

声道形状を単純化した模型による音声の音響教育

荒井 隆行

上智大学理工学部情報理工学科 〒102-8554 東京都千代田区紀尾井町 7-1

E-mail: arai@sophia.ac.jp

あらまし 音声科学を中心とする分野において、声道模型の音響教育教材としての有効性が確認されている。千葉・梶山による声道形状に基づく Arai の筒型声道模型は、その形状が固定されているがゆえに瞬時にデモンストレーションができるという手軽さが特徴のひとつである。ところで、声道形状を単純にしても、同じ韻質を持つ母音を生成することは可能である。本研究では、異なる母音において声道形状の何が韻質の違いを生むポイントなのか、その本質を教えるための教材として、声道形状がさらに単純化された模型を開発した。その結果、Fant の 3 音響管や Arai のスライド式 3 音響管のように、断面積が異なる任意の長さの 1 様音響管をいくつか接続するような「円筒管接続式」の声道模型によって、日本語 5 母音が実現された。

キーワード 音声科学, 声道模型, 音響教育, 母音, 音響管

Education in Speech-Related Acoustics using Physical Models with Simplification of Vocal-Tract Shapes

Takayuki ARAI

Dept. of Information and Communication Sciences, Sophia University

7-1 Kioi-cho, Chiyoda-ku, Tokyo, 102-8554 Japan

E-mail: arai@sophia.ac.jp

Abstract In the speech-related field, vocal-tract models are confirmed to be effective tools for education in acoustics. Arai's cylinder-type vocal-tract models are based on Chiba and Kajiyama's measurement of vocal-tract shapes. One of their characteristics is that the models are effective for quickly demonstrating vowel production since their shapes are fixed for each vowel. The vocal-tract shapes are able to be simplified and yet the qualities of vowels can remain the same. In this study, we developed vocal-tract models with simplified vocal-tract shapes as educational tools to teach the essence what in vocal-tract shape produces the difference among different vowels. As a result, the five Japanese vowels were achieved by "tube-connected models," where several uniform tubes with different cross-sectional areas and lengths were connected as Fant's three-tube models and Arai's sliding three-tube model.

Keywords Speech Science, Vocal-Tract Model, Education in Acoustics, Vowel, Acoustic Tube

1. はじめに

音声科学における音響教育の目的で、今までにいくつかの声道模型が提案されている。例えば、千葉・梶山 (1942) による声道形状の測定 [1] に基づく声道模型 (筒型モデルとプレート型モデル) [2,3], 肺の模型と頭部形状模型 [3,4], ゲル素材による軟らかい舌を伴う声道模型 [5], スライド式 3 音響管モデル [6] などである。これらの音響教育教材を用いると、母音生成の基

本的な考え方である音源フィルタ理論を実践的に教えることが可能であり、また、母音によって声道形状が変わる様子をデモンストレーションすることができる [3]。肺の模型と声道模型を連動させ、呼吸から発声までの一連のメカニズムを提示したり、声道模型を頭部模型に組み込むことで、実際の人間の声道形状をイメージしやすくすることも可能である [3,4]。さらに頭部形状模型では、ゲル素材の軟らかい舌を操作して調音

位置を変化させられるもの[5]や、固定された声道形状に可動式の軟口蓋をつけて鼻音化させることができるもの[3,4]、軟口蓋から咽頭壁にかけて軟らかいゲル素材で構成し、鼻咽腔閉鎖パターンを模擬しながら口蓋裂患者の開鼻声を考察できるもの[7]などがあり、目的に応じて様々な使い分けをすることが可能である。音声科学を中心とする分野において、これらの物理模型を使用することによる教育上の有効性が確認されている（例えば、文献[3]参照）。

その中でも、母音の韻質を決定するフィルタ（声道形状）の特徴を端的に示せるのが、筒型声道模型である。回転体の形状をした筒状の模型は、千葉・梶山（1942）[1]によって測定ならびに直線近似された日本語5母音の声道形状に基づいており、母音ごとに独立した模型となっている。音源を入力するだけで簡単に母音生成のデモンストレーションを行うことができ、音声を聴きながら声道形状も観察することで直感的に理論を体得できるという点で優れている。Fig. 1 (a)に、初代の筒型声道模型を示す（写真は母音の/a/で、透明なアクリル製）。Fig. 1 (b)は、肉厚が一定（写真の場合、厚さ 2.5 mm）の声道模型で、内部の声道形状はFig. 1 (a)の筒型声道模型と同じである（写真は同じく母音の/a/で、ケミカルウッド製）。

