

修士論文

音声の構造的表象に基づく
英語発音評価システムの
実用化に関する研究



2008年1月29日

指導教員 峯松 信明 准教授

東京大学大学院 新領域創成科学研究科

基盤情報学専攻 66313

鎌田 圭

内容梗概

本研究は英語発音評価システムに関する研究である。現在実現されている発音評価システムは「不一致問題」を抱えており、教育現場の信頼を勝ち得ていないのが現状である。「不一致問題」とは、被験者の身長や年齢、性別により、評価結果が不安定になる問題である。本研究では、「不一致問題」を根本的に解決可能な表象である「音声の構造的表象」に基づいた英語発音評価システムの実現を目指す。

先行研究により、英語発音評価システムの要素がいくつか提案されている。

1) 学習者の発音状態の記述、2) 大規模学習者群の発音分類、3) 学習者が自由に選択した教師発音との比較による母音矯正度推定、である。いずれも、従来手法では難解な要素であり、「音声の構造的表象」を用いることにより、実現された要素である。

本研究では、これらの要素群に対して、手法の改良や教育的妥当性の評価を行い、「音声の構造的表象」に基づく英語発音評価システムの実用化を目指す。1) については、学習者への発音状態の提示手法、すなわち「発音状態の視覚化」について検討した。2) については、教育的観点から妥当性を評価するため、専門家の協力を仰ぎ、調音音声学の視点からの発音分類を行い、構造的表象に基づく自動発音分類と比較した。3) については、母音矯正度の推定手法の高精度化を提案した。また、本研究を通して構築、改良した要素技術は東京大学柏キャンパス一般公開の場などを活用して、実地試験を行ってきた。これらの実地試験での様子について述べる。

目次

第 1 章	序論	1
1.1	本研究の背景	2
1.2	本研究の目的	2
1.3	本論文の構成	3
第 2 章	本研究の背景	4
2.1	英語教育における発音教育の重要性	5
2.2	日本語発音と英語発音の差異	5
2.2.1	母音	5
2.2.2	子音	8
2.2.3	音節	8
2.2.4	アクセント	9
2.2.5	リズム	9
2.3	実用化されている英語教育システム	9
2.3.1	英語教育システムの区分	9
2.3.2	マルチメディア技術に基づく英語教育システム	10
2.3.3	音声情報処理技術に基づく英語教育システム	10
2.4	本研究の目指す英語発音評価システム	11
第 3 章	従来の発音評価システムが抱える不一致問題	13
3.1	不一致問題の原因	14
3.2	連続 5 母音認識における不一致問題	14
3.2.1	認識実験に用いた学習用データと評価用データ	14
3.2.2	認識手法と認識結果	14
3.3	発音評価における不一致問題	15
3.3.1	発音評価に用いる英語音声	15
3.3.2	認識手法と認識結果	16
第 4 章	音声の構造的表象	18
4.1	音響特徴量ケプストラム	19
4.1.1	FFT ケプストラム	19
4.1.2	Mel ケプストラム	19
4.2	非言語的特徴による音声の歪み	19

4.3	音響的普遍構造	21
第5章	音声の構造的表象に基づく英語発音評価システムに関する先行研究	23
5.1	音声の構造的表象に基づく英語発音状態記述	24
5.1.1	日本人英語学習者の模擬音声の収録	24
5.1.2	発音状態を示す発音構造の抽出	24
5.1.3	発音状態の記述	24
5.2	音声の構造的表象に基づく英語発音分類	27
5.2.1	日本人英語学習者の模擬音声の収録	27
5.2.2	構造間距離尺度	27
5.2.3	模擬英語学習者の発音構造の分類	27
5.3	音声の構造的表象に基づく英語発音の母音矯正度推定	30
5.3.1	日本人英語学習者の模擬音声の収録	30
5.3.2	母音構造間の要素差異に基づく母音矯正度の推定	30
5.3.3	構造的表象に基づく母音矯正度	30
第6章	音声の構造的表象に基づく英語発音状態の二次元表記	32
6.1	はじめに	33
6.2	多次元尺度構成法 (MDS)	33
6.2.1	計量的 MDS と非計量的 MDS	33
6.2.2	Torgerson の方法 (計量的 MDS)	33
6.3	英語学習者の音声の収録	34
6.4	二次元表記による情報歪みの評価	35
6.4.1	英語母音発音構造の抽出	35
6.4.2	英語母音発音構造の視覚化	35
6.4.3	英語発音構造の視覚化に伴う歪みの評価	36
6.5	母音グループの分割による情報歪みの是正	36
6.6	発音分類への MDS の適用	37
6.7	本章のまとめ	37
第7章	音声の構造的表象に基づく英語発音分類の妥当性の検証	39
7.1	はじめに	40
7.2	母音図	40
7.3	日本人英語学習者の模擬音声の収録	40
7.4	英語音声学者による英語音声の母音図化	40
7.5	母音図に基づく英語発音分類	41
7.5.1	母音図に基づく母音間距離の算出	41
7.5.2	母音図に基づく学習者間距離行列からの樹形図構築	41
7.6	2つの英語発音分類の比較	42
7.7	本章のまとめ	44

第 8 章	音声の構造的表象に基づく母音矯正度推定の高精度化	45
8.1	はじめに	46
8.2	英語学習者の模擬音声の収録	46
8.3	英語音声学者による英語音声の母音図化	46
8.4	母音矯正度推定の妥当性の検証	46
8.4.1	発音の絶対的基準による母音矯正度	46
8.4.2	発音の相対的基準による母音矯正度	47
8.4.3	絶対的母音矯正度と相対的母音矯正度の相関分析	47
8.4.4	考察	47
8.5	母音矯正度推定の高精度化	47
8.5.1	構造歪みに基づく重み付け	48
8.5.2	絶対的母音矯正度と重み付き相対的母音矯正度の相関分析	49
8.5.3	考察	49
8.6	構造的表象に基づく母音矯正度の再推定	50
8.6.1	母音矯正度の再推定の構造的表象への適用	50
8.6.2	考察	51
8.7	平方二乗話による母音矯正度推定手法の検討	51
8.7.1	平方二乗話による母音矯正度推定	51
8.7.2	考察	53
8.8	本章のまとめ	54
第 9 章	音声の構造的表象に基づく英語発音評価システムの実地試験	55
9.1	はじめに	56
9.2	実地試験に用いた英語発音評価システム	56
9.2.1	英語発音評価システムへの入力	56
9.2.2	英語発音評価システムからの出力（診断結果）	56
9.3	外国語学校での実地試験	58
9.4	東京大学柏キャンパス一般公開での実地試験	58
9.5	巣鴨学園での実地試験	58
9.6	考察	59
第 10 章	結論	61
10.1	本研究のまとめ	62
10.2	今後の課題	63
10.2.1	連続発声への対応と評価音素数の拡大	63
10.2.2	英語発音分類の実体音声への適用	63
10.2.3	母音矯正度推定の推定条件の検討	63
謝辞		64
参考文献		65

発表文献	68
付録 A 柏地域の英語発音分類	i
付録 B 実地試験で提供した発音カルテ	iv
B.1 1 ページ目 (カルテの説明)	v
B.2 2 ページ目 (発音状態と発音教示)	v
B.3 3 ページ目 (発音分類)	v
B.4 4 ページ目 (発音タイプ判定)	v
B.5 5 ページ目 (発音状態)	v
B.6 タイプ別コメント	v
B.6.1 タイプ A に対するコメント	vi
B.6.2 タイプ B に対するコメント	vi
B.6.3 タイプ C に対するコメント	vi
B.6.4 タイプ D に対するコメント	vi
B.6.5 タイプ E に対するコメント	vi

目次

1.1	英語発音評価システムの教師選択インターフェース	3
2.1	Bachman's Model	6
2.2	母音図 (日本語母音)	7
2.3	母音図 (米語母音)	7
2.4	マルチメディア技術に基づく英語 CAI システム:Microsoft ENCARTA	10
2.5	音声情報処理技術に基づく英語 CAI システム:English Now!	11
3.1	単語 HMM および構造的表象での認識率	15
3.2	$P(o M)$ と $P(M o)$ による発音評定	17
3.3	構造歪みスコアを用いた発音評定	17
4.1	音声波形からのケプストラム抽出	20
4.2	Mel 周波数とその軸上に等間隔で配置された三角窓	20
4.3	スペクトルに対する水平/垂直方向の音響歪み	21
4.4	音声に内在する音響的普遍構造	22
5.1	各状態における発音状態を表す樹形図	26
5.2	回転とシフトによる構造群の接近と構造間距離	28
5.3	構造的表象に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類	29
5.4	母音実体間の差に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類	29
5.5	自動推定された母音矯正順序	31
6.1	発音構造の樹形図による視覚化	35
6.2	発音構造の MDS による視覚化	35
6.3	表示母音数による視覚化歪みの変遷	37
6.4	分割した母音グループによる二次元 MDS 母音図	38
6.5	MDS による発音分類の二次元視覚化	38
7.1	母音間距離を求める際の次元の重み	41
7.2	構造的表象に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類 [再掲]	42
7.3	母音図に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類	42
7.4	学習者間距離の相関	43

図目次

7.5	母音間距離の相関	43
7.6	知覚的に等距離となる基本 8 母音	44
8.1	絶対的母音矯正度と相対的母音矯正度の相関係数	48
8.2	絶対的母音矯正度と重み付き相対的母音矯正度の相関係数	50
8.3	推定に伴う相関係数の推移	51
8.4	母音矯正度推定時 (10 回目) の重み	52
8.5	絶対的母音矯正度と重み付き相対的母音矯正度 (絶対値和 / 平方二乗和) の相関係数	53
8.6	推定に伴う相関係数の推移 (絶対値和 / 平方二乗和)	54
9.1	巣鴨学園関係者の発音分類 (1 回目)	60
9.2	巣鴨学園関係者の発音分類 (2 回目)	60
A.1	柏地域の英語発音分類	iii

表目次

2.1	日本語・英語の主な子音の対照表	8
2.2	日本語のモーラと英語のシラブルの構造的差異	9
3.1	実験でを使用した三種類の英語発音	16
5.1	日本語母音・米語母音の置換表	25
5.2	母音置換によって模擬された6つの日本人英語発音	25
5.3	音響分析条件(5.1節)	25
5.4	母音置換によって模擬された8種類の発音状態	28
5.5	音響分析条件(5.2節)	29
6.1	音響分析条件(6.4.1節)	35
6.2	母音群別の視覚化歪み	36
9.1	英語発音評価システムに入力する単語発音	57
9.2	外国語学校での実地試験回数	58

第1章

序論

1.1 本研究の背景

本論文の表題にある英語発音評価システムとは、英語発音学習を計算機により支援するシステムであり、英語教育の一翼を担うことが期待されるものである。しかし、現在実現されている発音評価システムは須らく「不一致問題」を抱えており、教育現場への導入に懐疑的な意見も報告されている [1]。

「不一致問題」とは、被験者の身長や年齢、性別により、評価結果が不安定になる問題である。原因は学習者によって声質が異なることにある。男性の声は低く太く、女性の声は高く細い。男女が同じ単語を発生した場合、人間は容易に同じ言語情報（単語）が存在すると判断できるが、計算機は時として大変な困難を要する。同じ単語発声でも観測されるデータ（スペクトル）は異なるからである。

従来のシステムは、「不一致問題」の解決を、音声を集めることによって試みた。多量の音声を統計モデル化することによって、声質の違いの相殺を図ったのである。だが、統計モデルの想定しない被験者の音声は正しく評価できないため、「不一致問題」は残ることとなった。声質を内包した物理表象を用いる限り、根本的解決には至らないのである。

「不一致問題」を解決するには、声質を除去した表象が必要である。近年、この条件を満たす「音声の構造的表象」が提案された [2]。声質の潜む絶対量を捨象し、声質を除去した相対量を特徴量としている。この表象に基づいてシステムを構築することにより、「不一致問題」を解決した英語発音評価システムが実現できる。

「音声の構造的表象」に基づく英語発音評価システムについては種々の検討がなされているが [3][4][5][6][7][8][9]、実用化については道半ばである。本研究は「不一致問題」を解決した英語発音評価システムを実現するため、「音声の構造的表象」に基づく英語発音評価システムの実用化を目指す。

1.2 本研究の目的

本論文では、研究に先当たり、先行研究で成されている「音声の構造的表象」に基づく英語発音評価システムの要素について整理する。「音声の構造的表象」を用いることにより、学習者の発音状態¹を年齢や性別の影響を受けずに、少量の音声データで記述できる [8][9]。また、大規模学習者群の発音分類が可能²である [9]。さらに、手本となる教師の発音を学習者の好みに合わせて指定して、学習者の発音を診断することができる。選択インターフェースの一例を図 1.1 に示す。診断により矯正すべき発音の順序を学習者に提示することが可能である。

本研究では、上述の先行研究群によって成されている評価システムの各要素について、実用化に向けた検討を行う。1) 先行研究で示された話者性を排除した発音状態の記述について、学習者に提示するための発音状態の視覚化について検討を行う。2) また、音声の構造

¹母語話者の場合は、「イギリス英語発音」や「アメリカ英語発音」といった方言に相当する。

²母語話者群に対して発音分類を行えば、方言分類となる。理論上は、地球上に存在する 20 億人の英語話者の発音分類が可能である。



図 1.1: 英語発音評価システムの教師選択インターフェース

的表象について，教育的観点からの検討が不十分であるとの指摘に答え，先行研究の発音分類を英語音声学の見地からも行い，音声の構造的表象の教育的妥当性を確認する．3) そして，先行研究の母音矯正度推定について，より高精度な推定手法について検討する．4) さらには，実際に英語音声評価システムを構築し，様々な場で実地試験を行った．

1.3 本論文の構成

まず，本章において，本研究の目的について述べた．第 2 章では，本研究の背景について述べる．第 3 章では，英語発音評価システムを構築する上で障害となる「不一致問題」について述べる．第 4 章では，英語発音評価システムの核を成す「音声の構造的表象」について説明する．第 5 章で，本表象に基づく英語発音評価システムの各要素について先行研究を紹介する．第 6 章から第 8 章では，先行研究で行われた英語評価システムの各要素に関して行った改善や妥当性の検討について述べる．第 6 章では，5.1 節で紹介する英語発音状態の記述について，多次元尺度構成法を採用し，視覚化（二次元化）に伴う情報の歪みについて考慮しつつ表記する方法を検討する．第 7 章では，5.2 節で紹介する英語発音分類について，英語音声学による音声資料の母音図化を行い，母音図を通して英語発音分類の妥当性を検証する．第 8 章では，5.3 節で紹介する母音矯正度推定を高精度化する手法について提案し，母音図を通してこの手法の評価を行う．第 9 章では，要素技術から構築された発音評価システムの実地試験について述べる．最後に，第 10 章にて，本研究のまとめと今後の課題について述べる．

第2章

本研究の背景

2.1 英語教育における発音教育の重要性

本論文の表題にある英語発音評価システムは、発音教育を担うシステムである。この発音教育が、英語教育の中で特に重要な位置を占める事実の指摘について以下に記す。

外国語学習における目標として、読む能力 (Reading)・書く能力 (Writing)・聞く能力 (Listening)・話す能力 (Speaking) の4つの技能の習得が挙げられるが、近年は、コミュニケーション能力の育成に主眼が置かれている。自分の意図を相手に伝えることや、状況に合わせて言い方を変えるなどといった現実世界において実際に言語を運用する能力である。コミュニケーション能力育成においては、特に話す能力・聞く能力を重視した教育が行われる。そして、語学学習に対して音声情報処理技術による支援を行う場合、主にこれら2つの能力の学習が支援の対象となる [10]。

外国語によるコミュニケーション能力を評価する場合には、上記のような言語能力以外にも様々な観点から評価がなされるべきであり、例えば Bachman によって図 2.1 のように外国語コミュニケーションにおける様々な要素の分類がなされている [11]。そして、これらの各要素の中で、発音 (Pronunciation) はコミュニケーション能力を評価する上で非常に重要な要素となっており、下記のようなモデル式が存在している [12]。

$$\text{comm.} \simeq \text{pron.} \bullet \text{lex.} \bullet (1 + \text{syn.} + \text{rhet.} + \text{illoc.} + \text{soc.}) \quad (2.1)$$

発音能力 (pron.) と語彙力 (lex.) の二つがコミュニケーション能力 (comm.) に大きく寄与していることが示されている。すなわち、英語学習を外国語によるコミュニケーション能力の習得として考えた場合、発音教育は非常に大きな位置を占めており、その技術支援には重要な意味を持つといえる。

2.2 日本語発音と英語発音の差異

2.2.1 母音

発音には音素やアクセント、リズムといった要素が考えられるが、本研究の英語発音評価システムで取り扱う要素は音素、特に母音であるので、本節において詳しく述べる。また、母音図について、第7章と第8章で利用するため、この節で言及する。母音図とは、調音音声学や英語音声学にて母音を記述するために用いる二次元図 (図 2.2, 図 2.3 [13]) のことである。まずは、母音図について述べるために、調音音声学における母音の分類基準について述べる。

i) 調音音声学における母音の分類基準

舌の高低 母音生成時の舌の位置によって、母音は分類可能である。日本語母音 /い/ ¹ /う/ を発音するときは、舌が高く盛り上がっているときは口腔内は狭くなる。/あ/ を発音す

¹日本語母音は /a/, /i/, /u/, /e/, /o/ と表記するのが慣習であるが、本論文では米語母音との混同を避けるため、/あ/, /い/, /う/, /え/, /お/ と表記した。

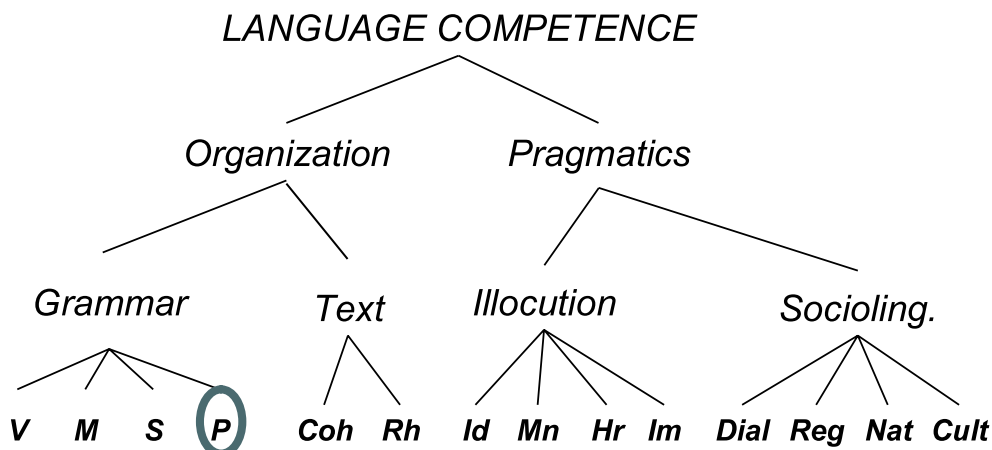


図 2.1: Bachman's Model

るときは、舌が下がって低くなり、口腔内が広がる。このように、母音は「舌の高低」によって分類することができる。

舌の高低によって母音は次の3つに分類される。日本語母音/い/、/う/のように、舌を高くして作られる母音を高母音 (high vowel) といい、/あ/のように舌を低くして作られる母音を低母音 (low vowel) という。/え/、/お/は、舌の高さは中程度で、中母音 (mid vowel) と呼ばれる。なお、中母音を、やや高めの中高母音 (mid-high vowel) とやや低めの中低母音 (mid-low vowel) に分けることもある。

舌の前後 舌の位置は、前述の舌の高低と、本節の舌の前後に分けることができる。日本語母音/い/、/え/を発音するときは、舌は全体的に前方に置かれ、舌の前部 (前舌面) が後部 (後舌面) よりも高くなる。対して、/う/、/お/を発音するときは、舌は全体的に後方に下がり、舌の後部は前部よりも高くなる。このように、母音は「舌の前後」によっても分類することができる。

舌の前後によって母音は次の3つに分類される。日本語/い/、/え/のように、舌の前部が高くなる母音を前舌母音 (front vowel) といい、/う/、/お/のように舌の後部が高くなる母音を後舌母音 (back vowel) という。また、前舌母音と後舌母音の中間の母音を中舌母音 (central vowel) といい、/あ/はこれに属する。ただし、ここでは/う/を後舌母音としたが、実際には舌が若干前寄りであり、純然たる後舌母音とは言いがたい (中舌母音に近い)。/あ/は基本的には中舌母音であるが、前後の音素によって、前舌母音寄りや後舌母音寄りとなる。

なお、高母音・中母音・低母音、前舌母音・中舌母音・後舌母音といった3段階の分類を行ったが、実際には中間段階が無数に存在する。

円唇性 母音の音色は、舌の位置のほか、唇の形にも左右される。日本語母音/お/は、唇は丸めて突き出して発音される。日本語母音/い/は、唇を左右に引いて発音される。/い/の舌の構えで、唇は丸めて突き出して発音すると、/い/とは異なる音色の母音となる。

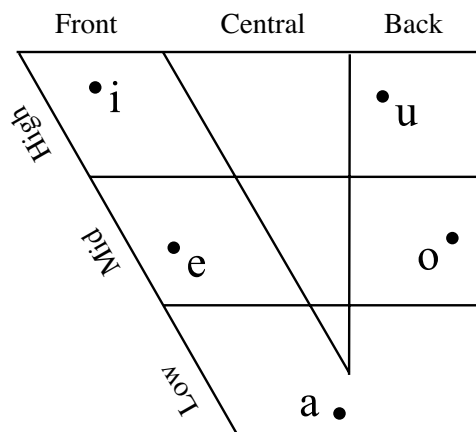


図 2.2: 母音図 (日本語母音)

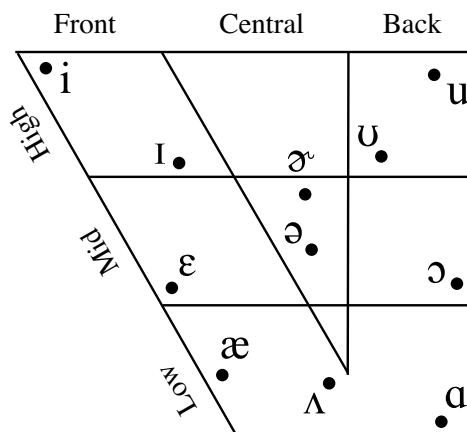


図 2.3: 母音図 (米語母音)

唇を丸めて発音する母音を円唇母音 (round vowel) といい、丸めずに発音する母音を非円唇母音 (unround vowel) という。日本語母音の /お/ は円唇母音であり、/あ/、/い/、/う/、/え/ は非円唇母音である。

ii) 母音図

i) 節で示した分類基準を、図で示したものが母音図 (vowel chart) である。母音図は、縦軸に「舌の高低」(上が高母音、下が低母音)、横軸に「舌の前後」(左が前舌母音、右が後舌母音)を取った四角形である。母音を母音図上に表すことにより、異なる音色の母音同士の位置関係が一目でわかるという利点がある。日本語で用いられる母音と、英語で用いられる母音を母音図上に表示したものが図 2.2、図 2.3 である。ただし、二重母音に関しては省略してある。なお、英語の母音は方言による違いが大きいですが、以下ではアメリカ英語の一般的な方言 (米語, General American) を扱う。

iii) 日本語母音と英語母音の差異

日本語母音はわずか5つであるのに対し、英語母音は、短母音5つ (/ɪ, ʊ, ε, ʌ, æ/) と長母音4つ (/i, u, ɔ, ɑ/) の合計9つ、二重母音は /eɪ, ɔɪ, aɪ, aʊ, oʊ/ の5つ、さらに弱母音 (/ə/) と、アメリカ英語独特の r 色の母音 (/ɝ/ など) を入れると全部で22種類になる [14]。

一般に、常用している言語の音素の範囲内の違いには鈍感であり、異なる音素にまたがる際には、それが仮に音響的に小さな差異であっても比較的敏感であるといわれている [15]。例えば低母音は、日本語では /あ/ のみだが、英語では /ɑ, æ, ʌ/ があるため、日本人学習者はこれら3つの母音の差異には鈍感である。/ɪ, i/ や /ʊ, u/ などの差異についても同様である。

日本語における長母音は音質的に見ると短母音が並んだものとしてとらえることが出来るが、英語の長母音は短母音を長く発音したものではなく、例えば beat の [i] と bit の [ɪ] は発音の仕方も音質も異なる別種の音である。二重母音も同様に、[aɪ] は [a]+[ɪ] ではなく、

表 2.1: 日本語・英語の主な子音の対照表

調音位置 調音方法		唇音		歯	歯茎	後部 歯茎	そり 舌	硬口蓋	軟口蓋	口蓋垂	声門
		唇	唇歯								
閉鎖音	日	p / b			t / d				k / g		ʔ
	英	p / b			t / d				k / g		ʔ
摩擦音	日	ɸ			s / z			ç			h
	英		f / v	θ / ð	s / z	ʃ / ʒ					h
破擦音	日				ts / dz	tʃ / dʒ					
	英					tʃ / dʒ					
鼻音	日	m			n			ɲ	ŋ	ɴ	
	英	m			n				ŋ		
弾き音	流音	日			r						
		英			l ɹ		ɹ				
閉鎖音	渡り音	日	w					j	w		
		英	w					j	w		

[aɪ] で1つの音である．そのため，長母音と二重母音は短母音とは別の音として取り扱われる．また，英語には弱母音 (schwa) と呼ばれる，弱く発音されるときのみ現れる母音が存在する．日本語では基本的にどの母音も同じ強さで発音されるため，日本人はこの弱母音の発音が特に苦手であると言われている．

2.2.2 子音

子音に関しては，調音位置と調音方法によって表 2.1 のように分類される [14]．無声/有声の対で示しており，発音記号により表記してある．日本人が特に苦手であると言われる英語の /l/ (発音記号では [l]) と /r/ (アメリカ英語は [ɹ]，イギリス英語は [ɹ]) は，日本語ではラ行の子音である [r] に置き換えられることが多い．また，/f, v/ や /θ, ð/ は日本語にない子音であるため，/v/ と /b/，/f/ と /h/，/θ/ と /s/，/ð/ と /z/ などの混同が起こりやすい．

2.2.3 音節

英語は音節，あるいはシラブル (syllable) と呼ばれるものを発声の基本的単位とする言語であり，日本語は音節より小さな発声の単位であるモーラ (mora) を基本的単位とする言語である [16]．表 2.2 にモーラとシラブルの構造的差異を示す．日本語の母音数は5種類であるが，英語の母音の種類は約20種類となっており，シラブル/モーラの構造的差異

表 2.2: 日本語のモーラと英語のシラブルの構造的差異

モーラ	基本は 母音 (V), 子音+母音 (CV) 他に 特殊拍 (撥音 (N), 促音 (Q), 長音) が存在する
シラブル	母音を中心にその前後に 0 個以上の子音が連結した形をとる。 最長シラブルは CCCVCCCC.

