

2014年度名古屋大学学生論文コンテスト

優秀賞 受賞

音響・調音音声学でのフォルマントによる多言語の母音比較分析

理学部 森 崇人

1. はじめに

私は合唱をやっており、外国語の曲と接する機会も多く、音声言語学(Linguistic Phonetics)にとっても興味を抱いていた。音声言語学で日本語およびそのほかの外国語を分析すれば、合唱に生かせるのではないかと思い、本研究を始めた。

言語の知覚に関して母音認識(vowel perception)はとても重要である。例えば、言語学習において、母音(vowel)を知ることは必須であり、様々な言語の母音を系統的に整理することでより系統的な言語教育をすることができる。また、吃音症などの発音障害においても、子音(consonant)および母音の発音訓練が重要であるが、そのときに主観的、独立的な訓練方法ではなく、客観的に言語音の区別・矯正ができるようなものがあれば、一般的に応用でき、安価でそのような治療を求めている人々に行きわたり、発音治療の大きな前進となるだろう。

この研究では、主に母音について実験および考察を行う。先行研究で、母音というのはフォルマント(formant)で区別されることがわかっている。本研究では、母音を特徴づけるためにフォルマントが有効であることを確認し、フォルマントを用いて複数言語間での母音比較をする。これまでの研究では、二言語間の比較や、単一言語内での特性の調査にとどまっていた(朝川 2005, 益子 2011, 楊 2005)。したがって、本研究では、複数の言語で母音を互いに比較し、包括的に共通の傾向および音声学的な差異を考察する。

2. 予備実験

2-1. 予備知識

[1] 音色、倍音、フォルマントの関係

音色(timbre)は倍音成分およびフォルマントで区別できることが知られている。倍音とはある音に含まれている、基底周波数以外の周波数の音のことであり、その倍音のピーク値がフォルマントである。

[2] アルゴリズム

入力信号からフォルマントを出すまでのアルゴリズムを簡潔に紹介する。

① 連続信号から離散信号に変換

連続信号(アナログ)を離散信号にするため、連続信号を標本化定理(sampling theorem)「原信号に含まれる周波数成分をすべて正確にサンプリングするためには、原周波数の2倍以上のサンプリング周波数が必要」にしたがって標本化(sampling)とよばれる音の時間方向の離散化と量子化(quantization)とよばれる音の強度方向の離散化をしたのち、符号化(encoding)をして、離散(デジタル)信号にする。

② 短時間フーリエ変換(Short-Time Fourier Transform: STFT)

STFTによって有限時間の信号をデジタルフーリエ変換する。窓関数をかける(窓かけ=フレーム化)ことによって範囲を区切る。これを分割統治法(D&C algorithm)という。ここで、離散フーリエ変換(Discrete Fourier Transform: DFT)を使い、倍音成分を得る。ただし、DFTは高速フーリエ変換(Fast Fourier Transform: FFT)というアルゴリズムを用いて N^2 の計算量を $N \log_2(N)$ の計算量に減らすことができるので、ここではDFTをFFTを使って計算する(使用するFFTアルゴリズムは

Cooley-Tukey 型。ただし FFT には他のアルゴリズムも存在する)。以下に STFT と DFT の数学的定義を示す。

<STFT>

– $w(t)$ を窓関数、 $x(t)$ を変換される信号とすると、STFT は以下のように表される

$$\text{STFT}_{x,w}(t, \omega) = \int_{-\infty}^{\infty} x(\tau)w(\tau - t)e^{-i\omega\tau} d\tau$$

– 離散時間に関する STFT は次のようになる

$$\text{STFT}_{x,w}[n, \omega] = \sum_{m=-\infty}^{\infty} x[n + m]w[m]e^{-i\omega m} dn$$

<DFT>

n 個の複素数列 x_0, \dots, x_{n-1} に対して DFT すると n 個の複素数列 f_0, \dots, f_{n-1} が得られる。

$$f_j = \sum_{k=0}^{n-1} x_k e^{-\frac{2\pi i}{n}jk} \quad j = 0, \dots, n-1$$

③ パワースペクトル密度 (Power Spectral Density: PSD)

PSD は信号の力が周波数についてどのように分布しているのか示している。定義はフーリエ変換の絶対値の二乗で、フーリエ変換を以下のように定義すると、

$$F_T(\omega) = \frac{1}{\sqrt{T}} \int_0^T f(t)e^{-i\omega t} dt$$

PSD は以下のように定義される。

$$\text{PSD}(\omega) = \lim_{T \rightarrow \infty} \mathbf{E}[|F_T(\omega)|^2]$$

④ スペクトル包絡 (spectral envelope)

