

人間の音声発話のしくみ

〔研究代表者〕

平山 亮

大阪工業大学情報科学部情報メディア学科／国立民族学博物館文化資源共同研究員

発話ロボット開発をめざして

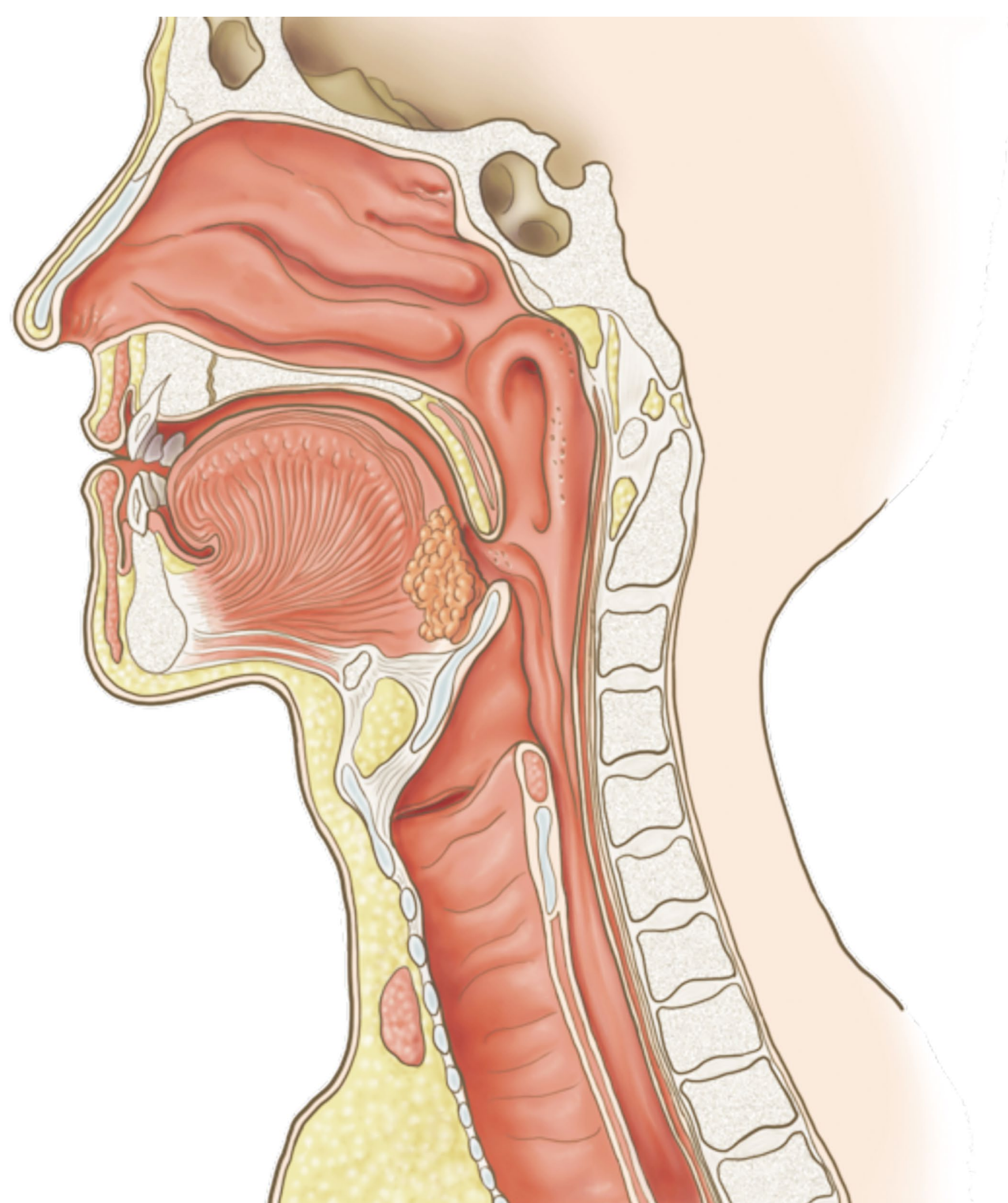
コンピュータに人間の言葉をしゃべらせる技術を音声合成技術といいます。音声合成技術はとても発展していて、音声合成技術によって作り出された音声は、様々なサービスで使われています。現在主流の音声合成技術では、アナウンサーや声優が話した声を録音し、細かく区切った音をつなぎ合わせたり、音声の特徴量を計算して音を合成したりして、人間が話すのに近い音声を作り出しています。