から, モーラの種類数は約 100 であるが, シラブルは約 10,000 種類数を持つと言われる。

2.2.4 アクセント

日本語では, 音の高さの変化によりアクセントを表現する「高さアクセント」であり, 音声情報処理においてはピッチのみを用いて記述される。一方の英語においては, 「強さアクセント」と呼ばれる。ピッチ, パワー, 持続時間, 母音の音質などがアクセントによって変化する。英語のアクセントを以降, 特に「強勢」(stress)という言葉を用いる。「強勢」とは, ある音節を発音するに当たって音源である呼気が強くなったりその量が多くなると喉頭や調音器官が緊張して調音のエネルギーが強くなり, 聞き手が感じる音の大きさ (loudness) が増大する現象をいう。強勢を受けた音節はピッチが高まり, 音が長めになる。強勢は強強勢 (strong stress) と弱強勢 (weak stress) とに二分され, 強さアクセントでは全ての音節はいずれかを受ける。語中にある音節に置かれた強強勢は単語強勢 (word stress) と呼ばれる。また, 文中の特定の音節が持つ強勢は文強勢 (sentence stress) と呼ばれる。語義を持つ内容語 (content word) 強い文強勢を受け, 機能語 (function word) と呼ばれる語義が希薄で主として内容語同士の文法的関係を示す働きをする語の文強勢は弱い。

2.2.5 リズム

リズムにおける等時性に関しても日本語と英語は異なる。日本語は一つ一つのモーラが等間隔で発音される「モーラ拍のリズム」と呼ばれる特徴を持っている。一方の英語では, ドイツ語, ロシア語と同様に強い強勢が等間隔に繰り返される「強勢拍のリズム」を持つといわれる。特に, 英語のリズムを形成するものとして, ある一つの強勢音節から次の強勢音節までを「脚」(foot) という。

2.3 実用化されている英語教育システム

2.3.1 英語教育システムの区分

英語教育システムの区分として, ここでは 2 つ紹介する。一つは CAI (Computer Assisted/Aided Instruction) システムである。CAI システムとは指導者が学習者を指導する時の主要な機能をコンピュータに代行させ, あらかじめ用意された教材 (コースウェア) に従って学習者個々に応じた最適な学習指導を行うシステムのことである。もう一つは CALL



図 2.4: マルチメディア技術に基づく英語 CAI システム:Microsoft ENCARTA

(Computer Aided Language Learning) システムである。CALL システムとは、コンピュータが生身の先生を取って代わるのではなく、コンピュータを使って言語学習の促進を図ることが主な目的のシステムのことである。

2.3.2 マルチメディア技術に基づく英語教育システム

一般的に商用の英語 CAI ソフトには音声情報処理技術を用いたものよりも、ビデオ映像などのマルチメディア技術を伴ったものが多い。その一例である Microsoft ENCARTA インタラクティブ英会話 [17] を図 2.4 に示す。他の教材と同様に、音声処理技術よりもビデオ映像を見た後にそれについての設問に答える形式が主な内容となっている。ビデオ映像の再生中にテキストを字幕として表示することが出来たり、問題文として提示されている英文について実際の音声を聞いて確認できるなど、マルチメディア技術を生かした構成になっている。本教材ではスピーキングのセクションも用意されているが、学習者の発音に対するフィードバックは無く、音声処理は特に行われていないと思われる。また、千葉大学の協力でメディア開発センターが英語の CALL 教材 [18] を発表しているが、これもリスニング中心で、リスニング後に設問に答える形式の教材となっている。

2.3.3 音声情報処理技術に基づく英語教育システム

音声情報処理技術を用いた英語 CAI ソフトの一例である English Now! [19] を図 2.5 に示す。この例のように、学習者が発声した音声を波形として表示し、教師音声の波形と比較して表示する形式のものが多い。「調音」「強弱」「高低」「リズム」等といった観点から比較が行われる。どの部分がどの程度異なっているのかは示されるが、分析結果によるフィード

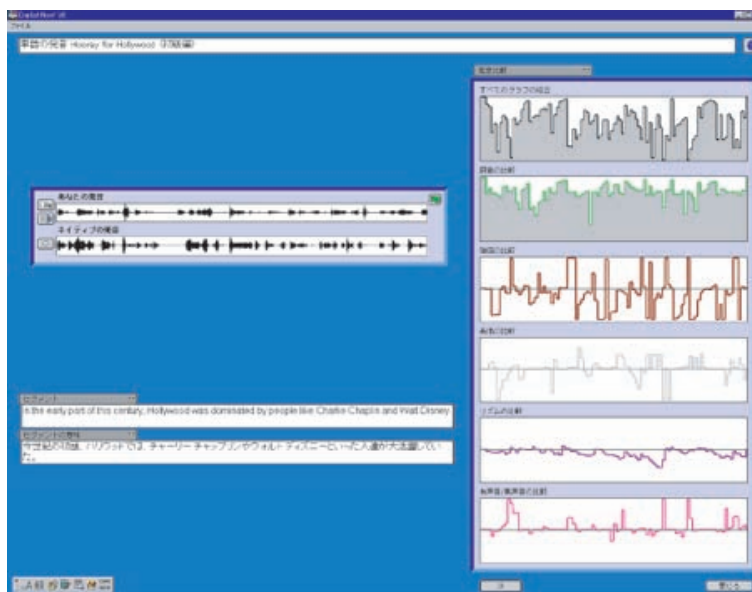


図 2.5: 音声情報処理技術に基づく英語 CAI システム:English Now!

バックの意味するところが不明確で分かりにくく効果は薄いように思われる。本教材では「リズム」の評価も行われるが、他の音声認識技術を用いた英語 CALL システムである [20] などでは、韻律的特徴よりも音韻学習が学習内容の中心になっている。学習者へのフィードバックに関しては、コンピュータアニメーショントーキングヘッド Baldi [21] 等のように、唇や舌、顎の動きを模擬した三次元アニメーションを用いてよりわかりやすい発音教示を行う例もある。

2.4 本研究の目指す英語発音評価システム

従来までの発音教育、特に英語教師 1 人に対して数十人の生徒という「一対多」の状況で行われる場合、そこで行われてきた教育は、生徒に対する英語という言葉の「紹介」であったと考察することができる。英語教師は発音に関して一般に音声学及び音韻論といった知識を習得しているが、音声学は言語を構成する個々の音を記述する科学であり、音韻論は一言語内の音のセットや並びに内在する関係や規則を記述する科学である。これらの科学は英語という言葉の音や規則を「紹介」のための非常に優れたツールではある。しかし、英語学習としては、言語の「紹介」だけでは不十分である。英語という言葉を外国語として習得する過程において、学習者がどのような状態にあるのかを記述し、その状態にある学習者にはどのような訓練が効果的であるのか、という「一対一」の個別指導が必要である。このような教育を行うためには学習者の発音状態を正確に記述できるような枠組みが必要である。

学習者の発音を波形やスペクトルで表示することは、現在の音響音声学 + PC というツールを用いれば可能であり、これらのツールを基にした発音学習アプリケーションは前節で

紹介したように広く開発が行われている。しかし、波形やスペクトルは発音の善し悪し以外にも性別・年齢・体格・マイク・収録環境などの様々な情報が含まれる。発音学習という観点から見ればこれらは全て雑音であり、利用者環境とシステムとの間での相性問題（不一致問題）が常につきまとう。このような技術を基にした発音学習支援は結果として不安定なシステムとなることは明らかである。さらに、音声学的な視点からの学習者の記述では、基本的には「単音」や「発話」を単位とした記述となる。つまり、個々の発話に対する善し悪しに対しては記述可能であるが、学習者が今どのような状態にあるのかという「学習者」を単位とした記述に対しては、音声学的な観察眼のみでは不十分であるといえる。学習者の記述において求められるものは、単音や発話の記述ではなく、学習者を「英語音生成システム」として見たときの、そのシステムを規定するパラメータ群を考え、そのパラメータを推定する形で学習者の発音状態を記述することである。それは即ち、音韻論的な視点から学習者を眺めることであり、これまでの音声学的観察眼から音韻論的観察眼への質的变化が必要であるといえる。

学習目標の設定についても検討する必要がある。従来までの語学学習システムは母語話者との音響的照合を行うことで評価を行っているが、近年の語学教育においては、*native-sounding* な発音ではなく、*intelligible* な発音が求められるようになってきている。学習者が目指す目標が多様となっているため、目指す目標をそれぞれの学習者が決めることができ、かつ、現在の状態から目標に到達するまでの道のりが学習者に対して明確に見える必要がある。

発音学習への技術支援の枠組みにおいて必要とされるものは下記の3点にまとめられる。

- 1) まず、学習者の「現在」の発音状態が記述可能ということである。学習者を「英語音生成システム」として捉え、学習者の発音状態を正確に記述することが求められる。同時に、学習者が見ても十分理解可能であるような形で記述する必要がある。
- 2) 次に、実環境において安定に動作可能ということである。相性問題が原則的に起こらない技術でなければならない。当然低年齢者などを対象とした場合にも問題なく動作する必要がある。
- 3) そして、現実的な学習目標が設定可能ということである。学習者全てが母語話者のような発音を目指しているわけではなく、学習者が各人の学習意欲に応じた目標を適切に設定可能であるものでなければならない。

このような英語発音評価システムを構築するには、音声から性別・年齢・体格・マイク・収録環境といった情報を除去し、発音状態を示す情報を抽出する必要がある。近年、上記のような枠組みの構築を目指して、「音声の構造的表象」という新しい音声の物理表象が提案されている。音響的普遍構造は、話者の性別・年齢や音響機器の違いなどの非言語的特徴を表現する次元を消滅させることで定義される新しい音声の物理表象であり、既に発音教育応用へ向けた構造的表象に着眼した研究がいくつか行われている。本研究が目指す英語発音評価システムは、「音声の構造的表象」を用いることにより上記の3点を実現した評価システムである。

第3章

従来の発音評価システムが 抱える不一致問題

3.1 不一致問題の原因

従来の発音評価システムは音響特徴量ケプストラムを直接参照するため、「不一致問題」を抱えている。ケプストラムには音声には話者の声道形状の特性，マイクロフォンなどの音響機器の特性，テレビ・ラジオなどの背景雑音といった非言語的特徴が内包されており，発音を純粋に評価することを妨げる要因となっている。

従来の発音評価システムでは，基本的に，大量の音声を用いて個々の言語事象を統計的にモデル化することで「不一致問題」に対処してきた。しかし，非言語的特徴が内包した特徴量を利用していることに変わりはなく，結果として，話者によっては正しく機能しないシステムとなっていた。結局，この手法では「不一致問題」の解決には到らなかった。

この「不一致問題」に対して，音声認識の分野では，種々の適応技術（例えば話者適応）が用いられている。しかし，これらの適応技術を発音評価に適用すると，下手な発音に対して高いスコアを与えるような適応が免れない。これは，話者の違いと発音の上手下手とがスペクトル包絡という同一物理現象に基づくにも関わらず，両者を切り分ける術を持っていなかったことが根本原因である。

3.2 連続5母音認識における不一致問題

本節では，従来手法（ケプストラムを特徴量に用いたHMM音声認識）による連続5母音音声認識実験を通して、「不一致問題」の一例を見る。また，後述する「音声の構造的表象」（第4章）に基づく音声認識の結果を示し，「不一致問題」を解消していることを付記する。

3.2.1 認識実験に用いた学習用データと評価用データ

認識モデルを作るための学習用データには，日本人の成人8名（男女各4名）の日本語5母音連続発声¹（120単語）を用いた。評価用データには，学習用データとは異なる日本人成人8名（男女各4名）の日本語5母音連続発声（120単語）を用いた。これに加え，不一致問題を引き起こす身長の高い人間の音声を用意するため，評価用データから音響分析による音声変換[22]によって，身長が低い人間の音声を作成した。

3.2.2 認識手法と認識結果

学習用データを用いて単語HMMを作成し，変換された評価用音声の認識を行った。結果は身長が低くなるに従って，認識率は低下した（図3.1，従来手法）。これは従来手法が話者の違いに対応できず，「不一致問題」を抱えることを示している。なお，「音声の構造的表象」に基づく音声認識[23]では，身長の変化による認識率の大きな変化は見られず，「不一致問題」に対応可能であることを示している（図3.1，構造的表象）。従来手法ではこの

¹ 「あいうえお」「うおいえあ」など， ${}_5P_5 = 120$ 通りの連続発声

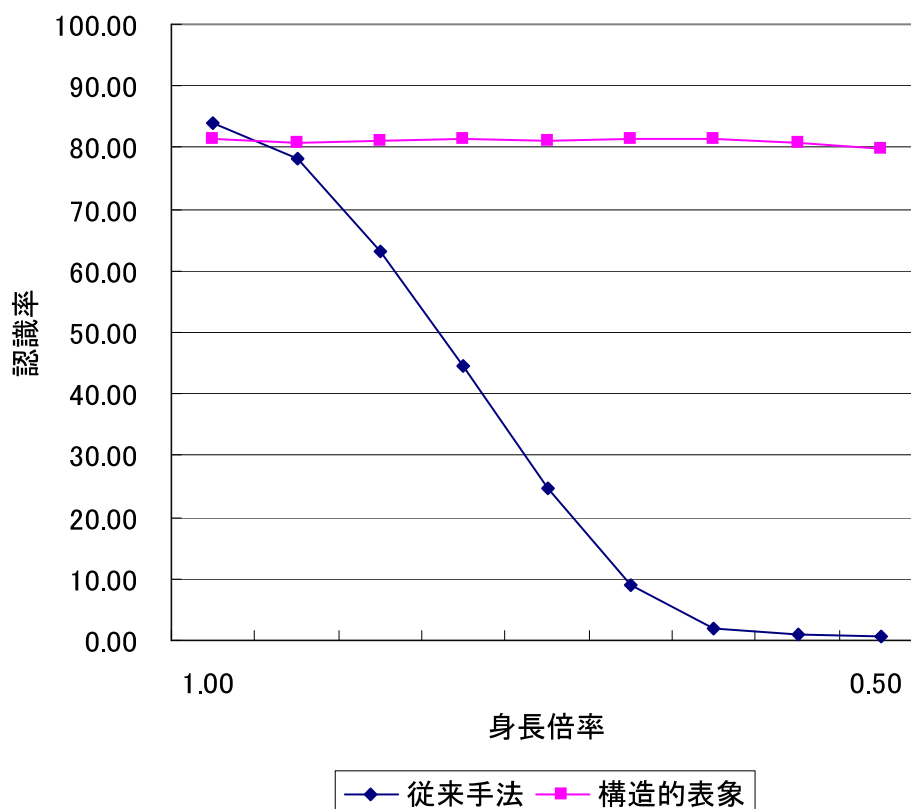


図 3.1: 単語 HMM および構造的表象での認識率

ような「不一致問題」に対応するために様々な話者適応技術が提案されており、音声認識率を改善することはできる。しかし、これらの話者適応技術が発音評価においては不自然な結果を残す場合がある。この例について次節で述べる。

3.3 発音評価における不一致問題

本節では、「不一致問題」解消のために広く用いられるパラメータによる発音評価を行い、発音分析の従来手法が「不一致問題」を抱えていることの一例を示す。また、「音声の構造的表象」に基づくことにより、「不一致問題」を解決可能である実験結果について述べる。

3.3.1 発音評価に用いる英語音声

3.2 節と同様「不一致問題」を起こす英語音声を用意して実験する。まず、評価対象の音声として、日本人男性 (NM) の熟練した英語音声 (A) を用意した。NM は、英語劇経験者であり、英語発音習熟度は極めて高い。次に、比較対象の音声として、この NM が故意に日本語訛り (カタカナ英語発音) で発音した習熟度の低い英語発音 (B) を用意した。さらに、もう一つの比較対象の音声として、英語教師である女性の米語母語話者 (USA/F12)

表 3.1: 実験で使用した三種類の英語発音

話者	USA/F12(F)	NM(A)	NM(B)
性別	女	男	男
年齢	約 50	36	36
マイク	Sennheiser	特価品	特価品
録音室	防音室	リビング	リビング
AD	SONY DAT	PowerBook	PowerBook
習熟度	perfect	good	Japanized

の英語音声 (F) を用意した。用意した音声を表 3.1 にまとめる。また、結果の考察用に、男性の米語母語話者の英語音声 (M) も用意した。

3.3.2 認識手法と認識結果

熟練した男性英語音声 A を、習熟度の一致しない B、話者が異なる M、話者も性別も異なる F と比較する。比較対象の音声から作成した HMM によりスコアを求める。なお、「不一致問題」を強く引き起こす比較が A と F である。スコアは、1) モデル (B, F, M) と A 間の尤度スコア $P(o|M)$ 、2) モデル (B, F, M) と A 間の事後確率スコア $P(M|o)$ を用いた。なお、2) の事後確率スコアは、モデルと入力話者間の相性や整合性を正規化する（「不一致問題」を解消する）目的で発音評価では広く使われている。

$P(o|M)$ 及び $P(M|o)$ による評価結果を図 3.2 に示す。それぞれのスコアを元に、評価対象 A の存在位置を内分点として示している。話者正規化前の $P(o|M)$ では、A は限りなく B に近くなっており、話者の一致 / 不一致を考えれば、当然の結果であると言える。さて、 $P(M|o)$ であるが、入力話者とモデルとの相性を正規化した後のスコアであると考えられる [24]。つまり、母語話者の性別に関わらず、同じ値が得られるように意図したスコアである。しかし、結果は図 3.2 のとおり、異なるスコアとなった。これは、従来の「不一致問題」を解決する技術が不安定であることを示している。

一方、5.2.2 節にある「音声の構造的表象」に基づく構造間距離を利用したスコアによる結果が図 3.3 である。母語話者の性別に影響されることなく、ほぼ同じ値が得られている。図 3.3 の下半分は、3.3.1 節の他に収録した米語話者、日本人の音声を評価した結果を同一軸上に示したものである。軸の上部は女性、下部は男性の評価結果を示している。米語話者の中に日本人が存在しているが、これは英語発音が堪能なバイリンガル話者である。図より分かるように、「バイリンガル以外のすべての日本人が話者 NM (A) 以上に、話者 NM (B) に近い」と判定されている。このような現象は、単純なスペクトル照合では不可能である。また、「音声の構造的表象」に基づく手法が発音習熟度のみに着眼し、「不一致問題」を解決していることを示す結果であると言える。

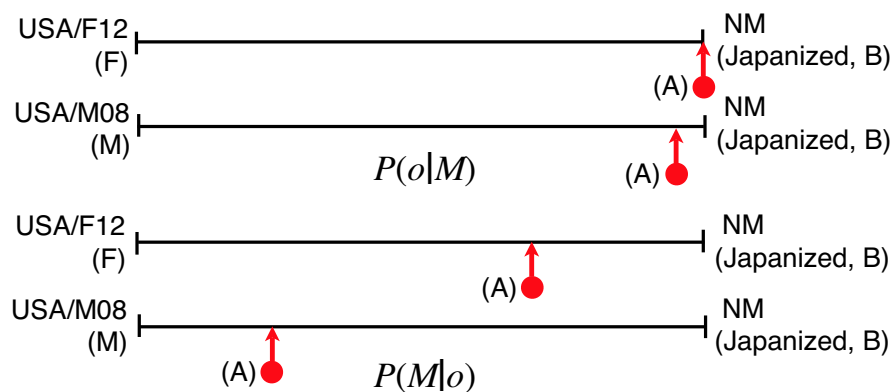


図 3.2: $P(o|M)$ と $P(M|o)$ による発音評定

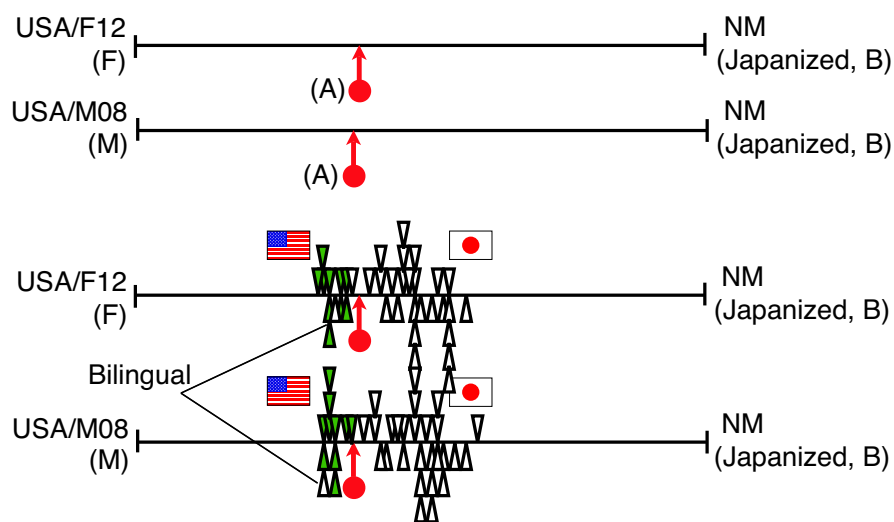


図 3.3: 構造歪みスコアを用いた発音評定

第4章

音声の構造的表象

4.1 音響特徴量ケプストラム

本節では、音響音声学で広く用いられる特徴量ケプストラム (Cepstrum) について述べる。「音声の構造的表象」においても、特徴量「音響的普遍構造」抽出の過程でケプストラムは利用する。

4.1.1 FFT ケプストラム

音声分析において、音声波形からケプストラムを抽出するまでの様子を図 4.1 に示す。まず音声波形から、数十ミリ秒程度のフレームを切り出し、その区間に対して離散フーリエ変換 (Discrete Fourier Transform; DFT) を施し、スペクトルを抽出する。その後、対数パワースペクトルに対して逆離散フーリエ変換 (Inverse DFT; IDFT) を施したものがケプストラムである。このケプストラムのうちの低次項のみを離散フーリエ変換すると、スペクトル包絡 (Spectrum Envelope) が得られる。声道管の共鳴によって強められた周波数をフォルマント周波数と呼び、これらはスペクトル包絡の山の部分におよそ相当するが、音声の音韻的特徴はフォルマント周波数によく表れる。つまりケプストラムは、音声の音韻的特徴を効率良く表すことのできるパラメータである。

