

音声発達研究に向けた乳児様発話ロボットの 高ピッチ発話可能な声帯の開発

○小島朋樹(阪大) 遠藤信綱(阪大) 小島友裕(阪大) 浅田稔(阪大)

1. 緒言

発声は人間の意思疎通において最も重要な手段の一つである。特に乳児が言葉を発するという行為を獲得するための前提となる音声発達のメカニズムについては、種々の試みはあるものの、十分な解明には至っていない。乳児の音声発達を、知覚・運動学習に関する神経的要因、他者との交流による社会的要因、発話器官の構造が声に与える影響に関する身体的要因の三つに分けて考えれば、前者二つは様々な計算モデル[1]や、マイクからの音声入力を受けて音声模倣する発話ロボット[2]などが開発されて研究が進んでいるが、乳児サイズの身体的要因については触れられてこなかった。この要因に関して解明を進めるためには、実際の乳児を解析対象とすることが望ましいが、任意の実験環境を再現する統制は非常に困難である。そこで我々は、統制可能な赤ちゃんとして、乳児様の身体的特徴を持つ発話ロボット Lingua [3] を開発している。

Sasamoto et al. [4]は乳児様発話ロボットを開発した。乳児様の高い声を出せるが、調音能力に乏しく、母音様の発声には至らなかった。また、Fukui et al. [5]の開発した発話ロボット WT-7RII は高い調音能力を持ち、母音も明瞭に発話できるだけでなく子音までも再現可能であったが、基本周波数が成人男性のものであり、乳児様の発話には対応していなかった。そこで、我々は Lingua を、これらの長所を兼ね備え、乳児様の発話が可能で高い調音能力を持つ発話ロボットとして開発した。

図1と表1に、Lingua の外観と各器官の自由度を示す。Lingua は肺、声帯、そして声道を持つ。肺と声帯の駆動機構は WT-7RII のものをそのまま用いており、声帯形状は Sasamoto et al. の乳児様発話ロボットと同様に乳児の解剖学データ[6]を基にしている。これまでに我々は、声帯の硬さと声の基本周波数との関連性について報告し、Lingua の声帯は、ヒト乳児の喃語における基本周波数の範囲 (300~410[Hz]) をカバーする発声が可能であることを確認した[7]。しかし、ヒトの乳児の声は泣きや笑いなどの感情を込めた場合、その基本周波数は 1000[Hz]を超えることが分かっており[8]、Lingua の声帯はそれらの発声を再現することができない (基本周波数域 250~500[Hz])。

そこで本研究では、ヒト声帯の解剖学的構造に基づき、ヒト乳児の泣きや笑いなどにおける 2000[Hz] もの高い基本周波数を安定して発声可能な乳児様声帯を開発した。

2. 発話のしくみ

2.1 発話器官

ヒトの発話に関わる器官には、肺、声帯、声道があり、声道は舌や鼻腔、口唇などにより構成される。肺の収縮により呼気圧が上昇し、それにより声帯が振動し喉頭原音が作られ、声道で反響して人間らしい音に変化し、外部に音波として放出される。ここで声帯に着目すると、声帯は図2(b)に示す喉頭の中に存在し、図2(a)のように筋肉・靱帯とそれを覆う粘膜が周りの軟骨と結合する形をとった2枚のひだからなる。その内部は図3のように、甲状披裂筋や声帯靱帯などからなる硬いボディ層と、粘膜からなる柔らかいカバー層の2層に分かれている。発話時は肺から送られる空気がこのひだの間(声門)を通り、ひだが振動して空気の疎密波が形成され、発声するしくみになっている。