スペクトル包絡は PSD の包絡線であり、そのピークがフォルマントである。声道の共振・反共振特性、音源のスペクトル概形、口唇・鼻孔などからの放射特性などを表す。スペクトル包絡を求めるには、主に二つの方法があり、それぞれ線形予測法 (linear prediction) / 線形予測符号 (Linear Predictive Coding: LPC)、ケプストラム法 (cepstrum) がある。今回は LPC を使った。LPC は以下のように定義される。

線形予測式: 信号 $X = \{x_0, x_1, x_2, \dots, x_{N-1}\}$

$$\hat{x}_t = - \sum_{i=1}^p a_i x_{t-i}$$

自己相関関数

$$v_\tau = \frac{1}{N} \sum_{i=\tau}^{N-1} x_i x_{i-\tau} \quad \tau = 0, 1, \dots, N-1$$

そして、Yule-Walker 方程式 (下) を Levinson-Durbin アルゴリズムで解く

$$\begin{pmatrix} v_0 & v_1 & v_2 & \cdots & v_{p-1} \\ v_1 & v_0 & v_1 & \cdots & v_{p-2} \\ v_2 & v_1 & v_0 & \cdots & v_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{p-1} & v_{p-2} & v_{p-3} & \cdots & v_0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{pmatrix} = - \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_p \end{pmatrix}$$

残差パワー

$$\sigma^2 = v_0 + a_1 v_1 + a_2 v_2 + \cdots + a_p v_p = \sum_{i=0}^p a_i v_i$$

そうすると、スペクトル包絡が得られ、そのピークがフォルマントになる（周波数が低いものから順に、第一フォルマント(F1)、第二フォルマント(F2)、第三フォルマント(F3)…という）。例えば、図1は男声（日本人男性16歳）のPSDとスペクトル包絡を示している。

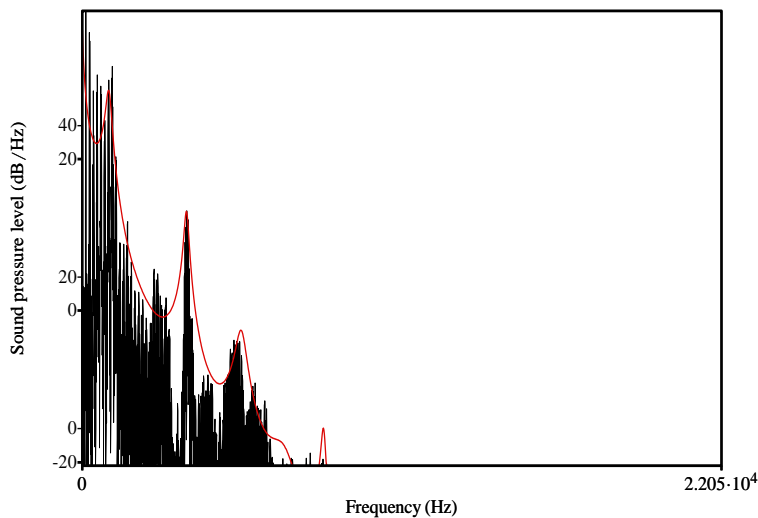


図1 男声の PSD およびスペクトル包絡（棒グラフ部分：PSD、包絡線：スペクトル包絡）

2-2. 予備実験

[1] 目的

同一母音のフォルマントにおける性差があるか調べる。

[2] 手法

OLYMPUS ボイストレック DM-10（記録形式：WMA(Windows Media Audio)、サンプリング周波数：44.1kHz、総合周波数特性：300～7,000kHz）を録音に用いた。フォルマント解析ソフトはおんかいくん、音の分析と Praat（すべてフリーソフト）を使用した。インフォーマント（実験協力者）は日本人男性と日本人女性で、日本語の/a/を2秒間発音してもらい、それを録音・解析し、基底周波数、倍音、フォルマントを比較した。

[3] 結果および考察

F0 は男性のほうが女性より周波数が低いですが、逆に倍音分布に関しては女性のほうが男性より相対的に低周波数域に分布していた。したがって、図2 でみられるようにF1 とF2 については男性と女性の間大きな差は見られなかった。したがって、母音のフォルマントのうちF1 とF2 に関しては、性差はないとみられる。