4.1.2 Mel ケプストラム

人間の音の高さの感覚は Mel 尺度と呼ばれるが、これは音の周波数に対してほぼ対数に近い特性を示し、人間の周波数分解能は低い周波数ほど細かく、高い周波数ほど粗いことが知られている。これをケプストラムに反映させた音声特徴量が提案されている。MFCC (Mel-Frequency Cepstrum Coefficient) はその一つである。MFCC は、図 4.2 に示すように Mel 周波数 (Mel 尺度化された周波数) 軸上に等間隔で配置された三角窓を用意し、フィルタバンク分析を行うことで求められる。尚、Mel 周波数 f_{mel} は周波数 f [Hz] に対して、

$$f_{mel}(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (4.1)$$

などの周波数ウォーピングを施すことで得られる。各窓毎に、対応する周波数帯域のパワースペクトルを求め、それに窓の大きさの重みを付けて和をとることで Mel スペクトルが得られる。これに離散コサイン変換を施すことで、MFCC が求められる。

これらの特徴量を直接参照することによって、「不一致問題」は引き起こされる。次節からは、特徴量「音響的普遍構造」の抽出過程を、「不一致問題」を回避する仕組みを交えて述べる。

4.2 非言語的特徴による音声の歪み

まずは、「不一致問題」を引き起こす、音声に混入する非言語的特徴の歪みについて整理する。音声に混入する非言語的特徴は、加算性雑音・乗算性歪み・線形変換性歪みの3種

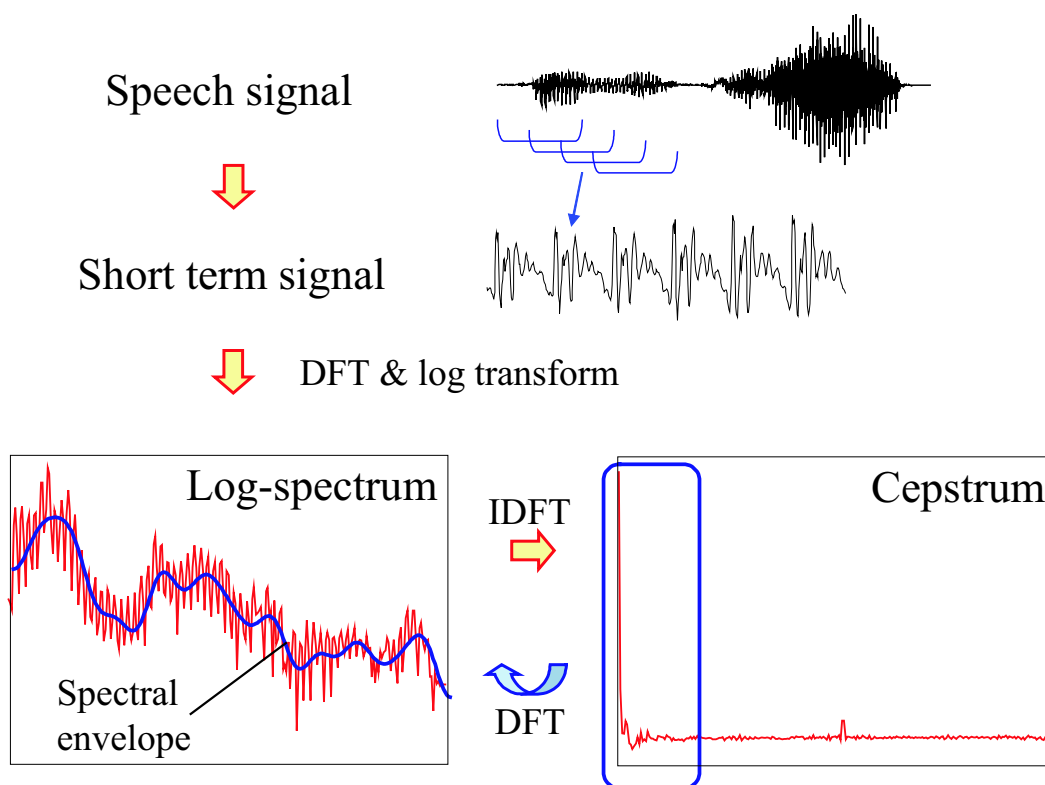


図 4.1: 音声波形からのケプストラム抽出

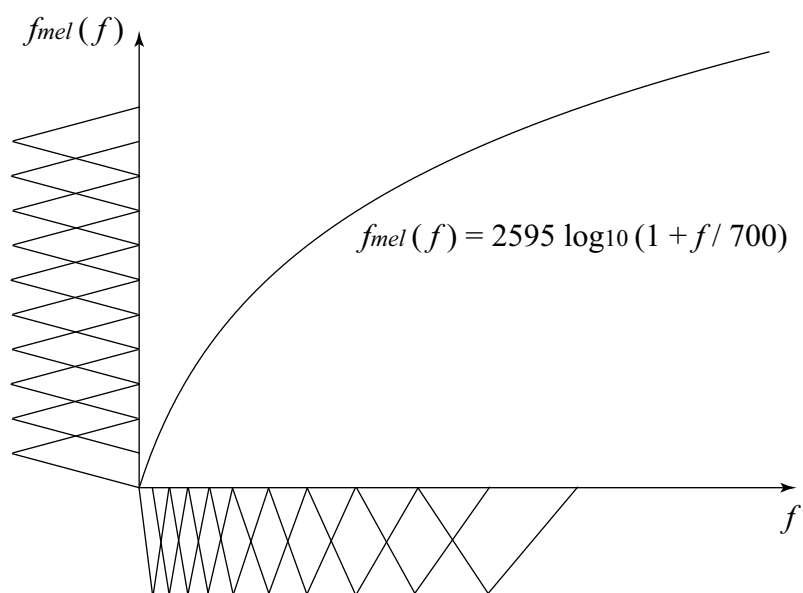


図 4.2: Mel 周波数とその軸上に等間隔で配置された三角窓

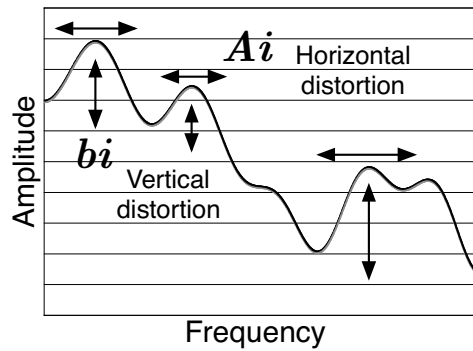


図 4.3: スペクトルに対する水平/垂直方向の音響歪み

類に大別される．このうち，音声に必ず混入する歪みは乗算性歪みと線形変換性歪みの 2 つである．

加算性雑音とは，時間軸上の加算で表現される雑音であり，テレビ・ラジオなどの背景雑音が挙げられる．これに関しては，発音評価システムでの収録環境を整えることで回避することが可能であるため，本研究では取り扱う対象とはしない．

乗算性歪みは，スペクトルに対する乗算で表現される歪みであり，元のスペクトル特性に対して伝達関数を一つかけ合わせたもの（すなわちフィルタ）に相当する．これには，マイクロフォンなどの音響機器の伝送特性が挙げられる．また，話者の声道形状の違いの一部も乗算性歪みとして扱われる [25]．乗算性歪みを音響特徴量ケプストラムへの影響で表現すると，ケプストラムベクトル c に対するベクトル b の加算 $c' = c + b$ となる．

線形変換性歪みは， c に対する行列 A の乗算 $c' = Ac$ で表現される歪みである．話者の声道長の差異，さらには聴取者の聴覚特性の差異を表すために，対数スペクトルに対して周波数ウォーピングが施されるが，単調増加かつ連続である周波数ウォーピングは， c に対する A の乗算で表すことができる [26][27]．すなわち，声道長の差異，聴覚特性の差異は近似的に線形変換性歪みとして扱うことができる．

音声を計算機で取り扱う場合，発生源である話者と収録機器であるマイクロフォンを介することを考慮すれば，回避不可能な音響的歪みは乗算性歪み ($c' = c + b$) と線形変換性歪み ($c' = Ac$) である．よって，音声に混入する避けることのできない非言語的特徴の歪みは， c に対するアフィン変換（一次変換） $c' = Ac + b$ で近似的に表現される．図 4.3 は，アフィン変換が対数スペクトルに与える影響を示したものである．対数スペクトルの垂直変化が乗算性歪み，水平変化が線形変換性歪みである．これが最も簡素な非言語的特徴による音声変動の数学モデルである．

4.3 音響的普遍構造

空間内に存在する N 点に対して， ${}_N C_2$ 個のすべての二点間距離を求めると，その N 点で張られる構造は一意に規定される [2]．この構造は距離行列の形で表現される．音声事象をすべてケプストラム空間上の点として捉え，すべての二点間距離の計算すると，音声事

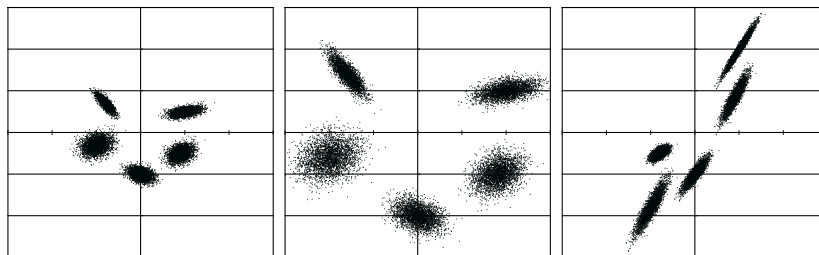


図 4.4: 音声に内在する音響的普遍構造

象群が成す幾何学構造を規定することになる．しかし，この幾何学構造には非言語的特徴（アフィン変換）による歪みを内包している．

このアフィン変換歪みを踏まえ，二分布間距離としてアフィン不変性を持つバタチャリヤ距離¹を採用する．二つの分布の確率密度分布関数をそれぞれ $p_1(x)$, $p_2(x)$ とすると，

$$BD(p_1(x), p_2(x)) = -\ln \int_{-\infty}^{\infty} \sqrt{p_1(x)p_2(x)} dx \quad (4.2)$$

と表される． $0 \leq \int_{-\infty}^{\infty} \sqrt{p_1(x)p_2(x)} dx \leq 1$ を確率として解釈すれば，式（4.2）は自己情報量となり，単位は [bit] となる．また，二つの分布がガウス分布で表現されているとき，

$$BD(p_1(x), p_2(x)) = \frac{1}{8} \mu_{12}^T \left(\frac{\sum_1 + \sum_2}{2} \right)^{-1} \mu_{12} + \frac{1}{2} \ln \frac{|(\sum_1 + \sum_2)/2|}{|\sum_1|^{\frac{1}{2}} |\sum_2|^{\frac{1}{2}}} \quad (4.3)$$

となる．なお， μ_{12} は $\mu_1 - \mu_2$ である．

二分布間距離としてバタチャリヤ距離を用いることにより，音声事象群が成す幾何学構造はアフィン変換に対して不変となる．先にも触れたが，バタチャリヤ距離はアフィン不変性を有しており，式（4.3）に対して共通の如何なるアフィン変換を施しても距離不変²である．例えば，五分布群は全て等しい距離行列を持つ図 4.4 のような三つの分布群を考える．これらの分布群はユークリッド平面上で観測すれば異なる歪んだ構造として観測されるが，バタチャリヤ距離行列（構造）は不変となる．これはバタチャリヤ距離を用いて事象間距離を定義することは，空間そのものを歪め，非ユークリッド空間で事象群を観測することに相当するからである．すなわち，構造が歪まないように，空間を歪めているわけである．このような話者 / マイクに普遍に観測される構造を，「音響的普遍構造」と呼ぶ．この「音響的普遍構造」を用いて英語発音評価システムを構築することにより，「不一致問題」を根本から解決したシステムを実現可能である．

¹同様にアフィン不変性を持つ距離尺度として，カルバック・ライブラ距離やヘリンガー距離が挙げられる．

²なお，バタチャリヤ距離はアフィン変換（線形変換）だけでなく，一対一対応，かつ，全空間で微分可能な非線形変換に対しても不変であることが示されている [29] ．

第5章

音声の構造的表象に基づく
英語発音評価システムに
関する先行研究

5.1 音声の構造的表象に基づく英語発音状態記述

本章では、英語発音評価システムの要素技術として3つの先行研究を紹介する。

5.1.1 日本人英語学習者の模擬音声の収録

英語劇及び発音指導経験を持つ日本人男性に米語11単母音を/bVt/の形で発声させた。ここで有意味語が存在しない場合は、/t/を/d/とするなどの処置を行い、11単母音の収録を1回ずつ行った。次に、/bVto/という形で日本語母音を5回ずつ収録した。これらの音声資料を用いて、一部の米語母音を日本語母音と置換する形で、日本人英語発音を模擬した。表5.1に用いた母音置換表を示す。母音置換の組み合わせは、日本人が混同しやすい組み合わせであり、混同の様子は図2.2、図2.3の母音図を見ることにより窺うことができる。異なる米語母音が同一の日本語母音と置換される場合は、同一母音・異発声のサンプルを用いた。例えば、/bʌt/と/bæʔ/を置き換える場合は、2回の/bato/の発声を用いてそれぞれの発音を置き換える。これらの置換パターンを用いることで、1)日本語母音に全て置換した発音、2)/あ/に近い母音を矯正、3)/い/に近い母音を矯正、4)/う/に近い母音を矯正、5)/え/に近い母音を矯正、6)全ての母音を矯正、と6種類の発音状態(S1~S6)を規定した。各置換パターンをまとめると表5.2のようになる。

5.1.2 発音状態を示す発音構造の抽出

収録音声から目視により母音部分を切り出し、表5.3に示す音響分析条件の下でケプストラムパラメータを求め、ガウス分布化した。データが少量であるため、発音構造が不安定となることが考えられる。これの対処として、各ガウス分布にはMAP推定(最大事後確率推定)を施した[8]。そして、すべての分布間のバタチャリヤ距離を求め、発音構造(11×11距離行列)を抽出した。

5.1.3 発音状態の記述

11×11の距離行列で表現される構造を、階層的クラスタリング手法の一つであるボトムアップクラスタリング(Ward法)により樹形図化したものが図5.1である。Ward法とは、クラスター融合の際に発生するクラスター間距離の歪みが最小になるように距離行列を更新する手法である。分析対象のデータはあくまでも模擬音声であるが、完全なる日本語母音の使用による英語発音(S1)から米語発音(S6)に至るまでの一つの遷移の様子が、明確に示されている。

表 5.1: 日本語母音・米語母音の置換表

日本語母音	↔	米語母音
/あ/		/æ/, /ʌ/, /ɑ/, /ɚ/, /ə/
/い/		/i/, /I/
/う/		/u/, /ʊ/
/え/		/ε/
/お/		/ɔ/

表 5.2: 母音置換によって模擬された 6 つの日本人英語発音

	ɑ	æ	ʌ	ə	ɚ	ɪ	i	ʊ	u	ε	ɔ
S1	J	J	J	J	J	J	J	J	J	J	J
S2	A	A	A	A	A	J	J	J	J	J	J
S3	A	A	A	A	A	A	A	J	J	J	J
S4	A	A	A	A	A	A	A	A	A	J	J
S5	A	A	A	A	A	A	A	A	A	A	J
S6	A	A	A	A	A	A	A	A	A	A	A

A : 米語母音を使用, J : 日本語母音で置換

表 5.3: 音響分析条件 (5.1 節)

サンプリング	16bit / 16kHz
窓	窓長 25ms, シフト長 4ms
パラメータ	FFT ケプストラム (1~10 次元)
HMM	1 混合 monophones (全角分散行列)
トポロジー	3 状態 1 ガウス分布

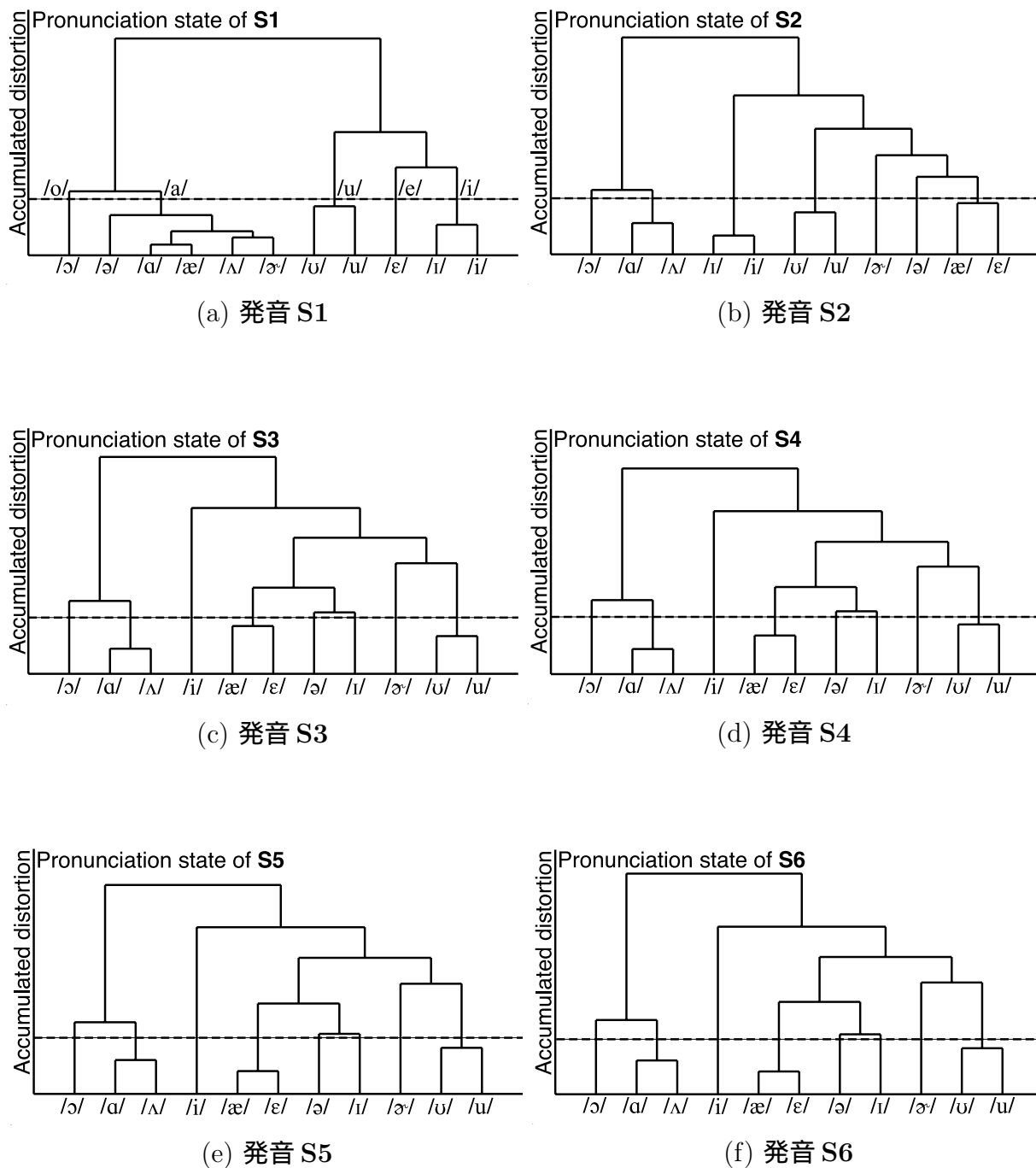


図 5.1: 各状態における発音状態を表す樹形図

5.2 音声の構造的表象に基づく英語発音分類

5.2.1 日本人英語学習者の模擬音声の収録

帰国子女、或いは英語劇経験者の日本人12名（男性6名、女性6名；各話者をA～Lとする）より、11種類の米語母音と5種類の日本語母音を5.1.1節と同様に、各々1回、5回ずつ収録した。これらの米語母音と日本語母音を同一話者内で置換することにより、様々な発音状態を規定する。本実験では、表5.4に示す置換パターンを用い、話者ごとに8種類の発音状態（P1～P8）を用意した。なお、米語・日本語間の母音置換の対応は表5.1の通りである。複数の英語母音を1種類の日本語母音で置き換える場合は、日本語母音は異なる発音を用いるものとした。このようにして模擬された96種類（12話者×8発音）の発音状態に対して、下記で定義される距離尺度に基づいて距離行列化し、樹形図化することで分類を試みた。

5.2.2 構造間距離尺度

学習者が発声する11母音の母音間距離をすべて求めることにより、 11×11 の母音間距離行列で表現される構造（11角形）を定義することが出来る。11角形と11角形の構造間距離を規定することができれば、それは学習者と学習者の距離を、純粋に発音訛に着眼した形で（非言語的特徴を無視した形で）定義することとなる。そこで、収録音声から目視により母音部分を切り出し、表5.5に示す音響分析条件の下でケプストラムパラメータを求め、MAP推定を用いて分布化した。そして、模擬された96種類（12話者×8発音）の発音状態に対して、すべての音素分布間のバタチャリヤ距離を算出し、合計96個の学習者発音構造（母音間距離行列）を得た。

構造間距離は図5.2に示すように、一方の構造を回転及びシフトして他方に近づけた際に観測される、対応する二点間距離の総和の最小値として規定できる¹。この最小値は、二つの距離行列間のユークリッド距離として近似できることが示されており、この距離尺度を使用する[7]。

$$D(S, T) = \sqrt{\frac{1}{m} \sum_{i < j} (s_{ij} - t_{ij})^2} \quad (5.1)$$

ここで、 m は母音数の11であり、 S_{ij} 、 T_{ij} は構造（ $m \times m$ 距離行列） S 、 T の (i, j) 成分である。

5.2.3 模擬英語学習者の発音構造の分類

学習者間距離が定義されたことにより、 96×96 の学習者間距離行列を算出することができる。この学習者間距離行列にWard法によるボトムアップクラスタリングを適用することにより、96人の学習者を分類することが可能である。

¹なお、一次変換によって距離行列が不変であるため、 A の乗算（声道長差異）は回転として、 b の加算（マイクなどの音響機器差異）はシフトとして解釈される。

表 5.4: 母音置換によって模擬された 8 種類の発音状態

	ɑ	æ	ʌ	ə	ɝ	ɪ	i	ʊ	u	ɛ	ɔ
P1	J	J	J	J	J	J	J	J	J	J	J
P2	A	A	A	A	A	J	J	J	J	J	J
P3	J	J	J	J	J	A	A	A	A	A	A
P4	A	A	J	J	J	A	A	J	J	A	A
P5	J	J	A	A	A	J	J	A	A	J	J
P6	A	J	A	J	A	J	J	J	J	A	A
P7	J	A	J	A	J	A	A	A	A	J	J
P8	A	A	A	A	A	A	A	A	A	A	A

A : 米語母音を使用, J : 日本語母音で置換

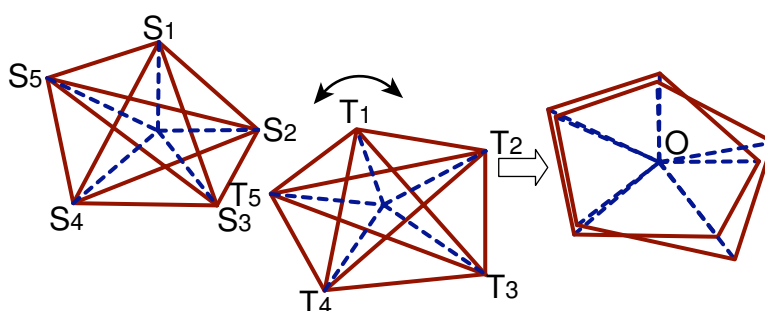


図 5.2: 回転とシフトによる構造群の接近と構造間距離

Ward 法によるボトムアップクラスタリングにより分類した結果を図 5.3 に示す。樹形図のリーフノードにおける数字 (1~8) が発音状態を意味し, A~L が話者を意味する。樹形図を見ると, おおよそ同じ発音状態が固まる形で 8 つのクラスが形成されている。この結果より, 構造的表象を用いることで, 話者や性別などの非言語的特徴に影響されることなく, 発音の分類, 即ち, 言語的な分類が可能となることが実験的に示された。

