftward along the x-axis in Figure 4.18 to align the pharyngeal regions of both conditions for comparison. With the alignment of the pharyngeal regions, the targeted condition area function (Fig. 4.18) suggests a laryngeal lowering (-1.8 cm from the glottis) and greater lip protrusion in the targeted condition. In the raw DICOM images, lip protrusion can be observed (Fig. 4.19) in the targeted condition.

In order to confirm whether the discrepancy in vocal tract length from Figure 4.18 was due to a lowered larynx and greater lip protrusion, the raw DICOM images (Figure 4.19 through 4. 24) were reviewed. Figures 4.19 and 4.20 are two axial slices from the /u/ no vibratory condition, and Figure 4.21 is the concurrent sagittal slice from the /u/ no vibratory condition. These raw images are a visual representation, in two different planes of the smaller oral cavity in the non-targeted condition. Figures 4.22 and 4.23 are two axial slices from the /u/ vibratory condition, and then Figure 4.24 is the concurrent sagittal slice showing greater A_o. The axial images were compared side by side, and the author and committee agreed regarding the smaller oral cavity secondary to less lip protrusion in the untargeted/no vibratory condition.

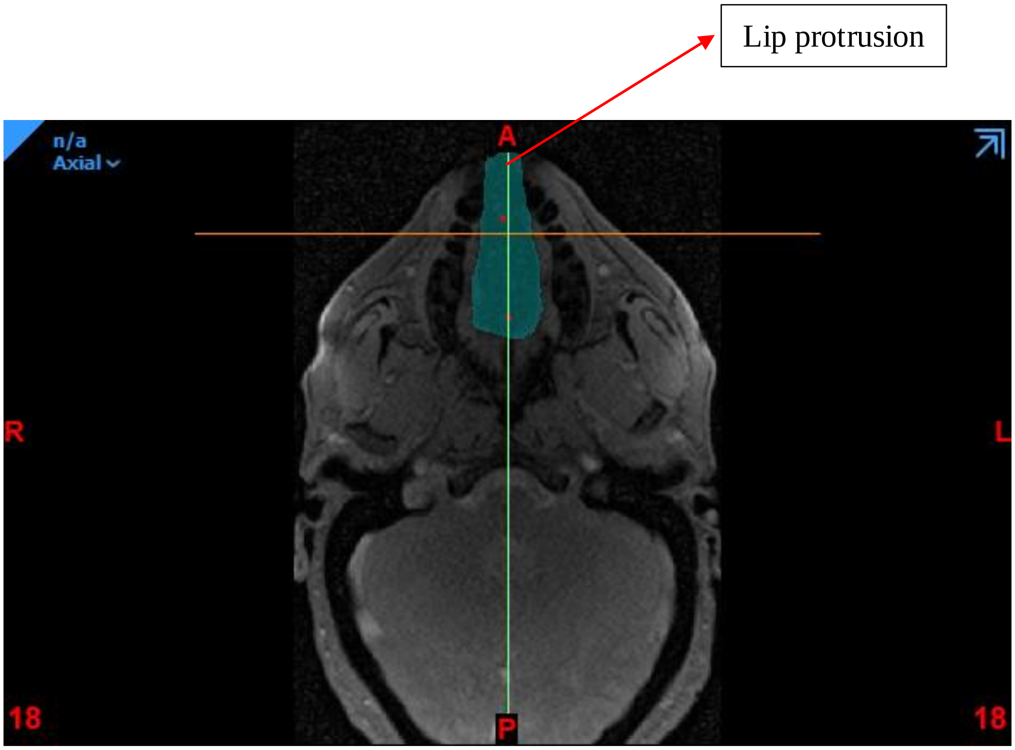


Figure 4.19: Raw DICOM in axial plane. Green highlighted area represents area of air in the oral cavity for Participant 03 /u/ no vibratory focus. Red arrow showing lip protrusion during production as seen as an elongation past the boundaries of the teeth.

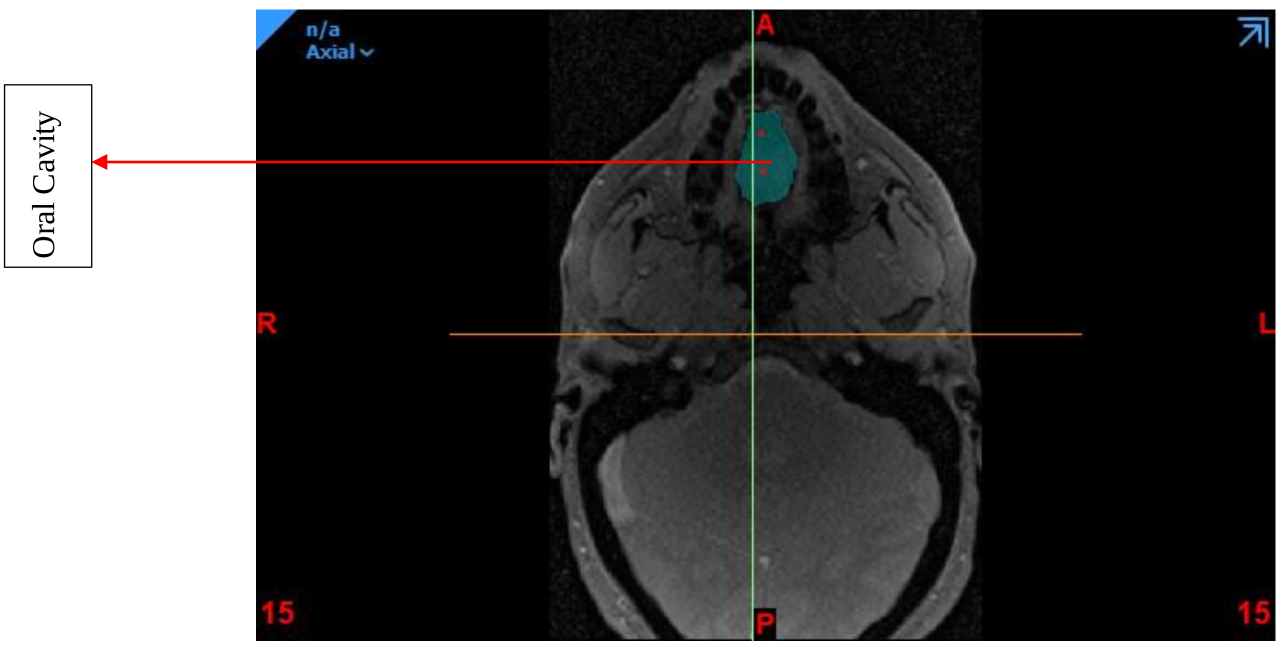


Figure 4.20: DICOM image of axial slice with oral cavity area highlighted in green for Participant 03 /u/ no vibratory focus.

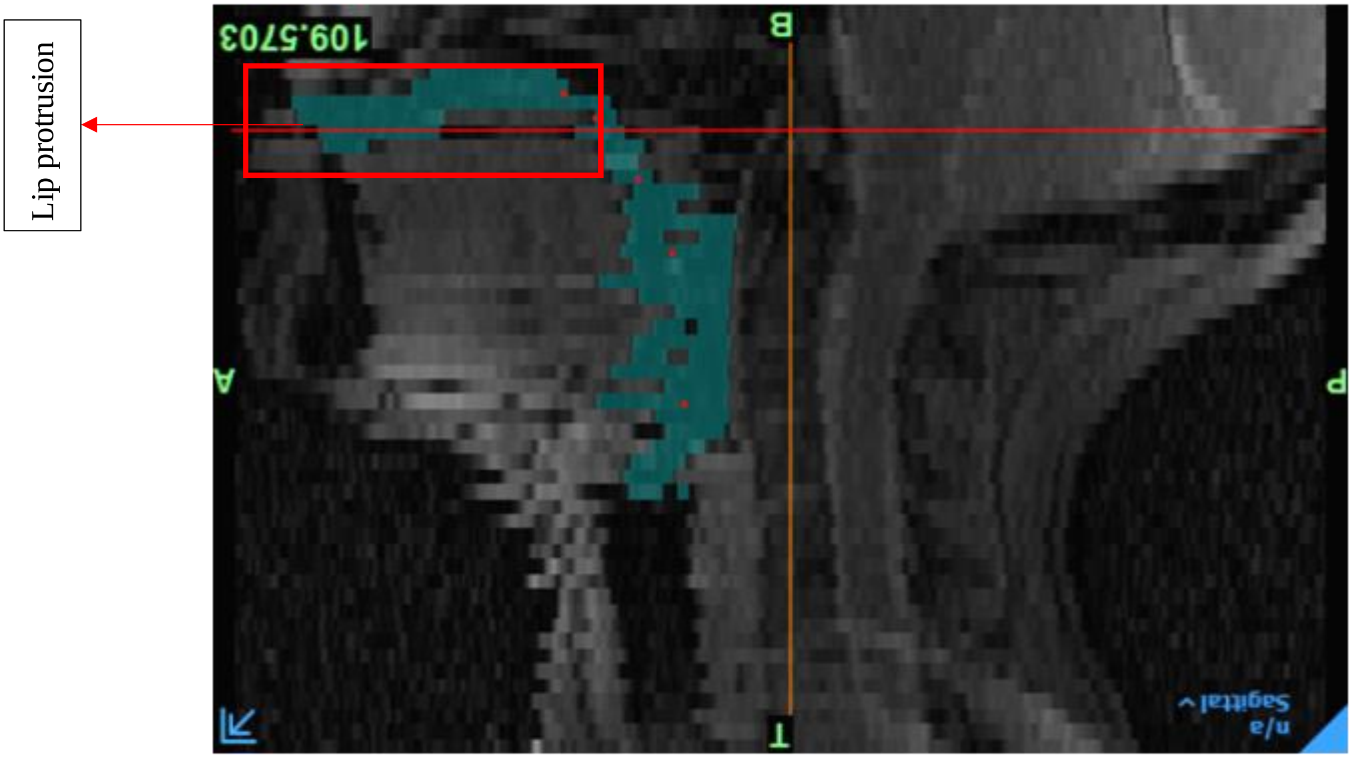


Figure 4.21: DICOM image of sagittal slice for Participant 03 /u/ no vibratory focus. Image showing sagittal plane of lip protrusion. Red box representing oral cavity.

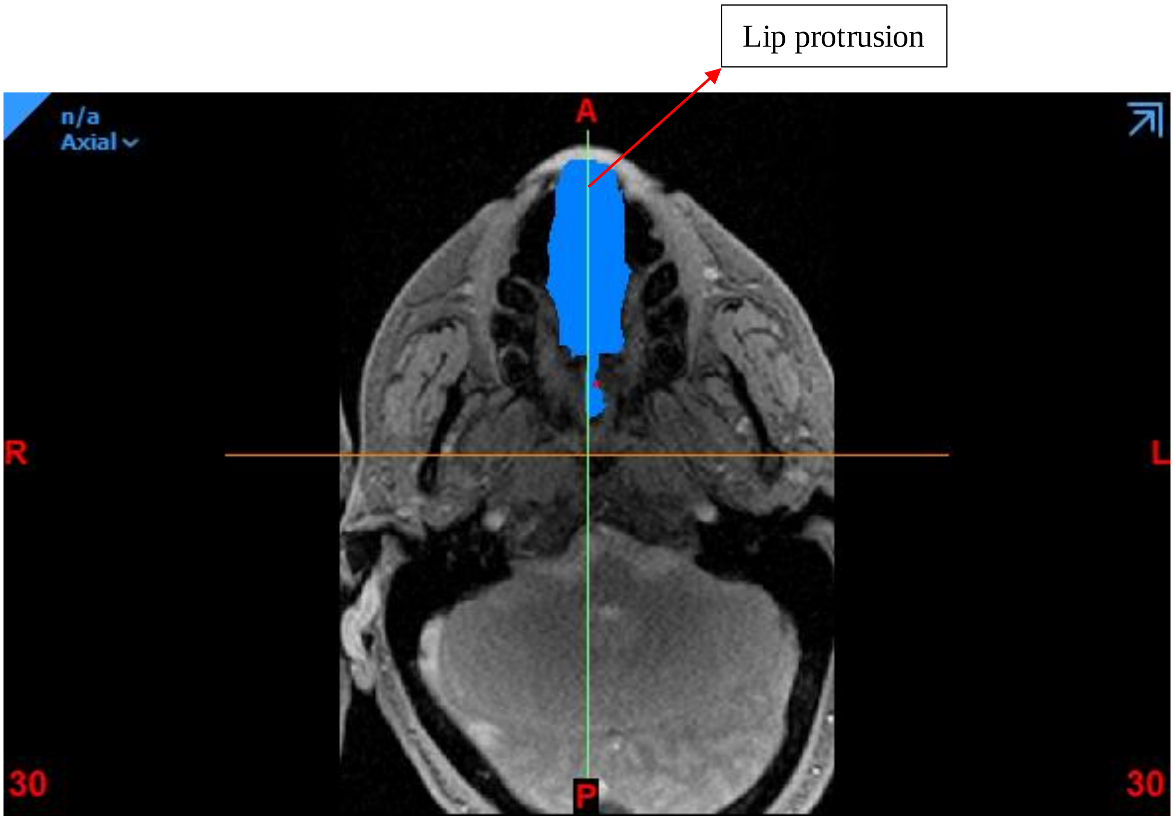


Figure 4.22: Raw DICOM in axial slice one. Blue highlighted area represents area of air in the oral cavity for 03 /u/ targeted. Red arrow shows lip protrusion with more lip rounding during production, as seen as an elongation past the boundaries of the teeth.

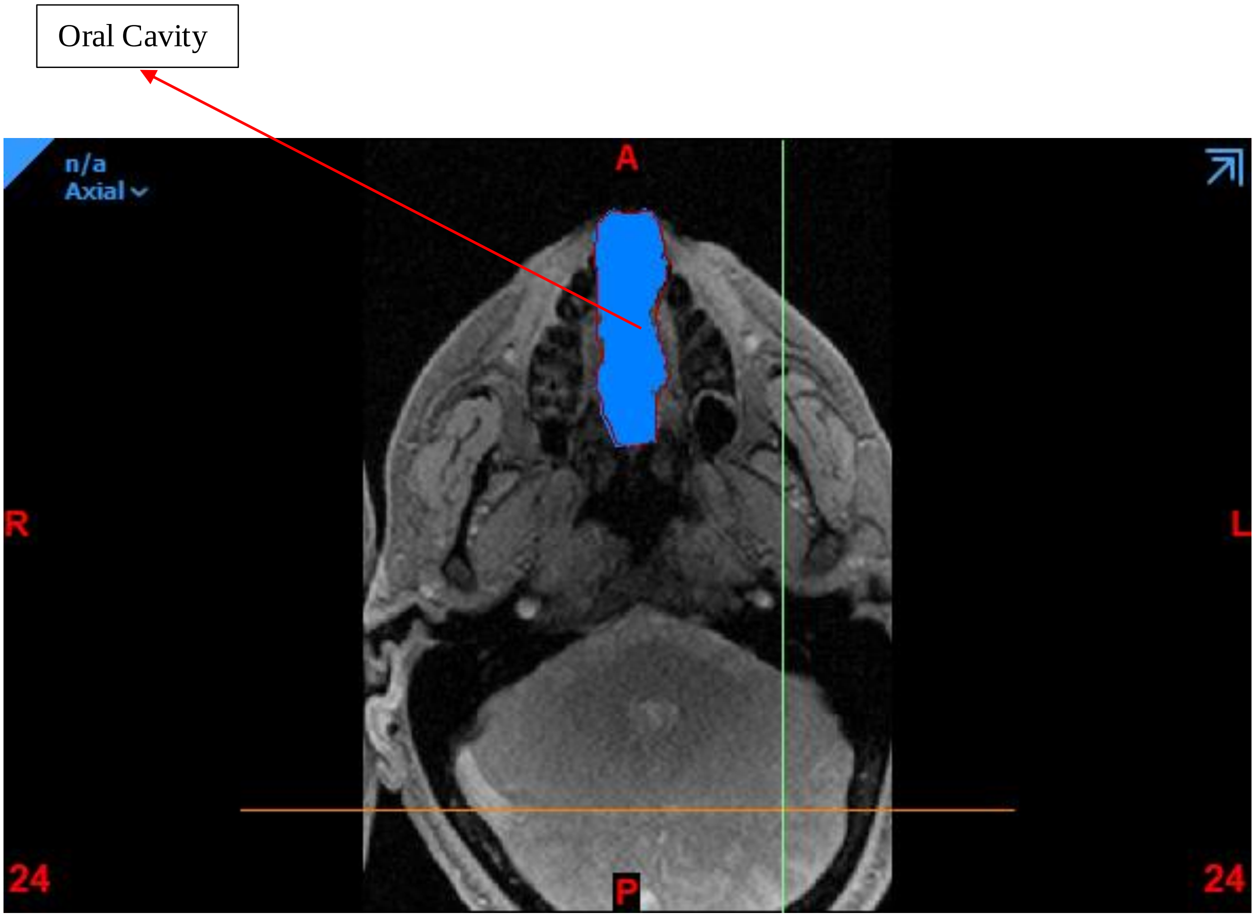


Figure 4.23: Raw DICOM in axial slice two. Blue highlighted area with outlined red represents area of air in the oral cavity for 03 /u/ targeted. Red arrow showing oral cavity area for one slice.

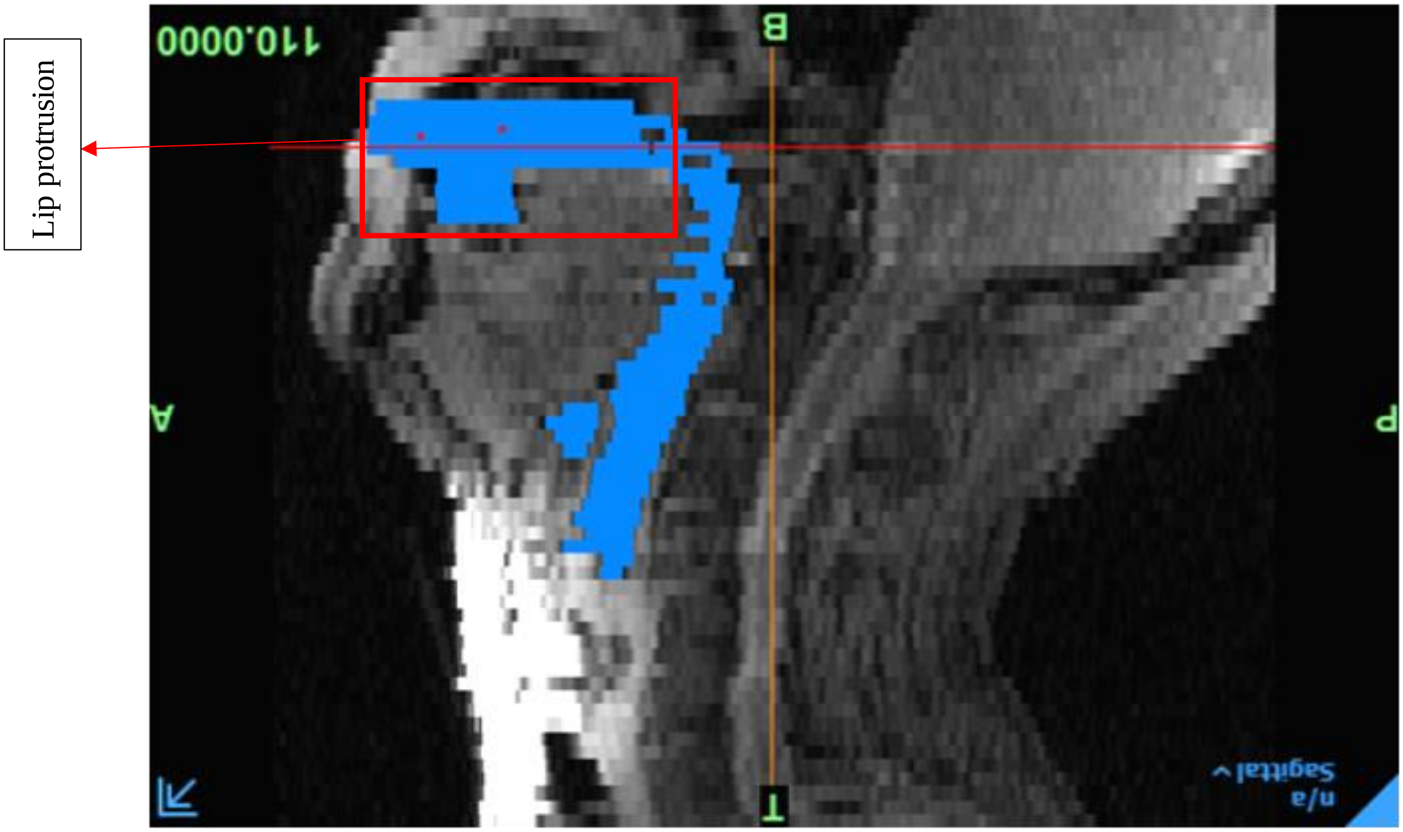


Figure 4.24: DICOM image of sagittal slice for Participant 03 /u/ targeted. Image showing sagittal plane of lip protrusion. Red box representing oral cavity.

Area comparison of /u/ with and without straw (Participant 03)

Straw without a target led to a more uniform epilaryngeal tube, greatest pharyngeal and oral cavity area, with the longest vocal tract. The greatest A_{epi} was in the /u/ straw targeted condition (Figure 4.25). Table 11 presents all the raw area values for Figure 4.25.

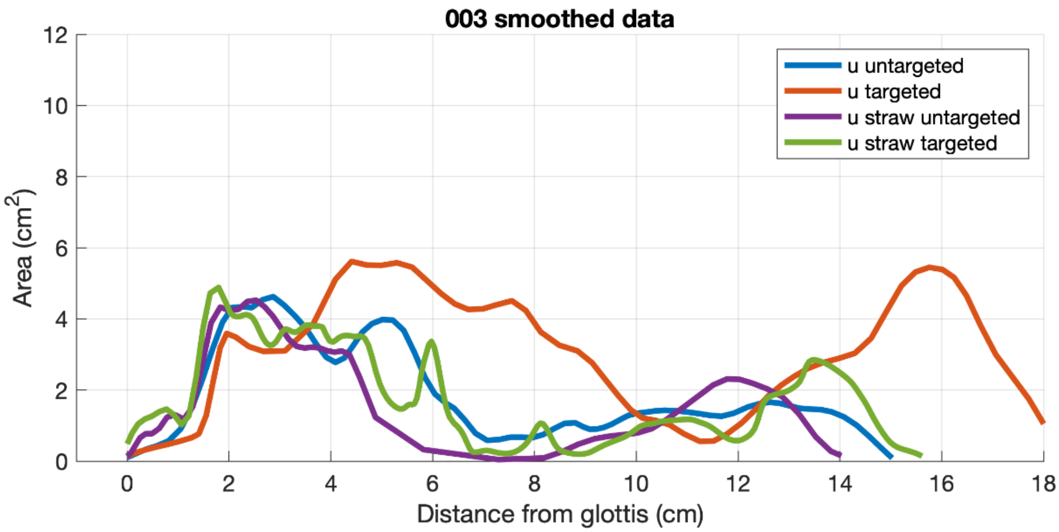


Figure 4.25: Combined area function smoothed participant 03.

Participant 03 Area Summary

For participant 03, both targeted conditions of /u/, irrespective of straw, had increased pharyngeal and oral cavity areas. For two of the three targeted conditions (/u/ with straw and /m/), the A_{epi} increased. Targeting did not improve epilaryngeal shape or length in this participant.

Table 11: Participant 03 area function raw data

Section number	/u/ untargeted	/u/ targeted	/u/ straw untargeted	/u/ straw targeted
1 (glottis)	0.08	0.15	0.13	0.47
2	0.32	0.34	0.68	1.17
3	0.47	0.5	1.04	1.39
4	0.82	0.77	1.24	1.18
5	1.82	2.28	1.77	3.11
6	3.27	3.33	3.91	4.53
7	4.3	3.13	4.26	4.11
8	4.3	3	4.43	3.63
9	4.62	2.99	4.39	3.5
10	4.37	3.58	3.81	3.7
11	3.84	5.43	3.26	3.74
12	3.08	5.55	3.19	3.47
13	2.82	5.51	3.1	3.49
14	3.39	5.61	3.01	2.9
15	3.91	5.07	1.97	1.75
16	3.9	4.55	0.42	1.74
17	2.93	4.25	0.01	2.44
18	1.88	4.34	0	1.2
19	1.52	4.49	0	0.35
20	0.94	3.76	0	0.26
21	0.57	3.25	0	0.28
22	0.66	3.11	0	0.66
23	0.67	2.42	0	0.75
24	0.75	1.56	0	0.3
25	1.03	1.17	0	0.24
26	1.05	1.03	0.02	0.41
27	0.93	0.61	0.16	0.63
28	1.08	0.56	0.42	0.9
29	1.3	0.92	0.61	1.07
30	1.39	1.45	0.7	1.15
321	1.44	2.05	0.75	1.14
32	1.39	2.47	0.8	0.97
33	1.32	2.75	0.96	0.69
34	1.26	2.9	1.34	0.72
35	1.38	3.11	1.66	1.43
36	1.61	4.02	2.06	1.91
37	1.64	5.08	2.33	2.27
38	1.56	5.45	2.3	2.75
39	1.47	5.3	2.13	2.61
40	1.45	4.5	1.95	2.18
41	1.3	3.17	1.66	1.46
42	1	2.36	0.88	0.71
43	0.54	1.46	0.3	0.34
44 (lip term)	0.09	0.23	0.15	0.14
45 (length)	0.34	0.42	0.33	0.35

Area comparison /m/ (Participant 03)

Figure 4.26 shows the 3D reconstructions for Participant 03 during the production of an /m/. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the right panel shows a gap from the lower portion of the pharynx through the laryngeal surface of the epiglottis. This resulted from missing 10 slices from the DICOM images (Figure 4.26). The data in both the upper and lower portions of the reconstructions were, however, deemed viable, and the centerline and area function analysis was carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

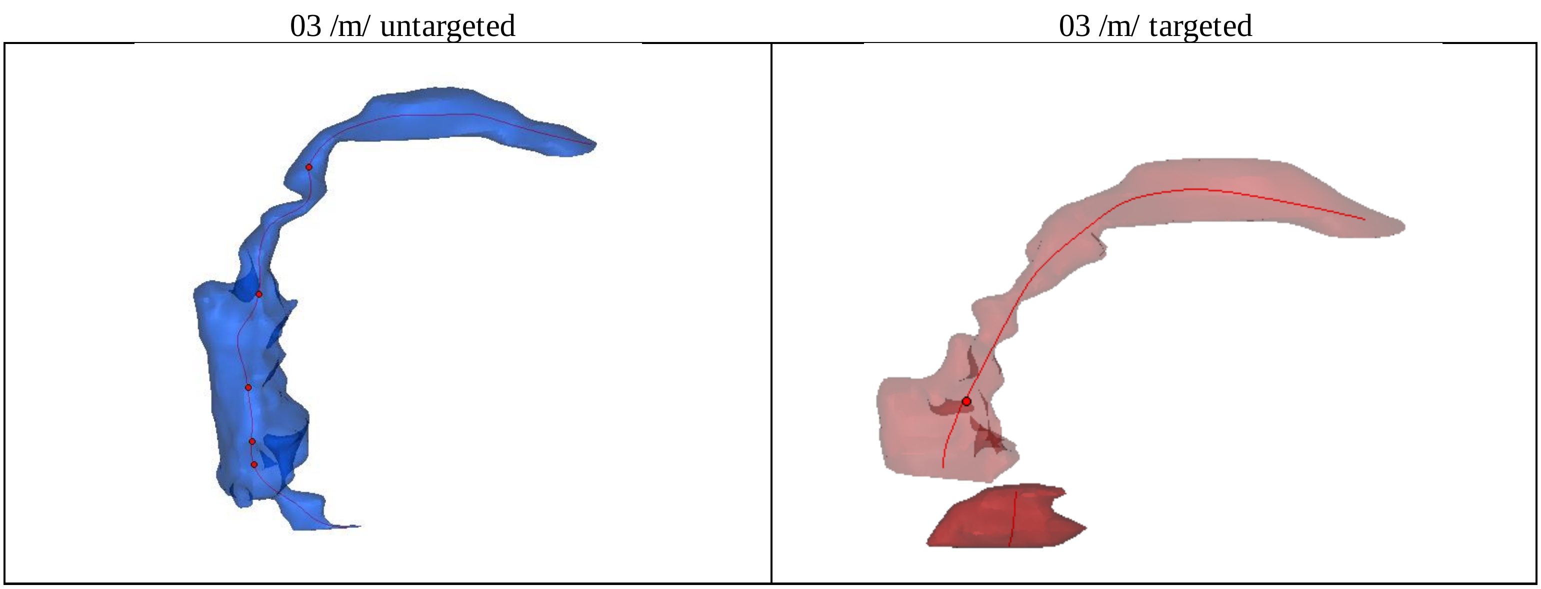


Figure 4.26: 3D model reconstructions for Participant 03 during production of an /m/ in the untargeted (left) and targeted conditions (right). 10 slices are missing from the DICOM images from the pharynx to epilarynx and appear as a gap in the 3D model for the /m/ targeted condition.

Figure 4.27 of the Mimics view shows how the missing data of the /m/ vibratory condition created a misalignment of the image. Due to the missing data, we used an interpolation of the available information to make assumptions about the lower portion of the pharynx area.

