

Title	脳活動測定による歌声と話声に関する非言語特徴の研究
Author(s)	中村, 友彦
Citation	
Issue Date	2009-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8117
Rights	
Description	Supervisor: 赤木 正人 教授, 情報科学研究科, 修士

修 士 論 文

脳活動測定による歌声と話声に関する
非言語特徴の研究

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

中村 友彦

2009 年 3 月

修 士 論 文

脳活動測定による歌声と話声に関する 非言語特徴の研究

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処理学専攻

0710053 中村 友彦

提出年月: 2009 年 2 月

概 要

本研究の目的は、異なる非言語情報を含む刺激音を聞いた際の脳活動を測定し、活動の違いを分析することで、非言語情報に関するどのような音響的特徴が脳に異なる活動を引き起こすのかを調べることである。音声には非言語情報と言語情報が含まれており、音声知覚には両方の情報が関係する。言語情報とは、音韻情報のことである。一方、非言語情報とは、声質、話者の性別や感情などの音韻情報以外の情報のことである。今日までの多くの研究により、言語情報における知覚の解明は進んでいるが、非言語情報に関する知覚は未だ解明されていないことが多い。そこで、本研究では、非言語情報が異なる音声の一例として歌声と話声に着目し、歌声に関する非言語特徴の異なる刺激音を聞いた際の脳活動を測定し、非言語情報の知覚の解明を試みる。脳活動測定実験の結果、歌声と話声を聞いた際の脳活動は異なることが判明した。歌声を聞いた際の脳活動は、話声を聞いた際に比べて、LORg (眼窩回側部) や MORg (眼窩回中央) などの情動系に関する部位により強い脳活動を示すことが判明した。また、歌声の基本周波数変化やスペクトル形状などの非言語情報の音響的特徴の違いにより、大脳辺縁系や大脳基底核などの本能や情動に関係する脳部位において活動差が起こることも判明した。この結果より、ヒトが歌声を知覚する際には、歌声特有の基本周波数変化・スペクトル形状による影響で、情動に関する脳部位が活性化することで、「歌声」と知覚する可能性があると考えられる。

目次

第1章	序章	1
1.1	はじめに	1
1.2	本研究の背景	1
1.2.1	歌声特有の音響的特徴に関する研究	1
1.2.2	脳活動測定実験による歌声知覚に関する研究	2
1.3	本論文の目的	4
1.4	研究方法	4
1.5	本論文の構成	5
第2章	歌声と話声の音声知覚に関する 脳活動測定実験Ⅰ	7
2.1	はじめに	7
2.2	実験Ⅰの目的	7
2.3	実験Ⅰで用いる刺激音	7
2.3.1	刺激音の作成方法	7
2.3.2	作成した刺激音	7
2.4	聴取実験	10
2.4.1	聴取実験の方法	10
2.4.2	聴取実験結果と考察	12
2.5	脳活動測定実験Ⅰ	16
2.5.1	脳活動測定実験Ⅰの方法	16
2.5.2	解析手法	17
2.5.3	解析結果	18
2.5.4	脳活動測定実験結果の考察	20
2.6	まとめ	22
第3章	歌声と話声の音声知覚に関する 脳活動測定実験Ⅱ	23
3.1	はじめに	23
3.2	実験Ⅱの目的	23
3.3	実験Ⅱで用いる刺激音	23
3.3.1	歌声に関する非言語情報のみ異なる刺激音の作成方法	23

3.3.2	作成した刺激音	28
3.4	聴取実験	28
3.4.1	聴取実験の方法	29
3.4.2	聴取実験結果と考察	30
3.5	脳活動測定実験Ⅱ	33
3.5.1	脳活動測定実験Ⅱの方法	33
3.5.2	解析手法	33
3.5.3	解析結果	33
3.5.4	脳活動測定実験結果の考察	45
3.6	まとめ	45
第4章	歌声と話声に関する非言語情報の違いが引き起こす脳活動の違い	46
4.1	はじめに	46
4.2	実際のヒトの歌声と話声を聞いた際の脳活動の違い	46
4.3	歌声特有の音響的特徴が与える脳活動の違い	47
4.3.1	F0の違いが与える脳活動の影響	47
4.3.2	スペクトルの違いが与える脳活動の影響	48
4.4	他の実験結果との比較	49
4.5	聴取実験結果と脳活動の関係	50
第5章	結論	51
5.1	本論文で明らかになったことの要約	51
5.2	今後の課題	51

第1章 序章

1.1 はじめに

音声を扱い相手に意見や意思を伝えることは、ヒトが行う重要なコミュニケーションである。音声には非言語情報と言語情報が含まれており、音声知覚には両方の情報が関係する。言語情報とは、音韻情報のことであり、非言語情報とは、声質、話者の性別や感情などの音韻情報以外の情報のことである。今日までの多くの研究により言語情報に関する知覚の解明は進んでいるが、非言語情報に関する知覚は未だに解明されていないことが多い。そこで本研究では、非言語情報が異なる音声の一例として歌声と話声に着目し、歌声に関する非言語特徴の異なる刺激音を聞いた際の脳活動を測定し、音声中の非言語情報知覚の解明を試みる。脳活動測定装置を用いて、知覚の主体である脳活動を測ることにより、ヒトの知覚をトップダウン的に解明していくことが出来ると考えられる。

1.2 本研究の背景

ヒトは、音声を知覚する際に、言語情報と非言語情報の両方を知覚する。コミュニケーションを取るには、これらの両方の情報の知覚が重要である。特に、非言語情報を知覚することにより、言葉に言い表せない感情や意思を相手に伝えることができ、言語・民族・文化に関係なく、ヒト共通のコミュニケーションを図ることが出来る。非言語情報の知覚を解明することは、言語や文化によらないコミュニケーション環境を構築することに繋がると考えられている。ヒトは歌声と話声に対して、言語情報に関わらず、非言語情報の違いを知覚し判別している。このため、歌声と話声は、非言語情報の異なる代表的な音声の一例であると言える。非言語情報の知覚を解明するには、非言語情報のみが脳活動に与える影響を知ることが重要である。そこで、非言語情報の異なる例として歌声と話声に着目し、脳活動を測定し、非言語情報の知覚の解明を試みる。本節では、歌声特有の音響的特徴に関する知見、また、脳活動研究の歩み、歌声に関する脳活動測定実験の知見について述べる。

