

3-3 人間の生物学的機能の解明に基づいた音声コミュニケーション機構の研究

3-3 Study of Speech Communication Mechanisms based on the Understanding of Human Biological Functions

本多清志 (株式会社 国際電気通信基礎技術研究所)
Kiyoshi HONDA

要旨

人間の音声には、思考、通信、情感及び身体情報の媒体としての機能があり、これらが一体となって人間の生活と社会の形成に大きな役割を果たしている。音声の意義はこの多機能性にあり、人間に生得的に備った生物学的機能にその基盤を見いだすことができる。これらすべての音声機能を代行する装置が実現されたならば、人間に等しい価値を認めることができるだろう。本プロジェクトでは、人間の生物学的機能の理解に基づいて音声言語を模擬する自動機械を創造することを将来の目標として、発話器官の形態的特性、音声生成機構のモデル化、音声に関わる脳機能の三領域について研究を開始した。

Human speech acts as vehicles of thought, communication, emotional state, and physical identity; these functions together contribute to our life and social organization. The value of speech is given by its versatility, which is derived basically from innate biological abilities equipped to humans. If the machinery emerged that handled all the speech functions, it would deserve human-equivalent significance. Towards evolution of spoken language automata as a prospective goal, we launch a research project to look into biological functions underlying human speech in the three areas; form and function of the speech organs, biophysical modeling of speech production mechanisms, and functional imaging of speech-related pathways.

[キーワード]

音声, 生物機能, 個人性, 話者正規化, 磁気共鳴画像法

Speech, Biological functions, Individuality, Speaker Normalization, Magnetic resonance imaging

1 はじめに

人間の音声は、様々な情報の媒体として人間の生活あるいは社会の形成に大きな位置を占めている。音声の機能を改めて考えてみるならば、まず内言語とよばれる思考の道具としての役割があげられる。この音波とならない音声の表象は論理の理解や文筆の準備に使用され、人間の知的活動と文化の伝承に貢献している。通信の媒体としての音声の機能は広く理解されており、その重要性には計り知れないものがある。音声による通信が一斉に断たれた場合には社会に著

しい混乱を与えることが容易に想像され、社会生活において音声通信は不可欠なものになっている。また、音声には情緒的意識を反映する働きがあり、声の調子などを使い分けることにより人間関係の醸成に用いられる。このほかに、音声には個人認証あるいは身体状況の情報源としての機能があり、咳払いを聞いただけでも誰であるかが分かったり、風邪をひいていることが感じられたりする。

これらの多様な音声の機能についての理解が進み、もし仮にすべての機能を工学的に代行する装置が出現したならば、おそらくそれは有機

的な人工物として人間に等しい価値を認めることができるだろう。今後の音声研究では、人間の音声の諸機能を模擬する自動機械の創造を目標として、その基盤となる生物的な問題を根本的に解決することが期待される。

2 音声コミュニケーション研究の課題

2.1 音声研究小史

音声の研究は1700年にフランスで行われた声帯振動の実験を始まりとすれば、ちょうど300年の歴史を経たことになる。18世紀後半には早くもフォン・ケンペレンが母音や子音の発音機構を分析し、機械式音声合成装置を製作している。19世紀になるとホイットストーンやヘルムホルツらが母音の生成理論を展開し、フォルマントが母音の音色を決定することを見いだしている[1]。その頃、声帯模型によって発音機構の研究を行ったJ・ミューラーは、音源・フィルタ理論を最初に説いた人といわれる。この理論の流れは、20世紀に入り千葉と梶山によるX線撮影に基づく母音の研究やファントによる音声生成の音響理論の研究に引き継がれてきた[2]。この音響管の理論は次第に単純化され、声道を一本の管で代表することが普通になっている。しかし、このような簡略化が音声生成機構の説明として必ずしも妥当とはいえ、音声を作り出す生体機構そのものを調べ直す時期にきているように思

われる。

2.2 残された研究課題

この300年の間に人間の音声について明らかになった事実も多いが、残された課題も多い。ここでは、21世紀に託されることになった二つの問題を取り上げ、当面の音声研究の課題とする。

第一には、音声の個人性の問題があげられる。顔の見かけが一人ひとり異なるように、音声の音色も個人ごとに異なる。顔と音声の個人性は本人の意志で変更できない身体的情報であり、個人認証技術への適用が期待されている。音声の個人性をもたらす生成要因については憶測の域を出ないので、声道の概形や分岐管の特性などから検討を進める必要がある。図1に示すように、声道には鼻腔と梨状窩という二つの分岐管があり、スペクトル上の起伏に個人差が反映されると考えられる[3]。このほかに、声道の長さも音声の個人性の生成要因と指摘され、咽頭腔の長さがフォルマントに影響することが推測されている[4]。