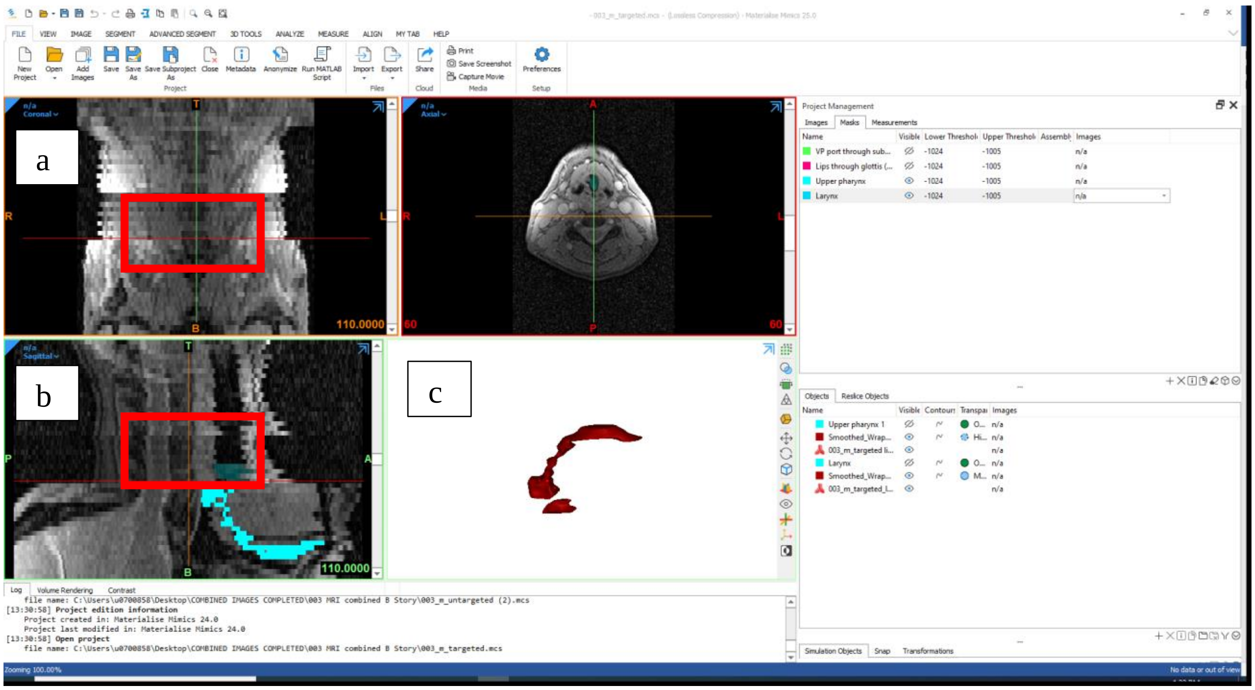


Figure 4.27: Mimics view of missing data highlighted with red box in both a) coronal view and b) sagittal. Missing section in 3D model for Participant 03 /m/ targeted.

In Figure 4.27, due to missing data in the targeted /m/ condition, the pseudo-midsagittal plot (Fig. 4.27) will appear different than in Figure 4.28. Figure 4.28 is the pseudo mid-sagittal plots for participant 03 /m/ conditions. In the untargeted /m/ condition in the left panel, there is evidence of a longer epilaryngeal section that opens into the lower pharynx. The vocal tract in the untargeted condition appears expanded from about 2 cm from the glottis to about 5 cm, before narrowing in the oropharynx. In the targeted condition, there expansion of the epilarynx but a shorter and less defined tube is evident. The /m/ targeted condition also had more pharynx narrowing at around 4 cm from the glottis, with a shorter but wider oral cavity.

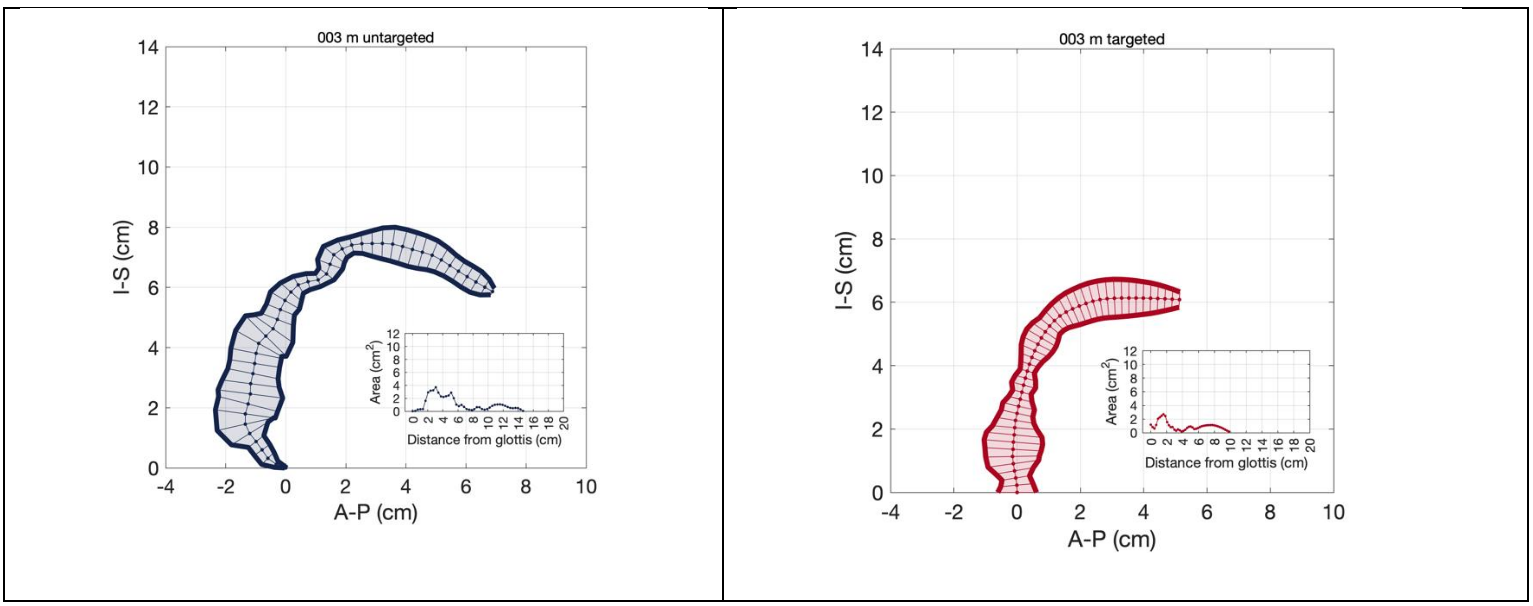


Figure 4. 28: Pseudo mid-sagittal plots for Participant 03 during production of an /m/ in the untargeted (left) and targeted conditions (right). /m/ targeted represents an interpolation of the vocal tract from approximately 1 cm to 3 cm due to missing data in the 3D model.

Figure 4.29 is the smoothed area functions for the overlaid /m/ untargeted and targeted conditions for participant 03. Targeting the /m/ resulted in a shorter epilarynx with a larger area than the untargeted production, with the epilaryngeal tube configuration more pronounced, defined, and uniform for the untargeted production of /m/ (Fig. 4.28 black box). Targeting the /m/ also led to smaller pharyngeal and larger A_o. To better understand the area functions, the author and co-advisors listened to the audio recorded during the MRI and looked at the raw images. We speculated that the tongue filled the anterior portion of the oral cavity (i.e., touched the hard palate) during the targeted production, decreasing the oral cavity's relative airspace and shortening the vocal tract during the targeted production. Table 12 presents all the raw area values for Figure 4.29.

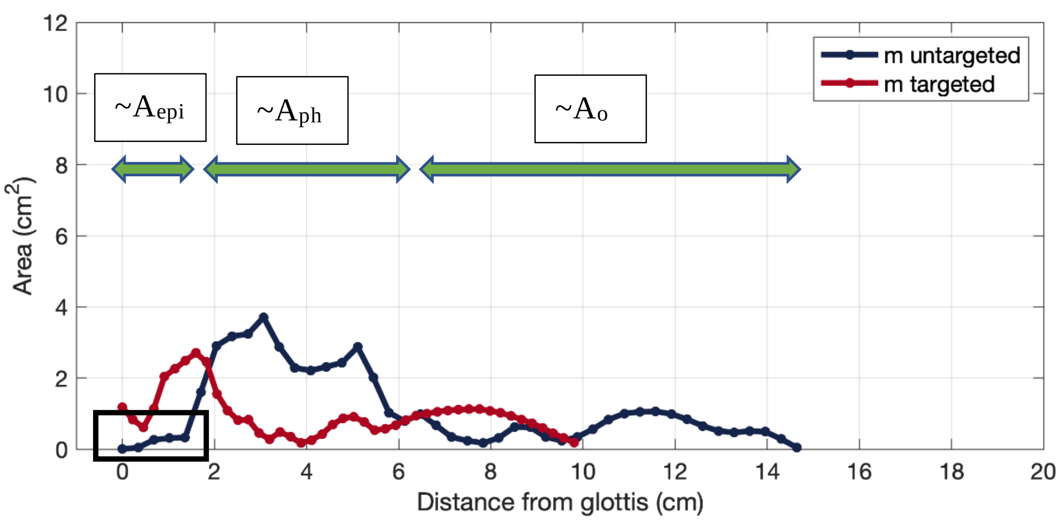


Figure 4.29: Smoothed area function for Participant 03 in the /m/ straw untargeted and targeted conditions. Interpolation of the /m/ targeted condition caused the appearance of a shortened vocal tract. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Table 12: Participant 03 area function raw data for /m/ in cm^2

Section number (02)	/m/ untargeted	/m/ targeted
1 glottis	0	1.18
2	0.04	0.83
3	0.26	0.61
4	0.31	1.15
5	0.33	2.04
6	1.61	2.26
7	2.91	2.49
8	3.18	2.71
9	3.23	2.45
10	3.71	1.55
11	2.88	1.09
12	2.29	0.82
13	2.21	0.83
14	2.32	0.45
15	2.43	0.28
16	2.87	0.48
17	2.01	0.35
18	1.03	0.17
19	0.79	0.26
20	0.99	0.42
21	0.67	0.69
22	0.34	0.87
23	0.24	0.91
24	0.17	0.77
25	0.32	0.53
26	0.63	0.57
27	0.62	0.67
28	0.34	0.82
29	0.24	0.94
30	0.34	1
31	0.57	1.05
32	0.83	1.09
33	1	1.11
34	1.05	1.12
35	1.06	1.12
36	0.99	1.08
37	0.84	1.02
38	0.65	0.94
39	0.51	0.83
40	0.47	0.73
41	0.51	0.59
42	0.49	0.45
43	0.28	0.31
44 lip termination	0.05	0.18
45 section length	0.34	0.23

4.3.3. Participant 04

Participant 04 /u/ straw

Figure 4.30 shows the 3D reconstructions for Participant 04 during production of an /u/ vowel productions with straw. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left panel shows a gap from the oropharynx from the uvula to the top portion of the pharynx. This resulted from missing 10 slices from the DICOM images (Figure 4.31). The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analysis was carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

04 /u/ straw untargeted | 04 /u/ straw targeted

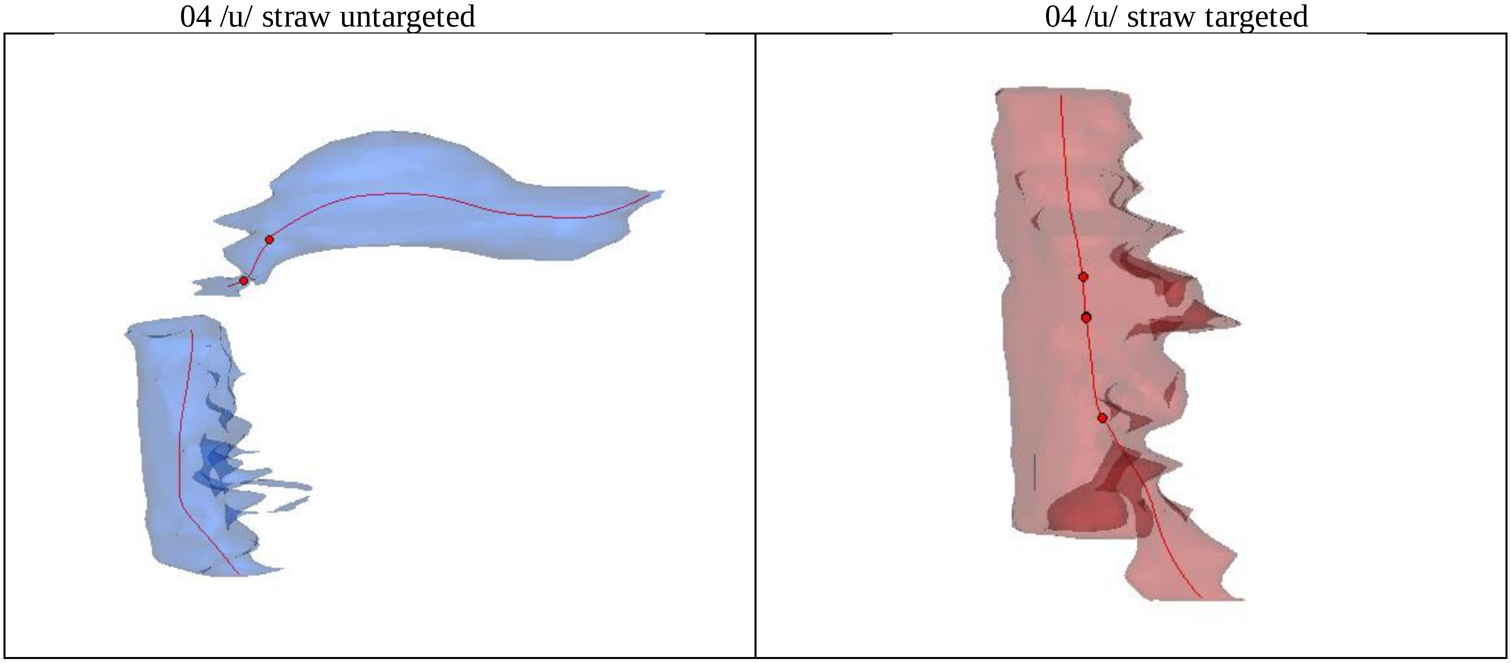


Figure 4.30: 3D model reconstructions for Participant 04 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). 10 slices from the DICOM images are missing from the oral cavity in the /u/ straw targeted condition.

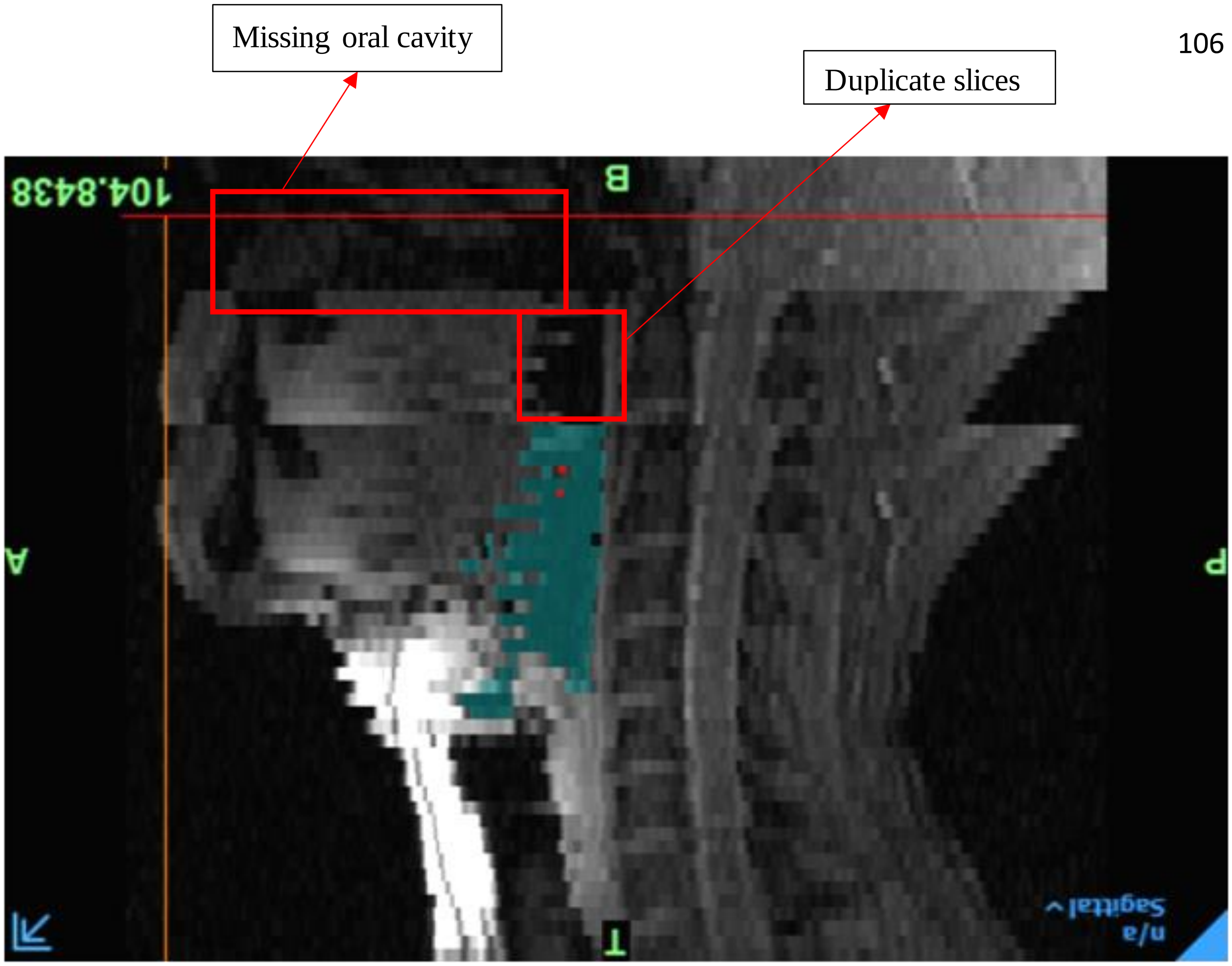


Figure 4.31: DICOM image from sagittal slice with no measurable air in the oral cavity (red box) or oropharynx (red box) for Participant 04 /u/ straw targeted.

Figure 4.32 is the pseudo mid-sagittal plots for the /u/ with straw in the untargeted (left panel) and targeted (right panel). Due to the lack of measurable air in the raw DICOM for the untargeted condition, an interpolation of the vocal tract area at approximately 5 cm from the glottis was created. In the interpolation of the /u/untargeted straw condition, a greater narrowing was observed in this region. There is no representation of the oral cavity for the targeted condition due to the missing data from the MRI image acquisition.

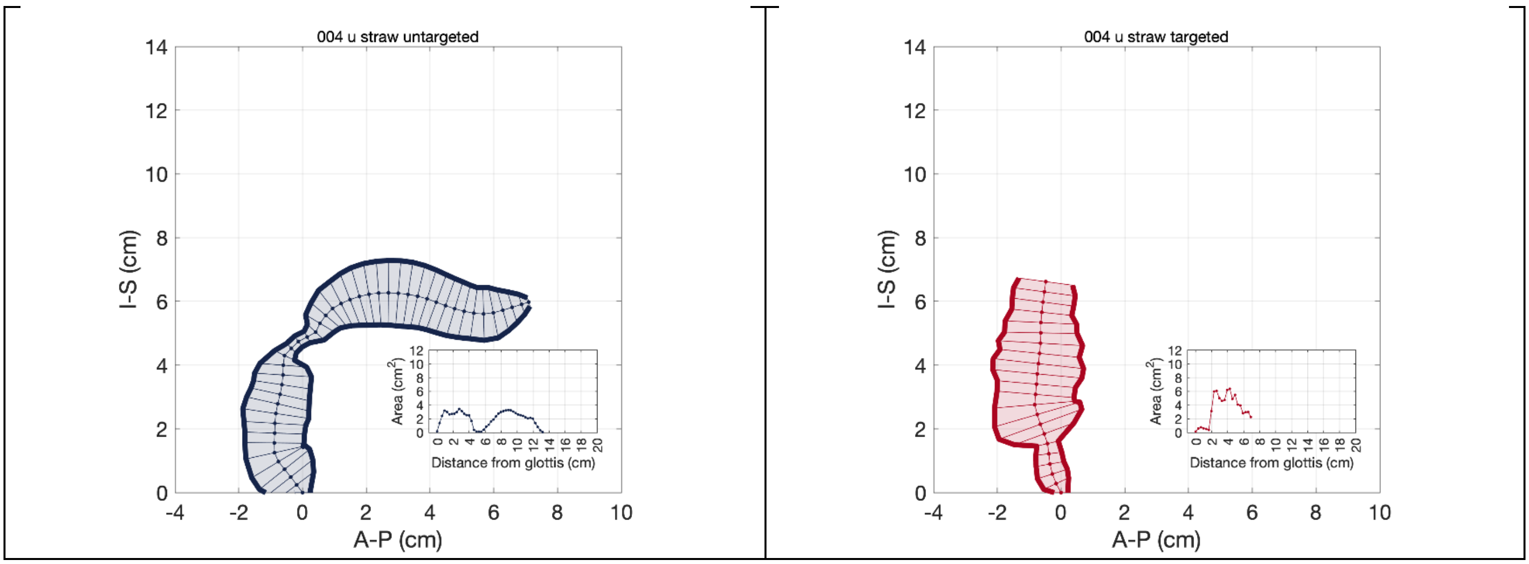


Figure 4.32: Pseudo mid-sagittal plots for Participant 04 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). /u/ straw targeted has no representation of the oral cavity due to missing data

Area comparison of /u/ with straw (Participant 04)

Figure 4.33 is the area function for /u/ straw with and without a vibratory focus. Targeting the /u/ during the straw condition resulted in a longer and more tube-like shape of the epilarynx (0 to ~1.9cm from the glottis) compared to the untargeted production. Targeting also led to a larger A_{ph}. Due to image acquisition errors, there are ten missing slices from the oral cavity and no measurable oral cavity area (Figure 4.31). No interpolation could be created for the oral cavity, given that there were no raw DICOM images to reference. Consequently, the area function shows "no oral cavity."

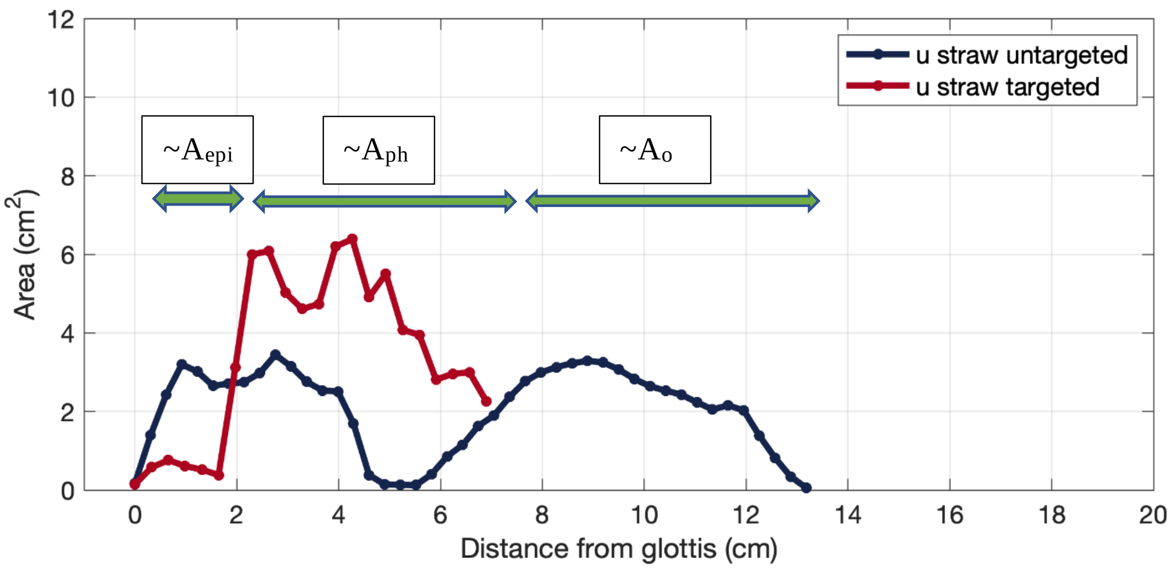


Figure 4.33: Smoothed area function for Participant 04 in the /u/ straw untargeted and targeted conditions. No area for the oral cavity due to missing data. The green arrows show approximate regions of A_{epi}, A_p, and A_o.

Participant 04 /u/ vowel

Figure 4.34 shows the 3D reconstructions for Participant 04 during production of an /u/ vowel productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left and right panels shows a gap in the oropharynx from the uvula to the top portion of the pharynx. This resulted from no measurable air in this region. The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analysis was carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation of the glottis to the lips.

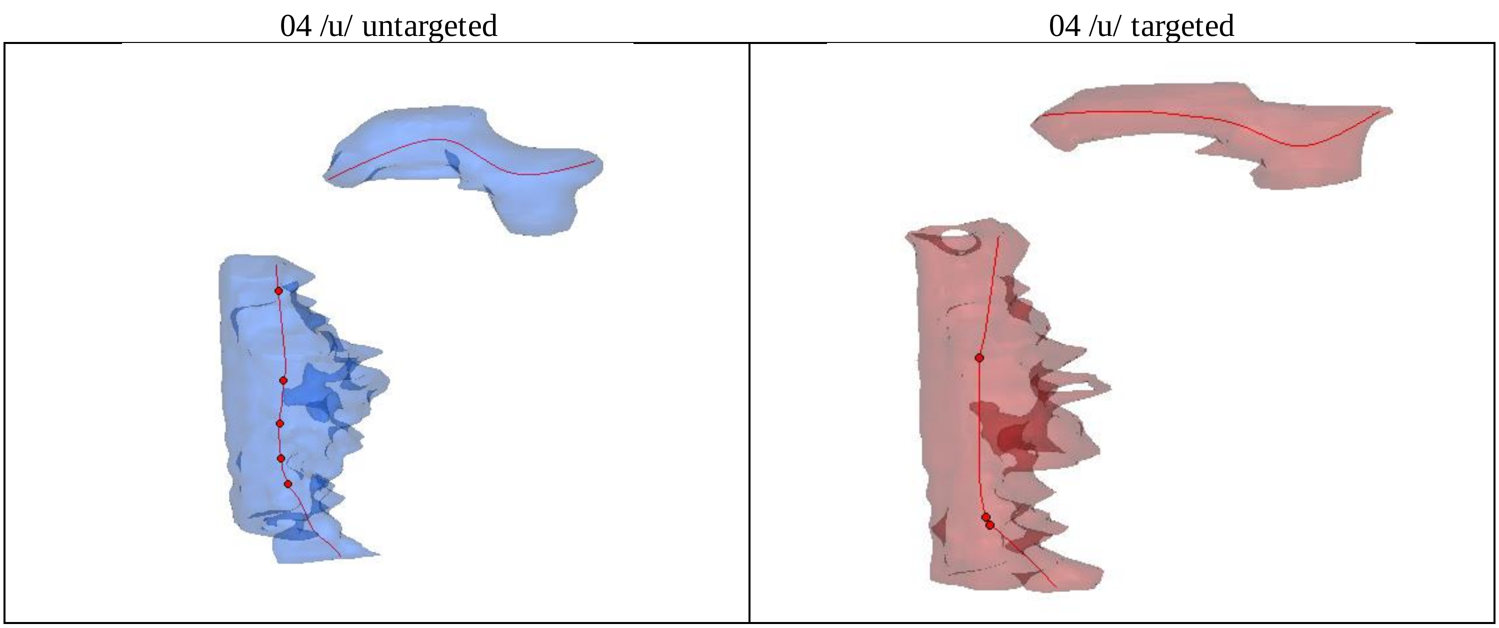


Figure 4. 34: 3D model reconstructions for Participant 04 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right). Gap in 3D model for /u/ straw untargeted and targeted conditions due to no measurable air in the oropharynx.

Figure 4.35 is the pseudo mid-sagittal plots for the /u/ vowel in the untargeted (left panel) and targeted (right panel). Due to the lack of measurable air in the raw DICOM, an interpolation of the vocal tract area from approximately 6 cm to 8 cm from the glottis was created. In the interpolation of the /u/ targeted condition shows a greater expansion in this region. This expansion is due to a greater area above and below the missing sections. The interpolation showing the expansion in the targeted condition at 6 cm to 8 cm from the glottis does appear plausible, given the general expansions seen elsewhere in the vocal tract for this condition. The lip terminations for both the untargeted and targeted /u/ vowel productions show different oral configurations. The lip termination and A_o area were greater for the untargeted condition. The narrowing seen at 6 cm to 8 cm from the glottis is likely the reason for the significant increase in the A_o for the /u/ untargeted condition.