1.2.1 歌声特有の音響的特徴に関する研究

音源フィルタモデル (source-filter model) [17] に基づけば、人間の音声は、声帯振動の情報を表す基本周波数と、声帯形状の情報を表すスペクトル形状の2つの音響パラメータ

によって表現される．歌声と話声の違いは，これら音響パラメータの違いとして現れる．実際に，歌声と話声それぞれにおける音響パラメータの特性が異なり，その違いを知覚することによって話声と歌声を容易に区別可能なことが先行研究で示されている [14]．歌声特有の基本周波数変化とスペクトル形状の特性を持つものを，ヒトは歌声として知覚すると考えられる．齋藤の歌声合成に関する研究により，歌声特有の基本周波数変化とスペクトルなどの音響パラメータが明らかとなった [6]．また，歌声知覚に最も大きな影響を与えている音響的特徴は，メロディ変化中に存在する F0 動的変動成分であることも明らかとなった．歌声特有の音響的特徴量と歌声知覚の関係を検討した齋藤，辻，鵜木，赤木らの研究では，歌声らしさにおいて基本周波数の準周期的な振動成分であるヴィブラートとそれに同期したホルマントの振幅振動成分，および 3 kHz 付近の顕著なスペクトルピーク成分と同帯域の強い高調波成分がそれぞれ寄与していることが明らかとなった [7]．

ヒトが歌声知覚する際に重要な音響的特徴は明らかになったが，それらは心理実験を用いて評価したものであり，知覚の主体である脳活動にどう影響を与えているのかまでは明らかになっていない．そこで本研究では，これらの歌声特有の音響的特徴を適宜変化させた合成音を作成し，それらを聞いた際の脳活動を測定して，ヒトが「歌」として知覚する音響的特徴を脳活動から解明する．

1.2.2 脳活動測定実験による歌声知覚に関する研究

脳研究は，19 世紀末にブローカが失語症の原因が脳の損傷にあることを発見し，盛んに解剖されるようになり進められてきた．失語症の患者が亡くなった際に解剖を行い，脳の損傷や異常から，脳の研究を行っていった．しかし，その当時は，生きているヒトの脳活動そのものを画像化する技術はなく，ヒトが音声知覚中にどのような脳活動を行っているのか知る術はなかった．

脳の機能についての研究は，心理実験や脳損傷患者の行動や脳の損傷部位の記録に基づいて行われてきた．特定の刺激や作業を与えたときの脳活動を測定する技術は，長い間，脳波計測 (EEG: electroencephalography) のみであった．しかし，特にここ数十年の間にさまざまなセンサ技術や計算機技術が向上し，脳の活動をより多面的に研究することが可能となった．ヒトの脳活動を解明する手法として，脳磁図 (magnetoencephalography: MEG)，陽電子断層撮像法 (positron emission tomography: PET)，機能的磁気共鳴画像法 (functional magnetic resonance imaging: fMRI)，光トポグラフィー (optical topography: OT) などの脳イメージング技術が発達し，それぞれの特徴を生かした研究が行われるようになった [18]．

本研究で用いる機能的磁気共鳴画像法 (functional Magnetic Resonance Imaging: fMRI) は人体の形態計測に用いられる MRI 装置を用いて，脳の機能 (function) に関わる部位を同定する手法である．MRI は組織を構成する分子の水素 (プロトン) の濃度の違いを画像化している．この濃度の違いは分子の構造や磁界の状態によって変化する．一般的な fMRI の手法といえる BOLD (blood oxygenation level dependent) 法はこれをうまく利用して

脳の活動部位を同定している．たとえば，刺激を呈示したときとしていないときの脳活動において，多数の脳の形態画像を高速に採取し，実験後画像の濃淡の程度を統計的に比較する．与えられた刺激に関連して活動した部位では，血液の成分（還元ヘモグロビンと酸化ヘモグロビンの濃度）の比や血流量が変化する．この変化を「ヘモダイナミクス」と言うが，これに伴い画像中の対応するピクセルではその濃淡レベルが変化する．しかし，これは目で見てわかる程の変化ではないので，刺激を呈示しているときとしていないときの間で差があるかどうかを，多数の画像を使い統計分析して調べる．そして変化のあった部分を着色表示することで活動部位を可視化する [3][22]．

また，fMRI では分析できる脳画像の空間分解能は画素 (voxel) 単位であるためミリ単位とすることができる．しかし，この BOLD 法では各条件下でのデータ採取時間は 30 秒程度かかるため，リアルタイムに脳の活動部位の変化を測るほどの時間分解能はない．時間分解能が EEG や MEG より劣るものの，空間分解能がそれらより優れている．fMRI の優れた特徴として，放射線被爆のなく非侵襲的に，空間・時間分解能が高く，繰り返して，通常の装置で簡易に画像化できることである．

ヒトの聴覚に関する脳活動画像による研究では PET が主に用いられてきたが，空間分解能が相対的に低いために限界があった．一方で，fMRI は撮影時に傾斜磁場コイルの発する音が大きく（約 100 dB），この音による聴覚関連野の賦活によって脳活動画像の S/N 比が悪くなることが問題であった．近年ヘッドフォンの遮音性の向上や撮影シーケンスの工夫により精度の良い画像化が可能となりつつある [1]．

これら脳活動測定装置の発達とともに，ヒトが音を知覚する際の脳活動を測る研究も多く行われるようになった．fMRI を用いた歌声と話声に関する脳活動研究として，Callan らの研究がある [30]．Callan らは，20 秒間で 8 小節ある日本語の童謡 (e.c. かごめかごめ，焚き火) を歌っている声と話している声を刺激音として用いた実験を行っており，それぞれ歌声と話声を聞いた際には，活性化する脳部位が異なることが判明している．

