

## 音声生成を直感的に理解するための声道模型とその教育応用\*

荒井 隆行<sup>\*1</sup>

【要旨】 今まで開発してきた声道模型とその音響教育への応用について述べた。まず代表的なものについて、模型の形状が屈曲しているか真っすぐかにより大きく分類し、動きの自由度によって各声道模型の特徴を議論した。そして声道模型を用いた講義の紹介と、合わせて行った試験やアンケート調査の結果を基にその教育効果について再評価を行った。その過程で模型を用いた音響教育の在り方について考察した結果、「音」そのものを教える音響教育に有効であるのはもちろん、音声教育にも有効であること、そして目的に応じて模型を使い分けることの重要性、更に多様な特性を持った模型を今後も開発していく必要性を確認した。

キーワード 声道模型, 音声生成, 音響教育, 音声教育, 音響音声学

Vocal-tract model, Speech production, Education in acoustics, Phonetic education, Acoustic Phonetics

## 1. はじめに

近代の音声科学, 音響音声学の分野において, 母音生成の機構に関する論争に終止符を打ったのは Chiba & Kajiyama の The Vowel [1] であった [2]。この著書が現代の音声科学の基礎となり, その後, Fant [3] や Stevens [4] などが中心となって母音のみならず, 子音を含む音声科学全体の体系が作られ, この分野は急速に発展した。その原点となった The Vowel [1] では, 世界に先駆けて母音生成時の声道形状に対する 3 次元計測も行われており, Artificial Vowels という節においてその 3 次元計測を更に簡素化した形状を持つ筒を物理的に実現している。更にその筒 (つまり「声道模型」) の喉頭側の端からある種の音源を入力することによって, 母音のような音が生成されることが示されている。このような機構で母音が生成されることは, 古くから von Kempelen (1791) [5, 6] などによっても機械仕掛けの模型により示されてきたが, Chiba & Kajiyama は更に声道の 3 次元形状と母音のスペクトル特性などにも言及し, 母音の質の違いはその形状の違いによることも証明している。その後, Fant [3] によってその原理は音源フィルタ理論 (source-filter theory) として定式化され, 更に Stevens はあらゆる音声は四つの音源と共鳴の組み合わせでモデル化されると述べた [7]。ここで, その四つの音源とは声帯振動源 (喉頭原音),

帯気雑音源, 過渡音源, 摩擦音源であり, この考え方に基づいて Klatt はフォルマント合成器を開発している [8]。

また, 声道模型を用いた研究も試みられている。梅田・寺西 (1966) では, 側面から角棒を抜き差しすることで形状を階段近似できるような声道模型を用いて, 声の韻質と声質の関係について調べている [9]。最近では, 本多らが MRI による精密な声道の 3 次元形状を基にその物理模型を製作している [10]。また, 誉田らは, 一連の発話ロボットを開発している (例えば, [11])。研究者によってはこれらの声道模型が教育上, 有効だと実感することもあったであろうと推察するが, 科学教育という観点で「声道模型を用いた音響教育」を研究対象とすること自身, ほとんどされてこなかったと言ってよいだろう。

一方, 著者らは The Vowel に登場する声道模型の復元 [12] をきっかけに, 多様な音響教育応用に特化した改良を重ね, 幾度もの試行を繰り返してきた [13-16]。その結果として開発された代表的なものを幾つか, 図-1 に示す。この図にあげられた声道模型を見ると, 一概に「声道模型」と言えども, 個々に様々な特徴を持ち合わせていることが分かる。そして注意深く考察していくと, それらをどう用いたらより教育効果をあげることができるか, あるいはいまだ足りない側面がどこにあるかなどが浮かび上がってくる。そこで本論文ではまず, これらの声道模型の特徴を改めて紹介し, それらを用いることによる教育効果についての議論を再検討する。この中で, 今後更にある特定の方向性を持って目的別に開発が集約されていく可能性があることについても触れる。

\* Vocal-tract models and their educational applications for intuitive understanding of speech production,

by Takayuki Arai.

<sup>\*1</sup> 上智大学理工学部  
(問合先: 荒井隆行 〒102-8554 千代田区紀尾井町 7-1  
上智大学理工学部 e-mail: arai@sophia.ac.jp)

