

タイトル	母音発話時の頸部皮膚振動の非接触測定
著者	高川, 昂太; 元木, 邦俊; TAKAGAWA, Kouta; MOTOKI, Kunitoshi
引用	北海学園大学工学部研究報告(39): 143-151
発行日	2012-02-14

# 母音発話時の頸部皮膚振動の非接触測定

高川 昂太\*・元木 邦俊\*

## Measurement of Skin Vibration of Neck by Noncontact Method

Kouta TAKAGAWA\* and Kunitoshi MOTOKI\*

### 要 旨

近年、音声合成技術は自動応答システムなどの様々な場面で使用されている。人の音声生成過程に基づいてより自然な音声を合成するには、高度な音声生成モデルが必要となる。音声生成モデルの高度化のためには、頸部皮膚からの放射音声について検討することが必要と思われる。本稿では、頸部皮膚及びその近傍の振動速度を2つの手法を用いて非接触で測定した結果について報告する。

## 1. はじめに

近年、音声合成技術は自動応答システムなどの様々な場面で使用されている。人の音声生成過程に基づいてより自然な音声を合成するには、高度な音声生成モデルが必要となる。従来の音声生成モデルは口唇と鼻孔からの放射音声のみを考慮している。しかし、実際の音声は声道壁の振動により頸部皮膚からも放射される。口唇と鼻孔からの放射音声だけでは、声道壁から漏れるバズバーなどの音を表現しきれないと考えられる。音声生成モデルの高度化のためには、頸部皮膚からの放射音声について検討することが必要と思われる。放射音声を測定する方法としては加速度センサを用いた方法が考えられる。しかし、この方法では加速度センサを皮膚に貼り付けるため、皮膚振動変位に影響があると思われる。ここでは頸部皮膚及びその近傍の振動速度を2つの手法を用いて非接触で測定した結果について報告する。始めにレーザー変位計を用いて皮膚振動特性を測定する方法について述べる。この方法は発話を行っている被験者の頸部に直接レーザー光を当てることにより、皮膚の振動変位を測定する。次に音圧の空間

---

\* 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

分布から測定する方法について述べる．この方法は被験者の頸部皮膚近傍に小型マイクロホンを複数設置し，音圧の空間分布から測定する方法である．これらの方法では，皮膚に非接触で測定できるので，本来の皮膚振動特性の測定が可能だと考えられる．

## 2. レーザー変位計による皮膚振動の非接触測定

頸部皮膚の放射音声を求めるためには皮膚の振動速度を測定することが考えられる．ここでは，レーザー変位計により得られる皮膚の変位信号を用いて振動速度を求める．

### 2.1 測定方法

図1に頸部皮膚の振動特性を求めるための測定系を示す．皮膚振動の変位を求めるためにレーザー変位計LK-H022（KEYENCE社）を使用した．最小分解能は $0.02\mu\text{m}$ ，測定レンジはレーザー変位計のヘッドの先端から20mmの基準位置から $\pm 3\text{mm}$ である．使用しているレーザーの安全基準はクラス2レーザーである．発話時の音圧レベルを計測するために精密騒音計NA-40（リオン社）を使用した．これらの信号はマルチチャンネルレコーダーLX-110（TEAC社）に記録した．量子化ビット数は16bit，サンプリング周波数は48kHzで測定を行った．なお，レコーダーは複数の信号を同時にサンプリングして記録可能である．

被験者（成人男性1名）は5母音/a/, /i/, /u/, /e/, /o/を連続して発話した．口唇から精密騒音計のマイクロホンまでの距離 $d_1=300\text{mm}$ になるように配置した．仰向けの状態で精密騒音計の音圧レベルが80dBになるように発話練習を行い，その後，測定を行った．レーザー変位計の測定レンジは $\pm 3\text{mm}$ と小さいので，被験者は僅かな動きで測定レンジから外れないように注意して発話を行った．また被験者自体の動きの影響を軽減させるために変位計からの出力信号に遮断周波数80Hzのハイパスフィルターを通した．

