

ブロック式声道模型による母音の韻質評価*

○荒井隆行（上智大・理工）

1 はじめに

これまで、人間の声道を模擬した物理模型は、特に音響音声学や音声科学分野において教育目的に応用した際、教材として非常に有効であることを示してきた[例えば, 1,2]. 今まで用いられてきた声道模型の多くは、

- 1) まっすぐな音響管モデル
- 2) 頭部形状型モデル

のいずれかであった. 1) の場合、声道形状から断面積関数を推定し、まっすぐで、かつ、シンプルな形状で近似している. 一方2) は、頭部の外形を保ちながら正中矢状断面等を参考に設計された声道を有する模型となっており、声道が頭部のどの辺りに位置しているかを学習者に示すのに役立っている. ただし、2) は模型のサイズが大きいため、簡便なデモンストレーションには1)が適していることも多い. しかし、1) はまっすぐだが実際の声道は曲がっているため、1) のようなシンプルさを保ちつつ、かつ屈曲したモデルを検討してきた.

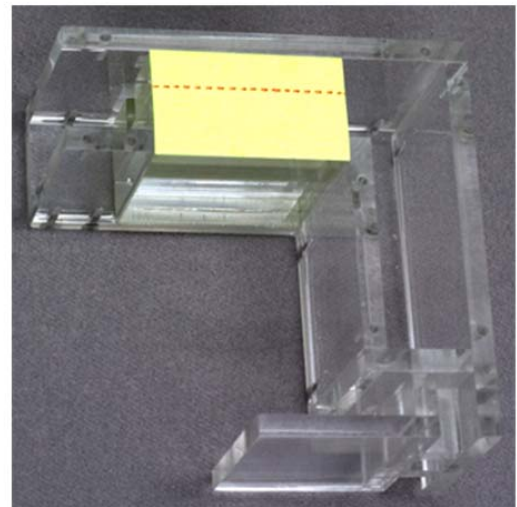
その結果、すでに開発していた軟らかい素材を用いた舌を有する模型[3]の延長として、Fig. 1 に示すようなブロック式の声道模型を開発し、そのテスト運用を行ってきた[4]. しかし、どこにどのようなブロック（狭め）が存在するときどの母音になるかなど、詳細についてはまだ分かっていないことも多い. そこで、本稿では途中で直角に屈曲する声道模型において、複数のブロックを少しずつ位置をずらしながら配置した際、生成される母音の韻質評価について報告する.

2 ブロック式声道模型

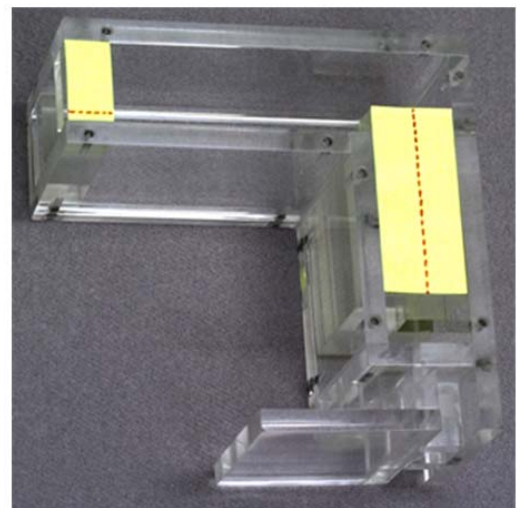
2.1 サイズ

Fig. 2 に、本稿で用いた声道模型の正中矢状断面を模式的に示す. その概形は口腔の長さが 90 mm, 咽頭腔の長さが 80 mm,

そしてその下に長さ 20 mm の喉頭部がある. また、口腔と咽頭腔において、中心線に直交する断面は、常に 45 mm × 20 mm である. そのような声道模型において、口唇ブロック、前舌ブロック、後舌ブロック



(a)



(b)

Fig. 1. Block-type vocal-tract models: (a) front vowels; and (b) back vowels.

* Evaluation of the quality of vowels produced by the block-type vocal-tract model, by ARAI, Takayuki (Sophia University).