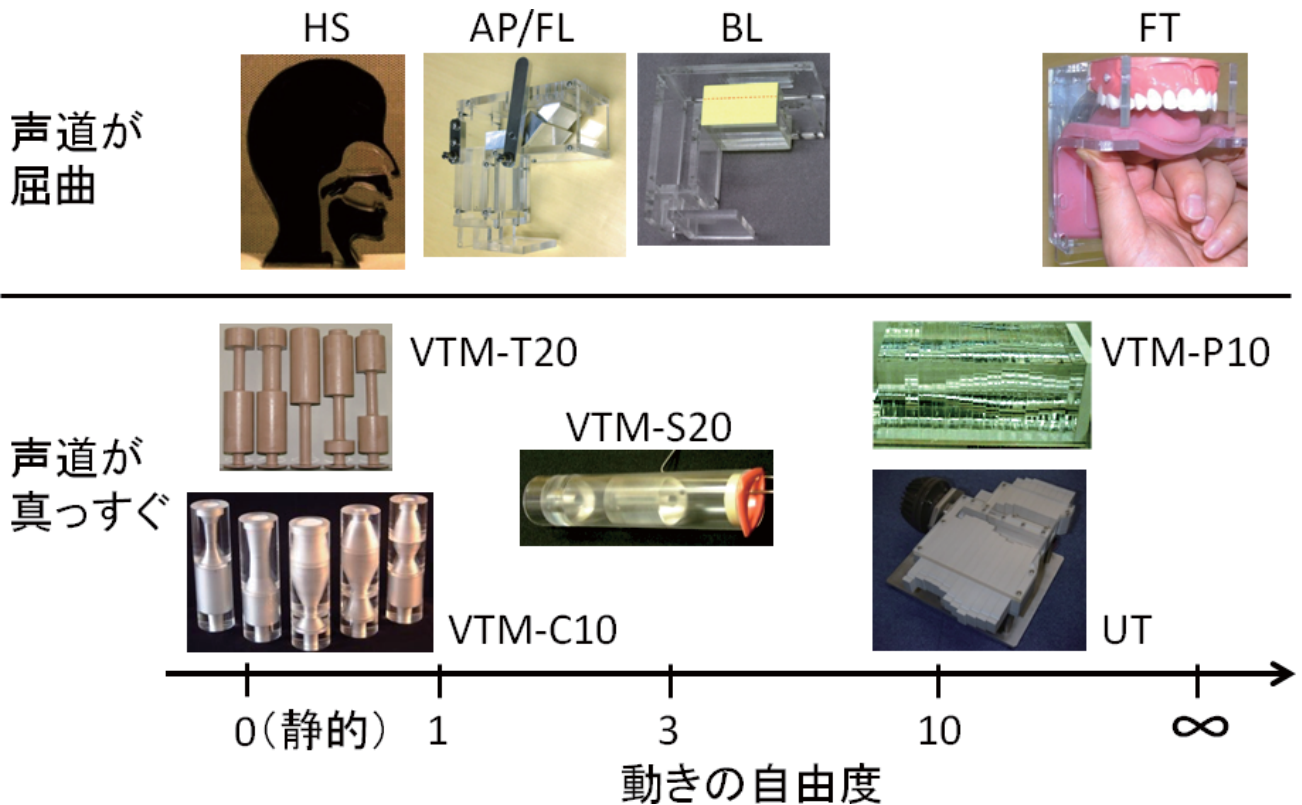


図-1 今までに開発した代表的な声道模型の例

## 2. 声道模型

図-1 に 9 種類の声道模型をあげた。まず上下段に分類し、下段に声道形状が真っすぐなもの、上段に声道形状が屈曲しているものが並べられている。横軸は各模型における「動きの自由度」を表しており、全く動かない静的なもの（自由度は 0）から、低い自由度で操作できるもの、そして高い自由度を持つものまでが存在する。このように 9 種類の模型を配置させた上で、以下ではそれぞれの声道模型の特徴について述べる。

### 2.1 声道が真っすぐな模型

図-1 の下段にあるのが、声道が真っすぐな模型群である。以下では下段の 5 種の模型について概説する。

#### 2.1.1 VTM-C10

Chiba & Kajiyama による The Vowel [1] にある日本語 5 母音に関する形状を基に復元したのが、VTM-C10 と書かれた筒型の声道模型である [12, 14]。この模型では、他の模型同様に、音源フィルタ理論や母音の質がその形状によって変わることが直感的に示すことが可能である。人工喉頭などの音源を喉頭側の穴から入力することによって、瞬時にデモンストレーションをすることが可能であるため、各種講義や講演、科学教室などに幅広く応用可能である。また、静岡科学館や日立シビックセンター科学館において、著者が監修した声のしくみに関する展示では、この VTM-C10 が用いら

れている。なお、アメリカ San Francisco にある科学館 Exploratorium に存在する声道模型の展示は、University of California, Berkeley の John Ohala 教授が監修されたもので、その形状は Chiba & Kajiyama [1] を参考にしたとされるものの、形状は文献 [1] のデータを忠実に復元したわけではなく、かなりの変形が施されている。

#### 2.1.2 VTM-T20

VTM-C10 の声道形状は、Chiba & Kajiyama が計測した声道の 3 次元データを、更に本人たちがシンプルにした形状を忠実に再現している。しかし、その形状が折れ線近似された断面積関数を基本とする回転体となっているため、模型の製作工程が複雑で製作費も必然的に高くなる傾向にあった。そこで、VTM-C10 の形状を更にシンプルにしつつたどりついたものが VTM-T20 である [16]。この模型では、基本的に外筒の直径や狭窄の内径など、揃えられるところは極力揃え、更にすべてが長さや直径だけが違う円筒管の接続で構成された設計となっている。そのため、製作費用も低く抑えることに成功した。更に、母音 /o/ についてはその母音の質も向上している。

また、VTM-C10 では筒の中の形状を外からも見えるようにするために透明素材（例えば透明アクリルなど）を用いる必要であったが、VTM-T20 では声道形状が一目瞭然であるため、必ずしも透明素材にこだわ

る必要がない。そのため、素材はアクリル以外に、プラスチック、金属（アルミなど）や木材、紙筒など、あらゆるもので実現可能である。更には、視覚に障害がある方々が触れることで形状を感じ取ることも可能となっている。このVTM-T20は、教具としての使い勝手がほぼVTM-C10と同じであること、そして母音の質は少しばかり改善されていることなどを総合的に考え合わせると、VTM-T20のほうがVTM-C10よりも普及には適していると言える。

### 2.1.3 VTM-S20

VTM-T20の声道形状は、見方を変えると長い外筒に小さい穴の開いた狭窄部をスライドさせることで近似できることが分かる。このような考え方を基に設計されたのがVTM-S20である[13,16]。VTM-C10やVTM-T20が静的なモデルであったのに対し、このVTM-S20は狭窄部がスライドするという自由度を持つ（第1自由度）。更に、狭窄部の穴の大きさ（狭窄の程度）を変えることによって、生成できる母音の幅が広がる（第2自由度）。更に、口唇の狭め（第3自由度）を変えられるようにすることで、結果的に「自由度が最大3」という少ない自由度でかなり広範囲な母音空間をカバーすることができる。素早いデモンストレーションを要求される場合は、口唇の狭めは使用せずに必要に応じて利用者の親指と人さし指を使って口唇部を模擬するなどの対応を取ることで代用できる。また、狭窄の程度が自由に換えられる可変式の狭窄部（文献[16]参照）が身近にない場合は、直径の異なる円柱で作られた狭窄部を複数、かわるがわるに使うことになる。例えば、外筒の内径が34mmの場合、狭窄部用の円柱として直径が32mmのものと27.5mmのもの二つを準備しておくだけでも、（狭窄部を最適な位置にさえ配置すれば）前者で母音の/i/, /a/, /o/, /u/を、後方で/e/を生成することができる。このように、口唇の狭めは設けず、狭窄の程度も数段階に限定することによって多少生成できる母音の種類に制限が加わったとしても、5母音を外筒1本と狭窄部2種だけでかなり明瞭に素早く出力できるという点で優れていると言える。

