Extended Abstract:
Development of vocal cord mechanism for a robot capable of infant-like speech and reproducing the pitch of a babbling and a shout

Tomoki Kojima[1], *Nobutsuna Endo*[2], *Minoru Asada*[1]

[1]Adaptive Machine Systems, Osaka University
[2]Graduate School of Science and Technology for Future Life, Tokyo Denki University
asada@ams.eng.osaka-u.ac.jp

Abstract

There are many mysteries related to the speech development process of a baby; the influence that the constraint imposed by the structure of the vocal organ such as the larynx / throat has on the speech development has not been elucidated. Therefore, we have developed a speech robot, specifically, the Lingua series, which has an infant's articulatory ability and can produce voices similar to infants [1, 2]. The purpose of development of Lingua is to develop a vocal robot platform that can be utilized in behavioral studies regarding the experiments of infant-caregiver interactions as a substitute for a real infant (Figure 1). In this article, we focus on an utterance platform that can precisely reproduce various utterances of real infants.

Index Terms: Lingua-R, pitch, babbling, shout, arytenoid cartilage

1. Lingua-R: a vocal robot platform

The Lingua can reproduce a pitch of 410–2000 Hz when the vocal cords are deformed manually [3]. However, several issues remain unsolved. The vocal cord folds could not satisfy the lower pitch of the babbling (300–400 Hz), the drive mechanism is not implemented, and the pitch control program is not coded. In this study, we propose the following (Lingua-R):

(i) a new vocal cord which satisfies the pitch range of the babbling/shout at the same time (300–1000 [Hz])
(ii) a drive mechanism to control the pitch
(iii) a pitch control code with auditory feedback

2. New vocal cords

First, regarding the vocal cord folds, the optimum values of shape parameters were determined in terms of five parameters related to the thickness and hardness of the fold [3] (Figures 2, 3). In this version, we improved the performance by adding two more parameters (angle of the fold and vertical thickness of the surface layer) in an attempt to keep the minimum and maximum pitch as low and high, respectively, as possible. As a result, the lowest pitch decreased from the conventional 410 Hz to 325 Hz, but at the same time, the maximum pitch dropped from 2000 Hz to 1028 Hz. However, these values cover 96 % of the pitch range of 300 (babbling)–1000 Hz (shout). Although it does not seem perfect, it seems to achieve the initial purpose.

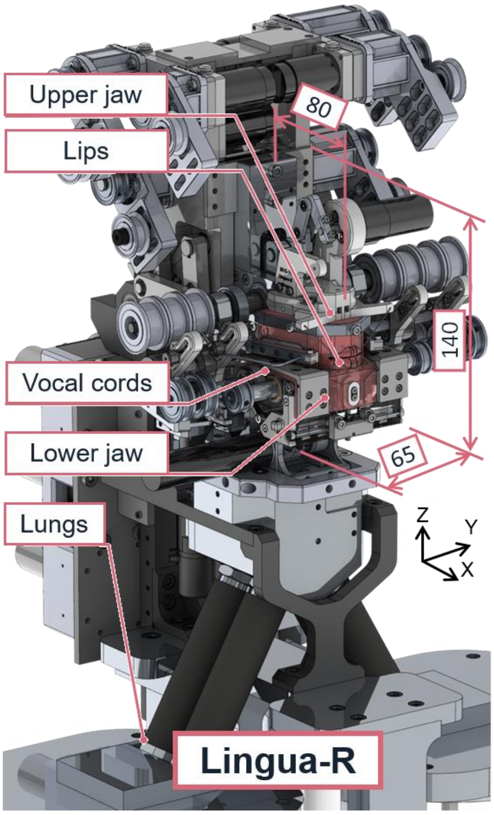

Figure 1: *Overview of Lingua-R.*

3. Arytenoid cartilage mechanism

To control the pitch of the voice, we implemented an arytenoid cartilage mechanism. The mechanism has two degrees-of-freedom (DOFs): open/close and stretch/relaxation of the vocal folds (Figure 4). Each DOF is driven by a DC motor. A motor driver controls the motor speed and an encoder measures the motor angle displacement. A Windows PC with I/O board controls the motor angle. The period of the position control of the motor is 2 ms. The control PC also controls the airflow into the vocal cords with the mechanism of the lungs. This system configuration enables simultaneous control between the vibration of the vocal cords by the airflow and stretching of the fold. Consequently, Lingua-R can produce voice and control its pitch.

4. Pitch control program

We also developed a pitch control program with auditory feedback. A PC measures Lingua-R's voice by a microphone, calculates the pitch, and sends the pitch value to the control PC. The control program changes the two-dimensional position of the distal end of the arytenoid cartilage mechanism according to a preliminarily calibrated pitch-position mapping. If there is an error between the desired and measured pitch, the program fine-tunes the position according to the gradient of the mapping.