マルチメディア応用研究室（平山研）では、現在の音声合成技術とはまったく異なる方法で音声を作ることに挑戦しています。つまり、舌や唇をもった発話ロボットを作成し、人間の音声発話のしくみと同じように、舌や唇を動かして音声を生成しようとしています。現在は、人間の舌、唇などの器官と筋肉の部品や模型を製作することに取り組んでいます。

人間の音声生成のしくみ

人間の音声生成では、しゃべろうと思う発話内容から、音声に関連するどの器官をどう動かせばしゃべれるのかということを脳が計算して、運動指令を筋に伝えて器官を動かします。

多くの器官の協調的運動により、肺で作られた呼気が、声帯で音となり、舌や唇などの形や動きによって、いくつもの子音や母音の言語音へと調整されて、言語音声となって聞こえてきます。



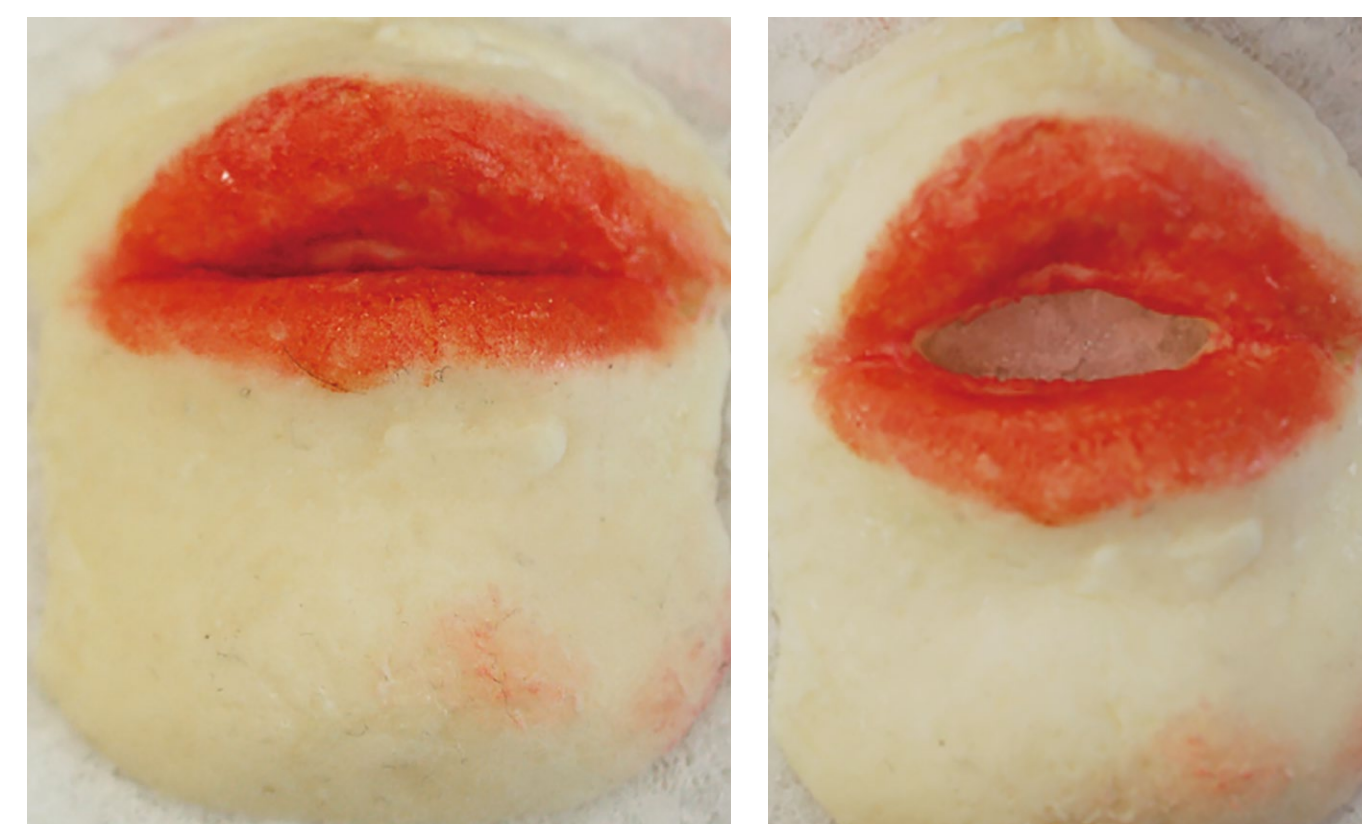
発声・発話器官

音声調音器官模型の作成

音声言語の教育用、音声生成のしくみの研究用、発話ロボットの部品の試作品として、舌と唇の模型を、粘弾性のあるウレタンゲル素材を用いて作成しています。模型製作には3Dモデリングマシン（切削機）や3Dプリンタを使います。