一方、スライド式3音響管モデル[6]は、外筒の中で狭めを作るための内筒をスライドさせるだけで、異なる母音を作り出すことができる。その概略図を、Fig. 2に示す。構造が単純であるがゆえに共鳴周波数の理論近似も容易に行うことができる[8]。模型全体を3つのパーツに分解し、1/4波長共鳴器、1/2波長共鳴器、ヘルムホルツ共鳴器のそれぞれの共鳴周波数を計算することで模型全体の共鳴周波数を推定することが可能である。そのようなことから、一般を対象に音源フィルタ理論や狭めの位置や狭めの程度と母音の韻質を示すのにも有効であるのに加え、大学生以上を対象に母音生成に関する音響理論の説明をする際にも適している。同時に、その構造のシンプルさゆえに子どもを対象とした科学教室等で「声道模型の製作実習」として取り入れることも可能である（科学教室などでの工作の実例については文献[9,10]などを参照のこと）。

本研究では、異なる母音において声道形状の何が韻質の違いを生むポイントであるのか、その本質を教えるための教材として、模型の声道形状をより単純化することを目指した。筒型声道模型[2,3]の特徴である「形状が固定で瞬時にデモンストレーションができる」という手軽さと、スライド式3音響管モデルのように「音響理論の解析も容易であるシンプルな形状」を合わせ持つ声道模型を実現することを試みた。



Fig. 1 筒型声道模型. いずれも母音の/a/(下が喉頭側, 上が口唇側). (a) 円柱の素材を中心軸に沿ってくり抜いた透明アクリル製（初代のモデル）, (b) 肉厚が一定（厚さ 2.5 mm）のケミカルウッド製.

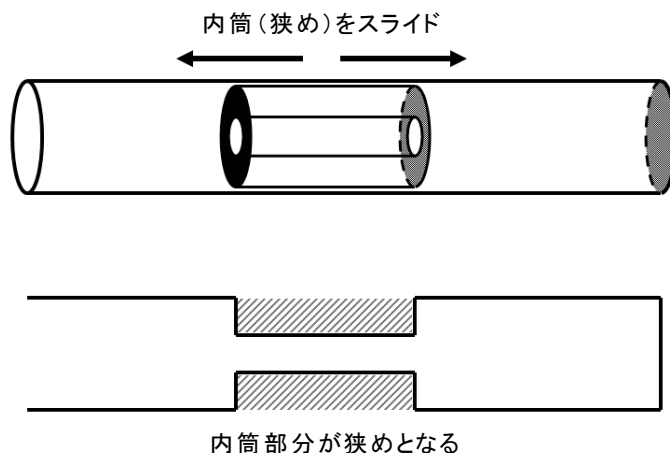


Fig. 2 スライド式3音響管の概略図（文献[6]より改変）. 長い外筒の中に挿入した内筒（狭め）を左右にスライドすることによって音響管の形状が変わる。音響管の形状によって共鳴周波数が変わるため、様々な母音を生成することが可能である。

2. 円筒管接続式声道模型

今までの筒型声道模型[2,3]は、千葉・梶山（1942）[1]に基づく声道形状がベースになっている。千葉・梶山は、X線などによって測定した声道形状を基に、その測定結果に対し母音の音響特性を損なうことなく

直線近似を行っている．筒型声道模型[2,3]では，直線近似後の形状を物理模型によって実現している．

一方，プレート型声道模型[2,3]は，その直線近似をさらに幅 10 mm で 0 次近似することで，階段状に声道形状を近似した．このような階段近似によっても，母音の高い韻質は確保されることを確認している[2]．



(a)



(b)

Fig. 3 円筒管接続式声道模型．左から母音の /a/, /i/, /u/, /e/, /o/ (下が喉頭側，上が口唇側)．(a) 透明なアクリル製，(b) ケミカルウッド製．

しかし，さらに単純な形状で声道をモデル化できることが知られている（例えば，文献[11]を参照）．Fant による 3 音響管モデル[11]によれば，断面積の異なる任意の長さの 3 つの一樣音響管を接続することによって，様々な母音を近似している．このようなシンプルな形状をベースに作られたのが，スライド式 3 音響管である．本稿では 3 音響管を参考に，断面積が異なる任意の長さの一樣音響管をいくつか接続するような「円筒管接続式」の声道模型を考えた．Fig. 3 にその新しい声道模型を示す．

