

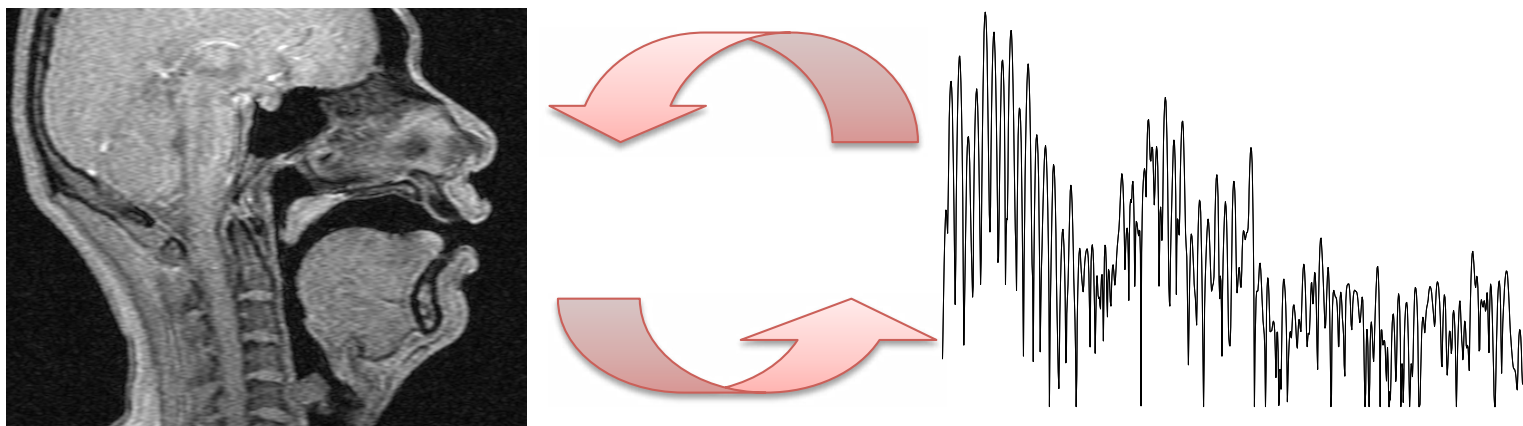
MRIにより計測した 声道長と音声データとの関係

波多野 博顕(神戸大)、北村 達也(甲南大)、
竹本 浩典、パーハム モクタリ(NICT)、
本多 清志(パリ第3大学)、
正木 信夫(ATR/ATR-Promotions)

はじめに

本研究の背景

- 形態的特徴(声道等)
 - 音響的特徴(出力音声)
- } 密接に結びついている



音声生成

音声認識技術(話者正規化等)

音声合成技術(声質変換等)

先行研究

- MRIを用いて声道長を計測した主な研究
 1. 楊 & 粕谷 (1995)
 - 子供・成人女性・成人男性を対象に、声道断面積関数 (vocal tract area function) の個人差を分析。
 2. Fitch & Giedd (1999)
 - 2～25歳の男女129名を対象に、成長に伴う形態変化や身長・体重・性差との関連を指摘。
 3. Vorperian et al (2005)
 - 子供63名・成人12名を対象に、形態的特徴を記述。
- いずれも声道の発達変化の観点からの分析
- 音声との関連性が明らかではない

本研究の特徴・目的

- ATR-Promotions 脳活動イメージングセンタによる日本語5母音発話のMRIデータ(最近公開)を使用

MRI画像は実際に発話中のものである
MRI撮像時の発話音声を収録している
被験者のBody size (身長) 情報あり

1. 声道長と身長との相関
2. 姿勢の相違におけるフォルマント周波数の比較 (母音/a/) →発表では割愛(予稿集に記載)
3. 声道長と5母音のフォルマント周波数の相関

