

論文・研究ノート

コミュニケーション・エイドとしての管楽器 ：電力の供給を必要としない音声合成器

池畑 光浩⁺、及川 靖広⁺⁺、永野 桃子⁺、山崎 芳男^{**}

1 はじめに

2004年3月3日にマレーシアマルチメディア大学にて開催されたワークショップにおいて、我々はイスラム科学の多様性とコミュニケーション技術の可能性について述べた¹。我々は快適な音環境の実現を目指し、「縦波コミュニケーション」をキーワードに基本原理に立ち返り、多機能・高能率なコミュニケーションシステムに関する研究を進めているが、中世のイスラム文化にも通ずることが多々ある。産業革命以来、我々は国内外の紛争の解決方法として科学技術を使用してきたが、古代ギリシャの理論に基づき技術を発展させ、一般市民のレベルで生活の質を向上させた中世イスラムの文化に学ぶべきことが多い。

例えばイスラム教徒が受け継いだ計算のシステムの中には指を使った計算方法があった。図1のように、それぞれの指の位置によって計算を処理する方法であり、右手の指は一の位と十の位を、左手の指は百の位と千の位を示している。両手を使用するこの計算方法は今日においても計算機に手慣れた人よりも速いことがある²。日本の「そろばん」を用いた暗算も、手の動きを利用して10桁を超える計算をも可能とした優れた計算方法であり、現代の日常生活においても大いに活用されている技術の一つである。

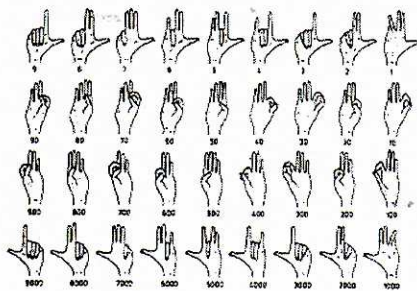


図1 指を使った計算の例を示している図解²

ところで、人の声はコミュニケーションにおいてテキスト表現される“文字”だけでなくイントネーションやアクセント、あるいは声の“かすれ”や“のび”と言った生理的な音響現象をも利用している。話し声はもちろん、歌や楽器なども人間のコミュニケーション手段の一つである。音楽はその表象の一つであり、人の声が多様な音色を生み出すことに長けた優れた楽器である。我々は多機能・高能率である人の声の発音原理に学び、人の声が出せる自然楽器の開発を行っている³⁻⁵。

指の動きや形とことばが連動したコミュニケーション・エイドとしての管楽器は、楽器本体のどの部位を押さえるかによって発音される音が決まるので、手の動きだけで音を想像することが出来る。また同時に、外部電力の供給を必要としないことも大きな特徴である。

西洋の楽器の歴史を振り返ると、それぞれの楽器がその音色を大切に継承しながら音域と音量の幅を増大させることに力を注いできた。前世紀の電子楽器の登場以来、様々な音色を表現するための楽器が登場したが、その一方では電力を用いないアコースティックな楽器が登場する余地がなくなってきた。本稿では優れた楽器の一つである人間の声に着目し、コミュニケーション・エイドとしての管楽器、つまり電力の供給を必要としない音声合成器の発音原理と開発経過について述べる。

2 発音原理

2.1 多機能かつ高能率である人の声

人の声は多機能であり、共振周波数を巧みに変化させられる楽器の一種と見る事ができ、我々は生み出された多様な音色をことばとして捉えている。音量に関しても表情豊かに調整することが可能で、歌手であれば pp から ff まで自由に調整する。人の声はまた高能率である。もともと肺と食道を分ける弁に過ぎなかった声帯を声の音源として転用しているし、音を生み出すための気流は呼吸における排気という、元々無駄であったものを利用しているのである。