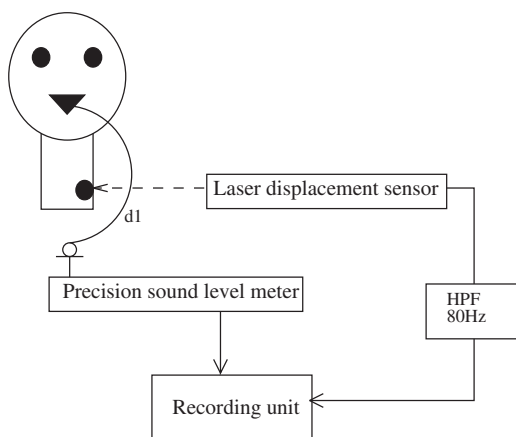


図1：測定系の概要

## 2.2 結果

測定時の/a/, /i/, /u/, /e/, /o/の発話強度は、それぞれ79.3, 72.8, 72.7, 77.6, 74.6 dBであり、一定の発話強度を保つことが困難であることがわかる。図2に振動速度のスペクトルを示す。この信号はレーザー変位計からの変位信号をFFTし、周波数領域で微分操作 ( $j\omega$  の乗算) をした後に得られたスペクトルである。図2を見ると1000Hz程度までの周波数で明確に調波構造が見られる。しかし、/i/では500Hz程度で高域成分が著しく減衰している。これ