第二の問題として、母音の等価性があげられる。親と子供は身体の違いがあるので母音の音色にも隔たりがあるが、この違いは意識されることなく、親子間の通信においても音声の互換性が保たれている。この性質は母音の正規化とよばれ古くから議論されてきたが、現在でも音声知覚理論では説明のつかない問題として

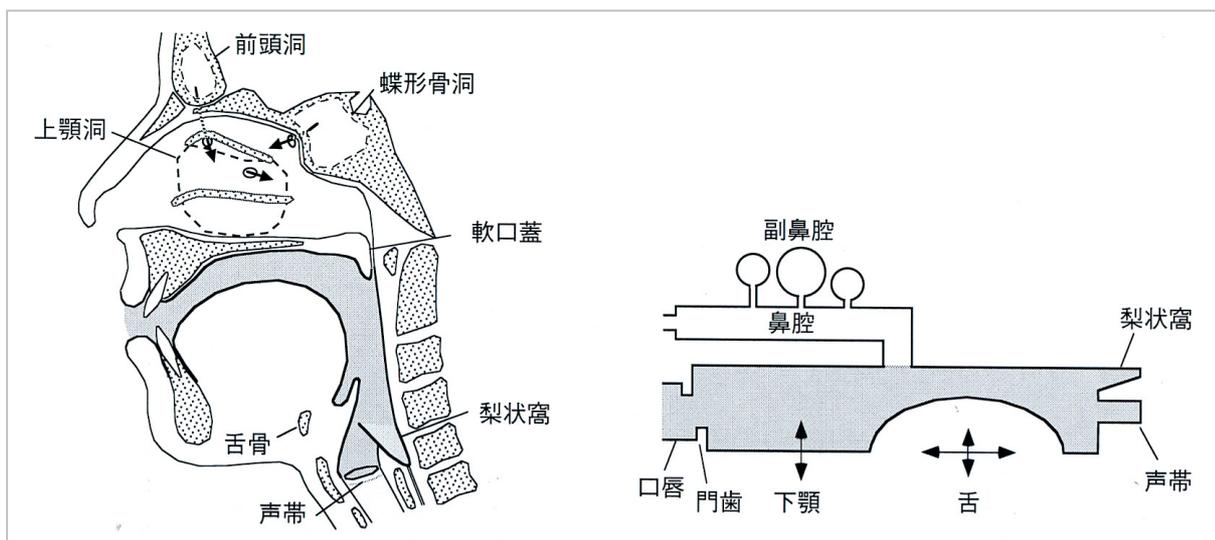


図1 発話器官の微細構造(左)とその音響管モデル(右)

取り上げられている。母音のフォルマントは声道長に逆比例して広い範囲で変動し、成人の/a/と子供の/o/がほぼ同じ位置に分布するが、我々はこの二つの母音を容易に区別できる。複数の話者の音声を人間がどのようにして対応付けしているかは音声認識に関わる問題であるが、その機構がどこにあってどのように機能しているかはいまだに分かっていない。

2.3 音声の生物学的背景

以上の二つの問題は、音声で個人を識別し得る程度に異なる信号でありながら通信の道具としての互換性を備えていることを意味している。この特性は音による情報交換が原始的な段階から進化する過程で備わったものと考えられ、音声で個体の識別にもメッセージの通信にも用いられる点は動物にも共通している。特に親子間の個体識別については鳥類や哺乳類で調べられており、鳥類の個体識別では視覚よりも聴覚が優位であることも知られている。動物の音声行動の研究は、人間の生物学的機能を理解する上で参考になる点が多い。

3 磁気共鳴画像法(MRI)を用いる音声研究

以上のような音声通信の生物学的側面を重視して、ATRにおける音声生物学プロジェクトでは、発話器官の形態計測、音声生成機構のモ

デル化、音声に関わる脳機能画像の研究を進めている。磁気共鳴画像法を主な観測法として用いてきたので、二、三の成果を紹介する。

3.1 発話器官の形態計測

声道の概形はX線撮影で観測されてきたが、3次元形状は磁気共鳴画像法(MRI)により初めて観測可能になり、歯列の造影法により観測精度が向上した。図2(a)は母音「エ」発話時のMRIデータより声道部分を抽出したものであり、単純な管とみなされてきた声道に大小の分岐管があることが分かる。また、撮像性能の進歩により3次元の動画を記録できるようになっている[5]。図2(b)は日本語の「アイウエオ」を発話したときの3次元動画より声道形状の時間変化を抽出したもので、分岐管を除いた声道部分の断面積関数を表示してある。MRIの利用は音響管理論の再検討と個人性要因の解明に威力を発揮すると思われる。