音響教育の教材であることから，音質には妥協したくない．「ア」らしさ，「イ」らしさ，といった生成音のクオリティ（母音の韻質）を確保しつつ，筒型声道模型の形状をどこまで単純化させることができるかが課題であった．スライド式 3 音響管モデルによる母音生成のシミュレーション実験をベースに，千葉・梶山（1942）[1]によって測定・直線近似された日本語 5 母音の声道形状を参考にしながら韻質の決定に影響を及ぼすと思われる「喉頭の狭め」や「口唇の丸め」などを取り入れた模型を実現した．

Fig. 4 に円筒管接続式声道模型で生成された母音のサウンドスペクトログラムを示す．録音に際しては，PCM レコーダ（ソニー，PCM-D1）と内蔵マイクロホンを用い，マイクロホンから約 20 cm 離れたところに声道模型の開口端がくるように配置した上，標本化周波数 48 kHz，16 ビット量子化でデジタル録音した．この図を見ると，各母音のフォルマントがはっきりと現われている様子が分かる．この図では，音源として電気喉頭（セコム，EL-X0010）を使用したがる，もともとの音源に含まれる低域側の成分が多少弱いために，このスペクトログラムでは特に第 1 フォルマント（F1）周波数が低い /i/ や /u/ で F1 のピークが薄くなっている．各母音について，音声分析ソフトウェア Wavesurfer を用いて，スペクトル分析を行い，第 1・第 2 フォルマント（F1，F2）周波数を求めた．その結果を，Table 1 に示す．この表には，従来の筒型声道模型によって生成された母音のフォルマント周波数[2]，ならびに男性日本語母語話者による平均実測値[12]を併記する．

結果を見ると，円筒管接続式声道模型による F1・F2 周波数は，各母音とも平均値に近いことが分かる．特に母音 /o/ については，F1，F2 共に，筒型声道模型に比べて円筒管接続式声道模型の方がより平均値に近付いている．聴覚印象としても，母音 /o/ の明瞭性が改善していることが確認された．/u/ に関しては，円筒管接続式声道模型に口唇の丸めを設けたことにより，筒型声道模型よりも F1，F2 が下がり，生成された音声は英語に近い円唇母音になっていることが裏付けられた．筒型声道模型の F2 の方がより平均値に近い値を示し

ている点に関しては、日本語の/u/が非円唇母音であることが関係していると思われる。/a/に関しては、筒型声道模型は F1 が多少低くやや/o/に近い響きがあったのに対し、円筒管接続式声道模型では改善が見られた。/e/に関しては筒型声道模型と円筒管接続式声道模型との間でほとんど韻質の違いがみられなかった。

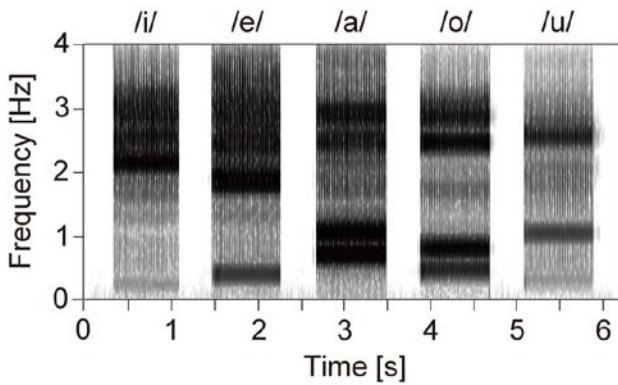


Fig. 4 円筒管接続式声道模型によって生成された母音のサウンドスペクトログラム. 左から/i/, /e/, /a/, /o/, /u/.

Table 1 円筒管接続式声道模型による第1・第2フォルマント周波数. 比較のため、筒型声道模型[2]と男性日本語母語話者による平均実測値[10]を併記する.

