

タイトル	母音発話時の皮膚からの放射音声の測定
著者	佐藤, 高雄; 元木, 邦俊; SATO, Takao; MOTOKI, Kuni toshi
引用	北海学園大学工学部研究報告(38): 79-85
発行日	2011-02-14

母音発話時の皮膚からの放射音声の測定

佐藤 高雄*・元木 邦俊*

Measurement of Sound from Skin Vibration for the Vowel Utterance

Takao SATO* and Kunitoshi MOTOKI*

要 旨

近年、MRIデータ等を元に構築された声道の3次元形状モデルを用いた広い周波数帯域に渡る音響解析が行われている。一般的な音声生成過程のモデルでは主に口唇と鼻孔からの放射音声を対象としている。音声生成モデルの高度化のためには、皮膚振動による放射音声も考慮すべきものと考えられる。本稿では予備的実験として皮膚からの音響放射と皮膚振動を直接測定した結果について述べる。

1 はじめに

近年、MRI (Magnetic Resonance Imaging: 磁気共鳴映像法) データ等を元に構築された声道の3次元形状モデルを用いた広い周波数帯域に渡る音響解析が行われている^{1,2)}。一般的な音声は声帯の振動によって生じた振動が声門から口唇までの空間(声道)を伝搬して口唇や鼻孔の開口端から放射された音波である。声道の3次元形状モデルでは主に口唇と鼻孔の開口端からの放射音声を対象としているが、実際の音声は頬や頸部等の皮膚振動に伴い開口端以外からも放射されていることが知られている。より高精度な音声生成シミュレーションを行うためには壁振動(皮膚振動)による放射音声の影響も考慮すべきものと考えられる。

本稿では、壁の振動特性を音声生成モデルに組み入れることを目的とし、開口端以外からの放射音声、及び皮膚振動の測定について報告する。

2 頸部皮膚振動による放射音声の測定

本節では頸部の皮膚振動によって生じる放射音声の測定を行った結果について述べる。

* 北海学園大学大学院工学研究科電子情報工学専攻

* Graduate School of Engineering (Electronics and Information Engineering), Hokkai-Gakuen University

2.1 測定方法

音声は口唇や鼻孔の開口端、及び頸部などの皮膚振動によって生じるものである。頸部の皮膚振動による音声のみを記録するため、**図1**に示すように被験者の口唇と鼻孔の開口端にアクリルパイプをあてがい、開口端からの放射音声の影響を受けないようにして記録を行い、通常の発話音声との比較を行った。

測定系は**図1**に示すものであり、放射音声の測定は無響室内の静音下で行われた。開口端からの放射音はアクリル製の管を開口端に密着させ他端はホースを介して室外に開放した。音声の記録には精密騒音計（リオン製：NA-40）のマイクロホンを、**図2**に示すように被験者の正中矢状面より水平方向に45°傾けた方向に設置し、頸部とマイクロホンの距離 d を変化させながら日本語の5母音/a/, /e/, /i/, /o/, /u/の記録を行った。録音データは量子化ビット数16 bit、サンプリング周波数48kHzのモノラルWAVE形式データであり、被験者は40代男性1名である。なお、被験者は口に管を当てずに発話した際、口の正面50cmに設置された精密騒音計が概ね80dBを示す強度になるよう留意して発話した。

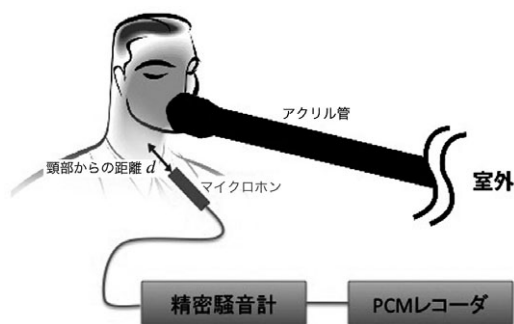


図1 測定系概要①

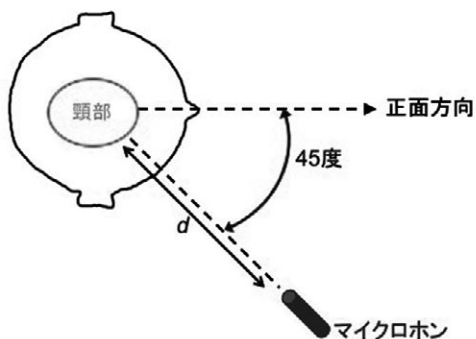


図2 測定系概要②

2.2 母音ごとの音圧レベル

距離 d の変化による母音ごとの音圧レベルの変化を表1に示す。母音により放射音圧レベルに相違が認められ、/i/が最も強く/a/が最も弱い。/i/の第1ホルマントは成人男性の場合一般に250~350Hzとされ、日本語の5母音の中では最も低い、また/a/の第1ホルマントは600~800Hzとされ最も高いことが知られている。今回測定された5母音の放射音圧レベルと第1ホルマントの周波数の関係とを比較すると、第1ホルマントの周波数が低い母音ほど頸部からの放射音圧レベルは大きいという結果が得られた。このような特徴は先行研究³⁾の結果ともよく一致している。