3.2 発話時の脳活動

発話時及び聴取時の脳活動は脳皮質について詳しく調べられているが、生物学的機能からみた場合には脳幹を含む領域に観測野を広げる必要がある。機能的磁気共鳴画像法(fMRI)により脳幹を撮像するには幾つかの技術的な問題があり、撮像方向などの検討を加えて発話時における脳幹を含む領域の活動を記録した。図3は自発発話時における一例で、視床、橋、小脳などの

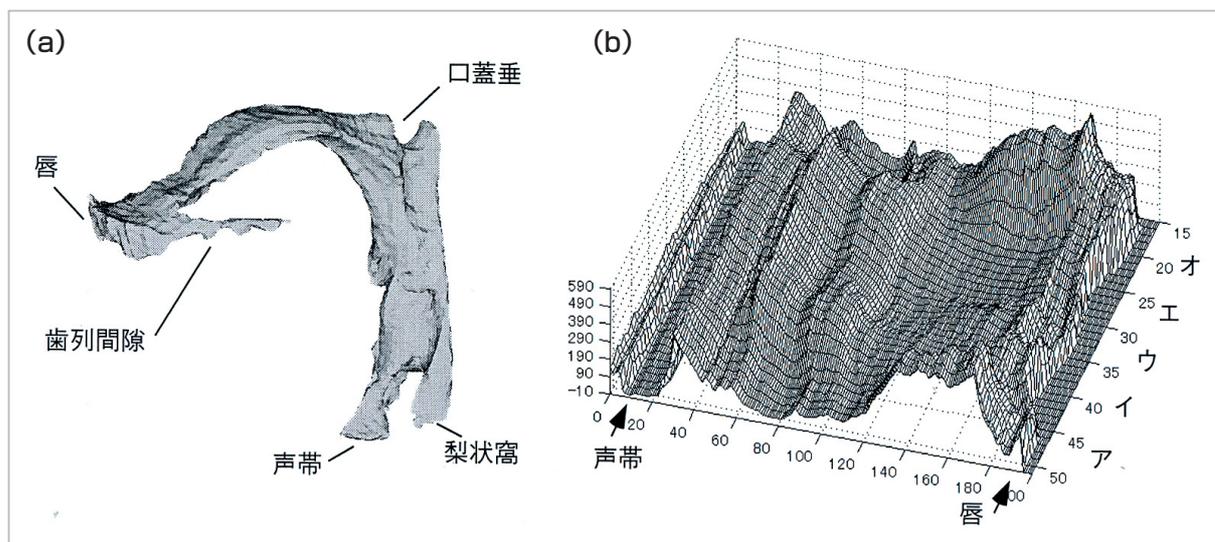


図2 声道の3次元形状(左)と「アイウエオ」における断面積関数の時間変化(右)

一般運動系の活動が視覚化されている。また、MRIを音声知覚の研究を利用するために装置から発生する騒音の抑制法を検討し、音声の生成と知覚とに関わる神経経路を明らかにする研究を計画している。

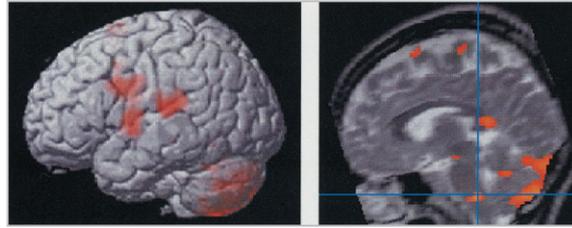


図3 文章発話時の大脳皮質及び小脳の活動部位(左)と脳幹運動系の活動部位(右)

参考文献

- 1 T. Chiba and M. Kajiyama, "The Vowel: Its Nature and Structure", Tokyo: Tokyo-Kaiseikan, 1941.
- 2 G. Fant, "Acoustic Theory of Speech Production", The Hague, Mouton, 1960.
- 3 本多清志, "音声の生物学的基礎", 講座言語の科学2, 音声, 東京, 岩波書店, 1998.
- 4 本多清志, "人の顔形と声質", 日本音響学会誌, pp.308-313, Vol.57, 2001.
- 5 竹本, 本多, 正木, 島田, 藤本, 高野, 武尾, "MRIによる発話運動の3次元動画記録", 日本音響学会講演論文集, pp.251-252, Mar., 2001.

本多清志

(株)国際電気通信基礎技術研究所(ATR)
先端情報科学研究部音声生物科学プロ
ジェクトリーダー 医学博士
音声科学