

Title	変形聴覚フィードバックによる発話系の補償動作に関する研究
Author(s)	内山田, 太一
Citation	
Issue Date	2007-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/3612
Rights	
Description	Supervisor: 党 建武, 情報科学研究科, 修士

修 士 論 文

変形聴覚フィードバックによる
発話系の補償動作に関する研究

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

内山田 太一

2007年3月

修士論文

変形聴覚フィードバックによる
発話系の補償動作に関する研究

指導教官 党建武 教授

審査委員主査 党建武 教授
審査委員 徳田功 准教授
審査委員 赤木正人 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

510016 内山田 太一

提出年月: 2007年2月

概要

音声生成・知覚過程がどのような相互作用を行っているのかを説明する研究が、これまで数多く行われている。しかし、十分な実験に欠くため、人間本来がもつこのメカニズムについて多くの謎が残っている。聴覚系と発話系の相互作用の存在を示すものとして、聴覚フィードバックがある。聴覚フィードバックとは発話音声を聴覚系にフィードバックしながら発話器官の制御を行うもので、この機能により、正常な発話を維持することが可能となっていると考えられている。聴覚フィードバックが音声生成過程においてどのような役割を果たしているのかについて検討するため、実時間の変形聴覚フィードバック実験を行った。本研究では、従来の音声スペクトルに加え、磁気センサシステムを用いて、調音器官の補償動作についても計測を行い、発話系の補償動作の解明を試みる。

目次

第1章	序論	1
1.1	研究背景	1
1.2	研究目的・手法	2
1.3	本論文の構成	2
第2章	音声生成・知覚過程に関する研究	3
2.1	研究背景	3
2.2	ことばの鎖 ~Speech Chain~	3
2.3	音声知覚の運動理論 ~Motor Theory~	4
第3章	変形聴覚フィードバック実験システム	6
3.1	聴覚フィードバック	6
3.2	変形聴覚フィードバック	6
3.3	実験要件	8
3.4	実験アプローチ	9
3.5	実験システム	10
第4章	聴覚フィードバックにおける非線形な摂動による補償動作の測定	12
4.1	聴覚フィードバック実験	12
4.1.1	実験目的	12
4.1.2	実験手順	13
4.1.3	分析方法	13
4.1.4	実験結果	14
4.2	考察	14
第5章	調音器官の補償動作の測定	18
5.1	変形聴覚フィードバック実験	18
5.1.1	実験目的	18
5.1.2	磁気センサシステム	18
5.1.3	実験手順	18
5.1.4	分析方法	21
5.1.5	主成分分析	21

5.1.6	実験結果	25
5.2	考察	32
第6章	まとめ	33
6.1	結論	33
6.2	今後の課題	33
	謝辞	37
	本研究に関する研究業績	38

目次

2.1	ことばの鎖	4
3.1	実験システム構成図	11
4.1	以前の摂動の与え方	12
4.2	本研究での摂動の与え方	12
4.3	トライアル毎の構成	13
4.4	F1 と F2 で異なる割合の摂動を与えた場合の反応 (-15°)	15
4.5	F1 と F2 で異なる割合の摂動を与えた場合の反応 (15°)	16
4.6	補償動作ありの場合 (15°)	16
4.7	人間がフィードバックする音声を認識する範囲 (赤線で太字ほど認識する)	17
5.1	磁気センサシステム	19
5.2	磁気センサ装着図	20
5.3	実験風景	20
5.4	磁気センサの位置	20
5.5	正常な測定結果	21
5.6	異常な測定結果	21
5.7	摂動を与えた場合の反応 (/e/ /i/)	27
5.8	補償動作ありの場合 (/e/ /i/)	27
5.9	摂動を与えた場合の反応 (/e/ /u/)	28
5.10	補償動作ありの場合 (/e/ /u/)	28
5.11	第1主成分の固有ベクトル	29
5.12	第2主成分の固有ベクトル	29
5.13	第3主成分の固有ベクトル	29
5.14	/i/の第1, 第2主成分の関係 (赤: 補償あり、青: 補償なし)	30
5.15	/i/の第1, 第3主成分の関係 (赤: 補償あり、青: 補償なし)	30
5.16	/i/の第2, 第3主成分の関係 (赤: 補償あり、青: 補償なし)	30
5.17	/u/の第1, 第2主成分の関係 (赤: 補償あり、青: 補償なし)	31
5.18	/u/の第1, 第3主成分の関係 (赤: 補償あり、青: 補償なし)	31
5.19	/u/の第2, 第3主成分の関係 (赤: 補償あり、青: 補償なし)	31
5.20	日本母音の調音位置	32

表 目 次

4.1	被験者の音響パラメータ	14
5.1	/e/ /u/のフォルマント変換量	19
5.2	音響分析と主成分分析の結果	26

第1章 序論

1.1 研究背景

音声生成と知覚は表裏一体の関係にあり、多様な音韻を生成するためには多様な音韻を聞き分ける能力が不可欠である。このため人間は、音声生成過程とその逆過程の音声知覚の繰り返しにより母国語を含む言語音声知覚機構の発達を遂げた。このような発達には音声の生成と知覚とを密接に結ぶ情報交換の通路（ことばの鎖）が大きく関与している [1]。このような “ことばの鎖” について話者内部（脳内）での音声生成と音声知覚との相互交信を説明するため、Liberman らは 1960 年代より音声知覚の運動理論（Motor theory of speech perception）を提唱してきた [2][3]。しかし、話者内部（脳内）において、音声生成・知覚がどのように関連し存在しているのかについてはまだ明らかにされていない。

これまでの先行研究では、発話時の音声生成・知覚の相互作用を明らかにするために、聴覚フィードバックに着目した研究が数多く行われてきた。その中で Lombard は雑音環境下では通常の会話より発話音声が大きくなり、基本周波数も高くなる現象を観測した [4][5]。その現象は Lombard 効果として知られている。また、発話音声を遅延して話者に聞かせる遅延聴覚フィードバック実験（DAF）では、吃音や発話速度が遅くなる等の現象が生じる [6]。これは、聴覚フィードバックの効果を表す顕著な例である。しかし、上記の報告は、定性的な性質を述べるにとどまっており、これらの現象を起こすメカニズムについての説明は不十分である。また、DAF のような発話過程を破壊するような実験では、発話過程自体が破綻してしまうため、発話への影響に対する定量的な分析は困難である。

そこで河原らは、非破壊的な実験パラダイムとして変形聴覚フィードバック（TAF）を提案した。河原らの一連の研究により、聴覚系が発話系の基本周波数制御に重要な役割を果たすことが明らかとなったが、フォルマントの影響については明らかにされなかった [7][8][9]。このため、佐藤と斉藤が、日本母音に関して短時間のフォルマント変形に対する発話器官の補償動作について、音声のスペクトルグラムを用いて検証を行った。その結果、スペクトルの変動は見られたが、それが補償動作か否かは確認できなかった [10][11]。そこで、松岡はさらにフィードバック音声の音質改善を行って、調音の違いが口唇の突き出しのみである中国母音の [i] と [y] を音声資料として用い、音声のスペクトルや、筋電信号、画像情報での分析を行ったところ、補償動作が確認された。さらに、複数の調音点の補償が相互に絡み合っている可能性が示唆された [12]。また、田中は日本母音を用い、日本語の中立母音 /e/ に対して第 1、第 2 フォルマントに摂動を加え、それぞれ /a/、/i/、/u/ の方向へ変化させ被験者にフィードバックすることにより摂動量と補償動作の関係につい

て考察した [13]。日本母音を音声試料として用いて摂動量を変化させたところ、摂動量により音声の曖昧さが増大した時に補償動作が大きくなっていることが明らかとなった。

これらの先行研究における知見により音声のスペクトル形状の変形により、補償動作の存在が示されている。しかし、音声スペクトルの分析が中心であり、調音器官の動作については調音の違いが口唇のみである中国母音を用いた場合だけでしか考慮していない。このため、日本母音についても調音器官がフォルマントの形状の変形を補うような方向で補償をおこなっている可能性が考えられる。

1.2 研究目的・手法

以上のような背景から考えられる可能性を明らかにするため、本研究では、フィードバック時に存在可能である瞬時的な調音器官の補償動作や応答特性を、音声信号以外の測定手法を加えて明らかにしていく。具体的には音声信号に加え、磁気センサシステム (EMA) を用いて調音運動からの観測も同時に行う。このように、本研究では、音声知覚の面からだけでなく調音運動のレベルでも考察できるような音声以外の計測手法を導入する。このような手法から、各々の調音器官の調音運動における補償動作を定量的に計測する。

また、発声・発話動作における聴覚フィードバックの影響に関して、基本周波数においては様々な知見が得られているが、フォルマントの変形において報告された例は少ない。そこで、本研究では、発話音声のフォルマントを変形したフィードバック音声を用い、短時間の変形聴覚フィードバック実験における発話器官の逐次的な応答特性を観測する。

以上のような手法を用いて、各発話期間の調音活動を定量的に測定することにより、聴覚フィードバックにおける補償動作の存在を明らかにし、聴覚系と発話系の相互作用を定量化することが本研究の目的である。

1.3 本論文の構成

本論文は6章により構成される。第1章では研究背景、目的、手法の本論文に関する導入部にあたる。第2章は第1章をより具体的に記述している。第3章では、本実験で用いた実験システムに関する記述である。第4章は、非線形なフォルマント変形を行い、それを被験者に与えた実験についての記述である。第5章では、磁気センサシステムを用いて、舌・口唇のような調音器官の反応を確認する聴覚フィードバック実験についての記述である。第6章はまとめとして、本論文の結論と今後の課題について述べた。

第2章 音声生成・知覚過程に関する研究

2.1 研究背景

音声生成と知覚過程は、人と人との音声コミュニケーションにおいて重要な機能を果たすものであり、コミュニケーションを円滑に保つためには両方の過程が一体となって働く必要がある。音声合成・認識は、人間の音声生成・知覚機構の工学的な表現で、人と機械とのコミュニケーションのインターフェイスとして期待されている。

近年、音声合成・認識は、コンピュータの発展と共に大きな進歩を遂げ、理想的な環境で応用できるレベルに近づいた。音声認識に関して言えば、現在用いられている音声認識方法は、理想的な環境ではほとんど行き着くところまで行き、次のステップとして、どのような環境においても高性能な音声認識を行う手法の開発が課題となってきた。ところが、現在の手法では、あらゆる実例を用意することが必要である。例えば、現在の音声認識システムは、雑音環境では認識率が著しく低下し、ほとんど無能に近い。それに対して、人間の音声認識能力は雑音環境にほとんど左右されない。

音声は本来人間により認知され人間により生成されるものである。“音（音声）を聴くとはどういうことであるのか”という原点にもう一度立ち帰り、音声生成と音声知覚という人間本来のメカニズムを解明することは、音声合成・認識の発展に大きく貢献すると考えられる。

2.2 ことばの鎖 ~ Speech Chain ~

音声は本来人間により生成され人間により知覚されるものである。人間は、音声生成過程とその逆過程の音声知覚の繰り返しにより母国語を含む言語音声を学習している。このような学習により音声の生成と知覚とを密接に結ぶ情報交換の通路（ことばの鎖）ができたのである [1]。“音（音声）を聴くとはどういうことであるのか”を解明するためには、“ことばの鎖”の中での知覚機構を考える必要があり、また同様に、“音声を生成するとはどういうことであるのか”を解明するためには、生成機構を考える必要がある。

“ことばの鎖 ~ Speech Chain ~”とは、図 2.1 に示すように、言語学的段階から生理学的段階を経て音響学的段階へと移行し、また逆の過程をたどる、人間同士がコミュニケーションを円滑に行うためのサイクルのことをいう。つまり、人間が相互にコミュニケーションを行う場合、言葉を発して相手に自分の考えを考え、感情などを伝えようとする一方で、相手が伝えてきたことを考え、感情などの情報を受け取り、理解して、そして適切