また距離 d の変化に対する音圧レベルの変化は母音ごとに異なっている。20mmから50mmへの距離変化、50mmから100mmへの距離変化に対する音圧レベルの変化量を倍率にすると各々2.5倍、2倍であり、音源が点音源であればそれぞれ-8dB、-6dBの距離減衰が生じる。表1か

らわかるように50mmから100mmへの変化においては母音*/i/*以外は概ね距離減衰に従った音圧レベル変化が見られる。

表1 母音ごとの音圧レベルの変化 単位：[dB SPL]

距離 <i>d</i> [mm]	<i>/a/</i>	<i>/e/</i>	<i>/i/</i>	<i>/o/</i>	<i>/u/</i>
20	76.7	78.0	93.3	81.5	89.7
50	66.3	65.2	81.0	70.1	73.3
100	60.1	59.3	71.2	63.5	68.2

2.3 スペクトルの比較

図3に母音*/i/*のスペクトルを示す。黒色の実線は通常の発話による放射音声（口からマイクロホンまでの距離50cmで、音圧レベルは概ね80dB SPL）であり、灰色の実線は頸部からの放射音声（2.1節に示す方法で記録し、頸部からマイクロホンまでの距離*d*=50mmの場合）である。頸部からの放射音声では2 kHz以上の高域で大きな減衰が認められた。これは音声の肉伝導において皮膚がローパスフィルタとして機能していることを示唆している。

図4～7に他の母音のスペクトルを示す。図4、5においては図3と同様に高域での大きな減衰が見られた。図6、7に示した母音*/o/*、*/u/*においては他の母音ほどその傾向は顕著では無い。これは他の母音に比べ高域のホルマントのピークが低く、相対的に特徴が現れにくかったものと考えられる。

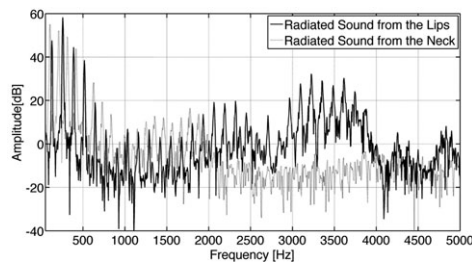


図3 */i/* のスペクトル

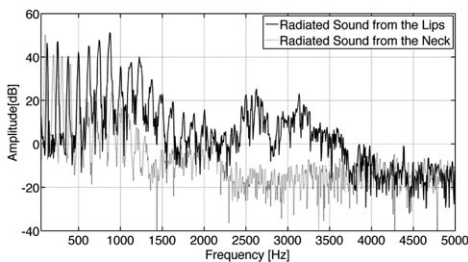


図4 */a/* のスペクトル

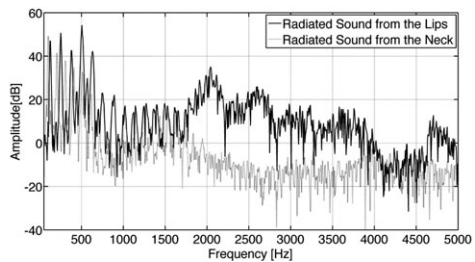


図5 */e/* のスペクトル

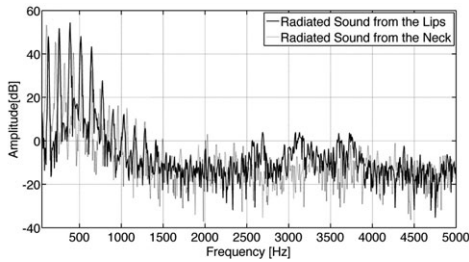


図6 /o/ のスペクトル

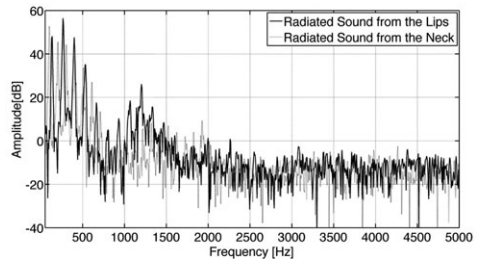


図7 /u/ のスペクトル

3 発話時における皮膚振動の測定

母音発話時の頸部皮膚振動を音声生成モデルに組み入れるためには、放射音声の音源となる振動部分の強度や分布等の特性を測定する必要がある。以下では基礎データの収集を目的として頸部の皮膚振動を直接測定した結果を述べる。

