JAIST Repository
https://dspace.jaist.ac.jp/

Title	脳活動測定による基本周波数の時間変化と感情知覚の関連性に関する基礎的研究
Author(s)	濱田，康弘
Citation	
Issue Date	2010-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8955
Rights	
Description	Supervisor:赤木　正人，情報科学研究科，修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

修 士 論 文

脳活動測定による基本周波数の時間変化と感情知覚の関連性に関する基礎的研究

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

濱田康弘

2010 年 3 月

修 士 論 文

脳活動測定による基本周波数の時間変化と感情知覚の関連性に関する基礎的研究

指導教官　赤木正人 教授

審査委員主査　赤木正人 教授
審査委員　鵜木祐史 准教授
審査委員　党 建武 教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

0810050 濱田康弘

提出年月: 2010 年 2 月

Copyright © 2010 by Yasuhiro Hamada

概 要

音声には，言語情報以外に，感情や個人性，性別などを伝える非言語情報が含まれる．その中でも，感情は人と人とのコミュニケーションを円滑に，また，豊かにするために重要な役割を担っている．感情を含む音声の知覚，生成機構の理解のためには，音声のどのような音響的特徴の違いが，どの脳部位を賦活させ，感情の差異として知覚されているのか，ということが重要である．過去の研究により，感情知覚と関係の深い音響的特徴として基本周波数の時間変化が知られている．本論文では，意図的に音響的特徴を制御した合成音声を用い，脳活動測定を行うことにより，基本周波数の時間変化と感情音声知覚の解明を試みた．

脳活動測定で用いた刺激音は，林の研究に基づいて，感動詞「ええ」の自然発話音声をそのまま分析合成した元音声 1 つ (S0) と合成前に基本周波数の時間変化に対応した音響的特徴を操作した合成音声 5 つ (S1 ∼S5) である．聴取実験において，基本周波数の時間変化の異なる各刺激音に対応する感情の特性を調べた．各刺激音の代表的な感情は，元音声 S0｛肯定，共感｝，合成音声 S1｛肯定，冷静｝，S2｛落胆，悲しい｝，S3｛聞き返し，驚き｝，S4 ｛疑い，否定｝，S5｛驚き，疑い｝であることが調べられた．脳活動測定では，5 つの合成音声 (S1 ∼S5) を聴いた際の脳活動それぞれから S0 を聴いた際の脳活動の差をとり，結果を得た．結果，刺激対 S1 - S0 では大脳皮質の前頭葉で，刺激対 S2 - S0 では大脳皮質の上頭頂小葉や角回で特徴が見られた．他 3 つの刺激対 (S3 - S0, S4 - S0, S5 - S0) では，大脳基底核の尾状核や被核で主な特徴が見られた．

聴取実験による感情の評価と脳の賦活部位の関係を結び付けるために，本研究では，感情階層説を参考とした．本研究で参考とした感情階層説とは，脳は進化論的に階層構造を成し，視床下部周辺において原始情動の処理を，大脳辺縁系において基本情動の処理を，大脳皮質において社会的感情の処理を行うと考え，また，注意覚醒系は感情とは別系統として発達してきたと考える仮説である．

各刺激音の感情と賦活部位の関係を調べた結果,｛肯定，冷静｝の感情 (S1) は，社会的感情として，大脳皮質の前頭葉に関係し,｛落胆，悲しい｝の感情 (S2) は，社会的感情として，大脳皮質において処理されている可能性が示唆された．また,｛驚き｝を含む注意覚醒系 (S3, S4, S5) では，原始情動を扱う，大脳基底核に関わってくる可能性が示唆された．

これらの結果より，基本周波数の時間変化に対応する音響的特徴の異なる刺激音は，異なる感情の階層に渡って処理されていることが示唆された．

目 次

第 1 章　序論 **1**
1.1　はじめに 1
1.2　本研究の背景 1
1.2.1　音響的特徴と感情知覚に関する研究 2
1.2.2　脳活動測定による感情知覚に関する研究 2
1.2.3　進化論的感情階層仮説 2
1.3　本論文の目的 3
1.4　研究方法 3
1.5　本論文の構成 4

第 2 章　感動詞「ええ」における音響的特徴の分析と音声合成 **6**
2.1　はじめに 6
2.2　脳活動測定で用いる刺激音の条件 6
2.3　音声分析合成系 STRAIGHT 6
2.4　感動詞「ええ」における音響的特徴の分析 7
2.4.1　音声データの収録 7
2.4.2　基本周波数の特徴 7
2.4.3　基本周波数時間変化とパワー包絡の関係 11
2.5　合成音声の制御方法 11
2.5.1　基本周波数の制御 11
2.5.2　パワー包絡の制御 14
2.5.3　フォルマントの制御 14
2.6　まとめ 14

第 3 章　感情音声に関する聴取実験 **15**
3.1　はじめに 15
3.2　聴取実験 1：合成音声の知覚的距離の測定と自然性の評価 15
3.2.1　目的 15
3.2.2　実験方法 15
3.2.3　実験結果，考察 16
3.3　聴取実験 2：脳活動測定で用いる刺激音の知覚特性の調査 21
3.3.1　目的 21

3.3.2　実験方法 21
3.3.3　実験結果，考察 23
3.4　まとめ 24

第 4 章　感情音声に関する脳活動測定実験　26
4.1　はじめに 26
4.2　実験の目的 26
4.3　実験方法 26
4.4　解析方法 27
4.5　解析結果 28
4.6　脳活動測定実験結果の考察 42
4.7　まとめ 44

第 5 章　脳活動測定結果と感情知覚に関する考察　45
5.1　はじめに 45
5.2　聴取実験の結果と感情階層説との関係 45
5.3　脳活動測定による感情知覚に関する考察 46
5.4　まとめ 46

第 6 章　結論　50
6.1　本研究で明らかになったことの要約 50
6.2　今後の課題 50

図 目 次

1.1 各章の相互関係 5

2.1 STRAIGHT による音響的特徴の分析過程 8
2.2 各感情音声における F0 継続時間長 10
2.3 各感情音声における基本周波数の変化：最大値，最小値，平均値 10
2.4 各感情音声における基本周波数の変化：勾配量 12
2.5 基本周波数の勾配とパワー包絡の相関関係 12
2.6 分析合成の過程 13

3.1 基本周波数の勾配と継続時間長 17
3.2 各刺激音における聞き返しの評価の平均値 17
3.3 各刺激音における驚きの評価の平均値 18
3.4 各刺激音における肯定の評価の平均値 18
3.5 各刺激音における時間稼ぎの評価の平均値 18
3.6 各刺激音における疑いの評価の平均値 19
3.7 各刺激音における落胆の評価の平均値 19
3.8 各刺激音における迷いの評価の平均値 19
3.9 クラスター分析の結果 20
3.10 各刺激音における自然性の評価の平均値 22
3.11 各刺激音の基本周波数の勾配と F0 継続時間長 22
3.12 各刺激音に対する感情語の評価の平均値 23
3.13 多次元尺度法による刺激音間の非類似性 25

4.1 S1 - S0 の脳活動の透し図 (左上：側面図，左下：上面図，右上：前面図であり，黒い部分が有意な活動差を示す．) 29
4.2 S0 - S1 の脳活動の透し図 30
4.3 S2 - S0 の脳活動の透し図 31
4.4 S0 - S2 の脳活動の透し図 32
4.5 S3 - S0 の脳活動の透し図 33
4.6 S0 - S3 の脳活動の透し図 34
4.7 S4 - S0 の脳活動の透し図 36
4.8 S0 - S4 の脳活動の透し図 37

4.9 S5 - S0 の脳活動の透し図 39
4.10 S0 - S5 の脳活動の透し図 40

5.1 感情階層説と各刺激音の分類．ref. 福田 正治「感情学入門」[1] 47
5.2 S2 - S0 における主な賦活部位 48
5.3 S3 - S0 における主な賦活部位 48
5.4 S4 - S0, S5 - S0 における大脳基底核における賦活 49

表 目 次

2.1 「ええ」を含んだ対話文 9
2.2 録音に使用した機器 9

3.1 聴取実験に使用する機器 16
3.2 クラスターを形成する刺激 17
3.3 感情語のリスト 23
3.4 stress 値 25

4.1 S1 - S0 における有意な脳活動部位 29
4.2 S0 - S1 における有意な脳活動部位 30
4.3 S2 - S0 における有意な脳活動部位 31
4.4 S0 - S2 における有意な脳活動部位 32
4.5 S3 - S0 における有意な脳活動部位 33
4.6 S0 - S3 における有意な脳活動部位 34
4.7 S4 - S0 における有意な脳活動部位 35
4.8 S0 - S4 における有意な脳活動部位 37
4.9 S5 - S0 における有意な脳活動部位 38
4.10 S0 - S5 における有意な脳活動部位 40

第1章　序論

1.1　はじめに

音声には，言語情報以外に，感情や個人性，性別などを伝える非言語情報が含まれる．その中でも，感情は人と人とのコミュニケーションを円滑に，また，豊かにするために重要な役割を担っている．感情における研究は，心理学，神経科学から哲学にいたるまで，広範囲にわたり，研究がなされてきた．近年では，fMRI などの装置の発達により，脳の活動が非侵襲的に計測できるようになり，脳科学の領域でも感情について研究されるようになってきている．

音声研究において，感情の研究は，知覚，生成の両面から研究がなされ，その応用として，音声合成，音声認識など様々な分野で研究がなされ始めている．Denes と Pinson によって提唱された「言葉の鎖」は人と人との音声コミュニケーションを模擬するモデルであり，人の調音運動から，音声生成，音響信号，音声知覚へと音声情報が伝わり，これが繰り返されていくものである [2]．その鎖の中で，脳の領域は未知の部分が多く，どのような音響信号がどのように脳に符号化されて，調音運動として出力されていくのかは明らかとなっていない．感情に関する知覚や生成の機構を解明するためにも，音声のどのような音響的特徴の違いがどのように人に知覚されているのかを調べる事は重要である．聴取実験により音響的特徴と感情知覚の関連性について調べられてきているが，知覚の主体である人の脳の知見はまだあまりない．

一方，神経科学者や心理学者によって，感情音声に関する脳活動は調べられ始めているが，音響的特徴が脳活動に与える影響については十分に考慮されていない．感情による音声コミュニケーションの本質を理解するためには，音声のどのような音響的特徴の違いが，どの脳部位を賦活させ，感情の差異として知覚されているのか，ということを調べることが重要である．本研究では，感情知覚と関係の深い基本周波数の時間変化に着目し，意図的に音響的特徴を制御した合成音声を用いて脳活動測定を行うことにより，感情音声知覚の解明を試みる．

1.2　本研究の背景

感情の研究は，心理学，神経科学など様々な分野で研究がなされている．その中で，まず始めに，聴取実験による音響的特徴と感情知覚に関する研究を述べる．次に，脳活動測定による感情知覚に関する研究を述べる．最後に，本研究で感情や情動の知覚と脳活動を

結び付けるための手段として参考とする，福田の進化論的感情階層仮説 (感情階層説) [3] について述べる．

1.2.1 音響的特徴と感情知覚に関する研究

非言語情報である感情音声の研究は，生成，知覚両面から様々な研究がなされている．その応用として，音声合成，音声認識などの分野で研究がなされ始めている．音声合成では，森山らにより，韻律の部分空間を用いた感情音声を合成する手法を提案され，感情が合成した通りに知覚されることが示されている [4]．また，小池らは，音素を合成単位とする波形合成方式による音声合成モジュールを用いて韻律情報を変化させることにより，音声に含まれる感情を表現できることを確かめている．音声認識では，刀根らは韻律情報を用いて HMM に基く感情モデルを構築し，動的な時系列パターンを扱うことで高い感情判別率を確認している [5]．このような感情音声合成，認識の研究から，韻律情報は感情知覚に密接に関係している音響的特徴であると考えられる．

韻律情報を形成する音響的特徴には，基本周波数の時間変化，パワー包絡，音韻継続長などがある．その中で，林は感動詞「ええ」を用いた聴取実験により，基本周波数の時間変化が多くの感情情報を伝えている，ということを報告している [6]．このように，聴取実験により，音響的特徴と感情知覚について調べられてきている．しかし，知覚の主体である人の脳の知見は数が少なく，どのような音響的特徴が人の脳内でどのように感情知覚と結び付くかは，明らかになっていない．

1.2.2 脳活動測定による感情知覚に関する研究

感情音声を含む多くの音声知覚, 音知覚の研究から, 音の情報処理は一次聴覚野を含む, シルビウス皮質に介在していることが知られている. 多くの研究で, 一次聴覚野に含まれる上側頭回 (STG) において音刺激の差異による活動が報告されている [7]. また, 上側頭溝, 中側頭回や, 溝の後ろにあたる視覚情報との統合もなされるとされる縁上回, 角回などの活動も報告されている [8]. 感情音声に関する脳活動測定では，Wiethoff らが平常の音声を聴いたときよりも，感情音声で右半球の上側頭回がより活動する，という報告 [9] や，Bach らによる音韻情報が主に左半球の下前頭回で処理されている，という報告などがある [10]．しかし，これらの報告では，音響的特徴が脳活動に与える影響については十分に考慮されていない．

1.2.3 進化論的感情階層仮説

感情や情動の知覚が人の脳内で，どのように結び付くのか，ということを議論するために，本論文では，福田の進化論的感情階層仮説 (感情階層説)[3] を参考とする．感情階層説では感情 (affection) は情動 (emotion) と高等感情 (feeling) に分類される．さらに，情動

は原始情動 (primitive emotion), 基本情動 (basic emotion) に分類され，高等感情は社会的感情 (social feeling), 知的感情 (intellcectual feeling) に分類される．

感情は進化的に階層構造を有しており，原始情動は進化の過程で最初に出現し，快-不快を表現するとされている．これは視床下部周辺の活動によるものと考えられている．基本情動は捉え方や言葉の定義などに多くの議論があるが，ここでは，喜び，受容・愛情，怒り，恐怖，嫌悪とされている．これは，進化の次の段階として, 大脳辺縁系に付け加えられたとされている．社会的感情は，サルから進化したホモ・サピエンスが言語を獲得するに至るまでの段階で身につけていったものと考えられている．社会的知性は推論，予測，問題解決能力，関係性の認知と長期間の記憶の保持，読心などの機能を表す．これは，個体が複雑な集団の中で社会的な問題解決に必要とされる社会的知性を遂行するための領域を大脳辺縁系や大脳皮質の領域で獲得していったと考えられる．知的感情は，ヒトが言葉を獲得し記録を残せるようになってから身につけた特有の感情である．また，驚きなどの注意覚醒系の機能は情動や感情とは別系統として発達したと考えられている．

1.3 本論文の目的

感情による音声コミュニケーションの本質を理解するためには，音声のどのような音響的特徴の違いが，どの脳部位を賦活させ，感情の差異として知覚しているのか，ということを調べる事が重要である．研究の背景より，現在の感情音声の研究の問題点として，

- 聴取実験により，音響的特徴と感情知覚の関連性について調べられてきているが，知覚の主体である人の脳の知見はまだあまりない．
- 感情音声に関する脳活動は調べられ始めているが，音響的特徴が脳活動に与える影響については十分に考慮されていない．

という問題が挙げられる．

そこで，本論文では，音響的特徴として感情知覚と関係の強い基本周波数の時間変化に着目し，これを様々に変化させ呈示したときの脳活動測定を行うことにより，基本周波数の時間変化と感情知覚の関連性を明らかにすることを目的とする．

1.4 研究方法

音響的特徴の影響を調べるために，林の研究に基づき，感動詞「ええ」における基本周波数の時間変化を意図的に制御した合成音声を用いる．音声のどのような音響的特徴の違いが感情の差異として知覚されているのか，ということに対し，聴取実験において，異なる基本周波数の時間変化に対応する感情の特性を調べる．そして，どの脳部位を賦活させるのか，ということに対し，脳活動測定により，各音声における脳の賦活部位を特定する．感情階層説を参考に，これらを統合し，異なる基本周波数の時間変化に対応する感情知覚と脳活動の関連性を検討する．

1.5　本論文の構成

本論文の構成を以下に示す.

第1章では，本論文が対象としている研究分野の背景と問題点を示し，本論文の目的と研究方法について述べる.

第2章では，感動詞「ええ」における音響的特徴の分析，および合成音声の制御について述べる.

第3章では，聴取実験を行うことにより，音響的特徴と心理特徴の関係を明らかにする.

第4章では，脳活動測定実験を行い，刺激音に対する脳の賦活部位を調べる.

第5章では，感情階層説を参考に，聴取実験の結果，脳活動測定の結果をまとめ，考察を行う.

第6章では，本論文で得られた結果を要約し，今後の展望を述べる.