な応答を行うサイクルのことである．自分自身の中でこのサイクルが上手くまわることにより，コミュニケーションが円滑に保たれている．

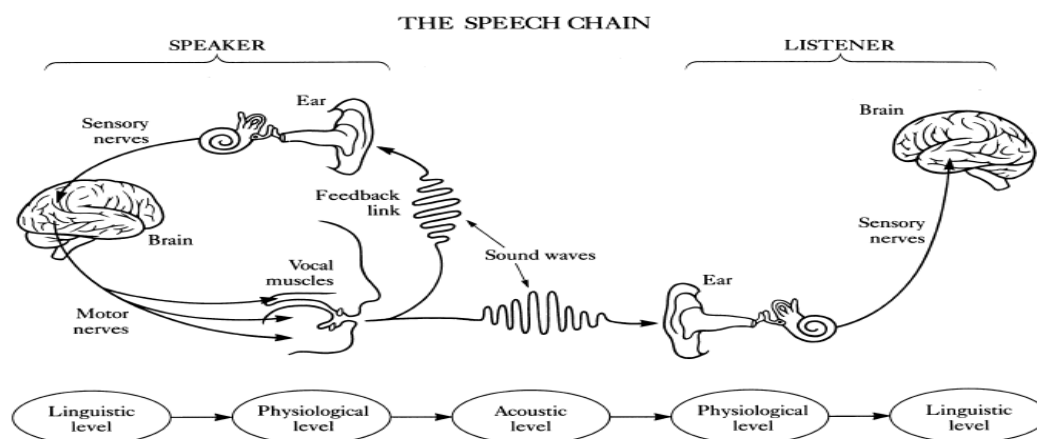


図 2.1: ことばの鎖

会話音声の場合，発話速度の変化や感情などの影響と，発話器官の生理学的な制約により生成された音声は変形され，最終的には音声波形と音素系列のはっきりとした対応関係は見られなくなる．しかし，人間には音声の生成と知覚との情報交換通路 "ことばの鎖" が存在しているため，音声と音素列の対応関係を正しく見つけ出すことが可能である．しかし，"ことばの鎖" について話者内部（脳内）での存在実態はまだ明らかにされていない．

2.3 音声知覚の運動理論 ~ Motor Theory ~

実験により様々な仮説が提案される一方で，"ことばの鎖" について話者内部（脳内）での音声生成と音声知覚の相互交信を説明するような，音声の知覚機構に関する多くのモデル化が提案されてきた．

中でも 1960 年代より Liberman らは，音声知覚の運動理論 (Motor theory of speech perception) を提唱してきた [2] [3]．運動理論の基本的な考え方は以下の通りである [14]．

- (1) 音声は他の音とは異なる特殊なものであり，人間には，音声のみを処理する特殊なモード（音声モード）が生得的に備わっている．
- (2) 音声モードでは，音声の生成と知覚に共通の符号が用いられる．その符号とは調音器官の働き（ジェスチャ）に関するものである．
- (3) 音声の知覚は，音響信号から，それを生成した発話者の調音ジェスチャを推定することに基づいてなされる．

(4) 調音ジェスチャの推定は，調音器官の内部モデルを用いた「合成による分析」に基づく．

しかし，Motor Theory の裏づけは非言語的な運動に対してから可能性を示されているが，音声知覚に特化していえるのかは明らかにされていない．一方，Action Theory では，音声に限らず一般に信号を形成する原因となった運動や事象を “直接” 検知することが知覚なのだと仮定しており，音声知覚では調音運動が “直接” 検知されると主張する．“直接” とは推測や仮説検証などの思考過程を経ずにという意味である．しかし，どのような機構によって “直接” 検知が可能であるのか，まだ十分に解明されていない．

第3章 変形聴覚フィードバック実験システム

3.1 聴覚フィードバック

”ことばの鎖”の実態を調べるため、これまで音声生成と知覚の相互作用に関して様々な実験を通して仮説が提案されてきた。

聴覚系と発話系の相互作用の存在を示すものの一つとして、”聴覚フィードバック”がある。“聴覚フィードバック”は発話音声を聴覚系にフィードバックしながら発話動作の制御を行うものでこの機能により正常な発話が可能となっていると考えられている。もし聴覚フィードバックが発話過程で常に働いているなら、フィードバック音声を人為的に変形し耳に提示した場合、その変形によって発話系に何らかの影響があるはずである。このような前提をもとに、古くから発話時における”聴覚フィードバック”の役割に関する実験が数多く行われてきた。

Lombardらは、初期の研究において、雑音によって発話音声をマスクした環境下でも発話が可能であることを示している[4, 5]。このことは、言語習得後に聴覚を失った後天性難聴者においても明瞭な発話が維持されることと一致する。また、言語獲得後に聴力を失った子供に訓練を施しても正確な発話を身に着けることが困難であることが報告されている。[15, 16, 17] これらの知見は聴覚フィードバックが言語獲得過程においては極めて重要な役割を果たすが、それ以降はあまり関与しないことを示す根拠となっている。しかし、Laneらは雑音環境下においても発話は可能であるが、発話者の発話音量を雑音の音量に応じて増大させている(Lombard効果)ことも示している[5]。また言語習得後においても聴力を失うと、/s/のような無声摩擦子音の生成や基本周波数の制御が球速に悪化すること[18]や、発話音声に数十～数百msの遅延を挿入する(遅延聴覚フィードバック: DAF)ことで発話動作を著しく崩壊させることから[6]、聴覚フィードバックが音声生成に必ずしも決定的に関与することは考えにくい、いくつかの知見は少なくとも正確に発話するうえで必要であることを示している。

3.2 変形聴覚フィードバック

これまで、上記で述べた”聴覚フィードバック”に関する実験を通じて、音声生成と知覚過程が相互にどのように関連しているのかを調べる研究が、様々な手法を用いて行われ

てきた．聴覚からの入力と音声生成過程及び言語習得過程での発話動作に関連があるならば，聴覚からの入力に変化が起こった場合，発話動作にも変化が現れることが予想される．変形聴覚フィードバックとは，聴覚への入力に対して変化（摂動）を与えることで発話動作に現れる変化を測定する実験方法である．聴覚と発話動作との関連を調べる方法として聴覚フィードバックによる実験が用いられてきた．Lombardら [4, 5] による雑音で発話音声を遮断する実験や Lee[6] による発話音声に遅延時間を挿入する実験（DAF）などは変形聴覚フィードバックの一種である．

しかし，これらの実験による報告は，聴覚フィードバックの定性的な性質を述べるにとどまっており，これらの現象を起こすメカニズムについての説明は不十分である．特に，DAFのような発話過程を破壊するような実験では，発話過程自体が破綻してしまうため，発話への影響に対する定量的な分析は困難である．

そこで，河原らは，音響パラメータを実時間で変換することにより非破壊的で定量的な分析を可能にする変形聴覚フィードバック（Transformed Auditory Feedback：TAF）を提案した．河原らは発声過程から遅延聴覚フィードバックに伴う発話への影響を避けて，正常な発話に近い状態で聴覚と発話との関係を探ろうとし，実時間で基本周波数に微小な摂動を与えた．この実験により発声された音声の基本周波数の変化による影響として摂動を与えてから約 150ms 程度の遅れを伴い，変化を打ち消す方向への応答がはたらくことが示された．このような河原らの一連の研究により [7, 8, 9]，基本周波数制御における聴覚フィードバックの役割を示す重要な知見が得られた．

一方，スペクトルの変動に対しての聴覚フィードバック実験も，様々な手法を用いて行われている．Houde と Jordan は，発話音声のフォルマント周波数を変化させた変形聴覚フィードバックの環境下で被験者に対し長時間発話させる実験を行った [19] [20]．その結果，話者はフォルマントの変化に対して，発話動作を調整，適応化させていることが示された．現在のところ，聴覚フィードバックと発話動作の関わりは，リアルタイムのフィードバック機能というより，むしろ発話動作の結果として起こる音響現象をモニタし，発話者が意図した音響現象と異なる場合に発話動作を調整（適応化）する機能，あるいは発話動作の運動計画を行う際に用いられる発話動作と音響現象との因果関係を学習によって獲得する機能にあるとされている．さらに，Houde らは，発話動作が適応化した後，マスキングノイズで聴覚フィードバックを遮断した状態においても発話音声のフォルマント周波数に同様な補償的な変化が見られたと報告している [20]．この結果は，聴覚フィードバック条件下の方が補償的な効果が大きいため，変形聴覚フィードバックに対する学習効果以外にも，瞬時的な発話動作の調整がなされているという可能性を示唆している．しかし，本来，このような長時間学習の実験には，人間に備わっている音声生成・知覚メカニズム解明に焦点はあてられてはいなかった．そこで，本学の音情報処理学講座では，フォルマントに変化を加える短時間の変形聴覚フィードバック実験を行ってきた．佐藤はノッチフィルタを用いて発話中に母音/i/の第2フォルマント付近である 2kHz 周辺の周波数成分を除去し，発話者の聴覚にフィードバックする実験を行った．その結果，除去された帯域，及びその周辺で通常発話より数 dB 上昇していることを確認した．佐藤はこの結

果より，スペクトルの変形による補償動作が存在する可能性を示した [10]．続いて斉藤も第1，第2フォルマントを変化させたフィードバック音を用いて聴覚フィードバックの実験を行った [11]．しかし，個人により分散が激しいため，定性的な知見は得られなかったが，この結果からフィードバック音声が発話動作に何らかの影響を与えている可能性が示唆された．そこで，松岡は口唇を調音点とした中国母音を音声資料として用いて，さらにフィードバック音声の音質の改善を行い，音声スペクトルや，筋電信号，画像情報での分析を行ったところ，補償動作が確認できた [12]．また，田中により日本語の中性母音 /e/ に対して第1，第2フォルマントに摂動を加え，それぞれ /a/、/i/、/u/ の方向へ変化させ被験者にフィードバックすることにより摂動量と補償動作の関係について考察した [13]．日本母音を音声試料として用いて摂動量を変化させたところ，摂動量により音声の曖昧さが増大した時に補償動作が大きくなっていることが明らかとなった．

これらの先行研究における知見より音声のスペクトル形状の変形により，補償動作の存在が示されている．しかし，日本母音における調音器官（舌，口唇など）に対する補償動作について定量的な観測，分析は行われていない．このため，音声スペクトルには現れないが，調音器官がフォルマント形状の変形を補うような方向で補償を行っている可能性が考えられる．

3.3 実験要件

佐藤・斉藤らは，実時間によるフォルマント変形聴覚フィードバック実験を行う上で以下の要件が満たされる必要があると報告している [10, 11] ．

要件 1 実時間による音声パラメータ変換を行う．

要件 2 遅延を最小限にする．

要件 3 発話者の音響物理量をできるだけ多く残した自然性の高い変換を行う．

要件 4 変化に対して被験者が修正可能な摂動を与える．

要件 5 被験者が変化を知覚できる摂動を与える．

要件の1つ目はフィードバック音声の変換は実時間処理が行われる必要があるというものである．その理由はフィードバック音声に基本周波数や振幅のような時間変化する音響物理量が保持されていることが重要であるためである．沢田らの報告によれば，基本周波数や振幅包絡の時間情報もフィードバック音声として重要な音響物理量であることから，これらの情報の損失はフィードバック音声として不適切となる可能性がある [21] ．例えば，フィードバック音声として予め被験者の音声を録音したものや合成したものを実験に用いたとすると，発話時の時々刻々と変化する基本周波数や振幅包絡の時間情報等がフィードバック音声に反映されない，という問題が生じる．そのため基本周波数の時間変化や振幅