表1 Lingua の自由度表

部位	自由度
顎	1
舌	7
軟口蓋	1
声帯	2
肺	1

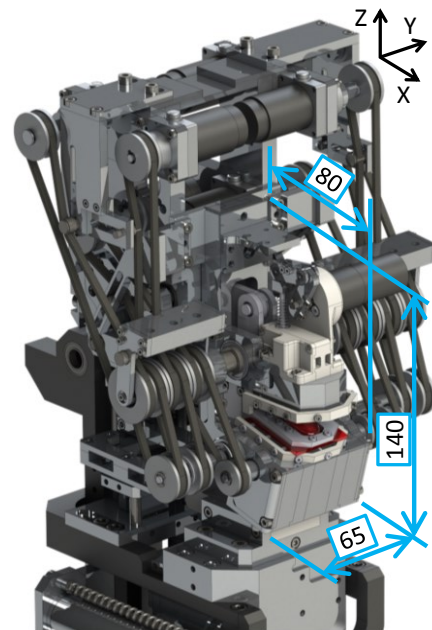


図1 Lingua の外観

2.2 声帯振動のしくみ

図4に声帯の振動中の振る舞いを示す。声帯の振動では主に声帯粘膜が下から上へ波打つように運動しており、内部の筋肉や中間層はほとんど振動しない。このようなひだの上下で位相差のある振動を粘膜波動と呼び、これにより声門閉鎖時にひだの間に強い負圧が生じ、低い呼気圧で安定した発声を可能にする。粘膜波動を実現するための条件は次の二つである[9]：

- (i) 声帯が内部の硬い筋肉層と表面の柔らかい粘膜層に分かれていること
- (ii) 粘膜層が低い粘弾性をもつこと

高ピッチ時、声帯は周囲の軟骨の動きで前後に伸展し、声帯振動に関わるひだの質量と接触面積が小さくなることで、より高速での粘膜波動が可能になる。しかし同時に、ひだの粘膜層の伸展により表面の粘弾性が上昇し、ひだ接触面の上下の位相差が徐々に小さくなり、最終的には位相差がゼロになるために十分な負圧を発生できず、粘膜波動を起こせなくなり、音が出なくなる。逆に、通常発声時に粘膜層が低い粘弾性を持ち、ひだ上下の位相差を十分に保てば、より高ピッチでの発声が可能になる。

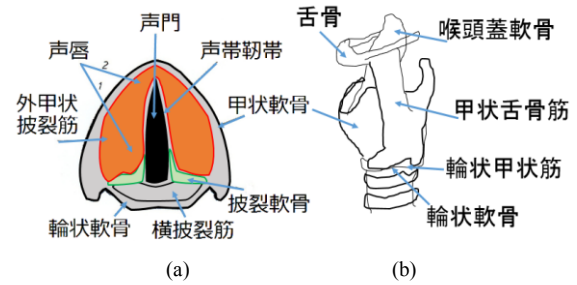


図2 (a) ヒトの声帯を上から見た図[10]
(b) 喉頭の外観図[11]

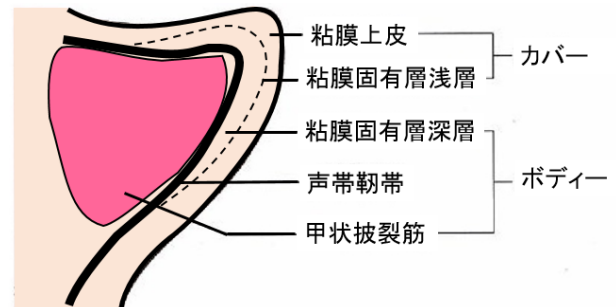


図3 ヒトの声帯の冠状面図[11]

3. 二層構造声帯ひだを持つ声帯の製作

3.1 Lingua の声帯の問題

これまでの Lingua の声帯 (図5) は、ヒトの声帯 (図2, 3) と比べて大きな隔たりがある。Lingua の声帯は全体が一様な材料からなり、ひだが硬い一層構造であるため発声しか行えなかった (250~500[Hz])。また駆動機構 (図6) は、筒を外側から押し潰して後方に引っ張る機構しか持たず、前方の固定点がないので、筒全体が後方に引っ張られ、ひだが十分に伸びなかった。

3.2 二層構造声帯ひだ

上述のヒト声帯の多層構造を基に、本研究では Lingua の声帯ひだを、硬い筋肉層と柔らかい粘膜層の二層構造とすることにした。その冠状断面図と、最適値を探索すべき形状パラメータを、図7に示す。ひだの各層はそれぞれ異なる硬さのシリコン樹脂 Dragon Skin FX-Pro (Smooth-On, Inc.) で成形されている。この樹脂は可塑剤の混合比によって硬さを変えることができる。

図8にこの声帯の成形過程を示す。まず、正中矢状面に沿って半分に切った状態の声帯を作るための雌型と、声帯ひだ外側の柔らかい粘膜層のための雄型を用い、粘膜層を成形する。次に、別の雄型を用いて、ひだ内側および声帯外壁をなす硬い層を成形する。最後に、雌型から半分の声帯を取り出し、対称に成形したもう半分の声帯と、声帯外壁の断面を合わせて接着し、二層構造のひだを持つ声帯が完成する。