	円筒管 接続式 声道模型		筒型 声道模型 [2]		男性話者 平均実測値 [10]	
	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]
/i/	239	2157	230	1992	263	2263
/e/	398	1835	405	1868	475	1738
/a/	717	1035	654	1121	775	1163
/o/	481	800	436	1054	550	838
/u/	318	1039	499	1276	363	1300

3. 考察

新しく提案された円筒管接続式声道模型は、従来の筒型声道模型[2,3]同様に「瞬時にデモンストレーションが可能」といった特徴を備えている。また、スライド式3音響管モデル同様の「形状のシンプルさ」も合わせ持ち、簡単な音響理論の近似によって共鳴周波数を推定することも可能である。形状の上でスライド式3音響管モデルとの違いは、1) いずれの母音も喉頭側が細くなっている、2) 母音の/u/と/o/については口唇の丸めを設けている、という2点である。

喉頭側の狭めと口唇側の丸めの有無による韻質への影響を、スライド式3音響管と円筒管接続式声道模型を比較することで調べてみた。その結果、喉頭側の狭めがある場合の方が音声はより明瞭になる場合があり、母音の/i/においてその影響が最も顕著であった。口唇の丸めについては、母音/u/と/o/に関して比較を行った結果、口唇の丸めがある場合の方がより明瞭であり、/o/においてその影響がより顕著であった。

筒型声道模型[2,3]は、円柱の亚克力素材を中心軸に沿ってくり抜いているため、声道を模擬するくり抜かれた空間の形状を外から見えるようにするためには、透明な素材で作る必要があった。Fig. 3 (a) に示す円筒管接続式声道模型は、筒型声道模型[2,3]と同様、透明な亚克力素材で出来ている。しかし、一定の厚み (Fig. 3 の場合、2 mm) の素材で出来ているため、内部の声道形状はそのまま外側の形状に反映される (この点は、Fig. 1 (b) と同じ)。つまり、円筒管接続式声道模型の場合、必ずしも透明な素材で作らなくても、その形状が一目でわかる、言わば外形を重視した物理模型であると言える。そのため、素材として透明アクリルの他、色の付いた亚克力素材、木やケミカルウッドなどでも製作を試みた (Fig. 3 (b) はケミカルウッド製)。また、その形状が外側に反映されていることから、手で持って触ることでその形状を感じる事ができ、デモンストレーションや展示等でもさらに効果を発揮すると同時に、視覚に障害を持った方々などにも音を聞いてもらいながら同時に「触って」形状を感じてもらえるようになる。

さて、本研究では、異なる母音において声道形状の何が韻質の違いを生むポイントであるのか、その本質を教えるための教材として、模型の声道形状をより単純化することを目指した。円筒管接続式声道模型は、そのひとつの答えであると言える。音声科学の分野において、例えば専門的に学ぶ学生を対象に音響理論を教授する場合や、小中学生や一般対象のサイエンス講座等での利用が考えられる。MRI などを使って詳細な声道形状を計測して作られた声道模型[13]、筒型声道模型、円筒管接続式声道模型など、それぞれの声道模

型に同じ音源を入力して生成音声を聴き比べたり、それを導入に用いて各対象に合わせた内容に発展させていくなどの方法が考えられる。「音声」という身近なテーマであるがゆえに、少なからず興味をもっていただけるであろうし、難解な理論の解説を伴わなくても目の前で起こる現象には不思議な説得力がある。コンピュータ・シミュレーションと異なり、いずれも触って感じることも可能であるため、視覚障害のある方も「手」で見て（触覚認知）、耳で聴くことができる。教育効果を考えても、教科書上の理解と経験に基づく知識ではやはり差があるようで、そのような点で物理模型の価値を実感している。まさに「百聞は一見にしかず」である。

また、音響教育の「教材・教具」として考えるとき、「何を教えたいか」という目的によって、用いる教材の選択基準は変わり、また、たとえ同じ教材であったとしても効果的な使い方が変わってくる。教材開発において、ある現象の中の本質的な部分を残して他の部分をそぎ落としていくという作業は、言い換えれば「単純化」と表現しても良いかもしれない。しかし、一口に「単純化」と言っても、音響理論を分かりやすく説明できるもの、人間の機能を模擬できるもの、容易に製作できて実際に操作することが可能なもの等、本質でもどこに注目するか、どのように使いたいのかで違ってくる。

例えば、母音の韻質を決定するフィルタの形状（声道形状）の本質的な部分を究めていくと、直線状のシンプルなモデルにたどり着く。「声道の狭めの位置や程度」という観点で考えていくと、本稿で紹介した筒型や円筒管接続式の声道模型、スライド式3音響管モデルが思い浮かぶ。さらに、スライド式3音響管モデルには一つで様々な母音が生成できるという利点もあり、「音声のダイナミクス」まで考える場合にはスライド式3音響管が勝っていることになる。