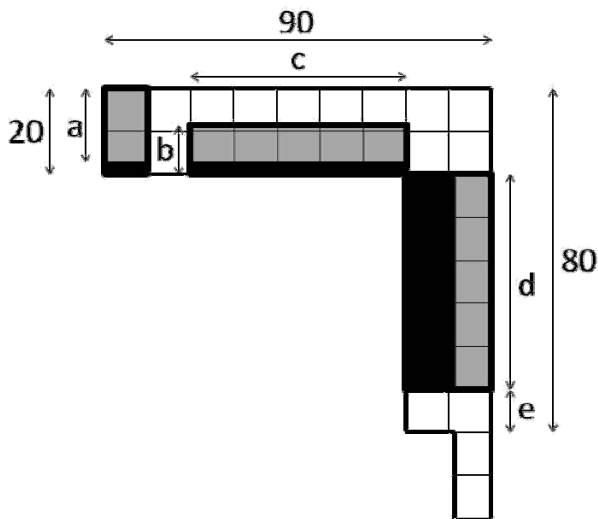


Fig. 2. Dimensions of the vocal-tract models with three blocks .

という3種類のブロックを Fig. 2 のように配置した. 口唇ブロックは, サイズが $45\text{ mm} \times 20\text{ mm} \times 10\text{ mm}$ で, 上方が $a\text{ mm} \times a\text{ mm}$ 切り取られている. 前舌ブロックは, サイズが $45\text{ mm} \times b\text{ mm} \times c\text{ mm}$ で, その上面には声道の長さ方向に沿って $9\text{ mm} \times 9\text{ mm}$ の溝が設けられている. 後舌ブロックは, サイズが $45\text{ mm} \times 20\text{ mm} \times d\text{ mm}$ で, その背面には声道の長さ方向に沿って $9\text{ mm} \times 9\text{ mm}$ の溝が設けられている. また, 後舌ブロックは, 喉頭部との境界から $e\text{ mm}$ だけ上方に位置しているものとした.

以上のように, 配置に関するパラメータが a から e まで5個存在する. それぞれのパラメータの値の組み合わせによって, 配置条件を表すものとし, また各パラメータの値が0の場合は, その該当するブロックを使用しない場合と定義した. そして, それぞれの条件におけるインパルス応答を測定すると共に, 音源信号を入力した際の出力信号を録音した.

2.2 録音

録音に際しては, ホーンスピーカのドライバユニット (TOA TU-750) を用いた. 入力信号を, オーディオインタフェース (RME, FireFace UC) とパワーアンプ

(FOSTEX AP1020) を介してドライバユニットに出力し, ドライバユニットからの出力信号が声道模型の声門側の穴に入力されるようにセッティングした.

入力信号として, 3つの信号を用いた. 1つめと2つめはインパルス列による入力信号で, 共にそれぞれ標準化周波数が 16 kHz で作成後, 48 kHz にアップサンプリングした. 1つめの信号では, 最初の 100 ms で基本周波数 f_0 を 100 から 125 Hz に上昇させ, その後の 200 ms で f_0 を 125 から 100 Hz に下降させた. 2つめの信号では, f_0 を常に 100 Hz とした. いずれの信号も, 持続時間は 300 ms であった. 3つめの信号は swept-sine 信号で, 標準化周波数は 48 kHz , サンプル数は 65536 であった.

ドライバユニットのネック部においては, 声道模型との音響的結合を実現するためにアルミ製の円盤を装着した. 円盤の片面中央にはネック部に挿入するための突起があり, その中央には面積が 0.13 cm^2 の穴が貫通している. ドライバユニットから発せられた音響信号は, その穴から声道模型へと入力されることになる.

模型から発せられる音は, 騒音計 (RION NL-18) とオーディオインタフェース (RME, FireFace UC) を介して標準化周波数 48 kHz にてデジタル録音された. マイクロホンは模型から約 20 cm 離れたところに位置し, 収録は防音設備のある部屋で行われた.

3 聴取実験

前節で録音された音声信号に対し, 聴取実験を行った. この聴取実験では, 実験参加者は日本語5母音のうち, どの母音に聞こえたかを, 満足度と共に答えてもらった.

3.1 刺激

聴取実験で用いる刺激は, 2節で録音されたもののうち, f_0 が山形の軌跡を有する音源を入力した際の出力音声である. 刺激は77個あった.