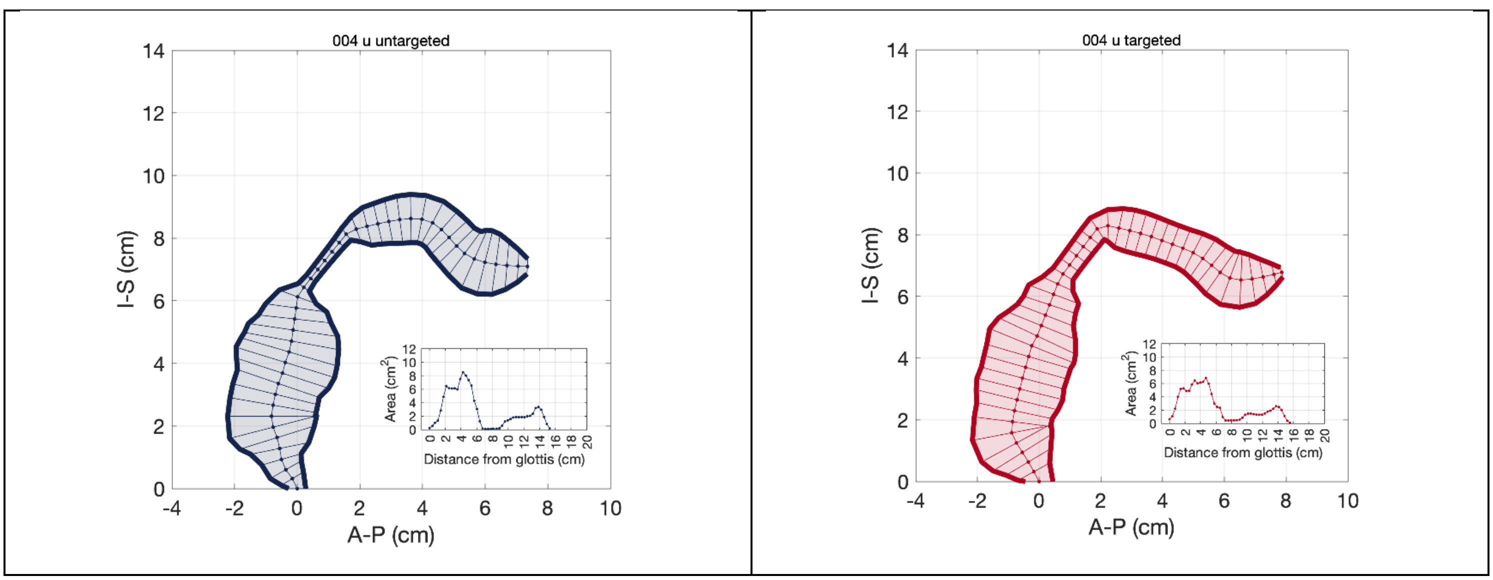


Figure 4. 35: Pseudo mid-sagittal plots for Participant 04 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right). Mid-sagittal plots show an interpolation of the oropharynx from approximately 6 cm to 9 cm for both conditions.

Area comparison of /u/ (Participant 04)

Targeting the /u/ resulted in a larger A_{epi} with a relatively similar geometric configuration to the untargeted production. Again, the zero values on the area function do not indicate a true area of zero. Rather, there was no measurable air in the raw DICOM; therefore, on the area function, a zero value is given. (Fig. 4.36 black box). Targeting also led to smaller pharyngeal and oral cavity areas.

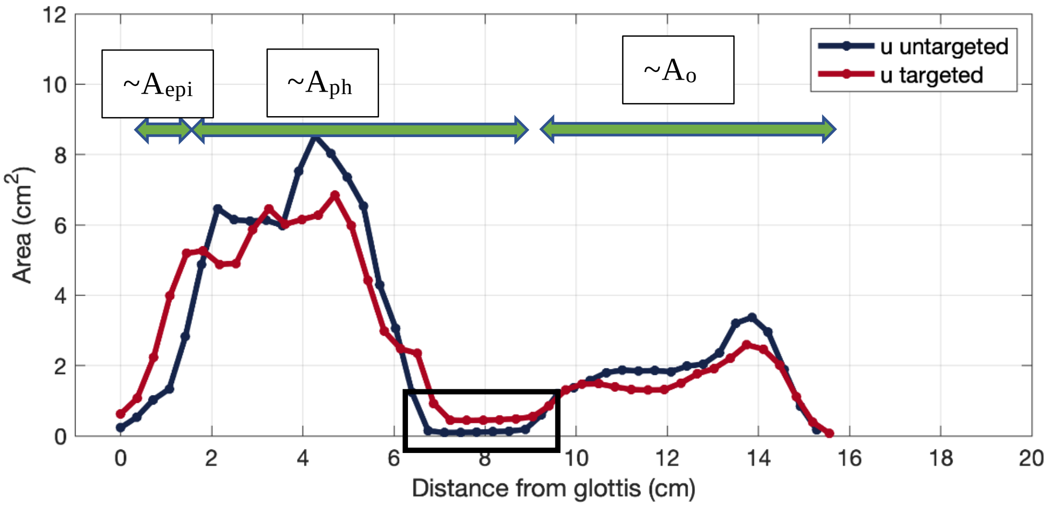


Figure 4.36: Smoothed area function for Participant 04 in the /u/ untargeted and targeted conditions. At approximately 7 cm to 9 cm represents the region with no measurable air in this region for both conditions. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Area comparison of /u/ with and without straw (Participant 04)

Straw with targeted resulted in the most uniform and longest epilaryngeal tube. Interestingly, the largest A_{ph} was seen in the /u/ untargeted without straw. The longest vocal tract length was not seen in the artificially lengthened conditions. Table 13 presents all the raw area values for Figure 4.37.

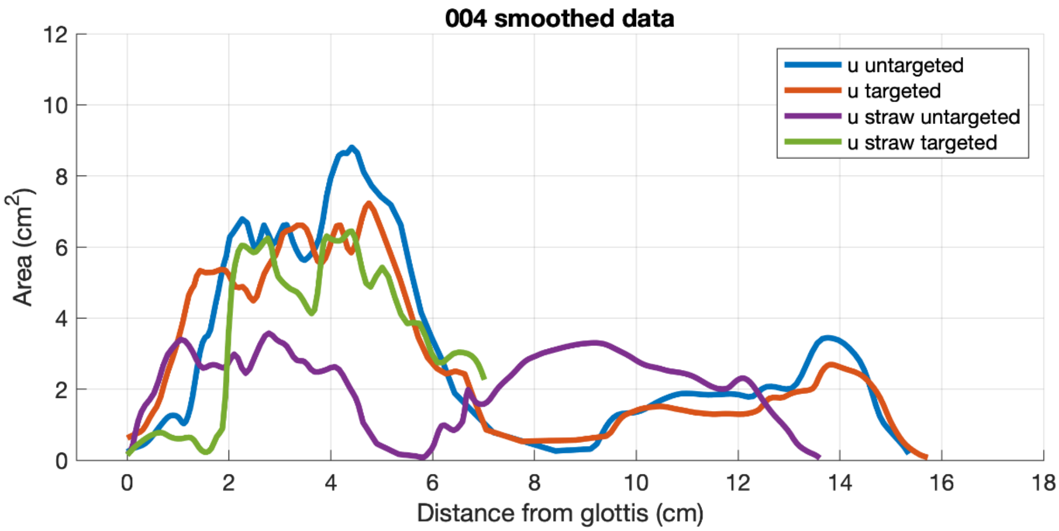


Figure 4.37: Combined area function smoothed participant 04.

Table 13: Participant 04 area function raw data.

Section number (04)	/u/ untargeted	/u/ targeted	/u/ straw untargeted	/u/ straw targeted
1 (glottis)	0.23	0.62	0.17	0.14
2	0.53	1.07	1.39	0.58
3	1.02	2.23	2.42	0.76
4	1.33	3.99	3.2	0.61
5	2.82	5.19	3.02	0.52
6	4.87	5.27	2.66	0.38
7	6.46	4.88	2.71	3.12
8	6.14	4.9	2.75	6
9	6.11	5.86	2.98	6.08
10	6.13	6.46	3.44	5.02
11	5.98	6.02	3.15	4.61
12	7.52	6.15	2.76	4.74
13	8.54	6.27	2.53	6.2
14	8.03	6.84	2.51	6.39
15	7.36	5.98	1.69	4.91
16	6.54	4.43	0.39	5.5
17	4.29	2.98	0.14	4.08
18	3.06	2.47	0.14	3.95
19	1.24	2.35	0.13	2.81
20	0.14	0.93	0.4	2.96
21	0.09	0.45	0.86	2.99
22	0.1	0.44	1.15	2.27
23	0.11	0.45	1.63	0
24	0.12	0.46	1.9	0
25	0.14	0.49	2.37	0
26	0.18	0.54	2.77	0
27	0.59	0.86	2.99	0
28	1.2	1.31	3.13	0
29	1.37	1.47	3.23	0
30	1.58	1.48	3.29	0
32	1.8	1.39	3.25	0
32	1.87	1.32	3.07	0
33	1.85	1.3	2.83	0
34	1.86	1.31	2.64	0
35	1.83	1.5	2.53	0
36	1.98	1.76	2.43	0
37	2.04	1.91	2.23	0
38	2.37	2.21	2.05	0
39	3.2	2.59	2.16	0
40	3.37	2.47	2.03	0
41	2.95	2.02	1.39	0
42	1.88	1.12	0.82	0
43	0.84	0.39	0.35	0
44 (lip termination)	0.18	0.08	0.06	0
45 (length)	0.36	0.36	0.31	0.33

Participant 04 /m/

Figure 4.38 shows the 3D reconstructions for Participant 04 during /m/ productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left panel shows a missing oral cavity. This resulted from no measurable air in the oral cavity. While physiologically impossible to have zero air present in the oral cavity, participant 04 likely occluded the oral cavity with their tongue and directed all airflow through the nose.

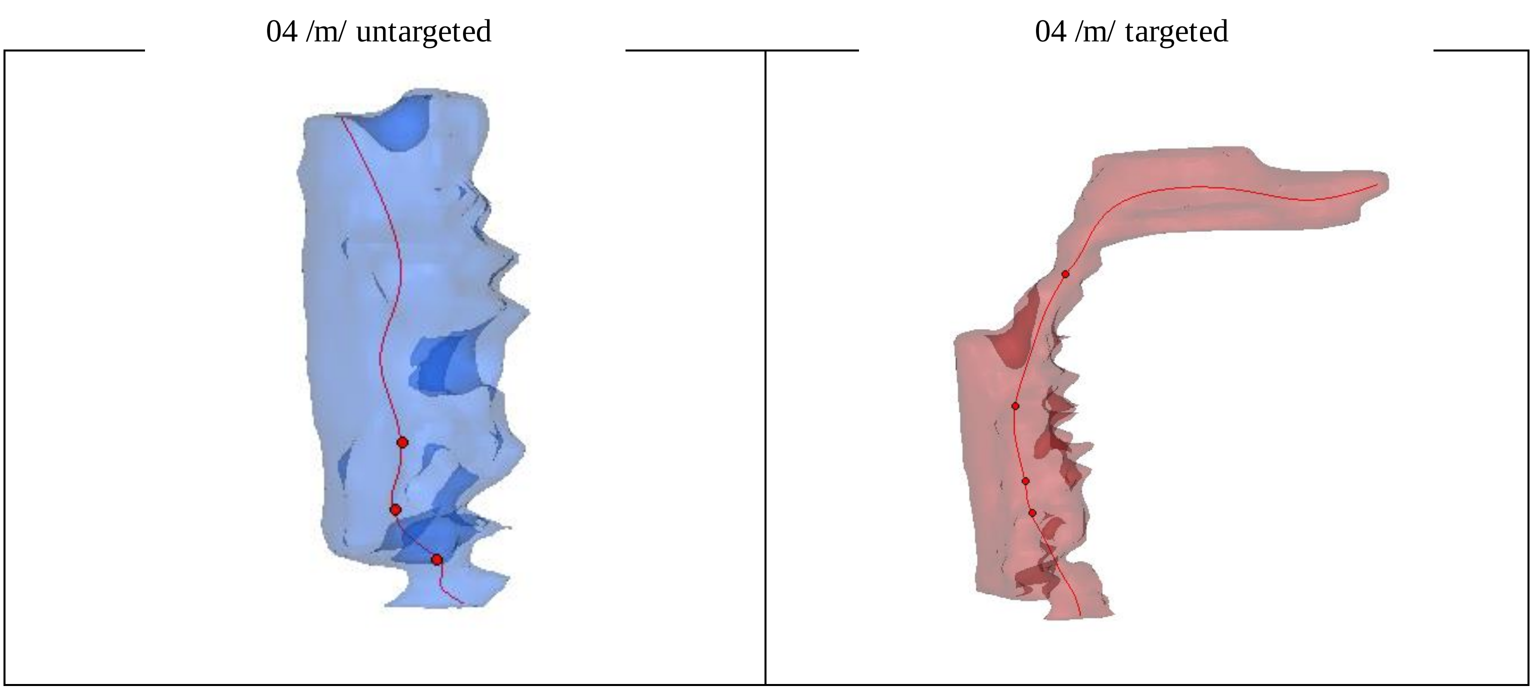


Figure 4.38: 3D model reconstructions for Participant 04 during production of an /m/ in the untargeted (left) and targeted conditions (right). Gap in 3D model for /m/ untargeted condition due to no measurable air in the oral cavity.

Figure 4.39 is of the pseudo mid-sagittal plots for the /m/ productions for both the untargeted (left panel) and targeted (right panel) conditions. The untargeted condition had a larger A_{ph} with no measurable air beginning 5.5 cm from the glottis. The targeted condition has a more defined epilaryngeal shape, regions of greater pharyngeal expansions, and a larger oral cavity. The lip termination for the targeted condition does not show complete lip occlusion. It is possible for this participant that a breath was taken or a spreading of the lips was captured during the MRI image acquisition.

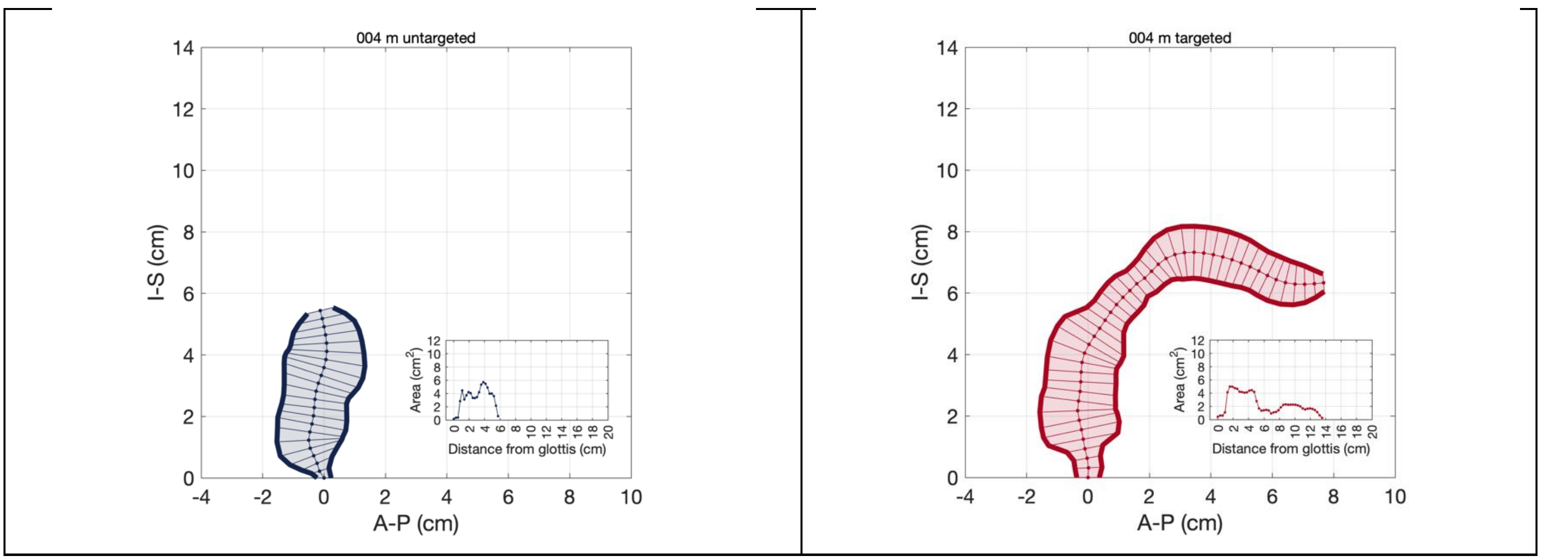


Figure 4.39: Pseudo mid-sagittal plots for Participant 04 during production of an /m/ in the untargeted (left) and targeted conditions (right). /m/ untargeted has no representation of the oral cavity due to missing data.

Area comparison for /m/ (Participant 04)

Figure 4.40, the area function for the /m/ conditions, showed that targeting the /m/ resulted in a larger area and similar length and shape of the epilarynx compared to the untargeted production. The cross-sectional areas of the pharynx are similar in both conditions. However, there were localized expansions closer to the glottis (1.8cm to 3.8cm from the glottis) in the targeted condition. Participant 04 did not have any measurable A_o for the /m/ untargeted production that was not the result of missing data. Targeting resulted in a longer vocal tract. Table 14 presents all the raw area values for Figure 4.40.

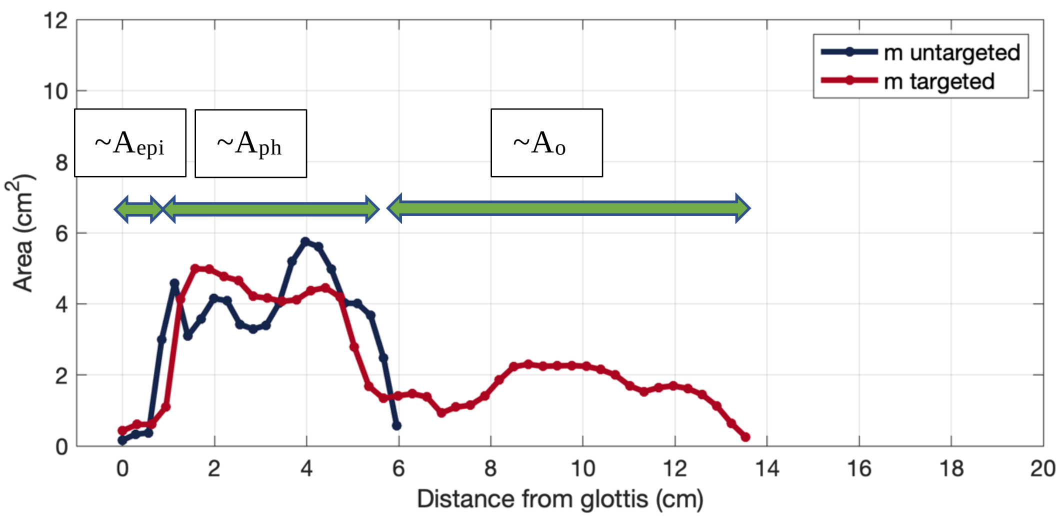


Figure 4.40: Smoothed area function for Participant 04 in the /m/ untargeted and targeted conditions. No area in the oral cavity for untargeted condition due to missing data. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Table 14: Participant 04 area function raw data for /m/ in cm^2

Section number	/m/ untargeted	/m/ targeted
1 glottis	0.16	0.43
2	0.33	0.61
3	0.37	0.61
4	2.84	1.1
5	4.46	4.13
6	3.09	4.99
7	3.73	4.97
8	4.21	4.77
9	4.06	4.65
10	3.34	4.21
11	3.31	4.16
12	3.45	4.07
13	4.19	4.11
14	5.35	4.37
15	5.71	4.45
16	5.48	4.19
17	4.89	2.79
18	3.97	1.68
19	3.99	1.34
20	3.61	1.41
21	2.17	1.47
22	0.57	1.38
23	0	0.93
24	0	1.09
25	0	1.15
26	0	1.4
27	0	1.85
28	0	2.23
29	0	2.3
30	0	2.24
31	0	2.26
32	0	2.27
33	0	2.24
34	0	2.15
35	0	2
36	0	1.69
37	0	1.53
38	0	1.65
39	0	1.69
40	0	1.62
41	0	1.44
42	0	1.13
43	0	0.64
44 lip termination	0	0.26
45 section length	0.27	0.31

Participant 04 Area Summary

For participant 04, a larger A_{epi} was observed in two of the three targeted conditions (/u/, /m/). Interestingly, for this participant, two of their targeted productions resulted in either no measurable A_o or a decrease in A_o. This participant likely interpreted the direction of feeling vibratory focus by pulling the tongue forward or allowing the tongue to fill the majority of the oral cavity space.

4.3.4. Participant 06

Participant 06 /u/ straw

Figure 4.41 shows the 3D reconstructions for Participant 06 during /u/ vowel productions with straw. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left panel shows a gap from the oropharynx from the uvula to the top portion of the pharynx. The reconstruction in the right panel shows a smaller gap region with no measurable airspace from the soft palate to the hard palate. The data from the centerlines in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the area function analyses were carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

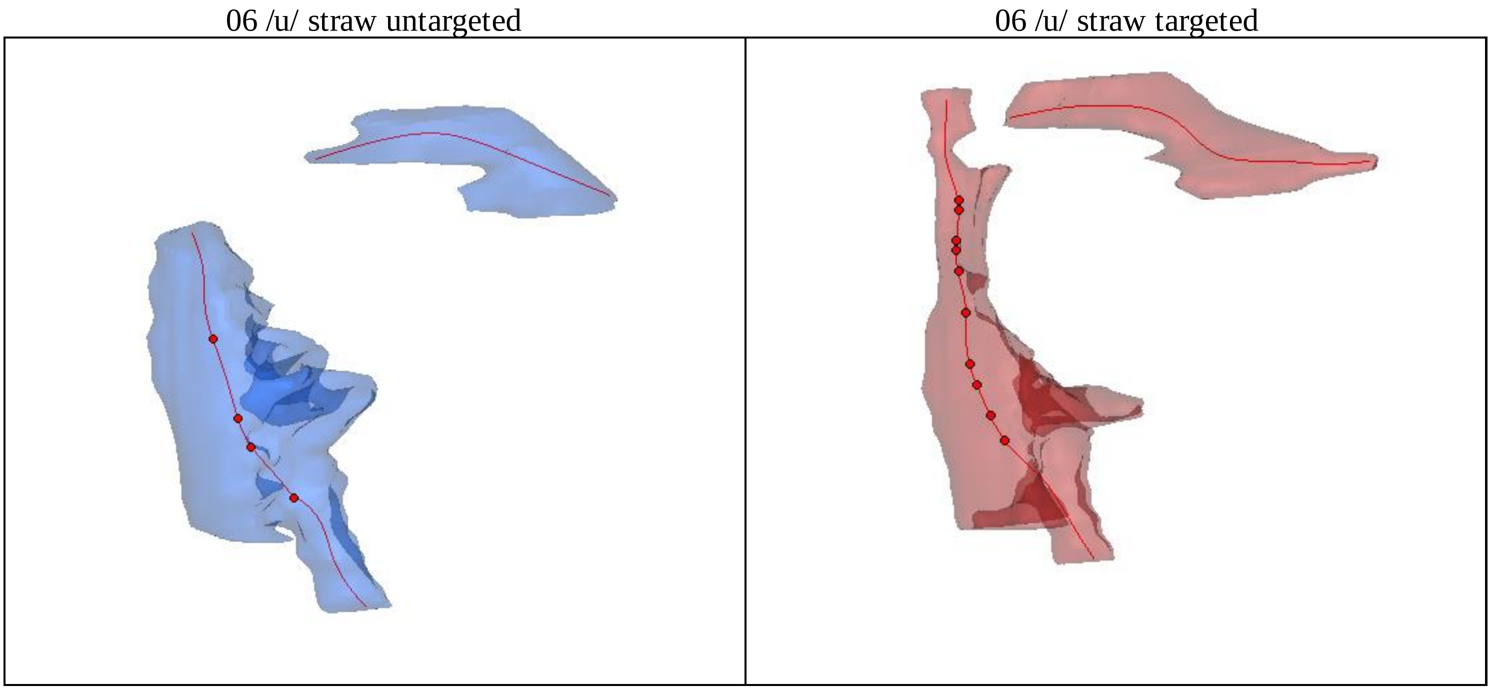


Figure 4.41: 3D model reconstructions for Participant 06 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). Gap in 3D model for /u/ straw untargeted condition due to no measurable air in the oropharynx.

Figure 4.42 is the pseudo mid-sagittal plots for the /u/ vowel with straw in the untargeted (left panel) and targeted (right panel). Due to the lack of measurable air in the raw DICOM for the untargeted condition, an interpolation of the vocal tract area from approximately 6 cm to 7 cm from the glottis was created. The interpolation of the /u/ straw untargeted condition showed a narrowing in this region. This narrowing appeared to be due to a significant narrowing in the sections just below and just above the missing region. The lip terminations for the untargeted and targeted /u/ straw productions show different oral configurations. The lip termination and A_o were greater for the targeted condition due to greater lip protrusion.

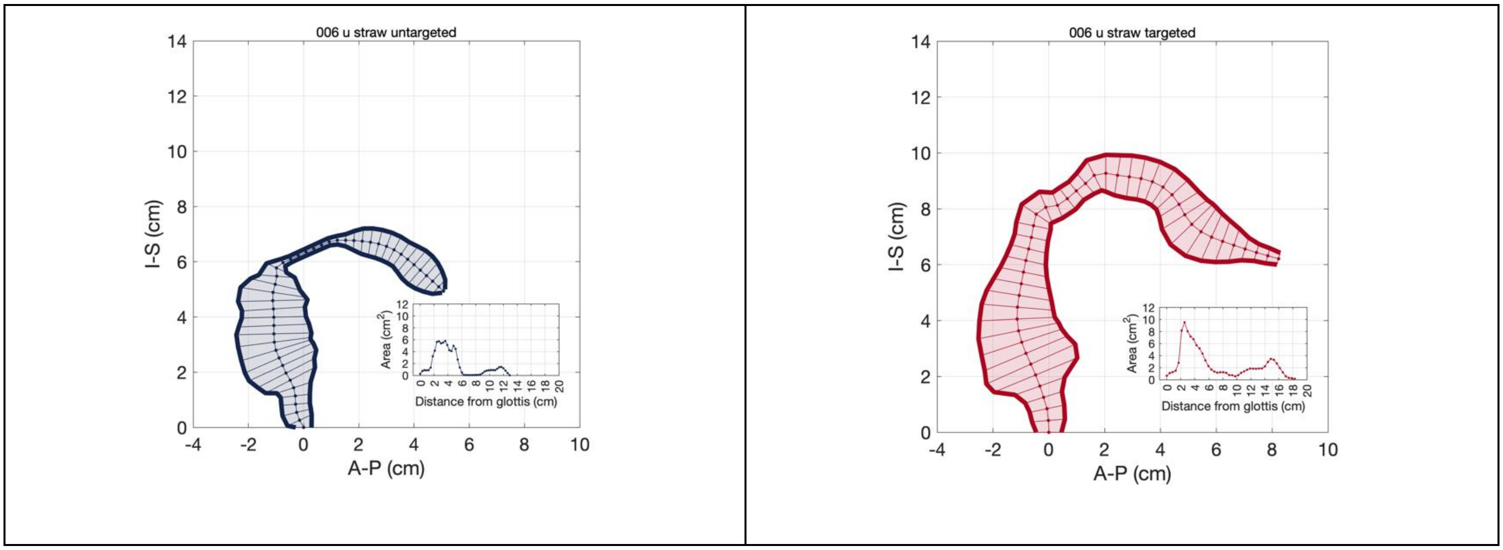


Figure 4.42: Pseudo mid-sagittal plots for Participant 06 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). Mid-sagittal plots show an interpolation of the oropharynx from approximately 6 cm to 9 cm for the untargeted condition and approximately 9 cm for the targeted condition.

Area comparison of /u/ with straw (Participant 06)

Targeting the /u/ during the straw condition resulted in a larger area of the epilarynx with the same length as the untargeted production (Fig. 4.43). Epilaryngeal shape was similar for both conditions. Targeting led to larger pharyngeal and oral cavity areas. In the untargeted condition, a lack of measurable air space in the raw DICOM images caused zero values in the cross-sectional area. Therefore, the pseudo mid-sagittal plot also had a zero value from 6.2 cm to 8.8 cm from the glottis. While we cannot confirm air was present in this space, it is likely the area in this region would have been similar to the /u/ straw targeted condition. The overall length of the vocal tract was longer (18.2 cm) than the untargeted condition (13 cm), secondary to lip protrusion.

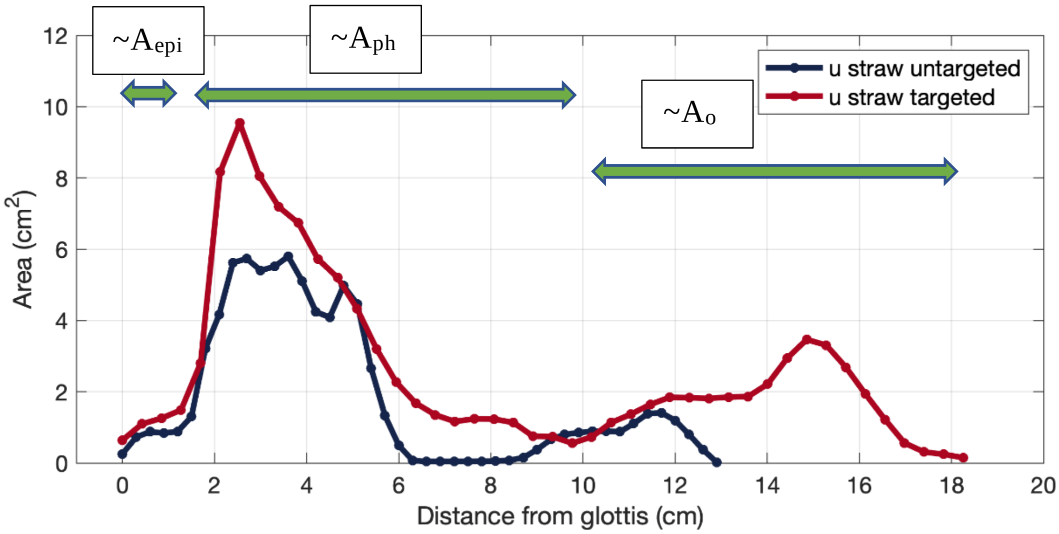


Figure 4.43: Smoothed area function for Participant 06 in the /u/ straw untargeted and targeted conditions. At approximately 6 cm to 9 cm represents the region with no measurable air and therefore no area. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Participant 06 /u/

Figure 4.44 shows the 3D reconstructions for Participant 06 during /u/ vowel productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus").

