

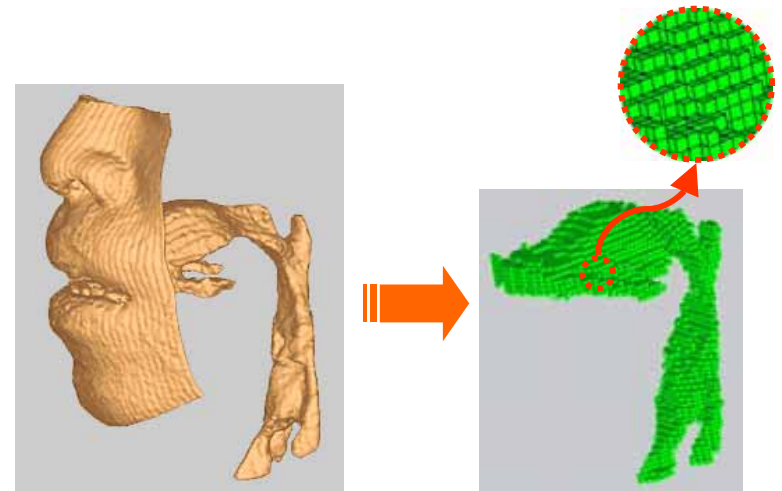
# ボクセルメッシュを用いた有限要素法による声道の音響解析

ATR人間情報科学研究所


北村達也 竹本浩典 本多清志

kitamura@atr.jp

本研究は情報通信研究機構の研究委託「人間情報コミュニケーションの研究開発」により実施したものである。

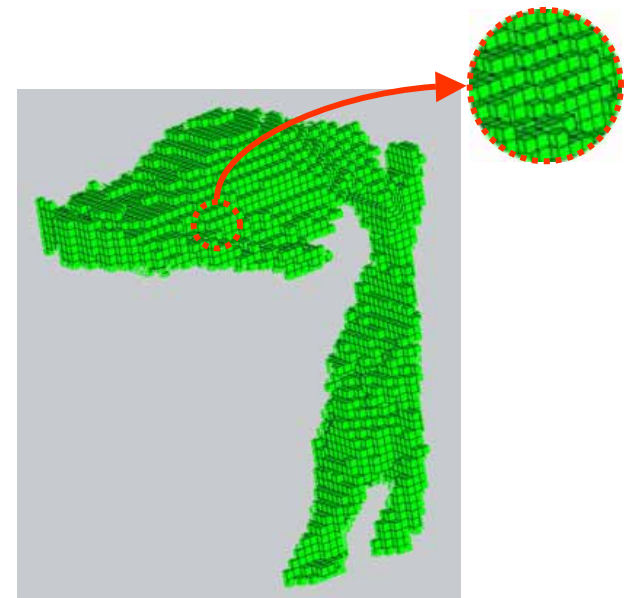


# 研究の背景と目的(1)

- 核磁気共鳴画像法(MRI)の進歩によって発声時の声道形状が詳細に観測できるようになった。
  - 得られた声道形状に対して有限要素法による音響解析が行われている。
  - 有限要素法は声道内の平面波伝搬を仮定する必要がない。複雑な形状を持つ声道の音響解析に有効。
- 
- 有限要素法による解析では要素分割がボトルネック。  
声道の形状が複雑すぎて自動的に要素分割できないことがある。  
半自動的な要素分割には膨大な時間を要する。

## 研究の背景と目的(2)

- 3次元MRIデータはボクセルの集合体であるボリュームデータで処理される．
- 有限要素としてボクセル要素(立方体の6面体要素)を用いれば，MRIデータからそのまま有限要素モデルが作成できる．
- この考えにもとづいて声道の音響解析を行う．
- 解析結果(声道伝達特性)と実音声のスペクトルの比較も行う．



# 研究の背景と目的(3)

- 声道の微細構造の影響も調査する。
  - 歯列間隙：上下の歯列間の空間
  - 喉頭蓋谷：喉頭蓋と舌との間の空間
  - 喉頭腔
- 竹本ら2003
  - 喉頭腔はヘルムホルツ共鳴器としてはたらしき，その形状変化は音声スペクトルのF4にのみ大きな影響を与える。
- 本多ら2003, 2004
  - 下咽頭腔(喉頭腔 + 梨状窩)の音響特性は主声道の音響特性とほぼ独立にはたらくとする音声生成モデル。
- 上記の微細構造を除去した有限要素モデルを作成し，除去の影響を調査する。

# MRI撮像

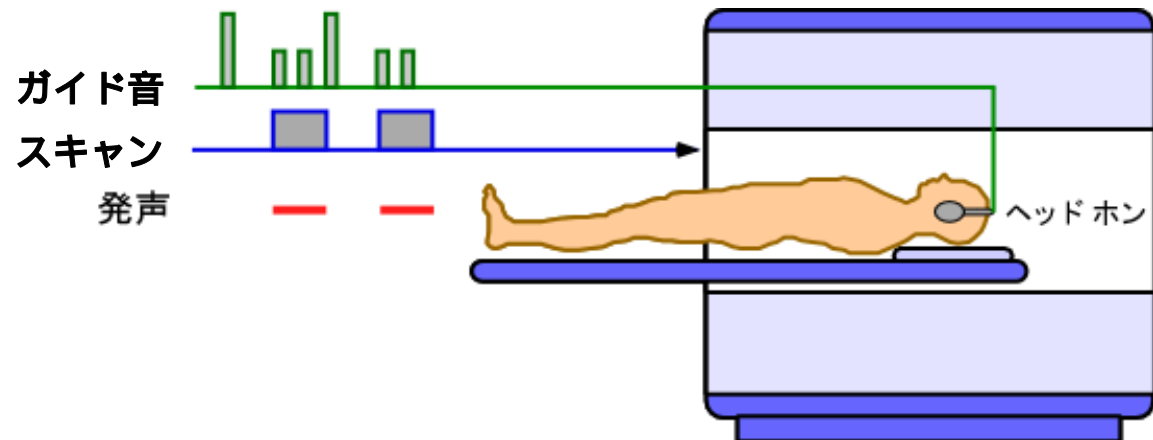
## 骨伝導と体伝導を用いたMRI同期撮像(1)

- MRI同期撮像法 [Masaki et al. 1999]

被験者に周期的なガイド音を呈示し，それに合わせて発声させ，発声区間のみで撮像．

吸気や唾液の嚥下に起因するアーチファクトを抑制し高S/NのMR画像が得られる．

ガイド音：120 Hzの倍音から成る調波複合音．



# MRI撮像

## 骨伝導と体伝導を用いたMRI同期撮像(2)

- 骨伝導と体伝導によるガイド音の呈示

MRI撮像時には大きな騒音が発生．ピーク音圧レベルが110 dB以上[宮地ら1998]．

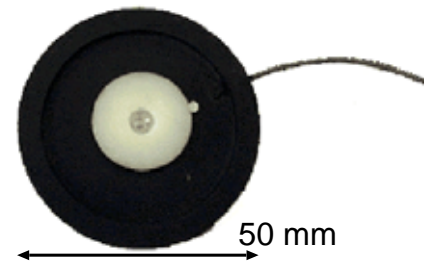
自分の音声をフィードバックできず**調音がひずむ**．  
被験者に耳栓とイヤマフをさせる．

**自分の声を骨伝導でフィードバック**できる．  
ガイド音は骨伝導と体伝導で呈示する．

- ピエゾスピーカ（非磁性体）を使用．



普段の調音に近い状態を高S/Nで撮像可能．  
撮像ノイズによる心的ストレスを軽減．



ガイド音の呈示に用いた  
ピエゾスピーカ

# MRI撮像条件

- 条件

シーケンス：Fast Spin Echo法

TE = 9 ms, TR = 4900 ms, FA = 90度 .