このように多機能・高能率である人間の声を楽器開発に生かし、演奏者の声質や音域に関わらず歌を歌うことが出来る楽器、声帯を失った人でも歌を歌う感覚を味わえる楽器、シンギング・エイドとしての管楽器を開発した。本稿では図2のように人間の声の出る仕組みを単純化し、図3のように音源部・共鳴部・放射部の3つに分け、それぞれを過渡的に形状変化させられる機械的な仕組みを解説する。

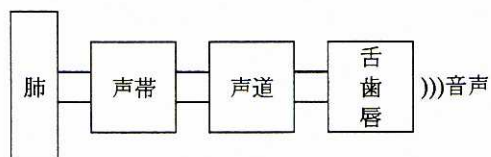


図2 人間の音声発生モデル

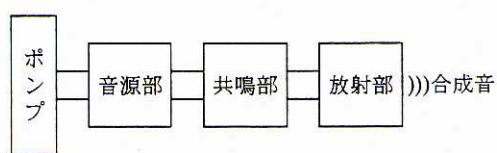


図3 音声合成器の発音モデル

2. 2 過渡的な形状変化が可能な声道物理モデル

機械系による発声物理モデルの研究は18世紀の Von Kempelen による Speaking Machine から現在まで様々に行われてきた。現在では声道の過渡的な形状変化による聴覚フィードバックに関する研究報告もある⁶。我々は電力を用いない楽器として音声合成器の開発を進めており、この管楽器は楽器としての用途だけではなく発話器としても用いることができる。

人間は自由に音色やピッチを変えながら声に様々な表情を持たせている。そのようなことが可能なのは人間が音源部・共鳴部・放射部を巧みにコントロールしているからである。従って理想の声道物理モデルに基づく音響管は音源部・共鳴部・放射部の全てが過渡的に形状を変化させられるものであり、「声の出る管楽器」はそのモデルに基づき開発を進めている。

また音源部においては、人間の歌声を発声しているときの声帯、笛式人工声帯、および管楽器の音源部を比較し、歌声を生成することが出来る自然楽器への応用を検討した。

3 音源部

3. 1 声帯の物理モデル

音源部では主に基本周波数を決め、声のピッチを生成する。母音の音源は声帯の振動であり、声帯は肺からの定常（直流）気流を周期的な気流パルス列に変換する。声帯振動を開始し維持するメカニズムは、解剖的な仕組み、筋肉による調整、流体の作用による非線形の振動、声道内圧によって生じる声門体積流 (glottal flow) と声帯運動との相互作用などからなる複雑な現象である⁷。

声帯の物理モデルには低次元のものから2質量モデル、3質量モデルのような高次のものがあるが、本稿では過渡的に音程と音量をコントロールすることを目標に音源を求めた。トロンボーンのように管の長さを調整し音程を取る方法もあるが、この楽器の場合は共鳴部と放射部の制御に手を用いたいため、口元だけで音程を調整できることが望ましい。笛式人工声帯を口の中に加え舌で弁を押さえて音程を取るという簡単な方法もあるが、笛式人工声帯では流入する空気の量のほとんどが音程を左右するので、クレッシェンドなどの音量の変化を作るには困難である。そこで最も単純な声帯の1質量モデルに基づく機械系での再現から開発を試みた。1質量モデルとは、左右の声帯組織をそれぞれ一つの質量で近似した側方変位のみモデルである。この楽器の音源部では、左右の声帯組織の運動を、唇または歯の上下運動により再現する。声門を完全に開いた状態で空気を流入すれば無声音を出すことが可能になる。

母音と子音、あるいは有声音と無声音との区別は主に音源の種類に依存することが大きいので音源部の役割は重要である。つまり声帯の開閉によって生み出されるパルス音によって有声音が合成され、声帯の非振動時に生ずる白色雑音によって無声音が合成されるので、音源部ではパルス音と白色雑音を切り替えられなければならない。また声帯の閉め具合によって過渡的に数オクターブもの音域で音程を変化させていることも人間の声が非常に優れている点である。