舌の粘弾性模型



唇の粘弾性模型



舌骨の樹脂模型

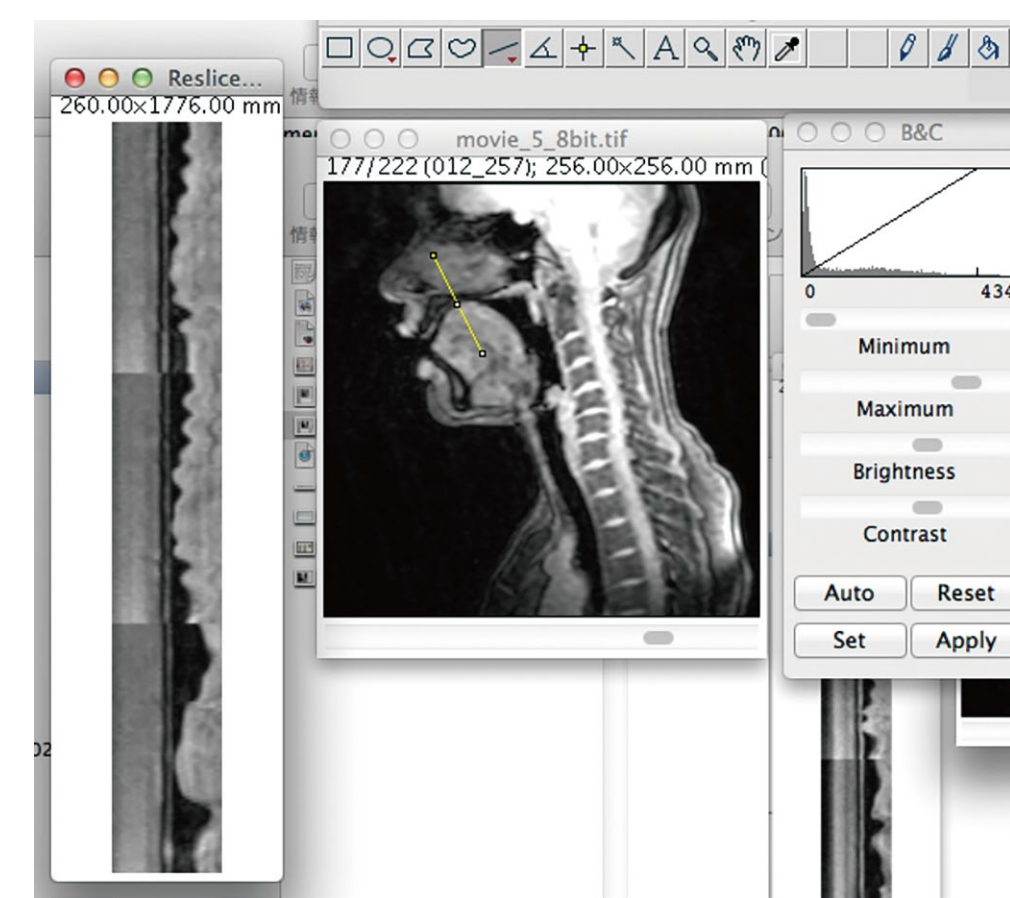


音声生成器官の計測実験

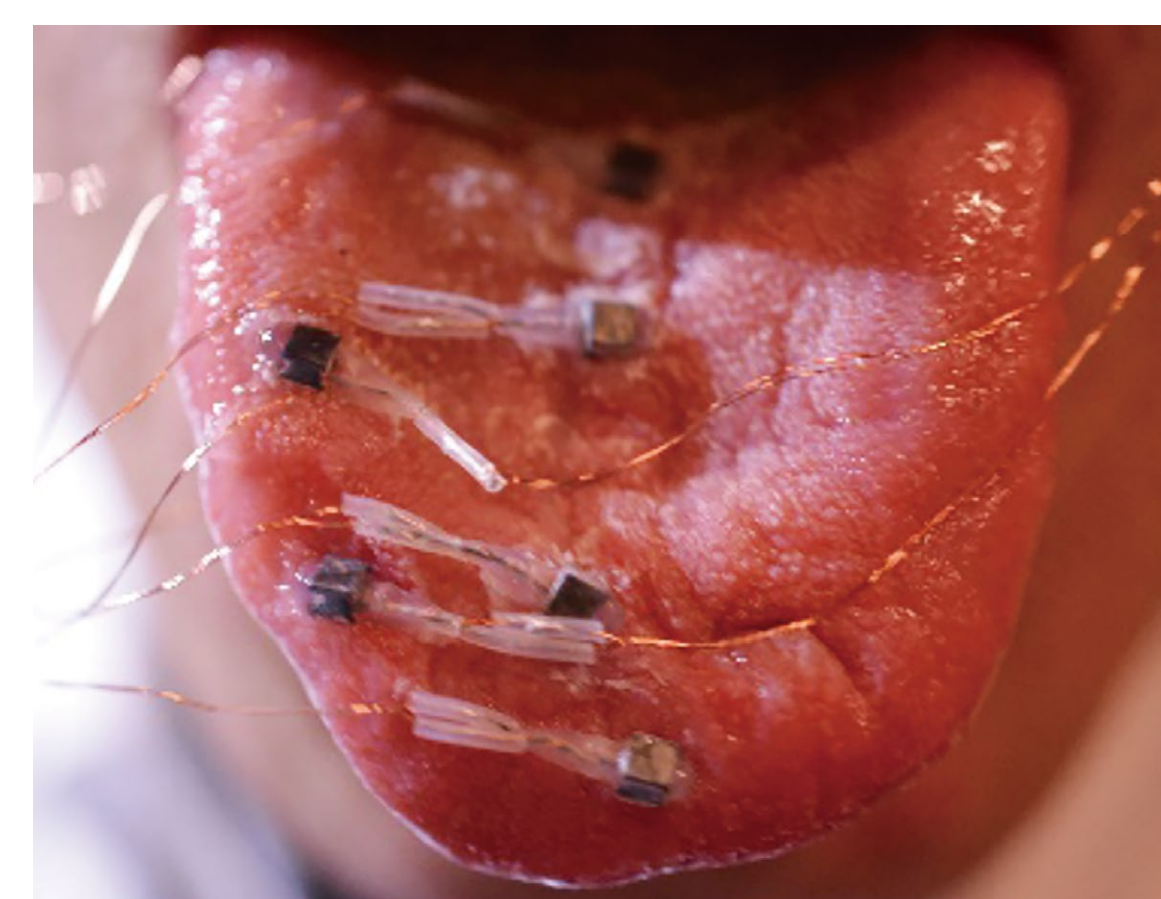
音声生成のしくみを研究するため、人がしゃべっている時の舌や唇の計測実験を行っています。MRI、3D位置計測装置などを使います。



MRIによる舌運動の計測と分析



磁気計測装置による舌運動の計測



The Human Speech Production Mechanism

[Principal Investigator]

HIRAYAMA Makoto J.