撮像領域：256x256 mm , 分解能：512x512 pixels

スライス：矢状方向 , 加算回数1回 .

厚み：2 mm , 間隔：2 mm , 41スライス .

撮像時間約8分 .

- 被験者

成人男性1名

- タスク

日本語5母音の持続発声 .



撮像に用いたATR脳活動イメージングセンターのMRI装置 . 7

# 音声の収録と分析

- 収録

MRIの撮像とは別の日に防音室にて。

仰臥位にて発声。撮像時に用いた枕を使用。

標本化周波数：48 kHz，量子化16 bit。

- 分析

標本化周波数16 kHzにダウンサンプリング。

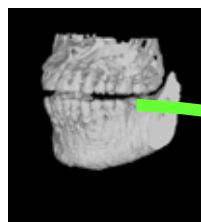
FFTケプストラムにもとづくスペクトル包絡。

- 不偏推定法[今井ら1987]
- プレエンファシス：0.98，Hanning窓，フレーム長：64 ms，フレーム周期：32 msec，ケプストラム次数：60次。
- 母音中心前後20フレームを加算平均。



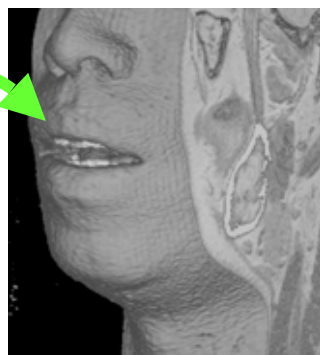
# 有限要素モデルの作成(1)

- MR画像をボリウムデータに変換し，歯列を補填．
- 声道領域および放射領域を抽出し，6面体で要素分割．  
声道および顔近傍の領域：2x2x2 mmの要素．  
放射領域：2x2x4 mmの要素(上下約100 mm，左右約70 mm，前後約50 mm)．  
要素サイズは解析ソフトが扱える要素数の上限から決定．  
前鼻孔，鼻咽腔開口部(母音/a/)をふさいだ．  
ボクセル要素を用いることで要素分割が容易．



歯列補填

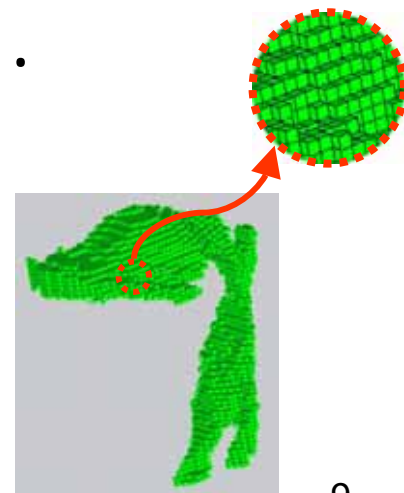
[Takemoto et. al 2004]



領域抽出

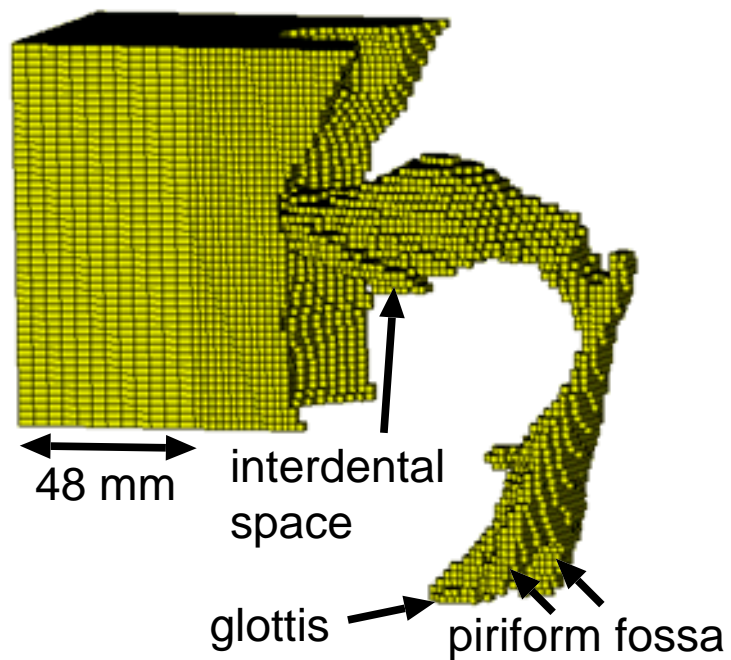


要素分割

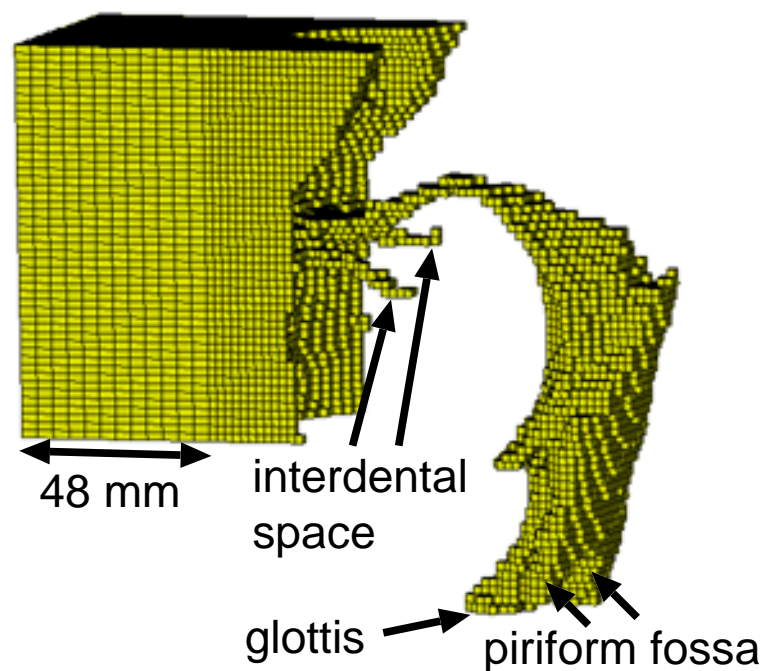


# 有限要素モデルの例

- ボクセル要素のサイズはFEMの解析ソフトが扱える要素数の上限から決定．



(a) 母音 /a/



(b) 母音 /i/

# 有限要素モデルのサイズ

有限要素モデルのノード数および要素数.

母音	ノード数	要素数
/a/	52,362	43,927
/i/	43,425	36,296
/u/	50,846	42,742
/e/	52,825	44,596
/o/	52,481	43,903

- 松崎ら(2003)の有限要素モデルの2倍以上.
- 1周波数の計算に要する計算時間: 平均18分.  
CPU: Pentium 4, 2.6 GHz  
メモリ: 2 GB  
OS: Windows 2000

# 有限要素モデルの作成(2)

- 声道の微細構造を除去したモデル

歯列間隙および喉頭蓋谷を除去したモデル(母音/e/) .