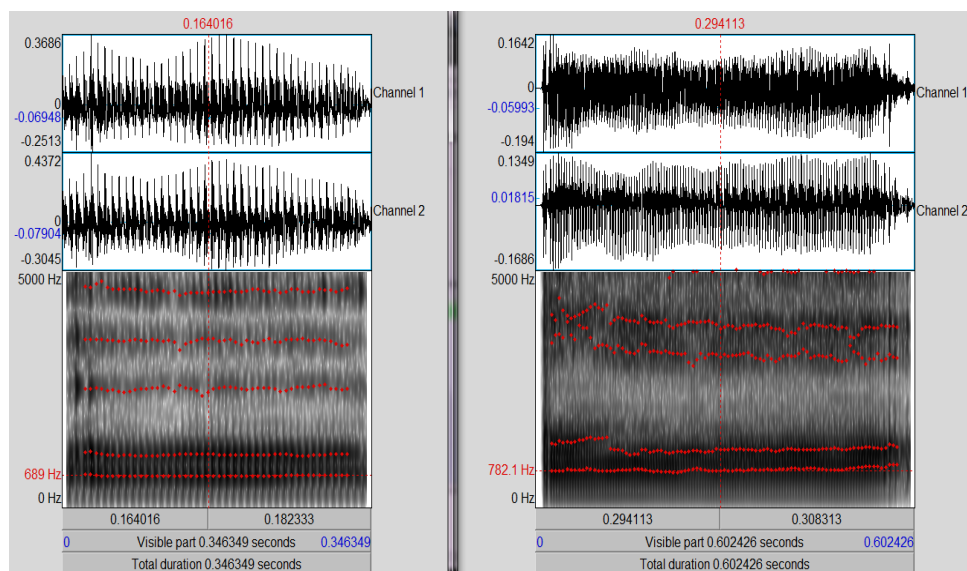


図2 男性（左）と女性（右）の/a/のフォルマント（図中の点）比較（下から順にF1、F2）

3. 実験

3-1. 実験 1

[1] 目的

人種の違いによる口腔および声道(vocal tract)の身体的、器質的差異または母国語の違いでフォルマントに影響が出るか調べる。

[2] 手法

録音機は予備実験と同様のものを使用した。フォルマント解析に当たっては、Praatを用いた。アメリカ英語を母国語とするオハイオ州出身の16歳男性（以後 Eng-M）と日本語を母国語とする16歳男性（以後 Jpn-M）から日本語の5母音/a/, /i/, /u/, /e/, /o/を録音し、フォルマントを解析した。これを一回行った。その後、F1-F2 散布図にして比較した。ただし Eng-M は日本語を流暢に話すのに十分な教育である4年間の日本語教育を受けている。また、Eng-M は General American を話し、特徴的な訛りはみられない。フォルマントの解析におけるサンプリングは25ms ごとにおこなった。

[3] 結果

図3のフォルマント散布図において、X軸はF2でY軸はF1である。どちらも左下に向かうにつれ、大きくなっている。なお、ここでの母音表記は国際音声記号(International Phonetic Alphabet: IPA)に従っている。そして同じ母音の部分を丸で囲んだ。赤（色が薄い方）は Eng-M、黒（色が濃い方）は Jpn-M を表している。二人のフォルマント分布はだいたい一致しているが、/o/においてかなり異なる結果となった。また、Eng-M のほうが、若干F1がJpn-Mより低い傾向にある。

[4] 考察

/o/に関して、「F1は、ほぼ開口度と対応する。狭母音ほど値が低く、順次開口度が増大するにつれて値が大きくなる。」「F2は、ほぼ舌位置と対応する。前舌母音ほど値が高く、順次舌が後退するにつれて値が低くなる。」(城生 1998, p. 76) ので、それを考え、基本母音図(図4)を参考にして母音を比較すると、Eng-Mは/o/をほぼ英語の[o]で発音しているのに対し、Jpn-Mの/o/は[ə]と[o]の中間母音である/o/だと推測される。言うならば円唇後舌め中央母音や中央準後舌円唇母音であり、これは新しい仮説としてもっと検証してみる必要があるだろう。また、Eng-Mのほうが、若干F1がJpn-Mより低い傾向にあるのは、Eng-MのほうがJpn-Mより少し狭母音傾向で発音したからだと思われる。しかし、全体的にみれば、図3のフォルマント散布図はほぼ同じ母音ごとにまとまっているので、人種・母国語による口腔および声道の身体的、器質的差異は個人間のものと同じくらいか、それより小さいと考えられる。