また，Brown らは，メロディとハーモニーの異なるピアノ音を刺激音として用いた実験を行い，歌声の複雑性や調和に関する脳部位の活動も判明している [29]．

Stefan は，シンタックスとセマンティックスの異なる音楽を用いた実験を行い，シンタックス処理やセマンティックス処理に関する脳部位が判明している [38]．

しかし，これらの研究では，用いた刺激音に言語情報の違いが含まれており，スペクトル形状や基本周波数変化といった非言語情報のみが脳に与える影響を明らかに出来ない．

他にも，聴覚や発声に関する脳活動研究も多く行われている．例えば，Jeremy らは，発声している画像と音声を見聞きしているときと，音声だけを聞いているとき，話している顔画像だけを見ているときのそれぞれの脳活動の違いを調べる実験を行い，聴覚や視覚を統合し活性化する脳部位や音声知覚で運動皮質が活動するという報告を行っている [43]．Friedemann らは，被験者に調音素性を変化させて発声をしてもらい，その時の運動皮質の脳活動を調査している [41]．歌声の発声に関しては，Gunji らが，MEG (脳磁図) を使い，ヒトの歌唱中の脳活動を調べる研究を行っている [35]．

このように、聴覚系の音声知覚や音声生成に関する脳活動研究は多数行われているが、歌声の非言語特徴の知覚にのみ着目した脳活動研究は行われていない。

1.3 本論文の目的

研究背景で述べたように、音声知覚に関する脳活動研究は多く行われているが、用いた刺激音に言語情報の違いが含まれており、非言語情報のみに関する脳活動(知覚)を明らかにすることはできない。本稿の目的は、歌声と話声に関する非言語情報を含む刺激音を聞いた際の脳活動を測定し、活動している脳部位の違いを検討し、歌声と話声それぞれに特有の脳部位を調べ、非言語情報の知覚の解明を試みることである。また、歌声特有のスペクトル形状や基本周波数変化などの音響的特徴を変化させた合成音を用い、どのような脳活動を示すのかを調べ、非言語情報に関する音響的特徴のみ変化することによって起こる脳活動部位を調べる。脳活動における違いを調べることにより、ヒトの脳が「歌」と判断することに関わる音響的特徴を考察する。

1.4 研究方法

非言語情報に関する知覚を解明するために、本研究の脳活動測定実験において用いる刺激音はすべて同一の言語情報 /a/ を持つものとし、実際の歌声と話声、そして歌声特有の非言語情報のみ異なる合成音を用いる。合成音は、高品質な分析合成系 STRAIGHT (Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrum) [5] を用いて作成する。

非言語情報のみ異なる刺激音を聞いた際の脳活動を測定することで、非言語情報のみに影響された脳活動を解析することが出来る。なお、刺激音がどのくらい歌声として意識して聞かれているかを調べるために「歌声らしさ」について聴取実験を行う。また、刺激音の品質の評価として、ヒトが発した声として聞こえるかという尺度である「自然性」についても評価を行う。それらの評価結果と脳活動結果を比較し、関係を考察する。

本研究では、用いる刺激音セットの異なる2つの脳活動測定実験Ⅰ、Ⅱを行う。実験Ⅰは、歌声と話声の脳活動の違いと、「歌声」と知覚される具体的な音響的特徴の物理量を調べるために、歌声合成に関する研究で重要であると判明した音響的特徴を話声に適宜加えて合成音を作り、脳活動にどのような影響があるのかを調べる。そこで、実験Ⅰで用いる刺激音は、実際の歌声と話声、さらに話声に歌声のスペクトル形状やヴィブラートを加えた合成音を用いる。実験Ⅱは、歌声の重要な要素である基本周波数(F0)、スペクトル形状、振幅エンベロープが、脳において「歌声」と知覚される上でどれほど寄与するのかを調べるために行う。そこで、実験Ⅱで用いる刺激音は実際の歌声と話声を構成するそれぞれ音響的特徴を入れ替えて、歌声と話声を補間する合成音を用いる。

最後に、2つの脳活動測定実験の結果をまとめ、歌声に関する脳活動、非言語情報に関する知覚に関する考察を行う。

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，本論文の位置づけと目的を示す．

第2章では，実験Ⅰで用いる歌声に関する非言語情報のみ異なる刺激音の説明と作成方法と，それらの刺激音に対する聴取実験の説明と実験結果，そして，脳活動測定実験Ⅰの説明と実験結果，考察を記述する．

第3章では，実験Ⅱで用いる歌声に関する非言語情報のみ異なる刺激音の説明と作成方法と，それらの刺激音に対する聴取実験の説明と実験結果，そして，脳活動測定実験Ⅰの説明と実験結果を記述する．