喉頭腔を2 mmだけ残して除去したモデル(母音/a/ , /i/ , /e/ , /o/) .



(a) 母音/a/のオリジナルモデルの正中矢状断面 .



(b) 母音/a/の喉頭腔を除去したモデルの正中矢状断面 .

# 伝達関数の推定法(1)

- 音圧  $p$  に関する波動方程式

$$\nabla^2 p + k^2 p = 0$$

k : 波数 ( $k = \omega / c$ )  
ω : 周波数  
c : 音速

- 境界条件

入力端：声門部．正弦波振幅 1.0 m/secの体積速度で励振．

出力端：放射領域の境界面．空気の固有インピーダンス  $\rho c$

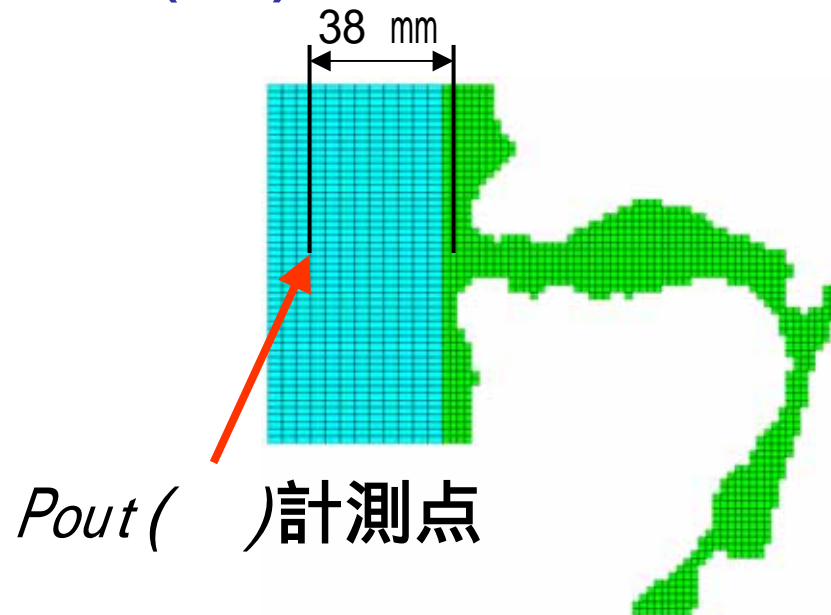
入力端と出力端以外は剛壁．

- 音速  $c = 352.33$  m/sec , 空気密度  $\rho = 1.1255$  kg/m<sup>3</sup>
- この条件で上の波動方程式をガラーキン法で解いた．
- 解析プログラム：LMS社SYSNOISE ver. 5.6

# 伝達関数の推定法(2)

- 伝達特性  $H(\omega)$

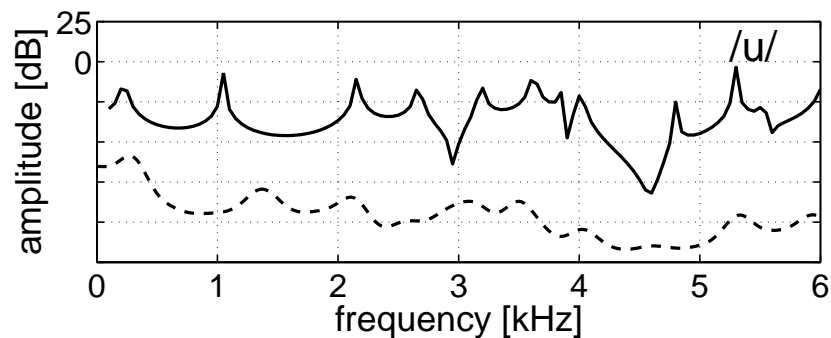
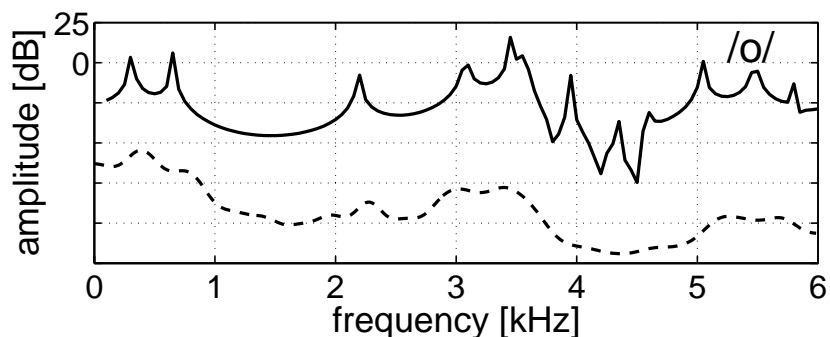
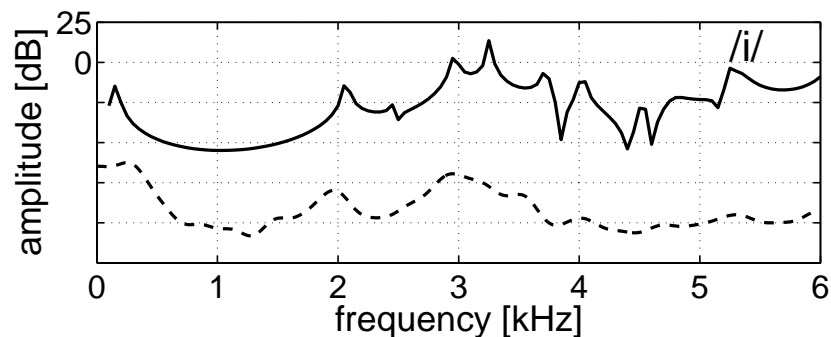
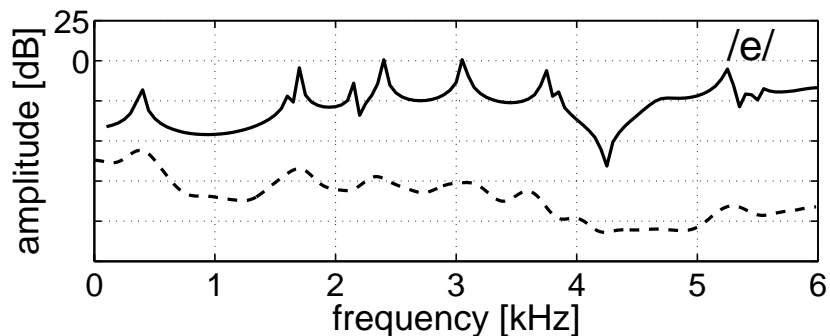
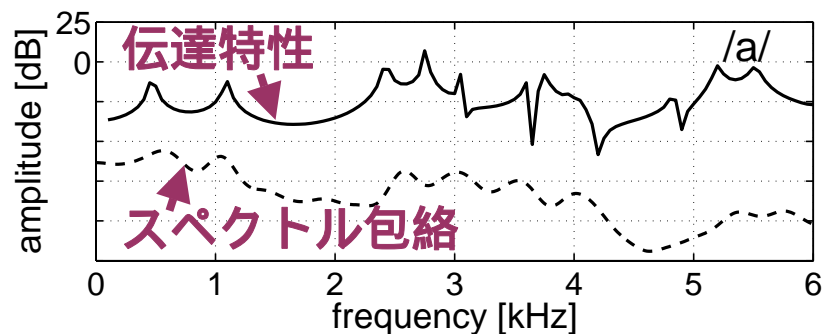
$$H(\omega) = \frac{P_{out}(\omega)}{U_{in}(\omega)}$$



- $P_{out}(\omega)$ : 正中矢状面における口唇先端から水平方向 38 mmの点の音圧 .
- $U_{in}(\omega)$ : 入力端の速度 .
- 100 Hzから6 kHzまでを50 Hzきざみで計算 .