分析对象・方法

被験者とMRI撮像パラメータ

被験者

日本人成人男性15名

年齢範囲 24～55 歳 平均 37.2 歳 SD 9.1

身長範囲 164～184 cm 平均 172.9 cm SD 5.6

MRI

Shimadzu-Marconi ECLIPSE 1.5T Power Drive 250

FOV 256 x 256 mm、image size 512 x 512 pixels
(1 pixel = 0.5 x 0.5 mm)、TE = 3.9 ms、TR = 15ms、
スライス厚 5 mm、加算回数 2回

発話タスクと音声収録

発話タスク

- 日本語5母音(/a/, /e/, /i/, /o/, /u/)の単独発話
- 撮像の間(約5秒)母音を一息で発話し、もし息が切れた場合は口の構えを維持するよう指示。

撮像時の音声収録

標本化周波数 16 kHz 量子化ビット数 16 bit
光マイクロフォン: Phone-Or SOM
レコーダ: Marantz PMD-670

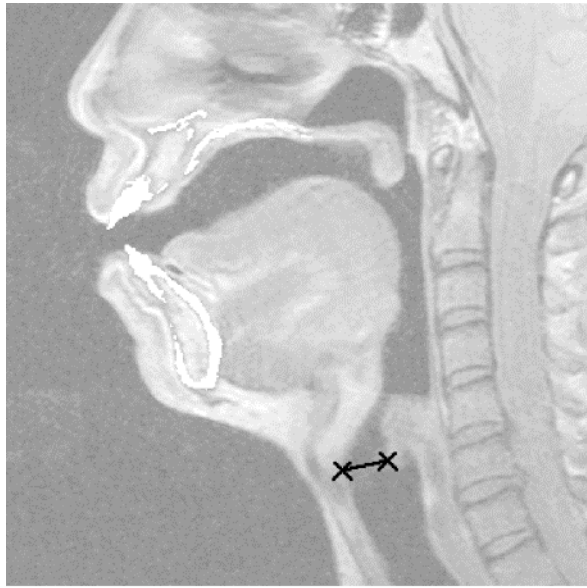
声道長とフォルマント周波数の抽出

声道長の抽出

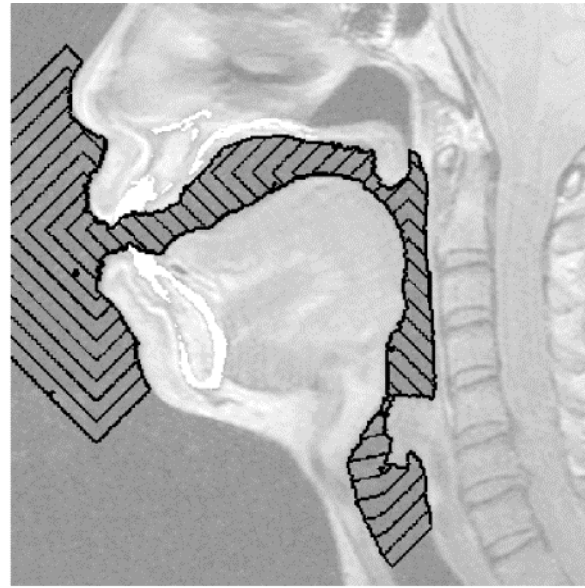
- Takemoto *et al.* (2006)の抽出法による
- (a)目視にてMRIデータ上の声門位置を決定し、(b)声道領域内にコンターマップを作成
- (c)コンターマップに基づき声道中心線を引き、(d)2.5 mm単位で計測