なお, 学習者間距離を (v_i^S を話者 S の母音 i の分布とする)

$$D'(S, T) = \sqrt{\frac{1}{m} \sum_i \text{BhattacharyaDistance}(v_i^S, v_i^T)} \quad (5.2)$$

のように, 異なる話者間で母音を直接的に比較する形で定義すると, 図 5.4 に示すように完全なる話者分類となった [28]。言い換えれば, 絶対的物理量の「差」を見ることで話者分類となり, 「差の差」を見ることで発音分類となった訳である。

表 5.5: 音響分析条件 (5.2 節)

サンプリング	16bit / 16kHz
窓	窓長 25ms , シフト長 1ms
パラメータ	FFT ケプストラム (1~10 次元)
HMM	1 混合 monophones (全角分散行列)
トポロジー	3 状態 1 ガウス分布

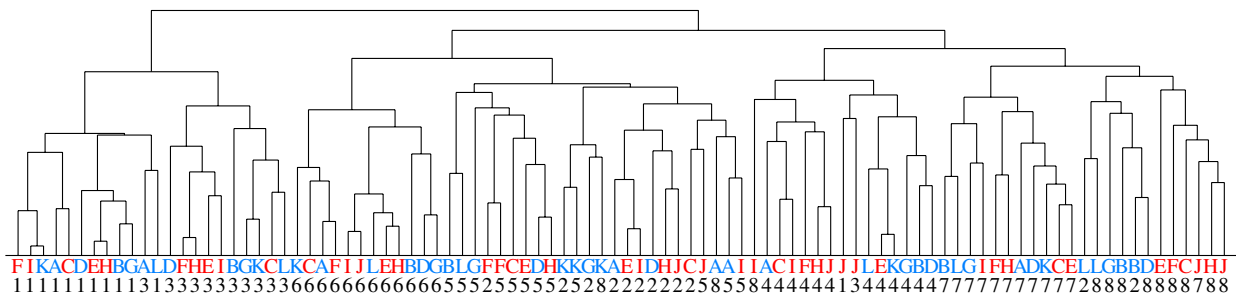


図 5.3: 構造的表象に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類

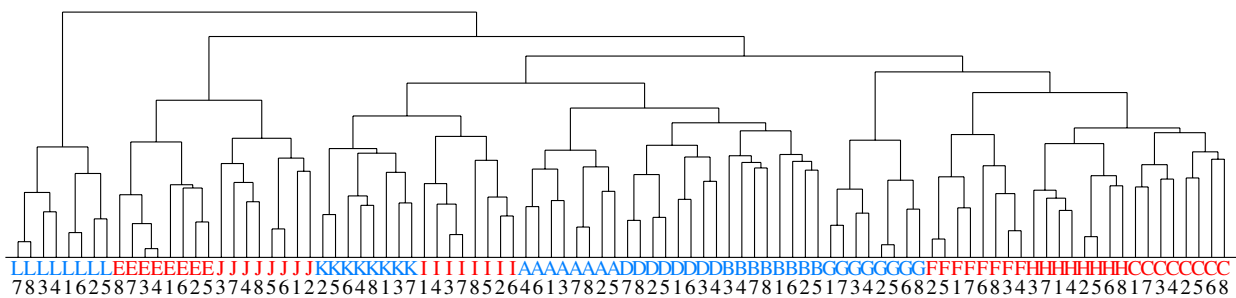


図 5.4: 母音実体間の差に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類

5.3 音声の構造的表象に基づく英語発音の母音矯正度推定

5.3.1 日本人英語学習者の模擬音声の収録

使用した模擬音声は、5.2.1 節と同様の 96 種類の 11 発音である。本節では、表 5.4 (母音置換表) の P1 ~ P7 は日本人米語学習者の発音状態に相当し、P8 は米語教師の発音状態に相当することを念頭に置く。

5.3.2 母音構造間の要素差異に基づく母音矯正度の推定

構造抽出の手順は 5.2.2 節と同様である。96 個の発音構造を得る。

異なる二つの構造を比較し、構造を構成する各要素 (構造の頂点) の構造歪みを求める。各要素が持つ構造歪みは各要素の構造を歪ませることへの貢献度を表し、二つの構造を一致させるための障害の大きさを表す指標である。米語教師の発音構造と日本人米語学習者の発音構造を比較した場合、各要素の構造歪みは当該母音に関する母音矯正度と捉えることができる。

二つの構造 S と T の、要素 v の構造歪みを次式で定義する。

$$d(S, T, v) = \sum_{j=1}^m |s_{vj} - t_{vj}| \quad (5.3)$$

m は構造を構成する要素の数 (本節では 11 となる)、 s_{ij} と t_{ij} は構造 (距離行列) S と T の要素 i と要素 j の要素間距離 (バタチャリヤ距離) を表す。

5.3.3 構造的表象に基づく母音矯正度

式 (5.3) を前述の模擬音声の発音構造に適用した結果が、図 5.5 である。図中の P1 ~ P8 は、12 話者平均構造である。各図は米語教師相当の構造 P8 と日本人学習者状態の P1 ~ P7 を比較したときの各母音の構造歪みを表している。教師と学習者の比較であるので、各母音の構造歪みは各母音に対する母音矯正度と捉えることができる。図の縦軸は各母音の構造歪み (母音矯正度) を表し、色づけられているものは日本語母音に置換した母音、白抜きのは米語発音の母音である。

半分の母音が置換された発音状態である P2 ~ P7 について見てみると、日本語に置換された母音の多くが、米語母音よりも高い母音矯正度を示している。また、すべての母音が置換された P1 に関して、日本人が発音を苦手とする /ɔ̃/、/æ/、/ə/、/ɪ/ が高い結果となった。このことより、構造的表象に基づいて各母音の母音矯正度を導出可能であることが実験的に示された。

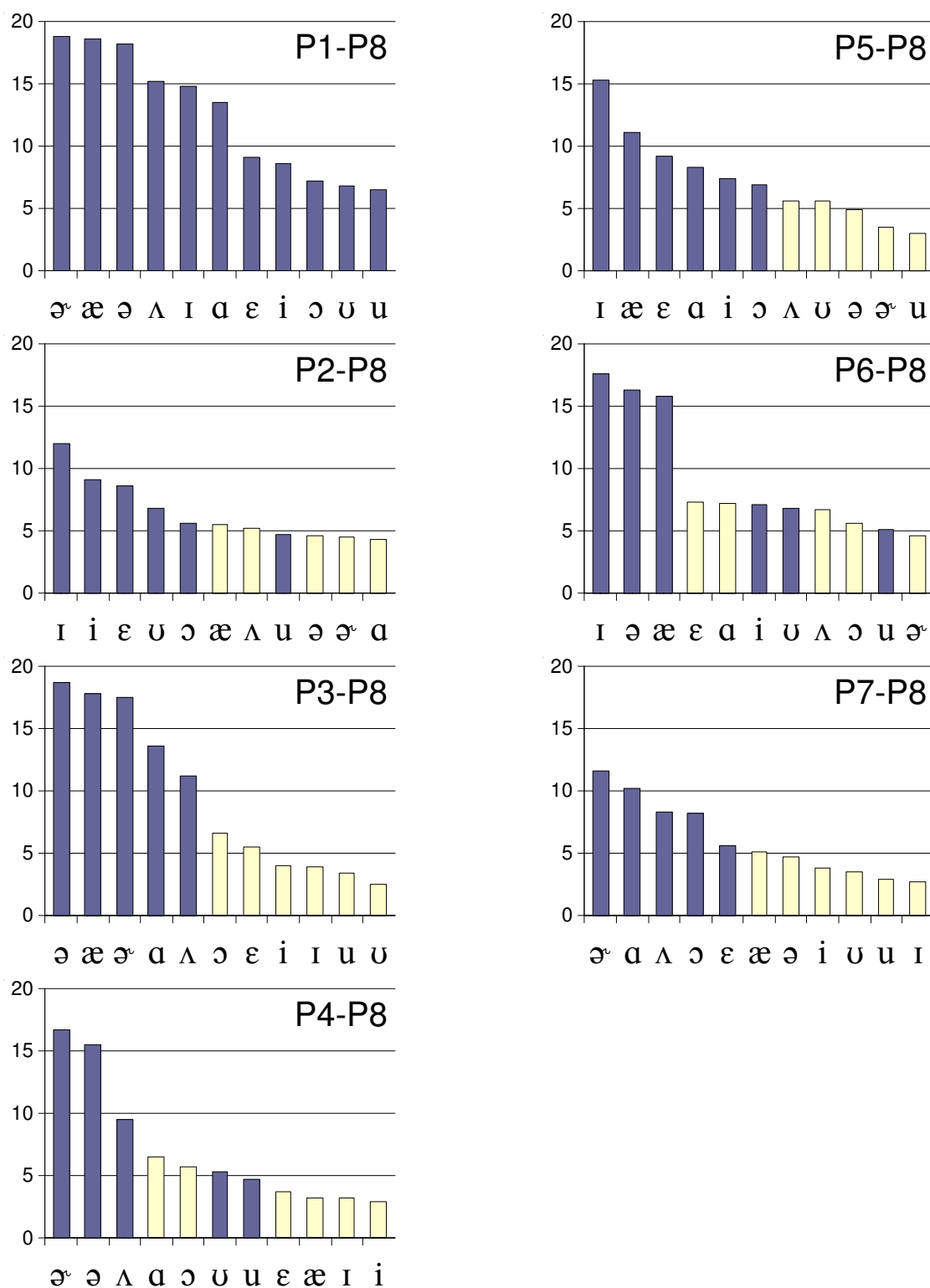


図 5.5: 自動推定された母音矯正順序

第6章

音声の構造的表象に基づく 英語発音状態の二次元表記

6.1 はじめに

音声の構造的表象を用いることで、英語学習者の発音状態を非言語的特徴に影響されることなく記述できる。この発音状態は母音間距離行列の形で表現されるため、図 6.1 のような樹形図の形で表現することができる。しかし、発音状態を表す母音間距離行列（発音構造）は元来、高次元（ $n > 2$ ）空間における幾何学構造であるため、図 6.1 のように二次元平面上で表示する際には、情報の欠落や情報の歪みが不可避である。この二次元化に伴う情報の劣化により、間違った情報を英語学習者に提示する可能性がある。

本章では、英語発音構造を表す距離行列を二次元視覚化する方法として、計量的多次元尺度構成法（計量的 MDS）を採用する（図 6.2）。樹形図は読解のための知識を必要とするため難解であるとの学習者の声に配慮し、母音間の遠近が直接読み取れる MDS を選択した。また、樹形図を用いた場合、情報の欠落についての評価が難しいが、MDS を用いた場合、情報の歪みを定量的に評価することは容易である。本章では、計量的 MDS を用いて、情報の歪みを抑えつつ、かつ、教育的に妥当な発音構造を視覚化する方法について考察する。

6.2 多次元尺度構成法（MDS）

6.2.1 計量的 MDS と非計量的 MDS

多次元尺度構成法（multidimensional scaling; MDS）とは多変量解析の手法の一つである。各標本を低次元（本章では二次元）の座標に配置する。この際に、可能な限り次元の少ない多次元空間に、標本間の距離が非類似度（本研究では母音間距離が相当する）に最もよく一致するように配置する [30]。

MDS は再現する非類似度の距離尺度によって二分される。間隔尺度・比例尺度を再現する場合は計量的 MDS、順序尺度を再現する場合は非計量的 MDS と区分される。計量的 MDS には、後に紹介する Torgerson の方法のほか、INDSCAL（Carroll & Chang）、数量化 IV 類、K-L 型数量化（林）が挙げられる。非計量的 MDS には、MDSCAL（Kruscal）、Shepard の方法、SSA（Guttman）、ALSCAL（Takane, Young & de Leeuw）が挙げられる。

6.2.2 Torgerson の方法（計量的 MDS）

本節では、計量的 MDS の一つである Torgerson の方法を紹介する [31]。対象とする非類似度行列を S とし、行列の大きさを $m \times m$ 、 S の要素を s_{jk} とする。

i) 加算定数問題

非類似度行列 S は $s_{jk} \geq 0$ 、 $s_{jj} = 0$ を満たすことが求められる。しかし、非類似度の距離尺度が間隔尺度の場合はこれらを満たさない場合がある。この場合、条件を満たすため

の前処理を行う必要がある（加算定数問題）．本稿で扱う非類似度行列（距離行列）は比例尺度であり，この作業を行う必要はないため，この問題の詳細については省略する．

ii) 座標行列の獲得

行列 S から，次式で b_{jk} を計算し， b_{jk} を要素とする行列 B を求める．

$$b_{jk} = \frac{1}{2} \left(\frac{1}{m} \sum_j s_{jk}^2 + \frac{1}{m} \sum_k s_{jk}^2 - \frac{1}{m^2} \sum_j \sum_k s_{jk}^2 - s_{jk}^2 \right) \quad (6.1)$$

なお，行列 B を求める式を行列演算で表すと次のようになる．

$$B = -\frac{1}{2} H^T S^{(2)} H \quad (6.2)$$

$$S^{(2)} = (s_{jk}^2), \quad H = E - \left(\frac{1}{m} \right)$$

式(6.1)と式(6.2)は等価である．行列 B を求めた後，次式からなる行列 B に関する固有値問題を解く．

$$B \mathbf{x}_t = \lambda_t \mathbf{x}_t \quad (t = 1, 2, \dots, m) \quad (6.3)$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq \dots \geq \lambda_m$$

得られた固有値 λ_t と固有ベクトル \mathbf{x}_t から座標行列 A を次式で求める．

$$\mathbf{a}_t = \sqrt{\lambda_t} \mathbf{x}_t \quad (t = 1, 2, \dots, r) \quad (6.4)$$

$$A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r)$$

これにより r 次元ユークリッド空間における座標行列 A が定まる．本稿では，二次元座標の獲得を行うため r は 2 である．

6.3 英語学習者の音声の収録

3年に渡る東京大学柏キャンパス一般公開(9.4節)を通して，英語学習者(主に日本人)の音声をのべ697名から収録した．収録内容は英語11単語と日本語5単語を1回ずつである．収録した単語は5.1.1節と同じである．収録した話者は4歳から84歳まで広い年齢層に渡る．録音環境が万全でなかったため，事前に聴取を行い，録音状態が良好な音声として564名分を選別した．

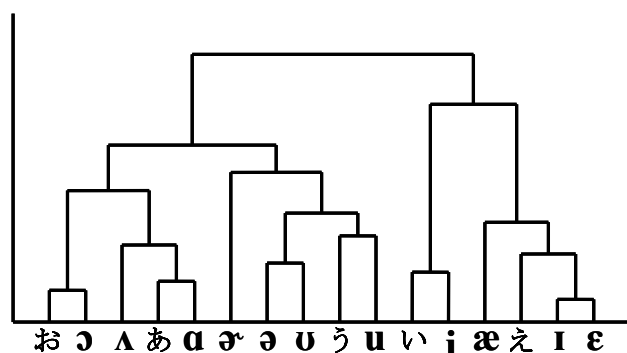


図 6.1: 発音構造の樹形図による視覚化

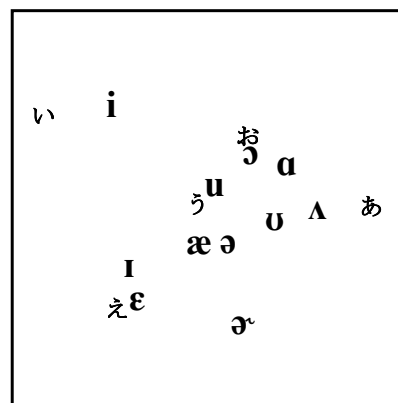


図 6.2: 発音構造の MDS による視覚化

表 6.1: 音響分析条件 (6.4.1 節)

サンプリング	16bit/16kHz
窓長, シフト長	25 msec, 1 msec
特徴量	Mel Cepstrum (1-10 次元)
推定法と分布	MAP 推定によるガウス分布

6.4 二次元表記による情報歪みの評価

6.4.1 英語母音発音構造の抽出

得られた音声資料から母音区間を HTK[32] を用いた自動切り出しにより求め、各母音区間を MAP 推定により分布化する。用いた音響分析条件を表 6.1 に示す。日米 16 母音の母音間距離 (バタチャリヤ距離) をすべて求め、 16×16 の母音間距離行列で表現される英語発音構造 (16 角形) を得る。

6.4.2 英語母音発音構造の視覚化

564 名の発音構造 (母音間距離行列) に MDS を適用し、564 枚の二次元 MDS 母音図を得る。MDS は計量的 MDS と非計量的 MDS に分けられるが、本稿では距離情報の再現を目指す計量的 MDS を用いる。計量的 MDS にはいくつか手法が存在するが、6.2.2 節で紹介した Torgerson の方法を採用する。これは、Torgerson の方法以外の計量的 MDS は、配置を求める際の最適化基準として距離行列を直接用いないからである (このため、Torgerson の方法以外の計量的 MDS を「準計量的 MDS」とする分類がある [31])。

表 6.2: 母音群別の視覚化歪み

	母音群		表示 母音数	平均 相関値	相関値		
	日本語母音	英語母音			0.85 以上	0.90 以上	0.95 以上
全部	5 母音全部	11 母音全部	16	0.907	91.5 %	64.5 %	9.06 %
/あ/周辺 母音群	/あ/,/お/	/ɑ/,/ʌ/,/æ/,/ə/,/ɚ/,/ɔ/	8	0.933	94.3 %	78.5 %	45.6 %
	/あ/,/お/	/ɑ/,/ʌ/,/æ/,/ə/,/ɔ/	7	0.948	97.2 %	88.3 %	58.1 %
	/あ/	/ɑ/,/ʌ/,/æ/,/ə/,/ɚ/	6	0.948	95.7 %	87.4 %	61.8 %
/い/周辺 母音群	/い/,/え/	/æ/,/ə/,/ɪ/,/i/,/ɛ/	7	0.975	99.8 %	99.6 %	87.6 %
	/い/,/え/	/æ/,/ɪ/,/i/,/ɛ/	6	0.981	100.0 %	98.9 %	94.3 %
/う/周辺 母音群	/う/,/お/	/ɑ/,/ə/,/ʊ/,/u/,/ɔ/	7	0.949	97.3 %	89.9 %	60.9 %
	/う/,/お/	/ɑ/,/ʊ/,/u/,/ɔ/	6	0.963	98.4 %	94.1 %	77.4 %

6.4.3 英語発音構造の視覚化に伴う歪みの評価

MDS によって作られた二次元 MDS 母音図は母音間距離行列を完全には再現しておらず、低次元化に伴う歪みを内包している。本稿では、この視覚化による歪みを相関係数を用いて定量的に計測した。n 次元空間における母音間距離と二次元 MDS 母音図上の母音間距離の相関を、564 枚の二次元 MDS 母音図の各々について調べた。16 種類の母音があるため、二次元 MDS 母音図一枚につき、120 個の距離が存在する。これらの相関係数を母音図に対する相関値とした場合、平均相関値は 0.907 となった。相関値 0.90 以上の二次元 MDS 母音図は全体の 64.5 %、0.95 以上は 9.06 % であった。

6.5 母音グループの分割による情報歪みの是正

一般的に、一枚の二次元 MDS 母音図に表示する母音の数が少なければ、視覚化による歪みも小さくなる。母音数の減少に伴う歪みの減少を示したものが図 6.3 である。この図は 6.4.1 節で得られた構造からすべての多角形部分構造を取り出し (n 角形部分構造の場合、 $564 \times {}_{16}C_n$ 個を取り出すことになる)、6.4.3 節と同様の手法で計算した歪み (平均相関係数) の変遷を表したものである。

そこで、16 母音を分割し、一枚の二次元 MDS 母音図に表示する母音数を減らすことで、歪みを低減することを考える。実際に 16 母音全てを同時に提示しても、どの部分に焦点を当てて分析結果を見るべきなのかが不明瞭であり、教育的な観点からも示すべき内容を絞った方が相応しい。しかし、英語教育的に意味のある母音グループにおいて、二次元視覚化に伴う歪みが十分小さいことが必要条件となる。英語音声学の専門家である牧野武彦准教授 (中央大学経済学部) により、教育的に意味のある母音グループの候補を幾つか挙げてもらい、個々のグループに対して歪みの大きさを算出した。結果を表 6.2 にまとめる。16 母音全てを視覚化した場合に比べ、平均相関係数は高くなり、0.90 以上の相関係数を示す二次元 MDS 母音図も多くなった。実際に分割した母音グループによる二次元 MDS 母音図の一例が図 6.4 である。

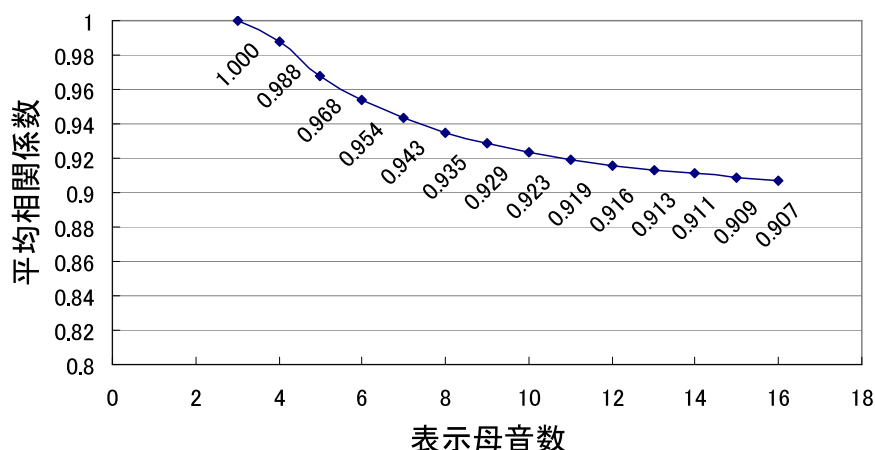


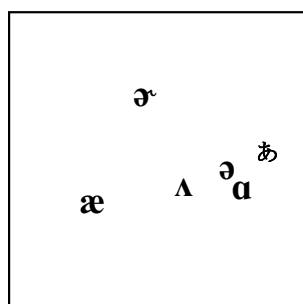
図 6.3: 表示母音数による視覚化歪みの変遷

6.6 発音分類への MDS の適用

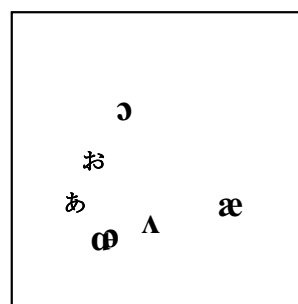
5.2 節で紹介した発音分類に，MDS を適用することが可能である．発音分類は学習者間距離行列にボトムアップクラスタリングを適用することによって実現されており，この学習者間距離行列に MDS を適用することにより，学習者間の関係を二次元平面上に投影することができる．発音分類に用いる学習者間距離行列に MDS を適用した一例が図 6.5 である．

6.7 本章のまとめ

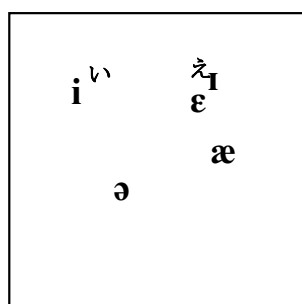
本章では，音声の構造的表象に基づく英語母音発音構造に対して，計量的 MDS を用いることで，母音間の関係性が理解しやすい二次元 MDS 母音図の作成を試みた．しかし，この場合，二次元化することにより高次元情報である母音間距離情報は歪む．そこで，相関分析により二次元化に伴う歪みを定量的に測定した．そして，母音群を分割し，一度に表示する母音数を減らすことにより，歪みを効果的に減少できることを確認した．