# 結果および考察

## 声道伝達特性と音声スペクトル包絡の比較(1)



第1から第4ホルマントは  
母音 /u/ を除いて大局的  
に一致 .

# 結果および考察

## 声道伝達特性と音声スペクトル包絡の比較(2)

- 第1から第4ホルマントは母音/u/を除いて大局的に一致。

周波数の差の最小値：3 Hz (母音/e/のF2)

周波数の差の最大値：444 Hz (母音/u/のF4)

- ホルマントごとのピーク周波数との差の平均

F1: 23 %

F2: 9 %

F3: 3 %

F4: 7 %

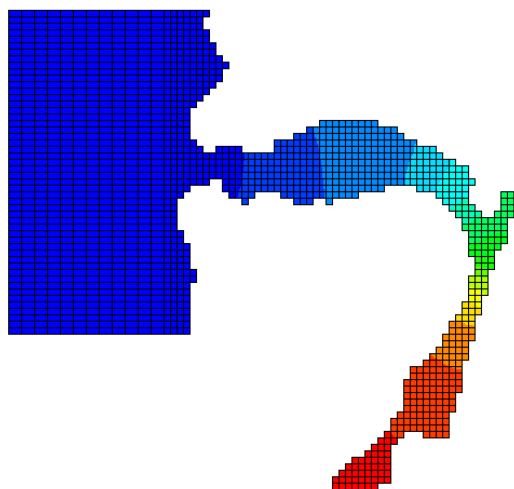
Nishimoto et al. (2004)の解析結果と同じ傾向。  
声道壁を剛壁としていること、鼻腔の影響が考慮されていないことなどが原因か。

- MR画像と同程度の分解能(0.5x0.5 mm/pixel)の有限要素モデルを計算できるようになれば、より実音声と一致すると期待できる。

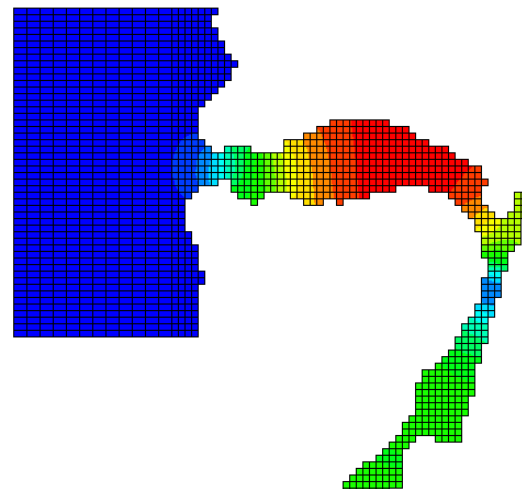


# 結果および考察

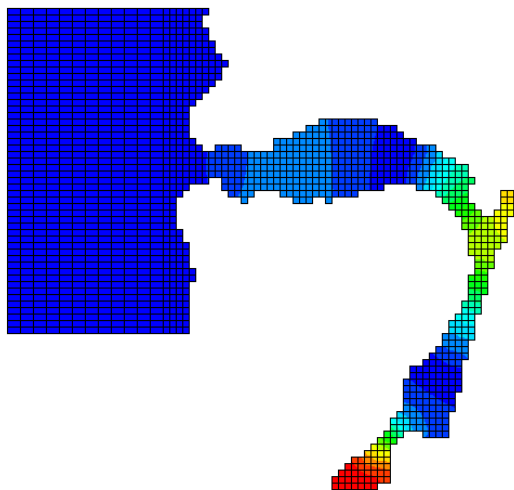
ピーク周波数での音圧分布 (母音 /a/)



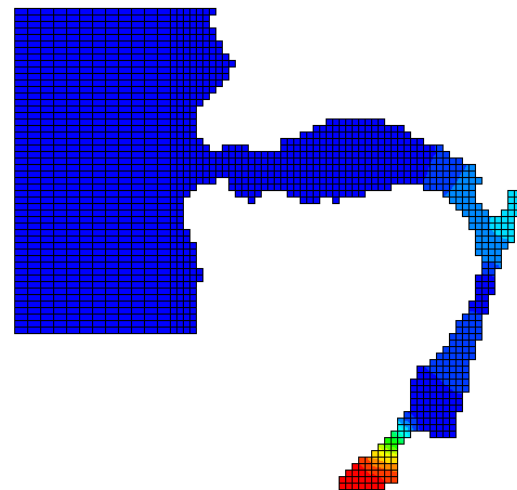
(a) 第1ピーク (450 Hz)



(b) 第2ピーク (1,100 Hz)



(c) 第3ピーク (2,400 Hz)

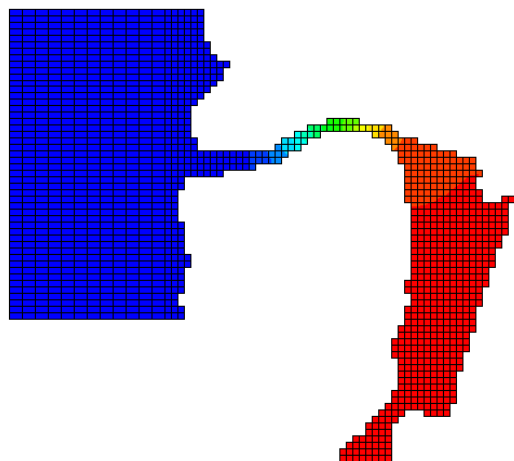


(d) 第4ピーク (2,750 Hz)

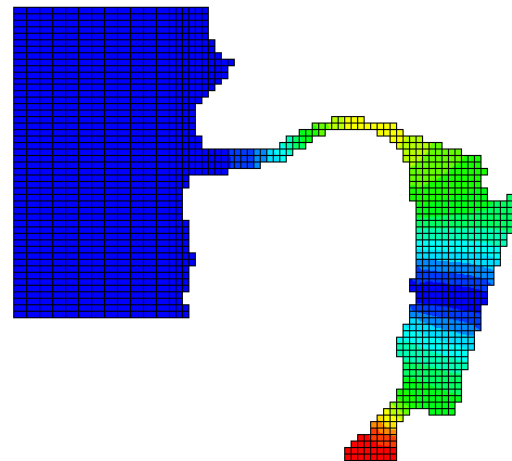


# 結果および考察

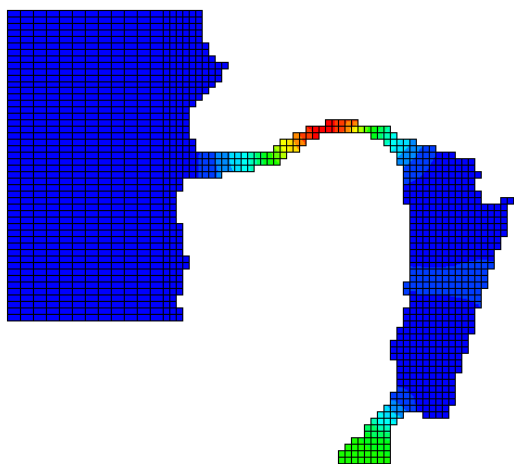
## ピーク周波数での音圧分布 (母音 /i/)