フォルマント周波数の抽出

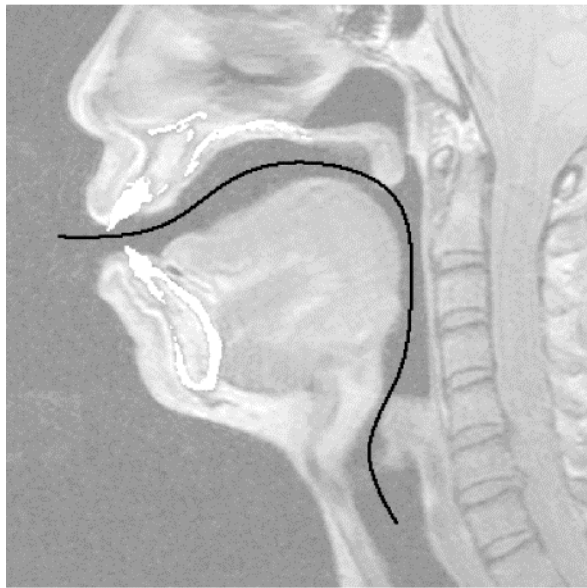
- 今井・古市(1987)による対数スペクトルの不偏推定を利用
- ハニング窓、フレーム長64 ms、フレーム周期16 ms、繰り返し3回、ケプストラム次数 60 次でスペクトル包絡を求め、フレーム間で加算平均
- 目視にて第1～第4フォルマント周波数を計測



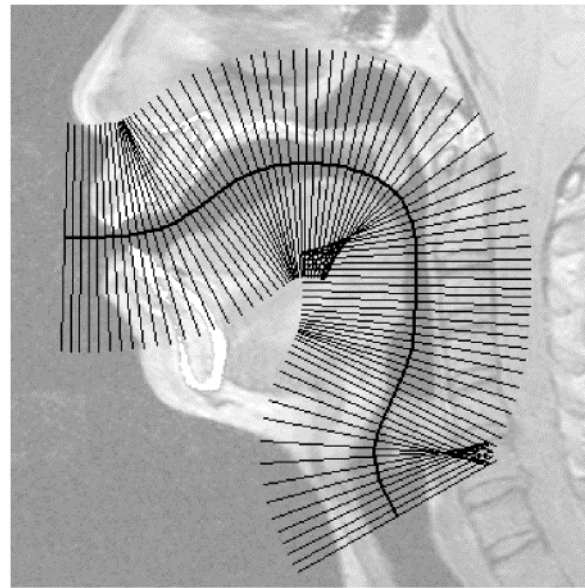
(a)



(b)