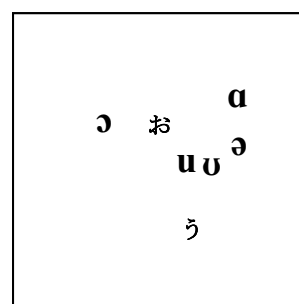
(a) /あ/周辺の二次元 MDS 母音図-1



(b) /あ/周辺の二次元 MDS 母音図-2



(c) /い/周辺の二次元 MDS 母音図



(d) /う/周辺の二次元 MDS 母音図

図 6.4: 分割した母音グループによる二次元 MDS 母音図

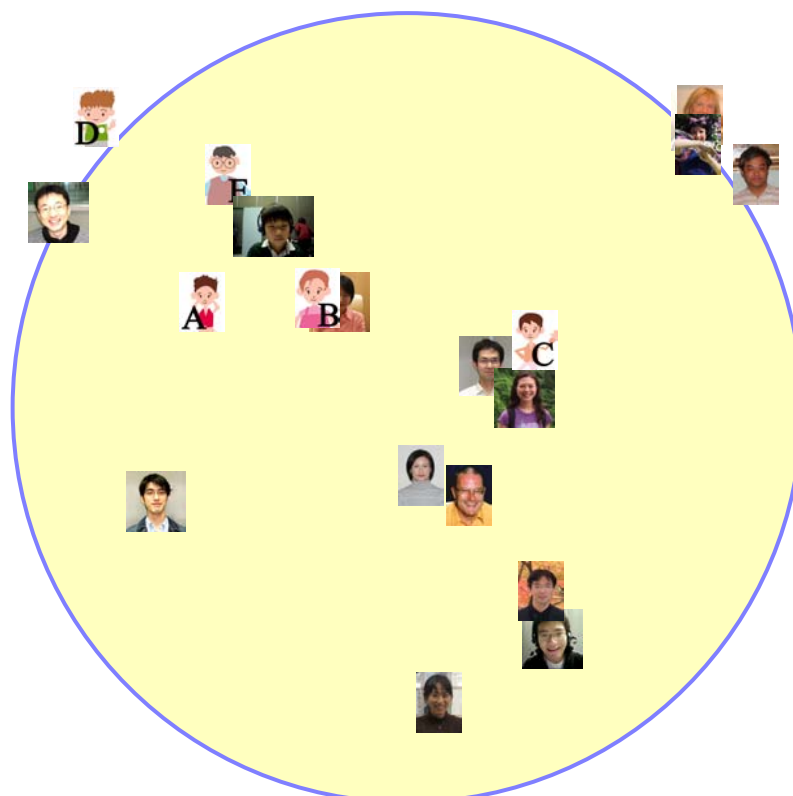


図 6.5: MDS による発音分類の二次元視覚化

第7章

音声の構造的表象に基づく 英語発音分類の妥当性の検証

7.1 はじめに

本章では、5.2節で紹介した英語発音分類の妥当性について検討する。5.2節は技術的観点からの検討であり、教育的観点からの検討については不十分との指摘があった。発音評価システムの教育現場への導入を考えた場合、現場に立つ教師達の理解を得る必要があるため、教育的観点からの検討は不可欠である。そこで、本章では、5.2節で使用した音声資料に対して、英語音声学を専門とする音声学による母音図作成を行う。そして得られた母音図と、提案手法に基づく構造表象との比較を通して、提案手法の教育的妥当性について検討する。

7.2 母音図

本章では、音声学者の感覚を定量化するため、母音図を用いる。調音音声学では、発音時の調音動作を母音図にプロットすることで記述する。母音図の詳細については、2.2.1節で述べたので、ここでは割愛する。

7.3 日本人英語学習者の模擬音声の収録

使用した模擬音声は、5.2.1節と同様の96種類の11発音である。この模擬音声に対して、構造的表象に基づく発音分類と母音図に基づく発音分類を行い、両分類を比較する。

具体的には、96種類の学習者発音を全て母音図化する。次に、母音図に基づく母音間距離行列（11角形）、及び、母音構造間距離行列（96角形）を求め、図5.3と同様の分類を行って、両者の類似性について検討する。これは、図5.3の分類は、8クラスタ時には妥当な発音分類が行われているが、8クラスタに至る分類のプロセスの妥当性の検証が、図5.3のみでは困難だからである。

7.4 英語音声学による英語音声の母音図化

母音図化を担当した英語音声学者は、第6章と同様、牧野武彦准教授である。母音図化は、次の4つの基準で行った。1) 舌の前後、2) 舌の高低（口の開き具合）、3) 円唇性、4) r音性、である。通常の（2次元平面としての）母音図は舌の前後/高低の位置を意図されているが、ここでは、円唇性、r音性についても評価すべきであるとの音声学者の指摘により、これらも含めて母音図化した。舌の位置（前後/高低の位置）については母音図上に位置を指定することで評価した。円唇性については、無円唇、弱い円唇、強い円唇の3段階で評価した。また、r音性はr音性有り（+）、無し（-）の2段階で評価した。通常、米語でr音性があるのは、/ɹ/のみである。以上、各母音を4組の値で表現した。

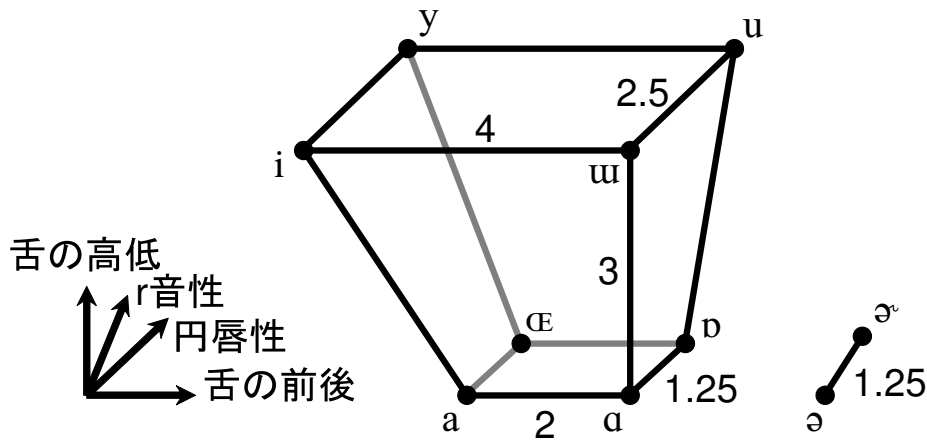


図 7.1: 母音間距離を求める際の次元の重み

7.5 母音図に基づく英語発音分類

7.5.1 母音図に基づく母音間距離の算出

母音間距離として、母音を表す4組の値のユークリッド距離を用いた。本来ならば、各母音間の知覚的距離などを参考すべきであるが、ここでは、母音図という従来より調音音声学において広く使われる母音群の構造的表示形式(4:3:2の台形/2次元表示)を重要視し、舌の前後/高低位置については台形表示におけるユークリッド距離を用いることとした。これに対して、円唇性、r音性を付与することで四次元空間の二点間距離として母音間距離は定義できるが、各次元の尺度が異なるため、追加する2次元には適当な重み付けが必要である。円唇性、r音性の重みに対しては調音音声学における構造的表示形式からの要請に基づき重みを議論すべきだが、これについては参考となる議論が見つからなかったため、各母音間のケプストラム距離を参考にして、円唇性、r音性の距離における重みを決定した。

最終的に4:3:2の台形/2次元表示に対して、円唇性のための次元は無円唇・強円唇の距離を2.5に、r音性の次元は、r音性の有無の距離を1.25として設定した(図7.1)。

7.5.2 母音図に基づく学習者間距離行列からの樹形図構築

7.5.1節の母音間距離の算出により、母音図に基づく 11×11 の母音間距離行列が得られ、各学習者の発音に対する構造を定義することができる。また、すべての学習者間で構造間距離を算出することにより、 96×96 の母音構造間距離行列(学習者間距離行列)を作ることが可能である。構造間距離の算出には式(5.1)を利用し、母音構造間距離行列(学習者間距離行列)を樹形図化した(図7.3)

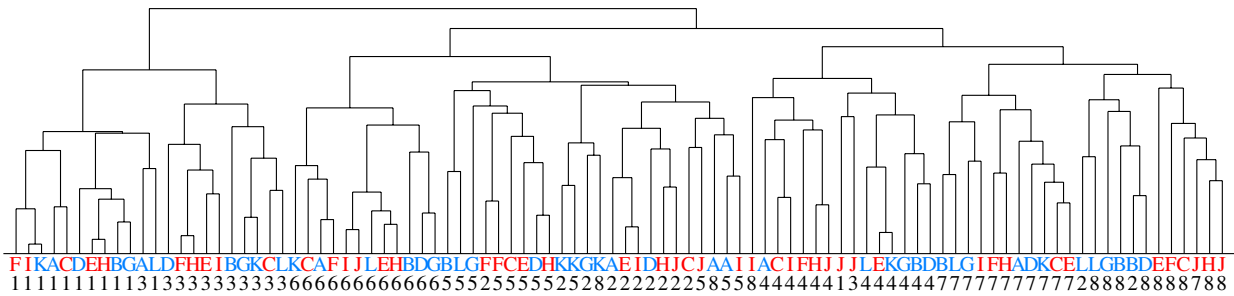


図 7.2: 構造的表象に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類 [再掲]

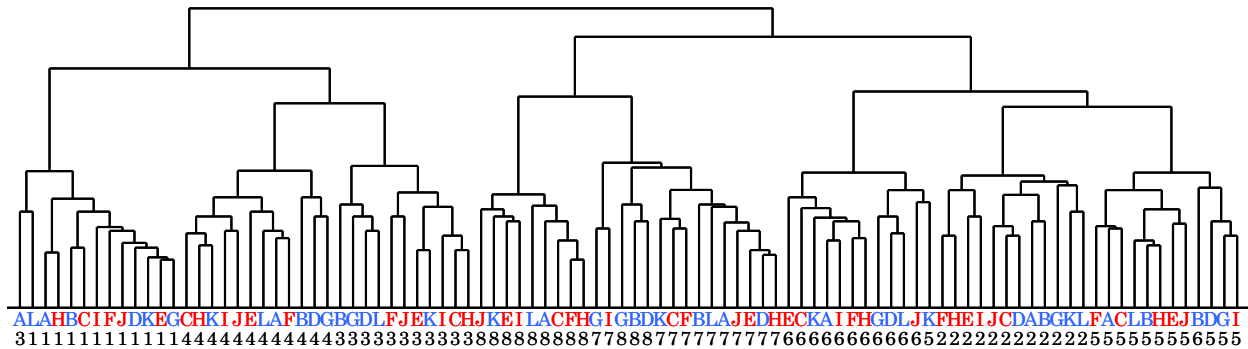


図 7.3: 母音図に基づく 96 発音構造 ([12 話者 A ~ L] × [8 発音状態 1 ~ 8]) の分類

7.6 2つの英語発音分類の比較

母音図に基づいて構築された樹形図を見ると、おおよそ発音状態ごとに固まった 8 つのクラスタに分かれている。この特徴は構造的表象に基づいて構築された樹形図にも共通している。また、クラスタの結合については、両樹形図において 8 つのクラスタの結合順序は一部異なっているが極めて似た結合となっている。構造的表象に基づく分類については、{1,3}, {2,5,6}, {4,7,8} の 3 つの大クラスタが作られ、その後、{2,5,6} と {4,7,8} が結合する。対して、母音図に基づく分類については、{1,3,4}, {2,5,6}, {7,8} の 3 つの大クラスタが作られ、その後、{2,5,6} と {7,8} が結合する。比較すると、状態 4 を除いて同じ結合順序となっている。

さらに、樹形図化において使用するデータである母音構造間距離行列 (96×96) で比較を行った。96 種類の学習者発音に対して、 ${}_{96}C_2$ 個の母音構造間距離が得られる。構造的表象に基づく母音構造間距離と、母音図に基づく母音構造間距離を算出し、プロットを行った。結果を図 7.4 に示す。横軸は母音図に基づく母音構造間距離を、縦軸は構造的表象に基づく母音構造間距離を表す。相関係数は 0.628 となり、相関が確認できた。しかし、樹形図の構造を一致させるほどの強い相関ではなく、クラスタの結合順序が一部異なるのはこの歪みが原因であると考えられる。

また、母音構造間距離行列の算出には近似式が使われているので、構造間距離行列の算出に使われた母音間距離そのものを用いた相関分析を行った。96 種類の学習者発音に対し

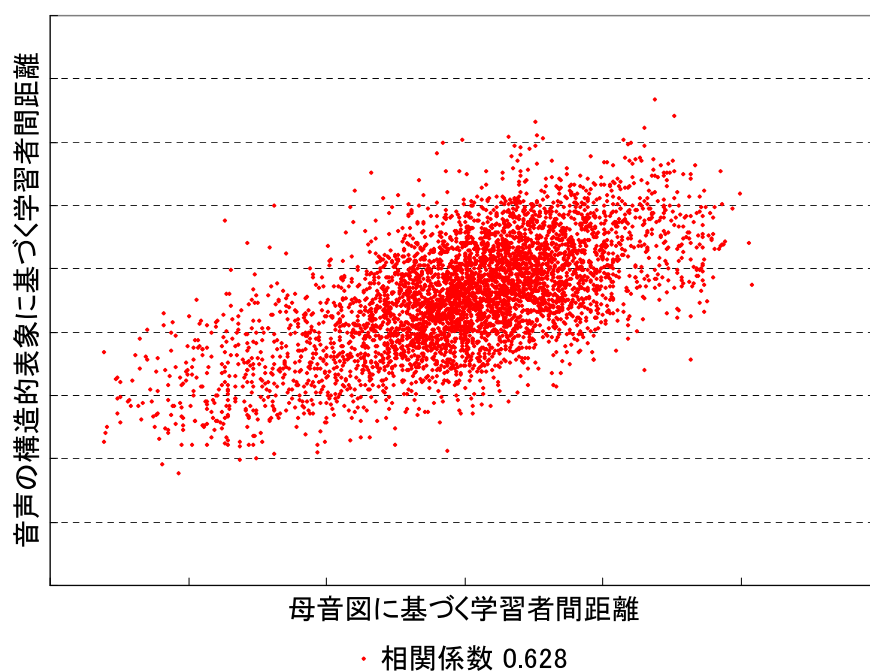


図 7.4: 学習者間距離の相関

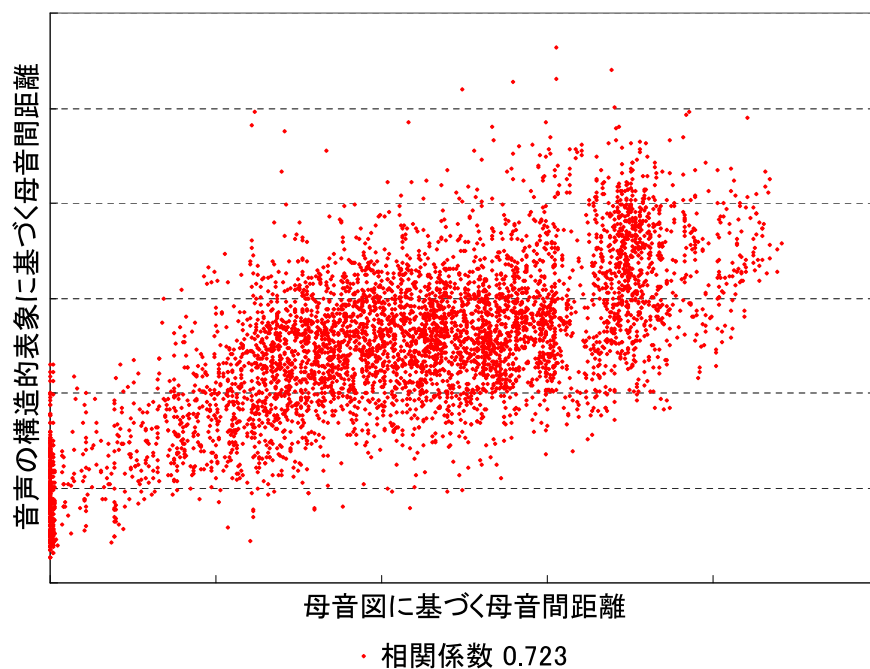


図 7.5: 母音間距離の相関

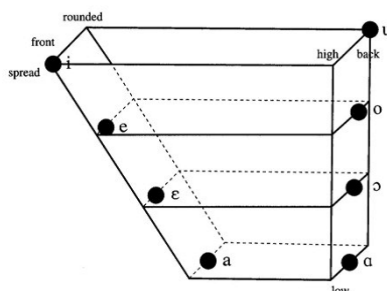


図 7.6: 知覚的に等距離となる基本 8 母音

て、それぞれ ${}_{11}C_2$ 個の母音間距離が得られる．よって、2 組の $96 \times {}_{11}C_2$ 個のデータの相関を見る．音響的な母音間距離（母音ケプストラム分布間のバタチャリヤ距離）と、母音図に基づく母音間距離を算出し、プロットを行った．結果を図 7.5 に示す．横軸は母音図に基づく母音間距離を、縦軸は音響的な母音間距離を表す．相関係数は 0.723 になり、より強い相関が確認できた．調音音声学の分野では、今回用いた母音表示形式（4:3:2 の台形 / 2 次元表示）のほかに、基本 8 母音（図 7.6）が知覚的に等距離である [33] という考えもあり、この場合、母音間の距離尺度は変わってくる．母音図空間が一意に定義されていないという事実が母音図と構造的表象の相関を弱める一因であると考えられることもできる．

7.7 本章のまとめ

本章では、5.2 節で紹介した先行研究について、調音音声学の観点から妥当性を検証した．調音音声学で用いられる母音図を空間と見立て、母音間距離を定義し、様々な分析を試みた．

96 種類の母音構造を樹形図にして比較した結果、クラスタの結合順序は一部異なったものの極めて似た樹形図となり、母音構造間距離や母音間距離では強い相関が確認された．

これにより、音声の構造的表象と母音図表象には相関があることが実験的に示され、実際の調音行動と音声の構造的表象に関連があることが間接的に示された．

第8章

音声の構造的表象に基づく 母音矯正度推定の高精度化

8.1 はじめに

本章では、5.3節で紹介した母音矯正度推定について検討する。母音矯正度推定の一般的な手法として考えられるのは、学習者の母音と教師の母音をなんらかの絶対的な座標軸で評価し、この評価値に基づいて学習者の母音と教師の母音の差を矯正度とするという手法である。しかし、5.3節は、母音事象間の距離のみを用いて母音矯正必要度の推定を行っており、母音自体に与えられた絶対的な評価を用いない特殊な手法である。特殊な手法であるため、一般的な手法と比較を行い、妥当性の検証を行う必要がある。

そこで、絶対的な座標軸を持ち、かつ非言語的特徴を有さない母音図に着目し、母音矯正度の推定方法の妥当性を検証する。5.3節にて使用された音声を母音図化し、これを用いて、絶対的な座標軸を基準として母音矯正度を得る。さらに、5.3節で提案されている推定手法を母音図空間で適応し、母音事象間の距離に基づいた母音矯正度を得る。そして、この二つの母音矯正度の相関分析を行うことにより、妥当性を検証する。また、この提案されている推定手法を見直し、より高い妥当性が得られる推定手法を検討する。

8.2 英語学習者の模擬音声の収録

使用した模擬音声に関しては、5.3.1節と同様の96種類の11発音である。本節では、5.3節と同様、表5.4(母音置換表)のP1~P7は日本人米語学習者の発音状態に相当し、P8は米語教師の発音状態に相当することに相当することを念頭に置いて検討を行う。

8.3 英語音声学者による英語音声の母音図化

7.4節同様に音声資料を母音図化する。96枚の四次元母音図を得る。

8.4 母音矯正度推定の妥当性の検証

前述の作業によって得られた96枚の母音図を用いて母音矯正度を推定する。母音矯正度の推定は、以下の2通りの方法で行う。

8.4.1 発音の絶対的基準による母音矯正度

ある母音図を教師の発音、異なる母音図を学習者の発音とする。この二つの母音図を用いて母音矯正度を考える場合、発音を表す座標値を用いて、教師の発音と学習者の発音を直接比較し、その距離をもって母音矯正度とする方法が考えられる。つまり、二つの母音図を母音図の軸(棒)を重ねて比較し、母音矯正度を推定する方法である。本稿では、この方法で推定された母音矯正度を、母音図の軸(棒)という絶対的な基準に基づいてることから、「絶対的母音矯正度」と称する。

8.4.2 発音の相対的基準による母音矯正度

次に、構造的表象で用いた式(5.3)により、母音矯正度を推定することについて考える。これは母音図の軸(枠)が消失し、母音だけがプロットされた状態で、二つの発音群を比較し母音矯正度を推定する問題となる。この場合、図5.2のように構造を回転&シフトさせ、二つの構造を重ね合わせた後に、母音矯正度を推定することとなる。構造的表象と同様にすべての発音間距離を求めることにより、母音図に基づく発音間距離行列を得る。そして、この距離行列に対して式(5.3)を適用し、母音矯正度を推定する。本稿では、この方法で推定された母音矯正度を、発音間の相対的な関係に基づいてることから、「相対的母音矯正度」と称する。

8.4.3 絶対的母音矯正度と相対的母音矯正度の相関分析

前述の絶対的母音矯正度と相対的母音矯正度に対して相関分析を行うことで、構造的表象における母音矯正度の推定式である式(5.3)の妥当性について検討する。妥当性の評価には相関係数を用いる。

まず、ある話者の P8 (米語教師相当の発音) と、同一話者の P1 (日本人米語学習者の発音) を比較し、母音矯正度を 2 通りの方法 (絶対的母音矯正度と相対的母音矯正度) で求める。この 12 組 × 11 母音に対して絶対的母音矯正度と相対的母音矯正度の相関分析を行う。同様に、P8 と同一話者の P2, P3, ..., P7 を比較し、相関分析を行う。最後に、教師も学習者も限定せず、96 人すべてを教師・学習者として相関分析を行う (${}_{96}C_2 \times 11$ 母音)。

8.4.4 考察

各比較の相関係数についてまとめたものが図8.1である。横軸は相関係数である。また、“96x96”は96人すべてを教師・学習者とした場合を表す。

96人総当たりの場合の絶対的母音矯正度と相対的母音矯正度の相関係数は0.772であり、強い相関が確認された。この事実から、式(5.3)は一定の信頼性があると考えられる。

P8とP1, P2, ..., P7を比較した場合の相関係数について見てみると、P2~P7は相関係数0.750~0.885と高い水準を示しているのに対し、P1のみが相関係数0.677と他の学習者群よりも弱い相関であることがわかる。これにはP1がP2~P7と比べてP8から乖離した発音構造であることが起因していると考えられる。P1はすべての母音を日本語発音に置換しており、置換量が半分であるP2~P7よりも歪んだ発音構造となっている。

8.5 母音矯正度推定の高精度化

この節では、母音矯正度を推定する式(5.3)を改善する方策について検討する。まず、式(5.3)の問題点について述べ、これを改善する推定式を示す。そして、改善式によって推定された母音矯正度に対して、前節と同様の相関分析を行い、改善の効果を確認する。

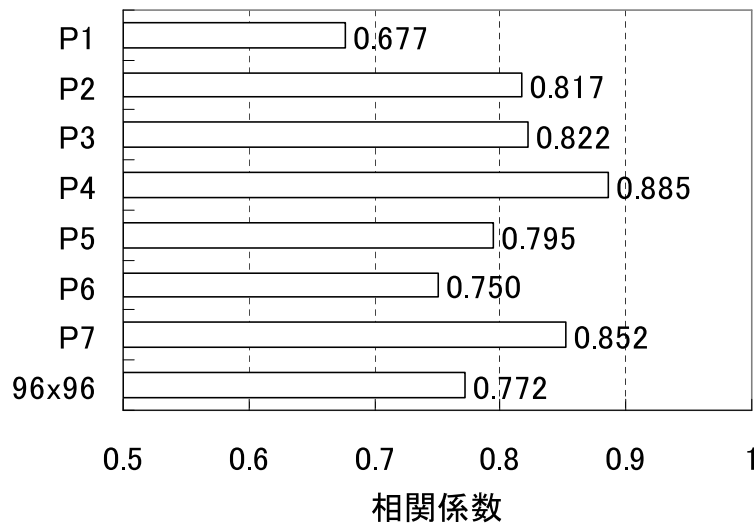


図 8.1: 絶対的母音矯正度と相対的母音矯正度の相関係数

8.5.1 構造歪みに基づく重み付け

構造歪みを測る式 (5.3) は, 構造のすべての要素を等価に扱っている. これは式 (5.3) を行列の演算に書き換えた式 (8.1) を見れば明らかである.

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{pmatrix} = \begin{pmatrix} 0 & \delta_{12} & \dots & \delta_{1m} \\ \delta_{21} & 0 & & \delta_{2m} \\ \vdots & & \ddots & \vdots \\ \delta_{m1} & \delta_{m2} & \dots & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad (8.1)$$

ここで, d_v は構造 S と構造 T を比較したときの要素 v の構造歪みを表す. また, δ_{ij} は, 要素 i と要素 j の要素間距離の差 ($\delta_{ij} = |s_{ij} - t_{ij}|$) である.