3.2 参加者

実験には, 20名の日本語母語話者 (男性10名, 女性10名) が参加した. 年齢は20歳から29歳で, 平均年齢は 24.1 歳であった.

3.3 手続き

聴取実験は防音設備のある室内で行われた。刺激音は、オーディオインタフェース（RME Babyface）とアンプ（FOSTEX AP1020）を介して、スピーカ（NAE NESmini）から提示された。提示された刺激音の騒音レベルは、平均 70.1 dB であった。本実験に先立ち、練習用として 8 個の刺激が使用された。

本実験では、各試行につき 1 刺激音ずつ提示され、その刺激音に対して次に示す 2 つの質問それぞれについて選択肢を 1 つ、GUI (graphical user interface) を介して回答してもらった：

- Q1) 聞こえてきた母音が日本語の「ア」「イ」「ウ」「エ」「オ」のどれに聞こえたか、1 つ選んでください。
- Q2) Q1 の回答に対し、どれほど満足しているかを、100%、75%、50%、25%、0%の中から 1 つ選んでください。

3.4 結果

Table 1 に、実験の結果を示す。表の数値は、3.3 節の Q1 と Q2 の結果を両方反映させた数値である。つまり、Q1 に対しある母音として判定された場合すべてについて満足度を累積し、その作業を参加者 20 名分について各母音ごとに行う。そして、累積された満足度を参加者の人数である 20 で割った。全員が同じ母音に回答し、さらにその満足度がすべて 100%であったとき、Table 1 の累積値は 100%を示すことになる。

この表を見ると、影を付けた条件においてそれぞれの母音に対する累積満足度が高い値を示していることが分かる。

4 考察

母音/a/については、後舌ブロックだけで十分であり、かつブロックが喉頭部との境界から 5 mm とわずかに上方にシフトしたときに最も/a/らしく聞こえることが分かる。また、その後舌ブロックの長さも 45 mm あるいは 50 mm が最適であった。

母音/i/については、逆に前舌ブロックだけで十分であり、かつブロックの高さが 20 mm のものときに/i/として聞こえることがわかった。特に前舌ブロックの長さも、

50 mm か 55 mm の場合が最適であった。

母音/u/については、後舌ブロックに口唇ブロックが組み合わさった際に最も/u/らしく聞こえた。特に長さ 50 mm の後舌ブロックのなるべく上方に配置させ、かつ 9 mm × 9 mm の穴の開いた口唇ブロックを用いた場合に/u/らしく聞こえた。

母音/e/については、母音/i/と同様に前舌ブロックだけでよく、かつブロックの高さも 12 mm でよい。前舌ブロックの長さは 45 mm から 55 mm の範囲で最も/e/らしく聞こえた。

母音/o/については、後舌ブロックと口唇ブロックを組み合わせる必要があった。後舌ブロックは、長さが 50 mm あるいは 55 mm のものを、喉頭部との境界から 20 mm 上方にシフトさせた。口唇ブロックは 18 mm × 18 mm の穴の開いたものを用いた。このとき、最も/o/らしく聞こえた。

5 おわりに

ブロック式声道模型において、ブロックの配置をいろいろ変えながら、日本語 5 母音に最適な配置を検討した。その結果、前舌ブロック、後舌ブロック、口唇ブロックという 3 種類のブロックを決められたサイズで適切な位置に配置することによって、日本語 5 母音が生成されることを確認した。

今後は、得られた声道形状を有する静的な模型を製作したり、また限られたブロックをスライドさせるだけで複数の母音が生成される声道模型を開発し、教育現場での応用を目指したい。

謝辞

内容の一部は JSPS 科研費 24501063 の助成を受けた。

参考文献

- [1] Arai, T., "Education system in acoustics of speech production using physical models of the human vocal tract," *Acoust. Sci. Tech.*, 28(3), 190-201, 2007.
- [2] Arai, T., "Education in acoustics and speech science using vocal-tract models," *J. Acoust. Soc. Am.*, 131(3), Pt. 2,

2444-2454, 2012.

- [3] Arai, T., “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(2):188-190, 2008.
- [4] Arai, T. “Mechanical bent-type models of the human vocal tract consisting of blocks,” *Proc. of ICA*, 2013.