包絡の時間情報を保持したまま，フォルマントのみが変換されなければならない。

2つ目の要件は，フィードバック音声の遅延は発話動作の破壊 [6] など実験に望ましくない影響を与える可能性があるため，遅延を最小限に抑ええることも重要である．佐藤は実験により，遅延が 30ms 以下であれば，発話に影響しないと報告している [10]．

3つ目の要件は，フィードバック音声は被験者の音声に出来るだけ近い音声を利用すること，つまり，出来る限り音声に話者の個性を保持したまま実時間でフォルマント変換することが望ましい．Shimon らが指摘しているように，聴覚フィードバックの実験で用いるフィードバック音声の自然性は重要である [22]．自然側音により発話が妨害されないことや [5, 23]，多くの情報が失われたフィードバック音声が発話に影響しないことから分かるように [22]，人はフィードフォワード制御により自分の声とそれ以外の音声について弁別する能力が高い [24]．そのため，音響物理量あるいは自然性が損なわれることで，音声の話者の個性を失ってしまい，他人の音声として認知されるという問題が生じる．これは，聴覚フィードバック実験としては不適切な条件となり得る．

4つ目の要件は，摂動として与える変化が被験者により修正可能であることである．もし摂動に対する応答が補償動作であった場合，被験者の発話動作は発話の変化に対して元に戻そうと働くことになる．この時使用する摂動，あるいは音韻，音節によっては，補償の方向が発話機構などの物理的な制約や言語習得時に獲得したモデル内に存在しないことにより，応答が十分に現れない可能性がある．また，発話動作が破綻するような摂動も応答の測定にはふさわしくない．このため，これらの問題を避けるような摂動，及び対象とする音声資料（音韻，音節）の選択が重要となってくる．

5つ目の要件は，摂動が被験者によって知覚可能なことである．摂動に対する応答が，反射運動のような自動レベルでの応答であるのか，意識レベルでの応答であるのかは実験結果により判別できるが，それ以前に被験者に摂動が知覚されれば，当然発話動作での応答は確認できない．つまり，観測すべき応答が摂動に対する随意運動のような意識レベルによる応答であるか，摂動がなくても反応が起こる自律レベルによる応答であるのか区別するためには，被験者が摂動を検知する必要がある．さらに検知感度などについても考察する必要がある．例えばフォルマントの変化に対する検知感度は基本周波数のものほど高くなく [21]，観測対象とする調音器官の調音活動を充分考慮した摂動を与える必要がある．

3.4 実験アプローチ

上記の実験要件を考慮し，以下のような実験アプローチを満たす計測システムの構成が必要となる．

まず，実験アプローチの一つ目としてフィードバック経路の存在を示すにはフィードバック経路を遮断するか外乱を与えることで発話動作に現れる変化を観測することが可能である．そこで本研究では，フィードバック経路に外乱を与えることにより，調音器官の活動を含めて発話動作に現れる変化の観測を行う．また，非破壊的な実験パラダイムを用いて定量的な実験を実現させるために，実時間にフォルマント変形聴覚フィードバック実

験を行う。

さらに、本研究では、観測対象として追加する調音情報を取得するため、田中 [13] の実験システムに磁気センサシステム (Electro-magnetic-Articulography : EMA) を追加する。

また、松岡 [12] による実験では、中国母音の一部から調音器官について考察が行われているが、日本母音については田中 [13] までの先行研究では調音器官についての考察が行われていない。そこで、本研究では先行研究で用いられた日本母音を音声資料として選択を行った。

3.5 実験システム

本研究では、上記の実験要件と実験アプローチを踏まえ、先行研究の用いたシステム [13] を基本にしてシステムの構築を行った。実験システムの概要図を図 3.1 に示す。実験は防音室内で行う。被験者により発話された音声はリアルタイム OS である RT-Linux 上のプログラムにより実時間で変換が行われ、被験者にフィードバックされる。骨導音、自然側音等をマスクするため、フィードバック音声には 60dB 程度のピンクノイズを付加する。

発話音声、及び変換された聴覚フィードバック音声はそれぞれ記録用の計算機で記憶され、分析に用いる。被験者はヘッドフォン (ATH-PRO700) を身に着けた状態で、発話された音声をマイクロホン (EGM-G5M)、マイクロホンアンプ (MA-8) を経て、一方は計算機内の AD 変換ボード (PCI-3155) に入力され、もう一方は記録用の計算機に接続された AD 変換機 (DF-2021) を経て記録用の計算機に記録される。計算機内の AD 変換ボードに入力された音声は実時間処理によりフォルマントの変換処理が行われ、DA 変換ボード (PCI-3336) を通じて外部に出力される。

DA 変換ボードから出力された音声は一方はミキサ (AT-MX50) によりピンクノイズが付加され、防音室内のアンプ (AT- α 907MR) を経て被験者にフィードバックされる。もう一方は、AD 変換機を経て記録用の計算機に記録される。

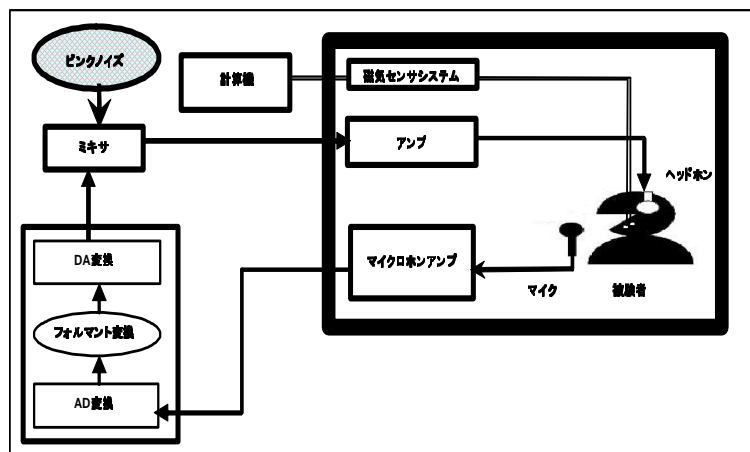


図 3.1: 実験システム構成図

第4章 聴覚フィードバックにおける非線形な摂動による補償動作の測定

4.1 聴覚フィードバック実験

4.1.1 実験目的

田中が行った研究では、被験者に与える摂動の変化量を図 4.1 のように第 1 フォルマント (F1) と第 2 フォルマント (F2) を対称に、それぞれ同じ割合 (0, 20, 40, 60, 70, 80, 100 % の 7 通り) で行っていた [13]。本研究では、人間がどのエリアの音声からフィードバックし、制御するための音声と認識するかを確認するために、日本語の中性母音 /e/ から /a/ へ非対称な摂動を与え、被験者の反応を調べた。図 4.2 のような非対称な摂動を加えてから F1, F2 が共に対称な 100 % の摂動の方向へ近づけるといふ実験を行い、それに対する人間の発話音声を測定し、どのような傾向が見られるのか確認した。

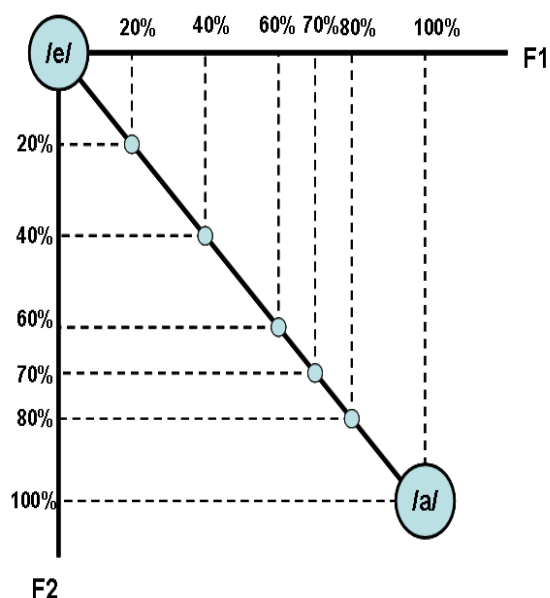


図 4.1: 以前の摂動の与え方

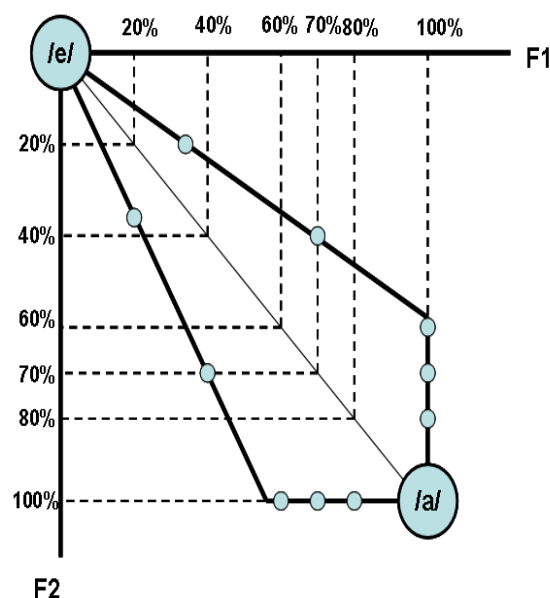


図 4.2: 本研究での摂動の与え方

4.1.2 実験手順

被験者に日本語の中性母音である/e/を発話するように指示し、発話途中で被験者の母音/e/のF1とF2を/e/から/a/へ非対称な摂動をフィードバックした。それを実現するために、被験者の音声/e/及び/a/を事前に分析してF1とF2を求めた。被験者の音響パラメータを表4.1に示す。

変形された音声ヘッドフォンを通じて被験者に提示される。摂動ありの場合、摂動を与えるまで被験者の音声をフォルマント変形なしでフィードバックし、その後2秒間でフォルマントの変形による”摂動あり”の音声をフィードバックした。その後再び”摂動なし”の音声を呈示した。2種類の異なる摂動開始時間のパターンと摂動なしパターンのうち一つをランダムに被験者に発話途中に呈示した。トライアルの構成を図4.3に示す。

1回の発話を1トライアルとし、1セットは3トライアルからなり、10セット連続して実験を行った。

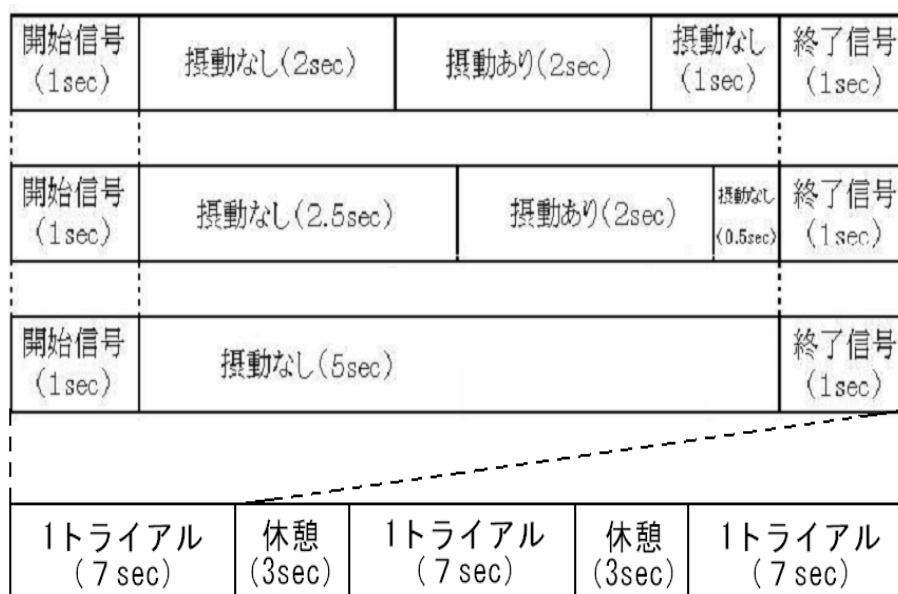


図 4.3: トライアル毎の構成

4.1.3 分析方法

サンプリング周波数 $16kHz$ で収録した音声に関して、フォルマント周波数は音響解析ソフトウェア WaveSurfer を用いてフレーム長 $49ms$ 、フレームシフト $10ms$ で算出した。摂動を与える前の被験者の発話した (F_{11}, F_{12}) 始点、摂動を与えている最中に被験者が発話した (F_{21}, F_{22}) を終点としている。松岡の研究の知見 [12] より、摂動開始から $290ms$ で補償が最大となることが示されており、本実験の摂動中のフォルマントはこれを考慮して