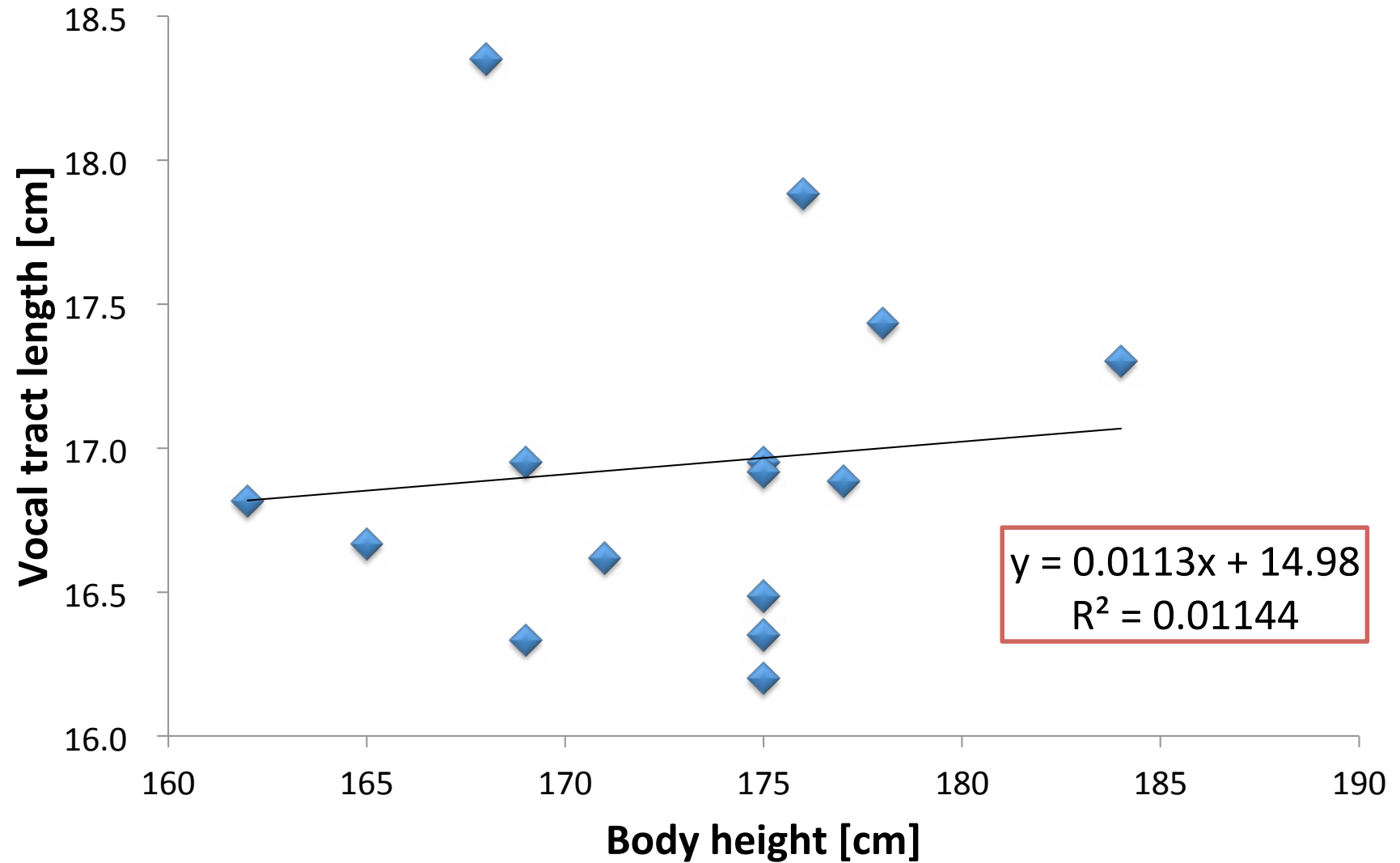
(c)