第4章では，実験Ⅰと実験Ⅱの結果をふまえ，歌声や話声，基本周波数変化やスペクトル形状などの音響的特徴の違いによる脳活動の違いについての全体考察を記述する．

第5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．

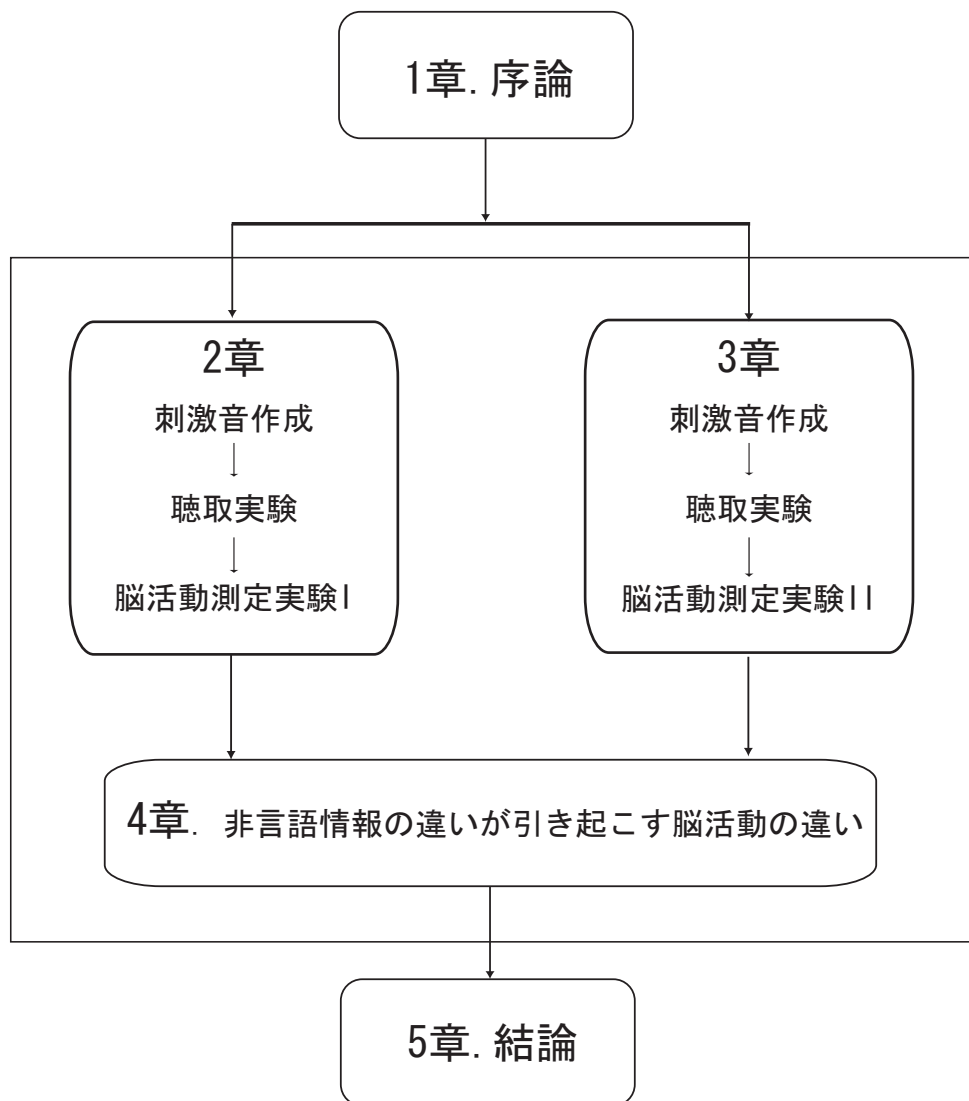


図 1.1: 各章の相互関係

第2章 歌声と話声の音声知覚に関する 脳活動測定実験Ⅰ

2.1 はじめに

本章では、非言語情報のみ異なる刺激音作成と、刺激音が意識上でどのように知覚されているのか評価を行った聴取実験、そして、刺激音を聞いた際の脳活動を測定する実験Ⅰについて記述する。

2.2 実験Ⅰの目的

脳活動測定実験Ⅰの目的は、実際の歌声と話声の脳活動の違いを調べることであり、また、「歌声」と知覚される具体的な音響的特徴の物理量を調べるため、歌声特有の音響的特徴を加えた合成音の脳活動に与える影響を調べることである。

2.3 実験Ⅰで用いる刺激音

2.3.1 刺激音の作成方法

非言語情報が与える脳活動を調べるために、歌声に関する非言語情報のみ異なる刺激音を作成する。刺激音作成には、高品質な分析合成系 STRAIGHT [5] を用いて作成する。歌声に関するスペクトル形状やヴィブラートは齋藤の手法 [6] を用いて作成した。

2.3.2 作成した刺激音

脳活動測定実験Ⅰの目的は、実際の歌声と話声の脳活動の違いを調べることであり、また、「歌声」と知覚される具体的な音響的特徴の物理量を調べるため、歌声特有の音響的特徴を加えた合成音の脳活動に与える影響を調べることである。そのために、実験Ⅰで用いる刺激音として、実際の歌声と話声に加えて、齋藤らの研究 [6][7] において、歌声の音響的特徴として重要であると判明したスペクトル形状とヴィブラートの変形を実際の話声に加えた合成音を作成する。実験Ⅰで用いた6つの刺激音の名称と作成方法を以下に記述する。

- Speak
ヒトが発した話声
- Sing
ヒトが発した歌声
- Base
Speak に対して 3 kHz 付近のホルマントピークを 18dB 強調し，歌声特有のスペクトル形状に変化させた合成音
- VR1
Base に 0.95-1.05 Hz の基本周波数変化を加えた合成音
- VR2
Base に 5.3-5.9 Hz の基本周波数変化を加えた合成音
- VR3
Base に 18.93-21.07 Hz の基本周波数変化を加えた合成音

すべての刺激音の音韻情報は /a/ である．実際の歌声と話声である刺激音 Speak と Sing は，歌声データベース「日本語を歌・唄・謡う」の中から選定した男性テノール歌手の音声を用いている [16]．Speak に対して 3 kHz 付近のホルマントピークを 18 dB 強調し，歌声特有のスペクトル形状に変化させた合成音 Base を作成する．次に，Base に対して，歌声の要素として重要なヴィブラートを加えて，VR1，VR2，VR3 を作成する．VR2 のヴィブラートのパラメータは，齋藤の実験で最も歌声の自然性の高い vibrato rate である 5.6 Hz と vibrato rate band width である 0.3 Hz とし，VR1，VR3 は，両極端の周波数 1 Hz と 20 Hz を中心周波数として，同じ割合の幅でヴィブラートを変化させ作成した．すべての刺激音の音圧は同一，継続時間は Sing のみが 1.79 sec であり，他の刺激音は 1.74 sec である．サンプリング周波数は 44.1 kHz である．

Sing と Speak をそれぞれ聞いた際の脳活動の違いを解析することで，実際の歌声と話声の脳活動の違いを調べることが出来る．また，Base と Speak や VR1 と Base の脳活動の違いを解析することで，スペクトル形状やヴィブラートが脳活動に与える影響を調べることが出来る．刺激音の概要を図 2.1 に示す．