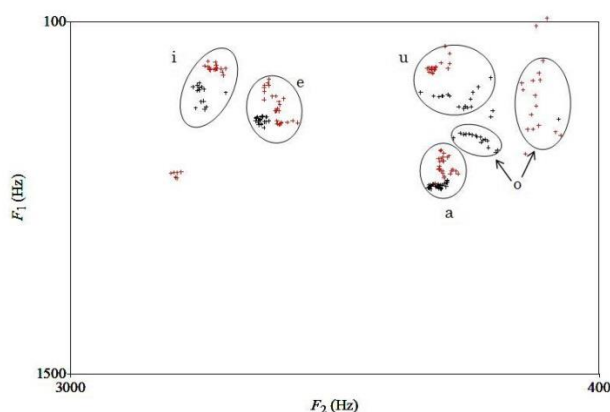


図3 Eng-MとJpn-Mの/a/, /i/, /u/, /e/, /o/のフォルマント散布図

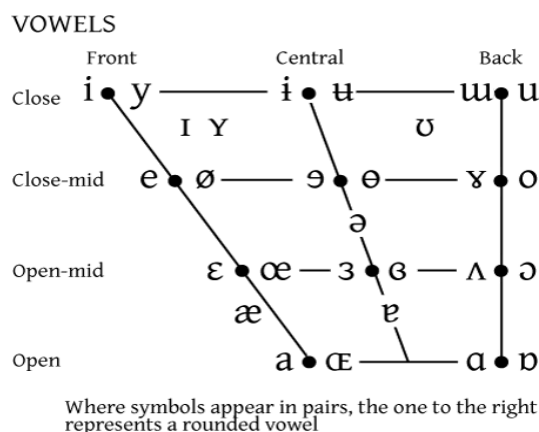


図4 基本母音図 (IPA)

3-2. 実験2

[1] 目的

多言語の母音のフォルマント分布から特徴、類似点、相違点を見つける。

[2] 手法

録音機および使用ソフトは3-1.と同じである。対象言語は、日本語、英語、中国語、韓国語、フランス語、ドイツ語、スペイン語、イタリア語である。インフォーマントはネイティブスピーカーの成人男性・女性で、各言語につき一名ずつである。ただし、日本語のインフォーマントは2名である。インフォーマントからそれぞれの言語の主に一重短母音(monophthong short vowels)を一つの母音につき、約1秒間録音して、フォルマントの解析を行い、F1-F2散布図にして比較した。これを一回行った。解析にあたっては、測定した1秒につき5ミリ秒単位でフォルマントの測定を行った。

[3] 母音

母音表記は IPA に従っている。測定する母音はその言語の基本的な母音を主に選び、母音が多い場合は一重短母音のみを選んだ。各言語ごとに測定した母音は以下の通りである。(注：日本語の u は IPA で [u] ではなく [ɯ] で表すこともある。)

| | |
|-------|--|
| 日本語 | [a], [i], [u], [e], [o] |
| 英語 | [æ], [ɑ], [ɔ], [ə], [ʌ], [ɛ], [i], [ɪ], [ʊ] |
| 中国語 | [a], [æ], [i], [o], [y], [u], [ɤ] |
| 韓国語 | [a], [ja], [ɯ], [i], [o], [jo], [ɔ], [jo], [u], [ju] |
| フランス語 | [a], [ə], [e], [ɛ], [i], [o], [y] |
| ドイツ語 | [a], [e], [ɛ], [i], [y], [o], [u], [ø] |
| スペイン語 | [a], [e], [i], [o], [u] |
| イタリア語 | [a], [e], [i], [o], [u] |

[4] 結果

それぞれの言語の母音のフォルマント散布図は図 5～図 12 に示した。

① 日本語

基本的に違う母音は F1, F2 が重なることはなかった。[a], [u] においては F1, F2 が二者間でかなり近く、あまり個人差がなかった。[o] については F1, F2 とともに 350Hz 前後の個人差があった。[i], [e] では、F1 は比較的個人差が小さかったが F2 が大きく異なっていた。

② 英語

母音の数が比較的多いためか、複数の母音が重なることはないが、かなり密集して分布していた。個々の母音は確かに F1, F2 で区別できた。[ɪ] と [ɛ] がかなり互いに近い関係にあった。また F2 においてはほとんど同じであった。/i/, /e/ の母音と /a/, /u/ の母音はかなり F2 が異なっていた。

③ 中国語

[y] と [i]、[u] と [o] の母音について、一部フォルマントの重なりがみられた。[æ] と [ɤ] ではフォルマントが狭く分布していたが、[a] では比較的広く分布しており、[i]、[u] において大きく、[y] において多少 F2 に広く分布しているのがみられた。しかし、F1 に関してはさほど広く分布していなかった。[i]、[y]、[u]、[o] がほぼ同 F1 周波数域に位置していた。

④ 韓国語

[i] を除くすべての母音でフォルマントの重なりがみられた。[a] と [ja] の一部、[ɔ] と [jo] の一部、[ɯ] と [jo] / [ju] の一部で重なりがみられたほか、[u] は [o] を内包していた。また、[jo] と [ju] のフォルマントはほぼ一致していた。[u] と [j-] (-は a, ɔ, u, o のいずれか) の母音については F2 でかなり広く分布していた。

