

エストニア語の母音を声道模型から生成する試み*

○荒井隆行（上智大・理工），木村琢也（上智大院・理工研／清泉女子大・文），
内田玲（上智大院・理工研）

1 はじめに

今まで，声道模型を用いた展示のいくつかに携わる機会を得てきたが，その多くは，筆頭著者がもともと開発してきた日本語 5 母音を中心とする声道模型であった．そのいくつかの例を以下に示す．

1.1 静岡科学館

静岡科学館「る・く・る」における展示に関しては，本学会の小特集記事[1]においても触れたが，「発声のしくみ」というコーナーに声道模型を用いた展示が設置された．この科学館の愛称である「る・く・る」は，科学館のキーワード「みる，きく，さわる」の語尾をつなぎ合わせたもので，ここでは我々が開発した声道模型[2]が子ども達に「見て，聞いて，触って」もらっている．展示では「人間の発声メカニズム」を意識して，声道模型と笛式人工喉頭（音源）とふいごを組み合わせ，空気を送って母音を出す仕掛けになっている．

なお，声道模型としては2種類をものを準備した．1 つめの種類は Chiba and Kajiyama に基づく声道形状を実現した筒型のもの（VTM-C10）で，もう1つの種類はその形状を0次近似したような階段状の断面積関数を実現するプレート型の声道模型（VTM-P10）である．このプレート型の声道模型については，直径が異なる穴が中央に開いた厚さ 10 mm のアクリル板（透明）を複数枚並べ，声道形状を実現するものとなっており，子どもたちがパズル感覚で任意の形状を実現できることから体験型展示には適したものとなっている．

1.2 日立シビックセンター科学館

一方，日立シビックセンター科学館においても，同様の展示を監修した[3]．その際，採用した声道模型は静岡科学館の1つめと同じ筒型の声道模型（VTM-C10）であった．

また，ふいごと音源を組み合わせると喉頭原音を声道模型に入力する原理は同じであるが，音源として新たに設計したリード式音源を採用し，リード（プラスチックシート）を設置する「受け」については，細いアクリルの筒を斜めにカットしたものをを用いている[3]．

1.3 沖縄こどもの国・ワンダーミュージアム

沖縄こどもの国にあるワンダーミュージアムにおいても，同様の展示について監修を行った．その際，採用した声道模型はVTM-C10の内側の形状をさらに簡素化したVTM-T50である．このVTM-T50は，その内側の形状に関しては文献[4]において報告した円筒管接続式のものと同様であり，母音/u/についてのみ新しい形状を採用している．

1.4 本研究の目的

以上のように，日本国内の博物館・科学館展示においては，日本語5母音を対象とした声道模型を応用することが可能であった．一方，最近になって，外国の博物館・科学館における展示について協力を依頼されることがあった（スイスやドイツなど）．その際，今までは日本語5母音を対象とした声道模型で対応することがほとんどであった．しかし，外国語の母音に対する声道模型を設計するほうが望ましいケースがあるのも事実であった．

そのような状況の中，エストニア国立博物館からの声道模型の展示に関する協力の依頼があった際，エストニア語の母音について声道模型での実現の可能性を探ることとなった．そこで，本稿ではそのために設計した声道模型の形状と，そこから生成される母音の音質について報告する．

* Producing Estonian vowels by using physical models of vocal tract, by ARAI, Takayuki (Sophia Univ.), KIMURA, Takuya (Sophia Univ. / Seisen Univ.), and UCHIDA, Rei (Sophia Univ.).