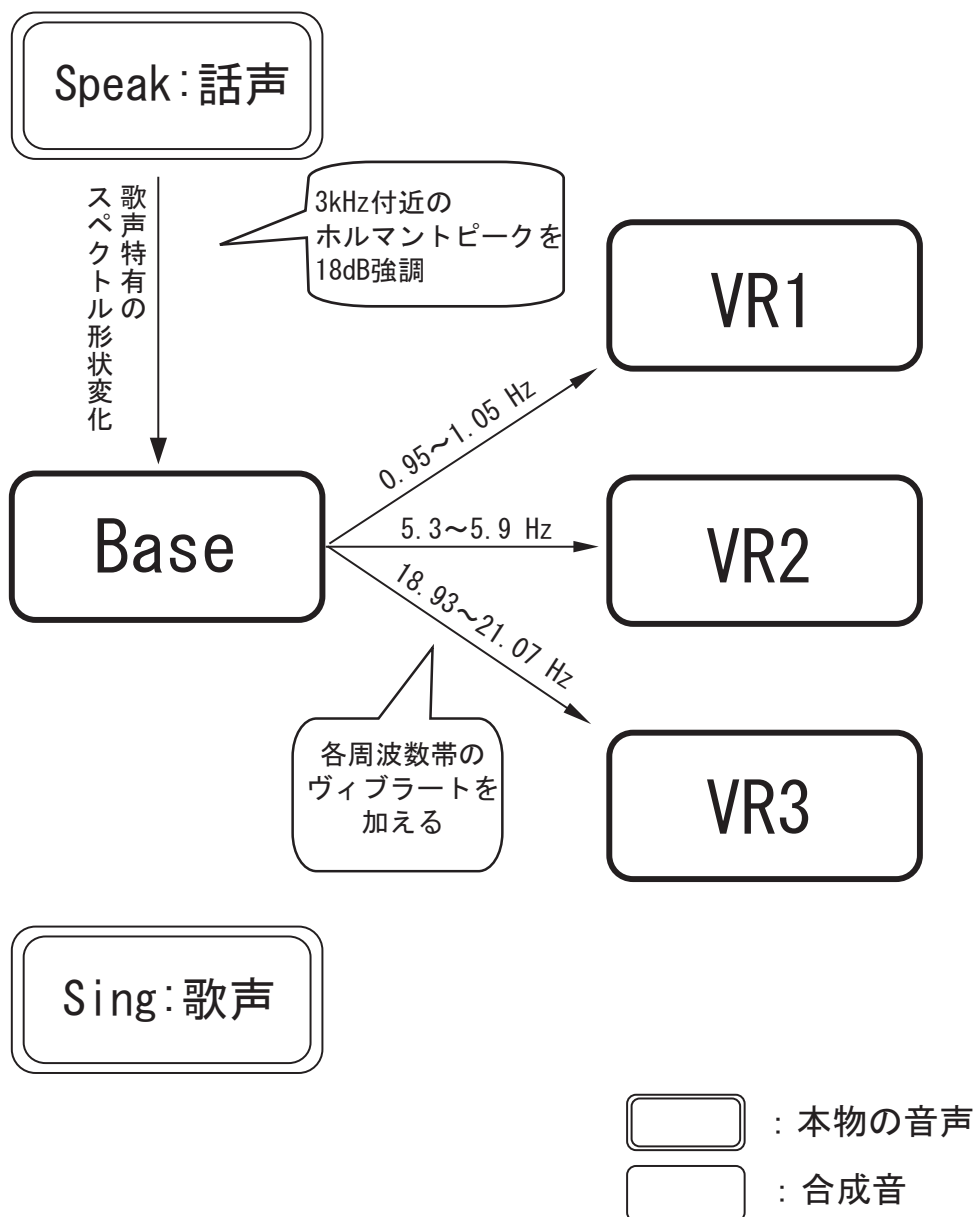


図 2.1: 刺激音の概要

2.4 聴取実験

本節では、聴取実験について記述する。本研究は、合成音を用いた歌声に関する脳活動の研究であるため、刺激音の「歌声らしさ」をあらかじめ評価する必要がある。また、刺激音がヒトの音声として聞こえるかを評価するため、「自然性」もあらかじめ評価する必要がある。聴取実験の評価結果と脳活動測定実験の結果を比較し議論する。

2.4.1 聴取実験の方法

聴取実験の方法は、上述の6種類の刺激音にたいして、「歌声らしさ」と「自然性」についてそれぞれ評価してもらう。評価方法は、2つ連続して呈示される刺激音に対して、どちらの刺激音がより「歌声らしい」かもしくは「自然性」が高いかを5段階評価でそれぞれ採点してもらい、シェッフェの一对比較法を用いて評価する[8]。心理物理実験において、一对比較法は、数個の刺激を2つずつ対にして判断を求める方法である。また、実験の所要時間は比較的短くて済み、2つの刺激に対して比較判断を求めるので、刺激間の差が微妙な場合にも適用できる。その中で、一对比較法にカテゴリー判断を取り入れたシェッフェの一对比較法は、被験者が対にして提示される刺激に比べて、どちらがどれだけ好きかなどの判断を求めるものである。一对比較法では、正規分布の仮定に基づいて序数尺度を間隔尺度に変化する手続きを行うため、多くの被験者を必要とするが、シェッフェの一对比較法では、被験者が判断した評価点を序数尺度のまま統計的検討を行うので、多くの被験者を必要としない利点がある。

刺激条件

実験で用いる聴覚刺激は、先に示した刺激音 Speak, Sing, Base, VR1, VR2, VR3 を2つずつ対にしたものである。刺激音が6種類あるので、刺激対の数は、順序効果も考慮した $6 \times 5 = 30$ 対である。図 2.2 に刺激の呈示順序を示す。

実験参加者

聴取実験の被験者は正常な聴力をもつ10人(男性9人,女性1人)である。被験者の年齢は20代(男性7人女性1人),30代男性1人,50代男性1人である。

実験環境

実験は、防音室において、実験被験者にはヘッドホン (Sennheiser HDA200) を介して刺激音対を両耳に呈示し、PCディスプレイ上の評価尺度をキーボードで入力させること

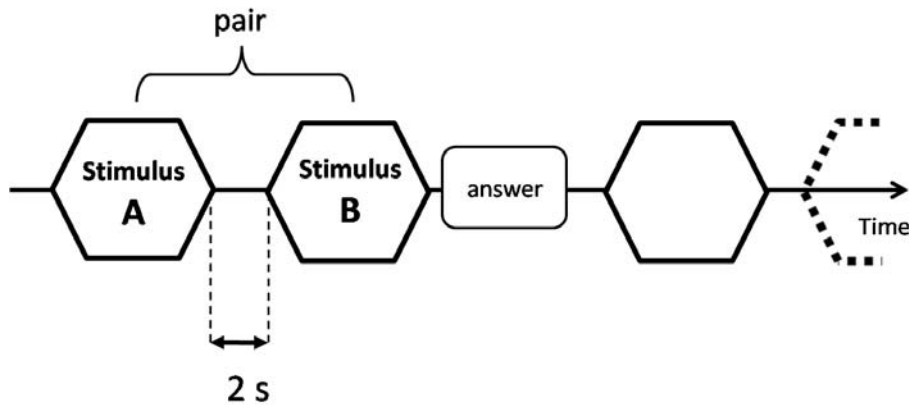


図 2.2: 刺激の呈示順序

で回答させた．また，音圧レベルは実験参加者の聴きやすいレベルに設定した．その際使用された主な機器を以下に列挙する．

ノート PC ：Panasonic CF-R6

D/A 変換器 ：YAMAHA DP-U50

ヘッドホンアンプ ：YAMAHA DP-U50

ヘッドホン ：Sennheiser HDA200

実験手続き

被験者には次のような教示を与え、「歌声らしさ」と「自然性」に関して評価してもらった．なお，各実験参加者につき，30 対の 1 セットを 3 回行い評価してもらう．