Graduate School and Faculty of Information Science and Technology, Osaka Institute of Technology/Cultural Resources Joint Researcher, National Museum of Ethnology

Toward the Development of Speech Robots

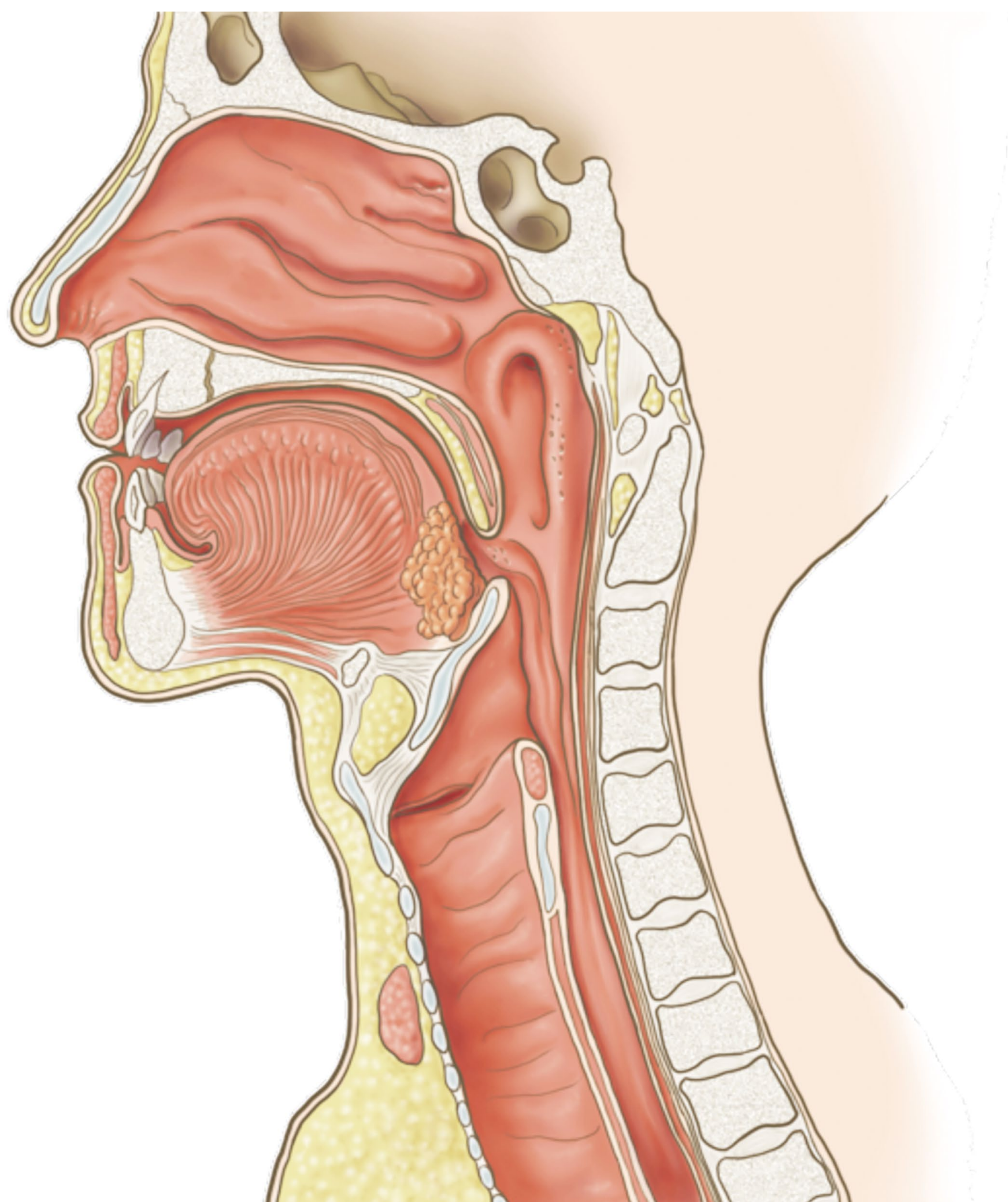
The technology that allows a computer to speak human words is called text-to-speech technology. This technology is highly advanced, and the voices produced by text-to-speech technology are used in a variety of services. Currently, mainstream text-to-speech technology produces speech similar to that of human speech by recording the voice of an announcer or voice actor and then joining finely segmented sounds together or synthesizing the sounds by calculating the speech features.

In the Multimedia Application Laboratory (Hirayama Lab), we are trying to create speech in a way that is completely different from current speech synthesis technology—we are attempting to create a speech robot with a tongue and lips and generate speech by moving them in the same way as in human speech production. Currently, we are working on producing the parts and models of the human tongue, lips, and other organs and muscles.