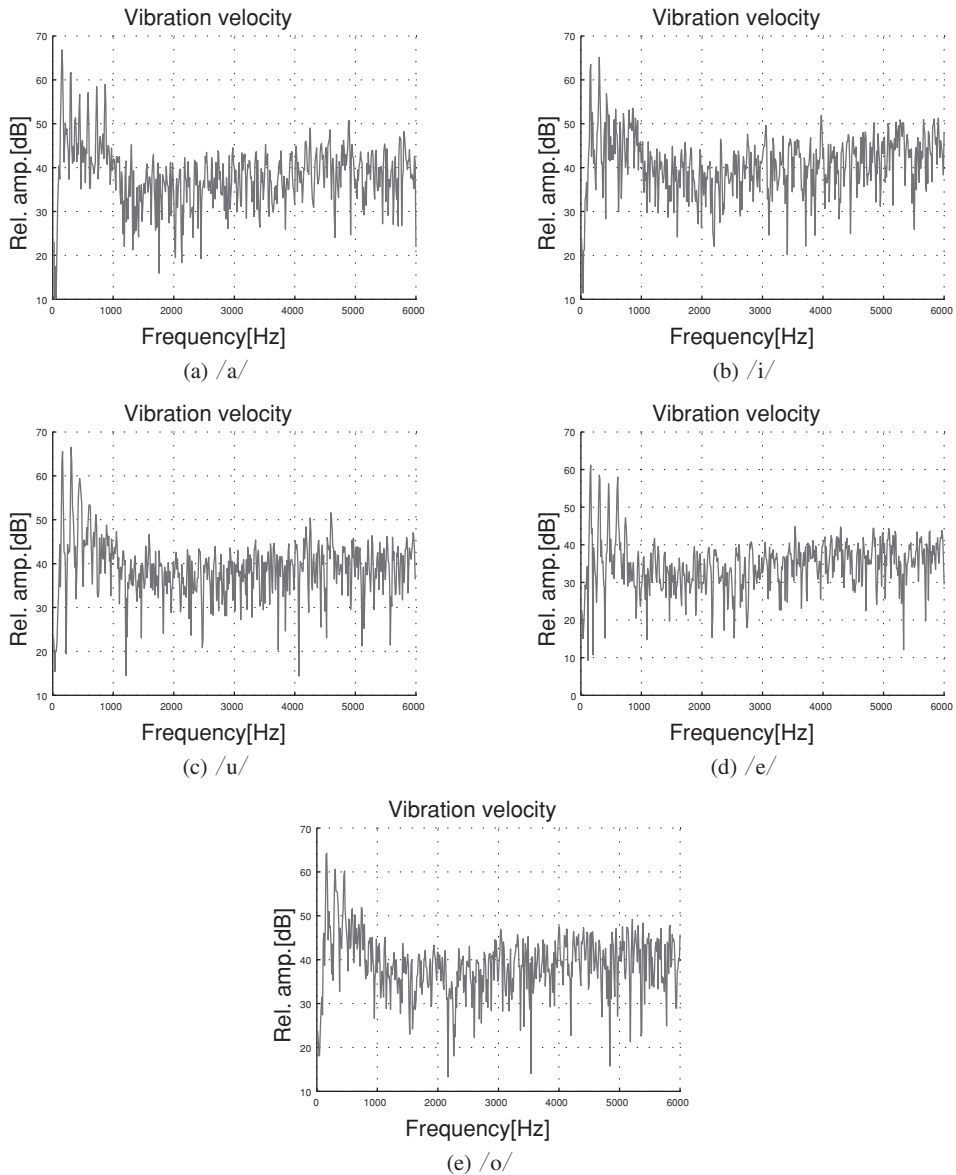


図2：レーザー変位計により得られた母音発話時の振動速度スペクトル

は他の母音よりも発話強度が弱いことが影響していると考えられる。口唇の音声と頸部皮膚からの音声を分離して測定すると頸部皮膚からの音声は高域が減衰する結果が得られており<sup>1)</sup>、これと同様の傾向が見られる。

### 3. 小型マイクロホンによる皮膚近傍の粒子速度の測定

頸部皮膚近傍の媒質粒子（空気）は皮膚振動と同じ動きをしていると考えられる。そこで粒子速度を測定することにより皮膚振動特性を求めることができる。しかし粒子速度を直接測定することは困難なので、ここでは小型マイクロホンを用いてマイクロホン間の音圧傾度から粒子速度を測定する。

#### 3.1 粒子速度と音圧の関係

音波伝搬の基本式として運動方程式は以下のように表される<sup>2)</sup>。

$$\rho \frac{\partial v(x,t)}{\partial t} = -\frac{\partial p(x,t)}{\partial x} \quad (1)$$

ここで $v(x,t)$ は粒子速度、 $p(x,t)$ は音圧、 $\rho$ は空気密度、 $x$ は位置座標を示す。 $V(x,\omega)$ 、 $P(x,\omega)$ を $v(x,t)$ 、 $p(x,t)$ のフーリエ変換とすると、運動方程式のフーリエ変換より次式が得られる。

$$V(x,\omega) = -\frac{1}{j\omega\rho} \frac{dP(x,\omega)}{dx} \quad (2)$$

ここで $\omega$ は角周波数を表す。式(2)より、音圧傾度 $dP(x,\omega)/dx$ がわかれば粒子速度 $V(x,\omega)$ が求められる。そこで音圧傾度を次のように差分で近似する。

$$\frac{dP(x,\omega)}{dx} \approx \frac{P_2(\omega) - P_1(\omega)}{\Delta x} \quad (3)$$

$P_1(\omega)$ 、 $P_2(\omega)$ は近接した2点間の音圧を表し $\Delta x$ は2点間の距離を表している。音圧傾度を差分で近似する場合は $\Delta x$ が波長より十分に小さい必要がある。

#### 3.2 測定方法

マイクロホンは、小型のエレクトレットコンデンサマイクロホンFG-23652-P16（Knowles Acoustics社）を使用した。このマイクロホンは直径が2.56mm、受音部が1.41mm、感度-53±3dB、周波数100～10kHzである。マイクロホンは図3のようにアクリル棒に固定した。測定系の概要を図4に示す。頸部に近いマイクロホンから順にM1、M2、M3としている。頸部からM1までの距離 $d_1=10\text{mm}$ 、M1からM2までの距離 $d_2=10\text{mm}$ 、M2からM3までの距離 $d_3=25\text{mm}$ となるように配置した。精密騒音計のマイクロホンは2.1節の実験と同じく口唇前方から300mmのところに配置した。粒子速度を求めるための2点の音圧の組み合わせはM1とM2、M2とM3、及びM1とM3の3通りの組み合わせで計算を行った。これによりアクリル