表 4.1: 被験者の音響パラメータ

	変形量	F1[Hz]	F2[Hz]
/e/	0%	491	1872
	20%	534	1715
	40%	576	1557
	60%	619	1399
	70%	640	1320
	80%	661	1242
/u/	100%	704	1084

抽出している．また，被験者の/e/の $(F1_e, F2_e)$ から/a/の $(F1_a, F2_a)$ への応答ベクトル $(F1_e, F2_e) - (F1_a, F2_a)$ との相関が 0.3 以上であるものを補償動作があると判断した．

4.1.4 実験結果

被験者に F1 と F2 が非対称な摂動をフィードバック音声として呈示した際の発話音声の音響分析の結果を報告する．

図 4.4，図 4.5 では，F1 と F2 が非対称な ($\pm 15^\circ$ ずらした) 摂動を与え，F1，F2 が共に 100 % の方向へ近づけたときの音響分析の結果である．F1 と F2 が非対称な摂動を与えたとき，補償動作を確認することはできなかった．また，4.6 はその中で補償動作が現れた場合の音響分析の結果である．図 4.6 より， 15° ずらした摂動では補償は現れず， $(F1, F2) = (60\%, 100\%)$ のとき 5 %， $(F1, F2) = (70\%, 100\%)$ のとき 15 %， $(F1, F2) = (80\%, 100\%)$ のとき 20 % と F1，F2 をの変化量が共に 100 % の方向へ近づくに連れてわずかにではあるが補償動作が確認されている．

4.2 考察

図 4.4 ~ 図 4.6 より，F1 と F2 が対称な摂動の場合より $\pm 15^\circ$ ずらし，F1，F2 が非対称な摂動を与えた場合では補償動作が確認されず，F1，F2 が共に 100 % の方向へ近づくに近づかずかであるが補償動作が増えていくことを確認した．

本実験で用いた，F1，F2 が非対称な摂動の範囲では，人間は摂動をフィードバックする音声だと認識していないことが推測される．また，F1 と F2 の変化量を同じ割合とする摂動の場合より -15° ずらし，F1，F2 が共に 100 % の方向へ近づける場合 F1 は 100 % で F2 のみが 60，70，80 % と変化していく．このとき，補償動作は確認されていない． 15° ずらし，F1，F2 が共に 100 % の方向へ近づける場合，F2 が 100 % で F1 のみが 60，70，80 % と変化していく．これを図 4.7 に示す．このとき，補償動作の起こる確率がわずかに増えている．これは，F1 は主に下顎や口唇が動きやすいので比較的補償動作が起こりや

すく、F2は舌の前後運動に深く関連するため、生理学的な拘束により補償動作が起こりにくくなるという田中の研究の知見 [13] から、被験者は F1 が一定で F2 のみの変化では補償動作が確認されず、F2 が一定で F1 のみの変化では補償動作が確認されたという結果が得られたと考えられる。

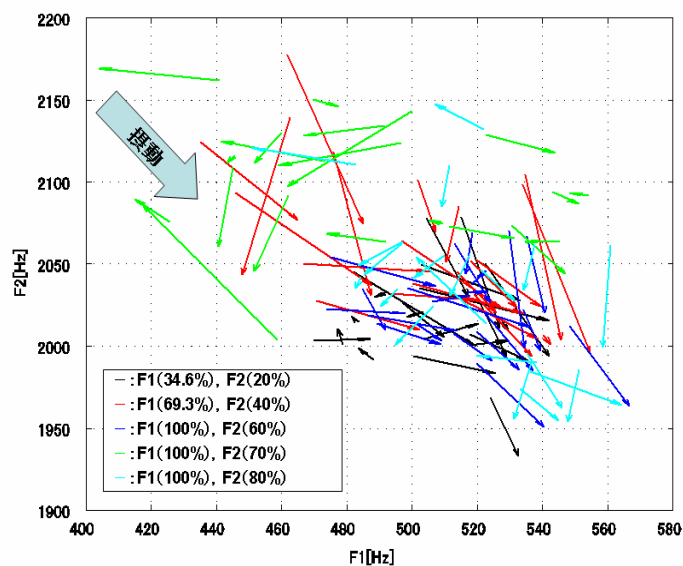


図 4.4: F1 と F2 で異なる割合の摂動を与えた場合の反応 (-15°)

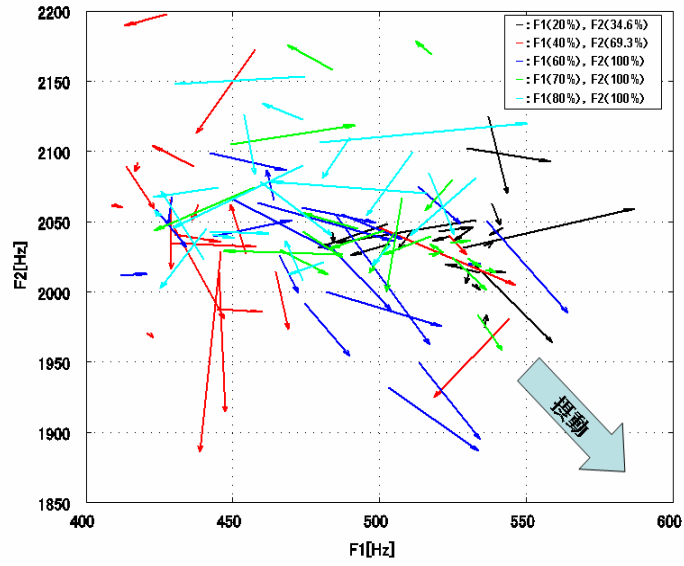


図 4.5: F1 と F2 で異なる割合の摂動を与えた場合の反応 (15°)

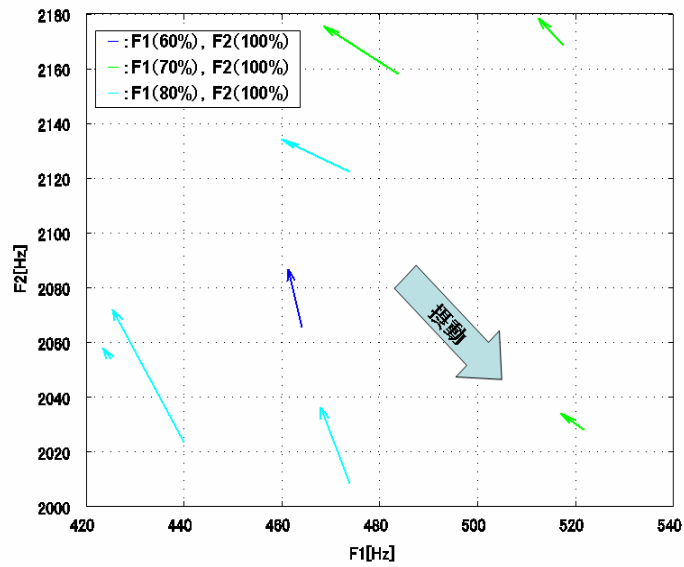


図 4.6: 補償動作ありの場合 (15°)

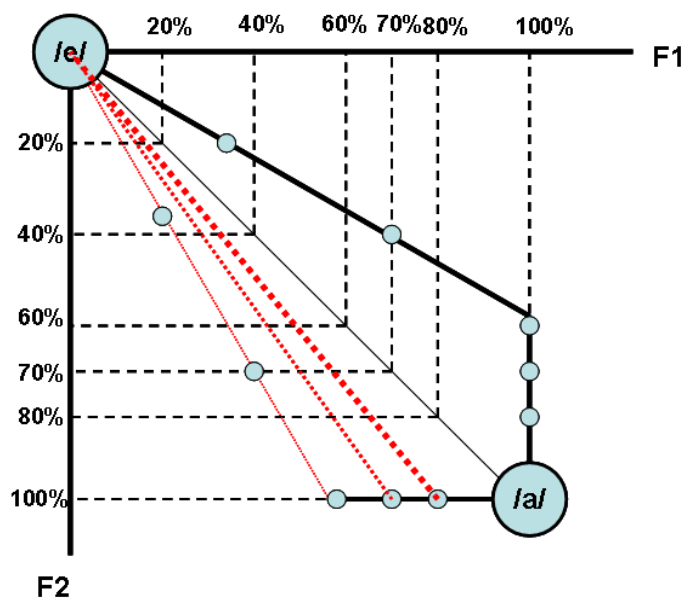


図 4.7: 人間がフィードバックする音声を認識する範囲 (赤線で太字ほど認識する)

第5章 調音器官の補償動作の測定

5.1 変形聴覚フィードバック実験

5.1.1 実験目的

松岡の行った実験では，調音器官の補償について口唇形状のみを考察し，用いた音素は中国母音/i/から/ü/であった．この実験から音響分析，口唇の動き共に補償動作が確認されている [12]．また，田中の行った実験から日本母音/e/から/a/，/i/，/u/においても音響分析では補償が確認された [13]．しかし，音響分析の結果から補償の起こる割合は60～70%であり，調音器官の動きにより打ち消されている可能性が示されている．そこで本実験では，日本母音について舌・口唇のような調音器官における補償動作について測定を行った．

5.1.2 磁気センサシステム

ここで，本実験で舌・口唇のような調音器官を測定するために用いた磁気センサシステム (Electro-magnetic-Articulography: EMA) について説明する．磁気センサシステムとは，発話器官上の離散的なポイントに磁気センサを貼り付けてその運動を観測する装置である．今回用いた磁気センサシステムは Casterns 社製の AG500 である．このシステムは x, y, z の3次元的位置にセンサの前後と左右の傾斜を合わせた5次元の測定が可能なシステムである [25]．この磁気センサシステムを図 5.1 に示す．

5.1.3 実験手順

実験の手順は 4.1.2 の図 4.1 と同様に行った．被験者の舌・口唇に磁気センサを装着し，に日本語の中性母音である/e/を発話するように指示し，発話途中で被験者の母音/e/の第1フォルマント (F1) と第2フォルマント (F2) を/a/，/i/，/u/の方向へ0, 20, 40, 60, 70, 80, 100%の割合でそれぞれ変動してフィードバックした．摂動量が0%である場合は/e/，100%である場合はフィードバック音声は完全に/a/，/i/，/u/となる．それを実現するために，被験者の発話音声/a/，/i/，/u/，/e/を事前に分析して F1 と F2 を求めた．/e/から/u/への変形を一例として表 5.1 に示す．図 4.3 と同様に1回の発話を1トライアルとし，1セットは3トライアルからなる．各摂動量 (フォルマント変形) に関して