⑤ フランス語

比較的それぞれの母音が独立して分布していた。[y] が比較的 [i] よりも [e] に近く位置していた。

⑥ ドイツ語

[i]と[e]で一部フォルマントの重なりがみられた。[y]はほぼ[i]に内包されていた。[a]はほかの母音からかなり独立していた。[i]と[e]、[u]と[o]は互いに近い関係にあった。[i]，[u]はF2に広く分布していた。

⑦ スペイン語

それぞれの母音はほぼ独立していた。[i]はF2に広く分布していた。

⑧ イタリア語

[i]，[e]が近くに位置していた。それ以外は相関性はみられない。

⑨ 言語間

[i]，[u]がF2に広く分布している言語が複数みられた。[i]が比較的F2に広く分布していた言語が8言語中3言語、[u]がF2に広く分布していた言語も8言語中3言語あった。基本的に今回のフォルマント散布図は基本母音図(図4)に対応していたが、一部英語の[a]と[o]の逆転現象や、韓国語の[u]と[o]、[ju]と[jo]の重なりなどは対応していなかった。同じ母音でも、日本語、イタリア語、スペイン語ではフォルマント分布はかなり異なっていた。

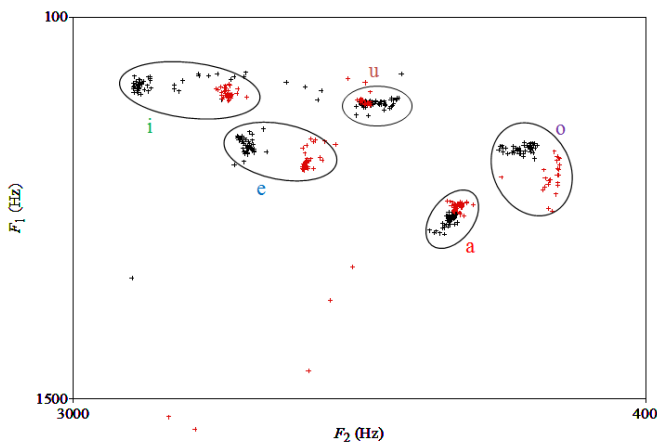


図5 日本語の母音[a]，[i]，[u]，[e]，[o]のF1-F2散布図

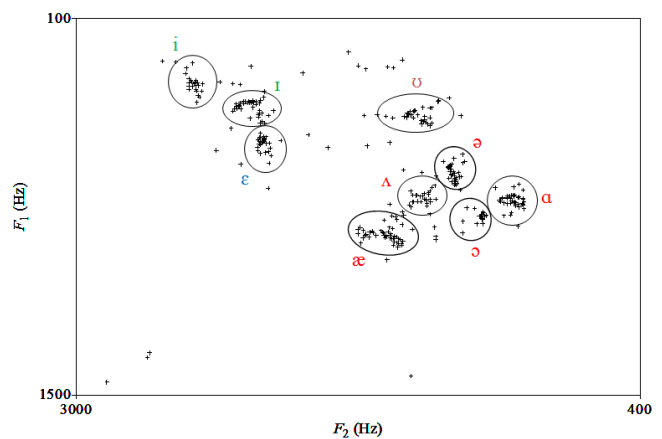