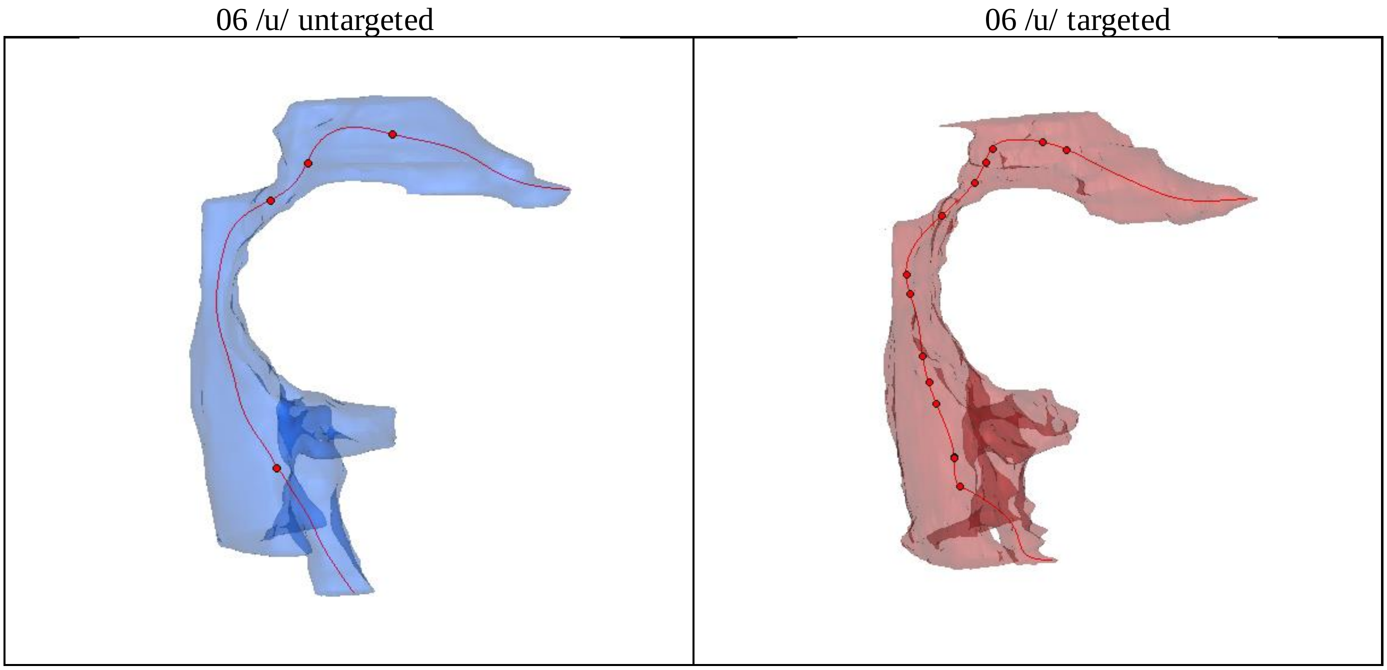


Figure 4.44: 3D model reconstructions for Participant 06 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right).

Figure 4.45 is the pseudo mid-sagittal plots for the /u/ vowel in the untargeted (left panel) and targeted (right panel). Similar vocal tract shapes were seen in both the /u/ vowel conditions, with the greatest difference between them being general expansions of each region of the vocal tract. The lip terminations for both conditions were similar, but more lengthening of the oral cavity due to lip elongation was seen.

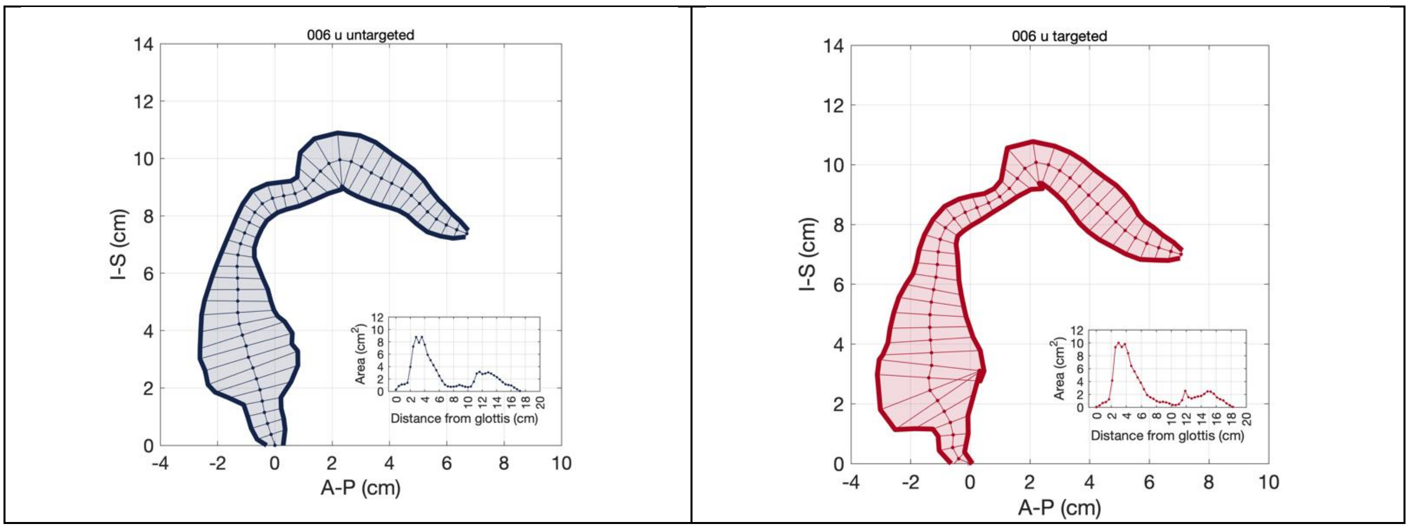


Figure 4.45: Pseudo mid-sagittal plots for Participant 06 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right).

Area comparison of /u/ (Participant 06)

Targeting the /u/ resulted in a smaller area and the same length of the epilarynx compared to the untargeted production (Fig. 4.46). Targeting the /u/ also led to an increase in A_{ph} and a decrease in the A_o. Overall, the vocal tract was longer (18.1 cm) in the targeted condition than in the untargeted condition (17.3 cm), secondary to lip protrusion.

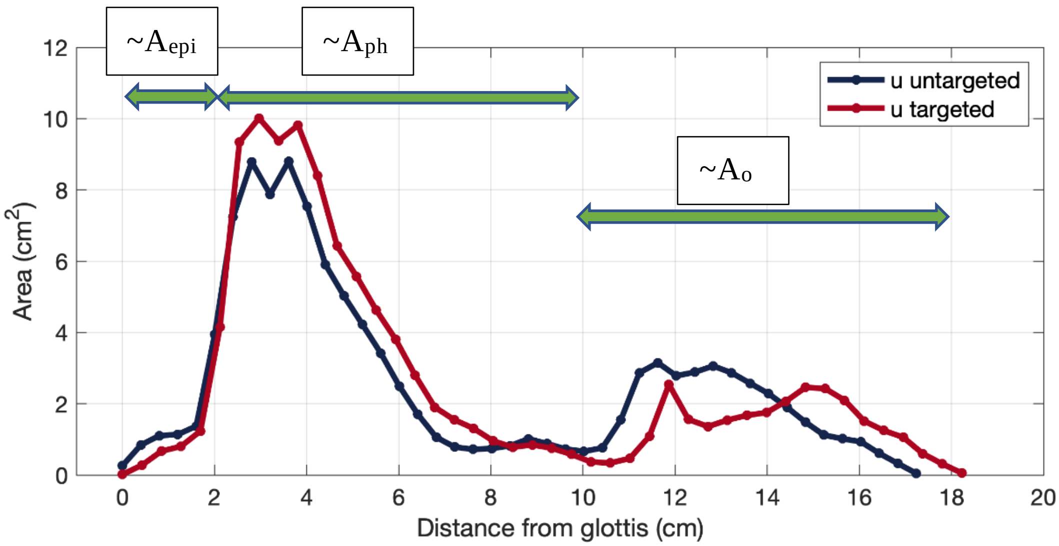


Figure 4. 46: Smoothed area function for Participant 06 in the /u/ untargeted and targeted conditions.

Area comparison of /u/ with and without straw (Participant 06)

Epilaryngeal shape and length were consistent across all /u/ conditions, irrespective of straw. The longest vocal tract was seen for the /u/ straw targeted condition but was close to the same length for the /u/ targeted condition. Targeting overall created the longest vocal tract length and increased pharyngeal and oral cavity area with and without straw (Fig. 4.47). Table 15 presents all the raw area values for Figure 4.47.

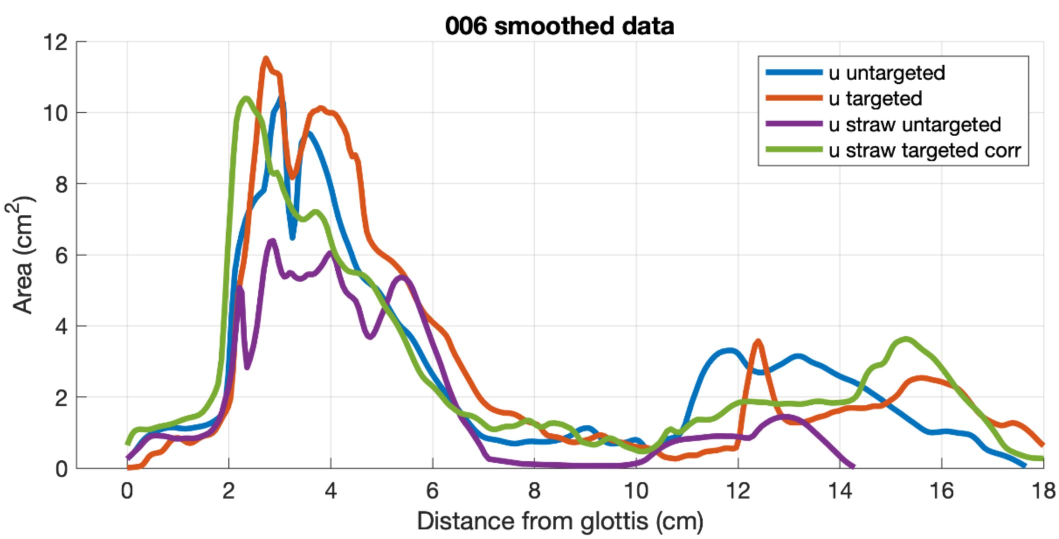


Figure 4.47: Combined area function smoothed participant 06.

Table 15: Participant 06 area function raw data.

Section number (06)	/u/ untargeted	/u/ targeted	/u/ straw untargeted	/u/ straw targeted
1 (glottis)	0.26	0.02	0.26	0.64
2	0.84	0.28	0.73	1.1
3	1.1	0.67	0.88	1.26
4	1.14	0.8	0.84	1.49
5	1.37	1.24	0.89	2.81
6	3.94	4.16	1.3	8.17
7	7.24	9.34	3.21	9.54
8	8.79	10.02	4.16	8.06
9	7.88	9.38	5.62	7.19
10	8.8	9.82	5.73	6.73
11	7.54	8.4	5.4	5.72
12	5.9	6.43	5.51	5.21
13	5.03	5.57	5.8	4.33
14	4.23	4.63	5.11	3.2
15	3.42	3.81	4.24	2.27
16	2.5	2.8	4.09	1.68
17	1.71	1.91	4.97	1.35
18	1.06	1.55	4.46	1.17
19	0.79	1.31	2.66	1.24
20	0.72	0.96	1.33	1.23
21	0.75	0.78	0.49	1.14
22	0.82	0.85	0.07	0.75
23	1.02	0.75	0.04	0.74
24	0.88	0.58	0.04	0.56
25	0.73	0.37	0.05	0.73
26	0.66	0.34	0.05	1.14
27	0.76	0.47	0.05	1.37
28	1.55	1.09	0.06	1.64
29	2.87	2.54	0.08	1.85
30	3.15	1.56	0.15	1.83
321	2.78	1.36	0.38	1.81
32	2.89	1.54	0.66	1.85
33	3.06	1.68	0.8	1.87
34	2.87	1.75	0.85	2.22
35	2.58	2.07	0.89	2.94
36	2.28	2.46	0.89	3.47
37	1.9	2.43	0.88	3.31
38	1.48	2.1	1.11	2.69
39	1.12	1.52	1.38	1.94
40	1.02	1.25	1.41	1.22
41	0.93	1.06	1.19	0.56
42	0.61	0.6	0.8	0.31
43	0.32	0.31	0.38	0.25
44 (lip termination)	0.05	0.06	0.02	0.15
45 (length)	0.4	0.42	0.3	0.42

Participant 06 /m/

Figure 4.48 shows the 3D reconstructions for Participant 06 during /m/ productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left and right panels shows a gap from the oropharynx from the uvula to the top portion of the pharynx. The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analysis was carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

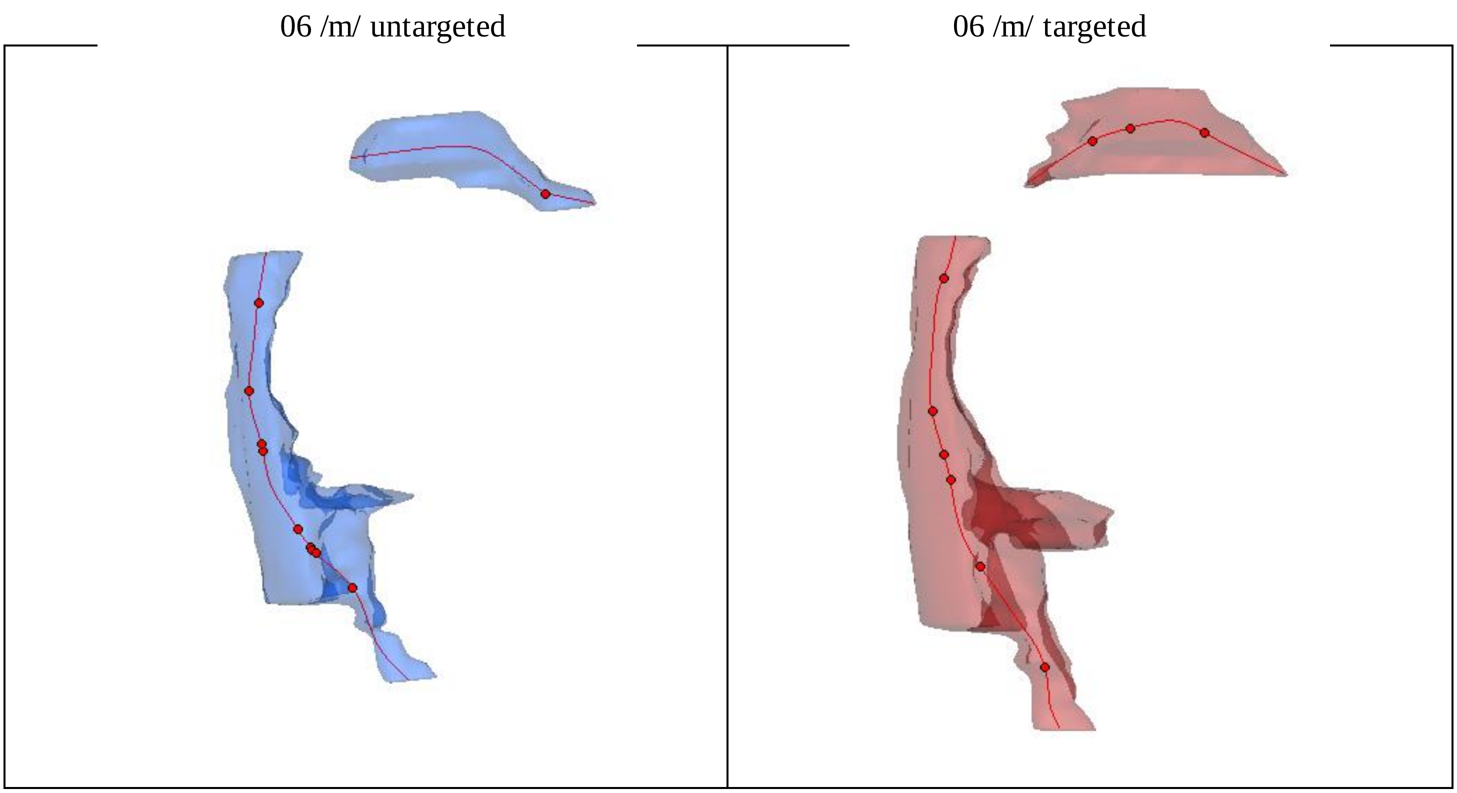


Figure 4.48: 3D model reconstructions for Participant 06 during production of an /m/ in the untargeted (left) and targeted conditions (right). Gap in 3D model for /m/ untargeted and targeted conditions due to no measurable air in the oropharynx.

Figure 4.49 is the pseudo mid-sagittal plots for the /m/ in the untargeted (left panel) and targeted (right panel) conditions. Due to the lack of measurable air in the raw DICOM, an interpolation of the vocal tract area from approximately 9 cm to 10 cm from the glottis was created. In the interpolation, narrowing in this region was seen for both untargeted and targeted productions. The lip termination shape for both the untargeted and targeted show different configurations, with neither case having a true occlusion at the lips. The A_o is greater in the untargeted condition.

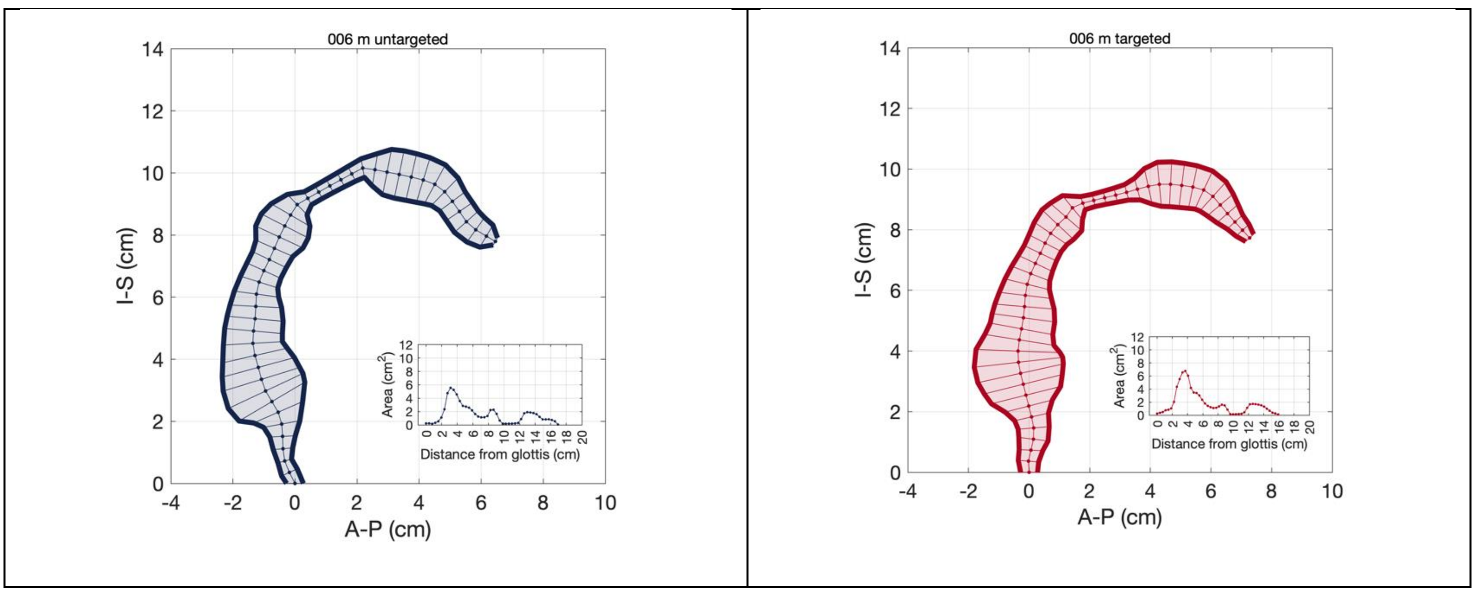


Figure 4.49: Pseudo mid-sagittal plots for Participant 06 during production of an /m/ in the untargeted (left) and targeted conditions (right). Mid-sagittal plots show an interpolation of the oropharynx from approximately 9 cm to 12 cm for both conditions.

Area comparison for /m/(Participant 06)

Targeting the /m/ resulted in a larger area and similar length of the epilarynx compared to the untargeted production. Targeting also led to a larger pharyngeal and comparable oral cavity area. There was little to no measurable air in the raw DICOM for either /m/ condition (9.8-12cm from the glottis); therefore, on the area function, a zero value is given. (Fig. 4.50). This can be seen visually on the axial and sagittal sections in figures 3.52-3.55. For the /m/ conditions, this participant likely put the majority of the air in the velopharynx/nasal cavity, which was not included in these area functions. Table 16 presents all the raw area values for Figure 4.50.

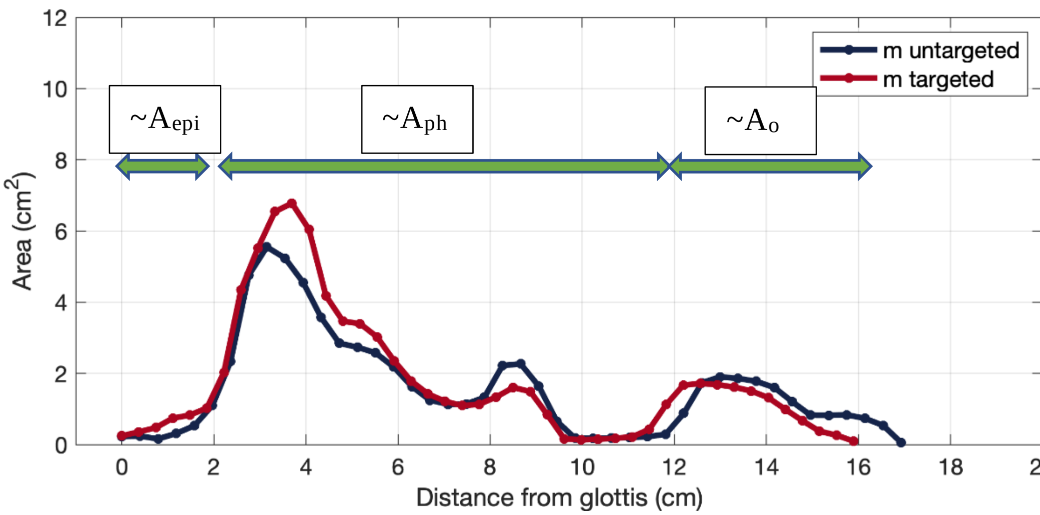


Figure 4.50: Are Smoothed area function for Participant 06 in the /m/ untargeted and targeted conditions. At approximately 9 cm to 12 cm represents the region with no measurable air and therefore, no area for both conditions. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

The untargeted production resulted in a longer vocal tract. The audio recordings from the MRI acquisition session and the DICOM images were examined to understand the differences in the vocal tract length between the two productions. Consensus was reached between the student and co-advisors that the tongue filled the oral cavity for these sections where no airspace was measured (Fig. 4.50) and nasal resonance was stronger for the increased nasal targeted /m/ production (Fig. 4.52 through Fig. 4.54).

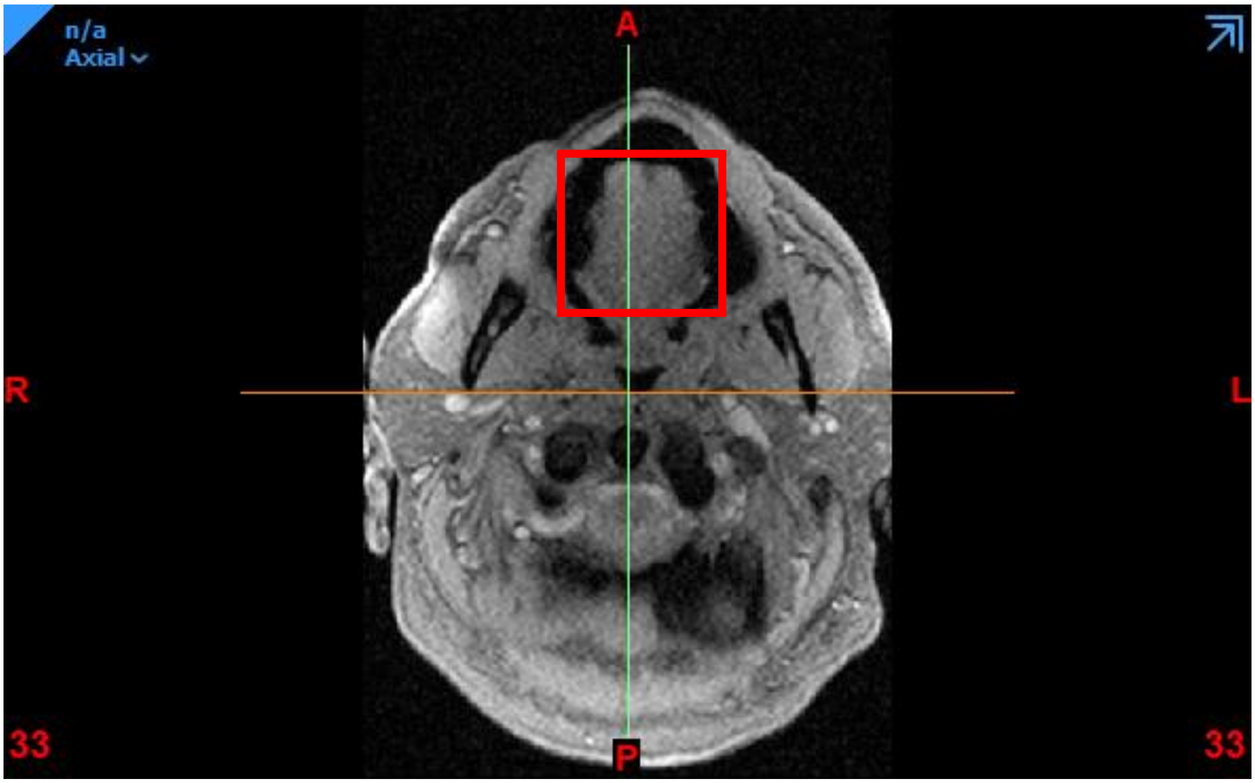


Figure 4.51: Raw DICOM in axial slice one. No measurable air in the oral cavity due to the tongue consuming the space highlighted in red box.

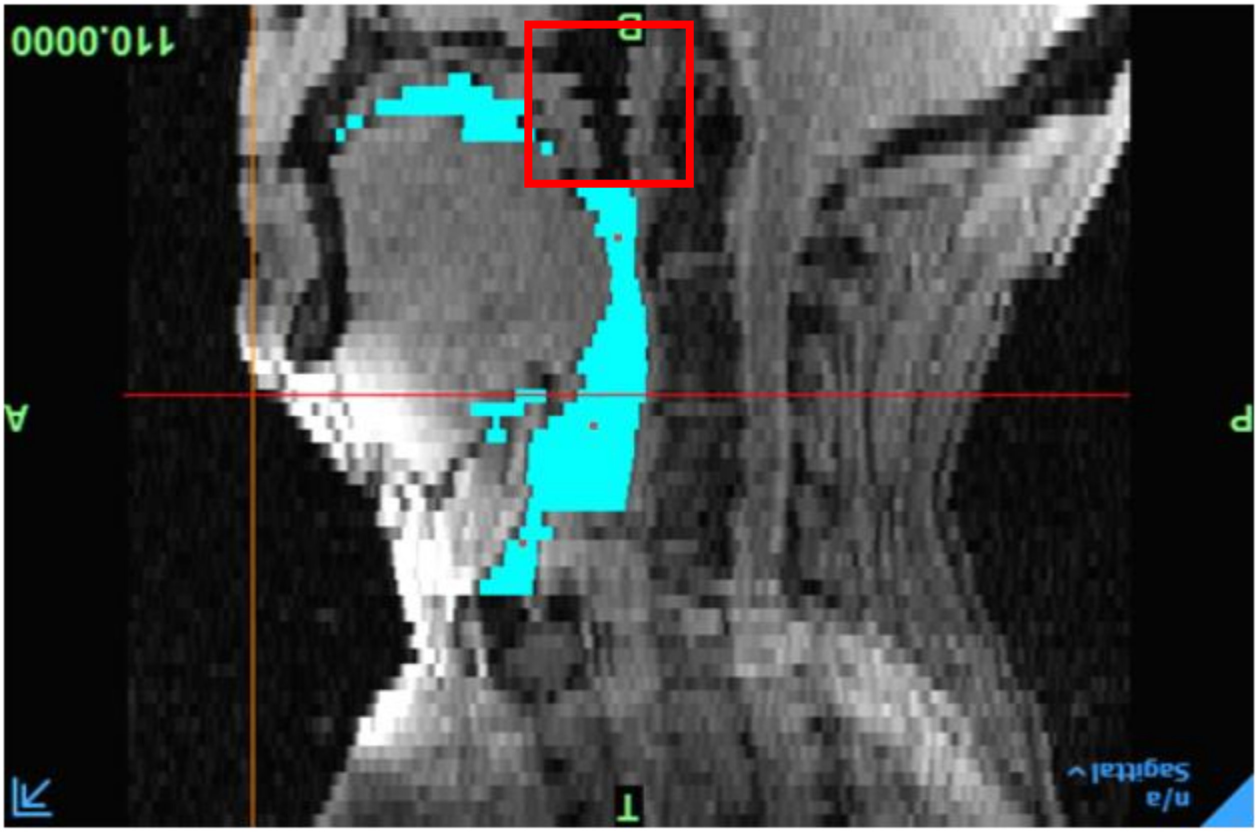


Figure 4.52: Raw DICOM sagittal slice. Blue highlighted area represents area of air for Participant 06 /m/ targeted. Red box shows the velopharyngeal port with air during /m/ production.