3. 2 開発した音源

音源に用いたのは伸縮可能なゴムであり、声帯音源波形に近い音を目指した。現段階では約4オクターブの音域を表現できる。図4、5に約900Hzの高い音と約60Hzの低い音の波形を示す。さらに図6に1オクターブ間における基本周波数と音階の変化の様子を示す。基本周波数に対する倍音が豊富な点や、微分音の表現も可能であるという点は、声帯の特徴に近い。

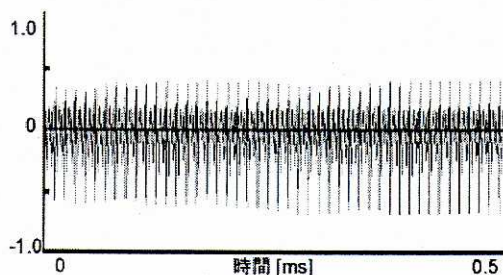


図4 高い音程 (約900Hz) の音源波形

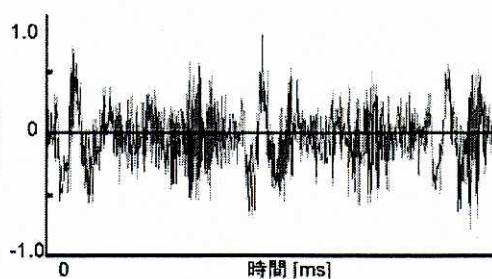


図5 低い音程 (約60Hz) の音源波形

この音源を管楽器に応用することで特定の音に限れば歌声に近い音を生み出すことができたが、より人間の歌声に近い音源を求めた。

Howardらは基本周波数変動と対応して声帯が閉じている(Closed Phase: CP)時間に着目したCQ値(Closed Quotient value)の分析を行った。様々な音声に関して声帯の振る舞いを観察した結果、図7のように歌唱の経験、あるいは訓練されたほど、声帯のCQ値が高いことを報告している⁸。また表1は歌唱の訓練・経験と話し声および歌声のCQ値の測定結果であるが、歌唱の訓練・経験とCQ値の相関が示されていると同時に、個々人の声でも話し声と歌声によってCQ値が違うということを示している。

以上を踏まえ、音声波形を声帯の開閉運動による二つの声道インパルス応答の重畳であると考え、音声波形のエネルギー減衰のみから声帯開閉周期を推定する手法⁹を用い、さまざまな音源を解析した。この手法は従来のElectrolaryngographを用いる手法とは異なり、発話者に測定器具を装着する必要がないので自然な発声を評価することができる。音源には笛式人工声帯、ダブルリードの管楽器、人間の話し声、および歌声の音声波形のエネルギー減衰曲線から声帯の開閉周期推定を試みた。人間の声に関しては、胸声、頭声、裏声(ファルセット)など出来るだけ様々な声に関して分析を行った。

図8は笛式人工声帯と声道模型VTM-10¹⁰を用いて発音された母音/a/の波形(上段)とエネルギー減衰率(下段)である。CPは声帯が閉じている段階であり、OPは声帯が開いている段階であることを示している。