このようにVTM-S20ではVTM-T20に近い声道形状が実現できるが、両者にはそれぞれの特徴があるため、どちらをどのように使用するか、あるいは両者をどのように組み合わせるかについては、実際のデモンストレーションに合わせて各モデルの特徴を生かした使い分けをすることが重要である。例えば、中性母音である「シュワ」をまず生成してから5母音を生成するステップに移行するというのであれば、VTM-S20において狭窄部を挿入しない外筒だけで声を作るとこ

ろからスタートすればよい。その際、音源だけの音と外筒に音源が入力されたときの出力音を比較するとなおよい。また、VTM-T20はそれぞれの母音に対して別のモデルが必要であるが、VTM-S20の場合には一つのモデルで十分である。しかし、VTM-S20の場合はどこに狭窄部がくるとどの母音になるかを、教える側が事前に練習をしておくなどの準備が必要である。しかし簡単な機構や少ない仕掛けで複数の母音が容易に生成されることから、VTM-T20のように母音ごとに模型を用意してはかさばる、というような場合にVTM-S20は重宝する。また、更にVTM-S20の場合は、母音生成における動的な音の変化を付けることも可能となる。その結果、聞こえ上の「オーバシュート」効果によって、より明瞭に聞こえる母音を提供できるようになる。つまり、もし仮に個々の母音を単独で生成した場合、聴取者全員が同じ母音として認識しないようなケースであったとしても、/a/ → /i/ → /u/などと母音を連続して生成することによって母音の識別率が上がることがしばしばある。これは人間の生成機構を見ても実際に同様の現象が起きており、仮にターゲット音の理想的な調音器官の構えに到達する前に別の構えに遷移したとしても（ターゲット・アンダシュート）、聴覚機構がそれを補ってターゲット音を聴取している[17]。そのような音声のダイナミクスに潜む現象をも、このような動的な模型を使ってデモンストレーションすることが可能である。

### 2.1.4 VTM-P10

この模型では、直径が異なる穴が中央に開けられた厚み10mmのプレートが複数準備されており、それを利用者が好きに組み合わせながら並べることによって、任意の声道形状を階段近似することができるように工夫されている[12,14]。10枚以上のプレートそれぞれに開けられた穴の直径を変えられるという自由度の高さの結果、かなりのバリエーションの母音を実現することが可能である。それに加えて、子どもがパズル感覚で能動的に活動に参加できるという利点がある。しかし、自由度があり過ぎて適当に組み合わせってしまう可能性もあるため、見本となる形状を準備しておくことが望ましい。なお、先述の静岡科学館の声のしくみに関する展示では、このVTM-P10がVTM-C10と共に体験型展示として併設されている。

### 2.1.5 UTモデル（梅田・寺西型モデル）

このモデルは梅田・寺西[9]による声道模型で、前述のとおり、側面から複数の角棒を抜き差しすることによって任意の声道形状を実現するものである。VTM-P10同様、自由度が高く、様々な母音を実現できる半面、目的の母音を達成しようとする、訓練を要する。

また、一つの角棒を動かすと隣接する棒も一緒に動いてしまうことがあり、精密な操作を手動で行おうとすると難しい。そこで、Arai [18] ではこの UT モデルの各棒にアクチュエータを接続し、PC 制御によって声道形状を実時間で変える試みも行われている。

## 2.2 声道が屈曲した模型

図-1 の上段にあるのが、声道が屈曲した模型群である。以下では上段の 4 種の模型について概説する。

### 2.2.1 HS モデル (頭部形状モデル)

このモデルは、外形が頭の形をしている上、声道形状も横から見ると丁度教科書によく出てくる正中矢状断面図になるように工夫された模型である。更に図-1 に示すように、口腔に加えて鼻腔も備わっており、鼻咽腔の結合部での気流の切り替えがツمامミの回転によって実現できるように工夫されている。その結果、鼻咽腔閉鎖が実現されているときは母音の/a/が、ツمامミを回転し鼻腔にも気流が流れるようになると鼻音化母音が生成される。このような模型が開発された背景には、図-1 の下段にあるような形状が真っすぐな声道模型だけを使って説明を行っている際に、「真っすぐな声道模型が実際はどのように頭の中に配置されているかがイメージしにくい」といった質問を受けたことが一つのきっかけとしてあった。この HS モデルは、そのような疑問に答えるためにも効果的な模型と言える。実際、著者が行う講義の中では、可能な限り必ずと言ってよいほど、真っすぐな声道模型に加えて、この HS モデルを併用するようにしている。更に、肺の模型 [14] とこの HS モデルを一緒に組み合わせることによって、呼吸や発声から調音まで母音生成を系統的に説明することができるようになり、また発せられる母音が人間らしいことも手伝ってデモンストレーションとしても見栄えが良い。その結果、学習者に深い印象を与え、理解を促進する効果につながることが多い。ただし、この HS モデルでは基本的に異なる母音には異なる模型セットが必要になる。