Mechanisms of Human Spoken Language Production

In human speech production, the brain calculates which organs related to speech sounds should be moved based on the content of speech that one intends to utter, and then it transmits motor commands to the muscles to move those organs.

Through the coordinated movement of many organs, the exhalation produced by the lungs becomes a sound in the vocal cords, which is coordinated into a number of consonant and vowel sounds by the shape and movement of the tongue, lips, and other parts, and is heard as a speech sound.



Speech Production Organs

Modeling of Speech Articulatory Organs

Models of tongues and lips are made from viscoelastic urethane gel material for use in teaching spoken language, research on the mechanism of speech production, and as prototypes for speech robot parts. We use a 3D modeling machine (cutting machine) and a 3D printer to make the models.



Viscoelastic model of the tongue



Viscoelastic model of the lips



Resin model of the hyoid bone

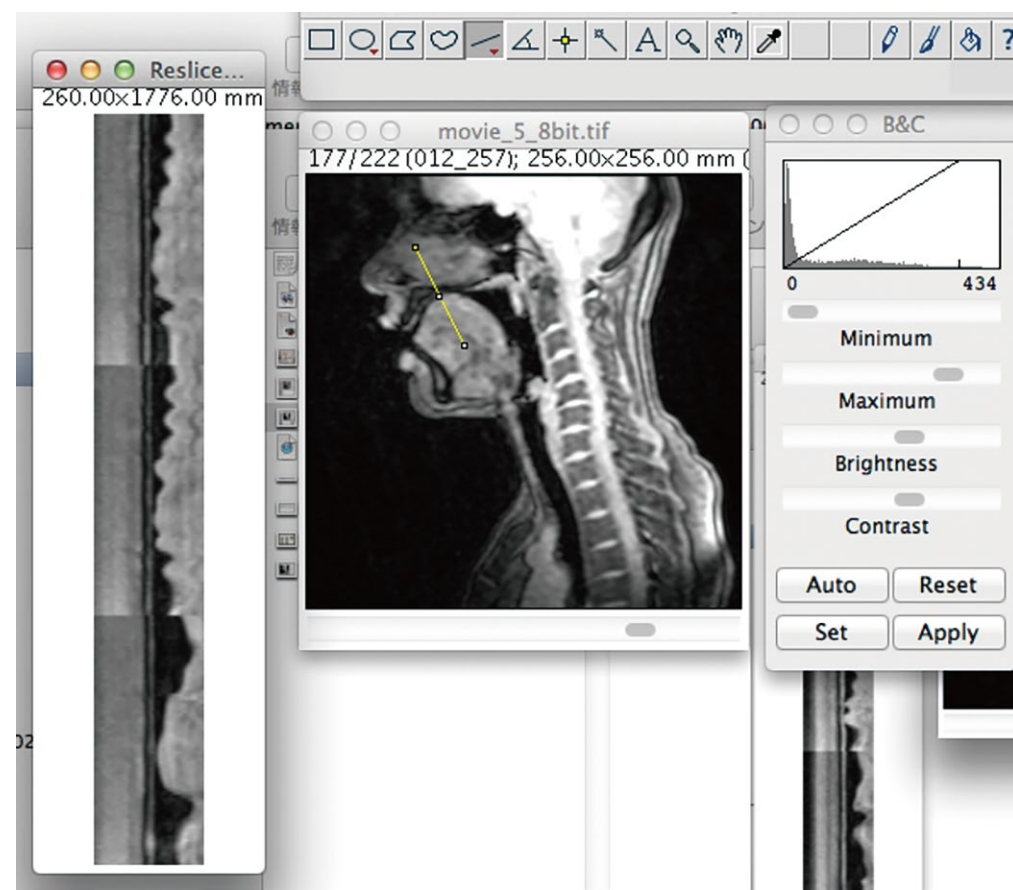


Experimental Measurement of Speech Production Organs

To study the mechanism of speech production, we are conducting experiments to measure the tongue and lips while a person is speaking, using equipment such as MRI and 3D positional measurement equipment.



Using an MRI to measure and analyze tongue movement



Using a magnetic measuring device to measure tongue movement