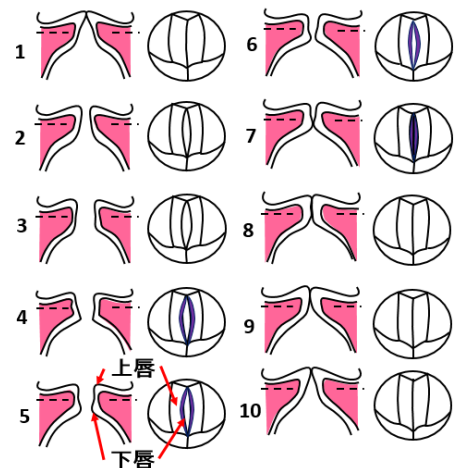


図4 ヒトの声帯の振動中の振る舞いの冠状断面図と水平面上方より見た図[12]

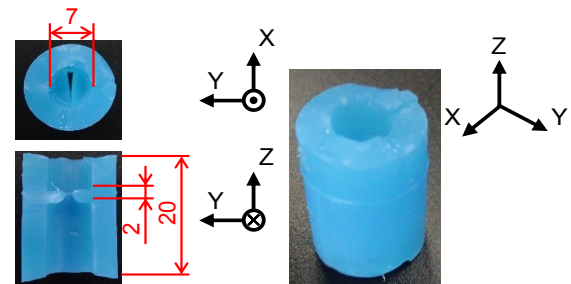


図5 Lingua の声帯

4. 発声実験

目標とする高い基本周波数域を発声可能な、二層構造声帯ひだの形状・硬さパラメータの組み合わせを決定するために、複数パターンの形状・硬さパラメータの組み合わせを持つ二層構造声帯を作り、発声実験により評価した。

4.1 実験手順

二層声帯において未確定であるパラメータは以下の5つである。

- ・ h_o : 粘膜層の柔らかさ
- ・ h_i : 筋肉層の柔らかさ
- ・ t_a : ひだ接触部の粘膜層の厚さ
- ・ t_b : ひだ接触部以外の粘膜層の厚さ
- ・ t_f : ひだ自体の厚さ

ここで各層の柔らかさは、Dragon Skin FX-Pro に対する可塑剤の質量比[%]で定義し、大きいほど柔らかい[7]。

これらのパラメータを変えた二層声帯を 49 種類製作した (h_o : 150, 200, 250, 300[%]. h_i : 25, 37.5, 50, 75[%]. t_a : 0.5, 0.75, 1.0, 1.25[mm]. t_b : 0.5, 0.75, 1.0, 1.25[mm]. t_f : 3, 4[mm])。これらを Lingua に搭載し、肺機構から空気を送り込みつつ、声帯駆動機構の代わりに指先で筒を変形させてひだを伸張させることで発声させ、この音声を計測し、基本周波数の変化を Praat [13]で解析した。

4.2 実験結果

実験で使用した 49 種類の二層声帯のうち、安定して発声した 28 通りのパラメータの組み合わせからなる声帯の、最高・最低の基本周波数を、図 9, 10 に示す。これらのパラメータは $t_f=3$ [mm], $h_o=200$ [%]で、 t_a , t_b , h_i は図内に示すように変化させた。

また、 $t_f=3$ [mm], $h_o=150, 250, 300$ [%], $h_i=50$ [%]の声帯は、そもそも発声しなかった。また $t_f=4$ [mm], $h_o=200$ [%], $h_i=50$ [%]の場合は、いずれも最低基本周波数が 360~400[Hz]程度、最高基本周波数は 450~500[Hz]程度であった。

実験の結果、最も広範囲の基本周波数で発声することができたのは、以下のパラメータの組み合わせの二層声帯であった。

- ・ 粘膜層の柔らかさ : 200[%]
- ・ 筋肉層の柔らかさ : 50[%]
- ・ ひだ接触部の粘膜層の厚さ : 1.25[mm]
- ・ ひだ接触部以外の粘膜層の厚さ : 1.25[mm]
- ・ ひだ自体の厚み : 3[mm]