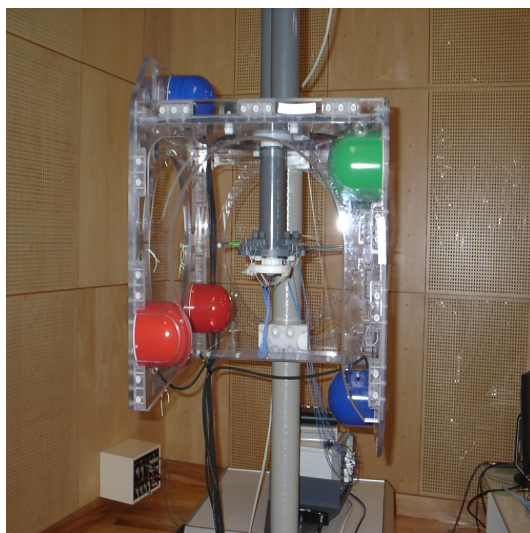


図 5.1: 磁気センサシステム

10 セット連続して測定し，音声収録はサンプリング周波数 $16kHz$ で行った．

本実験では被験者に図 5.2 のように磁気センサを装着した．図 5.2 を正中断面から見ると，磁気センサの位置は図 5.4 のようになる．また，実験風景を図 5.3 に示す．磁気センサシステムによる調音運動はサンプリング周波数 $200Hz$ で行った。

表 5.1: /e/ /u/ のフォルマント変換量

	変形量	F1[Hz]	F2[Hz]
/e/	0%	533	2059
	20%	499	1961
	40%	466	1863
	60%	432	1766
	70%	416	1717
	80%	399	1668
/u/	100%	365	1570

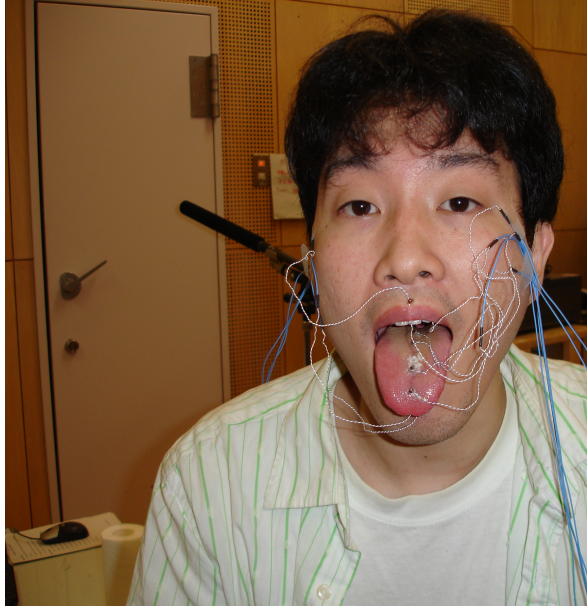


図 5.2: 磁気センサ装着図



図 5.3: 実験風景

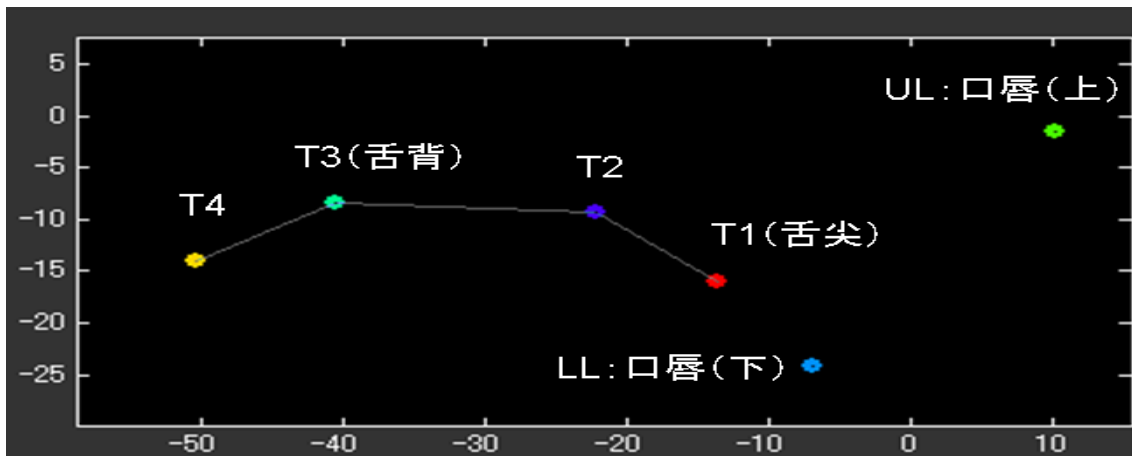


図 5.4: 磁気センサの位置

5.1.4 分析方法

本実験では、収録した音声による音響分析と磁気センサシステムで測定した調音運動について分析を行った。音響分析については、4.1.3と同様に補償動作を示すベクトルとの相関が0.3以上のものを補償動作とする。また、調音運動の分析に関しては主成分分析を行った。しかし、/a/の摂動を与えた場合の測定では、正常な測定結果である図5.5に比べ、図5.6のような磁気センサの一部が舌の動ける範囲とは考えられない動きを見せる正常とはいえない結果が得られなかったため、/a/の測定結果を除く/i/、/u/を摂動として与えた場合の測定結果を分析に用いた。なお、主成分分析については次に述べる。

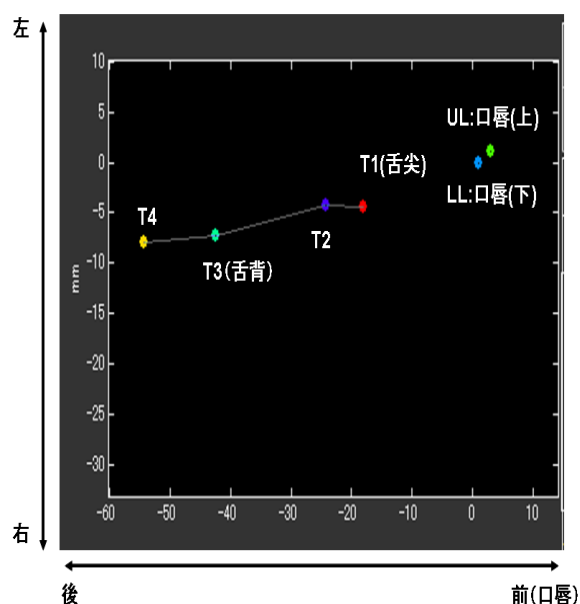


図 5.5: 正常な測定結果

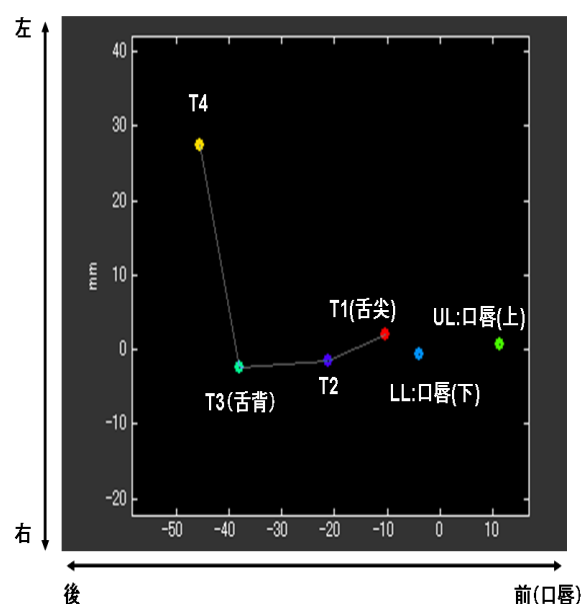


図 5.6: 異常な測定結果

5.1.5 主成分分析

本実験の調音器官の動きの分析に用いた主成分分析について説明する。主成分分析 (PCA:Principal Component Analysis) とは、ある対象について観測された多数の項目 ($x_1, x_2, x_3, \dots, x_p$) を個々に分析するのではなく、それらの項目全体が織り成す意味合いを解釈するために使用する手法である。項目を個々に分析する方法は一変量解析と呼ばれ、個々の項目ごとに平均や分散を求め各項目の持つ意味を解釈するが、主成分分析は多数の項目が全体として持っている意味合いを解釈するために使用するもので多変量解析の一方法論である。

主成分分析とは、 P 個の変数 $\{x_p\} (p = 1, 2, \dots, P)$ の持つ情報を、情報の損失を最大限に抑えながら、 $\{x_p\}$ の一次結合として与えられる互いに独立な $M (M \leq P)$ 個の主成分

(総合的指標) $\{z_m\}$

$$z_m = \sum_{p=1}^P w_{pm} x_p \quad (m = 1, 2, \dots, M) \quad (5.1)$$

を用いて表現する手法である。なお、 z_m は第 m 主成分と呼ばれ、その結合係数 $\{w_{pm}\} (p = 1, 2, \dots, P; m = 1, 2, \dots, M)$ は以下の条件を満足するように決定される。

<条件>

第1主成分 z_1 の分散は $\{x_p\} (p = 1, 2, \dots, P)$ のあらゆる1次式の持つ分散の中で最大であり、第 m 主成分 $\{z_m\} (m = 2, \dots, M)$ の分散は $\{z_{m'}\} (m' = 1, 2, \dots, m-1)$ の全てと無相関な1次式の持つ分散の中で最大である。ただし、

$$\sum_{p=1}^P w_{pm}^2 = 1 \quad (m = 1, 2, \dots, M) \quad (5.2)$$

とする。

この条件に従い、主成分の分散が最大となるような主成分を決定する方法について述べる。いま、 P 個の変数について N 個のサンプルがある場合を考え、測定値を $\{x_{np}^*\} (n = 1, 2, \dots, N; p = 1, 2, \dots, P)$ とする。各変数についてその平均値 $\{\bar{x}_p\} (p = 1, 2, \dots, P)$ からの偏差 $\{x_{np}\}$ を導入する(今回は簡単のため各変数についてその平均値を0と定義する)。すなわち、

$$x_{np} = x_{np}^* - \bar{x}_p \quad (n = 1, 2, \dots, N; p = 1, 2, \dots, P) \quad (5.3)$$

とする。このとき、測定データ全体は次の行列 X で与えられる。

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1P} \\ x_{21} & x_{22} & \cdots & x_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NP} \end{bmatrix} \quad (5.4)$$

第1主成分 z_1 の導出は(5.1)式で与えられるので、その結合係数を

$$\mathbf{w}_1 = \begin{bmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{P1} \end{bmatrix} \quad (5.5)$$

とすると、 n 番目のサンプル

$$\mathbf{x}_n = (x_{n1} \quad x_{n2} \quad \cdots \quad x_{nP}) \quad (5.6)$$

に対応する第1主成分 z_1 の値 t_{n1} は

$$\begin{aligned} t_1 &= \sum_{p=1}^P w_{p1} x_{np} \\ &= \mathbf{x}_n \mathbf{w}_1 \end{aligned} \quad (5.7)$$

となる．この第1主成分 z_1 の値 t_{n1} を第1主成分得点と呼ぶ． N 個のサンプルに対応する第1主成分得点を1つのベクトルにまとめ、

$$\mathbf{t}_1 = \begin{bmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{N1} \end{bmatrix} \quad (5.8)$$

とおくと、

$$\mathbf{t}_1 = \mathbf{X} \mathbf{w}_1 \quad (5.9)$$

が成り立つ．今回の定義から、第1主成分得点の平均値 \bar{t}_1 は

$$\begin{aligned} \bar{t}_1 &= \frac{1}{N} \sum_{n=1}^N t_{n1} \\ &= \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \mathbf{w}_1 \\ &= \frac{1}{N} \sum_{n=1}^N \left(\sum_{p=1}^P w_{p1} x_{np} \right) \\ &= \frac{1}{N} \sum_{p=1}^P w_{p1} \left(\sum_{n=1}^N x_{np} \right) \\ &= 0 \end{aligned} \quad (5.10)$$