### 2.2.2 BL モデル (ブロック式モデル)

HS モデルでは、母音ごとに異なる模型が必要になっていた。そこで、一つのモデルで複数の母音が生成可能な屈曲型の声道模型を設計することを試みた。その結果として生まれたのが、この BL モデル (ブロック式モデル) である [19]。このモデルでは、声道が中央付近で直角に屈曲し、かつ断面が長方形をした空間にブロックを入れて、そのブロックの位置を変えることによって声道形状を変える機構になっている。ブロックは、前舌ブロック、後舌ブロック、そして口唇ブロックを用意しておくことによって幅広い母音を再現することが可能となる。今までの多くの模型同様、外壁に

透明アクリルを使用することによって、中のブロックの位置などが視覚的にも確認できる。本モデルの周りに頭部形状を形取ったウレタン素材の「頭」をかぶせることによって、HS モデルのような効果 (頭部のどこに声道が配置されているかを示すなど) も得られる。

### 2.2.3 FT モデル (軟らかい舌を伴うモデル)

柔軟な舌の動きを追及した結果、たどりついたのがこの FT モデルであり、このモデルでは舌が軟らかい素材でできている [15]。舌の形状は下側から手で変えることができ、手で操作しながら微妙な位置の調整をしつつ、耳で生成される音を確認でき、また視覚的にもその形状を観察することが可能になっている。舌の素材が軟らかいことから、声道形状は実質的に無限の自由度で変形が可能であり、産み出される音の種類も豊富である。その一方で、目的の音を実現するにはそれなりの訓練が必要なことも事実である。ただし、訓練によっては母音のみならず子音も生成可能であるなど、その応用され得る可能性の幅も広い。

### 2.2.4 AP モデル (接近音モデル)

FT モデルを除き、これまでの模型のほとんどは母音を生成するためのものであった。しかし、子音に対するニーズも高い。特に英語の/r/や/l/などの接近音については、母音用の模型から発展させ易く、FT モデルでも接近音らしい音を作り出すことが可能である。しかし、FT モデルではあまりにも自由度が高いために目的の音を出すためにはそれなりの訓練が必要であった。そこで、出せる音が限られはするものの、少ない自由度で接近音を生成できるように設計された模型が、この AP モデルである [20]。特にこのモデルでは、英語の/r/として実現されることの多いそり舌接近音と、/l/の歯茎側面接近音を生成することに主眼が置かれ設計されている。/r/と/l/に関しては以前、文献 [14] でも静的な模型を製作していたが、静的な模型であるがゆえに生成音も定常的な音しか生成されず、なかなか明瞭な子音として聞こえないという問題もあった。そこで、この AP モデルでは母音 → 子音 → 母音という一連の動作を、舌の前方部が口蓋に向かって折れ曲がりまた元に戻るといった簡素化された動作で実現し、その動的な変化をレバーの回転という 1 自由度でのみ実現することを試みた。この AP モデルのほか、同じような発想で「はじき音」専用として開発された FL モデルもある [20]。

## 3. 声道模型による教育効果に関する評価

著者はこれらの声道模型を組み合わせることによって、様々なスタイルの授業や講演、科学教室などを行ってきた。対象も子どもから大学院生までと様々である。

また、今では我々が開発した声道模型を実際に複数の教育機関において使ってもらっている。本章ではそれらの授業等を通じて収集されたデータを基に、教育効果を定性的かつ定量的に見てみることにする。

### 3.1 音響学の授業の例

提案してきた声道模型の中でも、肺の模型、HS モデル（頭部形状モデル）、VTM-C10 や VTM-P10 などを用いた音声生成を含む音響学の講義を、言語聴覚士をめざす専門学校生を対象に実施した例について、以前、文献 [14] にて紹介した。この授業では、音響学の講義を行う中で筆記試験を行い、その結果について報告している。本論文ではその試験の結果を改めて検討し、更に統計的な分析に基づく教育効果について再評価を行う。

#### 3.1.1 実施した講義並びに試験について

受講者は言語聴覚士を養成する専門学校の学生 20 名で、全員が音響学の講義の既履修者であった。まず最初に前半の講義として、約 3 時間ほど音響学や音声生成を含む音声科学の基礎について一般的な授業形式による説明を行った（以後、講義 1）。その講義 1 では、スライドと言葉による説明をメインに、それに加えて、デモンストレーションも声道模型は一切使わず一般的な説明に出てくる程度の音声ファイルや動画ファイルの再生にとどめた（例えば、声帯が振動する様子を見せる動画の再生など）。その講義 1 の直後、第 1 回目の試験を行った。そして昼食後、後半の講義を行った（講義 2）。この講義 2 では、音声生成に関する授業を約 1.5 時間、声道模型によるデモンストレーションも交えながら実施した。その中の最後の 15 分には、VTM-P10（プレート型声道模型）を用いた実習の時間も含まれる。その講義 2 の直後、第 2 回目の試験を行った。

試験の内容については、第 1 回目と第 2 回目では全く同じものを用いた。設問は全部で 28 問あり、それらは講義 1 に関係するが講義 2 の内容には直接関係しない設問群（以後、G1 群）と、講義 2 に関係する設問群（以後、G2 群）という二つのグループに分けられた。そして G2 群は更に、講義 2 の中でも声道模型を使った解説部分とは間接的にしか関係しない設問群（以後、G2a 群）と、声道模型によるデモンストレーションと直接関連が深かったと思われる設問群（以後、G2b 群）、という二つの小グループに分けられた。各設問は短い文からできており、文の一部が穴埋め形式になっている。その穴に最も適当と思われる答えを選択肢から選ぶ形式で答えさせた。