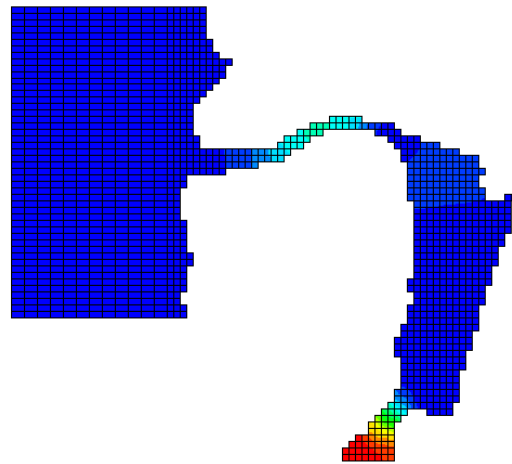
(a) 第1ピーク (150 Hz)



(b) 第2ピーク (2,050 Hz)



(c) 第3ピーク (2,950 Hz)



(d) 第4ピーク (3,250 Hz)



# 結果および考察

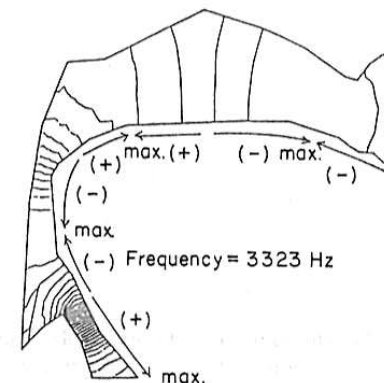
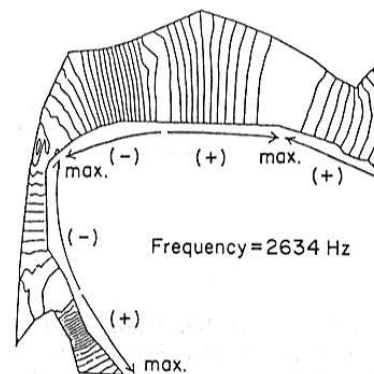
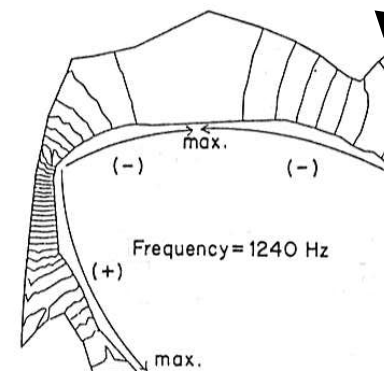
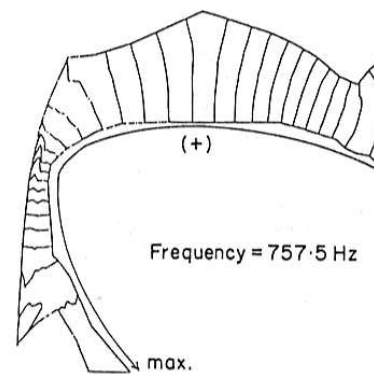
## ピーク周波数での音圧分布(1)

- Kagawaら(1992)が境界要素法で示した結果とおおむね一致.

母音/a/の第1, 2, 3ピークは声道を片開き管とみなしたときの1/4, 3/4, 5/4共鳴に対応.

母音/i/では第1, 2ピークは声道を片開き管とみなしたときの1/4, 3/4共鳴に対応.

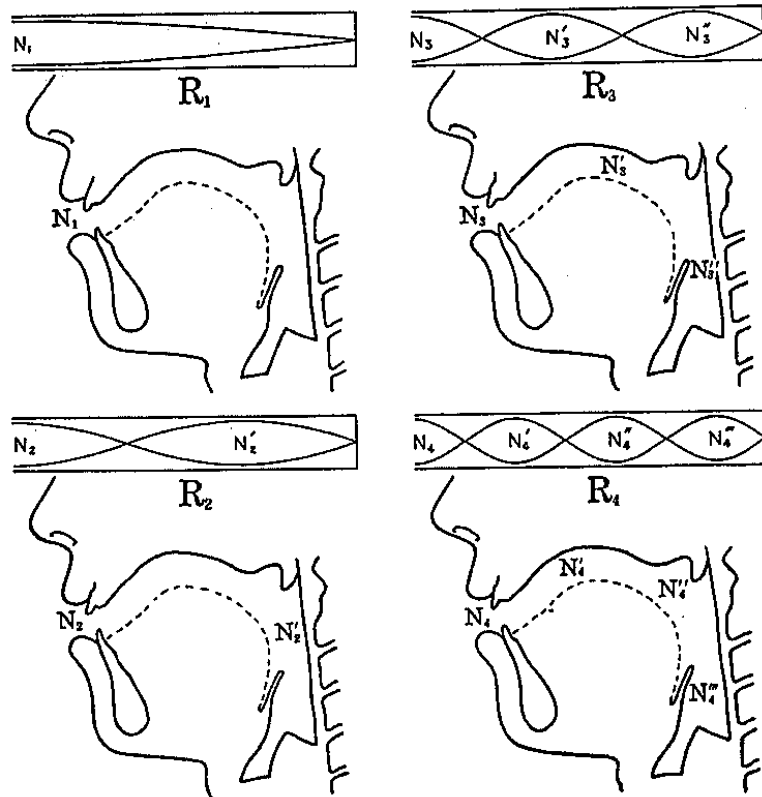
母音/i/の第3ピークでは音圧の高い部分がせばめに集中.



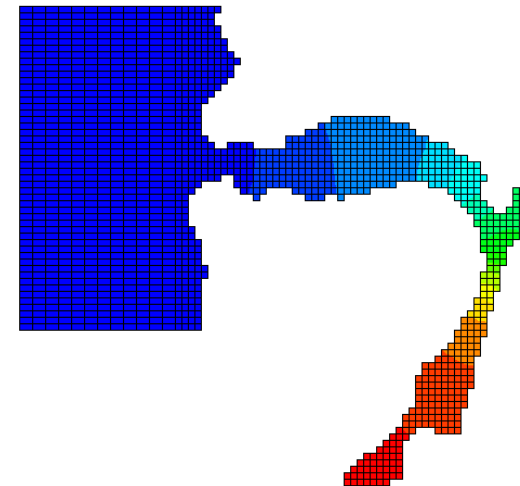
母音/a/のF1からF4までの等音圧面.  
[Kagawa and Shimoyama 1992]