したがって、第1主成分 z_1 の分散 $\sigma_{z_1}^2$ は

$$\begin{aligned} \sigma_{z_1}^2 &= \frac{1}{N-1} \mathbf{t}_1^T \mathbf{t}_1 \\ &= \frac{1}{N-1} (\mathbf{X} \mathbf{w}_1)^T (\mathbf{X} \mathbf{w}_1) \\ &= \mathbf{w}_1^T \mathbf{V} \mathbf{w}_1 \\ &\geq 0 \end{aligned} \quad (5.11)$$

となる．なお、行列 \mathbf{V} は共分散行列と呼ばれる非負定値行列であり、

$$\mathbf{V} = \frac{1}{N-1} \mathbf{X}^T \mathbf{X} \quad (5.12)$$

で与えられ，その要素 v_{ij} は

$$v_{ij} = \frac{1}{N-1} \sum_{n=1}^N x_{ni} x_{nj} \quad (5.13)$$

$$= \frac{1}{N-1} \sum_{n=1}^N (x_{ni}^* - \bar{x}_i) (x_{nj}^* - \bar{x}_j) \quad (5.14)$$

である．また， $v_{ij} = v_{ji}$ すなわち $\mathbf{V} = \mathbf{V}^T$ が成り立つ．

また〈条件〉より，第1主成分 z_1 は(5.2)式の下でその分散 $\sigma_{z_1}^2$ が最大となるように決定されなければならない．この最適化問題は Lagrange の未定乗数法より，Lagrange 乗数 λ を導入して

$$J_1 = \mathbf{w}_1^T \mathbf{V} \mathbf{w}_1 - \lambda (\mathbf{w}_1^T \mathbf{w}_1 - 1) \quad (5.15)$$

とおき， J_1 を最大にするような結合係数 \mathbf{w}_1 を求めればよい．そこで， J_1 を \mathbf{w}_1 で偏微分して0とおくと，

$$\begin{aligned} \frac{\partial J_1}{\partial \mathbf{w}_1} &= \begin{bmatrix} \frac{\partial J_1}{\partial w_{11}} \\ \frac{\partial J_1}{\partial w_{21}} \\ \vdots \\ \frac{\partial J_1}{\partial w_{P1}} \end{bmatrix} \\ &= \begin{bmatrix} 2 \sum_{p=1}^P v_{1p} w_{p1} \\ 2 \sum_{p=1}^P v_{2p} w_{p1} \\ \vdots \\ 2 \sum_{p=1}^P v_{Pp} w_{p1} \end{bmatrix} - 2\lambda \begin{bmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{P1} \end{bmatrix} \\ &= 2\mathbf{V}\mathbf{w}_1 - 2\lambda\mathbf{w}_1 \\ &= 0 \end{aligned} \quad (5.16)$$

となることから

$$(\mathbf{V} - \lambda\mathbf{I})\mathbf{w}_1 = 0 \quad (5.17)$$

という条件式が得られる．これは固有値問題に帰着することがわかる． λ が満たすべき条件は固有方程式

$$\det |\mathbf{V} - \lambda\mathbf{I}| = 0 \quad (5.18)$$

以上より，分散の最大値を与える主成分 \mathbf{w}_1 は共分散行列 \mathbf{V} の固有ベクトルの一つであり， λ はその固有ベクトルに対応した固有値であることがわかる．

ここで，を満たす固有ベクトル \mathbf{w} ，固有値 λ はそれぞれ， P 個あるが，そのうち，どれが分散の最大値を与えるかを考える．第1主成分 z_1 の分散 $\sigma_{z_1}^2$ は(5.11)式で与えられ， \mathbf{w}_1 は(5.17)式を条件として満足しなければならない．そこで，(5.11)式に(5.17)式を代入し， $\mathbf{w}_1^T \mathbf{w}_1 = 1$ に注意すれば，

$$\sigma_{z_1}^2 = \mathbf{w}_1^T \mathbf{V} \mathbf{w}_1$$

$$\begin{aligned}
&= \mathbf{w}_1^T \lambda \mathbf{w}_1 \\
&= \lambda
\end{aligned}
\tag{5.19}$$

であるから，分散の値は固有値に等しいことがわかる．よって，求める主成分は最大固有値に対応した固有ベクトルとして与えられることがわかる．第2主成分以下の結合係数 $\{\mathbf{w}_m\} (m = 2, 3, \dots, M)$ も帰納法を用いて第1主成分と同様な手順で求めることができる．

また，主成分分析とは少数の総合的指標（主成分）を用いて変数間の関係や特徴を把握するための統計的手法である．従って，各主成分が元のデータに含まれる特徴をどの程度表現しているのか，あるいは何個の主成分を採用すれば元のデータに含まれる特徴を十分に表現できるのかを知ることが必要となる．このための指標として，寄与率および累積寄与率がある．共分散行列 \mathbf{V} の (p, p) 要素 v_{pp} が変数 x_p の分散に等しくなるため，

$$\begin{aligned}
\sum_{p=1}^P \sigma_{x_p}^2 &= \sum_{p=1}^P v_{pp} \\
&= \text{tr}(\mathbf{V})
\end{aligned}
\tag{5.20}$$

で与えられる．一方，第 m 主成分の分散 $\sigma_{z_m}^2$ は共分散行列 \mathbf{V} の m 番目に大きな固有値 λ_m に等しいので，合計 P 個の主成分の分散の和は

$$\begin{aligned}
\sum_{p=1}^P \sigma_{z_p}^2 &= \sum_{p=1}^P \lambda_p \\
&= \text{tr}(\mathbf{V})
\end{aligned}
\tag{5.21}$$

で与えられる．すなわち，変数の分散の総和と主成分の分散の総和とは等しくなる．そこで第 m 主成分が元のデータに含まれる特徴をどの程度表現しているかを示す指標として，第 m 主成分の分散が分散の総和に占める割合

$$C_m = \frac{\lambda_m}{\sum_{p=1}^P \lambda_p} = \frac{\lambda_m}{\text{tr}(\mathbf{V})}
\tag{5.22}$$

を利用することができる． C_m は寄与率（proportion）と呼ばれる．また，第 m 主成分までの分散の和が分散の総和に占める割合

$$P_m = \sum_{i=1}^m C_i = \frac{\sum_{i=1}^m \lambda_i}{\text{tr}(\mathbf{V})}
\tag{5.23}$$

は累積寄与率（accumulated proportion）と呼ばれる．累積寄与率を用いて採用する主成分数を決定する場合，通常累積寄与率が80%となるように主成分数が決められる．

5.1.6 実験結果

まず，/e/から/i/へ摂動を与えた際の音響分析の結果を図5.7に示す．その中から，補償動作があると判断したものを図5.8に示す．補償動作の起こる確率は約31.7%となつて

いる。また、/e/から/u/へ摂動を与えた場合には、図 5.9 より補償動作が約 61.7% の確率で起こっている。/e/から/i/の場合では被験者が摂動につられてしまう場合が多く、/e/から/u/の場合では、田中の行った研究と同等の割合（60～70%）で補償動作を確認できた。

また、磁気センサシステムの測定結果から、各磁気センサの x 方向（前後）、z 方向（上下）における主成分分析による各主成分の固有ベクトルを図 5.11、図 5.12、図 5.13 に示す。これらの T1～LL は図 5.4 の装着した磁気センサの位置に対応している。この 3 つの図より、第 1 主成分は舌後部の上下の動き、第 2 主成分は下唇の前後、舌尖の上下、舌背の上下の動き、第 3 主成分は舌前部の前後の動きの指標となっている。更に、図 5.14、図 5.15、図 5.16 に/e/から/i/へ、図 5.17、図 5.18、図 5.19 に/e/から/u/へ摂動を与えたときの各主成分の関係を示す。ここで、図 5.14～図 5.19 のベクトルは全トライアルデータに対して求めた各主成分の係数を用いて、各トライアルの摂動前の平均係数が始点、摂動中の平均係数が終点となっている。3 つの図において各主成分が示す正の方向は、第 1 主成分が舌後部が上方へ、第 2 主成分は下唇が前方、舌尖が上方、舌背が下方へ、第 3 主成分は舌前部が前方へ動いていることを示す。よって、摂動が/i/である場合、補償動作は第 1 主成分が負、第 2 主成分が正、第 3 主成分が負の方向へそれぞれ動いた時となり、摂動が/u/である場合、補償動作は第 1 主成分が負、第 2 主成分が正、第 3 主成分が正の方向へそれぞれ動いた時となる。また、補償動作は各主成分の補償動作の方向を示す単位ベクトルから相関が 0.3 以上になるものとした。その結果、/i/へ摂動を与えた場合には、図 5.14 では全トライアル中の約 14.7%、図 5.15 では約 20.2%、図 5.16 では約 20.2% となっている。/u/へ摂動を与えた場合には、図 5.17 では全トライアル中の約 46.4%、図 5.18 では約 58.8%、図 5.19 では約 54.6% の割合で現れた。また、この 3 つの場合のいずれかで補償が生じる割合は約 79.4% となった。それに対して、/i/の摂動に対する補正動作が生じる割合は約 44.0% となった。以上の結果をまとめたものを表 5.2 に示す。

表 5.2: 音響分析と主成分分析の結果

摂動	/i/	/u/
音響分析で補償動作の起こる確率	31.7%	61.7%
第 1, 第 2 主成分で補償動作の起こる確率	14.7%	46.4%
第 1, 第 3 主成分で補償動作の起こる確率	20.2%	58.8%
第 2, 第 3 主成分で補償動作の起こる確率	20.2%	54.6%
3 つのいずれかで補償動作の起こる確率	44.0%	79.4%

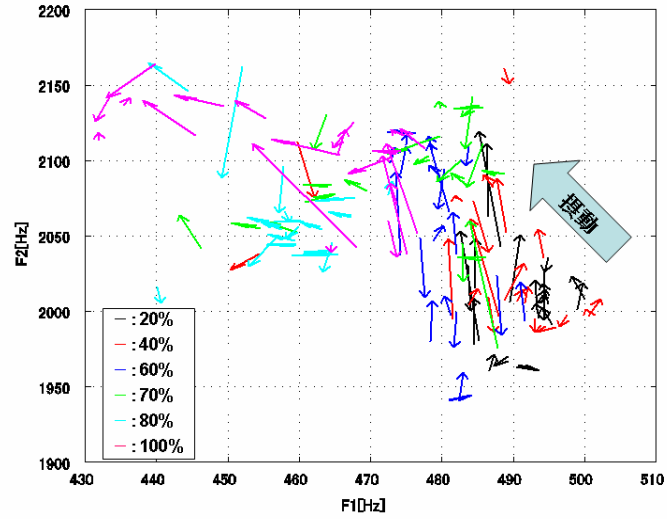


図 5.7: 擾動を与えた場合の反応 (/e/ /i/)

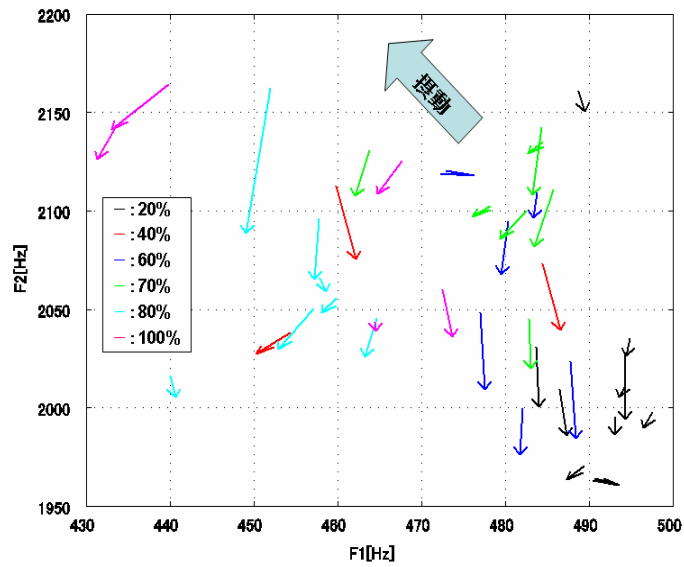


図 5.8: 補償動作ありの場合 (/e/ /i/)