式 (8.1) において, すべての要素が 1 からなる列が存在するが, この列は各要素の重みと考えられる. すべての重みが一致している式 (8.1) ではすべての要素を等価に扱っていると捉えることができる. 構造歪み推定対象 v を除く要素がおよそ正しい場合, この扱いは妥当であるが, 正しくない要素が多く存在する場合 (最も極端な例が P1 に相当する) は妥当性が低下することになる. そこで, 正しいと思われる要素に重きを置き, 正しくないと思われる要素の影響力を下げる推定手法を試行する.

ここでは, 式 (8.1) によって推定された構造歪み d_v を要素の正しさを測るための値として利用し, 構造の各要素の重み w_v を求める. 重み w_v を求める式を次式のように定める.

$$w_v = c \times \frac{1}{d_v} \quad (8.2)$$

c は重み列正規化係数 ($\sum_{i=1}^m w_i = m$) である. 式 (8.2) によって求められた重みを用い

て、構造歪み d'_v を再推定するには次式を用いる。

$$\begin{pmatrix} d'_1 \\ d'_2 \\ \vdots \\ d'_m \end{pmatrix} = \begin{pmatrix} 0 & \delta_{12} & \cdots & \delta_{1m} \\ \delta_{21} & 0 & & \delta_{2m} \\ \vdots & & \ddots & \vdots \\ \delta_{m1} & \delta_{m2} & \cdots & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} \quad (8.3)$$

さらに、 d_v の代わりに d'_v を用いて、重み w'_v を得ることにより、 d'''_v を定義することも可能である。式 (8.1) で推定された要素 v の構造歪み d_v を相対的母音矯正度としたことに倣って、式 (8.3) で再推定された構造歪み d'_v を「重み付き相対的母音矯正度」と本稿では称する。

8.5.2 絶対的母音矯正度と重み付き相対的母音矯正度の相関分析

8.4.3 節と同様に、絶対的母音矯正度との相関分析を行うことで、式 (8.3) の効果を測る。

まず、ある話者の P8 (米語教師相当の発音) と、同一話者の P1 (日本人米語学習者の発音) を比較し、重み付き相対的母音矯正度を求める。この 12 組 × 11 母音に対して絶対的母音矯正度と重み付き相対的母音矯正度の相関分析を行う。同様に、P8 と同一話者の P2, P3, ..., P7 を比較し、相関分析を行う。最後に、教師も学習者も限定せず、96 人すべてを教師・学習者として相関分析を行う (${}_{96}C_2 \times 11$ 母音)。

求められた相関係数をまとめたものが図 8.2 である。

8.5.3 考察

96 人総当たりのときの絶対的母音矯正度と重み付き相対的母音矯正度の相関係数は 0.807 となり、相対的母音矯正度の相関係数 0.772 と比べ、3.5 ポイント改善した。また、P8 と P1, P2, ..., P7 を比較して相関分析をした場合でも、すべての場合で改善を示した。これにより式 (8.3) の有効性が示された。

次に、重みを決定する式 (8.2) の相対的母音矯正度 d_v の代わりに重み付き相対的母音矯正度 d'_v を適用することにより、さらに相関係数を上昇させることを考える。この重み付き相対的母音矯正度の再々推定を行った結果が図 8.3 である。ここでは、再々推定で得られた d'_v を利用して d'''_v を推定する、というように再帰的に 10 回推定した。結果はさらなる改善を示したが、やがて相関係数は収束した。10 回推定したときの相関係数についてまとめたものが、図 8.2 である。96 人総当たりのときの絶対的母音矯正度と重み付き相対的母音矯正度 (推定 10 回目) の相関係数は 0.827 となり、相対的母音矯正度と比較して 5.5 ポイント、重み付き相対的母音矯正度 (推定 2 回目) と比較しても 2.0 ポイントの改善を示した。また、P8 と P1, P2, ..., P7 を比較して相関分析をした場合でも、すべての場合で改善を示した。

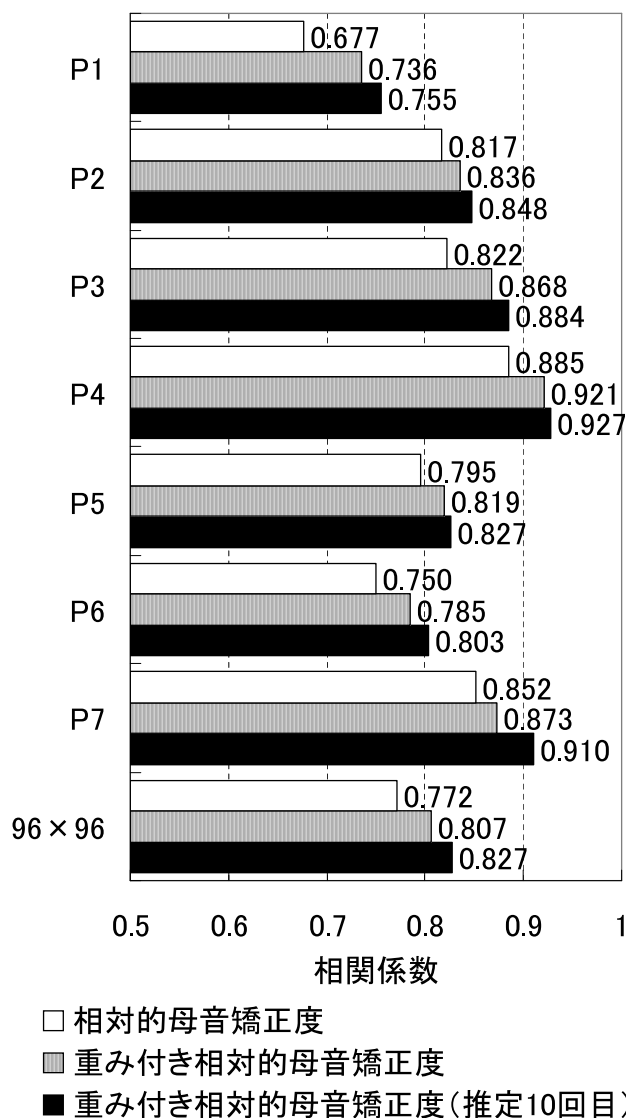


図 8.2: 絶対的母音矯正度と重み付き相対的母音矯正度の相関係数

8.6 構造的表象に基づく母音矯正度の再推定

前節では母音図表象に母音矯正度の再推定を適用した．本節では構造的表象に対して適用する．そして，適用した際の重みを調べることにより，再推定の効果について考察する．

8.6.1 母音矯正度の再推定の構造的表象への適用

5.2.2 節と同様に収録した音声資料から音声の構造的表象に基づく構造を抽出し，前節までに検討した手法で母音矯正度を推定する．推定時の重みについてまとめたものが図 8.4 である．図中の P1～P8 は，12 話者平均構造である．各図は米語教師相当の構造 P8 と日本人学習者状態の P1～P7 を比較したときの母音矯正度推定時（10 回目）の重みを表して

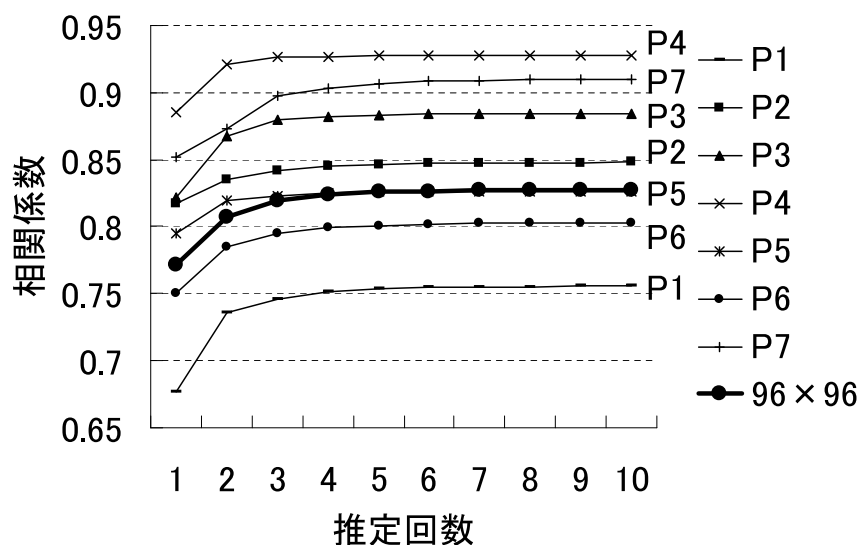


図 8.3: 推定に伴う相関係数の推移

いる．図の縦軸は重みを表し，色づけられているものは日本語母音に置換した母音，白抜きのは米語発音の母音である．

8.6.2 考察

母音矯正度の推定時に用いられる重みは，各母音の信頼度と捉えることができる．図 8.4 を見ると，米語発音母音（白抜き）の多くが重く評価されている．米語発音の母音は信頼できる母音であるから，これは妥当な結果と言える．

また，日本語母音に置換されている母音では， $/i/$ ， $/ɔ/$ が比較的大きい重みを保有している．これらの母音は，日本語の $/い/$ ， $/お/$ に近い母音とされており [33]，この事実と一致する結果となった．しかし， $/a/$ ， $/v/$ ， $/u/$ といった母音も大きめの重みとなっており，重みの算出方法について検討の余地がある．

8.7 平方二乗話による母音矯正度推定手法の検討

8.7.1 平方二乗話による母音矯正度推定

各母音に起因する構造歪みの大きさは式 (5.3) でこれまで検討されてきたが，次のような式も考えられる．

$$d_v = \sqrt{\sum_{i=1}^m (s_{vi} - t_{vi})^2} \quad (8.4)$$

この式 (8.4) は

$$D(S, T) = \sqrt{\frac{1}{m} \sum_{i>j} (s_{vi} - t_{vi})^2} = \sqrt{\frac{1}{2m} \sum_{i=1}^m d_i^2} \quad (8.5)$$

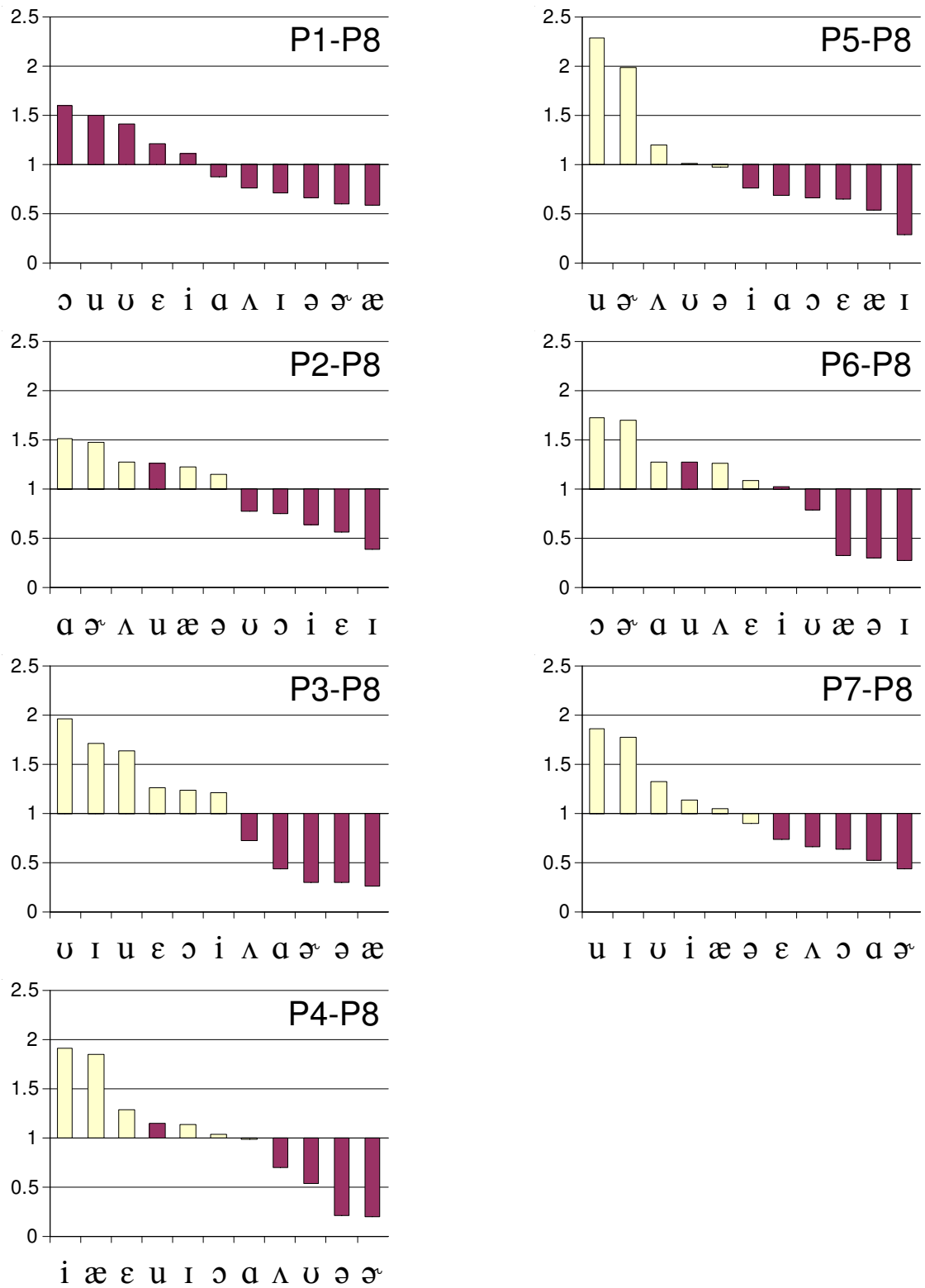


図 8.4: 母音矯正度推定時 (10 回目) の重み

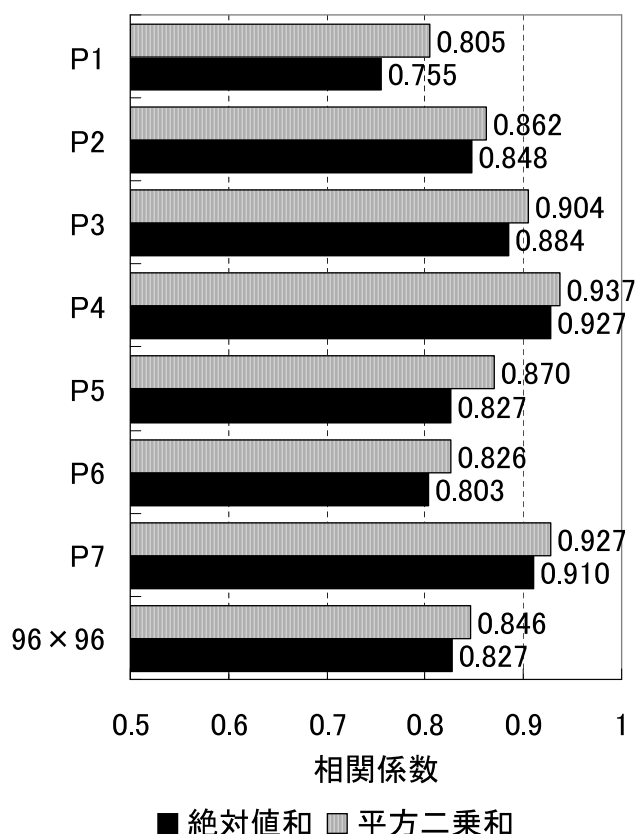


図 8.5: 絶対的母音矯正度と重み付き相対的母音矯正度（絶対値和 / 平方二乗和）の相関係数

が成り立ち，構造間距離と各要素の構造歪みの関係が明確となる．

8.5 節と同様に重みを

$$w_v = c \times \frac{1}{d_v} \tag{8.6}$$

とし（ $\sum_{i=1}^m w_i = m$ ），重みを付けた推定手法を

$$d'_v = \sqrt{\sum_{i=1}^m (w_i (s_{vi} - t_{vi}))^2} \tag{8.7}$$

とする．また，8.5.3 節と同様に再推定を繰り返し行う．

推定手法の評価は 8.5.2 節と同様の手法で行う．結果は図 8.5, 図 8.6 である．図 8.5 は推定 10 回目のときの各状態の相関係数をまとめたものであり，図 8.6 は 96 人総当たりのときの相関係数の推移を表したものである．各図の「絶対値和」は式 (5.3) に基づく推定結果であり，「平方二乗和」は式 (8.4) に基づく推定結果である．

8.7.2 考察

まず，重み付けをしない場合であるが，平方二乗和を用いた式 (8.4) の相関係数は 0.718 となり，絶対値和を用いた式 (5.3) の 0.772 を下回ったが，重み付けをした場合（推定 10

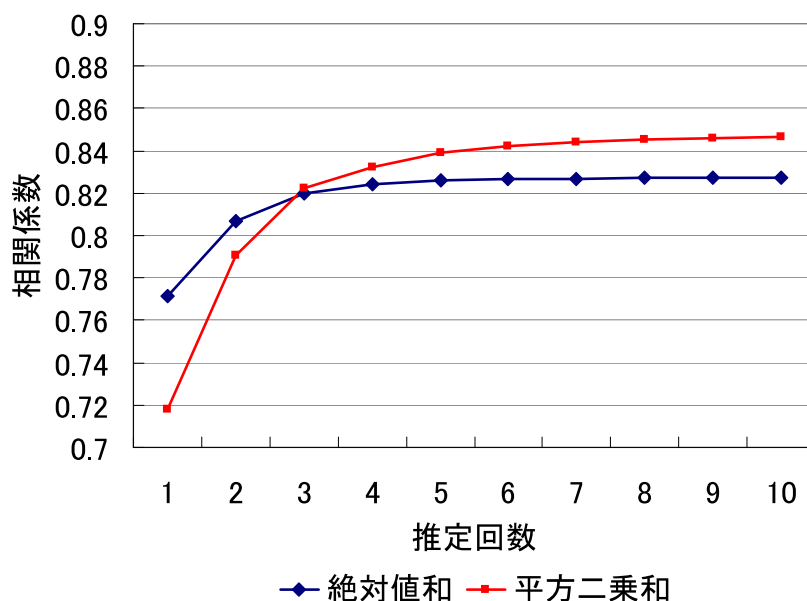


図 8.6: 推定に伴う相関係数の推移 (絶対値和 / 平方二乗和)

回目)の結果に関しては、平方二乗和は 0.846 となり絶対値和の 0.827 を上回った。

以上より、平方二乗和を母音矯正度として用いた場合、重み付けを行わない場合の精度は単純合算に比べて落ちるが、重み付け計算を施すことで単純合算と同じ精度を出すことが可能であると考えられる。今後、母音矯正度推定を検討するに当たり、平方二乗和による推定手法は選択肢の一つになると考えられる。

8.8 本章のまとめ

本章では、5.3 節で紹介した母音矯正度の推定手法について、母音図を通して妥当性を検証した。母音図に基づく母音間距離を用いて、座標軸がある場合に下される母音矯正度 (絶対的母音矯正度) と提案されていた推定手法による母音矯正度 (相対的母音矯正度) を求め、相関分析を行った。その結果、相関係数 0.772 と良好な相関を示した。しかし、比較する二つの構造が大きく異なると、この相関係数は低下することが確認された。

また、母音矯正度の推定手法について改善策を提案した。構造歪みを考慮した重み付けを取り入れた結果、前述の母音矯正度の相関はより強くなった (相関係数 0.807)。また、再推定された構造歪みを用いて構造歪みを再々推定することによって、相関が一層強くなることが確認された (相関係数 0.827)。さらに、この推定手法を構造的表象に適用し、推定時の重みを調べたところ、信頼されるべき母音の重みが高くなる、という妥当な結果を得た。

第9章

音声の構造的表象に基づく 英語発音評価システムの実地試験

9.1 はじめに

英語発音評価システムを構築するに当たり、実際の運用を通しての改善が不可欠である。本章では、音声の構造的表象に基づく英語発音評価システムを運用した実績について言及する。音声の構造的表象に基づく英語発音評価システム（単語発音評価）を、外国語学校や東京大学柏キャンパス一般公開（オープンキャンパス）、巣鴨学園で実際に運用したので記述する。また、実地試験を通して浮かび上がった問題点について言及する。

9.2 実地試験に用いた英語発音評価システム

9.2.1 英語発音評価システムへの入力

i) 目標教師音声の選択

図 1.1 に示すインターフェースにより、利用者は自由に目標教師を設定できる。目標教師を指定することは、従来の技術では難解であるが、「音声の構造的表象」が持つ「音声から話者性の除去が可能」という特長によって可能になる。設定できる目標教師は運用場所に合わせて変更した。巣鴨学園では実際に教鞭を執る英語教師の音声を選択できるようにした。

ii) 音声の収録

本システムでは単語発音（正確には単語に含まれる母音部分の発音）を評価する。収録する単語は表 9.1 であり、分析には各単語の母音区間のみを用いている。母音区間の切り出しは HTK による自動切り出しによって行っている。各単語の発声回数であるが、収録時間（日本語・英語・フランス語を収録）が長くなる外国語学校（9.3 節）では各単語 1 回、分析時間の短さが要求される柏キャンパス一般公開（9.4 節）では各単語 1 回、正確さが重要である巣鴨学園（9.5 節）では各単語 3 回¹である。

9.2.2 英語発音評価システムからの出力（診断結果）

診断結果を構成する要素は以下のとおりである。実際に利用者に渡した診断結果（カルテ）は付録として本論文の最後に付した。

i) 利用者の発音構造の樹形図

5.1 節に基づく学習者の発音状態を示す樹形図である。利用者の各母音発音の遠近を示す。

¹3 回発声の場合、母音 a と母音 b の母音間距離は（各発声を $a_1, a_2, a_3, b_1, b_2, b_3$ とする）、 $\overline{a_1b_1}, \overline{a_1b_2}, \overline{a_1b_3}, \overline{a_2b_1}, \overline{a_2b_2}, \overline{a_2b_3}, \overline{a_3b_1}, \overline{a_3b_2}, \overline{a_3b_3}$ の 9 個の距離の平均を取っている。

表 9.1: 英語発音評価システムに入力する単語発音

英語単語	評価される母音	日本語単語	評価される母音
beat	/i/	バト	/あ/
bit	/ɪ/	ビット	/い/
bet	/ɛ/	ブト	/う/
bat	/æ/	ベト	/え/
but	/ʌ/	ボト	/お/
pot	/ɑ/		
bought	/ɔ/		
put	/ʊ/		
boot	/u/		
bird	/ɜ/		
about	/ə/		

ii) 指定した目標教師音声と比較した場合の母音矯正順序

5.3 節, 第 8 章に基づく, 指定した教師の発音と利用者の発音を比較した場合の母音矯正順序である.

iii) 発音分類 MDS 図

5.2 節に基づく学習者分類に MDS を適用し, 二次元視覚化して表示した (6.6 節). この二次元 MDS 話者マップによって, 利用者と目標教師音声や後述する発音タイプの関係を知ることができる.

iv) 発音タイプ判定

2005 年, 2006 年の東京大学柏キャンパス一般公開にて集めた利用者の 400 名を越える音声 (実際に使用したのは録音状態が良好な約 350 名), 5.2 節に基づいて発音分類し, 5 つの発音タイプを定義した. 各発音タイプに属する過去利用者の発音構造を平均し, この平均発音構造と利用者の発音構造の距離を式 (5.1) によって求め, 最も近い平均構造を利用者の発音タイプとした.

v) 発音が最も近い話者

利用者の発音構造と前述の過去 2 年間の利用者と発音構造の距離を式 (5.1) によって求め, 最も近い過去利用者に付与した番号を表示した. 過去利用者の音声を聴くことができる端末を別途用意した.