しかし、実際の人間の発声・調音機構を教えたい場合、それが一番シンプルで適切な教材かというところとも言えない。「舌の動きを含めた単純化」「ダイナミクスまでを考えた単純化」を考えていくと、「真っ直ぐな音響管」より「曲がった音響管」の方がふさわしいからである。言語音の生成には、人間だけが持っている音声器官構造（つまり、声道が直角に近い形状で曲がっている、かつ喉頭が下がって口腔にも咽頭腔にもある程度の広がりがある、舌に厚みがあり上下・前後に動くなど）が欠かせない。それは、霊長類の中でも人類特有のものと言われている。例えば、「/u/では声道の中央付近に狭めが存在する」という事実は同じであっても、直線状に動くスライド式音響管モデルの「内筒」と、直角の管を上下・前後に動く「舌」では

全く違ったイメージになってしまうのは言わずもがなであろう。/a/ /i/と連続して発話する場合も、スライド式3音響管では「内筒（狭め）」を大きく移動させることになるが、直角の管と「舌」の場合はそこまで距離感を持った動きにはならない。単純な形状ではあるが「直角」という要素が入っているものでないと、微妙な舌の動きで様々な調音をしている様子をより実態に即した形で模擬することはできない。

単純化とは何か、どんな本質を教えたいのか。譲れる部分はどこで、譲れない部分はどこなのか。目的によって教材・教具は変わるし、同じ教材・教具でも効果的な使い方が変わってくる。

4. おわりに

声道形状を単純化した円筒管接続式声道模型を提案した。この声道模型は、音声科学における音響教育教材として、従来の筒型声道模型と同じように使うこともできる。しかし、その形状がより単純化されていることから、形状の中でも狭めの位置や狭さなど、声道形状が持つ「韻質」を決定付けるのに重要な「本質」を理解することに今まで以上に貢献することが期待できる。

5. 謝辞

内容の一部は日本学術振興会の科学研究費補助金（19500758）、及び文部科学省私立大学学術研究高度化推進事業上智大学オープン・リサーチ・センター「人間情報科学研究プロジェクト」の助成を得た。

文 献

- [1] T. Chiba and M. Kajiyama, *The Vowel: Its Nature and Structure*, Tokyo-Kaiseikan, Tokyo, 1942.
- [2] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," 音声研究, vol.5, no.2, pp.31-38, 2001.
- [3] T. Arai, "Education system in acoustics of speech production using physical models of the human vocal tract," *Acoustical Science and Technology*, vol.28, no.3, pp.190-201, 2007.
- [4] T. Arai, "Lung model and head-shaped model with visible vocal tract as educational tools in acoustics," *Acoustical Science and Technology*, vol.27, no.2, pp.111-113, 2006.
- [5] T. Arai, "Gel-type tongue for a physical model of the human vocal tract as an educational tool in acoustics of speech production," *Acoustical Science and Technology*, vol.29, no.2, pp.188-190, 2008.
- [6] T. Arai, "Sliding three-tube model as a simple educational tool for vowel production," *Acoustical Science and Technology*, vol.27, no.6, pp.384-388, 2006.
- [7] 荒井隆行, 田中希美, 片岡竜太, "軟らかい素材による軟口蓋を伴う声道の物理模型," 信学技報, vol.SP2008-103, pp.143-148, 2008.

- [8] K. N. Stevens, *Acoustic Phonetics*, MIT Press, Cambridge, MA, 1998.
 - [9] 荒井隆行, “音に関する子ども向け科学教室の実例,” 日本音響学会春季研究発表会講演論文集, pp.1483-1486, 2008.
 - [10] T. Arai, “Science workshop with sliding vocal-tract model,” *Proc. of Interspeech*, pp.2827-2830, 2008.
 - [11] G. Fant, *Acoustic Theory of Speech Production*, Mouton, The Hague, Netherlands, 1960.
 - [12] 粕谷英樹, 鈴木久喜, 城戸健一, “年令, 性別による日本語5母音のピッチ周波数とホルマント周波数の変化,” 日本音響学会誌, vol.24, no.6, pp.355-364, 1968.
 - [13] K. Honda, H. Takemoto, T. Kitamura, S. Fujita and S. Takano, “Exploring human speech production mechanisms by MRI,” *IEICE Trans. on Information and Systems*, vol.E87-D, no.5, pp.1050-1058, 2004.
-