3.1 測定方法

測定は小型の加速度センサーを被験者の皮膚に直接貼り付ける方法に加え、レーザー変位計を用いた非接触による変位測定法を用いた。実験系は図8に示す構成であり、実際の機材の配置例を図9に示す。また、測定に使用した加速度センサー、及びレーザー変位計の諸元を表2

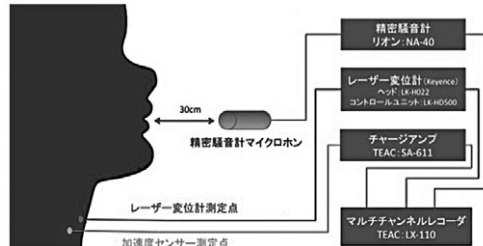


図8 実験系概要③

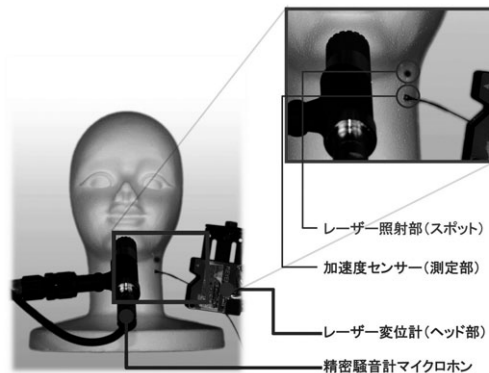


図9 実験系概要④

表2 使用機器諸元

加速度センサー		レーザー変位計 (ヘッド部)	
型番(メーカー)	MODEL611 (TEAC)	型番(メーカー)	LK-H022 (KEYENCE)
形式	圧電型加速度トランスデューサ	レーザー波長	650[nm]
測定レンジ	$\pm 1.0 \times 10^3$ [m/s ²]	出力	0.95[mW] ※クラスIIレーザ
質量	0.2[g] ※ケーブル含まず	最小分解能	0.02[μ m]
感度	0.035 \pm 20% [pC/(m/s ²)]	測定レンジ	ヘッド先端より20mmの位置を基準とし、基準位置より \pm 3mm
周波数帯域	0.2~100000[Hz]		

に示す。頸部（咽頭部）に加速度センサーを接着し、およそ10mm離れた位置をレーザー変位計によって測定した。被験者は母音を発話する際、口の正面方向30cmに設置された精密騒音計の音圧レベルが概ね80dBとなるよう留意して発話している。また被験者の粗動による変位の誤差を考慮し、レーザー変位計の出力にはカットオフ周波数80Hzのハイパスフィルタを挿入している。

データの採集は16チャンネルのマルチチャンネルレコーダによって行った。量子化ビット数は16bit、サンプリング周波数は48kHzであり、本稿記載の測定においては3チャンネルのみを使用した。なお、被験者は2節の実験と同一の40代男性1名、記録した音声は日本語母音/a/, /e/, /i/, /o/, /u/である。

3.2 測定結果

測定された母音ごとの変位 (RMS値)、さらに2.2節に示した母音ごとの放射音圧レベル ($d=50$ mm) と各母音の第1ホルマント周波数⁴⁾を表3に示す。母音により皮膚振動の変位に相違が認められ、/i/が最も大きく/a/が最も小さいことがわかる。第1ホルマントが低い母音ほど発

表3 各母音の特性

	/a/	/e/	/i/	/o/	/u/
変位 [μ m]	15.3	16.3	28.4	18.2	27.0
音圧レベル [dBSPL]	66.3	65.2	81.0	70.1	73.3
第1ホルマント周波数 [Hz]	600~800	350~550	250~350	420~500	300~480