#### 3.1.2 試験の大まかな結果

2 回行われた試験の結果について、文献 [14] でも述べているように、第 1 回目の試験の結果は正答率が

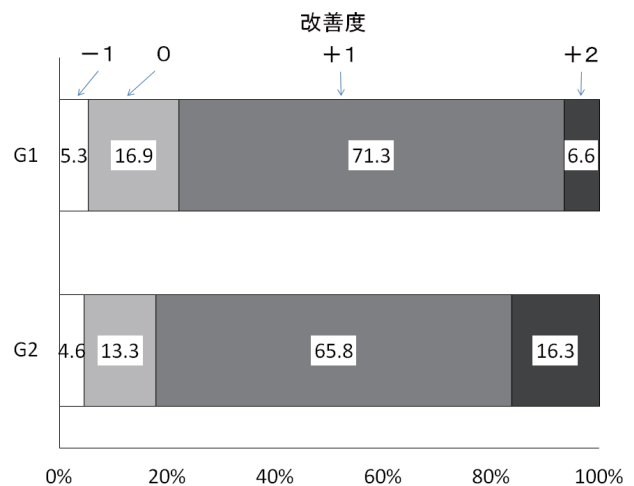


図-2 G1 群と G2 群における改善度の割合

73.9%であったのに対し、第 2 回目の試験は正答率が 79.6%と上昇していた。上昇した理由としては、同じ試験を 2 回実施しているために学習効果があったことが考えられる。しかし、仮に学習効果があったとしても G1 群と G2 群にて正答率の改善に差があるとすれば、講義 2 を実施したことの効果が確認できる。そこで、各群における改善度を評価することにする。

#### 3.1.3 改善度の定義

1 回目から 2 回目の試験の成績の変化を表す指標として、改善度を次のように定義した。1 回目が正解で 2 回目が不正解であった場合、理解度が悪化したとして改善度を  $-1$  とする。1 回目が不正解で 2 回目も不正解であった場合、理解度が不良ということで改善度を  $0$  とする。1 回目も 2 回目も正解であった場合、本来の改善度に変化はないが理解度は十分として、この指標での改善度を  $+1$  とする。1 回目は不正解であったが、2 回目で正解になった場合、著明な改善として改善度を  $+2$  とした。なお、この考え方は正解から不正解になった場合は講義が全く効果を示さず、不正解から不正解は講義が十分な効果を示さず、正解から正解は講義が更に理解を深める効果があり、不正解から正解は講義が十分に機能したと捉え、点数化したものである。ノンパラメトリックな扱いをする限り、点数の順序のみが重要であり、数値そのものは測定量としての意味はない。すなわち、 $+2$  は  $+1$  の 2 倍を表すものではない。

#### 3.1.4 G1 群と G2 群における改善度の比較

まず G1 群と G2 群の比較を行った。比較においては、G1 群と G2 群別に改善度を集計し、対応のない Mann-Whitney  $U$  test を実施した。検定の結果、G2 群は G1 群に比して、有意に高い改善度を示した ( $p = 0.004$ )。なお、図-2 は各群における改善度別の割合を示す。

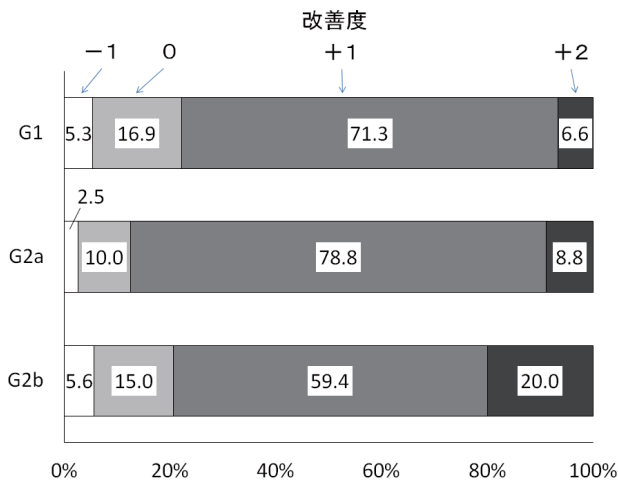


図-3 G1 群, G2a 群, G2b 群における改善度の割合

### 3.1.5 G1-G2a 群間と G1-G2b 群間との比較

次に G2 群を更に詳細に G2a 群と G2b 群に分けた上で、前節同様の検定を行った。対応のない Mann-Whitney *U* test を実施した結果、G2a 群は G1 群に比して、有意な傾向で高い改善度を示したのにとどまった ( $p = 0.054$ ) が、G2b 群は G1 群に比して、有意に高い改善度を示した ( $p = 0.011$ )。なお、図-3 は各群における改善度別の割合を示す。

## 3.2 音声学の授業の例

声道模型のうち、VTM-T20 と VTM-S20 を主に用いた講義をアメリカ Rutgers 大学にて実施した例が、文献 [16] に報告されている。そこで、この講義で実施したアンケート調査の結果について紹介する。

### 3.2.1 講義の内容

講義は前半に 1 回、後半に 2 回連続で行われ、20 名から 25 名の学生が参加した。前半 1 回の講義のテーマは調音で、声道模型を交えた説明が 30 分程度実施された。まず始めに VTM-S20 を使って狭窄部を使わずに外筒だけでシュワ音を生成した。そして狭窄部を挿入し、前舌母音と後舌母音を生成した。次に、VTM-T20 を使って、特に母音の /i/ と /e/ を比較することで舌の高さの違いによる音の変化を説明した。最後に、VTM-T20 の母音 /u/ と /o/ を用いて口唇の狭め (丸め) の説明を行った。

後半 2 回の講義のテーマは音響音声学であり、2 回のうち 1 回目は VTM-S20 の外筒だけで作られるシュワ音を対象に、片側が開いてもう片側が閉じた断面積が一樣な音響管において音響理論から求められる共鳴周波数と測定結果が一致するかどうかを分析した。次に VTM-T20 の母音 /a/ の形状が 2 音響管であることに着目し、やはり音響理論から求められる共鳴周波数と測定結果が一致するかどうかを分析した。後半 2 回の講義のうち 2 回目では、VTM-S20 で作られる母音