図6 英語の母音[æ]，[ɑ]，[ɔ]，[ə]，[ʌ]，[ɛ]，[i]，[ɪ]，[ʊ]のF1-F2散布図

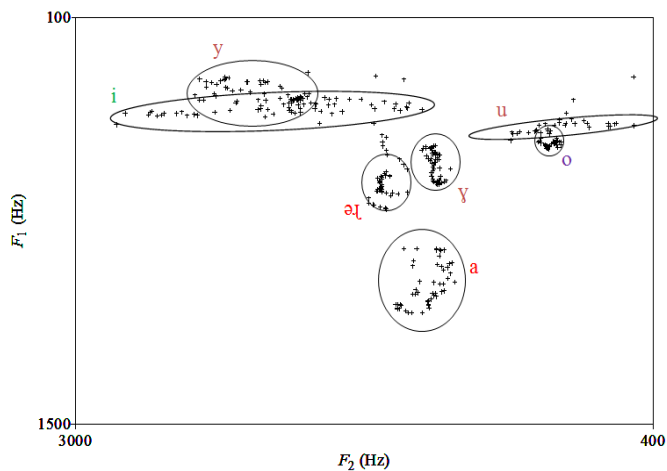


図7 中国語の母音[a]，[ə]，[i]，[o]，[y]，[u]，[ɤ]のF1-F2散布図

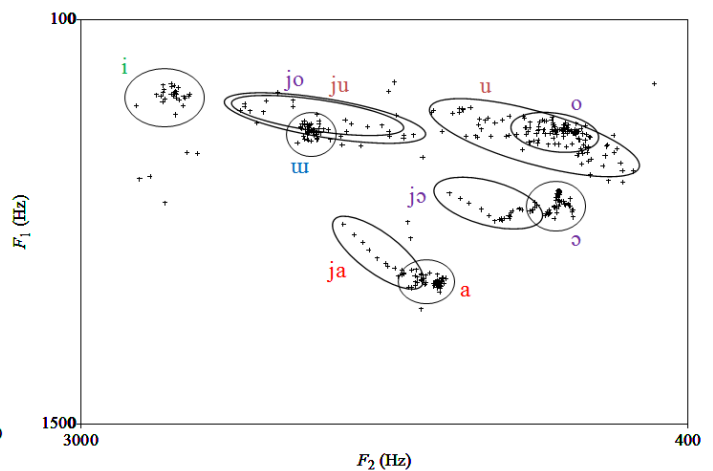


図8 韓国語の母音[a]，[ja]，[w]，[i]，[o]，[jo]，[ɔ]，[jɔ]，[u]，[ju]のF1-F2散布図

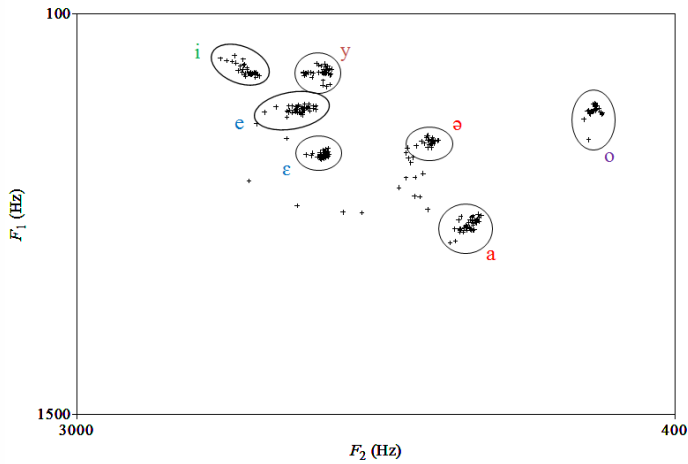


図9 フランス語の母音[a], [ə], [e], [ɛ], [i], [o], [y]のF1-F2 散布図

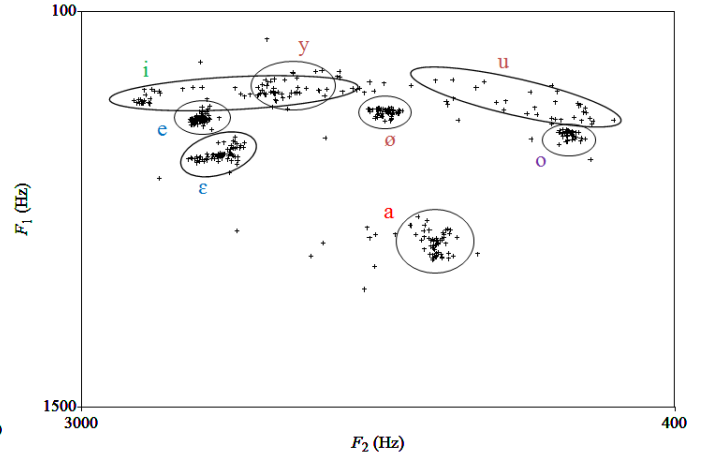


図10 ドイツ語の母音[a], [e], [ɛ], [i], [y], [o], [u], [ø]のF1-F2 散布図

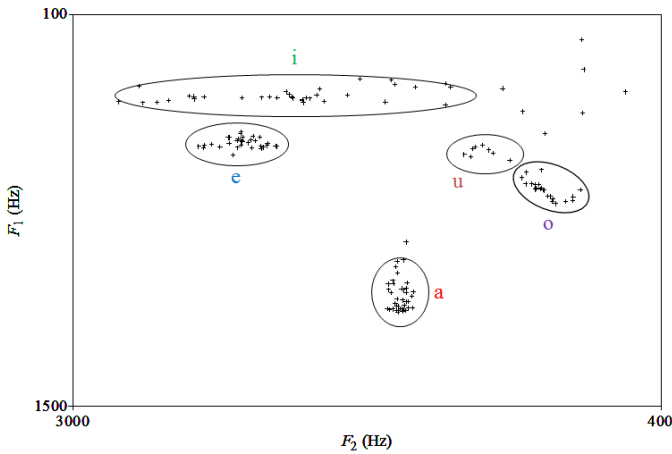


図11 スペイン語の母音[a], [e], [i], [o], [u]のF1-F2 散布図

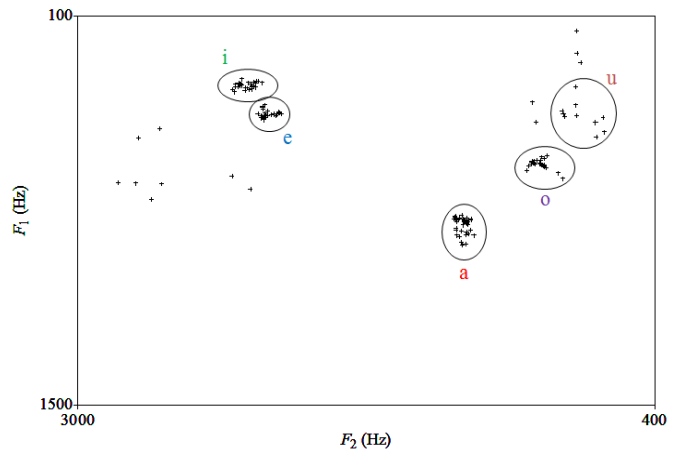