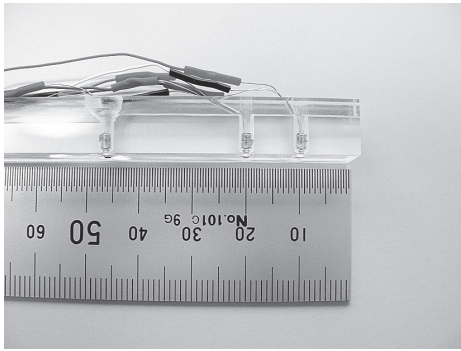


図 3 : マイクロホン固定の様子

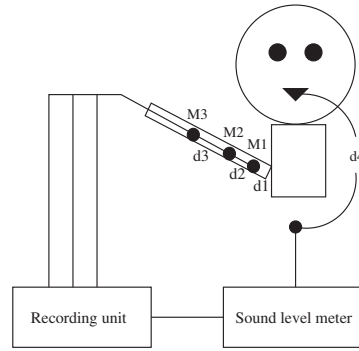


図 4 : 測定系の概要

棒方向の粒子速度を測定することができる。マイクロホン間の距離  $\Delta x$  はそれぞれ 10, 25, 35 mm となる。  $\Delta x$  が 35mm の場合、周波数 4000Hz において約 0.4 波長となる。発話音声と発話強度は 2.1 節の実験と同じ 5 母音で 80dB 程度で測定した。一定の発話強度を保つために被験者は次節に示す発話強度を表す LED 表示器を見ながら発話を行った。

### 3.3 一定の発話強度を保持するための補助装置

測定の際には一定の発話強度を保持することが望ましいが、指定された音圧で一定の発話強度を保つのは困難である。そこで被験者が一定の発話強度を維持することを容易にするための補助装置を作成した。装置の概要を図 5 に示す。この装置は音声信号の実効値を求め、適切な音圧レベルの時に大型の LED を点灯させる。被験者はこの LED の点灯状態を見ることによって、一定の音圧レベルになるように発話強度を調整する。以下に各部について説明する。

#### 3.3.1 RMS 変換回路と LED 表示器

精密騒音計の出力は交流信号電圧として得られているので、この信号の実効値 (RMS) を求めるために AD376 (アナログデバイセズ社) を使用して RMS 変換回路を作成した。

LED 表示には 24 ポイント連結型 LED バーグラフメーター AE-LM24LED-SANSAI (秋月電子通商) を使用した。しかし、これのみではどの程度の音圧を出力しているのかを視覚的に読み取るには不十分だったので、新たに 3 つの大型の LED 表示回路を作成した。LED の点灯状態と発話強度の対応関係は表 1 のようになっている。

$X, Y, Z$  をそれぞれオレンジ, グリーン, レッドの LED の点灯状態を表す論理値とすると、次の論理式で表される。

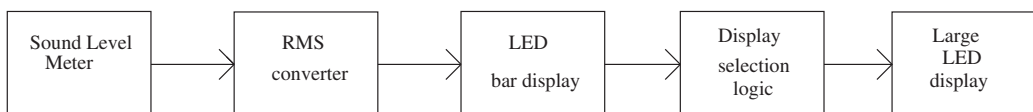


図 5 : 一定の発話強度を保持するための補助装置のブロック図

表 1：発話強度とLEDの点灯状態の対応関係

かなり弱い (~75.8dB)	消灯
少し弱い (75.8~78.4dB)	オレンジ点灯
適切な発話強度 (78.4~81.5dB)	グリーン点灯
少し強い (81.5~83.2dB)	レッド点灯
かなり強い (83.2dB~)	全点灯