を対象に、3 音響管に関する音響理論から求められる共鳴周波数と測定結果が一致するかどうかを分析した。

### 3.2.2 アンケート結果

前半と後半のそれぞれの講義の後に、自由記述によるアンケート調査を実施した。それらをまとめたものを、以下に示す。なお、括弧の中の数字は類似する回答があった場合の件数を示す。

#### 【前半の講義後のアンケート結果】

- 声道模型は興味深く、学ぶのに有益で理解の助けになった (10)。
- 声道模型は、声道形状と母音の質の関係を理解する助けになった (3)。
- 声道模型は人間の母音らしく聞こえ、母音による違いもはっきりと分かった (5)。
- 声道模型は音の可視化を可能にし、音声生成に対するより深い理解の助けになった (6)。
- 声道模型は、声道における狭窄の位置を視覚的に確認することを助けてくれた (1)。
- 声道模型によるデモンストレーションによって、喉頭原音がどのようなものであるかということについての理解を助けてくれた (1)。

#### 【後半の講義後のアンケート結果】

- 声道模型は興味深く、学ぶのに有益で理解の助けになった (9)。
- 声道は通常、見ることはできないが、声道模型はそれを可能にしてくれた (4)。
- 声道模型は 3 次元であり、図で見るよりも効果的であった (1)。
- 声道模型は、声道形状がいかに音に影響するかを示してくれた (3)。
- 声道模型は、いかに音が作られるかの理解を助けてくれた (2)。
- 音声分析ソフトウェアを用いて生成される音を録音・分析し、理論式から得られる結果と一致したという点で、声道模型は有益であった (3)。
- 声道模型から生成される音は本物のようであり、またフォルマント周波数も測定でき、測定結果は本物の母音とほぼ一致した (1)。
- 声道模型は、共鳴器に対する可視化の助けになり、接続される管の大きさによって共鳴周波数が変化することを理解する助けにもなった (1)。
- 声道模型は、五感に訴え、そして抽象的な概念をより具体的なものにするという点で、楽しかった (1)。

## 3.3 言語聴覚士を目指す大学院生に対する授業実施例

声道模型の中でも接近音専用が開発された AP モデルを用いた授業を、音響音声学に関する大学院の講義の中で行った。その講義には言語聴覚士や語学教師を

志す者も半数程度含まれていたが、デモンストレーションを交えた説明後に行ったアンケート結果について文献 [21] にて報告されている。そこで、その結果について以下に紹介する。

#### 【良い点】

- 音を聞きながら体感でき、興味が湧き易い。
- 舌のどこをどう配置すれば良いかが分かる。
- 調音の違いを明確に理解できる。
- いろいろな角度から見られる。
- 舌の動きと音の変化の直接的な関係を実感することができる。

#### 【改善点】

- 舌全体を色付けするとより良いだろう。
- よりリアルな舌のほうが自分と比較し易く効果も高まるかも。
- どこに注目してデモンストレーションを見たら良いかの詳しい説明が欲しい。
- /r/と/l/を瞬時に切り替えることができるとより良いかも。
- 模型自体が小さく、遠いと手元が見えにくい。
- 操作が大変のよう。レバーを速く操作する必要がある。

## 4. 総合考察

本稿では、著者が今まで開発してきた声道模型について述べてきた。まず代表的なものについて、声道模型が人間のように屈曲しているか、あるいは真っすぐ伸びた管なのかによって大きく分類後、模型における動きの自由度によって模型を並べた。そして、各声道模型に関して改めて紹介した後、それらを用いて行った講義の例の紹介と、教育的効果を示すデータ、アンケート調査の結果などを示した。それらのプロセスを通じ、本章では声道模型を用いた音響教育の在り方について考察する。

### 4.1 音響教育が意味するものと目的別声道模型

狭い意味での「音響教育」といった場合、音そのものを教育することが主眼となることが多い。そこでは、例えば音が何かの振動することによって作り出されること、そしてその振動が空気といった媒質を介して伝わっていくこと、音波の波長は振動数や音速などに関係すること、共鳴現象、音のスペクトルなどと様々なトピックが含まれる。音声は人間にとって大事なもので、かつ身近なものであるから、そういう「音そのものの教育」の際に声を例にすることは効果が非常に大きい。しかし、そのいわば「狭義の音響教育」の場合には、母音・子音の作られ方などを必ずしも説明するとは限らない。そのような説明がなかったとしても、音

声を使った「音に関する教育」は有効であり、また声道模型を介してそのような（狭義の）音響教育を行うことも有効であることが、我々の経験からも分かっている [14, 16]。

一方、単に「音響教育」といった場合には、もっと広い意味で使われることも少なくない。例えば、発音を教えることに主眼が置かれた「音声教育」あるいは「発音教育」もこの広い意味での「音響教育」に含まれるであろうし、音楽を教える「音楽教育」も同様である。特に本稿で取り扱った声道模型は、この音声教育にも有効であり、また言語聴覚士が言語障害者に施す言語治療においても有効であると考えられる。そこで重要になるのは、「いかにターゲットとなる『音』を出せるように指導、あるいは訓練するか？」ということになるが、その際に「いかにその音が作られるのか」に対する理解を、今まで著者が提案してきた声道模型が手助けしてくれることになる。その際、最終的には「その音」が出るようになることが目的であるので、必ずしもその背後に潜む難解な音響理論を説明する必要はない。そういう点で狭義の音響教育という色は必ずしも濃くはないものの、広義の音響教育には入ると考えていいだろう。背景にある音響理論まで理解する必要がある場合も声道模型を利用することが可能であり、大学院での講義や専門家が研究を進める上での実験装置にもなり得る教具である（学部や大学院での講義の例については、例えば文献 [16] を参照）。このことは、外国語を学ぶ第2言語学習者のための音声教育についても同様のことが言え、外国語の発音を指導・訓練する際に声道模型が有効になるものと考えられる。また、言語治療においても発音指導においても、教える側の指導者が理論を理解していることも重要であり、その際にも声道模型は有効と言える。