Table 1. Accumulated satisfaction scores of five vowels for each parameter setting of the block-type vocal-tract model.

a	b	c	d	e	/a/	/i/	/u/	/e/	/o/
0	0	0	0	0	0	0	3	55	8
0	0	0	45	0	29	0	0	23	0
0	0	0	45	5	81	0	0	0	0
0	0	0	45	10	55	0	0	0	0
0	0	0	45	15	34	0	3	0	14
0	0	0	45	20	11	0	4	0	16
0	0	0	45	25	3	0	28	3	13
0	0	0	45	30	15	0	40	0	0
0	0	0	50	0	38	0	6	1	1
0	0	0	50	5	79	0	0	0	0
0	0	0	50	10	49	0	0	0	9
0	0	0	50	15	21	0	18	0	4
0	0	0	50	20	25	0	18	0	0
0	0	0	50	25	1	0	41	0	1
0	0	0	55	0	51	0	6	0	1
0	0	0	55	5	56	0	5	1	0
0	0	0	55	10	73	0	3	0	0
0	0	0	55	15	28	0	15	1	0
0	0	0	55	20	3	0	20	0	6
0	0	0	60	0	51	0	18	0	0
0	0	0	60	5	60	0	18	0	0
0	0	0	60	10	39	0	0	0	19
0	0	0	60	15	8	0	14	26	1
0	12	45	0	0	0	0	8	66	4
0	12	50	0	0	0	0	15	69	0
0	12	55	0	0	0	0	9	65	1
0	12	60	0	0	0	23	0	61	0
0	20	45	0	0	0	58	16	0	0
0	20	50	0	0	0	64	0	25	0
0	20	55	0	0	0	86	0	0	0
0	20	60	0	0	5	60	0	0	5

9	0	0	0	0	0	0	0	61	0	9
9	0	0	45	0	21	0	41	1	0	0
9	0	0	45	5	0	0	51	0	6	6
9	0	0	45	10	21	0	35	0	1	1
9	0	0	45	15	1	6	14	0	26	26
9	0	0	45	20	0	0	24	0	21	21
9	0	0	45	25	1	4	18	0	16	16
9	0	0	45	30	0	1	28	0	14	14
9	0	0	50	0	1	0	50	1	15	15
9	0	0	50	5	0	0	65	0	0	0
9	0	0	50	10	0	1	24	9	6	6
9	0	0	50	15	0	20	15	0	15	15
9	0	0	50	20	0	3	14	0	19	19
9	0	0	50	25	1	0	19	11	16	16
9	0	0	55	0	11	0	53	0	0	0
9	0	0	55	5	6	0	43	6	0	0
9	0	0	55	10	0	0	24	0	23	23
9	0	0	55	15	0	3	9	0	39	39
9	0	0	55	20	0	0	14	0	34	34
9	0	0	60	0	1	0	54	0	8	8
9	0	0	60	5	19	0	38	0	1	1
9	0	0	60	10	0	0	13	24	19	19
9	0	0	60	15	0	0	31	0	16	16
18	0	0	0	0	4	0	28	5	18	18
18	0	0	45	0	3	1	21	25	0	0
18	0	0	45	5	29	0	14	0	0	0
18	0	0	45	10	18	0	0	0	26	26
18	0	0	45	15	11	0	1	1	41	41
18	0	0	45	20	4	0	15	0	38	38
18	0	0	45	25	1	0	21	0	28	28
18	0	0	45	30	4	0	16	0	16	16
18	0	0	50	0	16	0	11	11	0	0
18	0	0	50	5	56	0	3	0	0	0
18	0	0	50	10	9	0	10	0	33	33
18	0	0	50	15	10	0	0	4	41	41
18	0	0	50	20	1	0	10	0	54	54
18	0	0	50	25	11	0	24	0	15	15
18	0	0	55	0	28	0	16	1	0	0
18	0	0	55	5	30	23	1	0	5	5
18	0	0	55	10	15	0	19	0	11	11
18	0	0	55	15	25	0	0	0	25	25
18	0	0	55	20	4	0	3	0	59	59
18	0	0	60	0	38	0	3	24	0	0
18	0	0	60	5	54	0	0	0	6	6
18	0	0	60	10	13	0	21	0	21	21
18	0	0	60	15	9	0	3	0	23	23