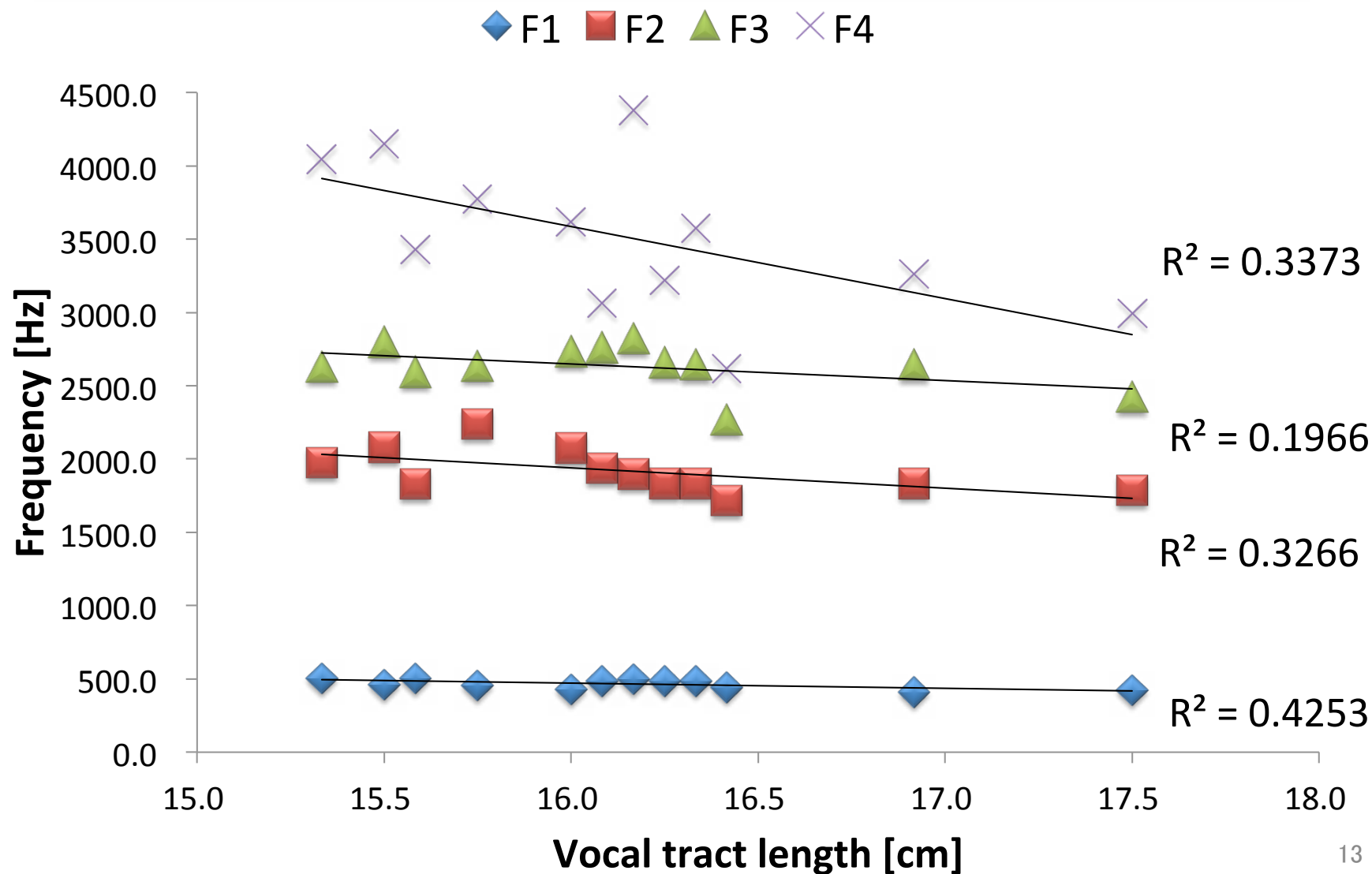
(d)

結果と考察

声道長と身長



声道長とフォルマント周波数 /e/ (予稿集には載っていません)