表 9.2: 外国語学校での実地試験回数

学校	人数	期間	回数
都立三田高校	29名	1年間	14回
日本大学	18名	1年間	30回
アテネ・フランセ 発音クラス	10名	3ヶ月間	9回
神奈川県立外語短期大学	40名	1年間	21回
	10名	半年間	13回

vi) 利用者と指定教師の発音構造の二次元 MDS 図

第6章に基づいて利用者の発音状態を MDS を用いて二次元視覚化したものである。

9.3 外国語学校での実地試験

竹内京子氏を通して、外国語学校で長期間利用してもらった。発音カルテの要素は発音構造の樹形図と母音矯正順序のみであるが、英語母音だけでなくフランス語母音も評価できるように拡張したシステムである。利用回数は表 9.2 のとおりである。

9.4 東京大学柏キャンパス一般公開での実地試験

2007年10月末に行われた東京大学柏キャンパス一般公開（柏オープンキャンパス）にて英語発音カルテと称してデモを行った。なお、先にも触れたが、この実地試験は2005年から行われているものである。3回目となる2007年は教師選択、発音分類 MDS 図、発音タイプ判定、発音が最も近い話者といった要素を追加し行った。期間中の天気は雨であり、来場者数は好天候であった前年の6割であったが、前年と変わらぬ200名を越える柏キャンパス近隣の住民（利用者の年齢は4歳から84歳までの広範囲に渡る）にシステムを利用してもらった。

9.5 巣鴨学園での実地試験

2008年1月に巣鴨学園にて中学生、高校生約20名に英語発音評価システム（英語発音カルテ）を利用してもらった。システム利用は2回である。1回目と2回目の間には1週間の間隔があり、1回目の際には発音カルテのほかに中央大学の牧野准教授による発音指導マニュアルを添付した。1回目と2回目の比較により、利用者は発音の上達具合を知ることが可能である。このデモから、学習者と指定教師の発音構造の二次元 MDS 母音図を追加し、前述のすべての要素を含むデモとなった。また、巣鴨学園の英語教師の音声収録し、巣鴨学園の英語教師の発音も目標教師音声として選択できるようにした。

9.6 考察

巣鴨学園の生徒と教師を話者分類したMDS図が図9.1(1回目)と図9.2(2回目)である。黄色が生徒であり、水色が教師である。2つの図を比較すると、1回目と2回目の間の練習によって、生徒が教師に近づいている様子が窺える。

また、収録に参加した巣鴨学園の生徒にアンケートを採ったところ、音声の構造的表象を用いることによる利点である「教師が選択可能である」ことは概ね好評を得た。希望する教師もいくつか挙げられた。

しかし、継続利用に関しては好不評が拮抗しており、課題を残すこととなった。診断結果(発音カルテ)を理解しやすくする、診断に要する時間を短くする、自宅からネットワークを通して分析可能にする、といった対策を行い、利用への障壁を少なくすることで、改善可能と思われる。

また、子音の評価といった機能拡張の要望もある。現在の技術検討では単語(孤立発声)の評価が限界であるので、連続発声(フレーズ)を評価する技術や、二重母音、半母音、子音といった音素への拡張を検討する必要がある。

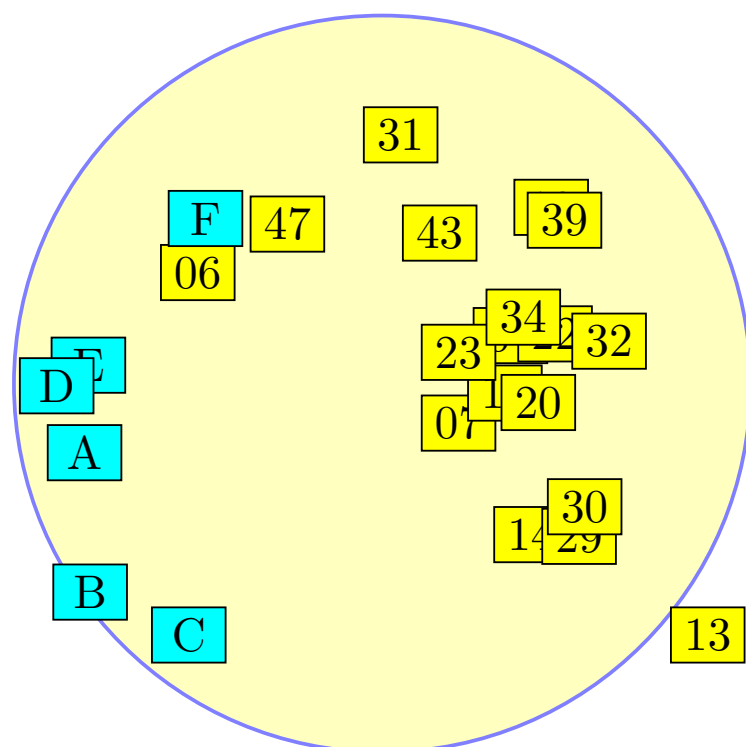


図 9.1: 巣鴨学園関係者の発音分類 (1回目)

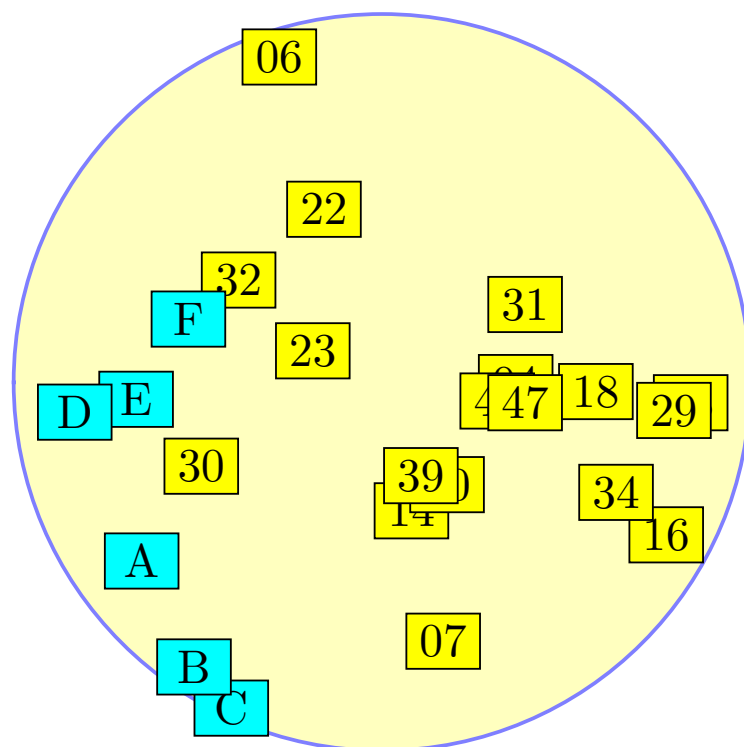


図 9.2: 巣鴨学園関係者の発音分類 (2回目)

第10章

結論

10.1 本研究のまとめ

本研究では、音声の構造的表象に基づく英語音声評価システムについて、実用化に向けた検討を行った。

まず、第1章において、本研究の目的について述べた。

第2章において英語発音評価システムが担う発音教育の重要性について述べ、日本語発音と英語発音の差異について説明した。また、現在実用化されている英語教育システムについて触れ、本研究が目指す英語発音評価システムについて述べた。

第3章では、英語発音評価システムを構築する上で障害となる「不一致問題」について述べた。「不一致問題」を引き起こす場合について実験を行い、従来手法による対応では不十分であることを示した。

第4章では、英語発音評価システムの核を成す音声の構造的表象について説明した。不可避な要因（発話者や収録機器）によってもたらされる音声の歪みについて整理し、この歪みを除去した特徴量「音響的普遍構造」について説明した。

第5章では、本表象に基づく英語発音評価システムの各要素について先行研究を紹介した。5.1節では、英語学習者の音声から話者性（年齢や性別など）を取り除き、純粋に英語の発音状態を評価し、記述可能であることを紹介した。5.2節では、多数の英語学習者を話者性に左右されることなく、発音状態に基づいて分類できることを紹介した。5.3節では、英語教師と英語学習者の発音を直接比較し、母音矯正必要度を推定する手法について紹介した。

第6章では、5.1節で紹介した英語発音状態の記述について、視覚化（二次元化）に伴う情報の歪みについて考慮しつつ表記する方法を検討した。まず、視覚化手法として多次元尺度構成法（MDS）を採用し、視覚化に伴う情報の歪みについて定量的に評価した。そして、表示する母音群を分割し、絞り込むことによって、情報の歪みを抑えつつ、表記の教育的意義の向上が図れる可能性を示した。

第7章では、5.2節で紹介した英語発音分類について、教育学的な見地から英語発音分類の妥当性を検証した。英語音声学による音声資料の母音図化を行い、母音図を通して英語発音を分類した。そして、発音分類樹形図、発音構造間距離（学習者間距離）、母音間距離の三つの観点から音声の構造的表象に基づく英語発音分類の妥当性を示した。

第8章では、5.3節で紹介した母音矯正度推定について、推定を高精度化する手法について提案した。第7章で使用した母音図を用い、母音図の枠を基準とした母音矯正度（絶対的評価）と、他の母音事象との関係を基準とした母音矯正度（相対的評価；音声の構造的表象に基づく母音矯正度推定はこちらの手法に属する）の相関を調べ、音声の構造的表象に基づく母音矯正度推定の妥当性を示した。さらに、母音矯正度を各母音の信頼度として捉え、この信頼度に則する各母音への重み付けをして母音矯正度を再推定した結果、妥当性の向上が確認された。

第9章では、第4章から第8章で示された「音声の構造的表象」に基づく要素技術を用いて英語発音評価システムを構築した。そして、このシステムを様々な場所で利用した。

10.2 今後の課題

10.2.1 連続発声への対応と評価音素数の拡大

第 9 章で述べたように、現在の音声の構造的表象に基づく英語発音評価システムは、単母音、孤立発声（単語発声）に限られている。

より実用的な評価システムを実現するためには、単語発音評価を拡張し、連続発声に対応することは不可欠である。連続発声へ対応することにより、フレーズの評価が可能になり、映画で使用された有名なワンフレーズを練習する、といった評価システムが構築可能になる。現在、音声の構造的表象の連続発声への適用が、音声認識の分野で行われている [23][35]。また、極めて整っていない連続発声であるシャドーイングを、頑健なセグメンテーションによって評価することが試みられている [36][37][38]。これらの技術の発音評価への応用が、連続発音評価への足がかりになるものと思われる。

また、評価音素の拡大に関しては、共鳴音である二重母音や鼻母音の評価には「音声の構造的表象」が利用可能と考えられる。二重母音に関しては、母音区間を前半と後半に分けるなどの工夫を要するが、現在の枠組みで実現可能であると思われるので、評価システムを構築し、実地試験を通しての検討が必要である。なお、フランス語母音（鼻母音含む）に関しては、現在実地試験が進行中である（9.3 節）。

10.2.2 英語発音分類の実体音声への適用

5.2 節や第 7 章はシミュレーション音声での検討であり、実在の英語学習者への発音分類の本格的な検討はまだ行われていない。実在の英語学習者への発音分類をするに当たって、検討すべきは次の 2 点である。

一つは、線形変換性歪み A の限定である。音声の構造的表象に基づく音声認識では、すべての線形変換性歪み A に対して不変性を成立させた結果、全く異なる単語が同一と見なされる可能性が指摘されている [23]。同様に、異なる発音状態が同一のカテゴリに分類される可能性があるため、線形変換性歪み A の限定 [23] は検討事項として挙げられる。

もう一つは、構造サイズの正規化である。発音構造サイズは調音努力や発音の明瞭度と捉えられるが [34]、これを揃えてから発音分類を行う方法である。構造同士を回転（線形変換性歪み A の消去）、シフト（乗算性歪み b の消去）、拡大縮小（構造サイズ差の消去）した後に、発音分類を行うという形になる。前述の線形変換性歪み A の限定と含めて、構造マッチングをどこまで行うのか、ということが検討事項である。

10.2.3 母音矯正度推定の推定条件の検討

5.3 節と第 8 章についても、10.2.2 節と同様に実在の英語学習者音声を用いた検討が必要である。また、母音矯正度推定をする前に行う構造への前処理についても、10.2.2 節と同様の検討を行う必要がある。特に構造サイズの正規化については、構造全体で正規化を行う、部分構造のサイズを一致させる、といった種々の条件について検討すべきである。

謝辞

二年間に渡る本研究を遂行するにあたり，多大なる御指導そして御協力を頂きました，指導教員である峯松信明准教授並びに広瀬啓吉教授に心より感謝いたします．特に，峯松信明准教授の熱意ある的確な指導は，研究を進める上での大きな推進力となりました．深く感謝いたします．また，快適な研究環境の維持に努めてくださった高橋登技官，秘書の武田祥子さん，笠島恵美さんに深く感謝いたします．

中央大学の牧野武彦准教授には，6.5 節にて母音群の分割を，7.4 節にて音声資料の母音図化を行っていただきました．准教授の英語音声学の見地からの意見は，英語発音評価システムの実用化を検討するに当たって不可欠なものでした．深く感謝いたします．

第 5 章の先行研究の研究者である博士課程の朝川智氏には，ご自身の研究で忙しいにも関わらず，研究方針に関する助言や研究用ライブラリの開発など，本研究を全面に渡りサポートしていただきました．ここに深く感謝いたします．

9.3 節にて多くの場で実地試験を行っていただいた博士課程の竹内京子氏を始め，実地試験に協力してくださった多くの方々に深く感謝いたします．特に，実地試験に用いる音声の選定を兼ねた 6.3 節の作業や，9.5 節の実地試験など，実地試験の遂行と開発で大きな尽力をいただいた高澤真章氏に深く感謝いたします．

峯松研究室で研究生活を共にした齋藤大輔氏，下村直也氏を始めとする広瀬・峯松研究室の皆様にも深く感謝いたします．修士課程の二年間が充実した日々となったのも，皆様のおかげであります．重ね重ね御礼申し上げます．

2008 年 1 月 29 日
鎌田 圭

参考文献

- [1] A. Neri, C. Cucchiarini, and H. Strik, “Automatic speech recognition for second language learning: How and why it actually works,” Proc. Int. Congress of Phonetic Sciences, pp.1157–1160 (2003–8)
- [2] 峯松信明, 西村多寿子, 西成活裕, 櫻庭京子, “構造不変の定理とそれに基づく音声ゲシュタルトの導出”, 電子情報通信学会音声研究会,SP2005–12,pp.1–8 (2005–5)
- [3] 峯松信明, “音声に内在する音響的普遍構造とそれに基づく語学学習者モデリング”, 電子情報通信学会音声研究会,SP2003–179,pp.25–30 (2004–1)
- [4] 峯松信明, “音声の音響的普遍構造の歪みに着眼した外国語発音の自動評定”, 電子情報通信学会音声研究会,SP2003–180,pp.31–36 (2004–1)
- [5] 峯松信明, “音響的普遍構造と言語的普遍構造の整合性に基づく発音明瞭度の評定”, 電子情報通信学会音声研究会,SP2003–181,pp.37–42 (2004–1)
- [6] 朝川智, 峯松信明, 広瀬啓吉, “音響的普遍構造のサイズと単語境界における音響的分離度に着眼した米語音声の音響分析”, 電子情報通信学会音声研究会,SP2004–28,pp.53–58 (2004–6)
- [7] 峯松信明, 志甫淳, 村上隆夫, 丸山和孝, 広瀬啓吉, “音声の構造的表象とその距離尺度”, 電子情報通信学会音声研究会,SP2005–13,pp.9–12 (2005–5)
- [8] 朝川智, 峯松信明, 村上隆夫, 伊勢井敏子・ヤッコラ, 広瀬啓吉, “音声の構造的表象に基づく非母語話者の英語発音分析”, 電子情報通信学会音声研究会,SP2005–24,pp.25–30 (2005–6)
- [9] 朝川智, 峯松信明, 広瀬啓吉, “発音の構造的表象に基づく学習者分類と発音教示生成”, 電子情報通信学会音声研究会, SP2005–156, pp.37–42 (2006–1)
- [10] 中川聖一, “語学学習における音声・言語処理技術の利用”, 電子情報通信学会誌,Vol.85,No.12, pp.942–943 (2002)
- [11] L. Bachman, “Fundamental considerations in language testing,” Cambridge University Press (1990)

参考文献

- [12] J. Bernstein, M. Lipson, G. Halleck, and J. Martinez, “Comparison of oral interviews and automatic tests of spoken language,” Language Testing Research Colloquium (LTRC’99) (1999)
- [13] International Phonetic Association, “Handbook of the International Phonetic Association,” Cambridge University Press (1999)
- [14] 川越いつえ, “英語の音声を科学する”, 大修館書店 (1996)
- [15] 竹蓋幸生, “日本人英語の科学”, 研究社 (1982)
- [16] 窪園晴夫, 太田聡, “音韻構造とアクセント”, 研究社 (1998)
- [17] Microsoft Corporation, “ENCARTA インタラクティブ英会話” (2000)
- [18] メディア教育開発センター, “Listen to Me!,” NHK エデュケーショナル (2000)
- [19] Transparent Language, “English Now! Ver.6.0J,” スリー・エー・システムズ (1998)
- [20] 山田玲子, “英語リスニング/スピーキング”, <http://exp.atrcall.jp/> (2000)
- [21] “UCSC Perceptual Science Laboratory,” <http://mambo.ucsc.edu/>
- [22] 江森正, 篠田浩一, “音声認識のための高速最ゆう推定を用いた声道長正規化”, 電子情報通信学会論文誌, Vol.J83-D-II, No.11, pp.2108-2117 (2000)
- [23] 朝川智, 峯松信明, 広瀬啓吉, “音声の構造的表象を用いた音声認識における特徴両空間分割とその効果”, 日本音響学会秋季講演論文集, 3-Q-10, pp.229-232 (2007-9)
- [24] Silke Witt, Steve Young, “Computer-assisted pronunciation teaching based on automatic speech recognition,” Proceedings of ICLSP (1998)
- [25] D.Reynolds and L.P.Heck, “Speaker verification:from research to reality,” Proc. Int. Conf. Acoustics, Speech, and Signal Processing, tutorial session (2001-5)
- [26] M. Pitz and H. Ney, “Vocal tract normalization equals linear transformation in cepstral space,” IEEE Trans. Speech and Audio Processing, vol.13, pp.930-944 (2005-9)
- [27] 江森正, 篠田浩一, “音声認識のための高速最ゆう推定を用いた声道長正規化”, 信学論 (D-II), vol.J83-D-II, no.11 (2000-11)
- [28] N. Minematsu, S. Asakawa, and K. Hirose, “Structural representation of the pronunciation and its use for CALL,” Int. Workshop on Spoken Language Technology (SLT’2006), pp.126-129 (2006-12)

- [29] 峯松信明, 朝川智, 広瀬啓吉, “線形・非線形変換不変の構造的表象とそれに基づく音響モデリングに関する理論的考察”, 日本音響学会春季講演論文集, 1-P-12, pp.147-148 (2007-3)
- [30] 平手小太郎, “空気調和・衛生工学における数学の利用(8) 多変量解析の基礎”, 空気調和・衛生工学 第77巻 第8号, pp.61-68 (2003)
- [31] 齋藤堯幸, “多次元尺度構成法”, 朝倉書店 (1980)
- [32] <http://htk.eng.cam.ac.uk/>
- [33] 牧野武彦, “日本人のための英語音声学レッスン”, 大修館書店 (2005)
- [34] 朝川智, 峯松信明, 広瀬啓吉, “強勢・弱勢母音における音響的普遍構造のサイズと発音習熟度に関する分析”, 日本音響学会春季講演論文集, 2-1-11, pp.225-226 (2005-3)
- [35] 朝川智, 村上隆夫, 峯松信明, 広瀬啓吉, “音声の構造的表象に基づく日本語母音系列連続発声の認識”, 電子情報通信学会音声研究会, SP2006-105, pp.119-124 (2006-12)
- [36] 下村直也, 朝川智, 峯松信明, 広瀬啓吉, “制約条件付きクラスタリングによる連続音声からのイベント境界検出”, 電子情報通信学会音声研究会, SP2007-12, pp.25-30 (2007-6)
- [37] 下村直也, 朝川智, 峯松信明, 広瀬啓吉, “時間制約を持つクラスタリングによる連続音声の自動セグメンテーション”, 日本音響学会秋季講演論文集, 3-4-4, pp.353-356 (2007-9)
- [38] 下村直也, 峯松信明, 山内豊, 広瀬啓吉, “ボトムアップクラスタリングを用いたシャドーイング音声の自動採点”, 電子情報通信学会音声研究会, (2008-3, 発表予定)

発表文献

- [1] 鎌田圭, 朝川智, 峯松信明, 牧野武彦, 広瀬啓吉, “音声の構造的表象を用いた英語学習者の分類に関する実験的検討”, 電子情報通信学会音声研究会, SP2006-77, pp.7-12 (2006-11)
- [2] 朝川智, 鎌田圭, 峯松信明, 牧野武彦, 広瀬啓吉, “音声の構造的表象に基づく学習者分類の教育的妥当性に対する実験的検討”, 日本音響学会春季講演論文集, 3-8-10, pp.243-244 (2007-3)
- [3] 朝川智, 鎌田圭, 峯松信明, 牧野武彦, 広瀬啓吉, “米語母音発音に対する絶対的評価と相対的評価間の相関分析”, 日本音響学会春季講演論文集, 2-2-1, pp.373-374 (2007-3)
- [4] 鎌田圭, 朝川智, 峯松信明, 牧野武彦, 広瀬啓吉, “学習者による米語母音発音に対する絶対的評価と相対的評価間の相関分析”, 電子情報通信学会音声研究会, SP2007-36, pp.73-78 (2007-7)
- [5] N. Minematsu, K. Kamata, S. Asakawa, T. Makino, T. Nishimura, and K. Hirose, “Structural assessment of language learners’ pronunciation,” Proc. INTERSPEECH, pp.210-213 (2007-8)
- [6] 鎌田圭, 朝川智, 峯松信明, 牧野武彦, 広瀬啓吉, “発音の構造的表象に基づく母音矯正度推定の高精度化”, 日本音響学会秋季講演論文集, 2-4-3, pp.327-330 (2007-9)
- [7] N. Minematsu, K. Kamata, S. Asakawa, T. Makino, and K. Hirose, “Structural representation of the pronunciation and its use for classifying Japanese learners of English,” Proc. ISCA Tutorial and Research Workshop on Speech and Language Technology in Education (SLaTE), CD-ROM (2007-10)
- [8] 鎌田圭, 高澤真章, 竹内京子, 朝川智, 峯松信明, 牧野武彦, 広瀬啓吉, “大規模英語学習者を対象とした音声の構造的表象に基づく発音分類とその応用”, 情報処理学会全国大会講演集 (2008-3, 発表予定)
- [9] 鎌田圭, 高澤真章, 朝川智, 峯松信明, 牧野武彦, 広瀬啓吉, “音声の構造的表象に基づく英語発音分析結果の視覚化に対する一考察”, 日本音響学会春季講演論文集 (2008-3, 発表予定)
- [10] 高澤真章, 鎌田圭, 竹内京子, 朝川智, 峯松信明, 牧野武彦, 広瀬啓吉, “大規模英語学習者を対象とした音声の構造的表象に基づく発音評価とその応用”, 日本音響学会春季講演論文集 (2008-3, 発表予定)

付録 A

柏地域の英語発音分類

東京大学柏キャンパス一般公開で収録した 564 名分の英語音声 (9.4 節) に対して発音分類を行い, 樹形図として表記した図が図 A.1 である. 東京大学柏キャンパス一般公開での収録は 3 年間行われており, のべ 697 名の英語音声収録されている. 録音環境が万全でなかったため, 事前の聴取によって録音状態が良好な音声として選別された 564 名分を分類している (6.3 節).

ここで分類した音声は柏周辺に居住する人々の音声のみであるが, 対象を世界中の人々に拡大することにより, 各地の英語発音を方言に基づいて分類した世界地図を作ることも理論的には可能である. ただし, 発音分類に用いているボトムアップクラスタリングは計算量のオーダーが $O(n^2)$ であるため, 大規模英語発音分類では分類が現実的な時間には終了しないことが予想される. 大規模英語発音分類に適した分類アルゴリズムを検討する必要がある.