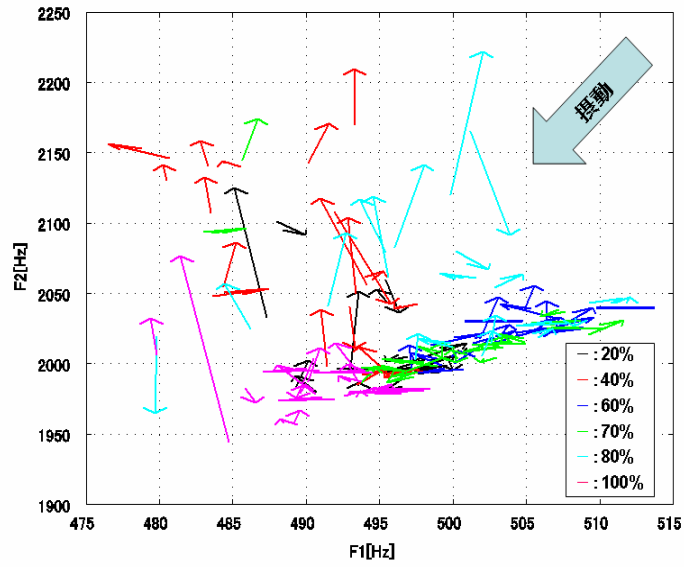


図 5.9: 摂動を与えた場合の反応 (/e/ /u/)

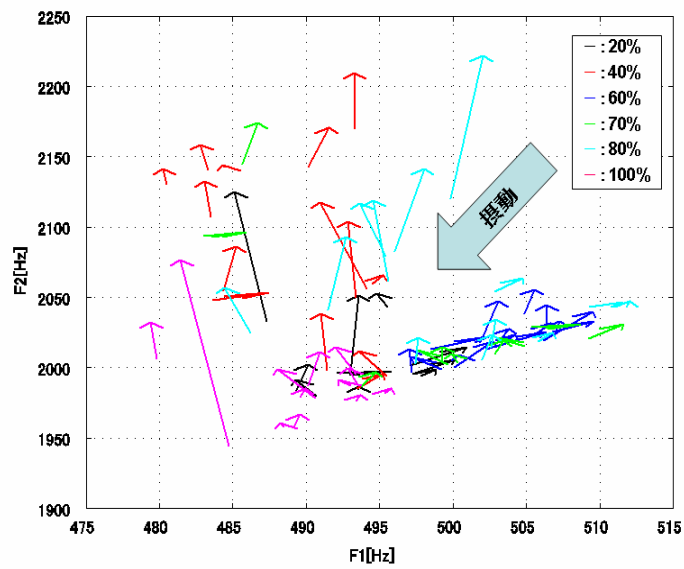


図 5.10: 補償動作ありの場合 (/e/ /u/)

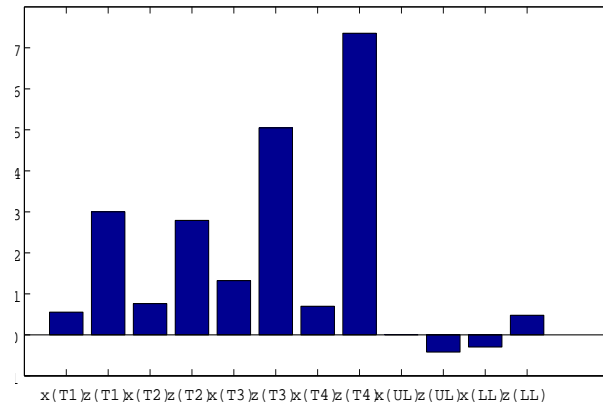


図 5.11: 第 1 主成分の固有ベクトル

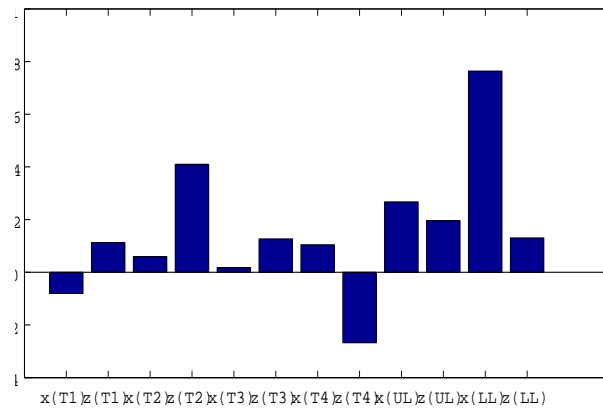


図 5.12: 第 2 主成分の固有ベクトル

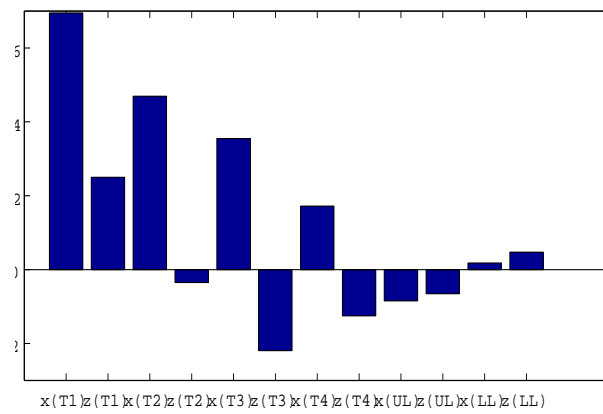


図 5.13: 第 3 主成分の固有ベクトル

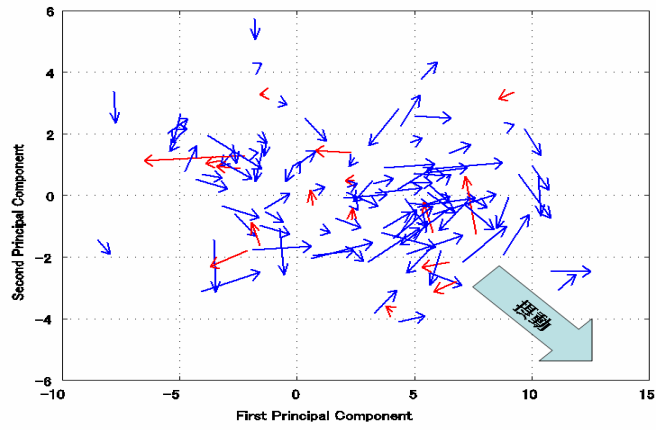


図 5.14: /i/の第1, 第2主成分の関係 (赤: 補償あり、青: 補償なし)

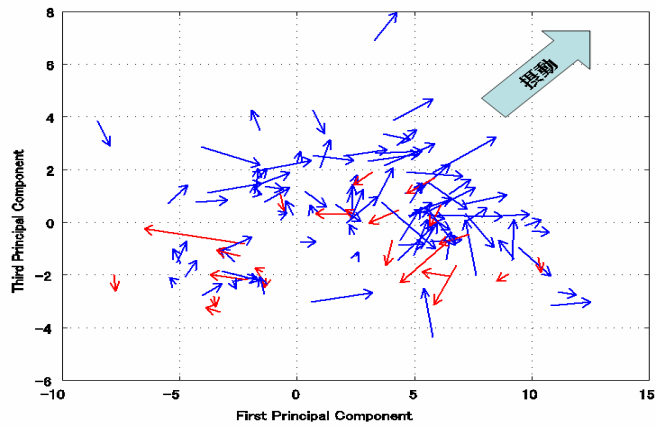


図 5.15: /i/の第1, 第3主成分の関係 (赤: 補償あり、青: 補償なし)

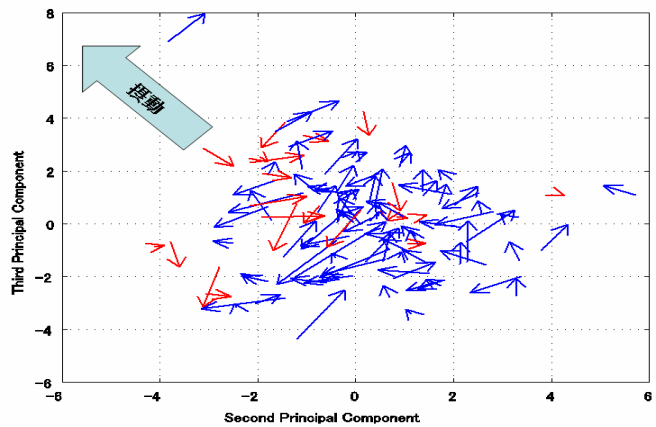


図 5.16: /i/の第2, 第3主成分の関係 (赤: 補償あり、青: 補償なし)

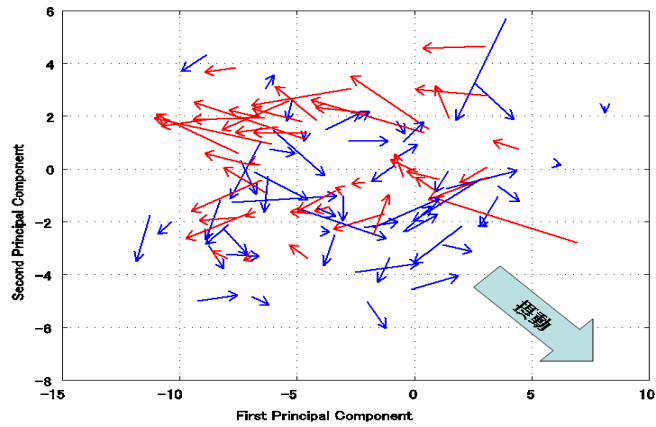


図 5.17: /u/の第1, 第2主成分の関係 (赤: 補償あり、青: 補償なし)

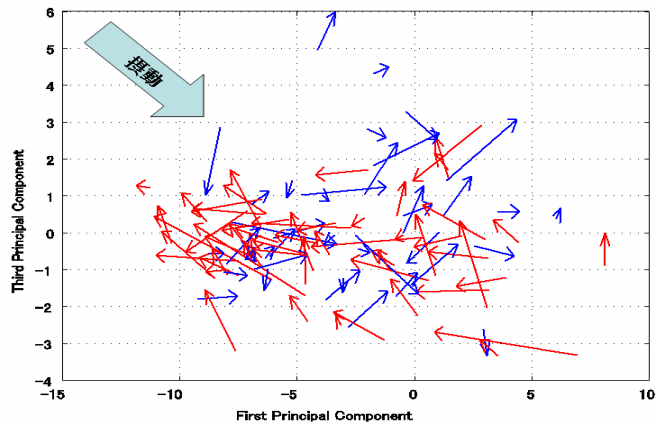


図 5.18: /u/の第1, 第3主成分の関係 (赤: 補償あり、青: 補償なし)

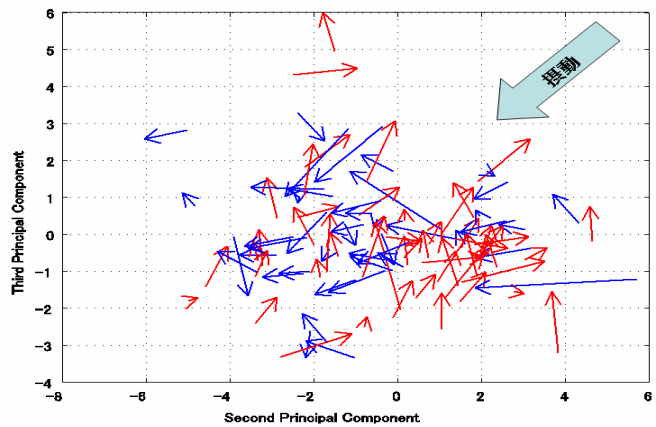


図 5.19: /u/の第2, 第3主成分の関係 (赤: 補償あり、青: 補償なし)