このように声道模型は、音響学そのものを教える狭義の音響教育にも有効であり、かつ発音教育・音声教育にも有効である。しかし、ここで問題となるのが、その目的によって最適な模型は必ずしも同一ではないということである。むしろ、その目的によって声道模型の種類は変わってきて当然であり、多様な目的に合わせて様々な模型を使い分けることが大切である。その点で、目的に対応できる模型のバリエーションを増やしておくことが重要となってくる。

一般に、ある現象を説明しようとした場合、直接的に関係のない要因はなるべく排除し、目的の現象に関係する部分だけを残したような教材・教具を開発することができれば、それが仮に実体とはかけ離れていても教育的価値は高い場合が多い。例えば、VTM-S20は声道は真っすぐで実際の声道とは異なる上、舌もス

ライドするわけではない。しかし、狭窄の位置が母音生成には重要であることを説明するにはかなり効果が高く、これは他の要因が極力排除されている効果であると考えられる。一方、発音訓練をしている訓練者にとって、よりリアルな舌の動きは発音上のイメージを作り易いものと考えられる。そのためには声道は屈曲していたほうがいだろうし、舌もある程度実体に近く、その動きも仮に簡素化されていたとしてもリアルな動的变化を模擬したもののほうがイメージがし易い。そのような観点からすると、AP モデルの有効性が期待される。もちろん、FT モデルは舌が比較的自由に動くので様々な音を作り出せる可能性はあるのだがその操作が難しい。そこで、ある程度、動きに制約をかけながらより限定された AP モデルのような声道模型も特に発音訓練などには有効になるものと考えられる。その観点からすると、言語障害者の構音障害に頻繁に現れる音をターゲットにした屈曲型の声道模型を開発することはこの分野において急務であると言っても過言ではない。例えば日本語において、ラ行子音は子どもにとっても獲得が難しい音だと言われている [22]。そこで、日本語ラ行子音として典型的に現れる「はじき音」を FL モデルを用いて示したり、その異音の数々を AP モデルを用いて示すことは有効であると考えている [22]。また、その他の音の例としては例えば/s/ や/j/などの音に関する模型など、実現されればその恩恵は大きいものと思われる。

以上のように、必ずしも自由度の低い模型のほうが自由度の高いものよりも劣っているということではなく、また人間のように複雑な音を出せる模型のほうがシンプルなものよりも勝っているということでもない。自由度が低いことによって目的の音を素早く簡単に（つまり誰もが）生成可能となると同時に、シンプルな模型であれば現象の本質を効果的に示すことが可能となる。このように、それぞれに長所があり、目的や利用方法によって価値が異なるので、それらをうまく活用することで教育上の効果も高まる。

#### 4.2 教育効果を示す方法について

3 章では、幾つかの授業にて行った試験による評価やアンケート調査の結果について紹介した。実際に声道模型を用いた教育上の効果を調べようとした場合、学習者とその授業内容を学習したかどうかを調べる必要がある。一番確実な方法は、二つのクラスに対して一方には声道模型を用いた授業を行い、もう一方は逆に声道模型を用いない授業を行う。そして同じ試験を実施することによって、その 2 クラス間に成績上、有意な差が存在するかどうかを問うものである。しかしここで問題となるのが、二つのクラス間で異なる講義

を行ったことによる不公平性である。そのため、声道模型を用いない講義を行ったクラスについては、後日、声道模型を用いる授業を行うことなどのフォローが必要となる。しかし、そういう時間が取れないケースも少なくないであろう。

文献 [14] で紹介した授業の場合は、1 クラス全員に対して声道模型を用いない講義を行い、まずは 1 回目の筆記試験を行った。その後には声道模型を用いた講義を行って、1 回目と全く同じ試験を再度行ったのである。この結果、1 回目よりも 2 回目の試験の成績のほうが統計的に有意に改善していた。しかし、2 回目の試験ではその改善の中に学習効果が含まれてしまっていることが予想される。そこで、前半の講義にしか関係しない問題群 (G1 群) と、後半の講義に関係する問題群 (G2 群) に分けて、その 2 群間で成績の改善度を調べた。3 章において図-2 に示される結果より、G2 群では、改善度=+2 が G1 群の 6.6% に比べて 16.3% と増加しており、不正解から正解に転じた学生の比率が多くなっている。その結果、クロス集計全体として有意差が認められた。これにより、追加で行った後半の講義 (講義 2) のように声道模型によるデモンストレーションを交えた授業を行うことによって、学習者の理解が促進されると考えられる。

ところで、講義 2 において声道模型によるデモンストレーションを交えないで同内容の講義を行った場合に、成績が同じように上がるのかどうかについては厳密には明確ではない。そこで、G1-G2a 群間と G1-G2b 群間との比較結果を考えてみる。G1 群で改善度=+2 が 6.6% であったのに対し、G2a 群では改善度=+2 が 8.8%、G2b 群では 20.0% であった。その結果、不正解から正解に転じた学生の比率が G2a 群に比べ、G2b 群でより多くなっている。その結果、G2a 群は G1 群に比して有意な傾向で高い改善度を示したのにとどまったのに対し、G2b 群では G1 群に比して有意に高い改善度を示したのと考えられる。つまり、講義 2 において追加の授業を受けたものの、声道模型による説明と直接的ではない内容を問われた場合 (G2a 群)、改善度はそれほどまでに伸びなかった。一方、声道模型によるデモンストレーションに直接的に関係する内容を問われた場合 (G2b 群)、改善度は著しく伸びていることから、ただ単に追加の講義を行ったから成績が上がっただけではなく、声道模型を用いたことの効果があつたと考えられる。