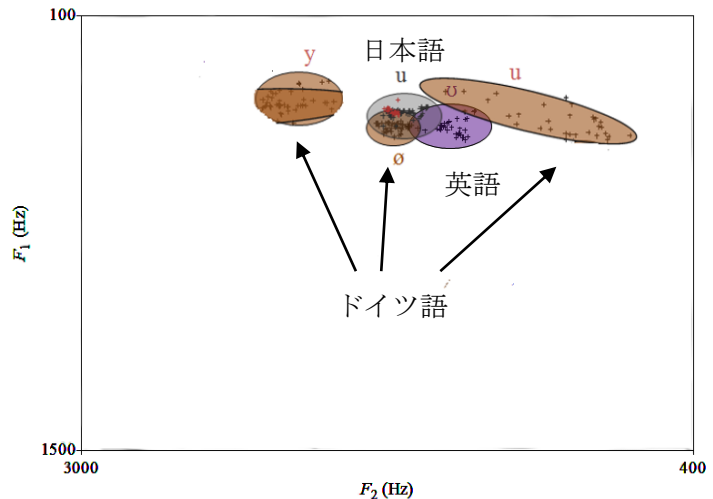
図12 イタリア語の母音[a], [e], [i], [o], [u]のF1-F2 散布図

[5] 考察

①音響音声学的考察

英語とフランス語を除く6言語は[i], [u]がF2に広く分布していたが、裏を返せばF1はあまり広く分布していなかったため、フォルマントと開口度・舌位置の関係より、それらの言語では[i], [u]は開口度で決まり、舌位置はあまり関係ないと考えられる。また、その理由としては、日本語

図13 日本語、英語、ドイツ語の/u/母音のフォルマント散布図(下)



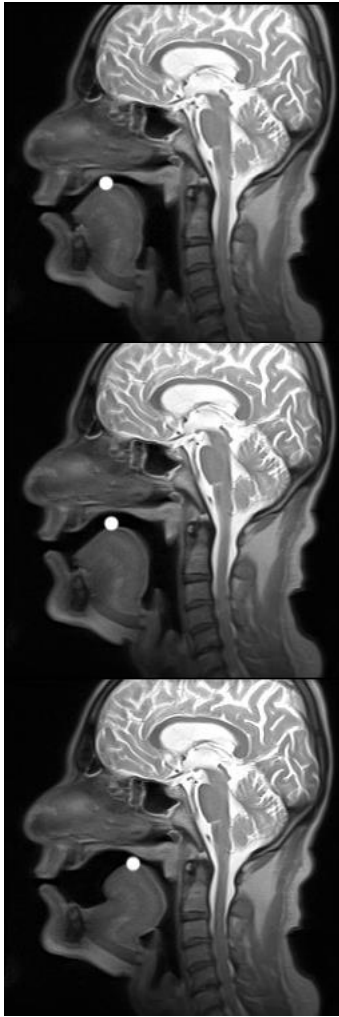


図 14 上から [y], [ø], [u] の MRI 画像

[Rietveld, A. C. M. et al., 2009]

を例に挙げれば、日本語は他言語に比べ、母音数が少ないため、母音一つ一つが範囲が広く、あいまいであるからだと考えられる。

日本語の [a] はフランス語の [a] とかなり近く、英語の [ɔ] や [ʌ] とは似ているが、 /a/ と /e/ の中間音である [æ] とはあまり近くない。F2 でも分かるように、[æ] は前舌母音なのに対し、日本語・フランス語の [a] は英語の [ɔ] や [ʌ] に近いので後舌か、あるいは中舌母音に近いからだと予測される。