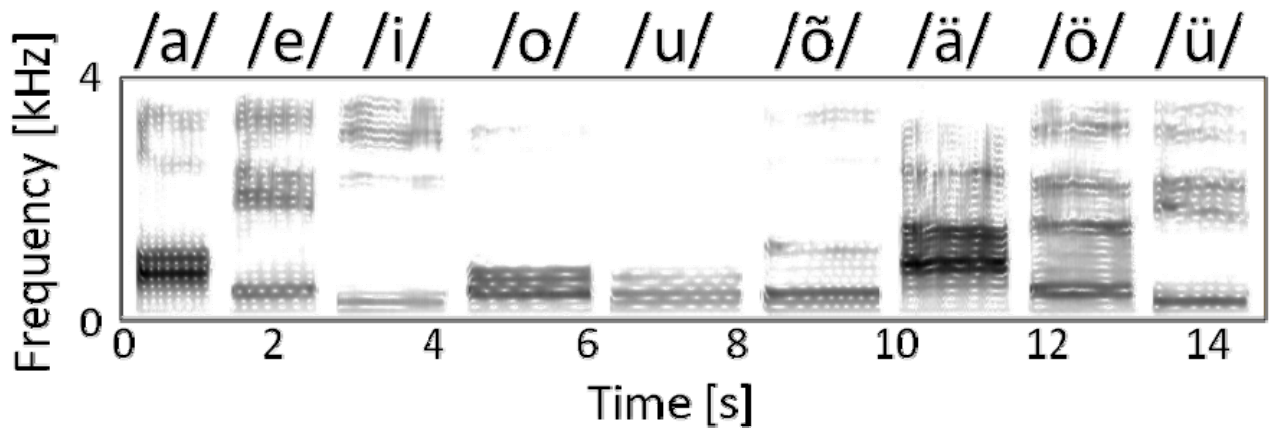


Fig. 1: Spectrograms of nine vowels in Estonian (male speaker).

2 エストニア語の母音

2.1 エストニア語について

エストニア語は、ウラル語族フィン・ウゴル語派に属する言語であり、次の9種の母音を有する（エストニア語の正書法に基づいて表記する）: /a/, /e/, /i/, /o/, /u/, /õ/, /ä/, /ö/, /ü/. この後、本稿では9母音を並べる際には、断りがない限りこの順番で表示することとする。

2.2 録音

エストニア語の9母音について、母語話者男女1名ずつに孤立発話、および単語に挿入した形で発話してもらった。その音声を、デジタルレコーダ（Marantz, PMD670）にて録音した（48 kHz 標本化, 16 ビット量子化）。また同時に、その際の舌の形状を超音波診断装置（日立アロカメディカル, ProSound 2）を用いて、そして顔正面の様子をビデオカメラを用いてそれぞれ撮影した。

2.3 スペクトログラム分析

男性話者が発話する9母音について、そのスペクトログラム分析した結果を Fig. 1 に示す。この図を見ることによって、大まかなフォルマントの位置が確認できる。

2.4 超音波画像の分析

同じ男性話者が母音/e/を発話した際の超音波画像を Fig. 2 に示す。この図は、プローブを顎の下から上方に向けて押し当てた際の正中矢状断面を表しており、左が前方、右が後方である。舌の上面で超音波が反射しているところが、図中で白く描かれている[5]。



Fig. 2: Ultrasonic image of the midsagittal plane when the male speaker uttered vowel /e/ (left: front).

3 声道模型による母音の生成

本節では、声道模型によってエストニア語の9母音を生成する試みについて述べる。

3.1 声道形状

母音の音響データから声道形状を推定することは、一対多の写像問題になるため、ここでは知識に加え、超音波画像などの補助的情報をもとに近い形状を予測し、その後、trial and errorを繰り返すことで近似解を求めることとした。より具体的な手順を以下に述べる。なお、使用する声道模型のタイプは、プレート型 (VTM-P10) とした。このプレート型では、厚さが 10 mm のアクリル板の中央に、直径が 8 mm から 38 mm まで（最小刻み幅は 2 mm）の異なる大きさの穴が開いている。

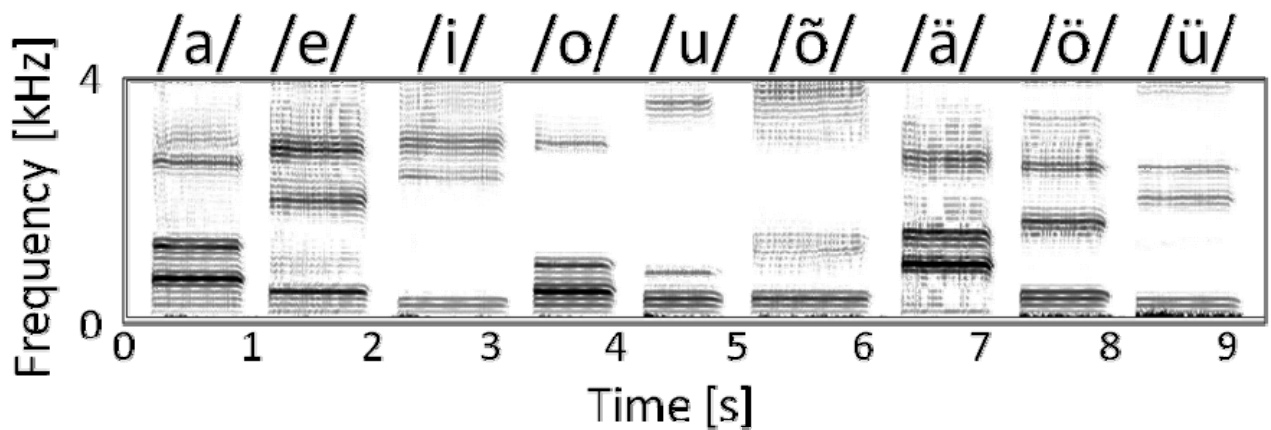


Fig. 3: Spectrograms of nine vowels in Estonian (vocal-tract model).