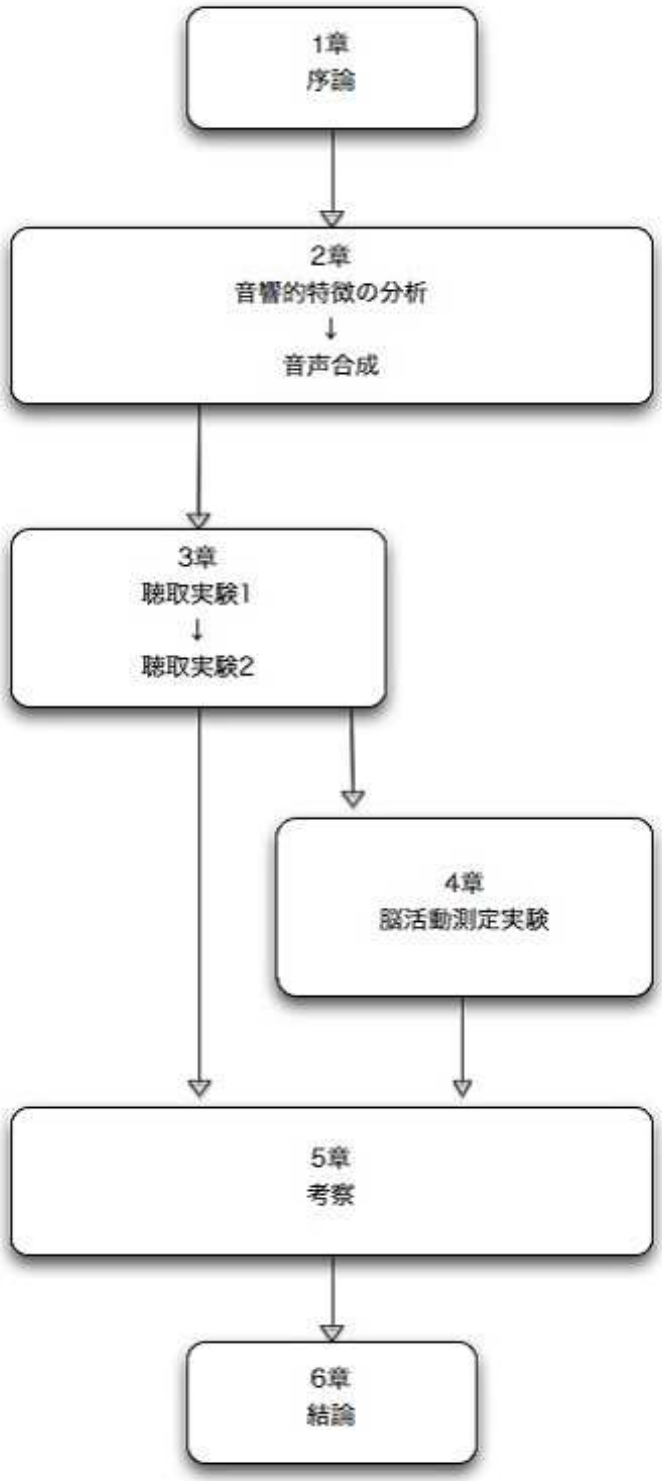

図 1.1: 各章の相互関係

第2章　感動詞「ええ」における音響的特徴の分析と音声合成

2.1　はじめに

本研究では，音響的特徴と感情知覚の関連性を議論するために，脳活動測定を行う際，感動詞「ええ」の基本周波数の時間変化を制御した合成音声を用いる．本章では，刺激音の音響的特徴を制御するための方法論を述べる．はじめに，脳活動測定で用いる刺激音の条件を述べ，次にその条件を満たす刺激音を作成する準備として，実際の音声の音響的特徴の分析について述べる．最後に，音響的特徴の分析結果を参考にした音響的特徴の制御方法について述べる．

2.2　脳活動測定で用いる刺激音の条件

研究の背景で述べたように，林は短い音節である感動詞「ええ」を用いることにより，基本周波数の時間変化と感情知覚の関係を調べることが出来た．本研究では，林の結果[6] を基に，感動詞「ええ」の音響的特徴を意図的に制御した合成音声を用いて脳活動を測定する．脳活動測定では，実験時間が制限されてしまうため，刺激音は実験時間を考慮して最適な数6種とした．脳活動測定では，自然性が低いと意図しない活動が示される可能性があるため，この刺激音は知覚的な差があり，自然性が高いことを条件とした．

音声を合成する際，声帯情報である基本周波数のみの制御を行うと，声道形状情報との相関が低くなり，自然性が損なわれると考えられる．そこで，基本周波数以外に，声帯情報と関わってくると考えられるパワー包絡，フォルマントについて着目し，自然性に留意して音声の合成を行うこととした．

6種の刺激音のうち，1つは自然発話音声をSTRAIGHT[11][12] によりそのまま分析合成した元音声で，他の5つは合成前に韻律情報を操作した合成音声である．

2.3　音声分析合成系STRAIGHT

脳活動で用いる刺激音を作成するためには，特有の音響的特徴を自由に操作でき，且つ精度よく音声を合成するシステムを用いる必要がある．音声分析合成系とは，音声生成過程モデルに基づき，音声から声帯振動情報である基本周波数と声道形状情報であるスペ

クトルの音響パラメータを抽出し，これらを用いて再合成するシステムである．本研究では，基本周波数の時間変化に着目し，それに対応した音響的特徴を議論する．そのため，音響的特徴の操作の自由度が高く，且つ高品質に音声を再合成することができる，河原によって提案された音声分析合成系STRAIGHTを用いる．STRAIGHTを用いることにより，基本周波数の時間変化やそれに対応する音響パラメータの制御を自由に，且つ精度良く音声を分析，再合成することが可能となる．

システムの音響パラメータの抽出過程を図2.1に示す.

2.4 感動詞「ええ」における音響的特徴の分析

STRAIGHTを用いて音声を分析合成する際，刺激音の数は条件より6種であり，その刺激音は知覚的に距離があり，且つ自然性が高いことが要求された．自然に発声される音声の特徴を損なう音響的特徴の操作を行うと，自然性が低下することから，まずはじめに音声データの採取を行い，その音声データの音響的特徴の分析を行うことにより，合成音声作成の指標とした．この節では，音声データの採取，音響的特徴の分析について述べる．

2.4.1 音声データの収録

音声データの収録に関する方法を述べる．発話者は男性5名，女性1名である．林の研究で用いたものと同じ7つの文脈(聞き返し，驚き，肯定，時間稼ぎ，疑い，落胆，迷い)で,「ええ」が発声される対話文(表1)をそれぞれ発話してもらった.

収録環境

録音は防音室にて行った．対話全体を標本化周波数44.1 kHzでDATレコーダに録音した．これを標本化周波数16 kHzにダウンサンプリングして,「ええ」の部分のみを切り出し，ワークステーションに保存し音声データとした．録音に使用した機器を以下に示す.

””

2.4.2 基本周波数の特徴

基本周波数の制御を行う際，自然性を保つために，実際に発声される音声の範囲内で基本周波数の制御を行う．音声合成の指標とするために実際の音声における基本周波数の特性を調べた．男性の発話者5名 ×7種の文脈における「ええ」について，STRAIGHTを用いて，基本周波数の抽出を行い，継続時間長，基本周波数の最大値，最小値，平均値，勾配の関係を調べた.

結果を図2.2から2.4に示す.

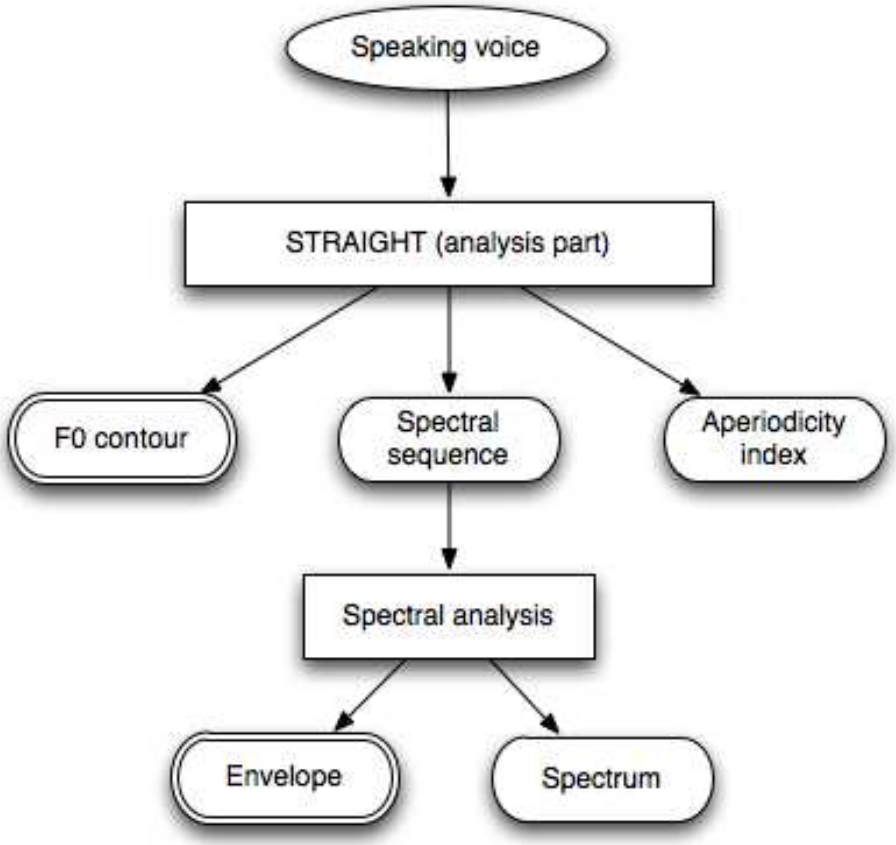

図 2.1: STRAIGHT による音響的特徴の分析過程

表 2.1: 「ええ」を含んだ対話文

(聞き返し)	A:庭には2羽ニワトリがいる.
	B:ええ？今なんて言った？
(驚き)	A:昨日朝青龍が優勝したじゃない.
	B:ええ !? 知らなかった.
(肯定)	A:先週の日曜は雨でしたっけ.
	B:ええ，どしゃぶりでした.
(時間稼ぎ)	A:では，審査委員長お願いします.
	B:それでは発表します．ええ，第1位は...
(疑い)	A:（いかすみスパゲッティーを前にして）これすっごくおいしいんだよ.
	B:（気持悪いと思って）ええ？
(落胆)	A:(合格発表の掲示板の前で) 番号ないね.
	B:ええ.
(迷い)	A:ねぇ，一緒に行こうよ.
	B:（決めかねて）ええ，どうしようかな.

表 2.2: 録音に使用した機器

機器	メーカー，機種
マイクロフォン	SONY C-536P
DAT レコーダ	Marants PMD671

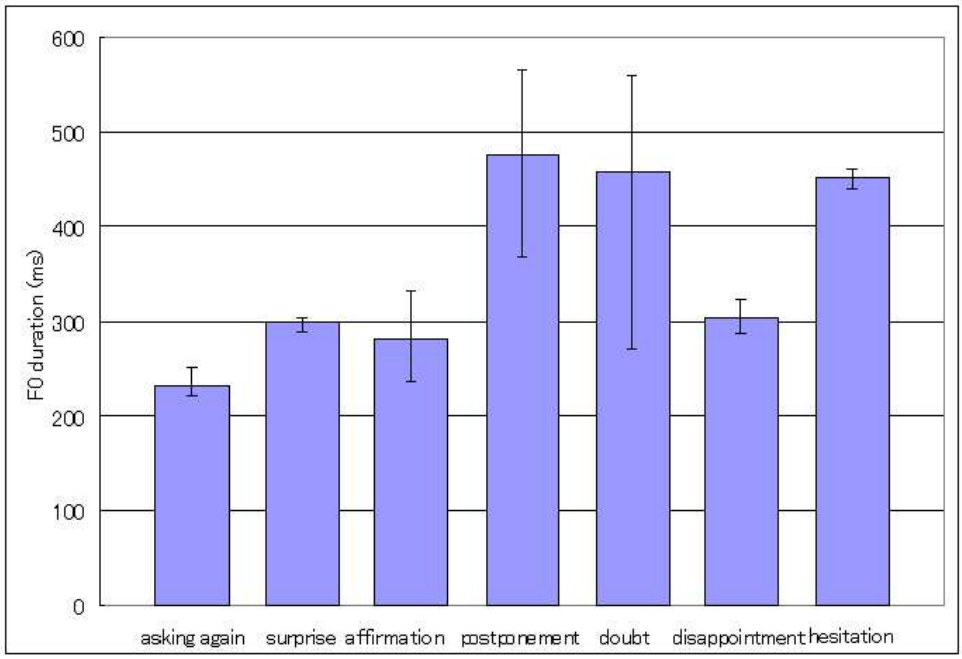

図 2.2: 各感情音声における F0 継続時間長

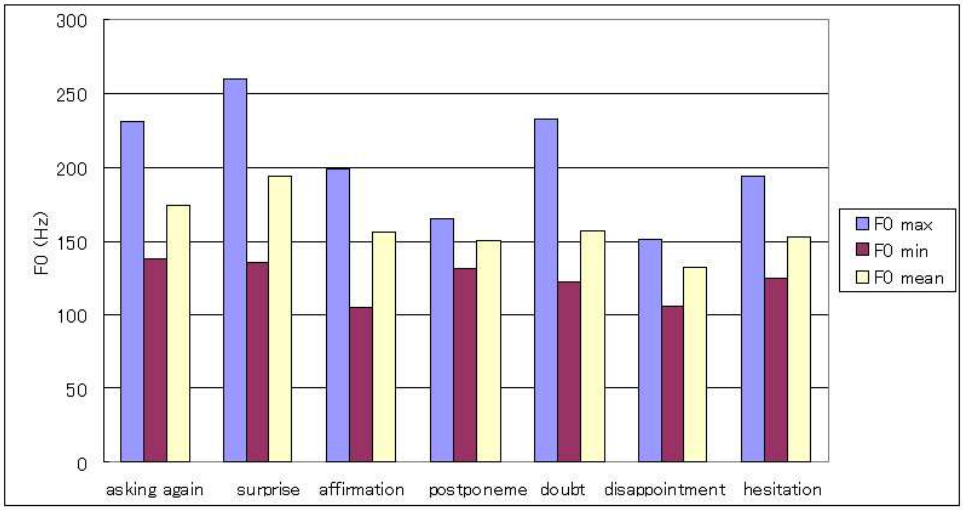

図 2.3: 各感情音声における基本周波数の変化：最大値，最小値，平均値

結果，各感情における継続時間長は，時間長の短いもの (聞き返し，驚き，肯定，落胆) と，長いもの (時間稼ぎ，疑い，迷い) とに特徴が分かれた．これは，林の研究でも同じような結果が報告されている．また，継続時間長が短いものは，負であるもの (肯定，落胆) と正のもの (聞き返し，驚き) に，継続時間長が長いものは，正 (時間稼ぎ，疑い，迷い) であるという特徴が見られた．

2.4.3 基本周波数時間変化とパワー包絡の関係

声帯の振動は声帯が引き延ばされると張力が増加し，また声帯が薄く軽くなるので，振動数の増加がもたらされる．このような声帯の伸長は，輪状甲状節の収縮や，輪状甲状節の回転を引き起こす外力が要因となる [13]．そのため，声帯の振動を表す基本周波数と，声道情報を表すスペクトル成分のパワーには相関関係があると考えられる．基本周波数の時間変化とパワー包絡の関係を調べるために，発話者 6 名 ×7 つの文脈における音声について，基本周波数，スペクトル成分のパワー抽出を行い，分析を行った．基本周波数の勾配とパワー包絡の勾配の関係を図 2.5 に示す．

相関計数は 0.88 となり，基本周波数とパワー包絡は相関が強いことが示された．回帰直線は，

$$y = 20.8x \tag{2.1}$$

で表された．(x: 基本周波数の勾配, y: パワー包絡の勾配)

2.5 合成音声の制御方法

この節では，合成音声の制御方法について述べる．分析結果を指標に，合成音声の作成を行った．音声データは ATR データベースの「ええ」を発話した録音を用いた (file name: mttsa414.wav)． データベースの「ええ」を入力とし，基本周波数，パワー包絡，フォルマント周波数を以下に記す方法で修正を行い，異なる感情が付加された合成音声を出力した．分析合成の過程を図 2.6 に示す．

2.5.1 基本周波数の制御

基本周波数の時間変化の異なる音声を合成するために，STRAIGHT を用いて抽出した基本周波数を修正し，新しい基本周波数を得た．基本周波数の修正は，点ピッチモデルを用いた．基本周波数の対数軸上で，始点，中間点，終点の 3 点の基本周波数の値を設定し，キュービックスプラインデータ補間を行った．また，声帯の振動は微細成分を含み，微細成分の変化は不自然な音声を生成することから，基本周波数をエンベロープと微細成分に分け，エンベロープのみを変化させ，再合成した．