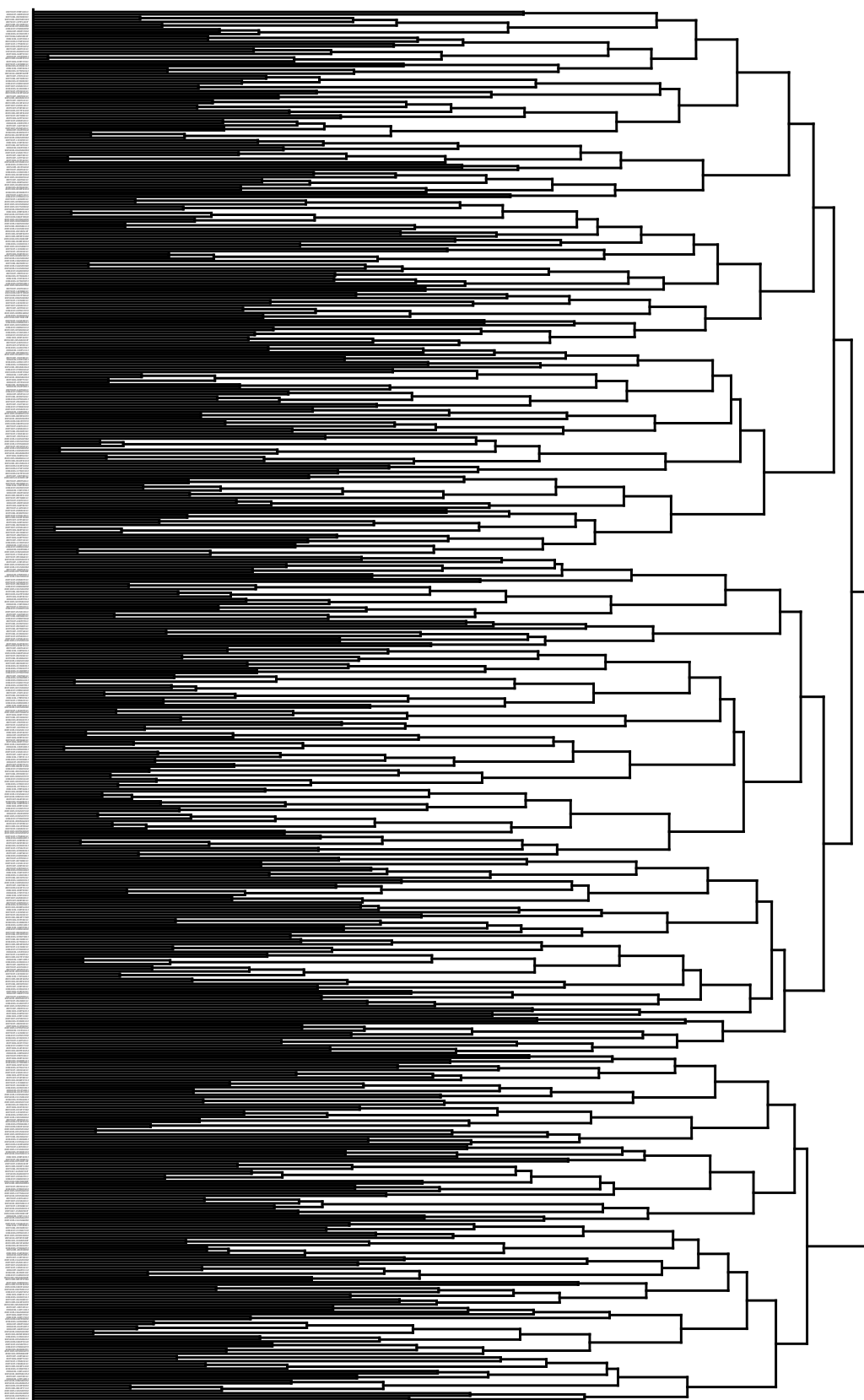


図 A.1: 柏地域の英語発音分類

付録B

実地試験で提供した発音カルテ

B.1 1 ページ目 (カルテの説明)

英語母音と日本語母音の差異と、発音状態を示す樹形図の見方について記述した。

B.2 2 ページ目 (発音状態と発音教示)

上部に発音状態を示す樹形図を (5.1 節)、下部に母音矯正度推定による発音教示 (5.3 節, 第 8 章) を記述した。母音矯正度推定による発音教示に関しては、バーが長い母音が練習を要する母音となる。また、母音矯正度推定の比較対象となる教師は、英語音声収録前に学習者が自由に選択したものである (図 1.1)。

B.3 3 ページ目 (発音分類)

学習者の発音、図 1.1 で選択できる教師発音、デモ会場にいるスタッフの発音、発音タイプ (B.4 節で後述)、の合計 20 人で話者分類したものを MDS を用いて表示したものである (5.2 節, 6.6 節, 第 7 章)。学習者は、自らの周囲にいる話者によって、おおよそ自分の発音を推測することができる。

B.4 4 ページ目 (発音タイプ判定)

学習者の発音タイプを判定し、発音タイプに合わせたコメント (B.6 節) を表示した。コメントの監修者は音声学者の牧野武彦准教授である。発音タイプは 5 つあり、2005 年、2006 年の東京大学柏キャンパス一般公開で収録した学習者音声 (約 400 名) を 5 タイプに発音分類 (5.2 節) することにより決定した。

また、発音が最も似ている話者を前述の約 400 名の学習者音声から探し出し、その最似話者の番号を表示した (5.2.2 節)。最似話者の発音は、別途端末を用意し聴くことができるようにした。

B.5 5 ページ目 (発音状態)

学習者の発音状態と、学習者が選択した教師の発音状態を、MDS を用いて表示した (第 6 章)。1 段目は /あ/, /ɑ/, /æ/, /ʌ/, /ɚ/, /ə/, 2 段目は /あ/, /お/, /ɑ/, /æ/, /ʌ/, /ə/, /ɔ/, 3 段目は /い/, /え/, /æ/, /ə/, /i/, /i/, /e/, 4 段目は /う/, /お/, /ɑ/, /ə/, /u/, /u/, /ɔ/ を二次元表示した。

B.6 タイプ別コメント

B.4 節にて用いたタイプ別のコメントを掲載する。なお、本カルテで判定されるタイプは、下手な日本人英語発音から上手な英語発音までであり、母国語話者のタイプ (方言)

は用意していないが、イギリス英語発音やアメリカ英語発音といった方言タイプを用意することは可能である。

B.6.1 タイプ A に対するコメント

bit/beat, put/boot における音質・音色の区別がまだまだ不完全です。それ以外はほぼアメリカ標準発音に近い発音になっています。

B.6.2 タイプ B に対するコメント

bird の発音は比較的良いのですが、その他はまだまだ日本語の「アイウエオ」の影響が色濃く出ている発音になっています。注意しましょう。

B.6.3 タイプ C に対するコメント

ほぼアメリカの標準発音が再現できています。

B.6.4 タイプ D に対するコメント

「日本人発音」の典型的なパターンです。英語の 11 の異なる音質の母音を、日本語の 5 母音で（これらに長短の変化をつけて）置き換えてしまっています。要注意です。

B.6.5 タイプ E に対するコメント

bit/beat の音質・音色が区別がまだまだ不完全ですが、アメリカ標準発音に比較的近い発音になっています。



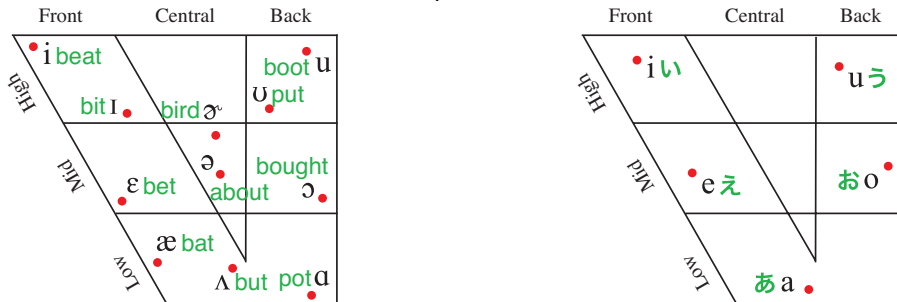
078 様に対する診断結果

難しい言葉があったら、お父さん、お母さんに聞いてみてね。



アメリカ英語と日本語の母音の配置の復習から始めよう！

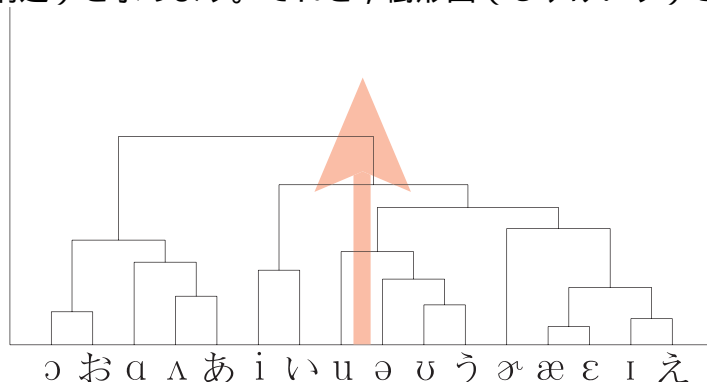
音声収録前の「音声ってどうやって作られるの？」で、アメリカ英語と日本語の母音の配置（顔を横から見た場合、舌の一番高いところが母音によってどう動くか）を説明しました。その図を以下に再度示します。分析結果はこの図と比較しながら、「にらめっこ」すると良いでしょう。



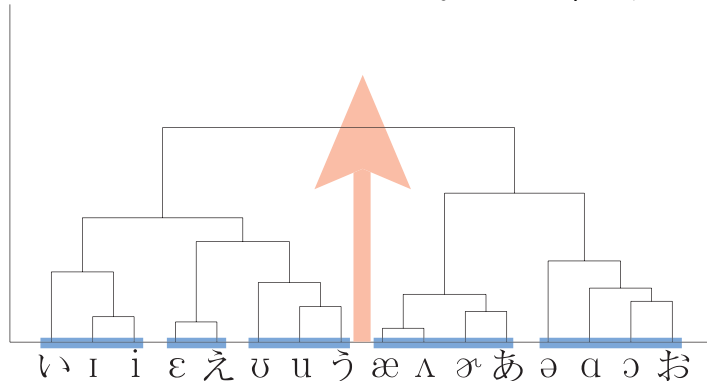
アメリカ英語の母音 ə は音色が不安定で、単語によってその位置が変わる音として知られています。

まずは、発音の上手な日本人の音声に対する分析結果から見てみよう！

発音の上手な日本人の音声に対する結果から見ていきましょう。収録した16種類の母音から、上の図のような母音配置（構造）を求めます。それを、樹形図（じゅけいず）と呼ばれる形で示します。



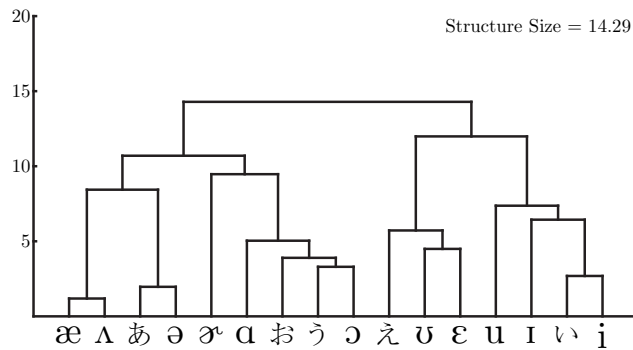
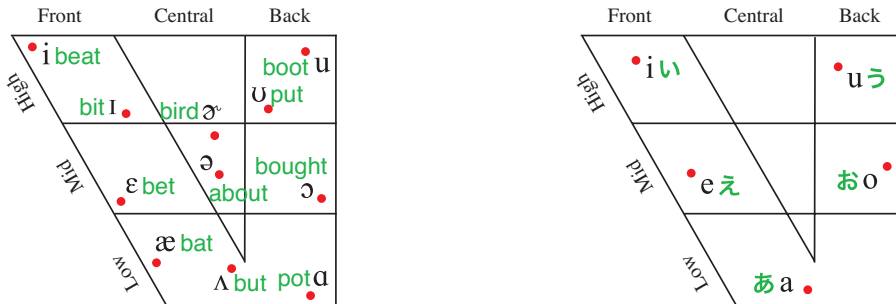
下から上に向かって眺めてください。似ているものほど、下の方でくっつきます。つまり、区別できてない、ということです。上の二つの母音図と比較してみてください。i(beat) と ɪ(bit) が離れている点、a(pot) に対して、ɔ(bought) や「お」が離れている点、ε(bet) と æ(bat) が殆ど同じである点など「へ～、意外～」っと思うところも多いのではないのでしょうか？今回は ə は u, ɝ, 「う」の方に寄っていますね。次に、その人が「わざと」日本語的な発音をした場合の結果を示します。bat を「バット」、but も「バット」...として発声した場合の結果です。さ～て、どうなるかな？



あははは。16個の母音があるけど、ほぼ5個の母音しか使っていない様子分かるよね。

お待ちかね。「あなた」の分析結果を見てみよう！

分析結果を二つの母音図と一緒に示します。チェックポイントはおよそ以下に示す通りです。

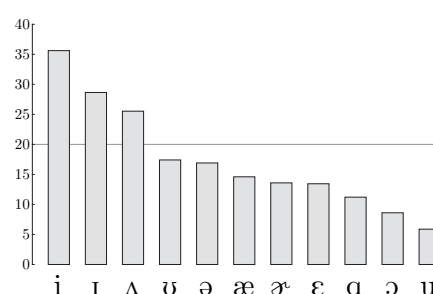
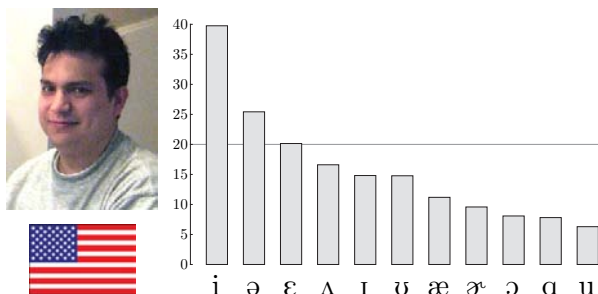


チェックポイントはここ!! さあ、自分の発音を「自分で」チェックしてみよう!!

- æとε この二つはいつも仲良しです。あなたの発音はどうなってますか？
- ɪとi これはかなり違う音ですよ。ɪは「イ」と「エ」の中間音とも言われます。
- ɪと, εやæ その結果, ɪはεやæに似くるはずですが, あなたの発音ではどうなってます？
- ɑとɔ アメリカ英語では, これらが似ている方言もあります。母音図でɔが下がってきます。
- əと, ʌやɑやæ 後ろ三つは舌を下げる音です。əがこれらと似てたら重症。救急車呼びます。
- əとɝ この二つが仲良し, という人もいるかもしれませんね~。どうです? あなたの発音。
- ɝ 実はrと殆ど同じです。母音として使われるのか, 子音として使われるのか, の違いだけ。

さてさて、どこから直していこうか。よ~く、考えてみよう！

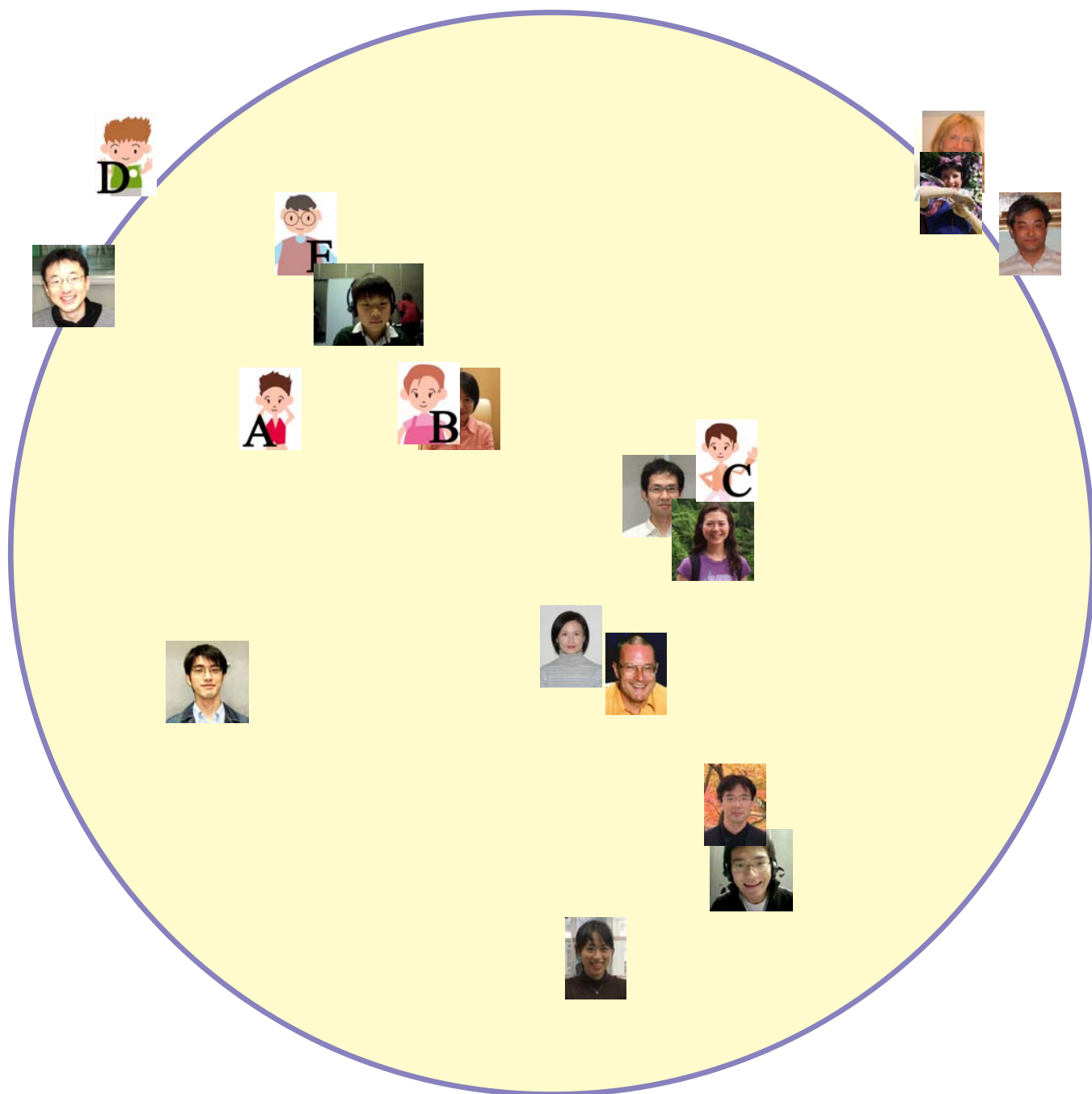
あなたの分析結果を、あなたが選んだ二人の先生と比較して「どの母音から直すべきなのか」を示します。目標とする発音に到達するための最短コース, という訳です。スコアの低い母音(より左に示されている母音)ほど「重症」で、緊急入院が必要な母音です。20程度であれば、ほおっておいて大丈夫ですよ。さてさて、どんな結果が出てますか? ゴールまでの道のりは長いですか? 短いですか?



どちらの先生に近づきたいかによって、結果が少し異なるでしょう。英語は方言によって母音の音質・音色が変わりますから、先生が、どの地方の出身なのかによって直すべき母音が変わってきます。もちろん先生の発音は両親(異なる方言話者かもしれない)の発音の影響を強く受けています。結局、厳密には、み~んな一人一人違う発音になっています。大切なのは、相手の心に届く「あなた自身」の英語発音を身につけることです。一人よがりじゃいけません。相手あっての言葉ですから。

みんなの英語を比較してみよう！

先生の英語と比較して直すべき母音を示しました。次は、13人の先生、6名の東大のお兄ちゃん・お姉ちゃん、5種類の発音（下段を参照して下さい）そしてあなたの英語を比べてみましょう。下の図は、これらの英語発音を比較した結果を顔写真で示したものです。写真が近いものほど、近い発音をしています。自分に似た発音をしているお兄ちゃん、お姉ちゃんがいたら、“Hello!! My name is”と声をかけてごらん。あなたの発音と、よ～く似た発音で返事してくれるかもしれません。



あなた（078）の英語発音のタイプは …

「英語は方言によって母音の音質・音色が変わる」と書きました。柏オープンキャンパスで2年間に渡り、400名以上の音声データを収録しました。これらの発音を分類し（柏英語の方言分類!?!）、5つのタイプA～Eを定義しました。その分類とあなたの発音を比べると、あなたは**タイプA**に分類されると判定されました。この**タイプA**の発音に分類される話者の中で、最も代表的な話者の発音を三省堂グランドセンチュリー英和辞典の執筆陣の一人である、音声学者の牧野武彦先生（中央大学）に聴取して戴いたところ、下記のコメントが寄せられました。なお、あくまでも「ある特定の話者」の発音に対するコメントであって、該当するタイプには微妙に異なる発音が沢山含まれております。

bit/beat, put/boot における音質・音色の区別がまだまだ不完全です。それ以外はほぼアメリカ標準発音に近い発音になっています。

また、約400人の発音データから、あなたそっくりの発音の持ち主を探し出しました。その人の話者番号は**0102**です。廊下に設置しているパソコンで、この「発音そっくりさん」の音声を聞くことができます。恐らくこの人と英会話すると、妙にヒアリング力が向上したかのように感じるはずですよ。自分とよく似た「訛り」を持った相手ですから。また、タイプA～Eの代表者、先生方・東大のお兄ちゃん・お姉ちゃんの発音など、2年間の間に収録した全ての音声データを聞くことができます。

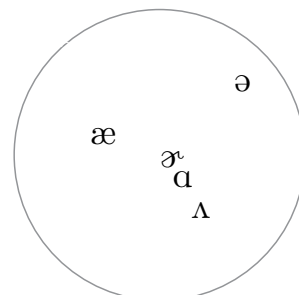
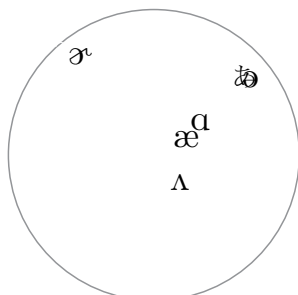
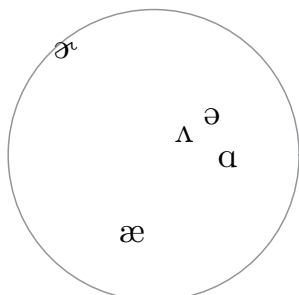
最後に。お父さん、お母さん、特に、英語の先生方へのメッセージ

小学校への英語教育導入が叫ばれています。その一方で、英語教育の環境作りに対する懸念が聞こえてきます。峯松研究室では、その環境作りに励んでいます。教育関係者の方（特に小・中学校の関係者）で、本技術開発にご興味のある方は、是非 mine@gavo.t.u-tokyo.ac.jp までご連絡下さい。

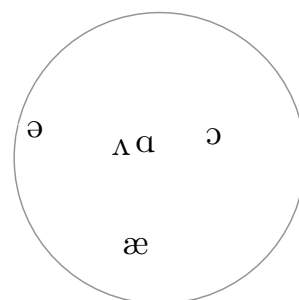
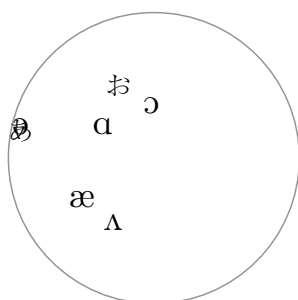
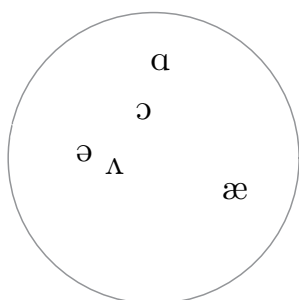
あなたの英語発音を分析した結果です。先生の英語発音を分析した結果と比較してみましょう。
先生の発音と似ているものは、記号の位置関係も似ているはずですよ。



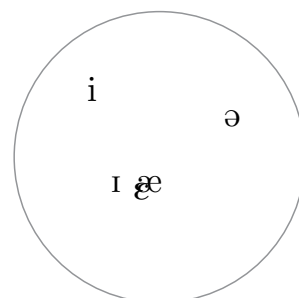
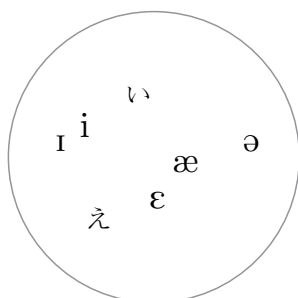
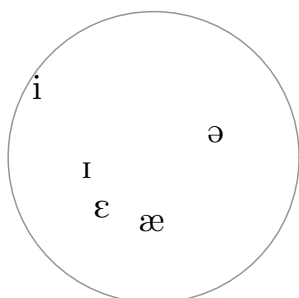
「あ」に近い英語発音の分析結果



「あ」「お」に近い英語発音の分析結果



「い」「え」に近い英語発音の分析結果



「う」「お」に近い英語発音の分析結果