ヘッドホンから 2 つの音を対にして流します．前の音と後の音を聞き比べて，どちらが「歌声らしい」か，下に記した 5 段階評価尺度に従って判断してください．前の音の方が歌声らしく聴こえたら負の値 (-2 ~ -1) に，後の音の方が歌声らしく聴こえたら正の値 (1 ~ 2) の当てはまるキーを入力してください．どちらも同程度の歌声らしさだと判断した場合は 0 に当てはまるキーを入力してください．

「自然性」の評価に関しては次のような教示を与えた．

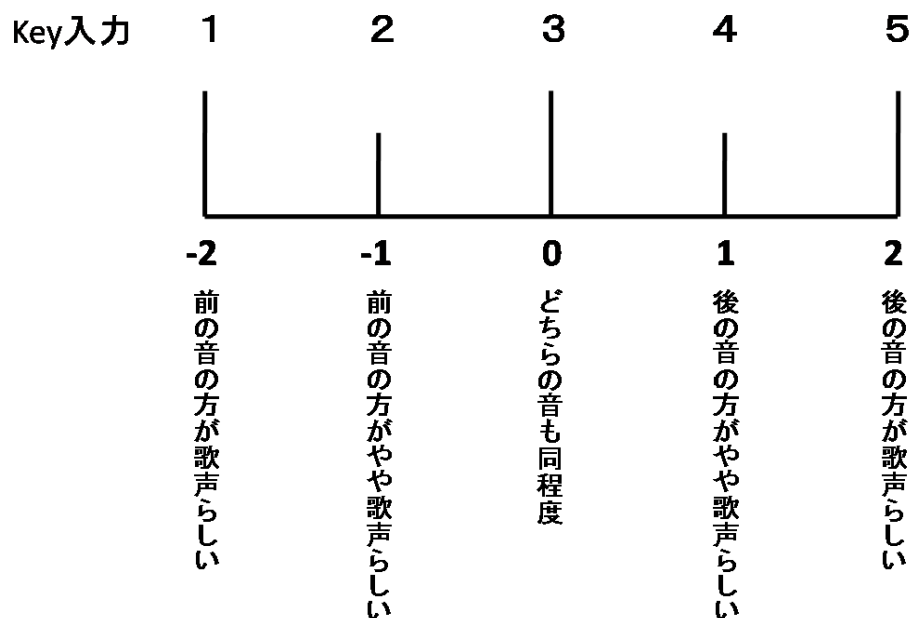


図 2.3: シェッフェの一対比較実験で用いた「歌声らしさ」に関する五段階評価尺度。

ヘッドホンから2つの音を対にして流します。前の音と後の音を聞き比べて、どちらが人間の発する声として「自然」か、下に記した5段階評価尺度に従って判断してください。前の音の方が歌声らしく聴こえたら負の値 (-2~-1) に、後の音の方が歌声らしく聴こえたら正の値 (1~2) の当てはまるキーを入力してください。どちらも同程度の歌声らしさだと判断した場合は0に当てはまるキーを入力してください。

2.4.2 聴取実験結果と考察

上記の実験方法で得られた「歌声らしさ」、「自然性」のデータを、浦の変法 [8] によって処理した結果を表 2.1, 2.2 に示す。

また、表に示した母数の値に従って、5つの刺激の距離関係を直線上で示したものが、「歌声らしさ」においては図 2.4、「自然性」では図 2.5 になる。また、「歌声らしさ」と「自然性」の評価結果をまとめて表した結果を Fig.2.6 に示す。母数の値は、刺激音がどれだけ「歌声らしい」もしくは「自然性」が高く聴こえたかを表す値であり、正の大きな値であるほど評価が高い。尚、F 検定法による刺激間の有意差検定を行った結果、「歌声らしさ」においては、Speak と VR3、Base と VR1 の刺激音間以外で、他の各刺激音間に 5% の水準で有意な差が確認された。「自然性」においては、Speak と Base、Speak と Sing、Base と Sing、VR1 と VR2 の刺激音以外で、他の各刺激音間に 5% の水準で有意な差が確認さ

表 2.1: 母数の推定 (歌声らしさ)

刺激音	母数
Speak	-0.76
Sing	1.48
Base	-0.55
VR1	-0.49
VR2	1.26
VR3	-0.94

表 2.2: 母数の推定 (自然性)

刺激音	母数
Speak	0.58
Sing	0.48
Base	0.54
VR1	-0.05
VR2	0.18
VR3	-1.73

れた。

図 2.6 からわかるように、「歌声らしさ」において、Sing と VR2 が高い評価を得ている。実際の歌声である Sing は予想通り最も歌声らしいという結果になった。合成音の VR2 が Sing に次いで高いことから「歌声らしさ」において、5.3-5.9 Hz 付近の基本周波数変化における音響的特徴が重要であることが判る。これは、ヴィブラートが 4~7 Hz 程度の周波数変化が一般的な特性とされていることと一致する [42]。実際のヒトの話声である Speak よりも、歌声特有のスペクトルを持つ Base の方が「歌声らしさ」が高く、また、スペクトル形状を変化させた Base よりもスペクトル形状と基本周波数変化を加えた VR2 の方が「歌声らしさ」が高くなるという結果は、齋藤の実験結果と一致する [6]。