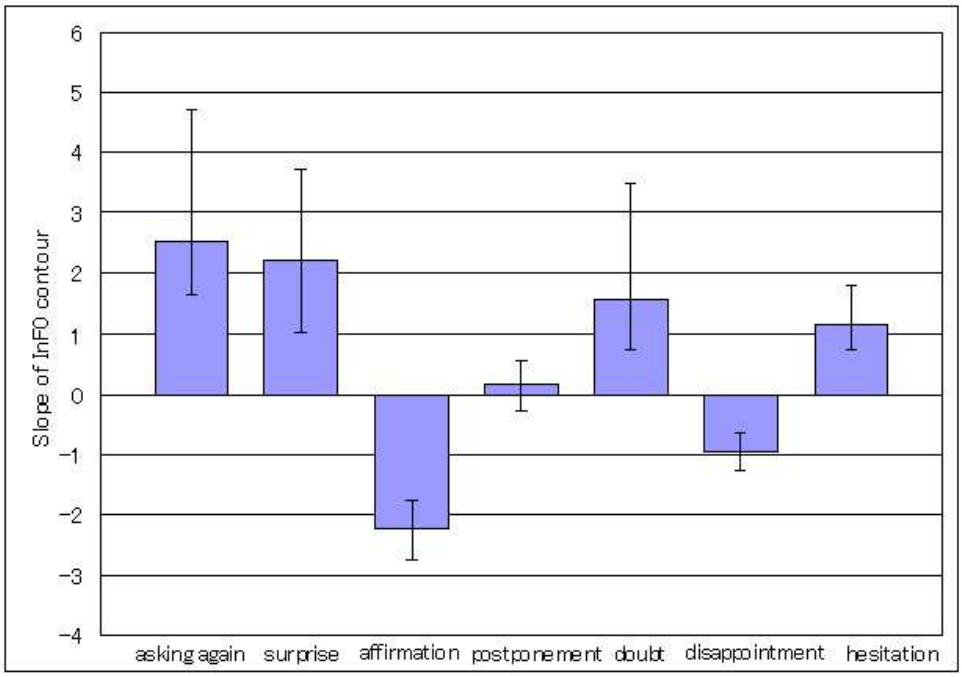

図 2.4: 各感情音声における基本周波数の変化：勾配量

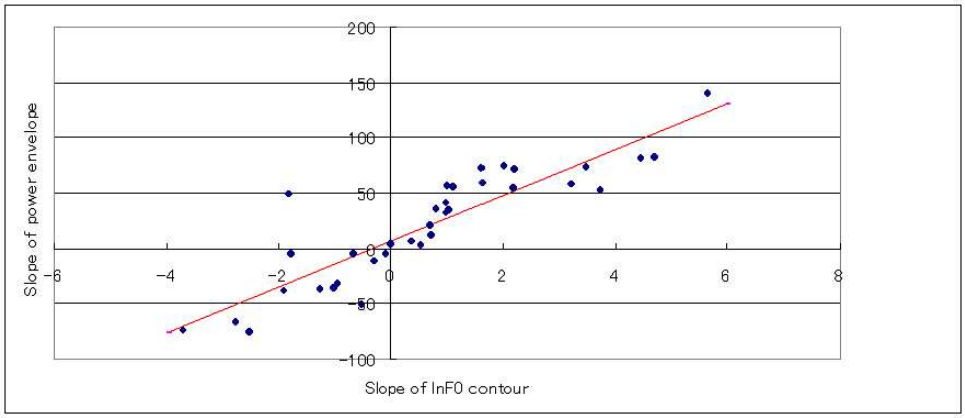

図 2.5: 基本周波数の勾配とパワー包絡の相関関係

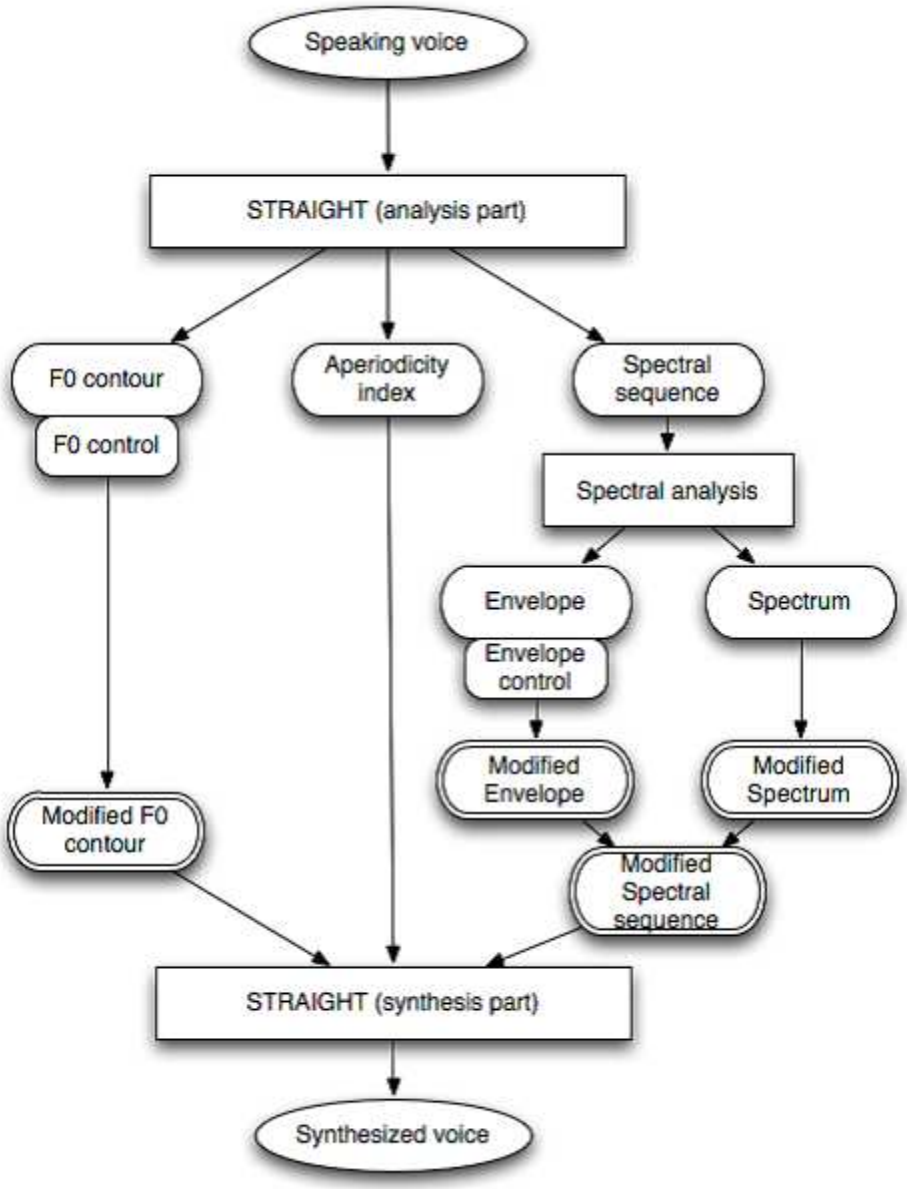
Speaking voice
STRAIGHT (analysis part)
F0 contour
F0 control
Aperiodicity index
Spectral sequence
Spectral analysis
Envelope
Envelope control
Spectrum
Modified Envelope
Modified Spectrum
Modified F0 contour
Modified Spectral sequence
STRAIGHT (synthesis part)
Synthesized voice

図 2.6: 分析合成の過程

2.5.2　パワー包絡の制御

前節で，基本周波数とパワー包絡には相関があることが判った．基本周波数のみの制御を行い，再合成すると声帯振動情報と声道形状情報の相関が低くなるため，合成音声の自然性が損なわれる．そのため，STRAIGHT を用いて抽出したスペクトル成分のパワー包絡を修正し，新しいパワー包絡を得た．パワー包絡の制御は前節で求めた基本周波数とパワー包絡の相関関係 (式 2.1) を用いた．スプライン補間を行い，基本周波数の時間変化に対応した制御を行った．

2.5.3　フォルマントの制御

声帯の振動には喉頭の上下動が関係しており，外喉頭が働くと声道の長さと舌の形状が変化し，フォルマント周波数も変動することが知られている [13]．そのため，自然性の高い合成音声を作るために，フォルマントの制御を行った．基本周波数の高低に対して，対数周波数軸上でフォルマント周波数をシフトさせ，フォルマントの修正を行った．

2.6　まとめ

本章では，脳活動測定で用いる刺激音の条件，音響的特徴の分析，音声の制御方法について述べた．刺激音は自然性が高いことが要求され，そのため，実際の音声から音響的特徴を分析し，それを基に STAIGHT を用いて基本周波数の制御を行った．基本周波数とパワー包絡には相関があるため，基本周波数に対応するパワー包絡の制御を行った．また，自然性の高い音声を合成するために，フォルマントの修正を行い，音声の合成を行った．

第3章　感情音声に関する聴取実験

3.1　はじめに

前章で述べたように，脳活動測定で用いる刺激音は知覚的に距離があり，自然性が高いことが必要であった．本章では，合成音声の知覚特性，自然性を調べるために行った聴取実験について述べる．まずはじめに，知覚的距離のある，自然性の高い合成音声6つを選択するための聴取実験について述べる．次に，選択した6つの刺激音について，その知覚特性を調べるための聴取実験について述べる．

3.2　聴取実験1：合成音声の知覚的距離の測定と自然性の評価

3.2.1　目的

刺激音の条件より，脳活動測定で用いる刺激音は実験時間を考慮して6つであり，知覚的距離があり，自然性が高いことが必要である．聴取実験1では，脳活動測定で用いる条件を満たす刺激音6つを用意するために，合成音声に対し，聴取実験を行い，感情の評価(聞き返し，驚き，肯定，時間稼ぎ，疑い，落胆，迷い)，および自然性の評価を行った．聴取実験の結果より，お互いに知覚的に距離のある，自然性の高い刺激音を5つ選び，選んだ5つの合成音声および元音声を脳活動測定で用いることとした．

3.2.2　実験方法

刺激音

音響的特徴の分析結果(図2.2 ～ 図2.4)より，各感情における継続時間長は，時間長の短いものと，長いものに分類された．また，各感情と基本周波数の勾配は，継続時間長が短いものは，負であるものと正であるもの，継続時間長が長いものは，正であるという特徴が見られた．これらの分析より，以下の規則により分類された3つのグループ，42種の合成音声について聴取実験を行った．

- (グループ1) 基本周波数の勾配が正（1 ～ 3 Hz/s）のもの (刺激音1 ～ 12)
- (グループ2) 勾配が負（-3 ～ -0.5 Hz/s）のもの (刺激音13 ～ 24)

表 3.1: 聴取実験に使用する機器

機器	メーカー，機種
ヘッドホン	SENNHEISER HDA200
ヘッドホンアンプ	YAMAHA OP-U50, SANSUI AU-α907MR
PC	PANASONIC CF-R6

- (グループ 3) 勾配が正（0 ～ 2 Hz/s) で，時間継続長を 2 倍に伸長したもの (刺激音 25 ～ 42)

用いる刺激音の基本周波数の勾配 (横軸) と継続時間長 (縦軸) を図**??**に示す.

実験環境

被験者は防音室内でヘッドフォンにより受聴する．使用する機器を以下に示す.

実験手続き

刺激音を呈示し，その音声が {聞き返し，驚き，肯定，時間稼ぎ，疑い，落胆，迷い} のどの感情に当てはまるか 0 ～ 5 の点数をつけて評価を行った．刺激音はランダムに呈示された．被験者は 10 名で，42 種の刺激音の呈示を各被験者について 2 回ずつ行った.

3.2.3 実験結果，考察

各感情に対する刺激音 1 ～ 刺激音 42 の評価の平均値を図 3.2 ～ 図 3.8 に示す．それぞれの刺激音間の距離を調べるために，クラスター分析を行った．図 3.9 にクラスター分析を行った結果のデンドログラムを示す.

刺激音間の距離は図 3.2 ～ 図 3.8 に示した評点を基にユークリッド距離を計算して求めた．クラスターを形成する際の距離の計算には最長距離法を用いた．図 3.9 において，距離 17 付近で分割することで，表に示す 5 つのクラスターを得た.

クラスター 1 は，基本周波数の勾配が負で変化が急峻なものであり，肯定の回答が多かった.

クラスター 2 は，基本周波数の勾配が負で変化が穏やかなものであり，落胆の回答が多かった.

クラスター 3 は，基本周波数の勾配が正のものであり，聞き返し･驚きの回答が多かった.

クラスター 4 は，継続時間長が 2 倍で，基本周波数の変化が緩やかなものであり，時間稼ぎ･迷いの回答が多かった.

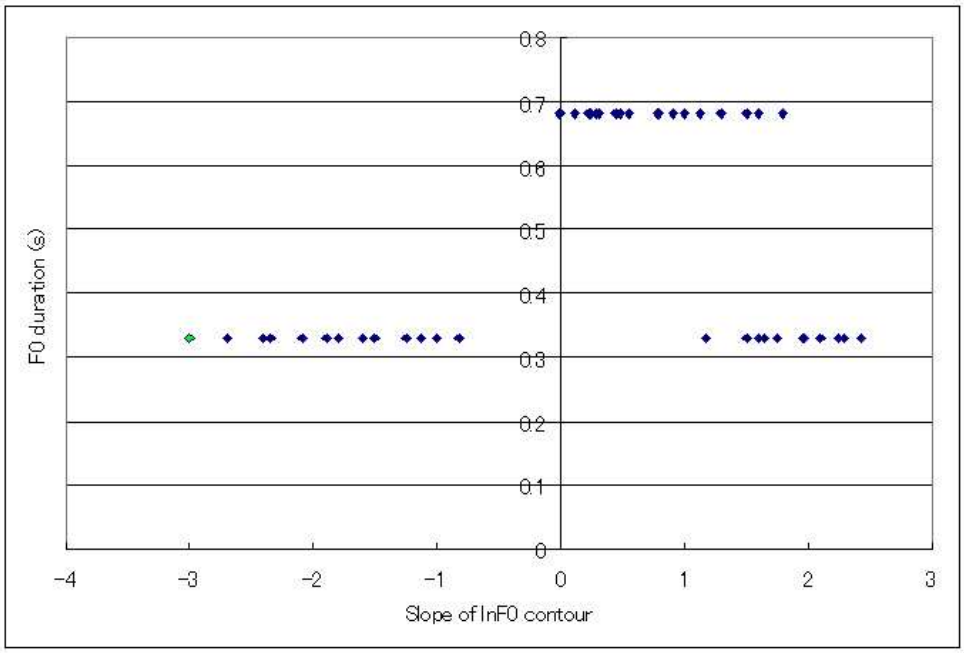

図 3.1: 基本周波数の勾配と継続時間長

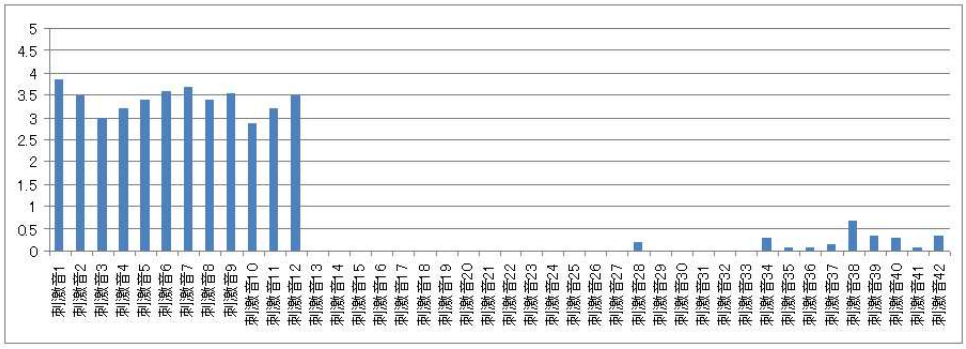

図 3.2: 各刺激音における聞き返しの評価の平均値

表 3.2: クラスターを形成する刺激

クラスター 1	刺激音 13 ～ 20(グループ 2 の一部)
クラスター 2	刺激音 21 ～ 24(グループ 2 の一部)
クラスター 3	刺激音 1 ～ 12(グループ 1)
クラスター 4	刺激音 25 ～ 33（グループ 3 の一部）
クラスター 5	刺激音 34 ～ 42（グループ 3 の一部）