5. Experiments and conclusion

As a result of the speech experiment, it was confirmed that the robot exhibited sufficient responsiveness, pitch followability, and stability to reproduce an infant's utterance (Figure 5). Although the pitch performance was numerically approved, the shout voice did not appear as realistic. Further improvement is required in addition to better coding for pitch control.

6. Acknowledgements

This research was supported by the JSPS Grants-in-Aid for Scientific Research (Research Project Number: 24000012).

7. References

[1] N. Endo, T. Kojima, Y. Sasamoto, H. Ishihara, T. Horii, and M. Asada, "Design of an articulation mechanism for an infant-like vocal robot "lingua"," in the 3rd Conference on Biomimetic and Biohybrid Systems (Living Machines 2014), pp. 389–391, 2014.

[2] N. Endo, T. Kojima, H. Ishihara, T. Horii, and M. Asada, "Design and preliminary evaluation of the vocal cords and articulator of an infant-like vocal robot "lingua"," in the IEEE-RAS International Conference on Humanoid Robots, pp. ThuI2–3.8, 2014.

[3] T. Kojima, N. Endo, T. Kojima, and M. Asada, "Development of Vocal Cords of an Infant-like Vocal Robot based on Anatomical Structure", Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2015), 2015.

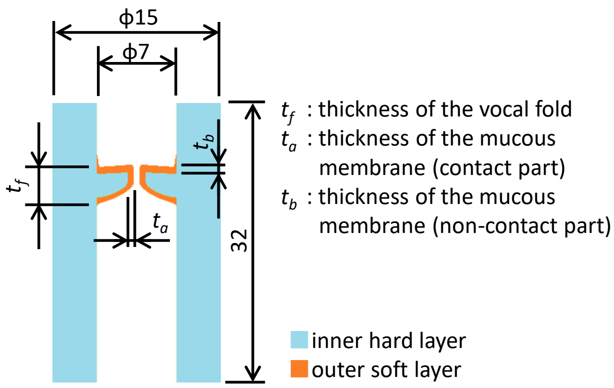

Figure 2: *Two-layered vocal cords for Lingua.*

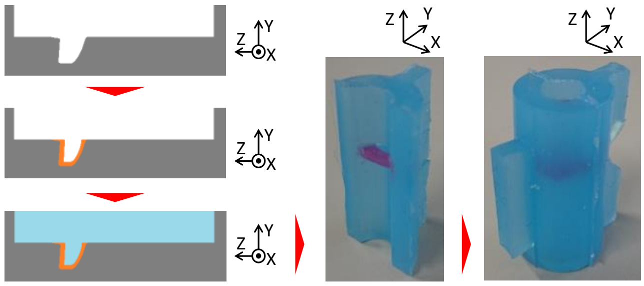

Figure 3: *Fabrication process of the vocal cords.*

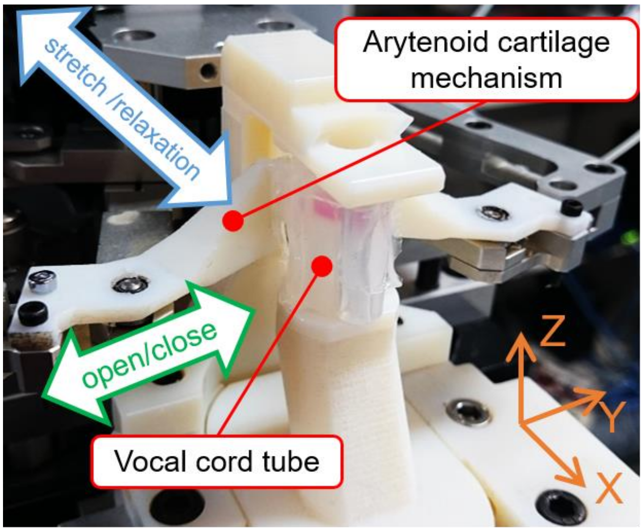

Figure 4: *Arytenoid cartilage mechanism.*

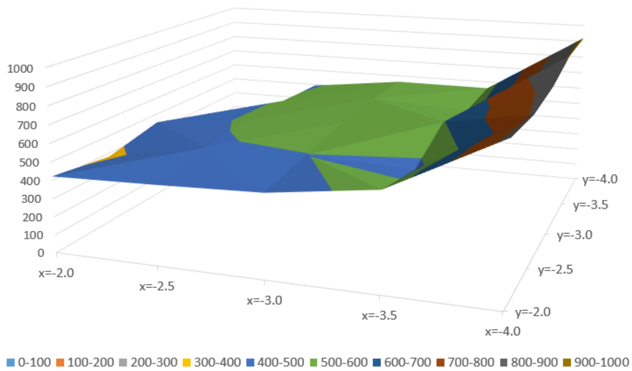

Figure 5: *Pitch performance in terms of two (x,y) parameters in driving the arytenoid cartilage arm.*