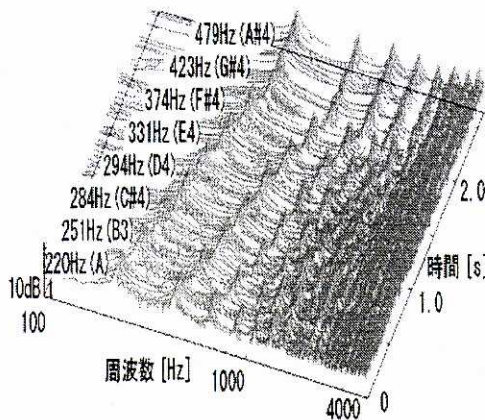


図6 1オクターブ間における基本周波数と音階の変化の様子

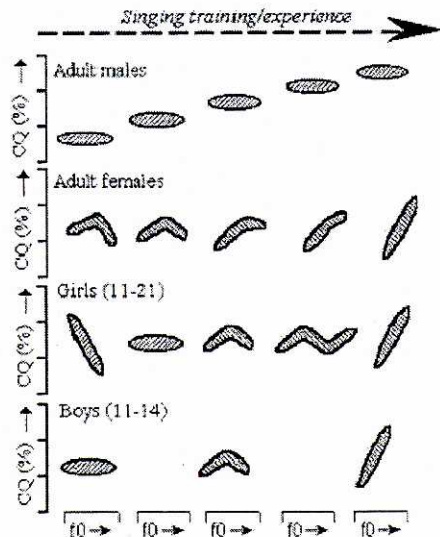


図7 歌唱経験・訓練および基本周波数によるCQ値の違い Adult data from Howard et al., (1990) and Howard (1995), children data from Barlow (1999)

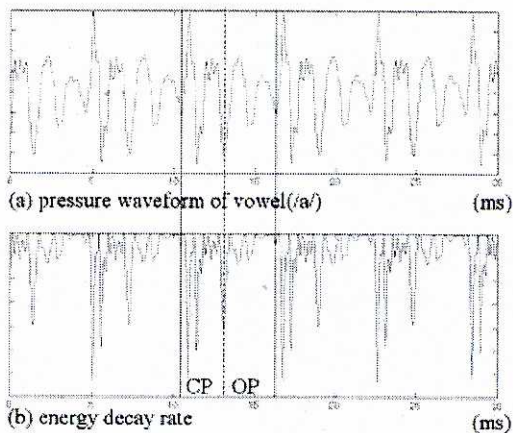


図8 笛式人工声帯による母音 /a/ の波形(上段)とエネルギー減衰率(下段)

表1 歌唱の訓練・経験と話し声 (speech) および歌声

subject	Singing background(years)		C Q (%)	
	Training	Experience	Speech	Scale
S1	6	30	57	63
S2	10	23	58	62
S3	10	13	52	60
S4	23	20	53	53
S5	1	15	43	51
S6	—	—	48	48
S7	2	11	46	48
S8	—	—	42	45
S9	—	5	46	45
S10	—	—	44	43
S11	—	—	33	43
S12	—	—	39	43
S13	—	2	43	43
S14	—	3	45	39
S15	—	—	41	38
S16	—	2	44	37
S17	—	2	44	37
S18	—	—	46	35

表2 笛式人工声帯 (tapia)、話し声 (speech) および歌声 (scale) の CQ 値

Sound Source (/vowel/)	C Q (%)
Scale (/e/)	62
Scale (/o/)	62
Scale (/a/)	60
Scale (/u/)	55
Tapia (/i/)	54
Speech (/u/)	53
Scale (/i/)	51
Speech (/o/)	50
Speech (/i/)	47
Tapia (/e/)	46
Speech (/e/)	45
Tapia (/i/)	45
Tapia (/a/)	45
Tapia (/o/)	44
Tapia (/o/)	44
Tapia (/i/)	43
Speech (/a/)	42
Tapia (/a/)	40
Tapia (/a/)	33
Tapia (/a/)	32

表2は笛式人工声帯(Tapia)、話し声(Speech)、および歌声(scale)のCQ値を示したものである。概ね笛式人工声帯よりも実際の話し声のほうがCQ値が高く、話し声よりも歌声のほうがCQ値が高いことが分かった。また、ダブルリード楽器の音源を用いて母音を生成した音源、および裏声の母音に関しては正確なCQ値を観測するに至らず、今後の課題とした。

4 共鳴部

共鳴部では主に母音を生成する。母音の共鳴部は声道であり、声帯で生成されたパルス（声門体積流）は声道の共鳴によって音響フィルタ処理が加えられ母音として生成される。声を楽器として見たときの最大の特徴は共鳴特性を自由に变化させられることである。電力を用いないアコースティックな楽器であり、且つ共鳴特性を变化させられる楽器としては口琴やディジリドゥーのようないわゆる民族楽器があるが、それらは結局人間の口の共鳴特性を利用している。