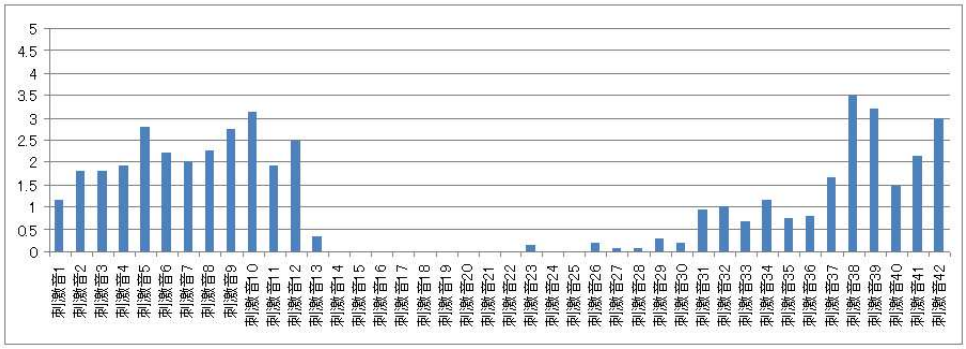

図 3.3: 各刺激音における驚きの評価の平均値

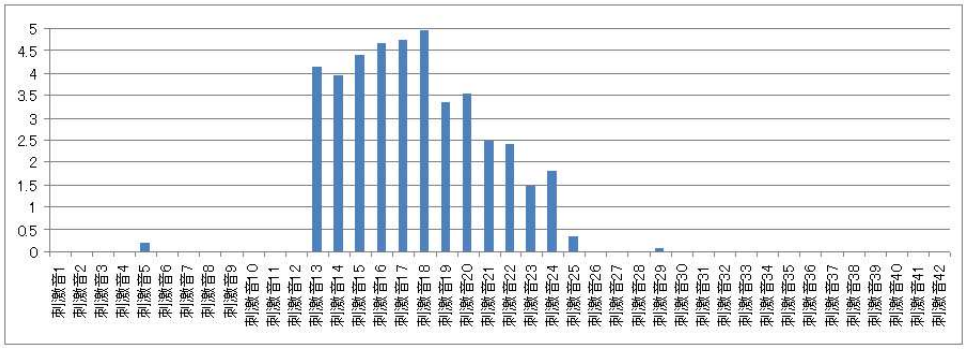

図 3.4: 各刺激音における肯定の評価の平均値

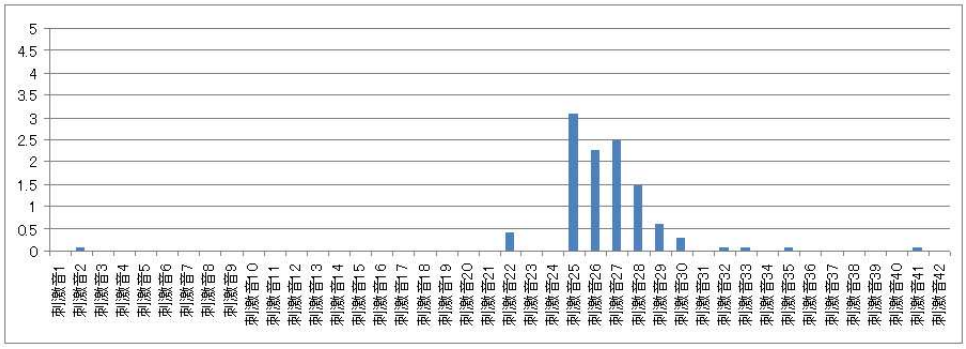

図 3.5: 各刺激音における時間稼ぎの評価の平均値

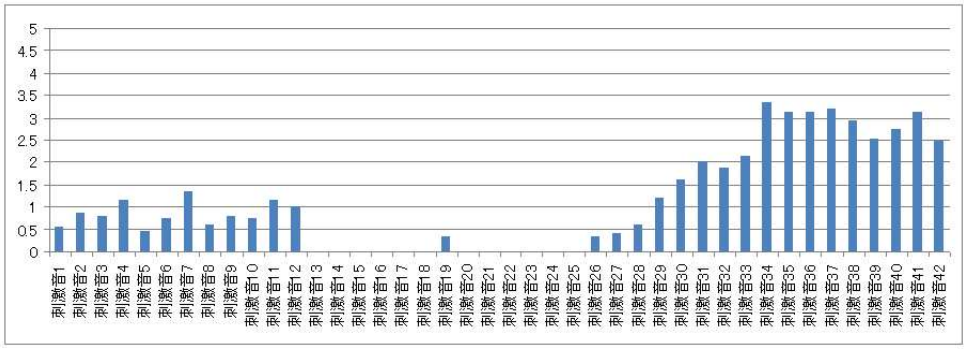

図 3.6: 各刺激音における疑いの評価の平均値

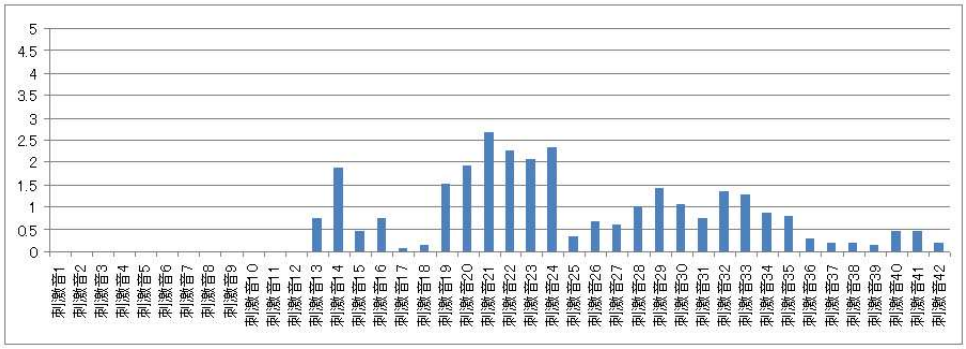

図 3.7: 各刺激音における落胆の評価の平均値

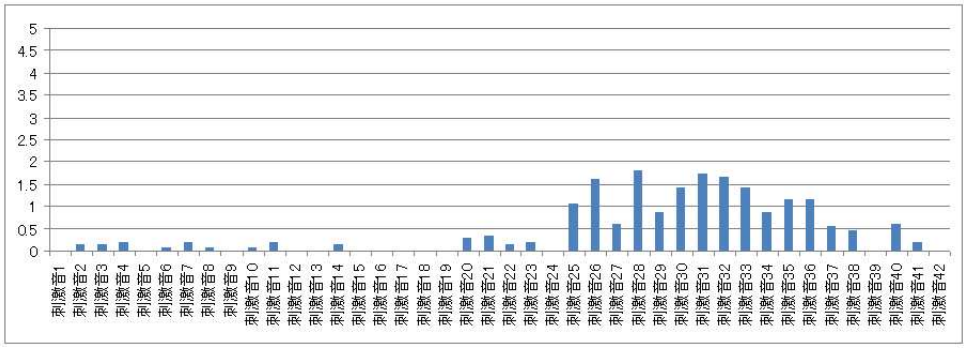

図 3.8: 各刺激音における迷いの評価の平均値

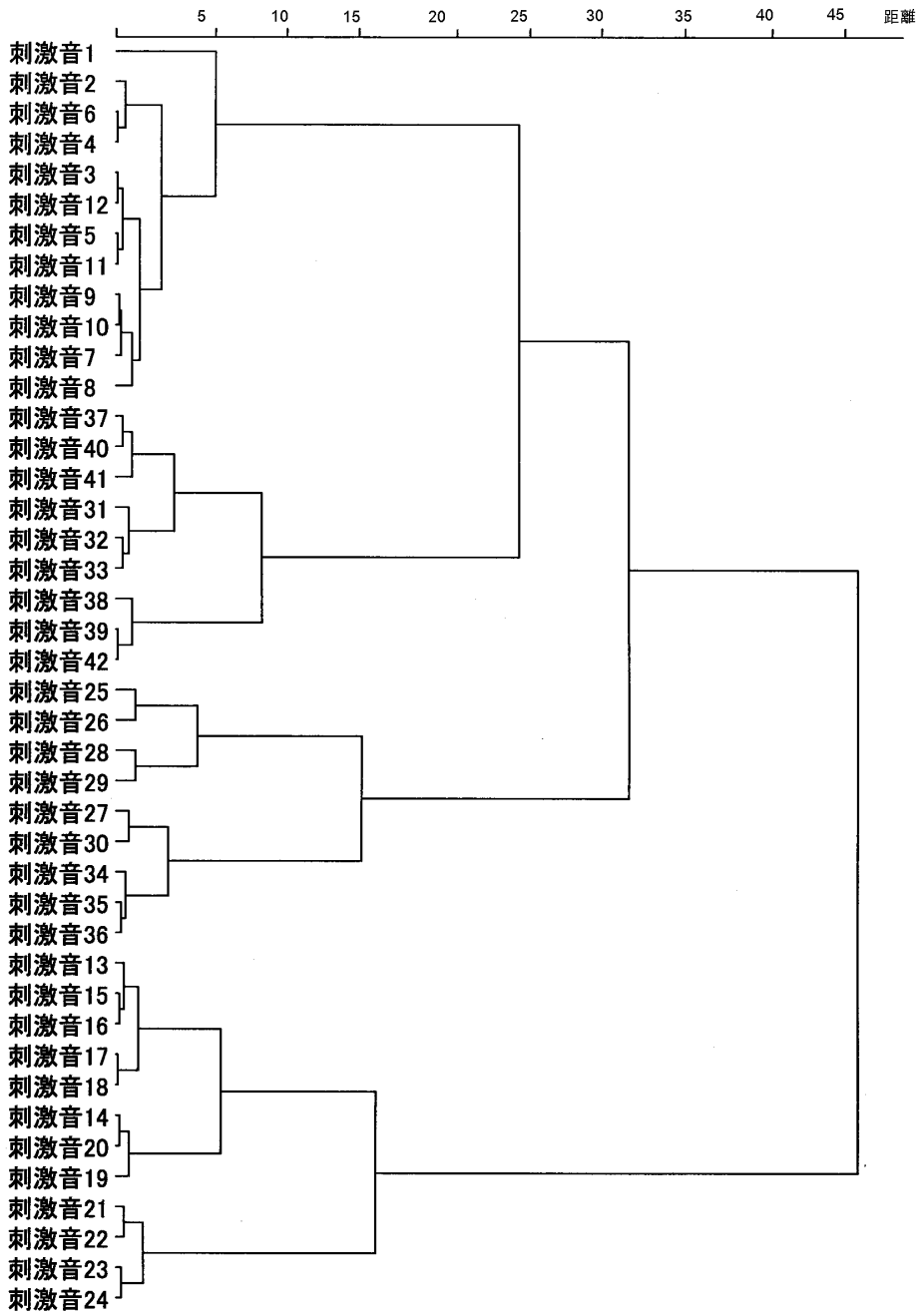

図 3.9: クラスター分析の結果

クラスター5は，継続時間長が2倍で，基本周波数の変化が急なものであり，驚き·疑いの回答が多かった．

自然性の評価の結果を図3.10に示す．

5つのクラスターの中で, 自然性の高い刺激音をそれぞれ1つずつ選択し，この5つの合成音声および元音声を用いて脳活動測定実験を行うこととした．選択した5つの合成音声および元音声の基本周波数の時間変化と継続時間長の関係を図3.11に示す．この際，クラスター1に属する刺激音をS1，クラスター2に属する刺激音をS2，クラスター3に属する刺激音をS3，クラスター4に属する刺激音をS4，クラスター5に属する刺激音をS5，元音声をS0と名前付けをした．

3.3 聴取実験2:脳活動測定で用いる刺激音の知覚特性の調査

3.3.1 目的

聴取実験1では林の研究に基づいた7つの感情語についてのみ，制限的に調べた．音声には先に調べた7つの感情以外に，多くの感情が含まれていることが考えられる．刺激音に対する脳活動測定の結果と感情知覚の関係を議論するためには，刺激音がどのような感情を含んでいるか，またそれぞれの刺激音が感情空間上でどのように付置されているのかを調べる必要がある．そのため，聴取実験2では，非制限的な感情語の評価を行い，選択した5つの合成音声と元音声について，それぞれの刺激音がどのような感情語を有しているか代表的な感情語を調べ，また，感情空間上でどのように付置されているのか，多次元尺度構成法を用いて調べた．

3.3.2 実験方法

評価する感情語は，表現する感情の幅が広く，刺激音が含む感情を的確に表現出来ることが望ましい．しかしながら，感情を表現する語彙は無数にあり，全ての感情語について調べることは困難である．そのため，多くの感情を含む，代表的な感情語を選定するために，予備実験としてKJ法を用いて感情語のリストを作成した．KJ法では，あらかじめ先行研究の文献[6] [4] [14]を参考にリストアップしたものを用意した．初めのプロセスにおいてリストに無い感情語の追加を行い，計85種の感情語を得た．次のプロセスにおいて，感情語は表3.3に示す27種に圧縮された．この際，聴取実験1で取り扱った感情語については恣意的に挿入された．この予備実験には計5名が参加した．

絞り込んだ各感情語について，感情が音声にどの程度含まれているかを7段階評価尺度(最左：「含まれていない」，最右：「非常に含まれている」)で評価してもらった．評価語の並び順が一定であると，一つ前に評価語の影響を受けると考え，その影響を統計的に回避するために，評価用紙の語の並び順はランダムにし，且つ，刺激音を聴かせる順番につ

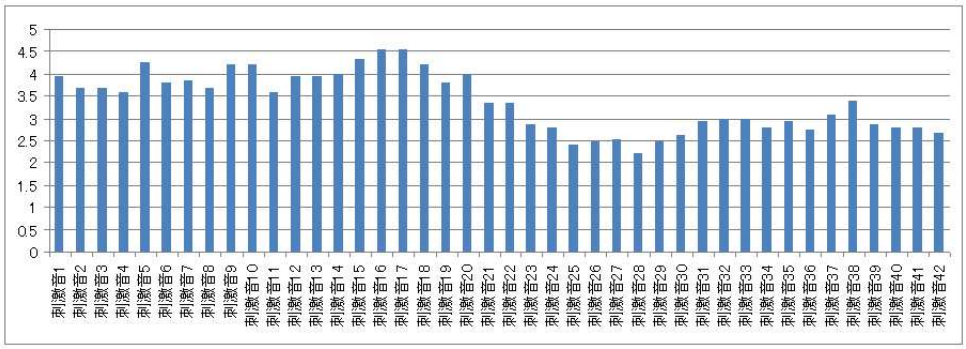

図 3.10: 各刺激音における自然性の評価の平均値

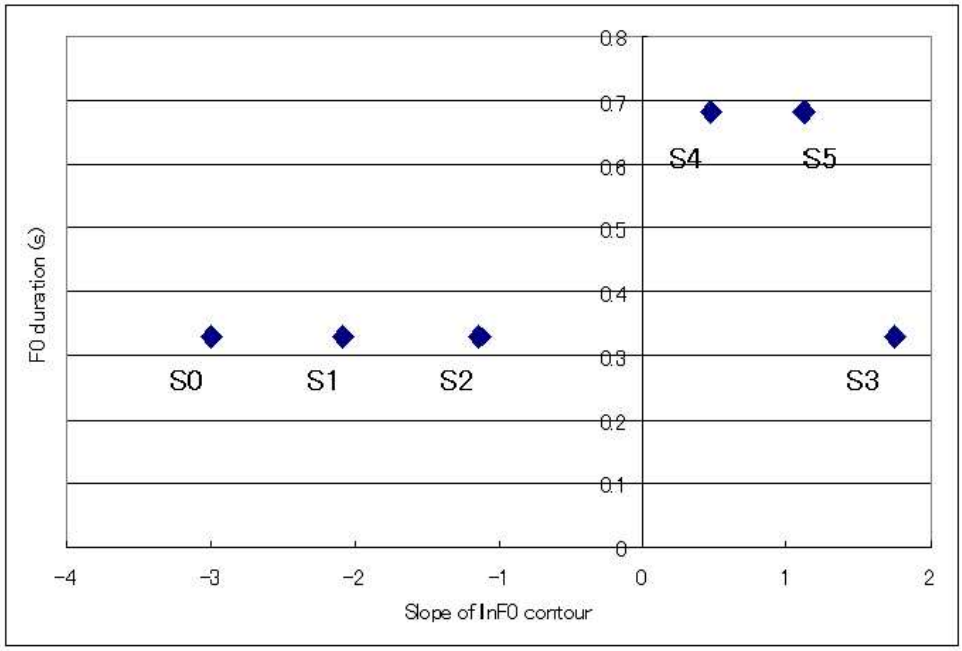

図 3.11: 各刺激音の基本周波数の勾配と F0 継続時間長

表 3.3: 感情語のリスト

1. 愛情	11. 嫌い	21. 驚き
2. 好き	12. 否定	22. 羞恥
3. 共感	13. 疑い	23. こび
4. 快い	14. 苦しい	24. 聞き返し
5. 喜び	15. 怒り	25. 時間稼ぎ
6. 楽しい	16. 恐れ	26. 肯定
7. 中立	17. 悲しい	27. 落胆
8. 冷静	18. 誇り	
9. 優しい	19. 迷い	
10. 期待	20. 焦り	