この組み合わせの声帯で発声実験を行った時の基本周波数変化を図 11 に示す。最高値は 2047[Hz]、最低値は 405[Hz] であった。

4.3 考察

ひだの厚み t_f が 3[mm] のときと比べると、4[mm] のときは低音が出やすかった (360[Hz] 出る瞬間も

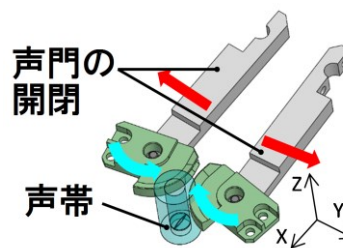


図6 Linguaのピッチ調節機構

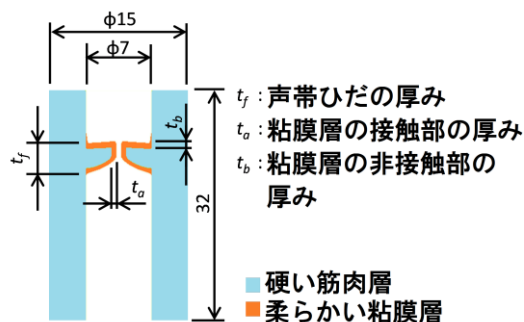


図7 二層構造のひだを持つ声帯

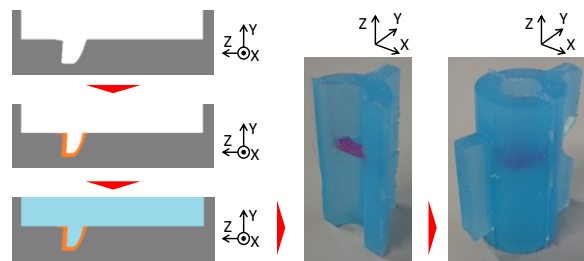


図8 二層構造の声帯の成形過程

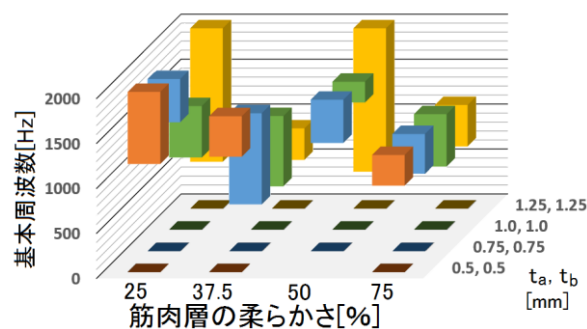


図9 実験結果 (厚み比固定. $t_f=3$ [mm], $h_o=200$ [%])

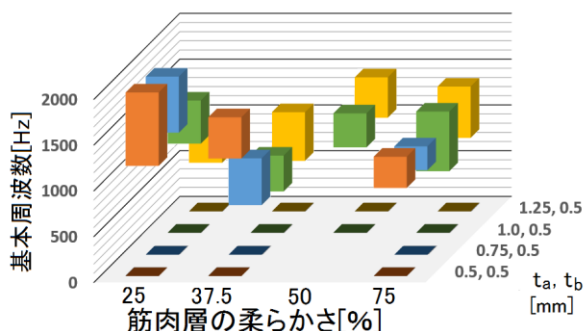


図10 実験結果 (t_b 固定. $t_f=3$ [mm], $h_o=200$ [%])

あった)が、高い基本周波数は出せなかった(高く
て800[Hz])。これは、ひだ伸展時に薄くなりきらず、
粘膜波動の周期を短くできなかったことが原因とし
て挙げられる。また、粘膜層の厚み t_a が0.5[mm]の
時は全く発声しなかった。粘膜層は1[mm]以上はな
いと粘膜波動を発生できないと考えられる。ひだ接
触部以外の厚み t_b も同様に1[mm]以上ないと発声し
にくかった。これは、ひだ接触部のみならず、接触
部の上下の厚みも振動に関わっており、特に粘膜移
動性を高めるという点で重要であると考えられる。

筋肉層の柔らかさ h_i が25[%]だと硬すぎて筒が変
形しにくかった。逆に筋肉層があまりに柔らか過ぎ
ても、高音は鳴りづらかった。2.2節に示した条件(i)
のように、粘膜層と筋肉層との柔らかさの差はある
程度確保されなければならない。また、粘膜層の柔
らかさ h_o を250[%]以上にしてみたところ、ひだ表
面に可塑剤の薄膜ができ、そこに小さな気泡が発生
し、発声が長続きしなかった。しかし200[%]だと安
定して発声したので、200[%]と250[%]の間に更なる
適値があると考えられる。これはひだの厚みも同
様で、3[mm]と4[mm]の間に適値があると考えられ
るが、さらに今後、パラメータを最適化していき
たい。また、シリコン樹脂および成形法を再検討す
る必要もある。