「自然性」においては、Speak、Sing、Base で有意差のない同程度の高い評価が得られた。実際の音声である Speak と Sing は、予想通り「自然性」が高い結果となった。Base の結果より、スペクトルの変化では「自然性」は大きく劣化しないことがわかった。一方、Base にそれぞれ 1 Hz、20 Hz を中心周波数としてヴィブラートを加えた VR1、VR3 に関しては「自然性」と「歌声らしさ」の両方とも高くない。「歌声らしさ」が低いものはヒトが発する声としても「自然性」が低いと知覚されると考えられる。また、VR2 は「歌声らしさ」では実際のヒトの歌声である Sing と同程度であったが、「自然性」では Sing に比

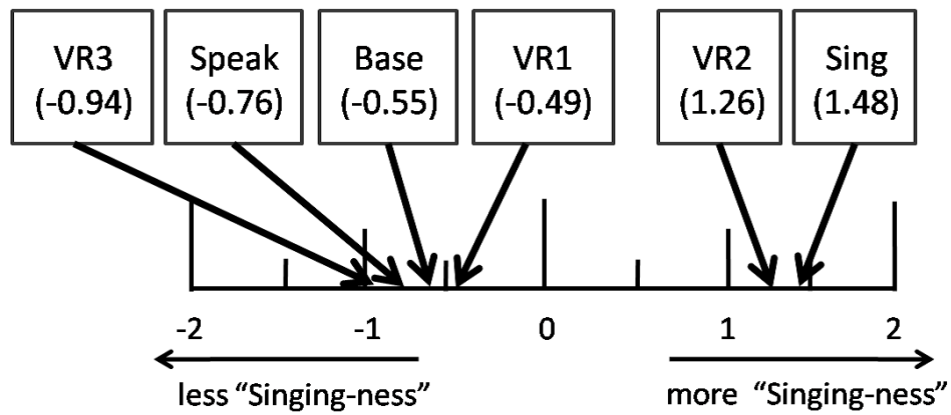


図 2.4: 「歌声らしさ」の関係 .

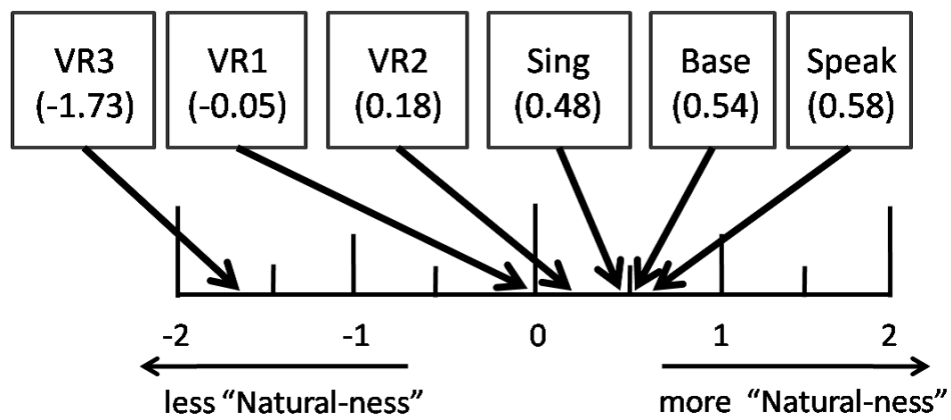


図 2.5: 「自然性」の関係 .

べ「自然性」が低い結果となった .

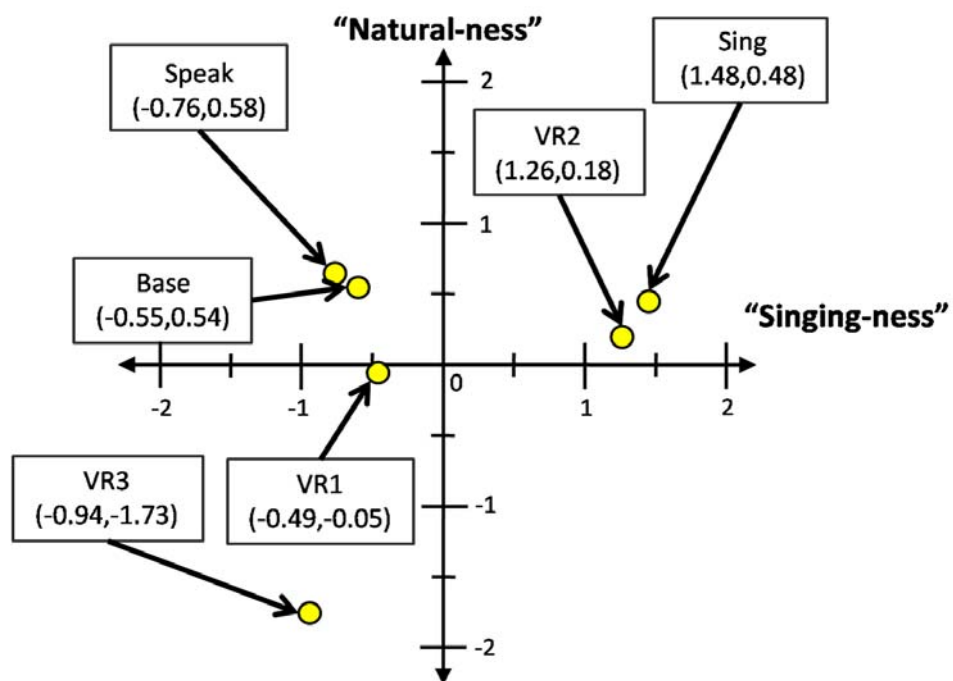


図 2.6: 「歌声らしさ」と「自然性」の関係 .