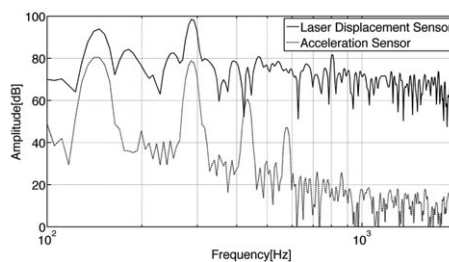


図10 /i/ のスペクトル

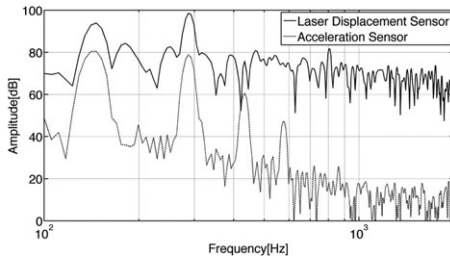


図11 /a/ のスペクトル

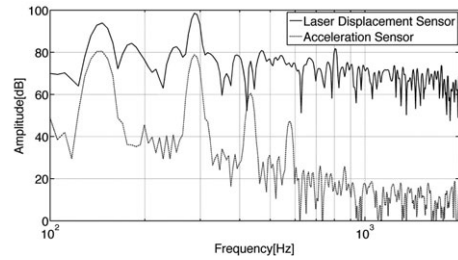


図12 /e/ のスペクトル

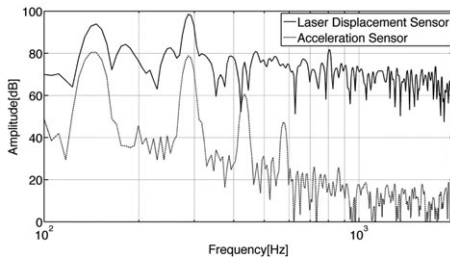


図13 /o/ のスペクトル

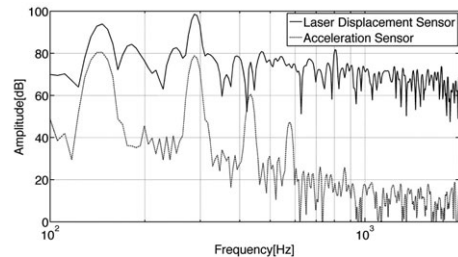


図14 /o/ のスペクトル

話時の頸部皮膚振動変位は大きい傾向があり、このような特徴は先の2.2節に示した結果と一致している。

次に、母音/i/を測定した加速度センサーの信号を積分（傾き補正-6dB/oct.）して得られた速度と、レーザー変位計の信号を微分（傾き補正+6dB/oct.）して得られた速度を図10に示す。実線はレーザー変位計の信号、破線は加速度センサーの信号を表しており、両者が重ならないよう縦軸方向にシフトして描画している。

レーザー変位計による非接触測定においても第2高調波までの調波構造が確認できるが、高域ではS/N比が悪化し母音のホルマント構造を確認するには至らなかった。また、加速度センサーによる測定では第4高調波までを観測できたが同様に母音のホルマント構造を確認するには至らなかった。他の母音のスペクトルを図11~14に示す。これらの結果から母音ごとの差異は確認できるが母音のホルマント構造を確認するには至らず、測定点の変更や被験者の発話方法の変更などによるS/N比の向上が必要と考えられる。

4 おわりに

本稿では頸部からの放射音声、及びその音源となる頸部皮膚振動の測定を行った。2節では頸部からの放射音声と通常の発話音声と比較し、両者の特徴、及び母音ごとの特徴を述べた。3節では小型加速度センサー、及びレーザー変位計による非接触測定法により頸部皮膚振動の母音ごとの差異を観測できることを示し、今後の課題として測定精度の向上を挙げた。頸部皮膚振動による放射音声の影響は音声生成において重要な要素であり、今後はより詳細な検

討が必要と考えられる。

本研究の一部は、北海学園大学ハイテクリサーチセンター（戦略的研究基盤形成支援事業）による補助を受けて行なわれた。

参考文献

- 1) H. Matsuzaki and K. Motoki: “Study of acoustic characteristics of vocal tract with nasal cavity during phonation of Japanese /a/”, *Acoust. Sci. & Tech.* **28**, 2, 124–127 (2007).
- 2) K.Motoki and H.Matsuzaki: “Computation of the acoustic characteristics of vocal-tract models with geometrical perturbation”, *Proc. INTERSPEECH2004–ICSLP*, TuB602p.16, 521–524 (2004).
- 3) 党, 中井, 鈴木: “音声発声時の口唇, 鼻孔, 咽頭部の音と振動の測定および非鼻音化音声における鼻腔口腔間の音響漏えいのシミュレーション”, *電子情報通信学会論文誌*, **J74-A**, 12, 1705–1714 (1991).
- 4) 早坂寿雄: 技術者のための音響工学, 丸善 (1986).
- 5) 佐藤, 元木: “声道壁振動による放射音声の測定”, 平成21年度電気・情報関係学会北海道支部連合大会, 17 (2009).
- 6) 佐藤, 元木: “母音発話時の声道壁皮膚振動の非接触測定”, 平成22年度電気・情報関係学会北海道支部連合大会, 183 (2010).