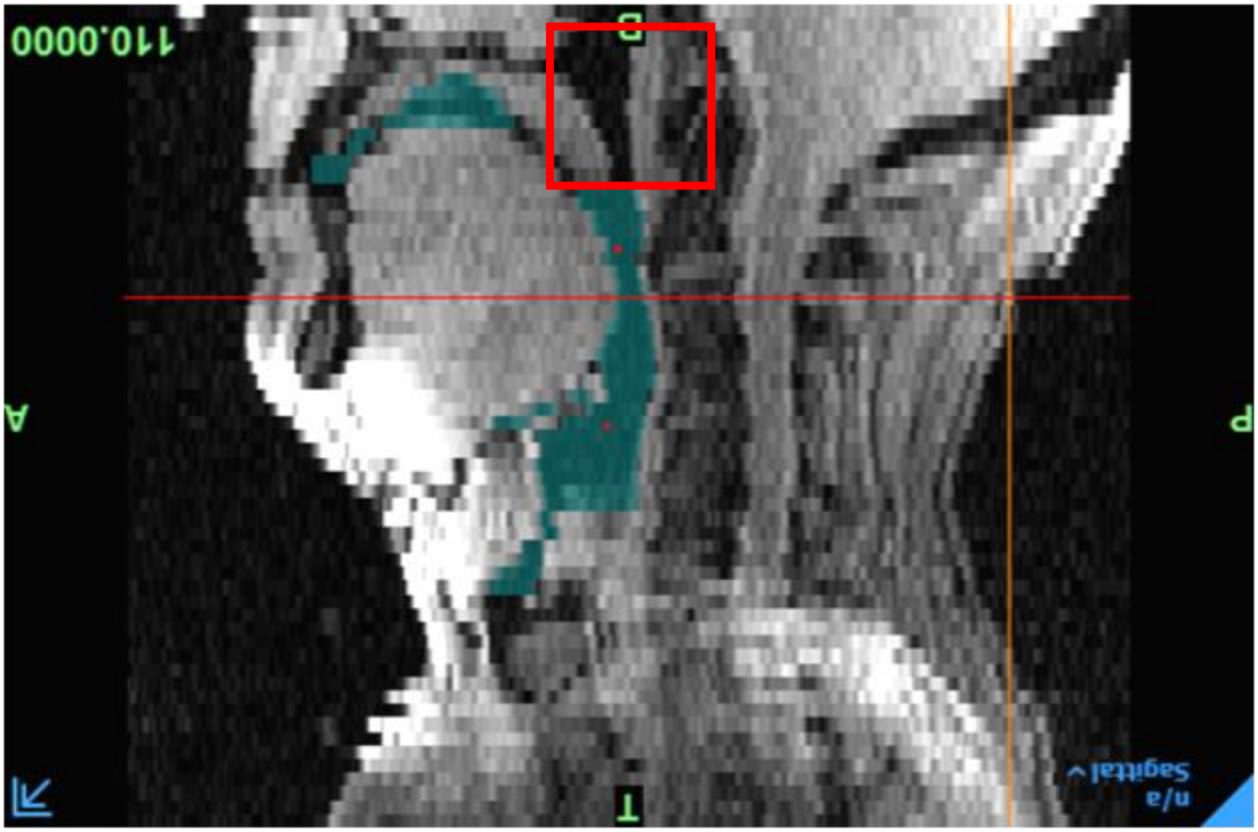


Figure 4.53: Raw DICOM sagittal slice. Blue highlighted area represents area of air for Participant 06 /m/ untargeted. Red box shows the velopharyngeal port with air during /m/ production.

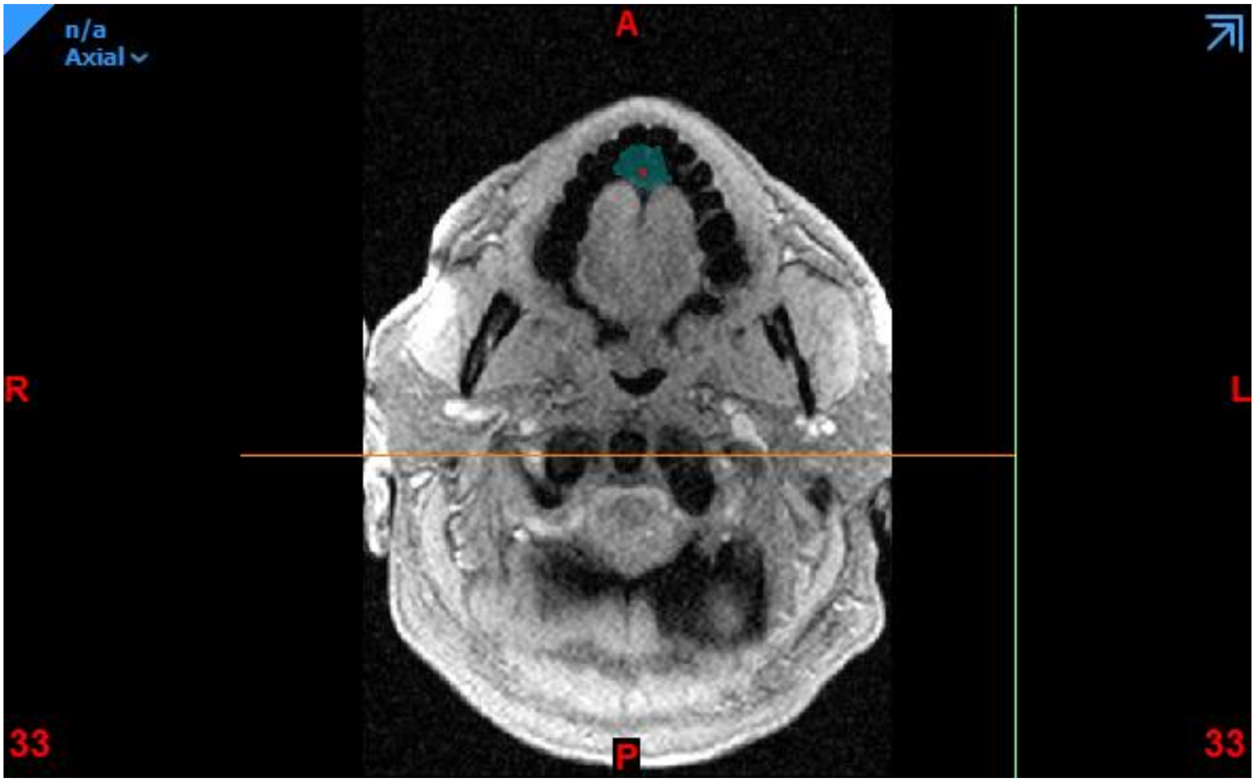


Figure 4.54: Raw DICOM in axial slice two. Blue highlighted area represents area of air in the oral cavity for Participant 06 /m/ targeted.

Table 16: Participant 06 area function raw data for /m/ in cm^2

Section number	/m/ untargeted	/m/ targeted
1 glottis	0.23	0.25
2	0.24	0.36
3	0.16	0.48
4	0.31	0.74
5	0.53	0.83
6	1.1	1.03
7	2.33	2.03
8	4.76	4.35
9	5.55	5.52
10	5.24	6.55
11	4.56	6.78
12	3.57	6.04
13	2.85	4.17
14	2.74	3.46
15	2.58	3.4
16	2.18	3.02
17	1.63	2.35
18	1.25	1.79
19	1.13	1.42
20	1.14	1.21
21	1.33	1.1
22	2.22	1.13
23	2.27	1.34
24	1.64	1.6
25	0.66	1.49
26	0.18	0.85
27	0.18	0.16
28	0.18	0.14
29	0.2	0.15
30	0.22	0.18
31	0.29	0.22
32	0.88	0.43
33	1.74	1.13
34	1.9	1.67
35	1.86	1.72
36	1.78	1.68
37	1.6	1.61
38	1.2	1.51
39	0.83	1.32
40	0.82	0.99
41	0.84	0.67
42	0.74	0.38
43	0.54	0.26
44 lip termination	0.05	0.1
45 section length	0.39	0.37

Participant 06 Area Summary

For participant 06, all targeted conditions increased the A_{ph} with no change in epilaryngeal length. Increased epilaryngeal and oral cavity areas were noted for two of the three targeted conditions. Targeting the /u/ production, irrespective of whether the straw was present, resulted in increased vocal tract length. The targeted production of /u/ with a straw resulted in the longest vocal tract.

4.3.5. Participant 07

<u>Participant 07 /u/ straw</u>

Figure 4.55 shows the 3D reconstructions for Participant 07 during /u/ vowel productions with straw. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the left panel shows a gap from the oropharynx from the uvula to the top portion of the pharynx due to no measurable air in this region. The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analysis was carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

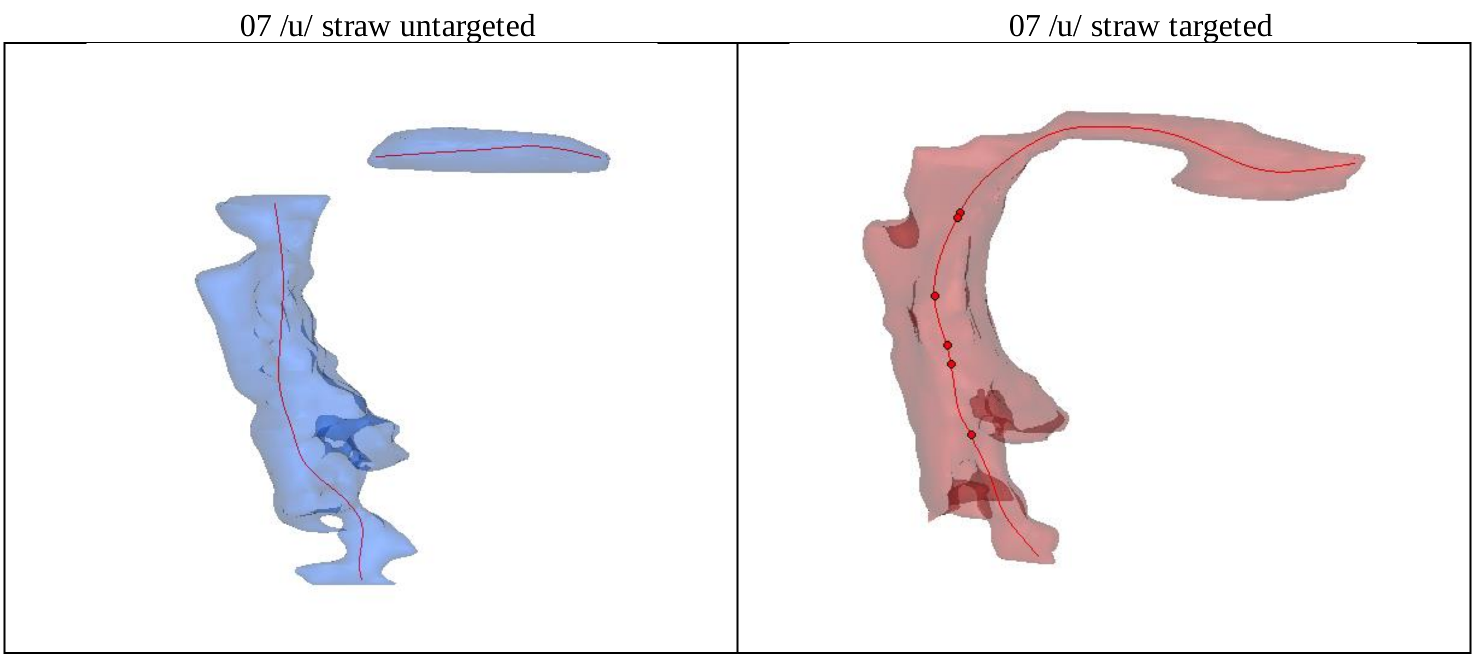


Figure 4. 55: 3D model reconstructions for Participant 07 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). Gap in 3D model for /u/ straw untargeted condition due to no measurable air in the oropharynx.

Figure 4.56 is the pseudo mid-sagittal plots for the /u/ straw conditions untargeted (left panel) and targeted (right panel). Due to the lack of measurable air in the raw DICOM for the untargeted condition, an interpolation of the vocal tract area from approximately 7 cm to 8 cm from the glottis was created. The interpolation of the /u/ straw untargeted condition shows a similar configuration to the targeted condition. The lip terminations for both the untargeted and targeted productions show different oral configurations. The lip termination and A_o area were greater for the untargeted condition at 9 cm to 11 cm from the glottis. However, a generally larger oral cavity area with more lip narrowing was seen in the targeted condition.

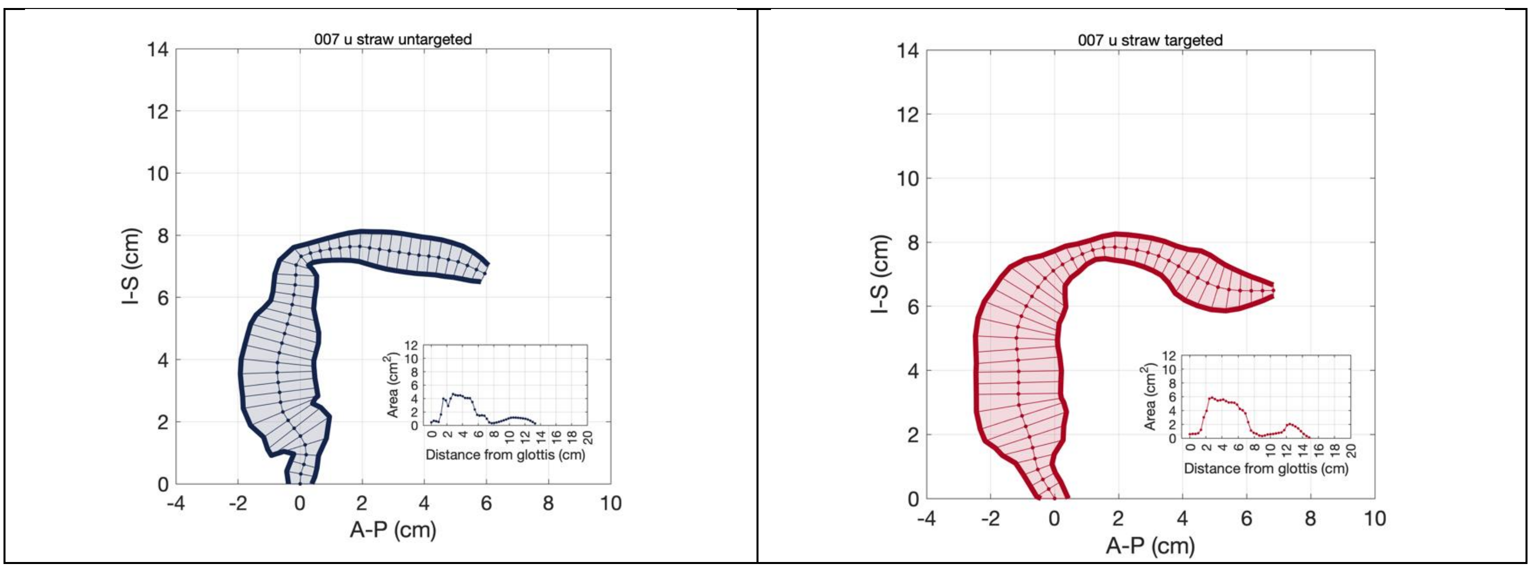


Figure 4. 56: Pseudo mid-sagittal plots for Participant 07 during production of an /u/ vowel with a straw in the untargeted (left) and targeted conditions (right). Mid-sagittal plots show an interpolation of the oropharynx from approximately 7 cm to 9 cm for untargeted conditions.

Area comparison of /u/ with straw (Participant 07)

Targeting the /u/ during the straw condition resulted in comparable epilaryngeal length and area (Fig. 4.57 black box). Targeting led to a larger pharyngeal and oral cavity areas. There was an overall lengthening of the vocal tract in the targeted condition (14.8 cm) compared to the untargeted production (13.8 cm), likely secondary to greater lip protrusion.

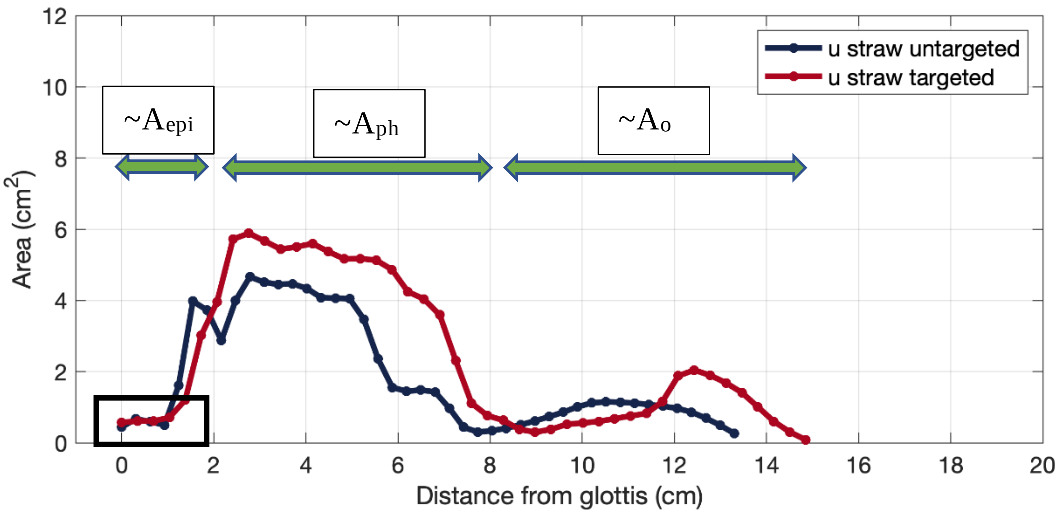


Figure 4.57: Smoothed area function for Participant 07 in the /u/ straw untargeted and targeted conditions. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

<u>Participant 07 /u/</u>

Figure 4.58 shows the 3D reconstructions for Participant 07 during /u/ vowel productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the right panel shows a gap in the oral cavity from the soft to the hard palate. This resulted from no measurable air in this region. The data in both the upper and lower portions of the reconstructions were, however, consistent with expectations regarding vocal tract shape, and the centerline and area function analyses were carried out separately for each section. The centerline data and area function data from each section were later concatenated to provide a contiguous representation from the glottis to the lips.

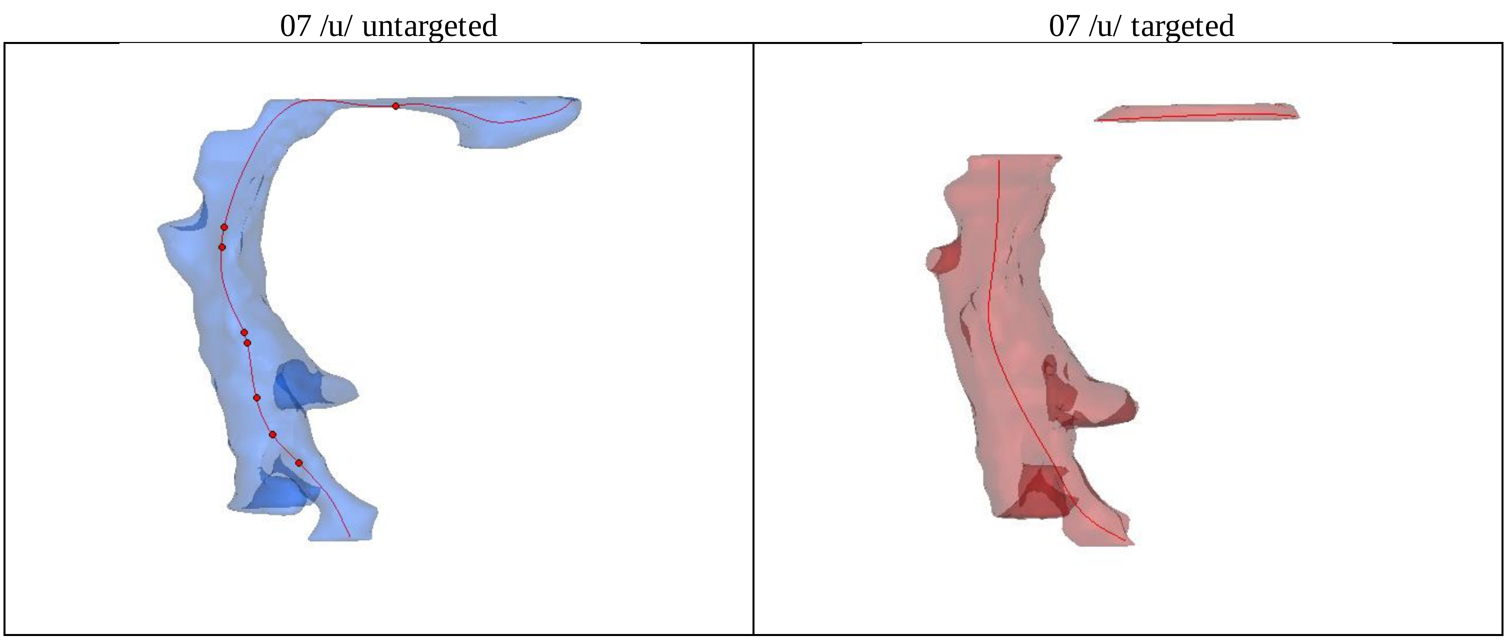


Figure 4.58: 3D model reconstructions for Participant 07 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right). Gap in 3D model for /u/ targeted condition due to no measurable air in the oral cavity.

Figure 4.59 is the pseudo mid-sagittal plots for the /u/ vowel in the untargeted (left panel) and targeted (right panel). Due to the lack of measurable air in the raw DICOM in the targeted condition, an interpolation of the vocal tract area from approximately 8 cm to 9 cm from the glottis was created. In the interpolation of the /u/ targeted condition shows a greater narrowing in this region. A more uniform oral cavity shape was observed in the targeted condition with a more narrowed lip occlusion. The lip termination and A_0 area were greater for the untargeted condition.

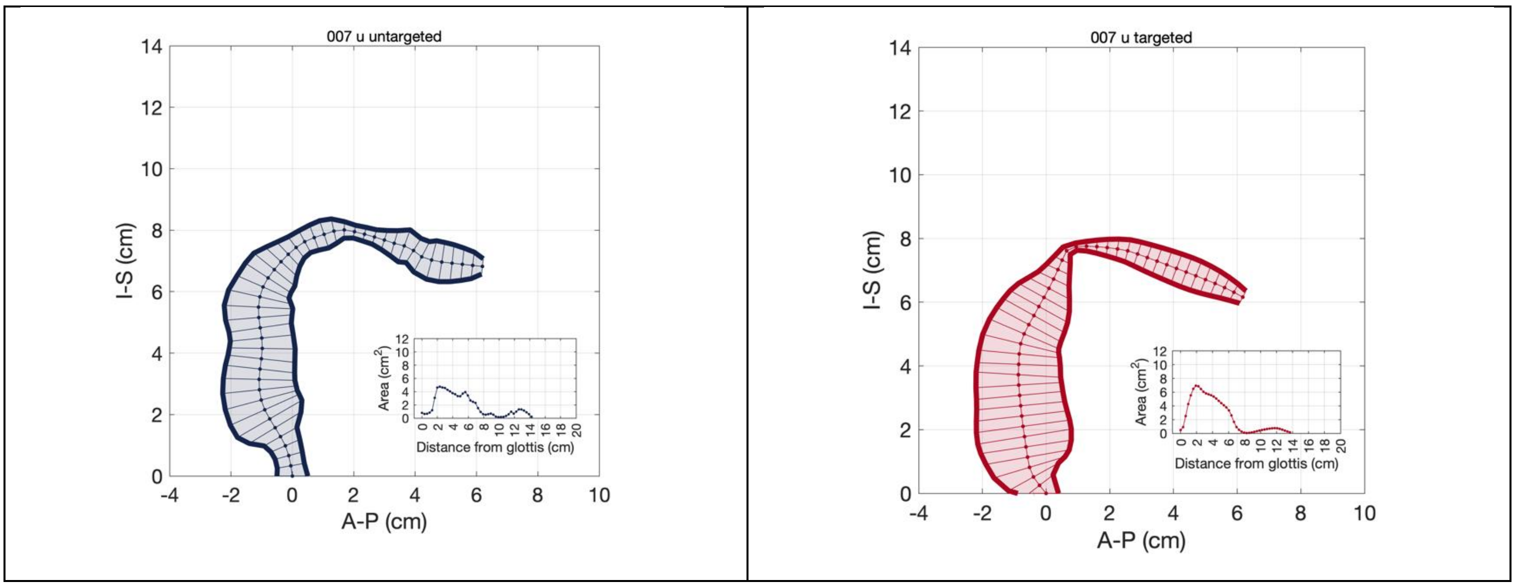


Figure 4.59: Pseudo mid-sagittal plots for Participant 07 during production of an /u/ vowel in the untargeted (left) and targeted conditions (right). Mid-sagittal plots show an interpolation of the oral cavity from approximately 7 cm to 9 cm for the targeted condition.

Area comparison of /u/ (Participant 07)

Targeting the /u/ resulted in a shorter epilarynx with the same area compared to the untargeted production. Targeting led to an increase in the A_{ph} and a decrease A_o. In the targeted production, the participant shortened the epilarynx, likely to increase the A_{ph}. (Figure 4.60).

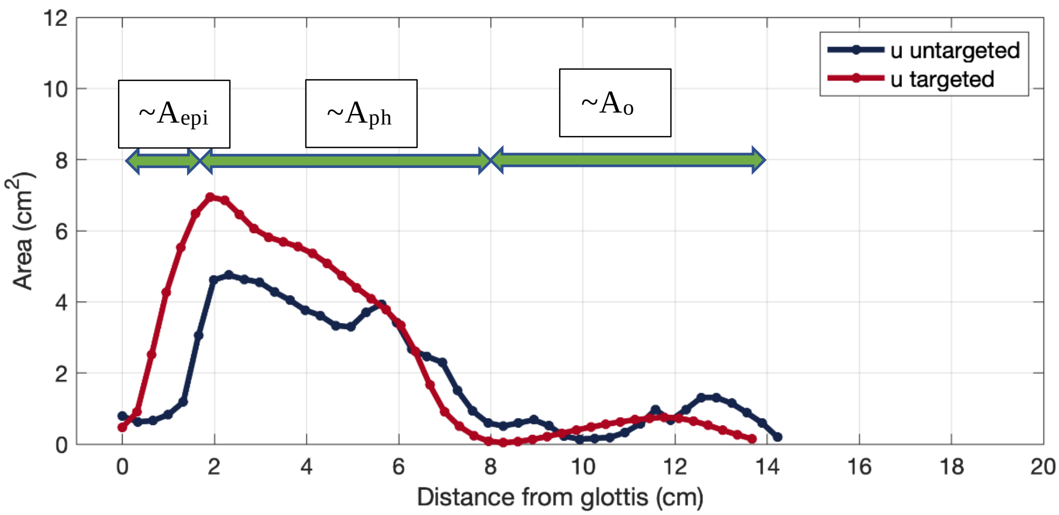


Figure 4.60: Smoothed area function for Participant 07 in the /u/ untargeted and targeted conditions. At approximately 7 cm to 9 cm represents the region with no measurable air in the targeted condition. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_o.

Area comparison of /u/ with and without straw (Participant 07)

The longest and most uniformed epilarynx, largest pharyngeal, and oral cavity area were achieved in the targeted /u/ with the straw condition (Fig. 4.61). Table 17 presents all the raw area values for Figure 4.61.

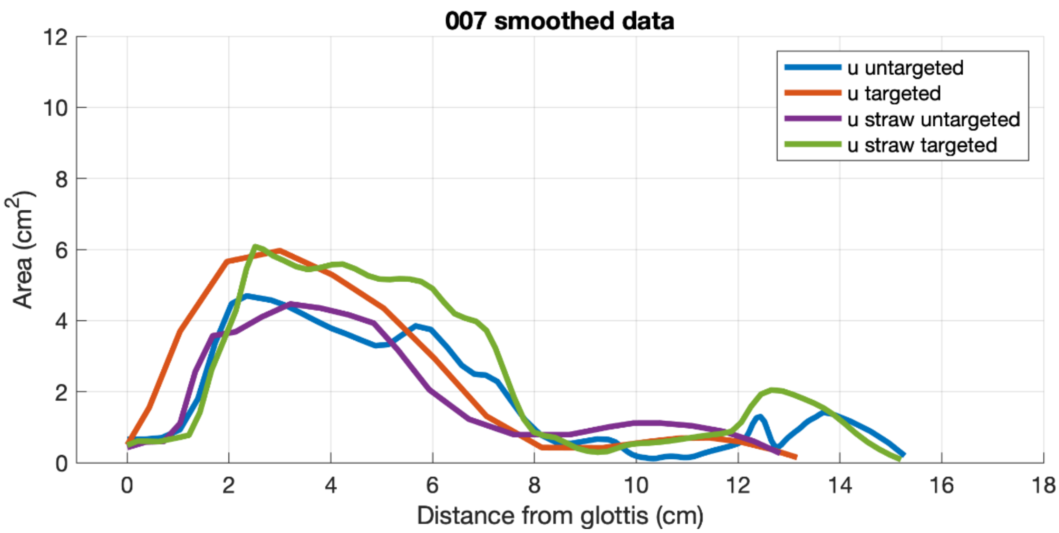


Figure 4.61: Combined area function smoothed participant 07.