また、英語だけを比較しても、[a] と [ɔ] は似ているが、[æ] とはあまり似ていない。これは F2 より、[a] と [ɔ] は後舌母音なのに対して、[æ] は前舌母音だからと考えられる。これは基本母音図とも一致している。韓国語の重なりについては、[j-] の形の母音は二つの単音から構成されており、純粋な一重短母音ではないため、調音運動・結合が関わってくると思われる。

ドイツ語の [y] と [i] の重なりについては、円唇・非円唇で区別されるはずの 2 母音が F1 ではあまり異ならなかったため、ドイツ語においては両者は結構あいまいな母音であると推測される。

/u/ 母音を見てみると、日本語は F2 が中くらいであり、これは [u] と [y] とともに似ておらず、特徴的である。そのなかでも日本語、ドイツ語、英語はドイツの 3 か国語に絞ってみると、日本語の [u] に、もっとも近いの語の [ø] で、英語では [ɔ] に近いことがわかる (図 13)。

基本母音図によると [ø] は前舌母音で、[ɔ] は後舌母音であり、日本語の [u] は IPA に従えば後舌母音なのだが、F2 の位置関係を考えると、日本語の [u]、ドイツ語の [ø]、英語の [ɔ] は中舌母音に近いと考えられる。

② 調音音声学的考察

ドイツ語の [y], [ø], [u] の MRI 画像を見てみると、確かに F2 の通り、[y] は前舌母音で、[ø] は多少中舌に近い前舌母音で、[u] は後舌母音であった (図 14)。これにより、F2 と舌位置の関連性が確認された。

4. 結論

今回の研究で、確かに性別や人種によらず、F1, F2 で母音が区別できることが確認された。また、基本母音図はフォルマント散布図に対応していた。そのことを利用して、母音を比較したが、F1, F2 分布では重なったところもあった。これは F1, F2 以外の母音決定因子が存在することを示唆している。また、一部、基本母音図の IPA で表記されているものからの予

測値とは違う結果が出た。特に日本語はあいまいな言語なので、完全に1対1で対応していないと考えられる。F1, F2だけでなく、F3, F4なども使えば、さらに細かく分類できると思われる。言語間で母音が同じIPAのものは、確かにフォルマント分布でも同じ位置にあった。

今後の最大の課題は、サンプル数を増やすことであり、それによって精度を上げていきたい。また、今回は母音を一重短母音に限ったので、今度は調音運動なども考慮しながら、子音や文中の言葉などでも調べたい。最終的には、発音障害の治療や、言語学習の教育現場に応用したい。

参考文献

- [1] Dick, B. et al. (2000) *MRI and X-ray evidence for commonality in the dorsal articulations of English vowels and liquids*, Proc. 5th Semin. on Speech Production
- [2] Gaowu Wang (2009) “A Study of Mandarin Chinese Using X-Ray and MRI”, *中国语音学报*, Vol. 2, pp. 51-58
- [3] Gaowu Wang (2008) “MRI-based Study of Morphological and Acoustical Properties of Mandarin Sustained Steady Vowel”, *J. Signal Processing*, Vol. 12, No. 4, pp. 311-314
- [4] Jacqueline Vaissiere (2011) *On the acoustic and perceptual characterization of reference vowels in a cross-language perspective*, pp. 52-59
- [5] John C.L. Ingram et al. (1997) “Cross-language vowel perception and production by Japanese and Korean learners of English”, *Journal of Phonetics*, Vol. 25, No. 3, pp. 343-370
- [6] Mokhtari Parham et al. (2000) “A Corpus of Japanese Vowel Formant Patterns”, *Bulletin of the Electrotechnical Laboratory*, Vol. 64, pp. 57-66
- [7] Pertti Palo (2011) *A wave equation model for vowels: Measurements for validation*
- [8] Rietveld, A. C. M. et al. (2009) *Algemene Fonetiek*, Bussum: Coutinho
- [9] 朝川智 (2005) 「音声の構造的表象に基づく非母語話者の英語発音分析」『電子情報通信学会技術研究報告』 Vol. 105, No. 132, pp. 25-30
- [10] 片寄晴弘 (2007) 「音楽音響信号の分析 (自動採譜概論と周波数解析)」関西学院大学理工学部 音楽情報処理, 第7回
- [11] 亀岡弘和 (2011) 「音声音響信号処理 (時間周波数解析)」東京大学大学院情報理工学系研究科, 第5回
- [12] 城生佰太郎 (1998) 『日本語音声科学』バンダイ・ミュージックエンタテインメント
- [13] 益子幸江 (2011) 「日本語の母音の音色とフォルマントについての一研究」『東京外国語大学論集』 No. 82, pp. 105-121
- [14] 楊曉安 (2005) 「日中単母音の音響音声学的分析: Vowel Formants の比較を中心として」『北海道文教大学研究紀要』 No. 29, pp. 55-64