# 結果および考察

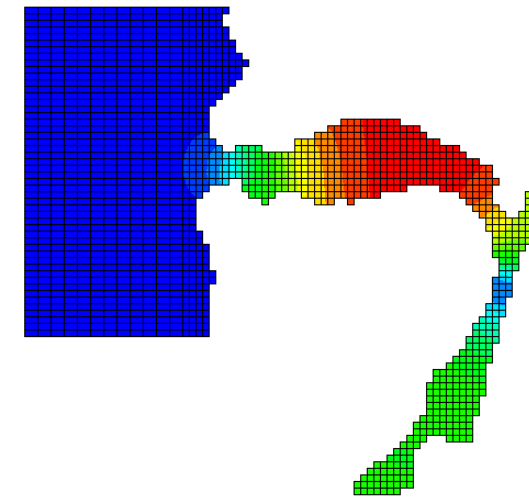
## 参考：体積流の分布



断面積が均一の音響管または声道における体積流の最大点の分布。  
 出典：千葉，梶山，“母音”（1942）



(a) 母音/a/の第1ピーク (450 Hz)



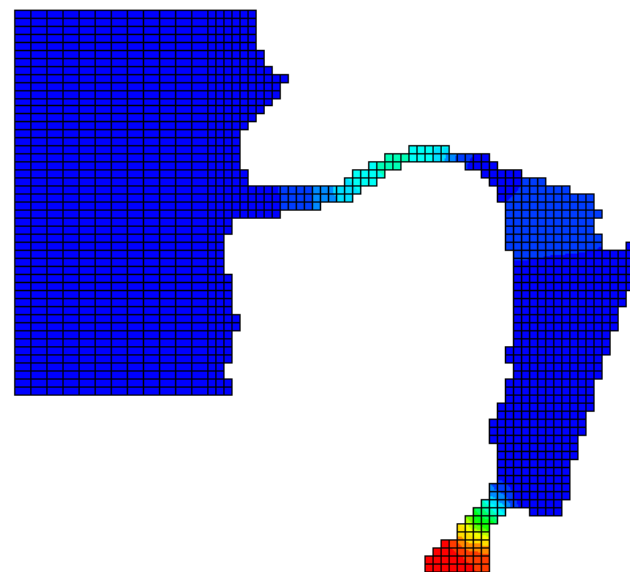
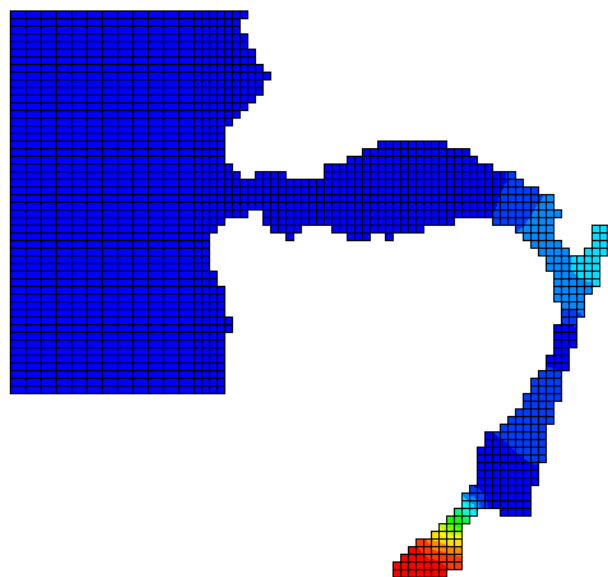
(b) 母音/a/の第2ピーク (1,100 Hz)

# 結果および考察

## ピーク周波数での音圧分布(2)

- 第4ピークの周波数では7/4共鳴の音圧分布が明確に観察されない。

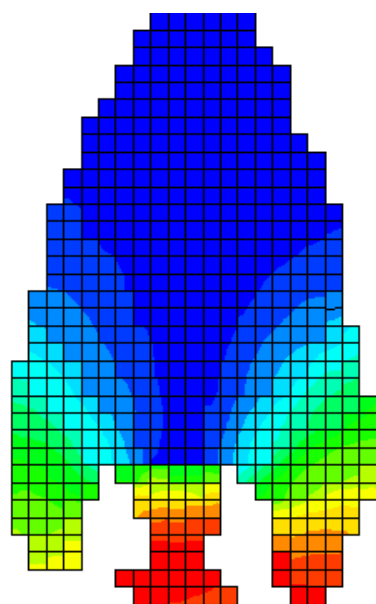
音圧の高い部分が喉頭腔に集中。



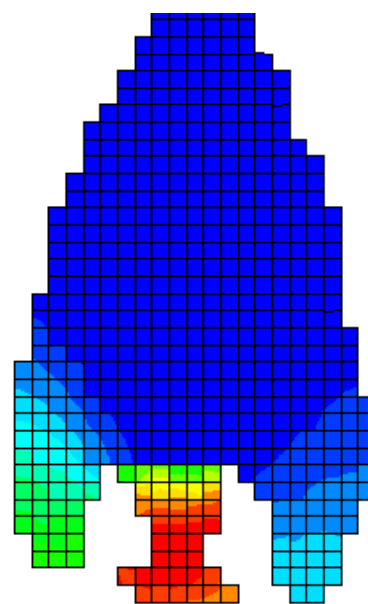
(a) 母音/a/の第4ピーク(2,750 Hz) (b) 母音/i/の第4ピーク(3,250 Hz)

# 結果および考察

梨状窩による谷の周波数での音圧分布 (母音 /a/)



(a) 3,650 Hz (右)



(b) 4,200 Hz (左)

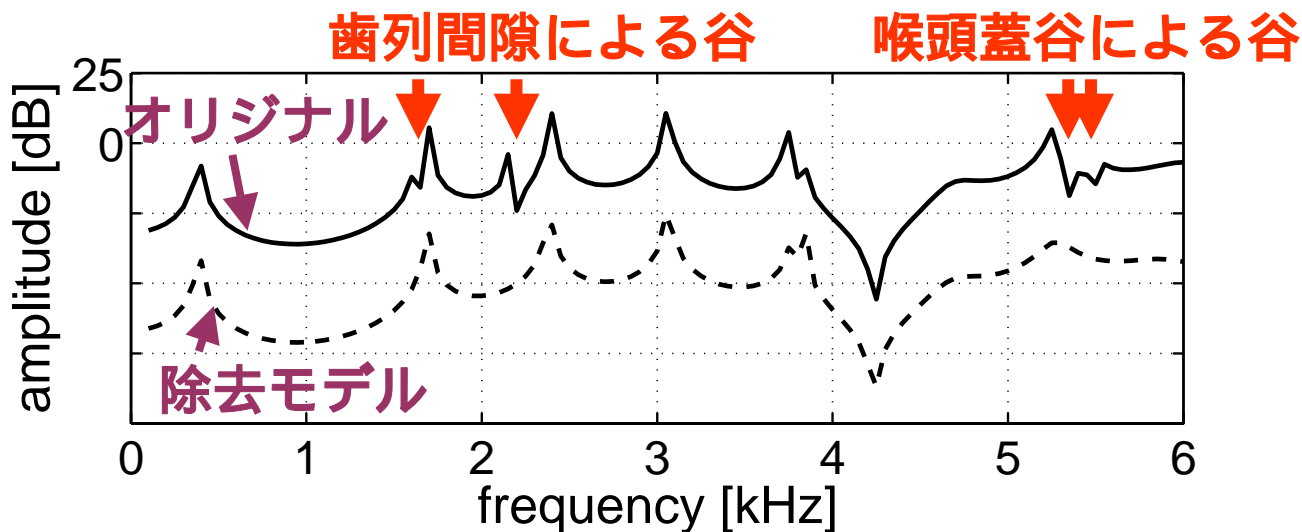


- 等音圧面は咽頭空の壁に沿って斜めに持ち上がっている。
  - 梨状窩の実効的な開口端は斜めに開口している。
- 河西ら (1996) のシミュレーション結果と一致。