Table 17: Participant 07 area function raw data.

Section number (07)	/u/ untargeted	/u/ targeted	/u/ straw untargeted	/u/ straw targeted
1 (glottis)	0.79	0.48	0.44	0.57
2	0.63	0.9	0.67	0.62
3	0.66	2.52	0.6	0.61
4	0.83	4.26	0.5	0.71
5	1.19	5.53	1.61	1.2
6	3.06	6.49	3.98	3.02
7	4.62	6.95	3.72	3.96
8	4.75	6.86	2.87	5.72
9	4.63	6.46	4	5.89
10	4.56	6.05	4.66	5.67
11	4.28	5.82	4.51	5.44
12	4.05	5.68	4.45	5.51
13	3.77	5.55	4.47	5.6
14	3.62	5.36	4.34	5.37
15	3.33	5.07	4.08	5.17
16	3.3	4.74	4.06	5.17
17	3.71	4.4	4.04	5.13
18	3.94	4.09	3.47	4.86
19	3.42	3.77	2.36	4.24
20	2.67	3.34	1.56	4.04
21	2.46	2.61	1.45	3.59
22	2.3	1.66	1.49	2.32
23	1.51	0.91	1.43	1.11
24	0.93	0.51	0.97	0.77
25	0.59	0.24	0.44	0.63
26	0.51	0.08	0.3	0.38
27	0.59	0.04	0.34	0.3
28	0.68	0.07	0.41	0.38
29	0.52	0.13	0.51	0.52
30	0.23	0.21	0.62	0.56
321	0.14	0.31	0.73	0.6
32	0.16	0.4	0.86	0.68
33	0.19	0.48	1.01	0.76
34	0.33	0.56	1.12	0.83
35	0.58	0.63	1.15	1.16
36	0.97	0.68	1.14	1.88
37	0.67	0.73	1.11	2.05
38	0.98	0.75	1.08	1.9
39	1.31	0.73	1.03	1.69
40	1.31	0.65	0.97	1.41
41	1.15	0.53	0.86	1.01
42	0.88	0.39	0.7	0.6
43	0.6	0.27	0.5	0.31
44 (lip termination)	0.2	0.15	0.26	0.09
45 (length)	0.33	0.32	0.31	0.35

Participant 07 /m/

Figure 4.62 shows the 3D reconstructions for Participant 07 during /m/ productions. The panel on the left is the untargeted, or "no vibratory focus" condition, whereas the right panel shows the targeted condition (i.e., "vibratory focus"). The reconstruction in the right panel shows a missing oral cavity. There was no measurable air in the DICOM images for this region.

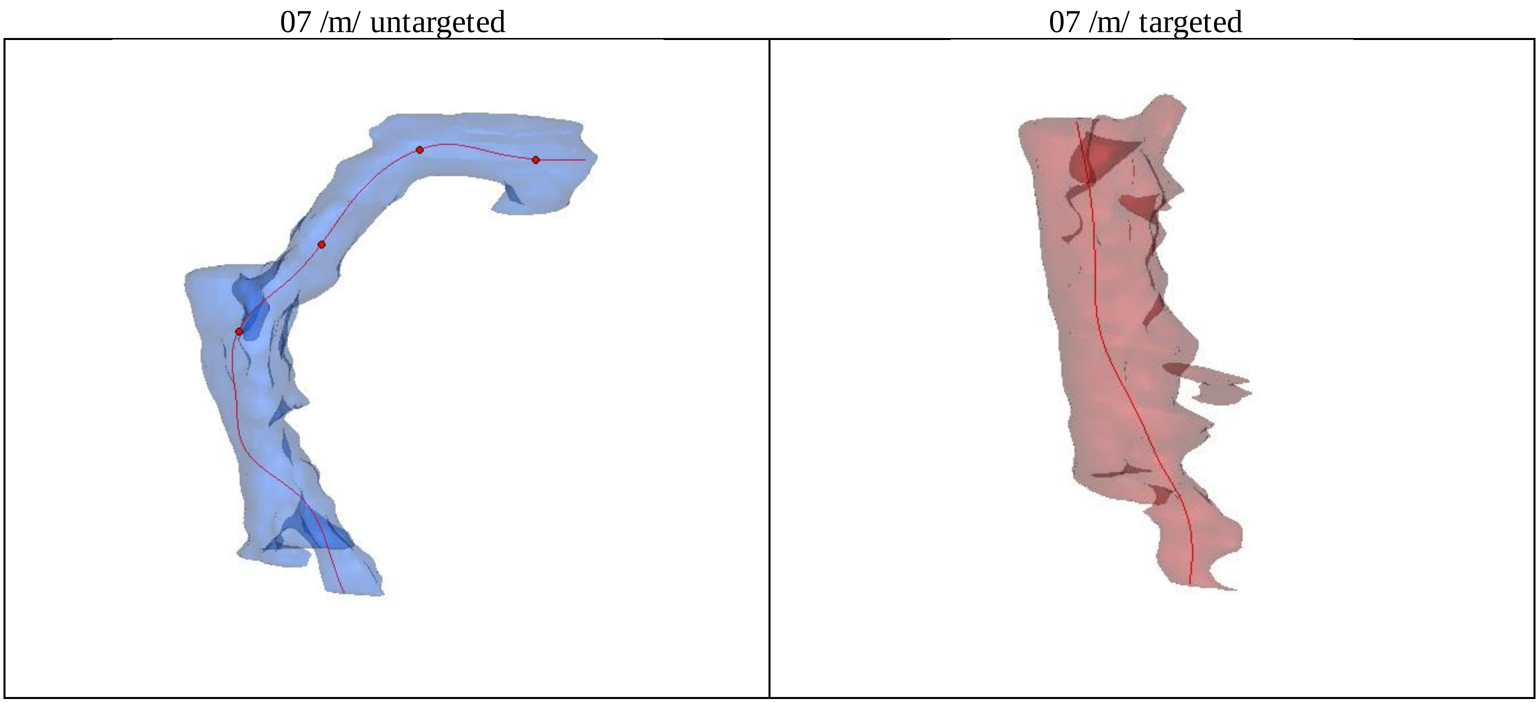


Figure 4.62: 3D model reconstructions for Participant 07 during production of an /m/ in the untargeted (left) and targeted conditions (right). Gap in 3D model for /m/ targeted condition due to no measurable air in the oral cavity.

Figure 4.63 is the pseudo mid-sagittal plots for the /m/ untargeted (left panel) and targeted (right panel) conditions. The targeted condition is represented without an oral cavity due to no measurable air being present in the DICOM images. The untargeted condition does not show a complete lip occlusion, like would be expected for the /m/ sound. The epilaryngeal tube shape is more uniformed in the targeted condition, with a general expansion at several regions of the pharynx.

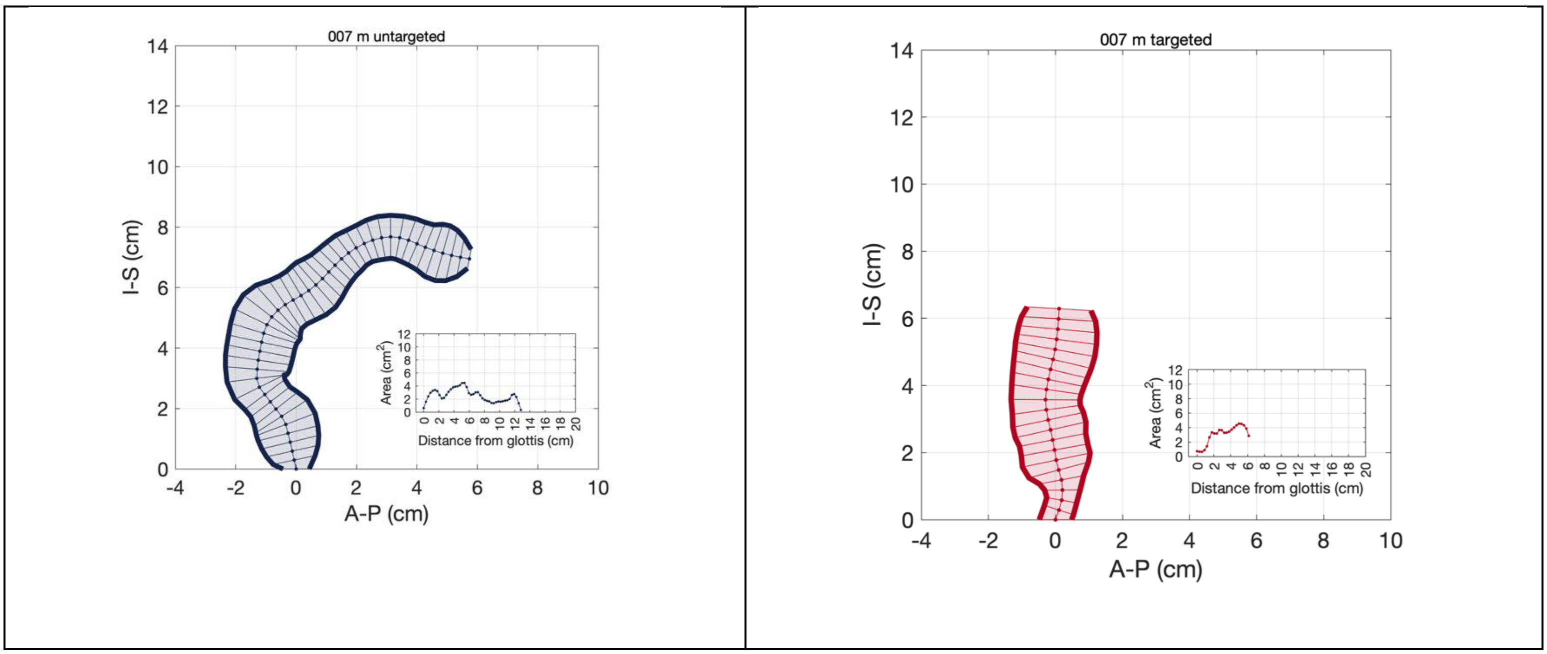


Figure 4.63: Pseudo mid-sagittal plots for Participant 07 during production of an /m/ in the untargeted (left) and targeted conditions (right). /m/ targeted has no representation of the oral cavity due to missing data.

Area comparison for /m/ (Participant 07)

Targeting the /m/ resulted in a smaller area with a more defined geometry of the epilarynx compared to the untargeted production (Fig. 4.64). Targeting resulted in a slightly larger A_{ph}, specifically around 3 cm from the glottis. Due to artifacts from the teeth in the oral cavity, no reliable area could be measured (Fig. 4.65- 4.67). In the targeted condition, the participant created more of an epilaryngeal tube directly above the glottis to expand the A_{ph}. Table 18 presents all the raw area values for Figure 4.64.

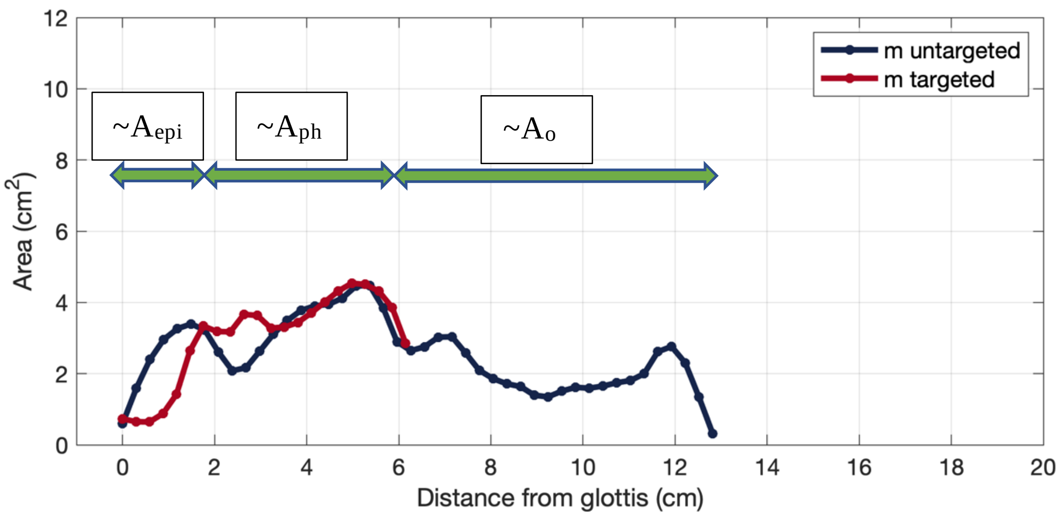


Figure 4.64: Smoothed area function for Participant 06 in the /u/ straw untargeted and targeted conditions. No measurable area in the oral cavity for the targeted condition. The green arrows show approximate regions of A_{epi}, A_{ph}, and A_{o}.

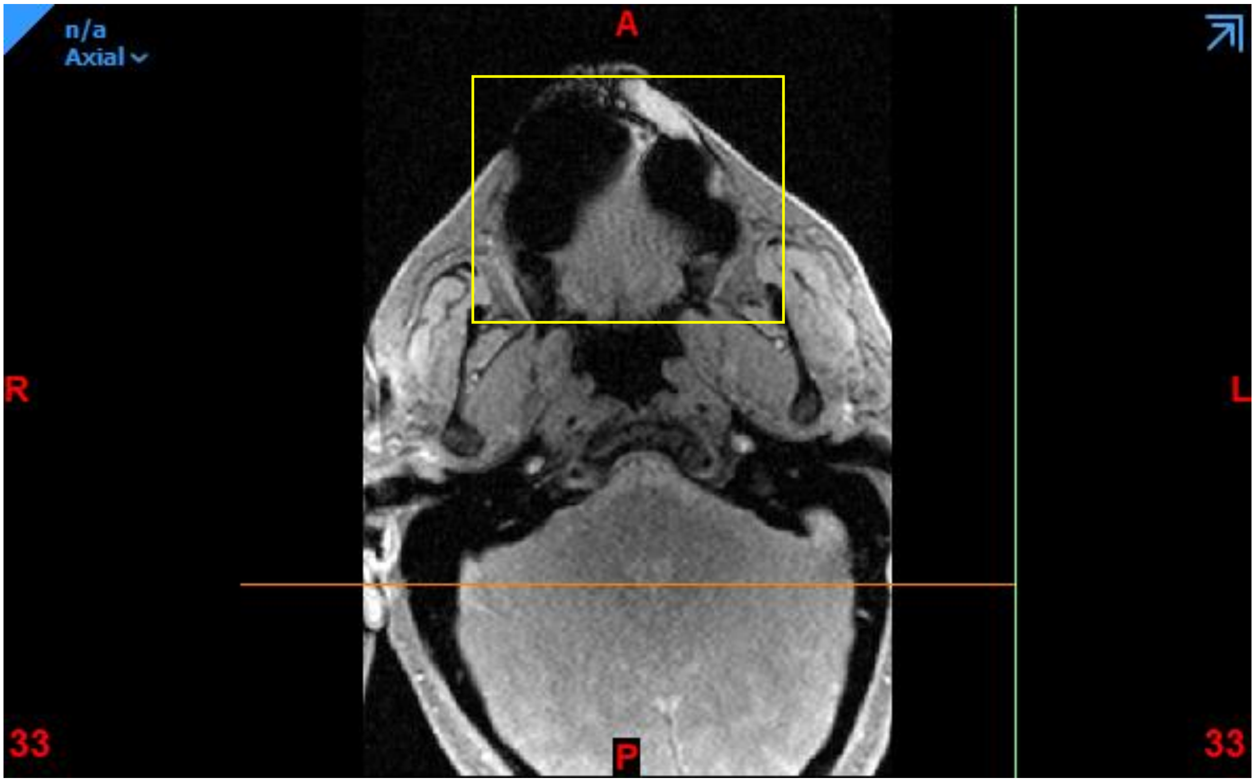


Figure 4. 65: Axial slice one for Participant 07 during /m/ targeted with teeth artifact outlined in yellow box impacting segmentation of oral cavity.

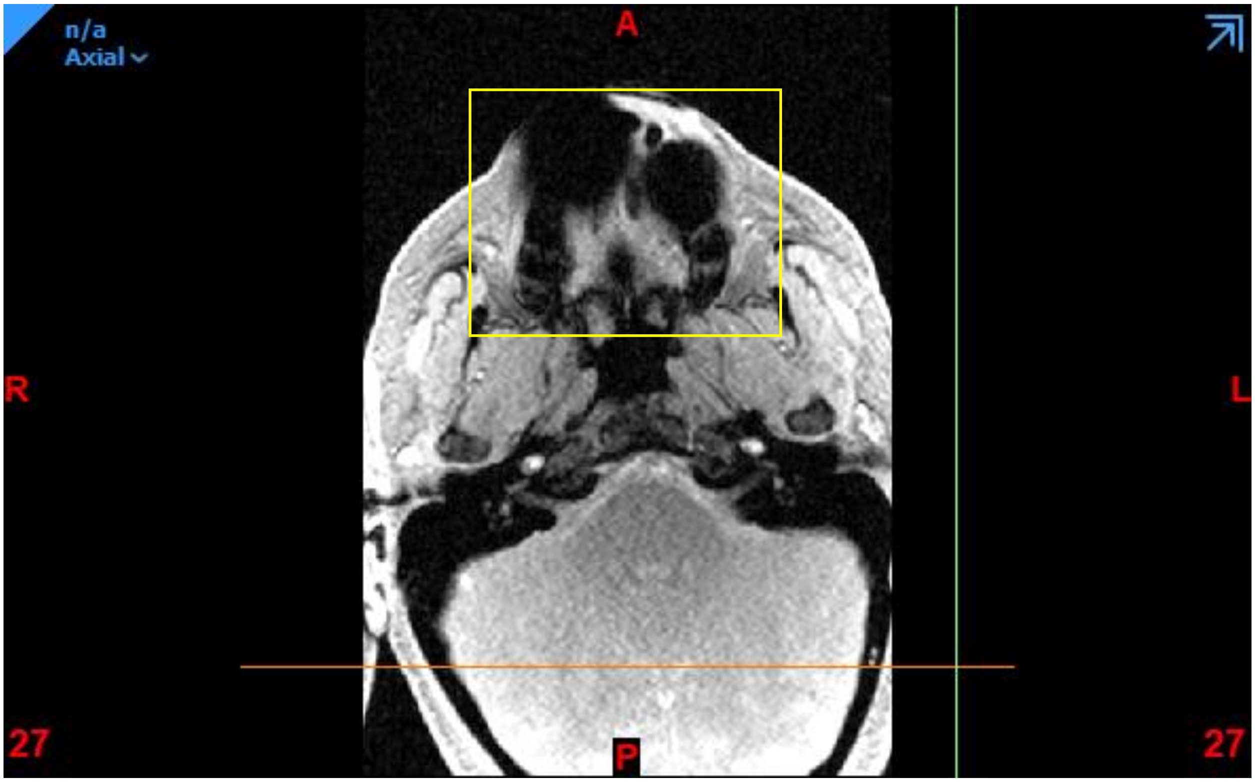


Figure 4. 66: Axial slice two for Participant 07 during /m/ targeted with teeth artifact outlined in yellow box impacting segmentation of oral cavity.

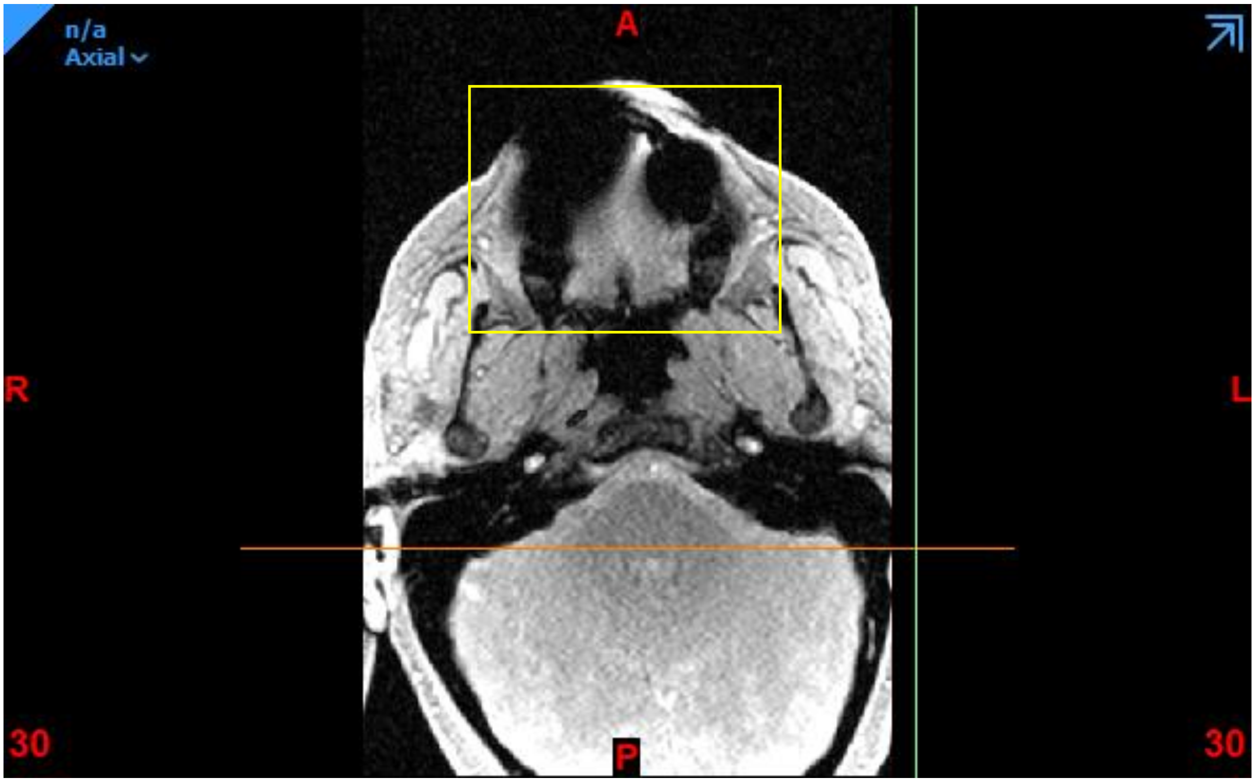


Figure 4. 67: Axial slice three for Participant 07 during /m/ targeted with teeth artifact outline in yellow box impacting segmentation of oral cavity.

Table 18: Participant 07 area function raw data for /m/ in cm^2

Section number	/m/ untargeted	/m/ targeted
1 glottis	0.59	0.73
2	1.59	0.65
3	2.4	0.64
4	2.96	0.88
5	3.26	1.42
6	3.4	2.64
7	3.22	3.34
8	2.6	3.18
9	2.08	3.17
10	2.17	3.66
11	2.64	3.63
12	3.11	3.28
13	3.49	3.31
14	3.78	3.43
15	3.9	3.71
16	3.95	4.01
17	4.11	4.32
18	4.46	4.54
19	4.48	4.51
20	3.84	4.32
21	2.9	3.86
22	2.65	2.85
23	2.75	0
24	3.02	0
25	3.04	0
26	2.58	0
27	2.09	0
28	1.86	0
29	1.72	0
30	1.64	0
31	1.39	0
32	1.35	0
33	1.51	0
34	1.62	0
35	1.59	0
36	1.65	0
37	1.74	0
38	1.81	0
39	2	0
40	2.62	0
41	2.76	0
42	2.3	0
43	1.34	0
44 lip termination	0.32	0
45 section length	0.3	0.29

Participant 07 Area Summary

For participant 07, all targeted conditions resulted in an increased A_{ph} and either the same or a smaller A_{epi}. Two out of three conditions resulted in a more uniform epilaryngeal tube shape.

4.4. Effects of Targeting: Group Trends

Group trends were based on a comparison of individual results from targeting (Table 19). All size changes were relative to the individual themselves (i.e., was the targeted area smaller than, greater than, or the same as the untargeted condition). For our five participants, the A_{ph} was greater for all targeted conditions across all participants, except for participant 03's /m/ condition. This finding was the only consistent trend observed in all participants. The A_{epi} trended towards being larger for targeted conditions in three out of five participants in /u/ with and without straw. Four out of five participants increased the A_{epi} in the /m/ condition. This trend of the increased area was also seen in the oral cavity for three out of five participants in /u/ with straw and /m/. Four out of five participants in the /u/ condition had decreased A_o.

Table 19: Group trends for area of targeted production relative to untargeted production.

Lip Occlusions		**Epilarynx**	**Pharynx**	**Oral**
Partially narrowed	/u/	3/5 decreased	5/5 increased	4/5 decreased
	/u/ with straw	3/5 increased	5/5 increased	3/5 increased
Complete	/m/	4/5 increased	4/5 increased	3/5 increased (2/5 no oral area)

There was no group trend regarding the influence of targeting on epilaryngeal tube shape or length (Table 20). The relative shape and length of the epilarynx remained unchanged across four out of five participants in the /u/ straw condition and three out of five participants in the /u/ without straw. The greatest variability in epilaryngeal tube configuration was seen in the /m/ conditions, where three out of five participants changed their epilaryngeal configuration in response to targeting.

Table 20: Group trends for epilaryngeal length with targeting for all participants

Lip Occlusions		**Participant**				
		002	**003**	**004**	**006**	**007**
Partially narrowed	/u/	Distinct tube	Unchanged	Unchanged	Unchanged	Distinct tube in untargeted
	/u/ with straw	Unchanged	Unchanged	Distinct tube	Unchanged	Unchanged
Complete	/m/	Shorter	Distinct tube in untargeted	Unchanged	Unchanged	Distinct tube

Vocal tract length (Table 21) increased for three out of five participants during /u/ with and without a straw. In four out of five participants, vocal tract length decreased with targeted productions of /m/.

Table 21: Group trends for overall vocal tract length change with targeting, relative to untargeted productions.

Lip Occlusions		**Vocal tract length**
Partially narrowed	/u/	3/5 longer
	/u/ with straw	3/5 longer
Complete	/m/	4/5 shorter

CHAPTER 5: DISCUSSION

This study aimed to understand the role of a vibratory-focused target during two commonly used semi-occluded vocal tract (SOVT) production to determine if targeting influenced cross-sectional areas and length of the vocal tract. Magnetic resonance imaging (MRI) was used to acquire axial images for the vocal tract during a series of SOVT productions from several participants; /u/, /u/ with straw, and /m/ with and without a vibratory focus at the level of the semi-occlusion, partial occlusion at the lips, partial occlusion at the lips with straw, and complete occlusions at the lips, respectively. Cross-sectional areas were measured across the entire vocal tract length, and interpreted through identification of area and length changes in the oral cavity, pharynx, and epilarynx that occurred with attention to vibratory focus. Although the vocal tract configuration has been well studied with regard to vowel and consonant production, as well as to understand changes in VT shape in response to therapeutic interventions, this study is unique in investigating whether cross-sectional area and vocal tract length changed as a function of directed targeting via a vibratory focus during SOVT productions. A commercially available, semi-automated 3D reconstruction software was used and compared to previously published vocal tract configurations based on different methodologies and analysis software (Story, 2005 and 2008). This study utilized a novel approach that improved the spatial resolution of the neck, specifically the glottis, during the acquisition of the MRI images by using a 7-channel laryngeal neck coil.

The discussion will be formatted to address objective two and the subsequent hypotheses.

Objective 2: To determine if cross-sectional area and vocal tract length change when participants are instructed to target a sensation of vibration at the location of three different lip semi-occlusions. Area functions (which included cross-sectional area and length) obtained for the targeted tasks were compared to those measured for the untargeted tasks.

<u>5.1. Hypothesis 1:</u>

Targeted tasks (/m/ and /u/) will decrease cross-sectional area measurements of the epilarynx compared to typical, untargeted productions.

5.1.1. Epilarynx

There is a substantial body of literature about the influence of the epilaryngeal shape and size on vocal fold vibration, vocal tract functioning, and overall voice quality (Rothenberg, 1981; Kent, 1993; Titze and Story, 1997; Story et al., 2001, Samlan et al., 2009, 2011, and 2014). Using computational modeling, Titze (2006) and others have shown that narrowing the epilarynx results in increased supraglottal inertive reactance, increased acoustical power when combined with an open mouth, and increased intraoral pressure when combined with a constricted mouth, increased vocal economy, and increased MFDR. Interestingly in the work of Titze (2006), the size of the epilarynx (i.e., narrowed, medium, or wide) directly results from the oral cavity/mouth size. Our work supported, in certain participants, similar findings. Several modeling studies (Titze and Story, 1997; Titze and Laukkanen, 2007; Moisik et al., 2015) have shown an inverse relationship between the epilarynx and the oral cavity. Titze (2006) found intraglottal pressures were similar during /u/ and /u/ with straw across three different epilaryngeal sizes.

There is literature to support that epilarynx and mouth opening have an inverse relationship. The current study was interested in how the epilarynx was influenced by targeting. We found participants narrowed their epilarynx and oral cavity during targeted /u/. The addition of straw created no consistent relationship between the epilarynx and the oral cavity. Also, the addition of straw caused epilaryngeal expansion more often than constriction. Generally, our findings showed a general trend towards epilaryngeal expansion for conditions with vibratory focus.