千葉・梶山は断面積が変化する不均一音響管を用いて共鳴周波数における声道内部の物理現象を論じ、声道物理模型を作った¹¹。これにより断面積が連続的に変化する管によって母音の生成が可能であることが示された。この原理を参考に、共鳴部にはシリコン製の柔らかい管を用いた。柔らかい管を外部から指で押さえることにより断面積を連続的に変化させ母音を生成する。共鳴部の過渡的な形状変化によって過渡的に母音を変化させられるのでウムラウトや中間母音の表現も可能である。なお、共鳴部はシリコンゴムで出来ており、管の直径は約4 cm、長さは約18cmである。図10に楽器の共鳴部の形状変化のみで合成された有声音母「あいうえお」に近い音のスペクトルを示す。

現段階では鼻腔がないので、鼻腔を用いて発声可能な母音の発音や、いわゆる声楽的な声、つまり頭部のあらゆる共鳴部を利用して音量と音色を豊かにするような表現も不可能である。今後はこのような表現を可能とするより高度な構造を導入することを検討している。



図9 共鳴部

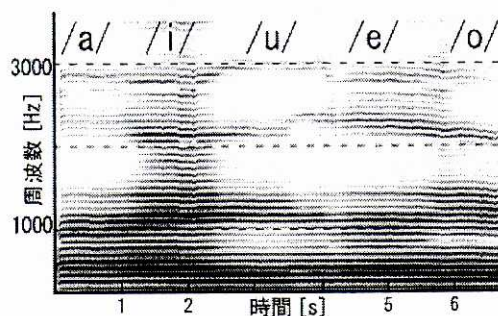


図10 生成された声のスペクトログラム

5 放射部

声道は口腔と鼻腔に分岐し口唇と鼻孔で外界へと放射される。口唇には舌・歯・唇がありここで主に子音が生成される。

舌・歯・唇の全てを機械系でコントロールするのは容易ではない。現段階では放射部に右手を用意して一部の子音が表現出来る段階である。また音源部を振動させずに息のみを用い、声道の形状を変化させることで雑音の周波数のある程度コントロールすることが出来るが、硬口蓋を用いる子音や破裂音、有声子音などの表現に達するには、より複雑な放射部の構造が必要である。

放射部が未完成なぶん、舌や唇の高度な調整が出来ない人間の赤ん坊が発音しやすいことばに近い音を再現することができた。ここでは音響管で生成した声「まんま」のサウンドスペクトログラムを図11に示す。

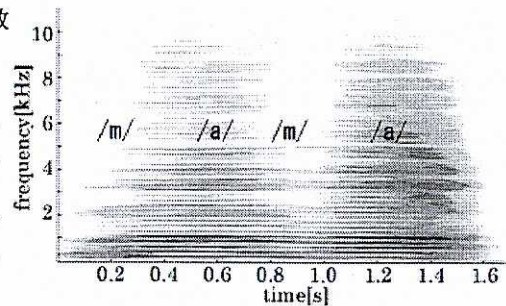


図11 音響管で生成した声「まんま」のサウンドスペクトログラム

6 むすび

人間の多様な発声法に学び、電力の供給を必要としない音声合成器の発音原理を音源部・共鳴部・放射部に分けて開発した。「あいうえお」などの母音の発音、放射部を工夫することによる /m/ や /p/ あるいは /b/ の子音の発音が可能である。また音源にトランペットのマウスピースを用いることにより赤ん坊の泣き声の表現が可能である。つまり「オギャー」「まんま」「パパ」といった、人間の幼児が発音可能な声の表現に達している段階である。