声道長とフォルマント周波数

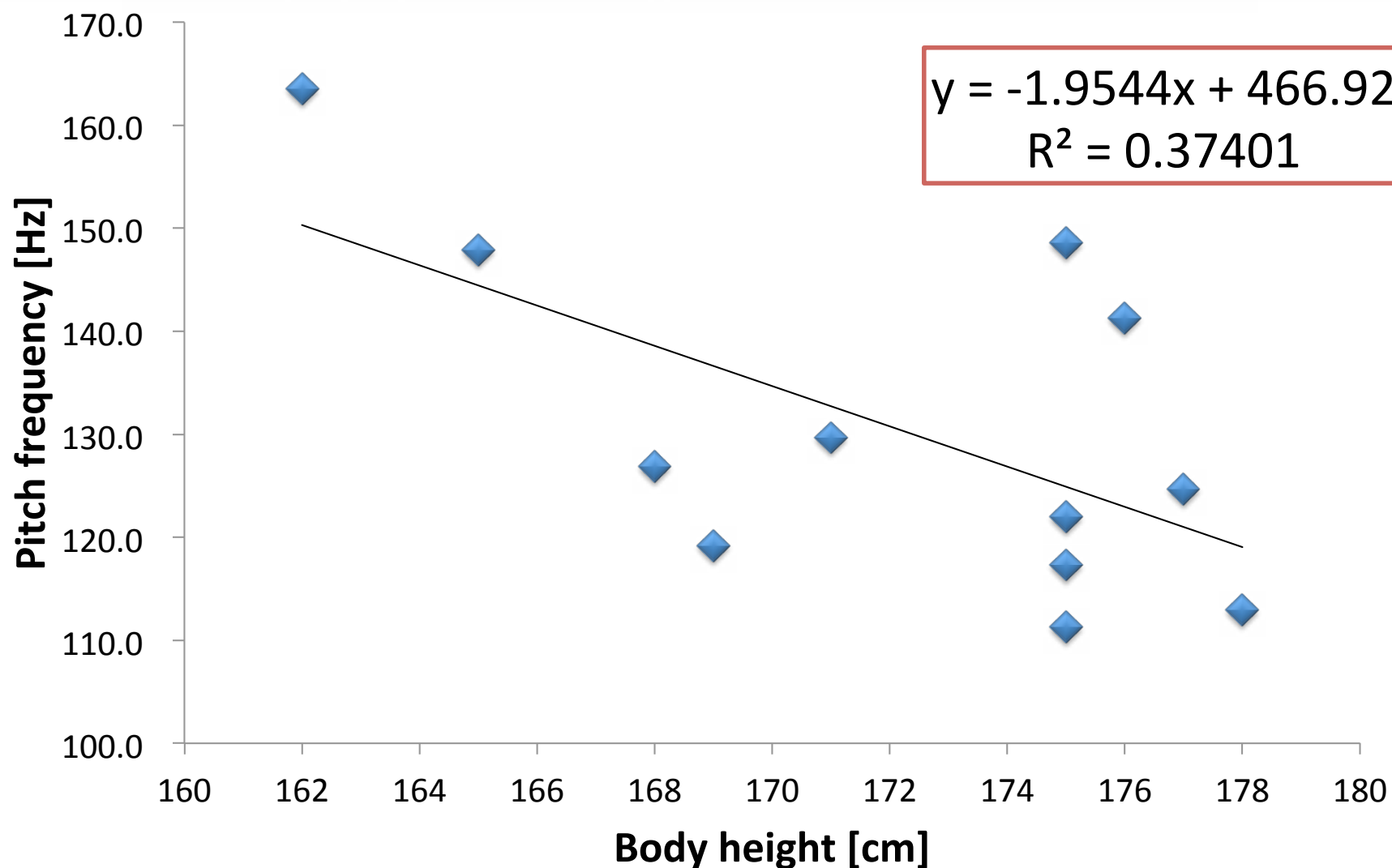
- r^2 値

訂正: 予稿集の数値が間違っていました

	/a/	/e/	/i/	/o/	/u/
F1	0.025	0.425	0.071	0.245	0.220
F2	0.022	0.327	0.007	0.249	0.001
F3	0.048	0.197	0.047	0.153	0.054
F4	0.062	0.337	0.145	0.080	0.057

⇒ /e/のF1, F2, F4の数値がやや高いものの、全体的に相関があるとはいえない。

基本周波数と身長 (予稿集には載っていません)



結果のまとめ

◆ 声道長と身長

相関なし ($r^2 = 0.01$)

◆ 声道長とフォルマント周波数

/e/の r^2 値はやや高いものの
全体的に相関ありとは言えない

◆ 身長と基本周波数

やや弱い相関あり ($r^2 = 0.37$)

考察

➤ 声道長と身長に相関は無い

Fitch & Giedd (1999)では非常に強い相関ありとしているが($r^2=0.86$)、声道を含む諸器官が発達している成人では相関がみられない。

➤ 声道長とフォルマント周波数に相関は無い

フォルマント周波数にみられる個人差は、声道長ではなく声道形状の違いに起因。(北村ら (2005)の結果を支持)

➤ 基本周波数は身長と相関する

声帯の長さは身長に比例して長くなり、そのために基本周波数が低くなる？

今後の課題

- ◆ 一般性の検証（更なる被験者の確保）
 - 全体の被験者は15名だが、うち撮像時の音声を収録できた被験者は12名のみであった。
⇒ /e/の特異性が解明できるかもしれない
- ◆ 性別による違い
 - 先行研究では性差による声道長の相違を指摘しているが、今回のデータは男性のみであった。

謝辞:本研究は平成23年度科研費(21300071, 21500184)によりおこなわれた