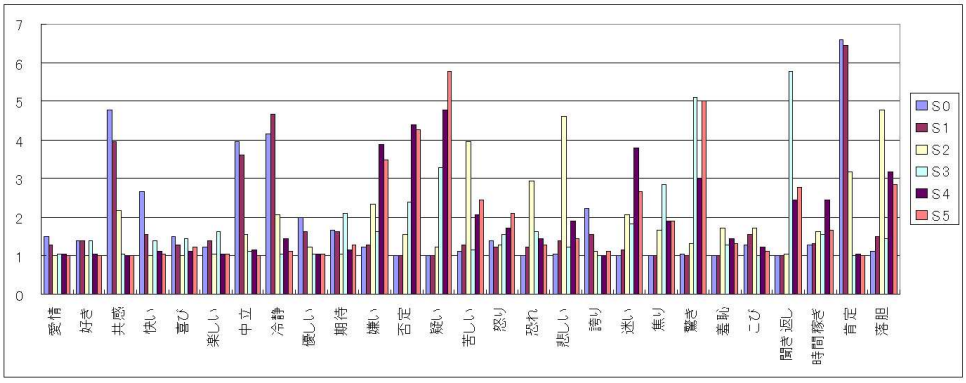

図 3.12: 各刺激音に対する感情語の評価の平均値

いてもランダムに呈示した．被験者は8名である．尚，実験に使用した機器は聴取実験1で用いたものと同じである．

3.3.3 実験結果，考察

各刺激音に対する感情語の評価を図3.12に示す．
評価の点数より，それぞれの刺激音に対する代表的な感情語を以下に示す．

S0：肯定，共感
S1：肯定，冷静
S2：落胆，悲しい
S3：聞き返し，驚き
S4：疑い，否定

S5：疑い，驚き

多次元尺度構成法による刺激音間の非類似性

脳活動測定では，基本周波数の時間変化の異なる合成音声から元音声との差を調べ，感情知覚との関係を議論する．この際，各合成音声と元音声との差を表す，その刺激音間の特性を知る必要がある．そのため，多次元尺度構成法を用いて，各刺激音の感情空間における付置を調べた．多次元尺度構成法による結果を図 3.13 に示す．アルゴリズムは ALSCAL を用い，距離行列は図 3.12 に示す評点を基に Minkowski の距離によって計算された．stress 値を表 3.4 に示す．

図 3.13 において，それぞれの刺激音の感情語の評価の対応関係より，それぞれの軸は，横軸が (正方向) 弛緩 - 緊張 (負方向)，縦軸が拒否 - 注目である Schlosberg の提案した 3 次元 [15] のうちの 2 次元 (もう 1 次元は快 - 不快である) に似た構造として表現できることが示唆された．各刺激音間の特性は，元音声 S0 との差において，S2 は拒否を，S3 では注目，緊張を，S4, S5 では緊張を表していることが示唆された．また，S1 は比較的付置が近いことが示唆された．

3.4　まとめ

本章では，脳活動測定で用いる合成音声の知覚特性，自然性を調べるために行った聴取実験について述べた．聴取実験 1 では，脳活動測定で用いる刺激音を選択するために，感情，自然性の評価を行った．クラスター分析を行い，知覚距離を求め，クラスターを 5 つに分類し，その中で自然性の高い合成音声を選択した．そのことにより，脳活動測定で用いる刺激音の条件に沿う，合成音声と元音声の 6 つの刺激音の準備が整った．聴取実験 2 では，6 つの刺激音において，その刺激音の感情の特性を調べた．結果それぞれの刺激音は S0｛肯定，共感｝，S1｛肯定，冷静｝，S2｛落胆，悲しい｝，S3 ｛聞き返し，驚き｝，S4｛疑い，否定｝，S5｛驚き，疑い｝の感情を含むことが示された．

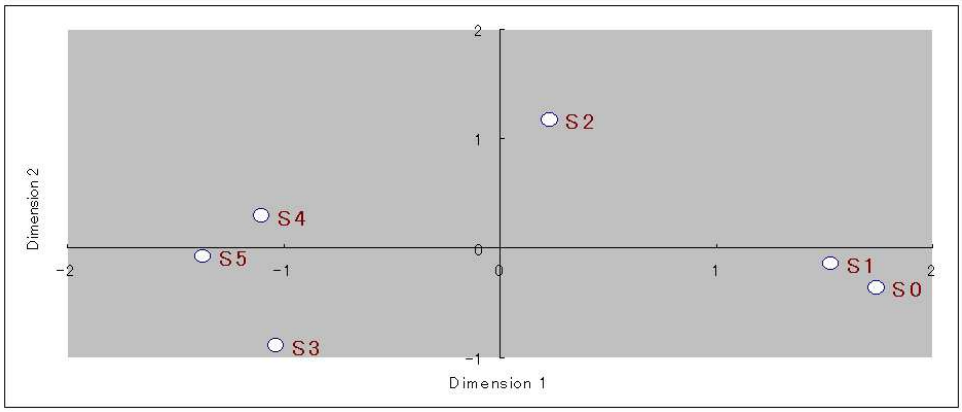

図 3.13: 多次元尺度法による刺激音間の非類似性

表 3.4: stress 値

Dimension	S-stress	Improvement
1	.05606	
2	.04951	.00655
3	.04911	.00040

第4章　感情音声に関する脳活動測定実験

4.1　はじめに

本研究では，音響的特徴の違いが脳にどのように符号化されるか，が重要である．本章では，前章で選択した感情音声に関する脳活動測定実験について述べる．実験の方法，解析方法について述べ，それぞれの刺激音との差における解析結果を示す．解析結果より，大脳皮質での活動，大脳辺縁系での活動，大脳基底核での，それぞれの活動差について考察を行う．

4.2　実験の目的

前章では，聴取実験によって，基本周波数の時間変化の異なる刺激音に対する感情知覚の特性が調べられた．基本周波数の時間変化と感情知覚の関連性を議論するために，基本周波数の時間変化が脳のどの部位を賦活させているかを調べる必要がある．脳活動測定を行うことにより，前章で選択した刺激音に対する脳の賦活部位を明らかにする．

4.3　実験方法

fMRIによる脳活動測定

近年，脳活動を測定する様々な装置が開発されてきている．中でも，functional magnetic resonance imaging(fMRI) は，

- 放射線被爆がない
- 空間分解能に優れる
- コントラスト分解能に優れる
- 非侵襲的である
- あらゆる角度の断面が撮像できる

などの長所があり，多くの研究で使用されている[16]．MR 装置を利用することで，刺激を与えた際に，脳内のどの部位が活動していたのか，その賦活部位を同定することができる．

実験装置

脳活動測定実験は，国際電気通信基礎技術研究所 (ATR) の脳活動イメージングセンタ (Brain Activity Imaging Center: BAIC) にて行った．主な実験装置を以下に記す．

MR 装置 ： Siemens 社製 MAGNETOM Trio, A Tim System (3 テスラ)

聴覚呈示装置 ： 日立アドバンストシステムズ社製 fMRI 装置用非磁性ヘッドセット

刺激呈示プログラム ： Neurobehavioral Systems 社製品 Presentation (ソフトウェア)

実験被験者

被験者は正常な聴力をもつ，右利きの日本人 18 名 (男性 10 名，女性 8 名) である．なお，右利きの条件は Edinburgh Inventory により決定された．

呈示方法

被験者には，撮像前に，この実験が感情音声に関するものであることを教示した．撮像中，被験者にはノイズ音が呈示された時にボタンを押してもらうというタスクを課した (オドボール課題)．1 セッションにつき各刺激音を 15 回ずつとノイズ音を 10 回呈示し，全部で 3 セッション行い，各刺激音は最適化された順番で呈示した．スキャンはスパース法 (TR = 4.0 s, TE = 6.4 ms) を用いた．スパース法とは，刺激音を呈示する際には撮像を行わず，刺激音呈示の後から撮像を行う方法である．このことにより，撮像する際に発生する実験装置によるノイズに影響されることなく，刺激音を聴くことができる．

4.4 解析方法

画像の解析は Statistical Parametric Mapping 5(SPM5) を用いた．脳画像は 30 スライスでボクセルの解像度は $3.0 \times 3.0 \times 4.0$ mm である．前処理は機能画像の位置補正，標準脳への正規化，ガウスフィルタ (FWHM = 8 mm) による平滑化を行った．統計解析は血流動態反応関数を畳み込んだ関数を使った一般線形モデルを用いた．脳活動には様々な要因が影響するため，2 つの刺激音を呈示した際の脳活動の差をとり，解析の結果を得る．そのことにより，特有の課題に対する活動差を調べることができる．各被験者内の各課題

に対する活動のコントラスト画像を求め，集団解析を行い，個人解析で得られた各被験者のコントラスト画像をt検定で検定を行った．賦活部位の同定は，Talairach 座標系 [17] に基づいて行った．

4.5 解析結果

5つの合成音声を聴いた際の脳活動それぞれから元音声を聴いた際の脳活動の差をとり，解析結果を得た．それぞれの結果を図 4.1 ～ 4.10 に示す．また，表 4.1 ～ 4.10 に活動領域の詳細を示す．それぞれの脳活動差の p 値は，S1 - S0，S0 - S1，S0 - S4，S0 - S5 で ($p < 0.01$ uncorrect)，S2 - S0, S0 - S2, S3 - S0, S0 - S3 で ($p < 0.001$ uncorrect.)，S4 - S0, S5 - S0 で ($p < 0.05$ FWE.) で検定された (FWE = Family - Wise Error)．それぞれの検定の厳しさは，($p < 0.01$ unc.) < ($p < 0.001$ unc.) < ($P < 0.05$ FWE.) である．

表 4.1: S1 - S0 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S1 - S0 (p< 0.01 unc.)				
SFG	R	14 18 58	3.13	30
FMG	L	-28 48 0	3.06	49
IFGTr	L	-50 26 6	2.97	52
AOI	L	-30 16 -2	2.94	33

Note. SFG, superior frontal gyrus; FMG, frontomarginal gyrus; IFGTr, inferior frontal gyrus pars triangularis; AOI, aria orbitoinsularis.

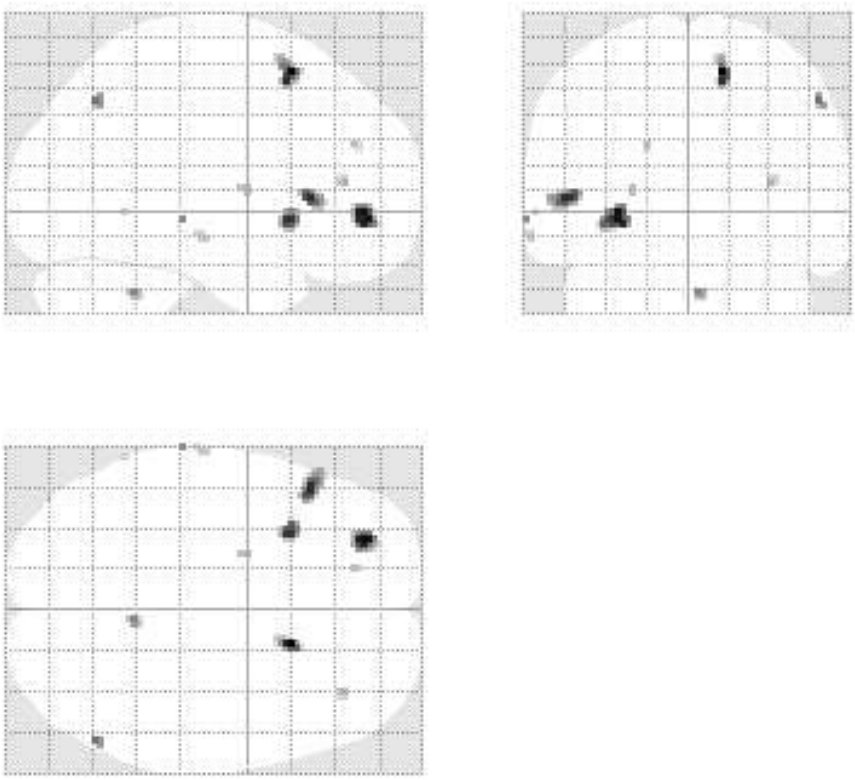

図 4.1: S1 - S0 の脳活動の透し図 (左上：側面図，左下：上面図，右上：前面図であり，黒い部分が有意な活動差を示す．)

表 4.2: S0 - S1 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S0 - S1 ($p < 0.01$ unc.)				
MPul	R	16 -32 6	4.41	184
CG		0 -20 24	3.81	904
bfx	L	-2 -26 16	3.5	
LPul	L	-16 -36 2	3.44	
POrG	L	-28 30 -6	3.8	135
MOrG	L	-20 36 -10	3.59	

Note. MPul, medial pulvinar; CG, cingulate gyrus; bfx, body of fornix; LPul, left pulvinar; POrG, posterior orbital gyrus; MOrG, midial orbital gyrus; LG, lateral geniculate nucleus; SMG, supramarginal gyrus.

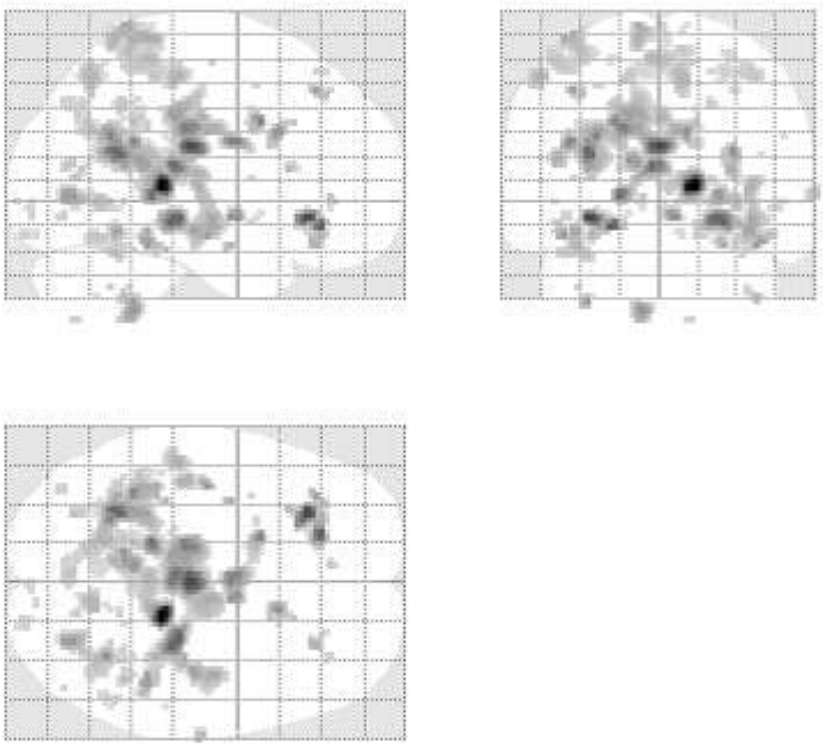

図 4.2: S0 - S1 の脳活動の透し図

表 4.3: S2 - S0 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S2 - S0 ($p < 0.001$ unc.)				
SPL	L	-18 -78 54	4.51	391
AnG	L	-28 -82 48	4.27	
SMG	L	-46 -42 48	3.82	43
STG	L	-68 -36 10	3.76	28
STG	R	68 -34 10	3.51	6
CG	L	-16 36 4	3.25	2

Note. SPL, superior parietal lobule; AnG, angular gyrus; SMG, supramarginal gyrus; STG, superior temporal gyrus; CG, cingulate gyrus.

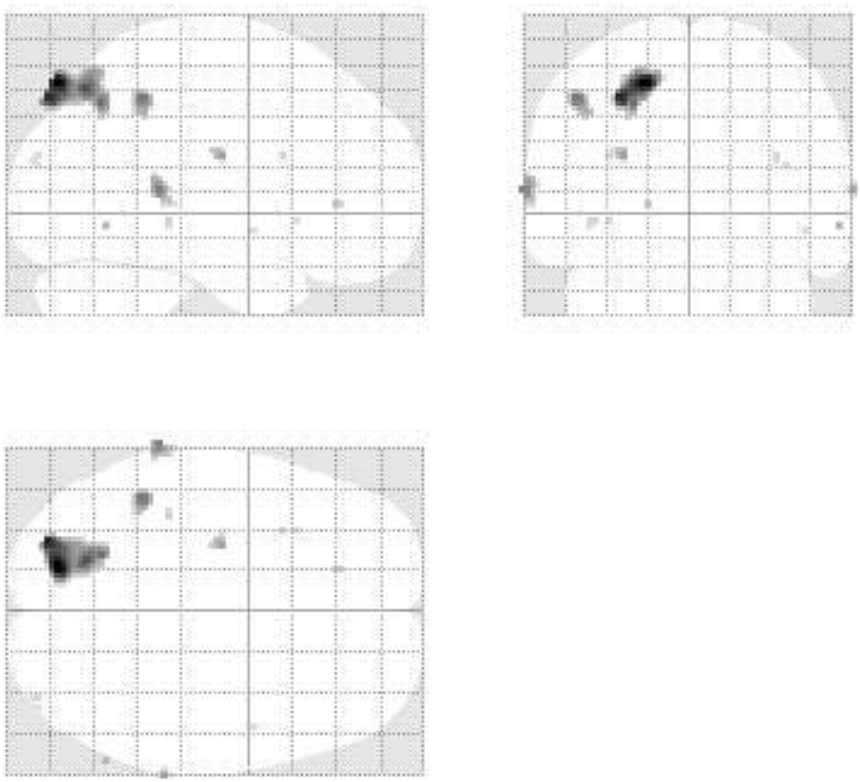

図 4.3: S2 - S0 の脳活動の透し図

表 4.4: S0 - S2 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S0 - S2 ($p < 0.001$ unc.)				
SFGL	L	-32 34 50	3.43	3
MTG	R	46 -66 24	3.26	4
PCun	R	8 -58 10	3.24	3
SFGL	R	26 46 22	3.14	1

Note. SFGL, superior frontal gyrus lateral part; MTG, middle temporal gyrus; PCun, precuneus.