最後に、基本周波数の最低値が旧声帯のものより
も低くなってしまった。旧声帯の振動モードと本研
究で製作した二層構造声帯の振動モードは異なると
考えられるので、高速度カメラで実際の振動の様
子を観察しつつ、さらに粘膜層の粘弾性を低め、よ
り低い周波数での粘膜波動も実現できる声帯を作
る必要がある。

5. 結言

本研究では、ヒト声帯の解剖学的構造に基づき、
声帯ひだを二層構造とした、2000[Hz]を超える高
い基本周波数の発声が可能な声帯を開発した。今
後は、高域での発声はそのままに、最低基本周波
数を改善し、Linguaの更なる性能向上を図って
いく。

謝 辞

本研究は日本学術振興会科学研究費補助金「特
別推進研究」(24000012)、同省「医・工・情報連
携によるハイブリッド医工学産学連携拠点整備事
業ー医工情報連携センター構築にむけてー」の助
成を受けた。また、早稲田大学理工学術院高西淳
夫研究室・スポーツ科学学術院菅田雅彰研究室より
人間形発話ロボットWT-7RIIの提供を受けた。

参 考 文 献

[1] K. Miura et al.: “How caregiver’s anticipation shapes infant’s vowel through mutual imitation”, IEEE T. Auton-

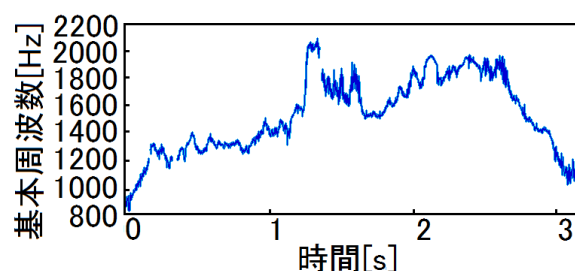


図 11 基本周波数の変化

- omous Mental Development, vol. 1, no. 4, pp.217-225, 2009.
- [2] Koh Hosoda et al.: “A constructivist approach to infants’ vowel acquisition through mother-infant interaction”, Connection Science, vol. 15, no. 4, pp.245-258, 2003.
- [3] 小島友裕ら: “音声発達過程の構成的理解のための乳児様発声プラットフォームの開発”, 日本機械学会ロボティクス・メカトロニクス講演会 2014, 2A1-X06, 2014.
- [4] Y. Sasamoto et al.: “Towards understanding the origin of infant directed speech: A vocal robot with infant-like articulation”, Proc. of the IEEE Third Joint Int’l Conf. on Development and Learning and Epigenetic Robotics, 2013.
- [5] K. Fukui et al.: “Speech Robot Mimicking Human Articulatory Motion”, Proc. of INTERSPEECH 2010, pp. 1021-1024, 2010.
- [6] Hans Edmund Eckel et al.: “Morphology of the human larynx during the first five years of life studied on whole organ serial sections”, The Annals of Otology, Rhinology & Laryngology, No.108, pp.232-238, 1999.
- [7] 小島友裕ら: “乳児様声帯における硬さによる発声特性変化の評価”, 第32回日本ロボット学会学術講演会予稿集, 1B3-09, 2014.
- [8] M. Robb and J. Saxman: “Developmental Trends in Vocal Fundamental Frequency of Young Children”, Journal of Speech, Language and Hearing Research, vol.28, pp.421-427, 1985.
- [9] 一色信彦(編): 喉頭機能外科-とくに経皮的アプローチについて-, 京都大学医学部耳鼻科同窓会, 1977.
- [10] 須永義雄ら訳, フレデリック・フースラーら著: うたうこと 発声器官の肉体的特質-歌声のひみつを解くかぎ-, 音楽之友社, 1987.
- [11] 山下敏夫(編): 新図説耳鼻咽喉科・頭頸部外科講座 第4巻 口腔・咽頭・喉頭・気管・食道, 廣済堂, 2000.
- [12] 板橋秀一(編): 音声工学, 森北出版, 2005.
- [13] “Praat: doing Phonetics by Computer”, <http://www.fon.hum.uva.nl/praat/>