【ステップ1】ターゲットとなる母音に近い日本語の5母音の声道形状を初期値とする。その形状は、文献[2]にある Chiba and Kajiyama [6]の声道形状に対する測定結果を、階段近似したものなどである。

【ステップ2】その声道形状をVTM-P10で実現し、そこから作られる音響出力から母音の音質を評価する。その際、調音音声学、音響音声学等の知識を用い、また超音波画像を補助的に利用しながら総合的に形状に補正を試みる。なお、声道模型を駆動するための音源は、リードの受けがR30の曲面を描くリード式音源を用いた（詳細は文献[7]を参照）。

【ステップ3】形状の補正が正しい音の変化を生んでいると判断された場合は、その補正を続け、もし音質が遠ざかっていると判断された場合は補正を中断する。この作業を繰り返し、最終的にターゲット音に十分近いと判断された時点で、形状を確定する。

上記のステップの結果、得られた声道形状をTable 1に示す。表中の数字は、それぞれの母音について、プレート中央に開けられた穴の直径を示す（単位はmm）。インデックス番号は喉頭側から順番に番号付けられている（声道長は、成人男性を想定して160mmにて統一）。

Table 1: Vocal-tract configurations of nine vowels in Estonian. The numbers are the diameters (in mm) of the hole for the 10-mm thick plate lining up from the glottis end through the lip end.

No.	/a/	/e/	/i/	/o/	/u/	/õ/	/ä/	/ö/	/ü/
1	12	12	12	12	12	12	10	12	12
2	12	12	12	12	12	12	10	12	12
3	26	30	32	38	32	32	10	30	32
4	16	30	32	18	38	38	10	30	32
5	12	30	32	12	38	38	10	30	32
6	14	30	32	14	34	34	12	30	32
7	20	28	32	14	12	32	12	28	32
8	26	24	32	16	8	10	12	24	32
9	30	18	24	30	8	10	30	20	24
10	34	16	10	38	8	10	30	18	10
11	38	16	8	38	32	12	30	16	8
12	38	16	8	38	34	30	30	14	8
13	34	16	8	34	34	34	30	14	8
14	30	18	10	30	32	34	32	16	10
15	32	22	14	18	16	32	32	24	14
16	24	24	24	14	14	30	32	10	8

3.2 スペクトログラム分析

リード式音源による母音の生成結果を録音した（機材等は2.2節と同じ）。9母音に対し、スペクトログラム分析を施した結果をFig. 3に示す。この図を見ることによって、Fig. 1と比べフォルマントの位置がほぼ同じであることが確認できる。

4 おわりに

本稿では、エストニア語の母音について声道模型で生成することを試みた。生成された音に対するエストニア母語話者による判定などは現在、進めているところである。そして最終的には博物館の展示に耐えうるかどうかなどの検討に入る予定である。

今後は、様々な言語において同様に個別の母音に対応できるような準備を進めていきたいと思っている。その際、今回のようなアプローチがいいのか、また別の手法がいいかについてなど、検討を進めたい。

謝辞

本研究にあたり、協力いただいた Timo Toots 氏には心から感謝申し上げます。内容の一部は日本学術振興会の科学研究費(15K00930)の助成を得た。

参考文献

- [1] 荒井隆行, “小中学生に対する音響教育,” 日本音響学会誌, 64(1), 29-34, 2008.
- [2] T. Arai, “The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity,” *Journal of the Phonetic Society of Japan*, 5(2), 31-38, 2001.
- [3] 荒井隆行, “日立シビックセンター科学館における声道模型の展示,” 日本音響学会春季研究発表会講演論文集, 1587-1590, Mar. 2011.
- [4] T. Arai, “Simple physical models of the vocal tract for education in speech science,” *Proc. of the INTERSPEECH*, 756-759, Brighton, 2009.
- [5] I. Wilson, “Using ultrasound for teaching and researching articulation,” *Acoust. Sci. & Tech.*, 35(6), 285-289, 2014.
- [6] T. Chiba and M. Kajiyama, *The Vowel, Its Nature and Structure*, Tokyo-Kaiseikan, 1942.
- [7] 荒井隆行, “スライド式声道模型のためのリード式簡易音源の比較,” 日本音響学会秋季研究発表会講演論文集, 1563-1566, Sep. 2008.

[Appendix]

付録として、エストニア語 9 母音に対する超音波画像から、舌上面をトレースした結果を示す(元の画像は Fig. 2 のような超音波画像のため、詳細は Fig. 2 を参照) :