$$X = A \cdot \overline{B} \cdot \overline{C} \cdot \overline{D} + A \cdot B \cdot C \cdot D \quad (4)$$

$$Y = A \cdot B \cdot \overline{C} \cdot \overline{D} + A \cdot B \cdot C \cdot D \quad (5)$$

$$Z = A \cdot B \cdot C \quad (6)$$

ここで  $A, B, C, D$  はLEDバールグラフメーターの出力において、それぞれ75.8, 78.4, 81.5, 83.2dB以上で1となる論理値である。この論理式から、表1のLEDの点灯状態になるように表示ロジックを作成した。

### 3.3.2 RMS値測定結果

被験者は仰向けの状態で5母音を80dBになるように発話を行った。精密騒音計のマイクロホンは口唇前方から約300mmの所に配置した。図6に母音を定常的に発話し、補助装置を使った時と使わない時の発話強度の時間変化を示す。図6を見ると補助装置を使わない時は、時間が経過するにつれ発話強度が低下する傾向にあるが、補助装置を使った時は、80dBに合せようとしている様子が伺え、一定の発話強度を保持することができている。

### 3.4 マイクロホンの補正

音圧傾度を求めるには位相と振幅が同じ特性のマイクロホンが必要となる。M1を基準にし

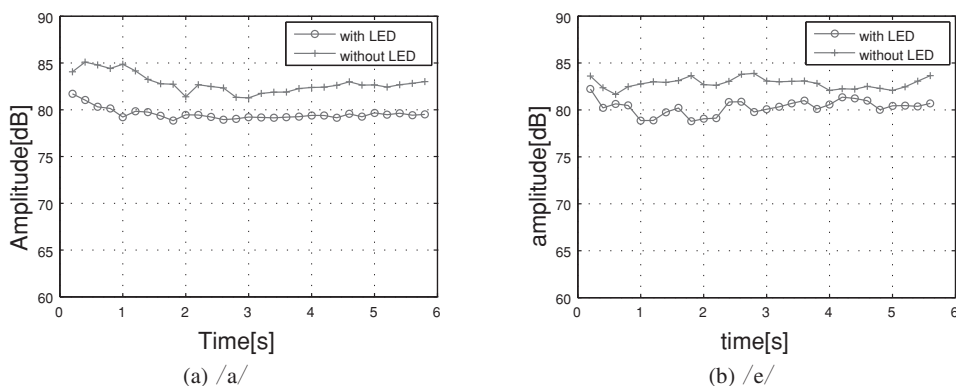


図 6：母音発話時のRMS値の変動

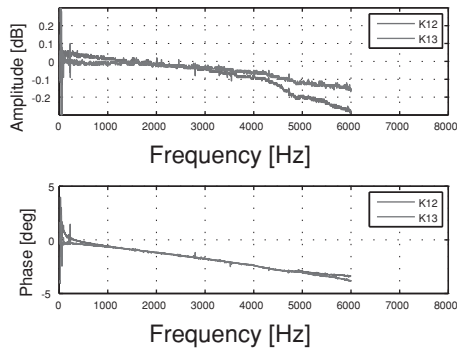


図 7：マイクロホンの振幅差と位相差

で相対的にマイクロホンの補正を行った。図 7 に振幅差と位相差を示す。ここで  $K_{12}$ ,  $K_{13}$  はそれぞれ M1 と M2, 及び M1 と M3 の振幅差と位相差を表している。3000Hz の時, 振幅差が 0.05dB 程度, 位相差が 1.8 度程度になっている。この特性を元にマイクロホンの補正を行った。