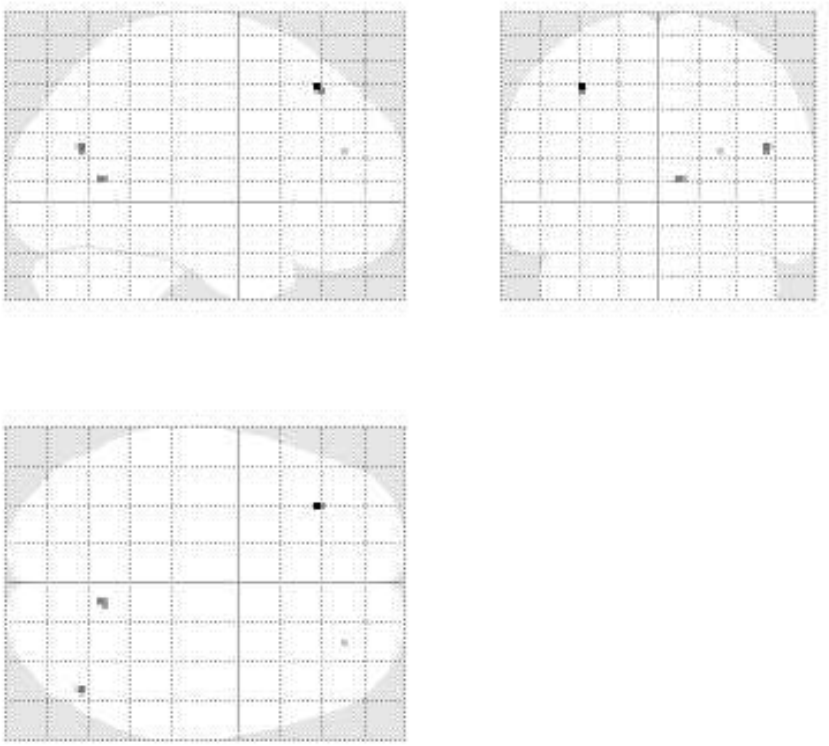

図 4.4: S0 - S2 の脳活動の透し図

表 4.5: S3 - S0 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S3 - S0 ($p < 0.001$ unc.)				
Cd	L	-24 -18 28	3.8	89
SFG	L	-22 -14 38	3.54	
SMG	R	34 -40 26	3.21	6
SMG	L	-34 -34 32	3.14	1

Note. Cd, caudate nucleus; SFG, superior frontal gyrus; SMG, supramarginal gyrus.

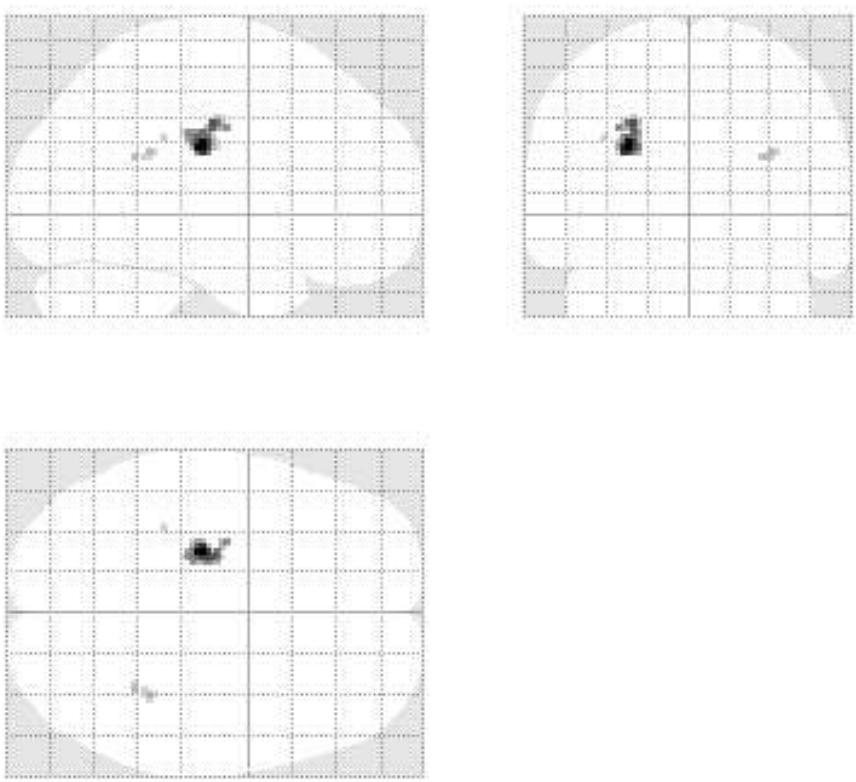

図 4.5: S3 - S0 の脳活動の透し図

表 4.6: S0 - S3 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S0 - S3 ($p < 0.001$ unc.)				
cerebullum	L	-6 -68 -46	4.18	100
Ins	L	-34 -16 6	4.05	50
cerebellum	R	38 -46 -36	3.57	18
acs	R	12 -50 8	3.53	69
SFGL	L	-34 34 48	3.34	15

Note. Ins, insula; acs, anterior calcarine sulcus; SFGL, superior frontal gyrus lateral part.

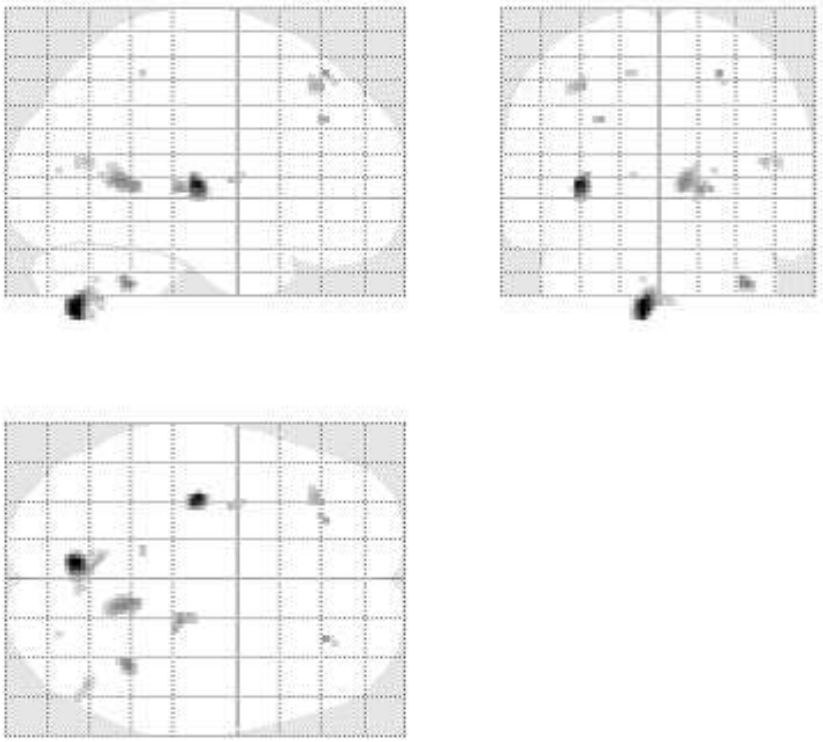

図 4.6: S0 - S3 の脳活動の透し図

表 4.7: S4 - S0 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S4 - S0 ($p < 0.05$ FWE.)				
cerebellum	L	-24 -70 -50	6.07	48
API	L	-30 12 -6	6.06	103
STG	R	48 -46 20	5.99	362
SFG	L	-22 18 18	5.71	37
Pu	L	-24 16 10	5.52	30
SPL	R	38 -48 44	5.67	28
SMG	R	42 -52 52	5.61	
cerebellum	R	12 -50 -16	5.56	125
IG	R	42 12 0	5.5	180

Note. API, area piriformis insulae; STG, superior temporal gyrus; SFG, superior frontal lobule; Pu, putamen, SPL, superior parietal lobule; SMG, supramarginal gyrus; IG, insular gyrus.

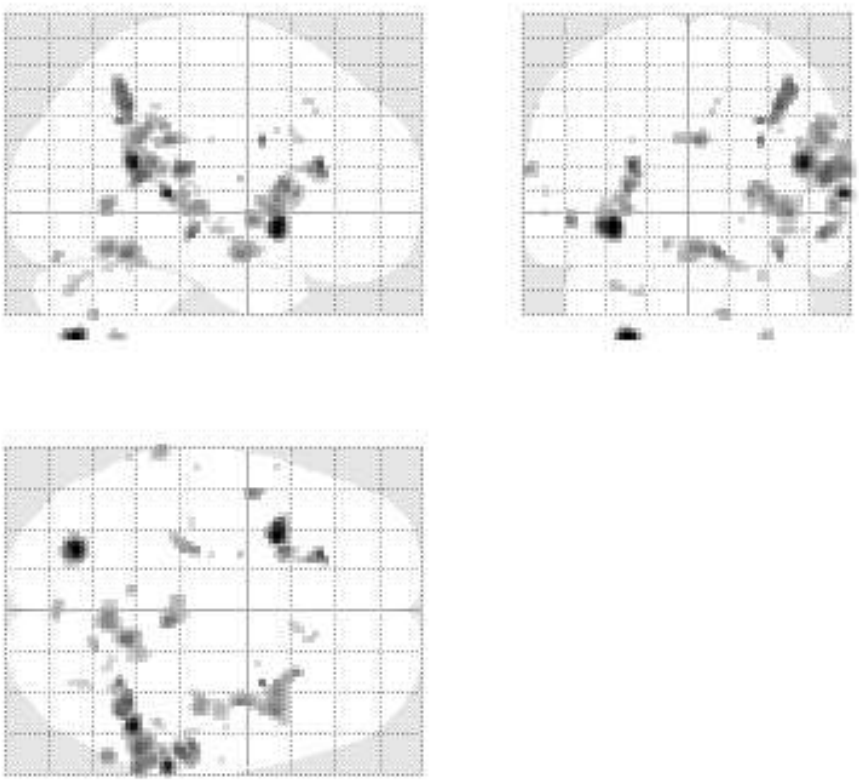

図 4.7: S4 - S0 の脳活動の透し図

表 4.8: S0 - S4 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S0 - S4 ($p < 0.01$ unc.)				
PrG	R	52 -10 36	2.65	70
PrG	L	-58 -4 30	2.64	37
PoG	R	50 -18 60	2.52	9

Note. PrG, precentral gyrus; PoG, postcentral gyrus.

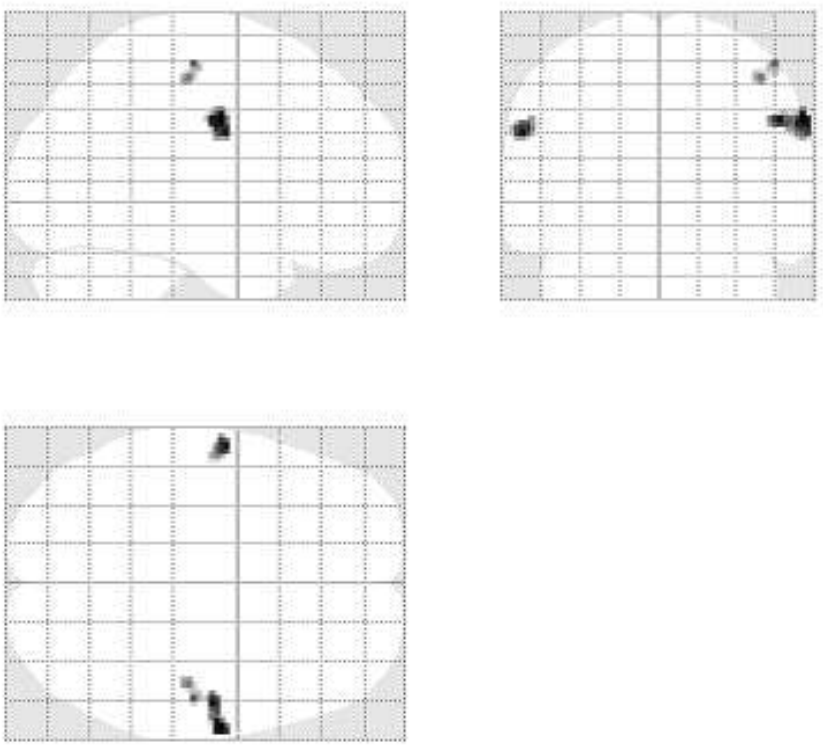

図 4.8: S0 - S4 の脳活動の透し図

表 4.9: S5 - S0 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S5 - S0 ($p < 0.05$ FWE.)				
Pu	L	-24 -26 8	6.32	77
Ins	R	38 -18 0	6.07	157
cerebellum	R	36 -52 -50	5.89	22
cerebellum	L	-24 -70 -32	5.79	54
BOp	R	32 18 -2	5.65	133
IG	R	38 12 0	5.25	
MTG	R	62 -56 0	5.53	25
STG	R	60 -38 20	5.49	163

Note. Pu, putamen; Ins, insula; BOp, basal operculum; IG, insular gyrus; MTG, middle temporal gyrus; STG, superior temporal gyrus.

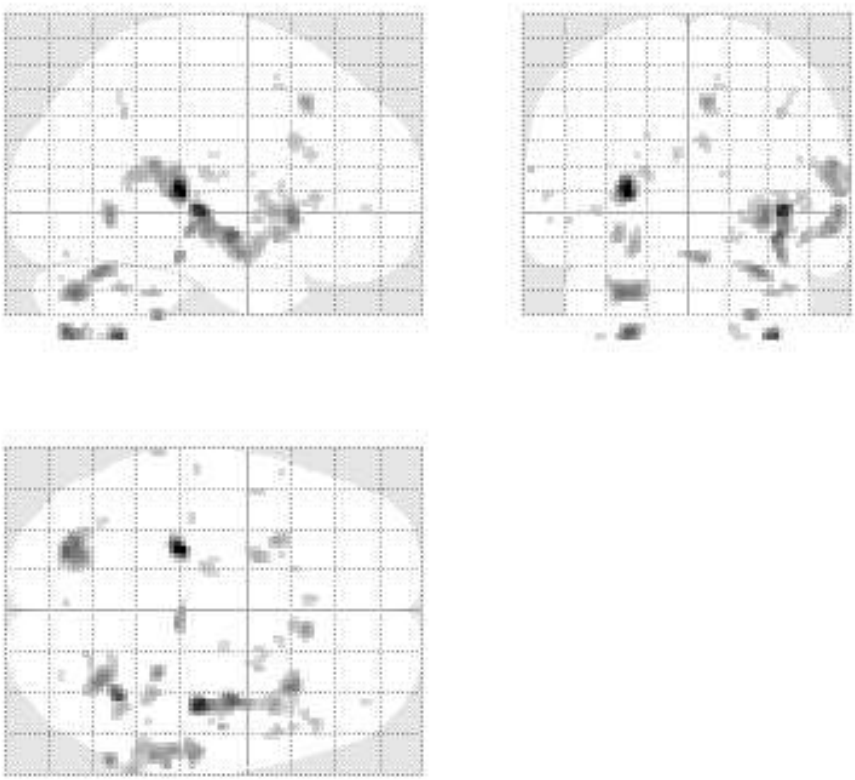

図 4.9: S5 - S0 の脳活動の透し図

表 4.10: S0 - S5 における有意な脳活動部位

Brain region	Hemisphere	MNI Coordinates	Zvalue	Cluster size
S0 - S5 ($p < 0.01$ unc.)				
PrG	R	64 -4 32	3.37	234
OcG	R	44 -70 -6	3.14	91
PrG	L	-58 -4 30	3.03	168
OcG	L	-14 -54 14	2.79	23
FuG	R	28 -52 -8	2.68	25
PoG	R	50 -18 60	2.67	23
ITG	L	-38 -78 -8	2.5	10

Note. PrG, precentral gyrus; OcG, occipital gyri; FuG, fusiform gyrus; PoG, postcentral gyrus; ITG, inferior temporal gyrus.