The length and shape of the epilarynx have not been well-researched, and many questions remain as to what influences, or creates a distinct shape or size change of the epilarynx. Additionally, what, if any, does this shape or lengthen changes of the epilarynx have in human phonation. In the current study, semi-occluded productions with or without a vibratory focus had little uniform influence over the epilarynx. The majority of participants kept the shape of their epilarynx the same across all three SOVT tasks. Arguably, it could be challenged that all the SOVT tasks selected would have been expected to perform the same in epilaryngeal length and configuration; they were all oral semi-occlusions. An interesting counterpoint is that very few of the trials did participants have a uniform or distinct epilaryngeal tube.

5.1.2. /u/ vowel

In the cases where participants were asked to feel vibrations or target placement of sound at the lips for /u/ vowels, we found a majority of participants (4 out of 5) had a direct relationship between the oral cavity and epilaryngeal area with vibratory focus. That is, if a person increased the oral cavity area during the vibratory focus, they also increased the epilaryngeal area and vice versa. The pharynx remained unchanged by the changes in the A_{epi} and A_o. In all conditions targeting resulted in an expansion of the pharynx. Pharyngeal expansions have been demonstrated to reduce paralaryngeal tension (Laukkanen et al., 2008; Guzman et al., 2013 and 2017), lower the larynx (Guzman et al., 2013 and 2017), and result in a perception of vocal ease (Kapsner-Smith et al, 2015). Our findings also showed that by targeting vibration during /u/

productions majority of participants actually decreased the area in the epilarynx and oral cavities rather than a general expansion like has been proposed (Guzman et al., 2013 and 2017; Laukkanen et al., 2012). It is interesting to note that majority of the literature looking at SOVT productions has used professional voice users or people with extensive knowledge of sound production. The production of a /u/ with a target was the closest approximation of area patterns that were modeled in Titze and Laukkanen (2007) and Titze and Palaparthi (2016) for greatest acoustic output.

5.1.3. /m/

The "resonant voice therapy" literature encourages a targeted "hum" that sets the teeth slightly apart with the lips held lightly together. The clinical training manual for the Lessac Madsen 3Resonant Voice Therapy (LMRVT) recommends supports the use of a relaxed jaw position, tongue neutral in the mouth during nasal productions, and easy/buzzy vibrations (Verdolini, 2000) Although these specific instructions were not given to participants, they were encouraged to feel vibrations in the front of the face and into the nose and teeth. It is important to note that perception of vibration in the nose/face is frequency dependent. Specifically, people feel more buzzy sensations at lower frequencies than at higher ones (Titze and Abbott, 2012). In the current work, participants were asked to phonate at a comfortable pitch and comfortable loudness and hold them steady at these comfortable levels, although not directly measured during the scanner session.

Significant variability in area was observed across all regions of interest. However, the majority of participants expanded the entire vocal tract during targeted productions. Four out of five participants increased the A_{epi} and A_{ph} during targeting, the only consistent trends with targeting /m/. The one participant who did not increase in A_{epi} had the same A_{epi} as for her untargeted production. Two participants increased the A_o, one participant decreased the A_o, and two presented no measurable airspace in the oral cavity. Additionally, the overall length of the vocal tract decreased with targeting, a consequence of the participants that had no measurable area in the oral cavity.

Our findings of targeting /m/ suggest that for a majority of participants, the instruction to feel vibrations was achieved by creating, based on perception, greater nasal resonance. In order to accomplish vibration on the /m/, this cohort placed or directed air up through the velopharynx, bypassing the oral cavity. This was visually observed in the DICOM images, as well. These findings support much of what has been found in the resonant voice literature (Andrade et al., 2014; Titze, 2002a; Detweiler, 1994). The pharyngeal expansion, much like the /u/ targeting, was again a finding consistent in the literature (Leppävuori et al., 2021; Guzman et al., 2013 and 2017; Laukkanen et al., 2012). With full lip occlusion, four participants had more uniform shaping of the epilarynx. This shape change could be the result of more accurate placement and possible synchronization of the vocal tract and glottis during /m/. Clinically /m/ is often used as a first attempt to train people to resonant voicing. It has been suggested (Verdolini, 2000) people subjectively feel more sensation with nasal sounds and can rely on auditory feedback to adjust sound quality. Clinicians additionally have used these sounds to contrast forward and "back focused" resonance. While further exploration of these nasal semi-occlusions is needed, it is reasonable from this work that using /m/ does have the most impact on epilaryngeal shape.

5.2. Hypothesis 2

Artificial lengthening, referred to as "straw," of the vocal tract (i.e., /u/ with straw) will result in the following changes compared to targeted /u/.

1. Increased vocal tract length due to expansion of the oral cavity and lower position of the larynx across all participants.
2. Increased area in the oral cavity and pharynx.
3. Decreased area of the epilarynx.

The cross-sectional area measurements obtained in the current study demonstrated that artificially lengthening the vocal tract with a straw directly influences the oral and epilaryngeal areas by increasing or decreasing the area, depending on how the person modifies the oral cavity shape. The pharynx, however, was consistently narrowed during straw conditions regardless of the changes in the oral or epilaryngeal areas. The effects of semi-occlusions using the /u/ vowel with and without straw are detailed below, specific to their influence on vocal tract shape and configuration. Each semi-occluded task for the /u/ productions will be discussed first as they relate to productions that did not have a target of vibratory focus, followed by those with a target. Our overall goal was to determine whether the addition of a straw to the /u/ production changed the vocal tract shape.

5.2.1. Effects of straw with no vibratory focused (/u/ untargeted to /u/ straw untargeted)

The findings indicated that the cross-sectional areas within the epilaryngeal, pharyngeal, and oral cavity regions were systematically different for /u/ vowel productions with and without straw. Adding the straw always decreased A_{ph}, but changes in A_o and A_{epi} were more complex. Specifically, the straw either increased or decreased the area at the epilarynx and mouth, but always in the same direction; that is, if the A_{epi} decreased, then the A_o decreased, and if the A_{epi}

increased, then the A_o increased. The direct relationship between changes in epilaryngeal and oral cavity areas indicates that the presence of a straw affected the participants' shaping of vocal tract shape in locations distant from the semi-occlusion. Titze (2006) proposed that to maximize acoustical power and vocal efficiency for /u/ vowel production, a narrowed mouth with a narrowed epilarynx is required. Although acoustical power and vocal efficiency were not investigated in the current study, it was found that, in the absence of instruction to feel vibration in the front of the mouth, how a person shapes the area of their oral cavity with the /u/ vowel productions determined the same shape/area in the epilarynx. The pharynx, interestingly, was always smaller with the straw. This result that the straw narrowed the pharyngeal area was in direct opposition to findings of previous MRI studies. Laukkanen et al., 2012 found the midsagittal width (cross distance) of the oral cavity, pharynx, and epilarynx increased during straw phonation and that the ratio of pharyngeal to epilaryngeal area increased during and after straw phonation. Additionally, Laukkanen et al. (2012) found that the straw assisted in establishing the speaker's formant cluster. Guzman et al. (2013 & 2017) suggested that SOVT productions, specifically those using straw or tubes, generally expanded the airway due to the steady pressure present in the vocal tract. In particular, they found that SOVT productions with artificial lengthening increased the area of the trachea, glottis, laryngeal vestibule, pharynx, and oral cavity. The current study did not find a consistent trend of airway expansion in untargeted productions. Note that those studies compared neutral vowels to SOVT with straw; the current study included a much subtler contrast of /u/ to /u/ with a straw.

Titze and Laukkanen (2007) investigated changes in acoustic impedance and the acoustical consequence of airway expansions and altered airway geometries. These authors suggested that specific airway expansions benefit vocal fold vibrations, and others dampened vibration of the sound source. Titze (2020) theorized that to overcome the impedance changes in the airway, constriction of the epilarynx during semi-occluded phonation tasks is required to counter the passive expansions that occur with an SOVT. Titze (2020) concluded that to gain the greatest inertance of the vocal tract during SOVT with a tube, a narrowed and lengthened epilarynx, pharyngeal expansion, and narrowed oral cavity are required. Although these predictions of vocal tract shape modifications were not found in the data collected from the participants in the current study, the suggestion that passive changes occur during SOVT productions was supported by what occurred in the untargeted conditions. Our findings additionally indicate that the passive nature of the pharynx appears to result in narrowing when the straw was added and does not appear to be influenced by the area of the oral cavity or epilarynx.

5.2.2. Effects of the straw with vibratory focus (/u/ targeted to /u/ straw targeted)

Adding a straw to the /u/ production when both were targeted resulted in all participants using different combinations of A_{epi} and A_o. The A_{ph} increased with straw for all participants. For two of the participants, the directions of A_{epi} and A_o change were opposite one another. When A_{epi} increased, A_o decreased, and vice versa. For the other two participants, the A_{epi} and A_o changed in the same direction as one another, like they did when adding the straw in the untargeted condition. The direction of change, though, was the opposite of what occurred in the non-vibratory focus conditions. Therapeutically, targeting with a vibratory focus is designed to allow a person to achieve ease of phonation while achieving perceptual improvements in voice quality. Cueing in therapy often relies on feeling a vibration in certain parts of the mouth or face (i.e., forward versus back sensations), reducing muscular effort or tension in the neck and throat, and modifying their exhale. In the current study, participants were only instructed to feel vibrations on the lips or end of the straw. With this cueing alone, the cross-sectional areas in the oral cavity and epilarynx did not trend in any specific direction across participants. Instead, each person interpreted vibratory cueing differently, creating a unique oral cavity and epilaryngeal shape. In voice thearpy, Speech-Language Pathologists guide patients to make changes following feedback and modification of the directions. Additional directions were likely needed to help participants identify and enhance the feeling of vibration in the straw. More variability occurred within the vocal tract when generalized instruction was given rather than the semi-occluded task alone. Given this variability with targeting, the current work supports the theory that the SOVT posture (i.e., straw) plays only a single part of a complex chain of possible modifications during therapeutic interventions.

5.3. Overall vocal tract length changes

Several factors influence measurements of overall vocal tract length changes. More importantly, many assumptions can be made in interpreting the directionality of vocal tract length changes. Guzman et al. (2007) suggested that during semi-occluded productions with artificial lengthening, the vocal tract was longer because of the lowering of the larynx. Others (Titze., 2006; Titze et al., 2021) have suggested that vocal tract length changes during semi-occluded productions occur in the oral cavity and that lip rounding/shape accounts for the overall length changes. Described in each of the following subsections are the effects of semi-occluded vocal tract productions on the overall length changes of the vocal tract.

5.3.1. /u/ vowel length changes (/u/ untargeted to /u/ targeted)

The vocal tract was the longest for all participants during /u/ vowel productions. Furthermore, the vocal tract was longer in the targeted /u/ productions than in the untargeted /u/ vowels for three out of five participants. For several participants, the lengthening of the vocal tract during these productions was achieved through lip rounding or modifications to the oral aperture. Clinical and computational literature has supported that changes in vocal tract length are contingent on lip/oral shaping. In a human subject's study with repeated productions, Guzman (2013 and 2017) found laryngeal lowering with more lip protrusion during straw semi-occlusions than ???.

5.3.2. /u/ straw length changes (/u/ straw untargeted to /u/ straw targeted)

In all cases where the participant produced an /u/ vowel with a straw at the lips, the vocal tract was shorter without cueing for a vibratory focus. When looking at the impact of targeting with a vibratory focus, the vocal tract was not always longer. More variability existed with vocal tract length changes than in the area for this condition. For most participants, three of five, targeting resulted in a longer vocal tract. This trend suggests that participants could interpret targeting to create an adjustment within the oral cavity (i.e., more significant lip protrusion) or a lowering of the larynx.

5.3.3. Straw influence on vocal tract length (/u/ targeted to /u/ straw targeted)

When comparing participants' targeted productions of /u/ with and without straw, the greatest overall length changes were during the /u/ without straw conditions. This finding is different than the results reported by Guzman et al. (2007), who showed that artificial lengthening via tube/straw produced the longest vocal tract compared to vowel productions without straw. From this cohort, participants adjusted the length of the vocal tract more dramatically, either by lowering the larynx or greater lip protrusion, without a straw. It is possible that this group's interpretation of the instruction to feel more vibration was more strongly felt when a straw was not present. However, it is important to note that the acoustic implications of this finding were not looked at. Furthermore, the additional length of the straw was not considered in the overall area functions for the current study.

5.3.4. /m/ length changes

Of all the semi-occluded tasks in the study, the complete lip occlusion (/m/) resulted in the shortest vocal tract lengths. A shortening of the vocal tract was observed with targeting for 3 of the 5 participants and was unchanged for 2 of the 5 participants. Given that the majority of participants decreased their A_o during targeting, it appears that this coincided with a shortening of the vocal tract. During informal listening and consensus among the authors, participants who did not have measurable area in the oral cavity were perceived to have greater nasal resonance. With cueing to feel vibrations in the nose or teeth, participants tended to allow the tongue to fill the anterior portion of the oral cavity (i.e., touched the hard palate) during the targeted production and decreased the oral cavity's relative airspace.

5.4. The role of the semi-occluded vocal tract

Using a semi-occluded vocal tract during sound production has become widely accepted as a standard of treatment for various voice disorders (Ford, 2021). Recently the scientific underpinnings of these exercises have begun to be investigated. It has been theorized that voicing using a semi-occluded vocal tract improve the overall efficiency of human vocalization by reducing paralaryngeal tension, lowering the larynx, and increasing space within the pharynx/larynx and subsequently reducing vocal effort and load (Guzman et al., 2012). Other sources using computational modeling have indicated that voicing with an SOVT physiologically "unpresses" the vocal folds secondary to the change in back pressure at the point of occlusion (i.e., lips) (Titze, 2006). Computer simulations and even physical model studies have systematically looked at the influence of an SOVT on airflow and pressures (Story &Titze, 1995; Titze & Laukkanen, 2007) by manipulating individual parameters of the vocal tract (i.e., muscle activation, marginal edge thickness of VFs, pressure, flow, symmetry of vibration, etc.). To date, the supporting data provided by computational and physical modeling work has provided the largest and most renowned evidence for how and why SOVT productions positively influence vocal tract functioning.

Acoustic, aerodynamic, perceptual, and quality of life studies have added to the literature supporting the benefits of productions using a semi-occluded vocal tract in both vocally healthy and disordered people. While necessary groundwork has been done on how the acoustic signal can "improve" following an intervention that focuses on using an SOVT, many questions remain. In a recent randomized control study, Fantini et al. (2017) found that the experimental, classically trained group who received SOVT intervention via ventilation mask demonstrated improvements in perturbation and resonance quality measurements. Improvements in acoustic power following SOVT intervention have been found in disordered voice populations (Frisancho et al., 2020; Guzman et al., 2013), but several others have also challenged this metric (Paes et al., 2013; Guzman et al., 2017; Guzman et al., 2018; Meerschman et al., 2018). The conflicting literature, specifically as it pertains to acoustic outcomes, is secondary to differences in methodology and, even more importantly, the type of SOVT intervention being tested. Attempting to classify and distill the mounting body of literature on SOVT treatments and how the acoustic chain has become extremely difficult due to the vast differences between studies.

A study examining clinicians' practice patterns and their use of SOVT interventions found that straw phonation and straw phonation in water were the most commonly used techniques used in therapy (dissertation of Michigan State). Ford (2021) additionally found that clinicians and those using SOVT treatments feel "exceedingly knowledgeable about the underlying physiology of SOVTs" (Ford, 2021). Professionals' resounding confidence in using SOVT productions therapeutically raises some concerns. An overgeneralization for the indications and the use of SOVT productions have dominated the voice rehabilitative and habilitative world. Despite documented changes occurring within the voice production system, almost no establishment or consensus has been reached regarding the directionality of these improvements in humans or disordered groups. A significant finding of this work was the inherent variability in the physiological consequence of SOVT productions.

Imaging studies, specifically CT and MRI, have confirmed that changes in vocal tract shape do occur in response to semi-occlusions (Guzman et al., 2007). These studies have provided useful insights into the relative area changes influenced during different semi-occluded production. Yet, this body of imaging literature has relied solely on the inherent/built-in principals at play with SOVT productions, with patients and clinicians often believing that the technique works without instruction by the SLP. In the current study, we showed that the instruction alters the vocal tract shape compared to the semi-occlusion alone. The literature has also failed to quantify these changes in vivo during active phonation nor demonstrate which region(s) of the vocal tract are most influenced by the semi-occlusion. No studies have systematically linked vocal tract configuration with clinically relevant placement cueing, like vibratory focus. The current study looked exclusively at the changes in the geometric shape of the vocal tract during semi-occluded production with and without directed instruction on vibratory sensation.

CHAPTER 6: CONCLUSIONS

6.1. Conclusion

Different semi-occluded productions influence the cross-sectional area and overall vocal tract length. It has been proposed that SOVT productions inherently create passive expansions of the vocal tract, and little instruction or cueing is needed during these productions to optimize the efficiency or organization of the system. The current study found several differences between SOVT productions with and without instruction. Without instruction on how to produce the sound during SOVT productions, the oral cavity and epilarynx consistently changed in the same direction of their expansion/widening or constriction/narrowing. The pharynx was uniquely independent of these regions and always narrowed during productions without instruction.

Only one distinct trend of targeting/instruction was observed: pharyngeal area increased when participants targeted lip vibration during /u/ vowel productions with and without straw. The epilaryngeal area was not consistently dependent on the oral cavity area or mouth shape. Still, it did tend to have a direct relationship with the oral cavity area without targeting and an inverse relationship with the oral cavity area with targeting. Significant variability existed among participants, with no definitive patterns in how they interpreted targeting. It was clear from the findings that vocal tract shape (i.e., area) was influenced by targeting vibratory focus during SOVT productions. Giving instructions to participants on how to create or target the SOVT production caused differences from what their system did without this instruction. Whether these changes optimized or diminished the functioning of the vocal tract system remains to be answered.

The cross-sectional area changes observed for the different conditions of the study can be summarized as:

Targeted tasks

1. Oral Cavity
   a. /u/ decreased area in four out of five participants.
   b. /m/ increased area in two out of five participants. The more significant change was seen in no measurable area (two participants) or a decreased area in the oral cavity (one participant).
2. Pharynx
   a. /u/ increased area in four out of five participants
   b. /m/ increased area in three out of five participants
3. Epilarynx
   c. /u/ increased the area of the epilarynx in three out of five participants.
   d. /m/ increased the area of the epilarynx in four out of five participants

Artificial lengthening of the vocal tract (i.e., /u/ with straw) resulted in the following changes compared to targeted /u/

1. Decreased vocal tract length due to constriction of the lips
2. Increased area in the oral cavity for three out of five participants
3. Increased area in pharynx for all participants
4. Increased area of the epilarynx for four out of five participants

6.2. Limitations and Future Directions

6.2.1. Limitations

This multiple single-subject, exploratory study described differences in vocal tract area and length across different semi-occluded productions. Several unforeseen limitations arose during the MRI image acquisition. During the scanner acquisition, the custom neck coils were not on during some scans, making the SNR in the acquired slices unusable for measuring area. Additionally, there were cases where no measurable air was present in the DICOM. This is physiologically impossible but did prevent us from being able to make any assumptions about oral or oropharyngeal configurations.

Several assumptions were made mathematically and in interpreting the pseudo-mid-sagittal reconstructions and, subsequently, the smoothed area functions. Consensus was reached at several time points throughout the data analysis portion of this dissertation. Dr. Story and the author visually inspected each semi-automated segmentation and agreed on anatomical boundaries or regions of air that appeared incorrect in the 3D model. The centerline that fit the 3D model in MIS was also inspected visually and mathematically by Dr. Story and the author. Several iterations were required to adequately fit the centerline due to the complex geometry of the vocal tract.

MRI images are obtained in the supine position. It is well known that vocal tract geometry changes positionally; therefore, interpretations from the current study are inherently different than what could be expected in the upright position. However, there has been a robust history in MRI of acquiring vocal tract shape changes, and several authors have generalized findings to what could be expected when upright.

6.2.2. Direction for future research

1. Vocal tract imaging was also obtained for the four corner vowels /a/, /i/, /ɑe/, /u/. Analysis of the remaining corner vowels will be completed (i.e., /a/, /i/, and /ɑe/.
2. Expand the scope of the current study to investigate other non-semi-occluded productions for comparison.
3. Calculate formant frequencies based on frequency response functions (from MRI data) and compare the formant frequencies obtained from the participant's acoustic recording of the scanner productions to determine the accuracy of the automated segmentation.
4. Compare the calculated formant frequencies from the MRI images to both the supine and upright acoustic recordings that were acquired in the initial session.
5. Sensitivity function analysis to determine the acoustical implication of targeting and whether targeting and the subsequent changes to the vocal tract optimize vocal functioning.

Appendix A: Initial Screening Questionnaire

1. Are you currently experiencing problems with your voice?
2. Do you have a history of voice problems?
3. Have you ever had an evaluation or received treatment for a voice disorder?
4. Are you a professional singer?
5. Have you ever received training or lessons for singing?
6. For female participants
    a. Are you pregnant? Or do you plan on becoming pregnant in the next month?
7. Have you ever undergone magnetic resonance imaging (MRI)?
8. The following questions with be specific to MRI eligibility.
    a. Do you have…
        i. pacemaker or cardiac defibrillator?
        ii. Breast tissue expanders?
        iii. Electronic stimulators (DBS, spinal, bladder, tens unit, gone growth other)?
        iv. Acticoat silver wound dressing?
        v. On-body injector?
        vi. Cochlear inner ear implants?

If the answer to any of the above questions is “yes,” participants will not be eligible for the study.

Appendix B: MRI screening questionnaire

HEALTH
UNIVERSITY OF UTAH

Today's Date ________________ MRN________________

Name ____

Date of Birth ______________________________________

IMPORTANT INFORMATION*: For purposes of MRI safety, ALL metallic, magnetic and electronic objects will need to be removed from your person. You will also be asked to change into hospital clothing. A locker will be provided. If you have any questions, please ask a technologist BEFORE entering the MRI scan room.*

☐ Yes ☐ No Pacemaker or Cardiac Defibrillator (ICD) ***(If yes, Stop** and inquire with MRI personnel)*
☐ Yes ☐ No Breast Tissue Expanders ***(If yes, Stop** and inquire with MRI personnel)*
☐ Yes ☐ No Electronic Stimulator (DBS, Spinal, Bladder, Tens Unit, Bone Growth, Other) (***If yes, Stop** and Inquire)*
☐ Yes ☐ No Acticoat Silver Wound Dressing ***(If yes, Stop** and inquire with MRI personnel)*
☐ Yes ☐ No On-body Injector (Neulasta® Onpro®) ***(If yes, Stop** and inquire with MRI personnel)*
☐ Yes ☐ No Cochlear Inner Ear Implant ***(If yes, Stop** and inquire with MRI personnel)*
☐ Yes ☐ No Stapes Inner Ear Implant: If yes, what year was it placed? ________________
☐ Yes ☐ No Implanted Medication Infusion device (Baclofen pump, etc): *Type* ____________
☐ Yes ☐ No Insulin Pump-Glucose Monitor (MiniMed® Dexcom®etc)*(Remove before entering MRI scan room)*
☐ Yes ☐ No Internal electrodes or electronic wires, including abandoned leads
☐ Yes ☐ No Brain Aneurysm Stents, Clips or Coils: *Type* ___________________
☐ Yes ☐ No Aortic or Carotid Artery Clips
☐ Yes ☐ No Heart Valve or Cardiac Stent
☐ Yes ☐ No Vascular Coil, Umbrella Filter, or other Stent
☐ Yes ☐ No Shunt (Intraventricular or Spinal) (programmable or non-programmable?)
☐ Yes ☐ No Port for I.V. Access or I.V. Catheter
☐ Yes ☐ No Prosthetic Device or Joint Replacement (Limb, Eye, Knee etc.): *Type, Location* ____________
☐ Yes ☐ No Metal (circle) Rods, Plates, Screws, Nails, Pins, Clips, Wire Suture, etc: *Location*_________
☐ Yes ☐ No Wire Mesh Implant
☐ Yes ☐ No Eyelid Spring, Wire, Gold Weight
☐ Yes ☐ No Shrapnel (Metal Fragments) or Gunshot Injury: *Location* ________________
☐ Yes ☐ No History as Grinder, Welder, Machinist (Possible metal fragments in eye past or present)
☐ Yes ☐ No IUD, Diaphragm, Penile Implant
☐ Yes ☐ No Breast (circle) Biopsy Markers or Implants (Silicon, Saline, Other)
☐ Yes ☐ No Hair Pins or Hair Piece / Wig *(Remove before entering MRI scan room)*
☐ Yes ☐ No Medication Patch *(Remove before entering MRI scan room)*
☐ Yes ☐ No Hearing Aid *(Remove before entering MRI scan room)*
☐ Yes ☐ No Body Piercings or Jewelry: *Location* _______________ *(Remove before entering MRI scan room)*
☐ Yes ☐ No Clothing: athletic, anti-microbial, anti-odor, heat retention, shapewear (Spanx) *(Must be removed)*
☐ Yes ☐ No Denture or Dental appliance, removable or permanent: *Type* _____________________
☐ Yes ☐ No Tattoos (Cosmetic, Body)
☐ Yes ☐ No Magnetic Eyelashes
☐ Yes ☐ No Bronzing / Tanning Lotions
☐ Yes ☐ No Any Other Implants (Pill Camera, PFO, etc.): *Type* ______________________
☐ Yes ☐ No Are you claustrophobic?
☐ Yes ☐ No Are you pregnant?

Person Completing This Form: ☐Patient ☐Guardian ☐Relative _____________ ☐Other ______________

___________________________________ ________________________________ _____________
Signature Print Name Date

FOR MRI DEPARTMENT USE ONLY

Information Reviewed by: _________________________ ________________________________
Print Name Signature

☐MRI Tech ☐MD ☐Other _________________________________

Okay to scan per Dr. ____

Owner: Nicole Cottam, MRI manager
Revision Date: 10/2020

Page **1** of **1**

Appendix C: Voice Handicap Index

Voice Handicap Index (VHI)

Please circle the word that matches how serious you feel your voice problem is **overall**:

No Problem Mild Problem Moderate Problem Severe Problem

Please circle the word that matches how serious you feel your voice problem is **today**:

No Problem Mild Problem Moderate Problem Severe Problem

Talkativeness: On a scale of 1-10, with 1 being least talkative and 10 being most talkative, rate yourself: ________**/10**

Because of your voice problem, has your talkativeness changed? If so, what is it now? ________**/10**

Loudness: On a scale of 1-10, with 1 being softest and 10 being loudest, rate yourself: ________**/10**

Because of your voice problem, has your loudness changed? If so, what is it now? ________**/10**

Percent of normal: 0 to 100%, with 0% being no voice and 100% being your normal voice, rate your voice: ________

Instructions: These are statements that many people have used to describe their voices and the effects of their voices on their

Check the response that indicates how frequently you have the same experience.

		Never	Almost Never	Sometimes	Almost Always	Always
F1	My voice makes it difficult for people to hear me					
P2	I run out of air when I talk					
F3	People have difficulty understanding me in a noisy room					
P4	The sound of my voice varies throughout the day					
F5	My family has difficulty hearing me when I call them throughout					
F6	I use the phone less often than I would like					
E7	I'm tense when talking with others because of my voice					
F8	I tend to avoid groups of people because of my voice					
E9	People seem irritated with my voice					
P10	People ask, "What's wrong with your voice?"					
F11	I speak with friends, neighbors, or relatives less often because of					
F12	People ask me to repeat myself when speaking face-to-face					
P13	My voice sounds creaky and dry					
P14	I feel as though I have to strain to produce voice					
E15	I find other people don't understand my voice problem					
F16	My voice difficulties restrict my personal and social life					
P17	The clarity of my voice is unpredictable					
P18	I try to change my voice to sound different					
F19	I feel left out of conversations because of my voice					
P20	I use a great deal of effort to speak					
P21	My voice is worse in the evening					
F22	My voice problem causes me to lose income					
E23	My voice problem upsets me					
E24	I am less out-going because of my voice problem					
E25	My voice makes me feel handicapped					
P26	My voice "gives out" on me in the middle of speaking					
E27	I feel annoyed when people ask me to repeat					
E28	I feel embarrassed when people ask me to repeat					
E29	My voice makes me feel incompetent					
E30	I'm ashamed of my voice problem					

For Clinician Use Only:

P Scale ________F Scale__________ E Scale__________ Total__________

Jacobson, Johnson, Grywalski, Silbergleit, Jacobson, Benninger, & Newman (1997). The Voice Handicap Index (VHI). *American Journal of Speech Language Pathology, 6,* 66-70.

Appendix D: Laryngoscopy Protocol

VELOPHARYNGEAL

1. Sustain /i/
2. Sustain /s/ or "sh"
3. Repeat /pi-pi-pi-pi-pi/
4. Repeat /mi-mi-mi-mi-mi/
5. Repeat "hamper – hamper…"
6. Say, "Peter will keep at the peak." "She speaks pleasingly."