5.2 考察

一般的に図 5.20 のように/e/は/i/に比べ舌の調音位置が後方かつ下方に移動する。また、/e/は/u/に比べ舌の調音位置が前方かつ下方に移動する。従って、/e/の発話時に/i/の方向へ摂動を与えた時、もし舌の後方かつ下方への動きが、同様に/u/の方向へ摂動を与えたとき、舌の前方かつ下方への動きが見られたときにそれぞれ補償動作があったと判断した。

今回、主成分分析を用いて各摂動に対する舌全体の補償動作について調査を行った。図 5.19~5.19 より、各主成分の関係から、/u/へ摂動を与えた際の補償動作を確認することが出来た。この補償動作は第 1、第 3 主成分での動きが大きく、第 2 主成分についての動きが小さいことから舌の上下前後の変化が大きく、口唇形状の変化が小さいことがわかる。よって/u/へ摂動を与えた場合には舌に補償動作が起きていることが示唆できる。また、舌の下方かつ前方、下方のみ、前方のみのいずれかの補償動作が起きている割合は約 79.4 % となり、先行研究で行われた音響分析の約 60~70 % に比べ高い割合で生じた。これは音声を生成する際に、聴覚系が常にモニタリングをしていることを示唆している。/i/へ摂動を与えた場合に、/u/と比較して、補償動作の割合が低くなっているが、今回の実験結果では、補償動作と逆方向の反応が目立ち、被験者が摂動/i/に対してつられている可能性が考えられる。しかし、舌の下方かつ後方、下方のみ、後方のみいずれかの補償動作が起きている割合は約 44.0 % とチャンスレベルであり、/u/と同様に音響分析の結果に比べ高い割合となっており、今後の課題として被験者を増やし、一般的な傾向を見ていく必要がある。

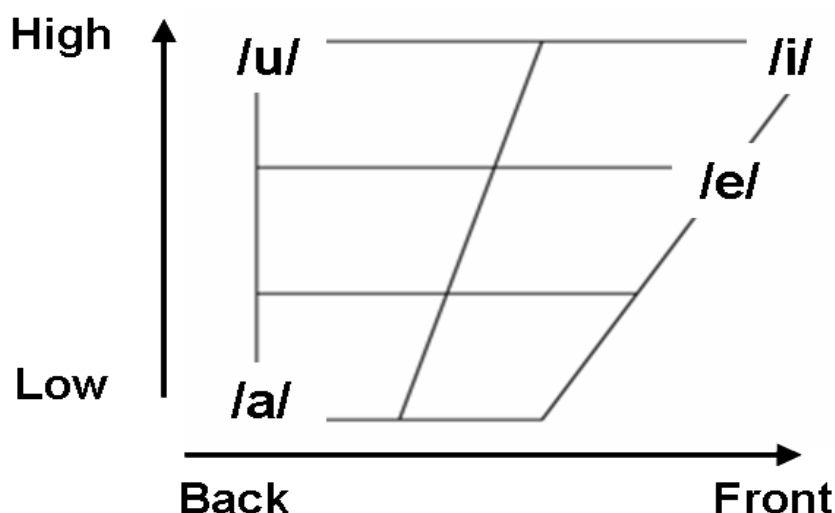


図 5.20: 日本母音の調音位置

第6章 まとめ

6.1 結論

本研究では、日本語の中性母音/e/の発話途中にF1とF2を変形し、摂動として与える実験を行った。/a/の方向へF1とF2を異なる割合の変化量で摂動を与える実験を行い、補償動作は起こりにくいことがわかった。また、F1とF2の変化量が共に100%に近づくにつれ、わずかであるが、補償動作の起こる割合が高くなっていることも確認できた。この結果、人間は摂動を与える際に摂動に対してF1とF2が同じ割合で変形された摂動に対して自分の発話を制御しようとする補償動作が起こるのではないかと考える。

また、磁気センサシステムを用いて、舌・口唇のような調音器官の反応を計測した。この位置情報から主成分分析を行った結果、摂動/i/に対しては、被験者が摂動につらわれている可能性が高い事が音響分析、磁気センサシステムの結果から確認された。しかし、調音器官の補償に関しては、舌の下方かつ後方、下方のみ、後方のみいずれかの補償動作が起こっている割合は約44.0%とチャンスレベルであった。摂動/u/に対しては、図5.17～図5.19各主成分の関係から、/u/へ摂動を与えた際の補償動作を確認することが出来た。この補償動作は第1、第3主成分での動きが大きく、第2主成分についての動きが小さいことから舌の上下前後の変化が大きく、口唇形状の変化が小さいことがわかる。よって/u/へ摂動を与えた場合には舌に補償動作が起こっていることが示唆できる。音響分析では61.7%と先行研究と同程度(60～70%)の割合で補償動作が確認されたが、磁気センサシステムからは、舌の下方かつ前方、下方のみ、前方のみいずれかの補償動作が起こっている割合は約79.4%となり、音響分析に比べ高い割合で補償動作を確認できた。

これは音声を生成する際に、聴覚系が常にモニタリングをしていることを示唆している。/i/へ摂動を与えた場合に、/u/と比較して、補償動作の割合が低くなっているが、今回の実験結果では、補償動作と逆方向の反応が目立ち、被験者が摂動/i/に対してつらわれている可能性が考えられる。

6.2 今後の課題

/i/へ摂動を与えた場合に、/u/と比較して、補償動作の割合が低くなっているが、今回の実験結果では、補償動作と逆方向の反応が目立ち、被験者が摂動/i/に対してつらわれている可能性が考えられる。しかし、舌の下方かつ後方、下方のみ、後方のみいずれかの補償動作が起こっている割合は約44.0%とチャンスレベルであった。本実験から調音器

官に関する分析方法を確立することができたので、今後の課題として被験者を増やし、一般的な傾向を見ていく必要がある。また、実時間の反応を計測する際に、被験者へ摂動を与える時間（本研究では2秒間）など、実験パラダイムについても更に検討していく必要があると考えられる。

参考文献

- [1] Denes,P. and Pinson,E. "The Speech Chain, 2nd Ed.", New York : W.H Freeman and Co.(1993)
- [2] Liberman,A.M., Cooper, F.S., Shankweiler, D.P. and Studdert-Kennedy, M., "Perception of the speech code", Psych.Rev.,74(6),pp.853-870.(1967)
- [3] Liberman,A.M. and Mattingly, I.G., "The motor theory of speech perception revised", Cognition,21,pp.1-36.(1985)
- [4] Lombard,E., "Le signe de l'elevation de la voix", Annuals Maladies Oreille, Larynx, Nez, Pharynx,37,pp.101-119.(1911)
- [5] Lane, H. and Tranel, B., "The Lombard sign and the role of hearing in speech", Journal of speech and Hearing Research, 14, pp.677-709.(1971)
- [6] Lee, B.S., "Effect of Delayed speech feedback", Journal of the Acoustical Society of America, 22, pp.824-826.(1950)
- [7] Kawahara, H., "Transformed auditory feedback : Effects of fundamental frequency perturbation", Journal of the Acoustical Society of America, Vol.94, No3, Pt.2, p.1883.(1993)
- [8] Kawahara, H., "Interactions between speech production and perception under auditory feedback perturbations on fundamental frequencies", J.Acoust. Soc. Jap, Vol.15, pp.201-202(1993)
- [9] 河原., "音声知覚・生成相互作用の伝達特性について", 音響学会聴覚研究会資料, H-95-35, pp.223-226.(1995)
- [10] 佐藤., "スペクトル変型聴覚フィードバックによる音声生成・知覚の相互作用に関する研究", 北陸先端科学技術大学院大学修士論文.(2003)
- [11] 斉藤., "音声生成過程におけるフォルマント変換音声フィードバックの影響に関する研究", 北陸先端科学技術大学院大学修士論文.(2004)

- [12] 松岡., ”聴覚系と発話系の相互作用に関する研究”, 北陸先端科学技術大学院大学修士論文.(2005)
- [13] 田中., ”変形聴覚フィードバックの摂動量と補正動作の関係について”, 北陸先端科学技術大学院大学修士論文.(2006)
- [14] 柏野., ”「スピーチ・チェイン」と脳”, 信学技報, SP2006-73.(2006)
- [15] Levitt,H., Stromberg, H., Smith, C. and Gold, T., ”The structure of segmental errors in the speech of deaf children”, Dec;13(6),pp.419-441.(1980)
- [16] Osberger, M. and McGarr, N., ”Speech production characteristics of the hearing impaired”, Speech and Language: Advances in Basic Research and Practice, Vol.8,pp.221-283.(1982)
- [17] Smith, CR., ”Interjected sounds in deaf children’s speech”, Jun:8(2),pp.123-128.(1975)
- [18] Cowie, R. and Douglas-Cowie, E., ”Speech production in profound postlingual deafness.”, In M.Lutman, and M. Haggard(eds.), Hearing science and hearing disorders. London : Academic Press, pp.183-230.(1983)
- [19] Houde, J. and Jordan, M., ”Sensorimotor adaptation in speech production”, Science, 279, pp.1213-1216.(1998)
- [20] Houde, J. and Jordan, M., ”Sensorimotor adaptation of speech I: Compensation and Adaptation”, Journal of Speech Language, and Hearing Research, 45, pp.295-310.(2002)
- [21] 沢田, 寛, ”聴覚フィードバックに利用される音声情報の物理的特徴”, 日本音響学会聴覚研究会資料, Vol.33, No.2, H-2003-21 pp.117-122.(2003)
- [22] Shimon Sapir, Elizabeth Derosier, Andrea M. Simonson, and Amy Wohlert., ”Effects of frequency modulated tones and vowel formants on perioral muscle activity during isometric lip rounding”, Jornal of Voice and Hearing, Vol.4, No.2, pp152-158.(1990)
- [23] Lane, H. and Webster, J., ”Speech deterioration in postlingually deafened adults”, Journal of the Acoustical Society of America, 89, pp.859-866.(1991)
- [24] 甘利, 外山, ”脳科学大辞典”, 朝倉書店 (2000)
- [25] <http://www.articulograph.de/>
- [26] 加納, ”「主成分分析」”(2002)
<http://www-pse.cheme.kyoto-u.ac.jp/kano/document/text-PCA.pdf>

謝辞

本研究を進めるにあたり、多大なる御指導並びにご鞭撻賜りました、北陸先端科学技術大学院大学 情報科学研究科 党建武教授、赤木正人教授、徳田功准教授、鷓木祐史准教授、Lu Xugang 助教、並びに党研究室、徳田研究室、赤木研究室、鷓木研究室の皆様へ深く感謝致します。また、この1年間共に研究生活を過ごすことが出来た同期のメンバーへ深く感謝致します。

最後に、本研究を遂行するにあたり、多忙な中、過酷な実験の被験者となってくださった松井崇倫君に心より感謝致します。日頃の研究生活からあたたかく御指導いただきました私の周りの全ての関係者の方々にもう一度深く感謝するとともに、何年も続くこの研究の一年に携わることができましたことに深く感謝の意を表します。

本研究に関する研究業績

口頭発表

内山田 太一, Lu Xugang, 党建武, 赤木正人, "変形聴覚フィードバックによる発話系の補償動作に関する研究" 日本音響学会春季講演論文, 3-P-9, September 2006

内山田 太一, Lu Xugang, 党建武, 赤木正人, "変形聴覚フィードバックに対する発話系の反応の計測", 日本音響学会春季講演論文, 2-Q-18, March 2007