# 結果および考察

## 歯列間隙および喉頭蓋谷の影響

- 母音 /e/ の有限要素モデルから歯列間隙と喉頭蓋谷を除去したモデルを作成．
- 除去したモデルでは1,650 Hz , 2,200 Hz , 5,350 Hz , 5,500 Hz の谷が消えた．



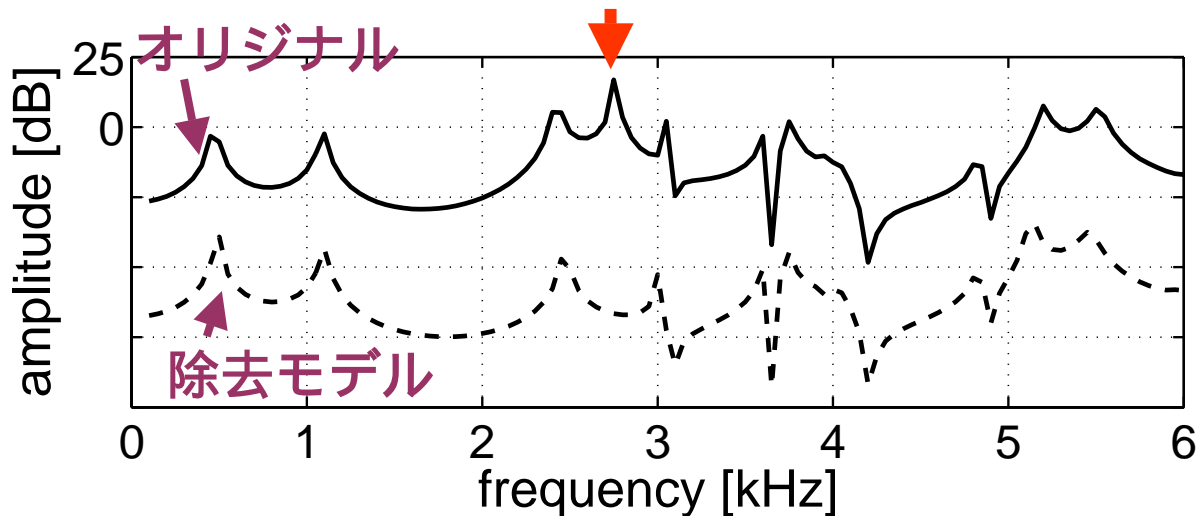
母音 /e/ の有限要素モデルと歯列間隙および喉頭蓋谷を除去したモデルの伝達特性．

# 結果および考察

## 喉頭腔の影響(1)

- 喉頭腔を除去することによって, 2,750 Hzに現れていた第4ピークが消えた.
- 2 kHzから3 kHzにかけてのパワーが減少した.

喉頭腔が第4ピークを生じさせる.

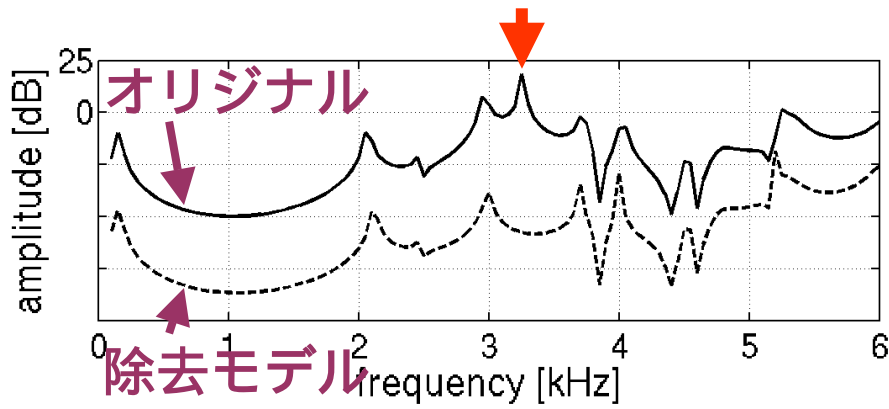


母音/a/の有限要素モデルと喉頭腔を除去したモデルの伝達特性.

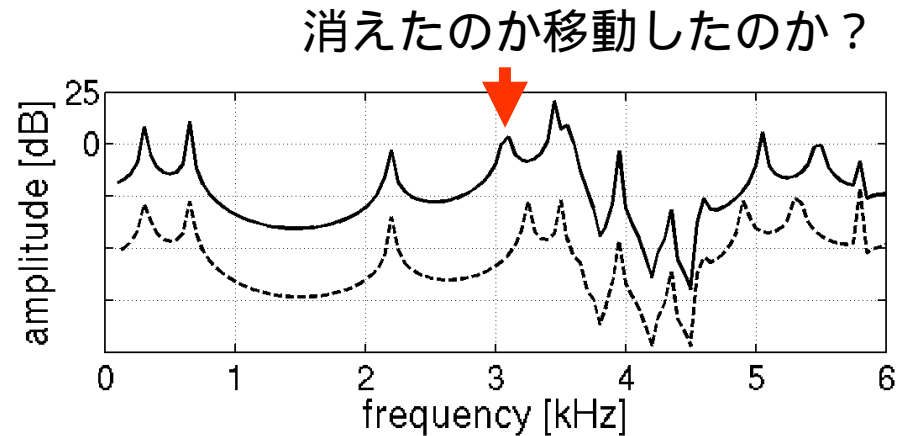


# 結果および考察

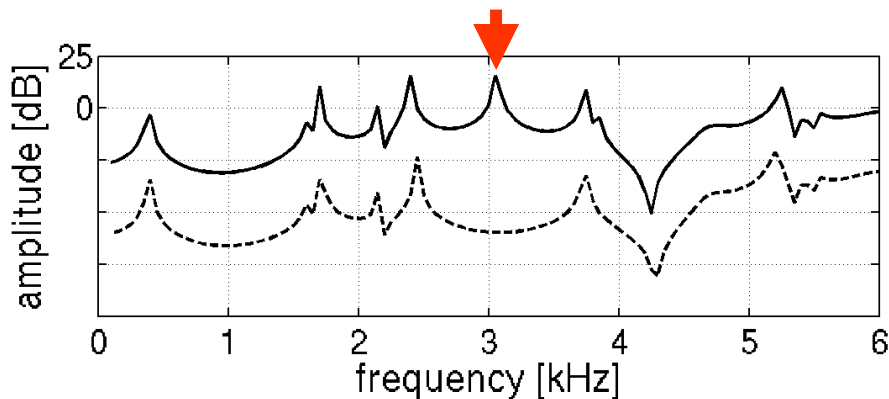
## 喉頭腔の影響(2)



(a) 母音 /i/



(c) 母音 /o/



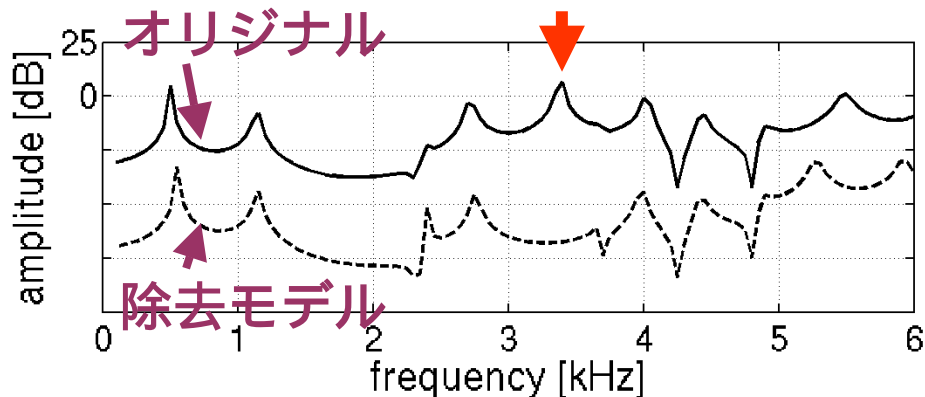
(b) 母音 /e/

喉頭腔が第4ピークの  
発生に大きく寄与.