PHARYNGEAL/LARYNGEAL

1. Sustained /i/ at comfortable pitch and loudness
2. 3 sniffs (while viewing the posterior cricoarytenoids)
3. Sniff-/i/ (repeated 3x)
4. Sniff-/i/ as quickly as possible without compromising task integrity
5. Sustained /i/ at comfortable pitch and loudness
6. Say, /i/-/i/-/i/-/i/-/i/ on one breath
7. Glide from low to high pitch on /i/
8. Glide from high to low pitch on /i/
9. Repeat the following sentences:
    a. The blue spot is on the key again.
    b. How hard did he hit him?
    c. We were away a year ago.
    d. We eat eggs every Easter.
    e. My mama makes lemon muffins.
    f. Peter will keep at the peak
10. Quiet respiration for 4 cycles of breathing
11. Close up view of the vocal folds during quiet respiration
12. Panting

STROBOSCOPY EXAMINATION OF LARYNX

1. Habitual pitch
2. High pitch
3. Low pitch
4. Loud phonation
5. Soft phonation
6. Slow glissando up
7. Slow glissando down

Appendix E: The Rainbow Passage

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

Appendix F: CAPE-V

Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)

Name:_____________________________ **Date:**___________

The following parameters of voice quality will be rated upon completion of the following tasks:
1. Sustained vowels, /a/ and /i/ for 3-5 seconds duration each.
2. Sentence production:
   a. The blue spot is on the key again.
   b. How hard did he hit him?
   c. We were away a year ago.
   d. We eat eggs every Easter.
   e. My mama makes lemon muffins.
   f. Peter will keep at the peak.
3. Spontaneous speech in response to: "Tell me about your voice problem." or "Tell me how your voice is functioning."

Legend: C = Consistent I = Intermittent
MI = Mildly Deviant
MO =Moderately Deviant
SE = Severely Deviant

				SCORE
Overall Severity	______ MI MO SE	C	I	___/100
Roughness	______ MI MO SE	C	I	___/100
Breathiness	______ MI MO SE	C	I	___/100
Strain	______ MI MO SE	C	I	___/100
Pitch	(Indicate the nature of the abnormality): ______________ ______ MI MO SE	C	I	___/100
Loudness	(Indicate the nature of the abnormality): ______________ ______ MI MO SE	C	I	___/100
__________	______ MI MO SE	C	I	___/100
__________	______ MI MO SE	C	I	___/100

COMMENTS ABOUT RESONANCE: NORMAL OTHER (Provide description):__________________

____

ADDITIONAL FEATURES (for example, diplophonia, fry, falsetto, asthenia, aphonia, pitch instability, tremor, wet/gurgly, or other relevant terms):

Clinician:____________________

Appendix G: Training Protocol (session 1)

We will practice a series of voicing tasks routinely used in voice therapy. All of the sounds that we practice today will be different from those you will complete in the MRI scanner. However, all the tasks that you will be asked to complete today and at the time of the MRI scanner are part of a group of sounds that "semi-occluded" or partially narrow your vocal tract. You can think of the vocal tract as the area from your lips to the back of your throat down to your airway. You will be asked to hold out each of these sounds for as long as you possibly can at a comfortable pitch and loudness. Comfortable pitch and loudness are the equivalent of you talking to a friend in a room with little background noise. For these tasks, I will be asking you to use an abdominal or belly breath, relax your shoulders, and remain as still as possible without moving any of your body. For the actual scanner session, you will repeat each task a total of six times. We will practice what that looks like today.

Let's start with holding out "ah" for as long as possible. This is just like the sound you made when we looked at your throat with the camera. It will sound like this (CLINICIAN TO PROVIDE MODEL). Let's have you try now. That was great! Now, I want you to try something for me. I want you to try to produce that sound with a creaky voice, like this (MODEL GLOTTAL FRY). Let's have you try it. Awesome! That sound is made in the back of your throat, sometimes called glottal fry. We can change the "location" of the sound by shifting where we place the air or sound. For example, I could make the "ah" in the front of my face like this (MODEL RESONANT VOICE). Let's have you try that. Great! Could you feel a difference between the sound that was far back in your throat and the other that was very forward in the front of the mouth? That concept of shifting the sound is called resonant voicing. This concept of a resonant voice is what we will call your "targeted" task. During the scanning session, I will ask you to produce some sounds with no target, just how you might say them in your everyday life, and produce some with a resonant voice. Let's practice a few more sounds!

Task /z/ untargeted: Let's first have you hold out a "z" sound normally. Great!
Task /z/ targeted: Now, let's have you focus on the "z" sound becoming buzzier and feeling vibration behind your teeth and on your tongue. Can you feel that!?

<u>Cues</u>

- Good, can you point to where you felt that sound coming from? If you felt that sound mostly coming through your throat, I want you to try again and see if you can't get the sound more on your lips, like this (then model).
- Good, can you relax your shoulders and try and get your breath more from the belly. It might look like this _____ (model abdominal breath)

Task /v/ untargeted:
Task /v/ targeted:

Although these are not the exact sounds you will be asked to produce during the MRI scanner session, what you will be asked to do will be the same. You will be asked to hold out some sounds normally and some with a target of resonance. Also, we will not practice this today, but one of the tasks you will do will involve using a straw. I will quickly show you what will be asked of you with the straw, but we will not be practicing this today. (MODEL STRAW). Do you have any questions or things I can answer for you? Do you have a good idea of what will be expected of you during the MRI scanner session? You will have plenty of time on the day of your MRI scan to ask me any questions you might have. You will also have a microphone during the entire scanning session so that you will be able to talk to the radiology technician and me. Please keep in mind that we are always here if you are uncomfortable, concerned, or need further explanation on something

Appendix H: Testing Protocol (Session 2)

<u>Task 1: /a/</u>
Take a comfortable breath and say /a/ for as long as possible.

<u>Task 2: /i/</u>
Take a comfortable breath and say /i/ for as long as possible.

<u>Task 3: /ɑe/</u>
Take a comfortable breath and say /ɑ/ for as long as you can. This sound will be similar to saying "park" in a Boston accent or a sigh saying "ah."

<u>Task 4: /u/</u>
Take a comfortable breath and say /u/ for as long as you can.

<u>Task 4: /u/ targeted (without straw)</u>
I want you to take a comfortable breath in and say "ooo" and hold this out for as long as possible without forcing the voice and feeling a buzz on your lips.

<u>Task 5: /u/ with straw</u>
Take that same comfortable breath in and say "ooo" into the straw. All of your air should come through the straw. Make sure no air is escaping out of your nose.

<u>Cues</u>

- Good, did you feel the air at the end of your straw?
- Awesome, did you feel any vibrations or the sound at the straw?

<u>Task 6: /m/ untargeted</u>
I want you to hold out a humming sound for as long as possible.

<u>Task 7: /m/ targeted</u>
I want you to hold out that same humming sound now, but think about making the sound buzzy in mouth.

Appendix I: MRI Sequencing Protocol 1

SIEMENS MAGNETOM Prisma_fit

\\USER\CAMT-Research\532_MRE532_Vocal Cords\Use this one Stark_MRI_002\3mm ax 10slices P H R-L oo straw untargated

TA: 0:23 PM: ISO Voxel size: 0.4×0.4×3.0 mmPAT: 2 Rel. SNR: 1.00 : fl

Properties

Prio recon	Off
Load images to viewer	On
Inline movie	Off
Auto store images	On
Load images to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Auto close inline display	Off
Start measurement without further preparation	Off
Wait for user to start	On
Start measurements	Single measurement

Routine

Slice group	1
Slices	10
Dist. factor	0 %
Position	R5.5 A49.7 H89.3 mm
Orientation	T > C15.1
Phase enc. dir.	R >> L
AutoAlign	---
Phase oversampling	0 %
FoV read	220 mm
FoV phase	68.8 %
Slice thickness	3.0 mm
TR	7.7 ms
TE	3.04 ms
Averages	3
Concatenations	10
Filter	Distortion Corr.(2D), Elliptical filter
Coil elements	BC

Contrast - Common

TR	7.7 ms
TE	3.04 ms
TD	0 ms
MTC	Off
Magn. preparation	None
Flip angle	6 deg
Fat suppr.	None
Water suppr.	None
SWI	Off

Contrast - Dynamic

Averages	3
Averaging mode	Short term
Reconstruction	Magnitude
Measurements	1
Multiple series	Each measurement

Resolution - Common

FoV read	220 mm
FoV phase	68.8 %
Slice thickness	3.0 mm
Base resolution	256
Phase resolution	90 %
Phase partial Fourier	Off
Interpolation	On

Resolution - iPAT

PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	24
Reference scan mode	Integrated

Resolution - Filter Image

Image Filter	Off
Distortion Corr.	On
Mode	2D
Unfiltered images	Off
Prescan Normalize	Off
Normalize	Off
B1 filter	Off

Resolution - Filter Rawdata

Raw filter	Off
Elliptical filter	On

Geometry - Common

Slice group	1
Slices	10
Dist. factor	0 %
Position	R5.5 A49.7 H89.3 mm
Orientation	T > C15.1
Phase enc. dir.	R >> L
FoV read	220 mm
FoV phase	68.8 %
Slice thickness	3.0 mm
TR	7.7 ms
Multi-slice mode	Sequential
Series	Interleaved
Concatenations	10

Geometry - AutoAlign

Slice group	1
Position	R5.5 A49.7 H89.3 mm
Orientation	T > C15.1
Phase enc. dir.	R >> L
AutoAlign	---
Initial Position	R5.5 A49.7 H89.3
R	5.5 mm
A	49.7 mm
H	89.3 mm
Initial Rotation	90.00 deg
Initial Orientation	T > C
T > C	15.1
> S	0.0

Geometry - Saturation

Saturation mode	Standard
Fat suppr.	None
Water suppr.	None
Special sat.	None

System - Miscellaneous

Positioning mode	ISO
Table position	H
Table position	89 mm
MSMA	S - C - T

Appendix J: MRI Sequencing Protocol 2

SIEMENS MAGNETOM Prisma_fit

System - Miscellaneous

Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Adaptive Combine
Save uncombined	Off
Matrix Optimization	Off
AutoAlign	---
Coil Select Mode	Default

System - Adjustments

B0 Shim mode	Tune up
B1 Shim mode	TrueForm
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Dominant Fat	Off
Assume Silicone	Off
Adjustment Tolerance	Auto

System - Adjust Volume

Position	Isocenter
Orientation	Transversal
Rotation	0.00 deg
A >> P	263 mm
R >> L	350 mm
F >> H	350 mm
Reset	Off

System - pTx Volumes

B1 Shim mode	TrueForm
Excitation	Slice-sel.

System - Tx/Rx

Frequency 1H	123.235043 MHz
Correction factor	1
Gain	High
Img. Scale Cor.	1.000
Reset	Off
? Ref. amplitude 1H	0.000 V

Physio - Signal1

1st Signal/Mode	None
TR	7.7 ms
Concatenations	10
Segments	1

Physio - Cardiac

Tagging	None
Magn. preparation	None
Fat suppr.	None
Dark blood	Off
FoV read	220 mm
FoV phase	68.8 %
Phase resolution	90 %

Physio - PACE

Resp. control	Off
Concatenations	10

Inline - Common

Subtract	Off
Measurements	1
StdDev	Off
Liver registration	Off

Inline - Common

Save original images	On

Inline - MIP

MIP-Sag	Off
MIP-Cor	Off
MIP-Tra	Off
MIP-Time	Off
Save original images	On

Inline - Soft Tissue

Wash - In	Off
Wash - Out	Off
TTP	Off
PEI	Off
MIP - time	Off
Measurements	1

Inline - Composing

Distortion Corr.	On
Mode	2D
Unfiltered images	Off

Sequence - Part 1

Introduction	On
Dimension	2D
Phase stabilisation	Off
Asymmetric echo	Allowed
Contrasts	1
Flow comp.	No
Multi-slice mode	Sequential
Bandwidth	320 Hz/Px

Sequence - Part 2

Segments	1
Acoustic noise reduction	None
RF pulse type	Normal
Gradient mode	Normal
Excitation	Slice-sel.
RF spoiling	On

Sequence - Assistant

Mode	Off
Allowed delay	0 s

REFERENCES

Alipour, F., Jaiswal, S., & Finnegan, E. (2007). Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models. Annals of Otology, Rhinology & Laryngology, 116(2), 135-144.

Andrade, P. A., Wood, G., Ratcliffe, P., Epstein, R., Pijper, A., & Svec, J. G. (2014). Electroglottographic study of seven semi-occluded exercises: LaxVox, straw, lip-trill, tongue-trill, humming, hand-over-mouth, and tongue-trill combined with hand-over-mouth. Journal of Voice, 28(5), 589-595.

Andrade, P. A., Wistbacka, G., Larsson, H., Södersten, M., Hammarberg, B., Simberg, S., ... & Granqvist, S. (2016). The flow and pressure relationships in different tubes commonly used for semi-occluded vocal tract exercises. Journal of Voice, 30(1), 36-41.

Baer, T., Gore, J. C., Gracco, L. C., & Nye, P. W. (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *The Journal of the Acoustical Society of America*, *90*(2), 799-828.

Beck, M. J., Parker, D. L., & Hadley, J. R. (2020). Capacitive versus Overlap Decoupling of Adjacent Radio Frequency Phased Array Coil Elements: An Imaging Robustness Comparison When Sample Load Varies for 3 Tesla MRI. *Concepts in Magnetic Resonance Part B, Magnetic Resonance Engineering*.

Bele, I. V. (2005). Artificially lengthened and constricted vocal tract in vocal training methods. Logopedics Phoniatrics Vocology, 30(1), 34-40.

Bergan, C. C., Titze, I. R., & Story, B. (2004). The perception of two vocal qualities in a synthesized vocal utterance: ring and pressed voice. *Journal of Voice*, *18*(3), 305-317.

Bickley, C. A., & Stevens, K. N. (1986). Effects of a vocal-tract constriction on the glottal source: Experimental and modeling studies. Journal of Phonetics, 14(3-4), 373-382.

Bonette, M. C., Ribeiro, V. V., Xavier-Fadel, C. B., da Conceição Costa, C., & Dassie-Leite, A. P. (2020). Immediate effect of semioccluded vocal tract exercises using resonance tube phonation in water on women without vocal complaints. Journal of Voice, 34(6), 962-e19.

Calvache, C., Guzman, M., Bobadilla, M., & Bortnem, C. (2020). Variation on vocal economy after different semioccluded vocal tract exercises in subjects with normal voice and dysphonia. Journal of Voice, 34(4), 582-589.

Clément, P., Hans, S., Hartl, D. M., Maeda, S., Vaissière, J., & Brasnu, D. (2007). Vocal tract area function for vowels using three-dimensional magnetic resonance imaging. A preliminary study. Journal of Voice, 21(5), 522-530.

Conroy, E. R., Hennick, T. M., Awan, S. N., Hoffman, M. R., Smith, B. L., & Jiang, J. J. (2014). Effect of variations to a simulated system of straw phonation therapy on aerodynamic parameters using excised canine larynges. Journal of Voice, 28(1), 1-6.

Costa, C. B., Costa, L. H. C., Oliveira, G., & Behlau, M. (2011). Immediate effects of the phonation into a straw exercise. Brazilian Journal of Otorhinolaryngology, 77(4), 461-465.

Croake, D. J., Andreatta, R. D., & Stemple, J. C. (2017). Immediate effects of the vocal function exercises semi-occluded mouth posture on glottal airflow parameters: a preliminary study. Journal of Voice, 31(2), 245-e9.

Dargin, T. C., & Searl, J. (2015). Semi-occluded vocal tract exercises: aerodynamic and electroglottographic measurements in singers. Journal of Voice, 29(2), 155-164.

Dargin, T. C., DeLaunay, A., & Searl, J. (2016). Semioccluded vocal tract exercises: changes in laryngeal and pharyngeal activity during stroboscopy. Journal of Voice, 30(3), 377-e1.

Detweiler, R. F. (1994). An investigation of the laryngeal system as the resonance source of the singer's formant. Journal of Voice, 8(4), 303-313.

Döllinger, M., Berry, D. A., & Montequin, D. W. (2006). The influence of epilarynx area on vocal fold dynamics. Otolaryngology—Head and Neck Surgery, 135(5), 724-729.

Echternach, M., Traser, L., Markl, M., & Richter, B. (2011). Vocal tract configurations in male alto register functions. Journal of Voice, 25(6), 670-677.

Eichhorn, J. T., Kent, R. D., Austin, D., & Vorperian, H. K. (2018). Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice*, *32*(5), 644-e1.

Engel, E. (1927). Stimmbildungslehre [Voice pedagogy]. Dresden, Germany: Weise.

Fant, G. (1986). Glottal flow: models and interaction. Journal of Phonetics, 14(3-4), 393-399.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106*(3), 1511-1522.

Frisancho, K., Salfate, L., Lizana, K., Guzman, M., Leiva, F., & Quezada, C. (2020). Immediate effects of the semi-occluded ventilation mask on subjects diagnosed with functional dysphonia and subjects with normal voices. Journal of Voice, 34(3), 398-409.

Guzman, M., Laukkanen, A. M., Krupa, P., Horáček, J., Švec, J. G., & Geneid, A. (2013). Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. Journal of Voice, 27(4), 523-e19.

Guzman, M., Castro, C., Testart, A., Muñoz, D., & Gerhard, J. (2013). Laryngeal and pharyngeal activity during semioccluded vocal tract postures in subjects diagnosed with hyperfunctional dysphonia. Journal of Voice, 27(6), 709-716.

Guzman, M., Calvache, C., Romero, L., Muñoz, D., Olavarria, C., Madrid, S., ... & Bortnem, C. (2015). Do different semi-occluded voice exercises affect vocal fold adduction differently in subjects diagnosed with hyperfunctional dysphonia?. Folia phoniatrica et logopaedica, 67(2), 68-75.

Guzmán, M., Castro, C., Madrid, S., Olavarria, C., Leiva, M., Muñoz, D., ... & Laukkanen, A. M. (2016). Air pressure and contact quotient measures during different semioccluded postures in subjects with different voice conditions. Journal of voice, 30(6), 759-e1.

Guzman, M., Jara, R., Olavarria, C., Caceres, P., Escuti, G., Medina, F., ... & Laukkanen, A. M. (2017). Efficacy of water resistance therapy in subjects diagnosed with behavioral dysphonia: a randomized controlled trial. Journal of voice, 31(3), 385-e1.

Guzman, M., Miranda, G., Olavarria, C., Madrid, S., Muñoz, D., Leiva, M., ... & Bortnem, C. (2017). Computerized tomography measures during and after artificial lengthening of the vocal tract in subjects with voice disorders. Journal of voice, 31(1), 124-e1.

Hadley, J. R., Roberts, J. A., Goodrich, K. C., Buswell, H. R., & Parker, D. L. (2005). Relative RF coil performance in carotid imaging. *Magnetic resonance imaging*, *23*(5), 629-639.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. Folia phoniatrica et logopaedica, 26(2), 89-94

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI) development and validation. American journal of speech-language pathology, 6(3), 66-70.

Kapsner-Smith, M. R., Hunter, E. J., Kirkham, K., Cox, K., & Titze, I. R. (2015). A randomized controlled trial of two semi-occluded vocal tract voice therapy protocols. Journal of Speech, Language, and Hearing Research, 58(3), 535-549.

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol.

Kent, R. D. (1993). Vocal tract acoustics. Journal of Voice, 7(2), 97-117.

Kreiman, J., Gerratt, B. R., Garellek, M., Samlan, R., & Zhang, Z. (2014). Toward a unified theory of voice production and perception. Loquens, 1(1), e009.

Kreiman, J., Lee, Y., Garellek, M., Samlan, R., & Gerratt, B. R. (2021). Validating a psychoacoustic model of voice quality. *The Journal of the Acoustical Society of America*, *149*(1), 457-465.

Laukkanen, A. M., Titze, I. R., Hoffman, H., & Finnegan, E. (2008). Effects of a semioccluded vocal tract on laryngeal muscle activity and glottal adduction in a single female subject. Folia Phoniatrica et Logopaedica, 60(6), 298-311.

Laukkanen, A. M., Horáček, J., Krupa, P., & Švec, J. G. (2012). The effect of phonation into a straw on the vocal tract adjustments and formant frequencies. A preliminary MRI study on a single subject completed with acoustic results. Biomedical Signal Processing and Control, 7(1), 50-57.

Leppävuori, M., Lammentausta, E., Peuna, A., Bode, M. K., Jokelainen, J., Ojala, J., & Nieminen, M. T. (2021). Characterizing vocal tract dimensions in the vocal modes using magnetic resonance imaging. Journal of Voice, 35(5), 804-e27.

Lim, Y., Zhu, Y., Lingala, S. G., Byrd, D., Narayanan, S., & Nayak, K. S. (2019). 3D dynamic MRI of the vocal tract during natural speech. Magnetic resonance in medicine, 81(3), 1511-1520.

Lingala, S. G., Zhu, Y., Kim, Y. C., Toutios, A., Narayanan, S., & Nayak, K. S. (2017). A fast and flexible MRI system for the study of dynamic vocal tract shaping. *Magnetic resonance in medicine*, *77*(1), 112-125.

Mainka, A., Poznyakovskiy, A., Platzek, I., Fleischer, M., Sundberg, J., & Mürbe, D. (2015). Lower vocal tract morphologic adjustments are relevant for voice timbre in singing. PLoS One, 10(7), e0132241.

Meerschman, I., Van Lierde, K., Ketels, J., Coppieters, C., Claeys, S., & D'haeseleer, E. (2019). Effect of three semi-occluded vocal tract therapy programmes on the phonation of patients with dysphonia: lip trill, water-resistance therapy and straw phonation. International journal of language & communication disorders, 54(1), 50-61.

Moisik, S. (2013). The epilarynx in speech (Doctoral dissertation).

Moisik, S., Esling, J. H., Crevier-Buchman, L., Amelot, A., & Halimi, P. (2015, August). Multimodal imaging of glottal stop and creaky voice: Evaluating the role of epilaryngeal constriction. In ICPhS.

Palaparthi, A., Riede, T., & Titze, I. R. (2014). Combining multiobjective optimization and cluster analysis to study vocal fold functional morphology. IEEE Transactions on Biomedical Engineering, 61(7), 2199-2208.

Palaparthi, A., Smith, S., & Titze, I. R. (2019). Mapping thyroarytenoid and cricothyroid activations to postural and acoustic features in a fiber-gel model of the vocal folds. Applied Sciences, 9(21), 4671.

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., ... & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American journal of speech-language pathology*, *27*(3), 887-905.

Pickup, B. A., & Thomson, S. L. (2010). Flow-induced vibratory response of idealized versus magnetic resonance imaging-based synthetic vocal fold models. *The Journal of the Acoustical Society of America*, *128*(3), EL124-EL129.

Radolf, V., Horáček, J., Dlask, P., Otčenášek, Z., Geneid, A., & Laukkanen, A. M. (2016). Measurement and mathematical simulation of acoustic characteristics of an artificially lengthened vocal tract. Journal of Sound and Vibration, 366, 556-570.

Samlan, R. A., Story, B. H., & Bunton, K. (2009). Kinematic modeling and acoustic measures of breathy voice. The Journal of the Acoustical Society of America, 126(4), 2221-2221.

Samlan, R. A., & Story, B. H. (2011). Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling. Journal of Speech, Language and Hearing Research, 56(4), 1209-1223

Samlan, R. A., Story, B. H., & Bunton, K. (2013). Relation of perceived breathiness to laryngeal kinematics and acoustic measures based on computational modeling. Journal of Speech, Language and Hearing Research, 56(5), 1267-1283.

Samlan, R. A., & Kreiman, J. (2014). Perceptual consequences of changes in epilaryngeal area and shape. The Journal of the Acoustical Society of America, 136(5), 2798-2806.

Samlan, R. A., Kunduk, M., & Ikuma, T. (2016). Vocal fold vibration and voice quality in the elderly. The Journal of the Acoustical Society of America, 140(4), 3443-3443.

Samlan, R. A., Kunduk, M., Ikuma, T., Black, M., & Lane, C. (2018). Vocal fold vibration in older adults with and without age-related dysphonia. American Journal of Speech-Language Pathology, 27(3), 1039-1050.

Sovijärvi, A. (1965). Die bestimmung der stimmkategorien mittels resonanzröhren. In Phonetic Sciences (pp. 532-535). Karger Publisher.

Stemple, J. C., Lee, L., D'Amico, B., & Pickup, B. (1994). Efficacy of vocal function exercises as a method of improving voice production. *Journal of voice, 8*(3), 271-278.

Story, B. H., Titze, I. R., & Hoffman, E. A. (1996). Vocal tract area functions from magnetic resonance imaging. The Journal of the Acoustical Society of America, 100(1), 537-554.

Story, B. H., Titze, I. R., & Hoffman, E. A. (1998). Vocal tract area functions for an adult female speaker based on volumetric imaging. The Journal of the Acoustical Society of America, 104(1), 471-487.

Story, B. H., Laukkanen, A. M., & Titze, I. R. (2000). Acoustic impedance of an artificially lengthened and constricted vocal tract. Journal of Voice, 14(4), 455-469.

Story, B. H., Titze, I. R., & Hoffman, E. A. (2001). The relationship of vocal tract shape to three voice qualities. The Journal of the Acoustical Society of America, 109(4), 1651-1667.

Story, B. H. (2002). An overview of the physiology, physics and modeling of the sound source for vowels. Acoustical Science and Technology, 23(4), 195-206.

Story, B. H. (2005). A parametric model of the vocal tract area function for vowel and consonant simulation. The Journal of the Acoustical Society of America, 117(5), 3231-3254.

Story, B. H. (2008). Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002. The Journal of the Acoustical Society of America, 123(1), 327-335.

Story, B. H. (2015). Mechanisms of voice production. The handbook of speech production, 34-58.

Story, B. H., Welch, G., Howard, D., & Nix, J. (2016). The vocal tract in singing. The Oxford handbook of singing.

Sulter, A. M., Miller, D. G., Wolf, R. F., Schutte, H. K., Wit, H. P., & Mooyaart, E. L. (1992). On the relation between the dimensions and resonance characteristics of the vocal tract: a study with MRI. *Magnetic resonance imaging*, *10*(3), 365-373.

Sundberg, J., & Sataloff, R. T. (2005). Vocal tract resonance (pp. 275-292). San Diego, California: Plural Publishing.

Titze, I. R. (1984). Parameterization of the glottal area, glottal flow, and vocal fold contact area. The Journal of the Acoustical Society of America, 75(2), 570-580.

Titze, I. R., & Story, B. H. (1997). Acoustic interactions of the voice source with the lower vocal tract. The Journal of the Acoustical Society of America, 101(4), 2234-2243.

Titze, I. R. (2002). Acoustic characteristics of resonant voice. J Voice, 15, 519-528.

Titze, I. R. (2002). Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. The Journal of the Acoustical Society of America, 111(1), 367-376.

Titze, I. R., Bergan, C. C., Hunter, E. J., & Story, B. (2003). Source and filter adjustments affecting the perception of the vocal qualities twang and yawn. *Logopedics Phoniatrics Vocology*, *28*(4), 147-155.

Titze, I. R. (2004). A theoretical study of F0-F1 interaction with application to resonant speaking and singing voice. Journal of Voice, 18(3), 292-298.

Titze, I. R. (2006). Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. Journal of Speech, Language and Hearing Research, 49(2), 448-459.

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. The Journal of the Acoustical Society of America, 123(4), 1902-1915.

Titze, I. R., & Laukkanen, A. M. (2007). Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. Logopedics Phoniatrics Vocology, 32(4), 147-156.

Titze, I. R., & Abbott, K. V. (2012). Vocology: The science and practice of Voice Habilitation. National Center for Voice and Speech.

Titze, I. R., & Palaparthi, A. (2016). Sensitivity of source–filter interaction to specific vocal tract shapes. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24(12), 2507-2515.

Titze, I.R., & Story, B.H. (2022). How well does the larynx “canal” match the ear canal? Journal of Singing, 78 (3), 359-361).

Tom, K., Titze, I. R., Hoffman, E. A., & Story, B. H. (2001). Three-dimensional vocal tract imaging and formant structure: Varying vocal register, pitch, and loudness. The Journal of the Acoustical Society of America, 109(2), 742-747.

Vampola, T., Laukkanen, A. M., Horáček, J., & Švec, J. G. (2011). Vocal tract changes caused by phonation into a tube: a case study using computer tomography and finite-element modeling. The Journal of the Acoustical Society of America, 129(1), 310-315.

Ventura, S. M. R., Freitas, D. R. S., Ramos, I. M. A., & Tavares, J. M. R. (2013). Morphologic differences in the vocal tract resonance cavities of voice professionals: an MRI-based study. *Journal of voice*, *27*(2), 132-140.

Verdolini, K. (2000). Resonant voice therapy. *Voice therapy: Clinical studies*, 46-61.

Wang, S. (1986). Singer's high formant associated with different larynx position in styles of singing. Journal of the Acoustical Society of Japan (E), 7(6), 303-314.

Xue, S. A., & Hao, G. J. (2003). Changes in the human vocal tract due to aging and the acoustic correlates of speech production. Journal of Speech, Language and Hearing Research, 46(3), 689-701. https://doi.org/10.1044/1092-4388(2003/054)

Yamasaki, R., Behlau, M., do Brasil, O. D. O. C., & Yamashita, H. (2011). MRI anatomical and morphological differences in the vocal tract between dysphonic and normal adult women. *Journal of Voice*, *25*(6), 743-750.

Yamasaki, R., Murano, E. Z., Gebrim, E., Hachiya, A., Montagnoli, A., Behlau, M., & Tsuji, D. (2017). Vocal tract adjustments of dysphonic and non-dysphonic women pre-and post-flexible resonance tube in water exercise: a quantitative MRI study. Journal of Voice, 31(4), 442-454.

Yiu, E. M. L., Lo, M. C., & Barrett, E. A. (2017). A systematic review of resonant voice therapy. International journal of speech-language pathology, 19(1), 17-29.

Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., ... & Glaze, L. E. (2011). Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). American Journal of Speech-Language Pathology, 20 (1), 14-22.