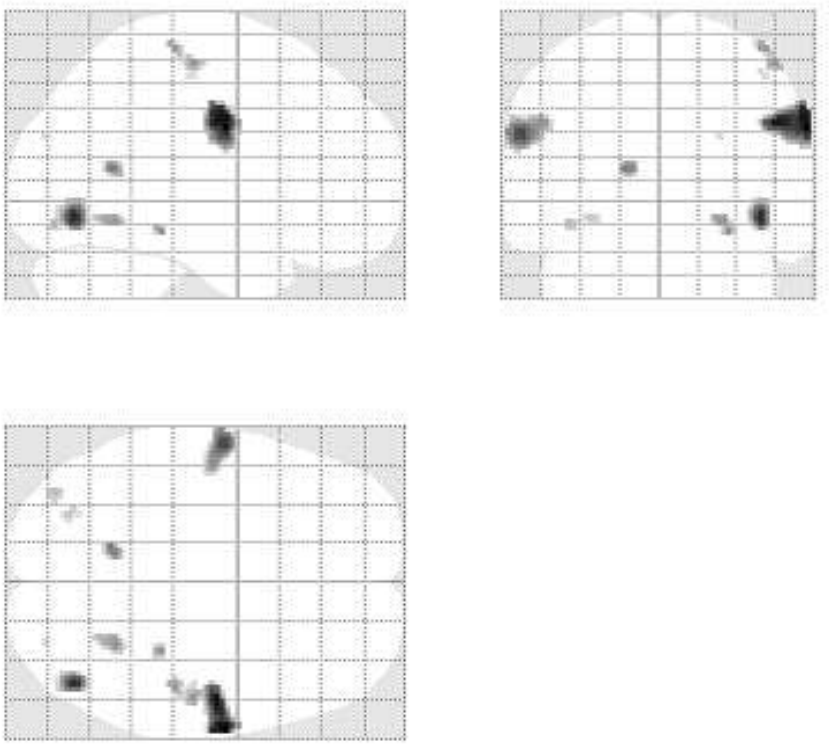

図 4.10: S0 - S5 の脳活動の透し図

S1 - S0における脳活動

S1 - S0 ($p < 0.01$ unc.) では，SFG(上前頭回)，FMG(前頭縁回)，IFG(下前頭回)，AOI などで主な特徴がみられた．

SFG や FMG, IFG などは前頭葉に属する部位である．前頭葉の損傷は認知や情動，運動などの異常を引き起こすことが知られている [18]．また，大脳辺縁系の活動を制御している，などの知見がある [19]．ブローカ野に含まれる左半球の IFG は発話に関して重要な部位だとされており，単語の生成や文理解の統語処理に関係している，などの知見もある [20] [21]．また，感情音声においても，韻律の処理として重要であることが示唆されている部位でもある [10]．

S0 - S1における脳活動

S0 - S1($p < 0.01$ unc.) では，Pul(視床枕核)，CG(帯状回)，bfx(脳弓体)，OrG(眼窩回)，LG(外側膝状核) などで主な活動が見られた．

S2 - S0における脳活動

S2 - S0($p < 0.001$ unc.) では，SPL(上頭頂小葉)，AnG(角回)，SMG(縁上回)，STG(上側頭回)，CG(帯状回) で主な特徴が見られた．

SPL は体性感覚野と頭頂連合野に関係する部位である．AnG は上側頭溝の上行枝の後縁を弓状に取り囲む脳回である．また，SMG は外側溝の上行枝の後縁を弓状に取り囲む脳回である．両半球でみられた STG は一次聴覚野に含まれ，音情報処理における様々な報告がなされている [22] [23] [24]．大脳辺縁系に含まれる帯状回は，感情や情動に関与していることが多くの研究から示唆されている [19] [25]．

S0 - S2における脳活動

S0 - S2($P < 0.001$ unc.) では，SFGL(上前頭回)，MTG(中側頭回)，PCun(楔前部) で主な活動が見られた．

S3 - S0における脳活動

S3 - S0($p < 0.001$ unc.) では，Cd(尾状核)，SFG(上前頭回)，SMG(縁上回) で主な特徴が見られた．

Cd は大脳基底核に含まれ，認識による運動に関係することが知られている [26]．

S0 - S3 における脳活動

S0 - S3 ($p < 0.001$ unc.) では，Ins(島葉)，acs(鳥距溝)，SFGL(上前頭回) で主な特徴が見られた．

S4 - S0 における脳活動

S4 - S0($p < 0.05$ FWE.) では，API(梨状島領域)，STG(上側頭回)，SFG(上前頭回)，Pu(被核)，SPL(上頭頂小葉)，SMG(縁上回)，cerebellum(小脳)，IG(島回) で主な活動が見られた．

Pu は大脳基底核に含まれ，運動や調音に関係することが知られている [27]．また，音声知覚において活動し，感情や情動において重要であることが報告されている [28] [29]．小脳は，運動調整機能を担うとされ，調音動作に関係することが報告されている [27]．

S0 - S4 における脳活動

S0 - S4($p < 0.01$ unc.) では，両半球の PrG(中心前回)，PoG(中心後回)，で主な特徴が見られた．

S5 - S0 における脳活動

S5 - S0($p < 0.05$ FWE.) では，Pu(被核)，Ins(島)，cerebellum(小脳)，BOp，IG(島回)，MTG(中側頭回)，STG(上側頭回) などで主な活動が見られた．

Ins や IG では，音声生成に関与するという知見や，左半球においてブローカ失語症に関係するなどの報告があり，DIVA Model においては，音声調音で重要とされている [30] [31] [32]．

S0 - S5 における脳活動

S0 - S5($p < 0.01$ unc.) では，PrG(中心前回)，OcG(後頭回)，FuG(紡錘状回)，PoG(中心後回)，ITG(下側頭回) で主な特徴が見られた．

4.6 脳活動測定実験結果の考察

神経科学や臨床の多くの研究により，脳の機能は調べられてきている．各刺激音における感情の機能がどの領域に属しているか調べるために，各刺激音の大脳皮質，大脳辺縁系，大脳基底核における活動を整理し，その機能について考察する．

第3章において，各刺激音の有する代表的な感情は，S0｛肯定，共感｝，S1｛肯定，冷静｝，S2｛落胆，悲しい｝，S3｛聞き返し，驚き｝，S4｛疑い，否定｝，S5｛驚き，疑い｝であった．

大脳皮質での活動

前頭葉の活動は，S1 - S0, S3 - S0, S4 - S0, S0 - S2, S0 - S3 において，SFG(上前頭回) や FMG(前頭縁回) の部位の活動差が見られた．前頭葉の損傷は認知や情動，運動などの異常を引き起こすことが知られている [18]．また，大脳辺縁系の活動を制御する [19]．などの知見がある．SFG や FMG の部位は感情を判別する上で，重要である可能性が高い．

また，感情や情動に関係すると言われている下前頭回，眼窩回での活動は，S1 - S0, S0 - S1 で左半球の IFGTr(下前頭回三角部), POrG, MOrG(眼窩回) で見られた．左半球や下前頭回や眼窩回はブローカ野に属しており，発話などに関して重要であると考えられている．感情音声の韻律情報の処理として，左の下前頭回が重要であるとの報告もある [10]．また，歌声と話声の差において，歌声で眼窩回が重要である，という報告もある [33]．S0 と S1 は比較的感情空間における付置が近く，S1 と S0 との差にのみ活動が見られたため，これらの部位は，より高等な感情の判断をしている可能性もあるかもしれない．

一次聴覚野付近の活動としては，両半球の STG(上側頭回) の活動が S2 - S0 で，右半球の STG(上側頭回) での活動が S4 - S0, S5 - S0 で見られた．感情や情動の処理として，右半球の STG(上側頭回) が重要であるという報告もあるが，継続時間長や音の大きさの影響がより強いという報告 [9] や，ピッチの変化の処理を行うなどの報告がある [7]．STG における活動は，音情報に関する初期の段階での処理をしているという報告もあり [34]，感情の差というよりは，音刺激全体の処理の差を表していると考えられる．

S4 - S0, S5 - S0, S0 - S3 においては，島皮質における Ins(島葉) や IG(島回)，API(梨状島領域) で活動がみられた．Ins や IG では，音声生成に関与するという知見や，左半球においてブローカ失語症に関係するなどの報告がある．また，DIVA Model においては，音声調音で重要とされている [30] [31] [32]．Ins が嫌悪の表情において賦活したという報告もある [35]．S4 や S5 において，｛嫌い｝における評価も高いことが示されている (図 3.12) が，S0 - S3 においても活動が見られるため，Ins が感情間の差として重要であるかは明確には示されなかった．

SMG(縁状回) の活動が S2 - S0, S3 - S0, S4 - S0 においておいてみられた．S2 - S0 においては，AnG(角回) での特徴もみられた．AnG は視覚，聴覚，触覚情報の統合と読字のシンボル統合をすることが知られている [36]．また，SMG や AnG の損傷において，失読や失書症の機能が損なわれるという報告もなされている [37]．

SPL(上頭頂小葉) の活動が S2 - S0，S4 - S0 で見られた．SPL は，特に触覚-視覚相関に関係することが知られている [38]．眼球運動の機能として，頭頂葉が感覚運動の変換として重要であるという報告 [39] や，音の位置の判断にとって重要であるという報告もある [40]．

大脳辺縁系での活動

amygdara(扁桃体) は，感情表現のトップダウンのコントロールとして重要であるという報告や [41]，恐怖における刺激に対して賦活している [42] などの報告があるが，今回の測定では，主な賦活は見られなかった．

また，帯状回の活動は，情動の処理として重要であると言われており [19]，今回の測定では，S0 - S1，S2 - S0 で見られたが，感情間の差として，どのような特徴の差であるかは判明しなかった．

大脳基底核での活動

大脳基底核に属する左半球の Cd(尾状核), Pu(被核) の活動が S3 - S0, S4 - S0, S5 - S0 において見られた．左半球の大脳基底核における活動は，感情や情動において重要であるという報告もある [28] [29].

全体の活動差

感情間の活動の違いを議論するために，以上の結果より，感情の差を表すと考えられるそれぞれの特徴を示す．

それぞれの活動差より，S1｛肯定，冷静｝では，大脳皮質に属する前頭葉 (SFG(上前頭回) など) が感情間の差として重要であり，また，S2｛落胆，悲しい｝では，大脳皮質における SPL(上頭頂小葉) や AnG(角回) で重要である可能性が高いことが考えられる．また，S3｛聞き返し，驚き｝では，大脳基底核に属する Cd(尾状核) で，S4｛疑い，否定｝や S5｛驚き，疑い｝では，大脳基底核に属する Pu(被核) などにおいて感情間の差異が表われている可能性があると考えられる．

4.7　まとめ

本章では，感情音声に関する脳活動測定について述べた．感情間の差は，S1｛肯定，冷静｝において前頭葉で，S2｛落胆，悲しい｝で SPL(上頭頂小葉) や AnG(角回) で，S3｛聞き返し，驚き｝や S4｛疑い，否定｝，S5｛驚き，疑い｝において大脳基底核に属する Cd(尾状核) や Pu(被核) において重要である可能性が示された．

第5章　脳活動測定結果と感情知覚に関する考察

5.1　はじめに

本研究では，どのような音響的特徴の違いが，どの脳部位を賦活させ，感情の差異として知覚しているか，という問題に対して，基本周波数の時間変化に着目し，研究を行った．第2章において，音響的特徴について述べ，第3章において，聴取実験によりその音響的特徴と感情知覚の関係を調べた．第4章において，脳活動測定により，刺激音に対する脳活動部位の同定を行った．本章では，感情階層説を参考することにより，これらの聴取実験の結果と賦活部位の関係を考察する．

5.2　聴取実験の結果と感情階層説との関係

各刺激音が感情階層説におけるどの階層に属するのか，それぞれの感情語に対する福田の感情階層説との対応関係を調べる．

S0における代表的な感情語は,｛肯定，共感｝であり，これらの感情は，社会的，知的感情の分類に当てはまる．また，S1における｛肯定，冷静｝も同様に，社会的，知的感情の分類に当てはまる．

S2にみられる｛落胆，悲しい｝における｛悲しみ｝に対する分類は未だ議論がなされている．表情を用いた心理学からの分類では，EkmanやIzardによる研究などによる多くの研究において，悲しみは基本情動に分類されている[43] [44]．また，Shaverなどによる研究においても基本情動(情動の基礎カテゴリー)に分類されている[45]．しかし，福田の提唱する感情階層説では，感情は情動に主観的体験が加わったものであると考え,｛悲しみ｝を社会的，知的感情であるとしている．

S3における代表的な感情語は｛聞き返し，驚き｝であった.｛悲しみ｝における分類と同様に,｛驚き｝においても，基本情動に属するとされる報告が多くあるが，感情階層説では,｛驚き｝は注意覚醒系の機能であり，情動とは別系統として進化してきているとされている．その進化の延長として，興奮，覚醒，注意，興味，好奇心などの機能があり，｛聞き返し｝もその系統にあると考えられる．

S4における代表的な感情語は｛疑い，否定｝であり，これらは，注意の操作，および読心の機能などにあたり，社会的，知的感情の分類に当てはまると考えられる．

S5 における代表的な感情語は｛驚き，疑い｝であった.｛驚き｝は S3 における代表的な感情語にもみられ，注意覚醒系の系統にあると考えられる．また,｛疑い｝は社会的，知的感情に分類されると考えられる.

感情階層説と各刺激音の分類を図 5.4 に示す [1].

5.3 脳活動測定による感情知覚に関する考察

感情階層説を参考に，聴取実験の結果と脳活動測定の結果における感情知覚の対応関係について考察する.

刺激音 S1 は主に社会的感情に分類される｛肯定，冷静｝の感情を含み，元音声 S0｛肯定，共感｝との差は，SFG(上前頭回) などの大脳皮質の前頭葉において主な活動差が見られた．大脳皮質は社会的感情や知的感情における処理とともに発達してきたと考えられ，結果における社会的感情の差異は，感情階層説における知見と一致する.

刺激音 S2 は主に社会的感情に分類される｛落胆，悲しみ｝を含み，元音声 S0 との差は，大脳皮質における，SPL(上頭頂小葉) や AnG(角回) に主な活動差がみられた．この結果は,｛落胆，悲しみ｝が社会的，または知的感情における処理として，触覚や視覚，聴覚を統合する様な部位に関わってくる可能性を示唆している.

刺激音 S3 は主に注意覚醒系の機能を持つ｛聞き返し，驚き｝を含み，元音声 S0 との差は大脳基底核における Cd(尾状核) において特徴がみられた．また，刺激音 S5 も同様に注意覚醒系の機能を持つ驚きの感情を含み，元音声 S0 との差は大脳基底核における Pu(被核) において主な活動がみられた．S4 においても，Pu における活動が見られたが，S5 と感情空間における付置が近く (図 3.13)，また，代表的な感情語には示されなかったが,｛驚き｝に対する評価も S0 と比べ高い (図 3.12) ことが示されている．大脳基底核における活動は，原始情動の処理をすると考えられており，身体の維持などの行動の調整システムを担っている．今後更なる知見を要するが，S3 や S4, S5 における注意覚醒系では，原始情動を扱うこのようなシステムに関わってくる可能性があるかもしれない.

5.4 まとめ

本章では，感情階層説を参考に，聴取実験の結果と賦活部位の関係を明らかにした.｛肯定，冷静｝の感情 (S1) は，社会的感情として，大脳皮質における前頭葉に関係し,｛落胆，悲しみ｝の感情 (S2) は，社会的感情として，大脳皮質において処理されている可能性が示唆された．また,｛驚き｝を含む注意覚醒系 (S3, S4, S5) では，原始情動を扱う，大脳基底核における処理に関わってくる可能性が示唆された.