本稿では特に音源部について詳しく紹介した。人間の男声から女声を合わせた音域に相当する約60Hzから900Hzまでの間で1/12オクターブの音階を取ることが確認できた。振動部の完全開口時にホワイトノイズを発生することができるが、低い音程に既にそのホワイトノイズの成分が現れている。声帯の優れた機能に学ぶという点では声帯の特徴にまで十分に追求できなかったので多くの課題を残した。今後は声楽的な声やファルセットといったような様々な歌声の表現や、操作性などに関しても開発を続けていく。

また、音声波形のエネルギー変化のみから声帯開閉周期を推定する手法を用い、人間の歌声を発声しているときの声帯、笛式人工声帯、および管楽器の音源部を比較した。CQ値の観測のみによれば、笛式人工声帯の音は話し声に近く、歌声を発音できる自然楽器の音源に用いるには改良が必要である。今後は音源そのものをストロボ画像やレーザーを用いて観察し、歌声を生成することが出来る自然楽器に適した音源はどのようなものであるかを探る。

共鳴部に十分な体積があれば様々な母音を生成することが確認できたが、放射部との組み合わせによる破裂音や摩擦音などの生成に関して今後検討を加え、より高度な物理モデルに基づいた機械系での再現を目指す一方、操作性の用意な構造などについても考察していく。

なお、トーキング・エイドとシンギング・エイドのどちらがより高次のものであるかという問いに対しては容易に答えは出ない。「歌声」が「声」全体の一種であり単なる“特殊な例”なのか、それとも「声」の範疇とは別に定義されるのか、という問題意識から、音楽の本質・コミュニケーションの本質、さらには人の言語獲得・表現獲得への糸口を模索中である。

註

- 1 Kiyoshi OGINO, Mitsuhiro IKEHATA, Hisao HATTORI, Makoto, Yasuhiro OIKAWA, Momoko NAGANO and Yoshio YAMASAKI: Diversity and Possibilities of Islamic Sciences: The Quality-of-Life Islamic Science and Communication, Workshop on ICT Innovations and Their Impacts in Socio-economic and Cultural Environments - Dialogue in Borderless World and Islamic Context (2004.3)
- 2 Howard R. Turner, 久保儀明訳: 図説 科学で読むイスラム文化, 青土社, p. 83 (2001)
- 3 池畑光浩, 山崎芳男: 人間の多様な声に対応した音響管. 音講論集, pp. 663-664 (2002. 9)
- 4 M. Ikehata and Y. Yamasaki: Audio-visual and Tactile Communication by Voice Generator without Electric Power Supply. 画像電子学会予稿集, pp. 97-98 (2003. 6)
- 5 池畑光浩, 小西雅, 山崎芳男: “シンギング・エイド”としての管楽器, 音講論集, pp. 689-690 (2003. 9).
- 6 澤田秀之, 橋本周司, “声道モデルの機械系による実現とその計算機制御”, 研究報告「音楽情報科学」 アブストラクト No. 016-002. (2001)
- 7 Brand H. Stor, “母音の音源に関する生理と物理およびモデル化”, 日本音響学会誌, No. 58(7), pp. 426-437 (2002)
- 8 D. M. Howard, G. A. Lindsey, and B. Allen: Toward the quantification of vocal efficiency. Journal of Voice, Vol.4, No.3, pp.205-212, (1990)
- 9 高橋義典, 東山三樹夫, David M. Howard: 音声波形のエネルギー減衰に着目した声帯開閉周期の抽出. 信学技報, EA 2003-156, pp. 7-12 (2004. 3)

- 10 T. Arai: The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity. *Journal of the Phonetic Society of Japan*, Vol.5, No.2, pp. 31-38, 2001.
- 11 T.Chiba and M.Kajiyama : The Vowel its Nature and Structure. Tokyo-Kaiseikan Pub.Co.Ltd., 1941

(+ 早稲田大学・国際情報通信研究科・博士課程、++ 早稲田大学・国際情報通信研究センター・講師、

* 早稲田大学・理工学総合研究センター・研究員、** 早稲田大学・国際情報通信研究科・教授)