### 3.5 結果

図 8 にマイクロホン M1, M2, M3 から得られた /a/, /e/ の音圧スペクトルを示す。M

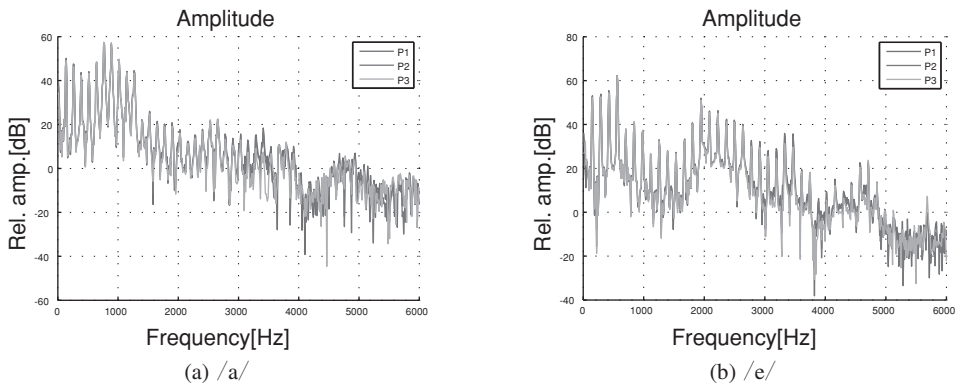


図 8：母音発話時の音圧スペクトル

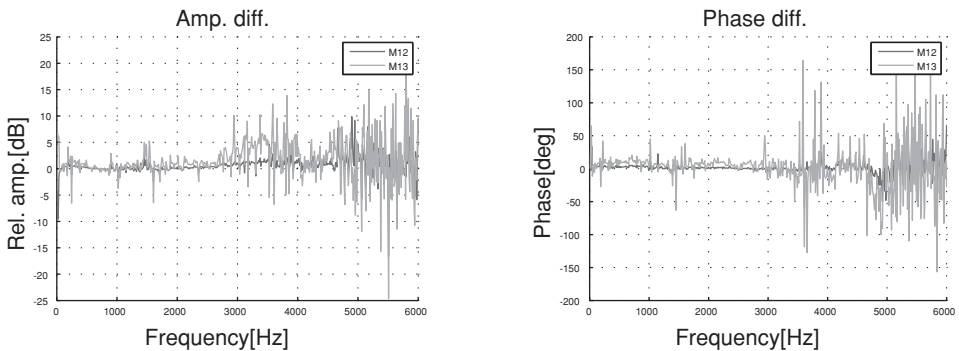


図 9：マイクロホン間の振幅差と位相差



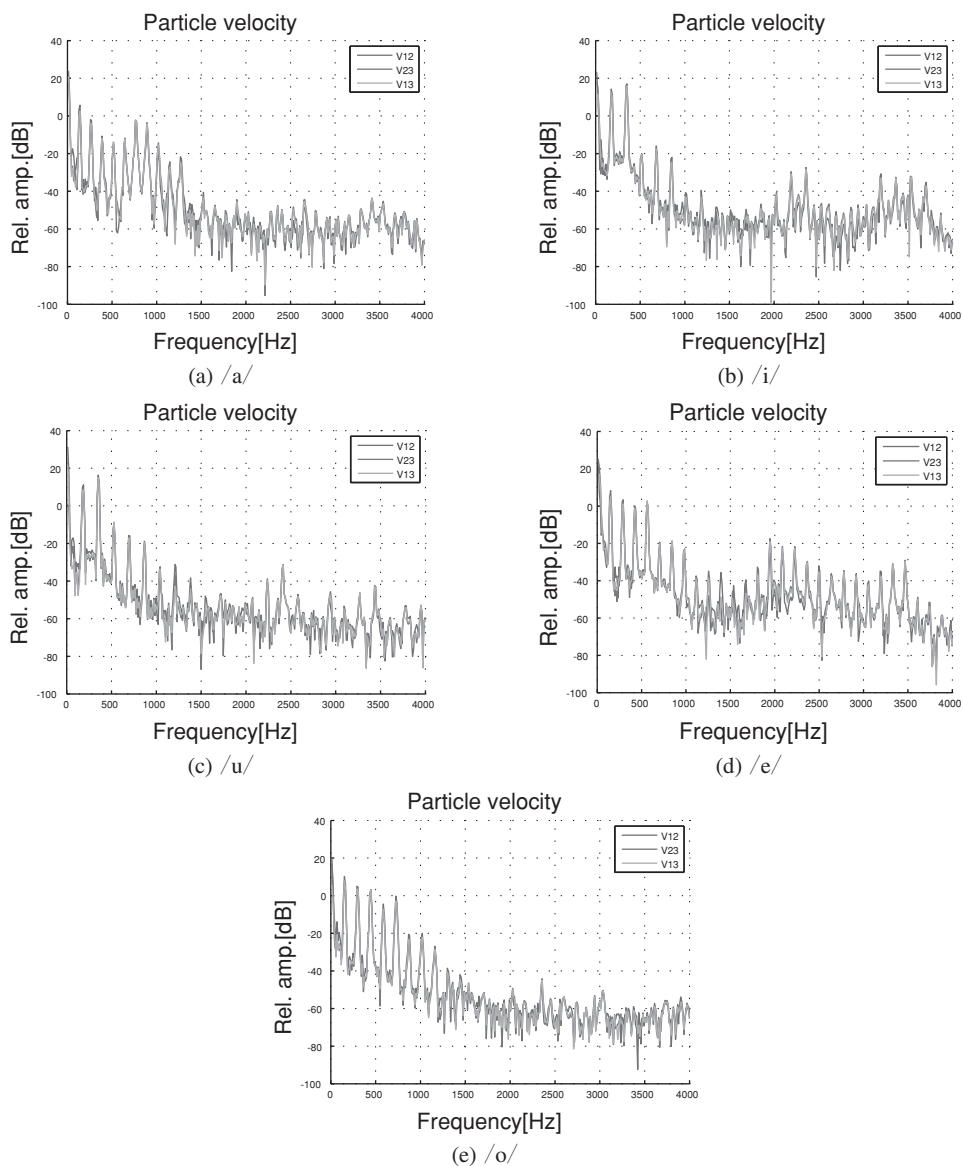


図10：音圧傾度から求められた母音発話時の粒子速度スペクトル

2, M3のスペクトルについては3.4節の補正を行っている。調波構造が明確に見られる部分では振幅と位相がよく一致していたが、3000Hz程度の周波数になると振幅差と位相差が大きくなっていることがわかる。図9に/e/発話時のM1とM2及びM1とM3のマイクロホン間の振幅差と位相差を示す。M1, M3の方がマイクロホン間の距離が長いので、振幅差と位相差がともに大きくなっている。また周波数が3000Hz程度から振幅差と位相差が大きくなっている。図10に式(2)から求めた母音ごとの粒子速度スペクトルを示す。V<sub>12</sub>はM1とM2の音圧

から求めた粒子速度スペクトルである。  $V_{23}$  と  $V_{13}$  も同様に求められた粒子速度スペクトルである。 1000~1500Hz程度までは明確に調波構造が見ることができるが、それより高い周波数ではスペクトルが大きく減衰している。 図2のレーザー変位計により求められた振動速度スペクトルと比較して、高域まで調波構造を見ることができる。

#### 4. まとめ

本稿では皮膚振動特性の非接触測定について述べた。 レーザー変位計を用いた測定では、高域が大きく減衰することや一定の発話強度を保持する必要があることがわかった。 そのため被験者が発話する際の頸部の固定の方法、発話方法などを工夫していく必要がある。 小型マイクロホンを用いた音圧傾度からの測定では発話強度を一定にするための補助装置を用いて測定を行った。 レーザー変位計よりも高い周波数域まで調波構造が確認できた。 今後はマイクロホン間の距離やアクリル棒の向きを変えるなど条件を変えて測定を行う必要があると考えられる。

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)による援助を受けて行われた。

#### 参考文献

- 1) 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).
- 2) P. M. Morse and K. U. Ingard: *Theoretical acoustics*, McGraw-Hill (1968).