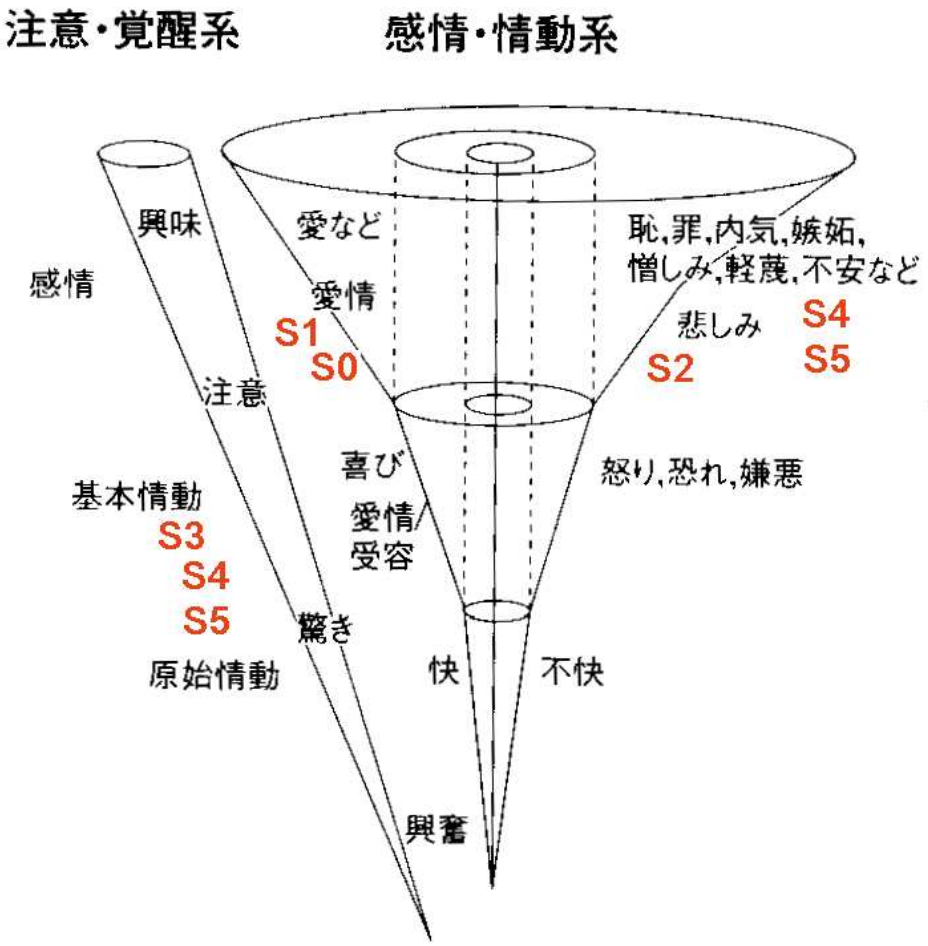

図 5.1: 感情階層説と各刺激音の分類．ref. 福田 正治「感情学入門」[1]

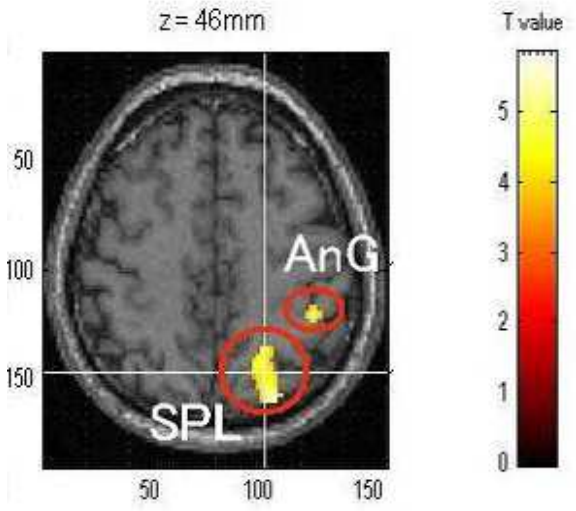

図 5.2: S2 - S0 における主な賦活部位

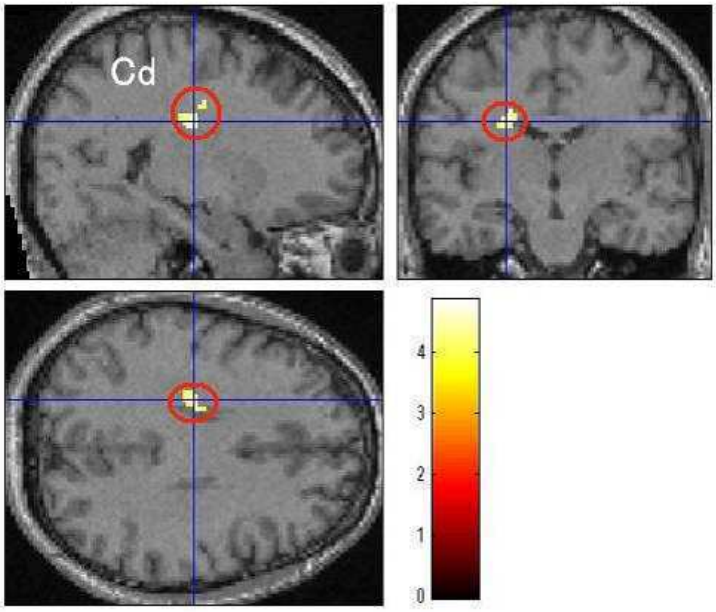

図 5.3: S3 - S0 における主な賦活部位

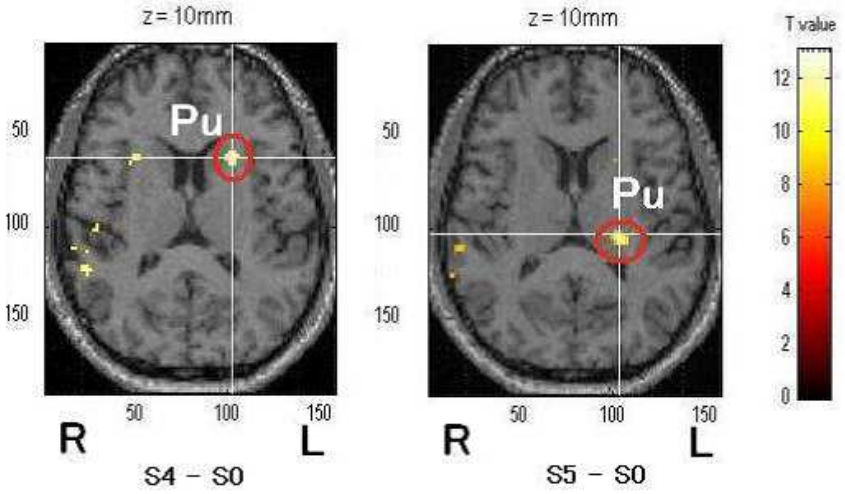

図 5.4: S4 - S0, S5 - S0 における大脳基底核における賦活

第6章　結論

6.1　本研究で明らかになったことの要約

本研究では，感情階層説を参考に，fMRI を用いて感動詞「ええ」を呈示した際の脳活動を測定し，韻律情報に対応する感情知覚と脳活動の関係を調べた．

｛肯定，共感｝を表す刺激音は，社会的感情を処理する大脳皮質に属する前頭葉で,｛悲しい，落胆｝を表す刺激音は，社会的感情を処理する大脳皮質に属する上頭頂小葉や角回で特徴がみられた.｛驚き｝などの注意覚醒系を表す刺激音は，原始情動を処理するとされる大脳基底核に属する尾状核や被核で主な特徴がみられた．これらの感情は異なる韻律情報によって引き起こされており，異なる脳領域，異なる感情の階層に渡って処理されていることが示唆された．

6.2　今後の課題

今回の実験では，基本周波数の時間変化に対応する音響的特徴の制御を行い，感情知覚との関連性を調べた．今後，感情知覚に密接に関係する韻律情報全体の制御を行った際の脳活動の賦活，感情知覚について詳しく調べていく必要がある．また，感情知覚には韻律情報以外に，声質の影響も重要であるという指摘もあり，これらを総合して調べていく必要もあるかもしれない．さらに，音声知覚の機構を解明するためには，感情に限らず，非言語情報処理一般において，本研究で用いたストラテジーと同じ様に，音響的特徴と脳活動を調べ，音声知覚における知見を得ていく必要がある．

謝辞

本研究を遂行するにあたり，多大なる御指導ならびに御鞭撻賜りました，北陸先端科学技術大学院大学 情報科学研究科 赤木正人 教授に深く感謝致します.

本研究を遂行するにあたり，多くの貴重な御助言を賜りました，甲南大学 知能情報学部 北村達也 准教授に心より感謝致します.

本研究を遂行するにあたり，貴重な御助言，御指導を賜りました，北陸先端科学技術大学院大学 情報科学研究科 鵜木祐史 准教授，李軍鋒 助教，宮内亮太 助教に心より感謝致します.

研究発表にあたり，熱心な議論並びに貴重な御助言を賜りました，北陸先端科学技術大学院大学 情報科学研究科 党健武 教授，徳田功 准教授に心より感謝致します.

また，日頃の研究生活において，多くの御協力，御討論をいただきました音情報処理学講座の皆様に厚く御礼申し上げます.

最後に，大学院での貴重な研究生活を与えてくれた両親に心から感謝し御礼を申し上げます.

参考文献

[1] 福田正治：“感情を知る -感情学入門-”, ナカニシヤ出版 (2003).

[2] P. B. Denes and E. N. Pinson: “The speech chain: The physics and biology of spoken language 2nd ed.” (1993).

[3] 福田正治：“感情の階層性と脳の進化：社会的感情の進化的位置づけ”, 感情心理学研究, **16**, 1, pp. 25–35 (2008).

[4] 森山剛, 森真也, 小沢慎治：“韻律の部分空間を用いた感情音声合成”, 情報処理学会論文誌, **50**, 3, pp. 1181–1191 (2009).

[5] 刀根優子, 萩原昭夫, 柴田浩：“音声対話システムのための hmm に基づく感情判別”, 電子情報通信学会, **22**, pp. 47–53 (SP2000).

[6] 林康子：“感動詞「ええ」におけるピッチ曲線と感情認知”, 電子情報通信学会技術研究報告, **98**, 177, pp. 65–72 (1998).

[7] 塚田裕樹, 入野俊夫, 大屋義和, R. D.Patterson, 河原英紀：“音声からの寸法情報処理の脳内部位の fmri による検討”, 日本音響学会講演論文集, pp. 571–572 (2009).

[8] B. Tillmann, N. E. S.Koelsch, P. L. E. Bigand, A. Friederici and D. von Cramon: “Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex”, NeuroImage, **31**, pp. 1771–1782 (2006).

[9] S. Wiethoff, D. Wildgruber, B. Kreifelts, H. Becder, C. Herbert, W. Grodd and T. Ethofer: “Cerebral processing of emotional prosody-influence of acoustic parameters and arousal”, NeuroImage, **39**, 2, pp. 885–893 (2008).

[10] D. R.Bach, D. Grandjean, D. Sander, M. Herdener, W. K.Strik and E. Seifritz: “The effect of appraisal level on processing of emotional prosody in meaningless speech”, NeuroImage, **42**, 2, pp. 919–927 (2008).

[11] 河原英紀：“聴覚の情景分析が生み出した高品質 vocoder: Straight”, 日本音響学会誌 54 巻 7 号, pp. 521–526 (1998).

[12] H. Kawahara, I. Masuda-Katsuse and A. de Cheveigne: “Restructuring speech representations using a pitch adaptive time-frequency smoothing and an instantaneous-frequency-based f0 extraction: Possible role of a repetitive structure in sounds”, Speech communication, pp. 187–207 (1999).

[13] 本多清志：“音声の生物学的基礎”, 音声 (岩波書店), **3 章**, p. 104 (2004).

[14] 石井カルロス寿憲, 石黒浩, 萩田紀博：“韻律および声質を表現した音響特徴と対話音声におけるパラ言語情報の知覚との関連”, 情報処理学会論文誌, **47**, 6, pp. 1782–1793 (2006).

[15] 福井康之：“感情の心理学”, 川島書店 (1990).

[16] 月本洋, 妹尾淳史, 渡邊修, 菊地吉晃, 安保雅博, 米本恭三：“脳機能画像解析入門-spm で fmri，拡散テンソルを使いこなす”, 医歯薬出版株式会社 (2007).

[17] J. K. Mai, G. Paxinos and T. Voss: “Atlas of the human brain”, ACADEMIC PRESS (2008).

[18] 甘利俊一, 外山敬介：“脳科学大事典”, 朝倉書店 (2000).

[19] 小野武年, 西条寿夫：“感情と知的情報処理の仕組み”, 高次脳機能研究, **25**, 2, pp. 116–128 (2005).

[20] D. J. Sharp, S. K. Scott, A. Cutler and R. J. Wise: “Lexical retrieval constrained by sound structure: The role of the left inferior frontal gyrus”, Brain and Language, **92**, pp. 309–319 (2005).

[21] 酒井邦嘉：“言葉の脳内処理機構”, 高次脳機能研究, **25**, pp. 153–164 (2005).

[22] F. Pulvermuller, M. Huss, F. Kherif, F. M. del Prado Martin, O. Hauk and Y. Shtyrov: “Motor cortex maps articulatory features of speech sounds”, Neuroscience, **103**, 20, pp. 7865–7870 (2006).

[23] S. Koelsch and W. A. Siebel: “Towards a neural basis of music perception”, Cognitive Sciences, **9**, 12, pp. 578–584 (2005).

[24] D. E. Callan, V. Tsytsarev, T. Hanakawa, A. M. Callan, M. Katsuhara, H. Fukuyama and R. Turner: “Song and speech: Brain regions involved with perception and covert production”, NeuroImage, **31**, pp. 1327–1342 (2006).

[25] 中村克樹：“情動情報の脳内処理”, 日本音響学会講演論文集, pp. 485–486 (2009).

[26] N. R. Carlson: “Physiology of behavior”, 丸善 (2008). (中村克樹，秦羅雅登 監訳. 神経科学テキスト -脳と行動-).

[27] 能田由紀子, 本多清志, 島田育廣, 藤本一郎：“fmriにより観測した音声生成時の大脳・小脳・脳幹の活動”, 日本音響学会講演論文集, pp. 451–452 (2001).

[28] S. Paulmann, M. D. Pell and S. A. Kotz: “Functional contributions of the basal ganglia to emotional prosody: Evidence from erps”, BRAIN RESEARCH, **1217**, pp. 171–178 (2008).

[29] S. A. Kotz, M. Meyer, K. Alter, M. Besson, D. Y. von Cramon, A. D. Friederici：“On the lateralization of emotional prosody: An event=related functional mr investigation”, Brain and Language, **86**, pp. 366–376 (2003).

[30] 辰巳格, 伏見貴夫, 佐久間尚子, 品川英朗, 石井賢二：“単語を想起中の脳の賦活部位, 特に前頭葉の機能について”, 認知神経心理学研究会 (2003).

[31] S. C. Blank, S. Scott and R. Wise: “Neural systems involved in propositional and non-propositional speech”, NeuroImage, 6, p. S509 (2001).

[32] F. H. Guenther: “Neural modeling of speech production”, Proceedings of the 4th International Nijmegen Speech Motor Conference (2001).

[33] T. Nakamura, T. Kitamura and M. Akagi: “A study on nonlinguistic features in singing and speaking voices by brain activity measurement”, NCSP, pp. 217–220 (2009).

[34] A. Schirmer and S. A. Kotz: “Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing”, Cognitive Sciences, **10**, 1, pp. 24–30 (2006).

[35] K. Meriau, I. Wartenburger, P. Kazzer, K. Prehn, A. Villringer, E. van der Meer and H. R. Heekeren: “Insular activity during passive viewing of aversive stimuli reflects individual differences in state negative affect”, Brain and Cognition, **69**, pp. 73–80 (2009).

[36] R. J. Love and W. G. Webb: “基礎神経学”, 西村書店 (2002). (田中 隆一, 相馬 芳明監訳).

[37] Y. Sakurai and T. M. Masahiko Asami: “Alexia and agraphia whth lesions of the angular and supramarginal gyri: Evidence the disruption of sequential processing”, Journal of the Neurologcal Sciences, **288**, pp. 25–33 (2010).

[38] J. C. Eccles: “脳の進化”, 東京大学出版会 (1997). (伊藤 正男訳).

[39] R. D. Rafal: “Oculomotor functions of the parietal lobe: Effects of chronic lesions in humans”, Cortex, **42**, pp. 730–739 (2006).

[40] C. Alain, S. R. Arnott, S. Hevenor, S. Graham and C. L. Grady: ""what" and "where" in the human auditory system", PNAS, **98**, 21, pp. 12301–12306 (2001).

[41] L. Pessoa, S. Kastner and L. G. Ungerleider: "Attentional control of the processing of neutral and emotional stimuli", Cognitive Brain Research, **15**, pp. 31–45 (2002).

[42] M. L. Phillips, L. M. Williams, M. Heining, C. M. Herba, T. Russell, C. Andrew, E. T. Bullmore and M. J. Brammer: "Differential neural responses to overt and covert presentations of facial expressions of fear and disgust", NeuroImage, **21**, pp. 1484–1496 (2004).

[43] D. Matsumoto and P. Ekman: "American-japanese cultural differences in intensity ratings of facial expressions of emotion", Motivation and Emotion, **13**, 2, pp. 143–157 (1989).

[44] C. E. Izard: "Translating emotion theory and research into preventive interventions", Psychological Bulletin, **128**, 5, pp. 796–824 (2002).

[45] P. R. Shaver, U. Murdaya and R. C. Fraley: "Structure of the indonesian emotion lexicon", Asian Journal of Social Psychology, **4**, pp. 201–224 (2001).

本研究に関する研究業績

国際会議

Yasuhiro Hamada，Tatsuya Kitamura，Masato Akagi，“A study on brain activities elicited by synthesized emotional voices controlled with prosodic features” Proc. NCSP10，Hawai，USA，March 2010.

口頭発表

濱田 康弘，北村 達也，赤木 正人，“fMRI による感動詞「ええ」呈示時の脳活動測定” 日本音響学会平成 22 年度春季研究発表会 講演論文集，1-R-4，March 2010.