# 結果および考察

## 喉頭腔の影響(3)

- 別話者の母音 /a/ の有限要素モデルで調査 .



- 喉頭腔とF4の間には強い関係がある .
- 竹本ら (2003) の結果を支持 .  
喉頭腔の断面積変化は音声のF4のみに大きな影響を与える .
- 本多ら (2004) の音声生成モデルを支持 .  
下咽頭腔の音響特性は主声道の音響特性に対して独立に働く .

# まとめ

- 日本語5母音発声時の声道形状をボクセル要素で分割して有限要素モデルを作成し音響解析を行った．
- 声道伝達特性のピーク周波数と音声のホルマント周波数は，母音/u/以外で大局的に一致．これらの差は数Hzから数百Hz．
- 歯列間隙，喉頭蓋谷および梨状窩は声道の伝達特性に谷を生じさせる．
- 喉頭腔は音声のF4生成に寄与する．
- 今後，声道模型を対象にした伝達特性の音響計測の結果と比較する予定．

# ATR母音MRIデータのご案内

- 日本語5母音発声時のMRI静止画像を公開。
- 話者：成人男性1名，東京方言
- 発話同期撮像法で撮像された高分解能 (0.5x0.5mm/pixel)，高S/Nな画像データ
- 話者による5母音音声データつき
- 2005年1月公開予定
- 問い合わせ先

ATR人間情報科学研究所 生体イメージング研究室

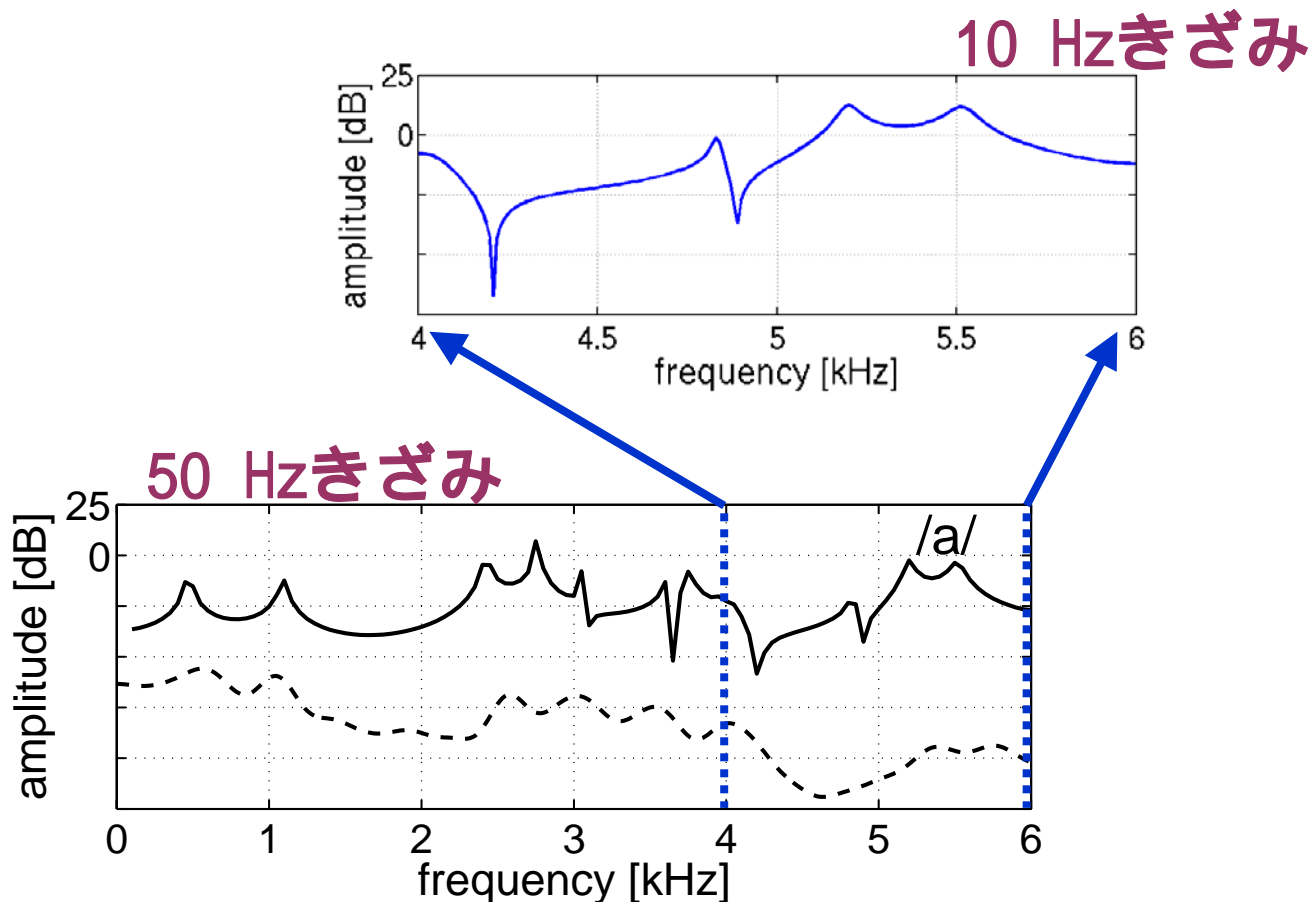
[hismridata@atr.jp](mailto:hismridata@atr.jp)

[www.his.atr.jp/bpi/](http://www.his.atr.jp/bpi/)

# 結果および考察

極零対は生じないのか？

- 母音 /a/ の有限要素モデルを10 Hzきざみで計算 .



# 結果および考察

## 声道伝達特性と音声スペクトル包絡の比較

声道伝達特性のピーク周波数と音声スペクトル包絡のホルマント周波数の比較 [Hz].

		F1	F2	F3	F4
/a/	FEM	450	1,100	2,400	2,750
	speech	563	1,047	2,578	3,016
/i/	FEM	150	2,050	2,950	3,250
	speech	250	1,969	2,953	---
/u/	FEM	200	1,050	2,150	2,650
	speech	267	1,359	2,109	3,094
/e/	FEM	400	1,700	2,400	3,050
	speech	375	1,703	2,344	3,094
/o/	FEM	300	650	2,200	3,100
	speech	391	719	2,266	3,016

FEMは50 Hzきざみで計算.