このような分析を含むことによって、1 クラス全体で同じ講義を実施する形式であっても、教育効果を見ることが可能となり、また学習者全員について同じ扱いをすることができるために不公平感は最小限に抑え



られる。ただし、この場合には学習効果への影響が問題群にわたって変わらないことが前提となっている点については、注意が必要である。

## 5. ま と め

声道模型を通じ、音響教育への応用について述べてきた。そのプロセスを通じて、声道模型は音そのものの教育に対して有効である他、発音教育・音声教育への有効性についても議論した。そして、目的に応じて声道模型を使い分けることが重要であること、そしてまた使い分けに対応すべく多くのバリエーションを持った模型を今後も開発していく必要があることが分かった。例えば、言語障害者にとって習得が遅い音に対する声道模型の開発も急務であることにも触れた。

これらの声道模型は現在、全世界の教育機関の協力を得ながら、その使用上の改善を含む模型に関するフィードバックを収集している。また、アウトリーチ活動の一環として博物館での科学教室（例えば [16]）等での使用のほか、新たに設置される展示へも協力している（日本では沖縄こどもの国ワンダーミュージアム、海外ではスイスの博物館など）。今後とも、様々な特性を持つ模型の開発やその応用を広げていきたい。

## 謝 辞

声道模型の開発並びに教育に関してお世話になった方々、特に前川喜久雄先生、佐藤大和先生、川原繁人先生、梅田規子先生、John Ohala 先生、Kenneth Stevens 先生、上智大学荒井研究室のメンバー、同大学言語聴覚研究センターの皆様方に感謝申し上げます。内容の一部は JSPS 科研費 24501063 の助成を受けた。

## 文 献

- [1] T. Chiba and M. Kajiyama, *The Vowel: Its Nature and Structure* (Tokyo-Kaiseikan Pub. Co., Ltd., Tokyo, 1941), pp. 115–154.
- [2] K. Maekawa, “From articulatory phonetics to the physics of speech: Contribution of Chiba and Kajiyama,” *Acoust. Sci. & Tech.*, 23, 185–188 (2002).
- [3] G. Fant, *Acoustic Theory of Speech Production* (Mouton, The Hague, Netherlands, 1960).
- [4] K.N. Stevens, *Acoustic Phonetics* (MIT Press, Cambridge, Mass., 1998).
- [5] J.L. Flanagan, *Speech Analysis, Synthesis and Perception* (Springer-Verlag, Berlin, 1965).
- [6] B. Gold and N. Morgan, *Speech and Audio Signal Processing* (John Wiley & Sons, New York, 2000).
- [7] K.N. Stevens, “The acoustic/articulatory interface,” *Acoust. Sci. Tech.*, 26, 410–417 (2005). K.N. Stevens 著, 荒井隆行訳, “音響と調音のインタフェース,” 音響学会誌, 61, 524–531 (2005).
- [8] D.H. Klatt and L.C. Klatt, “Analysis, synthesis, and perception of voice quality variations among female and male talkers,” *J. Acoust. Soc. Am.*, 87, 820–857 (1990).
- [9] 梅田規子, 寺西立年, “声の韻質と声質: 音響的声道模型による音声の合成,” 音響学会誌, 22, 195–203 (1966).
- [10] K. Honda, H. Takemoto, T. Kitamura, S. Fujita and S. Takano, “Exploring human speech production mechanisms by MRI,” *IEICE Trans. Inf. Syst.*, E87-D, 1050–1058 (2004).
- [11] T. Mochida, M. Honda, K. Hayashi, T. Kuwae, K. Tanahashi, K. Nishikawa and A. Takanishi, “Control system for talking robot to replicate articulatory movement of natural speech,” *Proc. Interspeech*, pp. 1533–1536 (2002).
- [12] T. Arai, “The replication of Chiba and Kajiyama’s mechanical models of the human vocal cavity,” 音声研究, 5(2), pp. 31–38 (2001).
- [13] T. Arai, “Sliding three-tube model as a simple educational tool for vowel production,” *Acoust. Sci. & Tech.*, 27, 384–388 (2006).
- [14] T. Arai, “Education system in acoustics of speech production using physical models of the human vocal tract,” *Acoust. Sci. & Tech.*, 28, 190–201 (2007).
- [15] T. Arai, “Gel-type tongue for a physical model of the human vocal tract as an educational tool in acoustics of speech production,” *Acoust. Sci. & Tech.*, 29, 188–190 (2008).
- [16] T. Arai, “Education in acoustics and speech science using vocal-tract models,” *J. Acoust. Soc. Am.*, 131, Pt. 2, 2444–2454 (2012).
- [17] R.D. Kent and C. Read, *Acoustic Analysis of Speech* (Singular Publishing, San Diego, California, 2001). 荒井隆行, 菅原 勉 監訳, 音声の音響分析 (海文堂, 東京, 1996).
- [18] T. Arai, “Mechanical vocal-tract models for speech dynamics,” *Proc. Interspeech*, pp. 1025–1028, Makuhari (2010).
- [19] T. Arai, “Mechanical bent-type models of the human vocal tract consisting of blocks,” *J. Acoust. Soc. Am.*, 133, 3364 (2013).
- [20] T. Arai, “Physical models of the vocal tract with a flapping tongue for flap and liquid sounds,” *Proc. Interspeech*, 2019–2023 (2013).
- [21] 荒井隆行, “英語 /r/–/l/ 音に対する物理模型の試作—音声教育のための新しい教材の提案—,” 日本音声学会全国大会予稿集, pp. 53–58 (2013).
- [22] 荒井隆行, “構音の獲得に潜む音響的側面を探る—日本語ラ行音を中心に—,” 日本コミュニケーション障害学会学術講演会予稿集, p. 51 (2013).