2.5 脳活動測定実験Ⅰ

本節では、脳活動測定実験Ⅰについての実験内容、解析方法、解析結果と考察について記述する。上述した刺激音を聞いた際の脳活動を測定し、歌声と話声、スペクトル形状やヴィブラートなどの音響的特徴を加えた合成音を聞いた際の脳活動の違いを調べるため脳活動測定実験Ⅰを行う。歌声と話声の脳活動は異なるのか、音響的特徴の違いで脳活動は異なるのか、聴取実験の結果と脳活動結果に関係があるのかどうかを調べる。

2.5.1 脳活動測定実験Ⅰの方法

脳活動測定実験とは

ある認知課題を施行中の局所脳血流量を測定し、他の課題を施行中の血流量との違いを統計処理を加えて比較することにより、ある認知過程に関与する脳部位を同定できる。これを脳活動測定実験という。

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。他の主な実験装置は、以下の通りである。

脳活動測定装置 : 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 Presentation (ソフトウェア)

呈示方法

Presentation ソフトを用いて、呈示プログラムを作成する。被験者には、実験に入る前に文章で実験の概要を説明し、同意してもらった上で実験を行う。被験者に歌声の実験であることを意識させないように、上述した刺激音以外にノイズ音を呈示し、ノイズ音の時にボタンを押してもらうというタスクを課している (オドボール課題)。1 セッションにつき 6 種類の刺激音を 15 回、ノイズ音を 10 回呈示し、全部で 3 セッションを行う。1 セッションは 7 分程度 (432 秒) であり、合計 20 分程掛かる。スキャン法はスパース法を用いる。スパース法とは、刺激音を出す際には撮像をせずに、刺激音を出した直後から撮像する方法である。この方法の特長は、刺激音を出す際には撮像しないことから、撮像をする際に実験装置から生じるノイズに邪魔されることなく、刺激音を聴けることである。し

かし、撮像時間が短いため、繰り返し刺激音を呈示しなければならず、ブロックデザインに比べて時間が掛かる。呈示方法の概要を図 2.7 に示す。

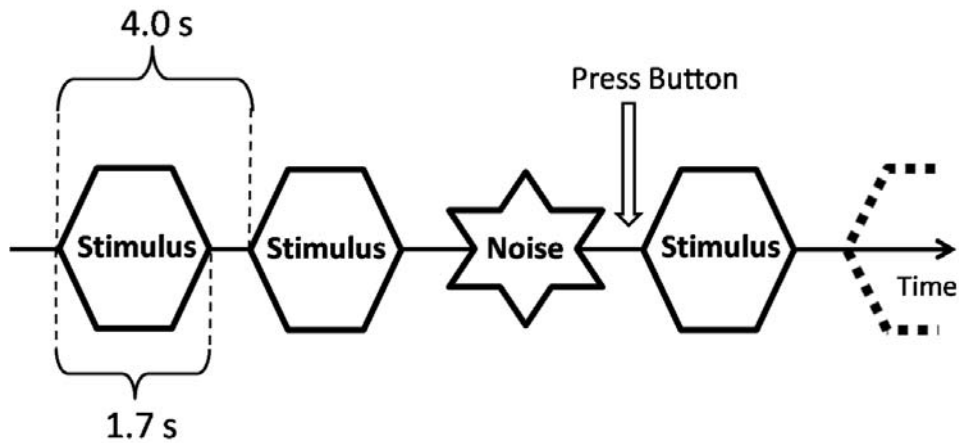


図 2.7: 脳活動測定実験の呈示順序

実験参加者

被験者は正常な聴力をもつ 15 人 (男性 11 人, 女性 4 人) であり, 年齢は 50 代の男性 1 人, 30 代の女性 1 人, 20 代男性 10 人, 女性 3 人である。一部の被験者が聴取実験の被験者と同じである。

2.5.2 解析手法

脳解析ソフト SPM5 (Statistical Parametric Mapping 5) を使い, 各刺激音を聞いた際の脳活動の違いを解析する。脳画像は 30 スライスで $3.0 \times 3.0 \times 4.0$ mm ボクセルの解像度を持つ。スパースデザインであるので, 1 スキャンに対して 1 刺激音を呈示する。1 セッションにつき, 6 種類の刺激音がそれぞれ 15 回, ノイズ音が 10 回呈示されるので, $6 \times 15 + 10 = 100$ スキャン。さらに, スキャンが安定するまで時間が掛かるので, セッションのはじめに 4 スキャンと, 刺激音呈示後に 4 スキャンを撮像するため, $100 + 4 + 4 = 108$ の合計 108 スキャン行われる。得られた fMRI データは, 位置補正, 標準テンプレートに当てはめる標準化を行い, $6.0 \times 6.0 \times 6.0$ mm の FWHM のガウシアンフィルタを通しスムージング化するなどの前処理を行う。

1 次処理として被験者の個人ごとの脳活動を解析する個人別解析を行う。fMRI を用いてヒトの脳機能について研究する場合, 臨床例の研究を除いて, 多くの場合被験者 1 人だけの脳活動画像だけを吟味してもあまり意味がない [13]。そこで, 個人別解析の結果を

基に集団解析を行い，母集団について統計的検定をおこなう．集団解析の結果を，刺激音を聞いたときの脳活動結果とする．今回の解析における有意な活動とは， t 検定において P 値を 0.001 以上で，活動のクラスタ数が 3 以上の部位とする．それらの有意な活動部位を，脳アトラス (『ATLAS OF THE HUMAN BRAIN』[28]) で調べて，脳部位名と脳座標を記述する．脳座標は，Talairach 座標系による x, y, z 軸の値であり， x 軸は左右 (左がマイナスの値)， y 軸が前後， z 軸が上下に通っている．Talairach 座標系とは，Talairach Tournoux (1988) によって提唱された脳の定位座標系であり，皮質下の前交連 (anterior commissure: AC) と後交連 (PC) とを結んだラインを基準線としている．なお，ある刺激音を聞いた時の脳活動から，他の刺激音を聞いたときに脳活動を差し引き，より活動している部位を求め，脳活動の解析結果とする．

2.5.3 解析結果

各 6 種類を聞いた際の脳活動の差 (コントラスト) を取ることにより解析を行った結果を記述する．全コントラストは 30 対あるが，実験 I で調べたい脳活動差は，歌声と話声の脳活動の違い，また歌声に重要な音響的特徴が加えられることによる脳活動の違いである．以下の 3 つのコントラストに絞り，解析結果を示す．

- 歌声 (Sing) と話声 (Speak) の違い : Sing - Speak , Speak - Sing
- スペクトル形状のみの違い : Base - Speak
- ヴィブラートのみの違い : VR1 - Base , VR2 - Base , VR3 - Base

歌声と話声の違い

歌声である Sing を聞いたときの脳活動から，話声の Speak を聞いたときの脳活動を差し引いたコントラスト (Sing - Speak) において，脳活動差を解析した結果，有意な脳活動の差が見られた (図 2.8，図 2.9)．歌声である Sing を聞いたときの方が，話声の Speak を聞いたときよりも，LOrG (側部眼窩回)，SPL (上頭頂小葉)，PrG (中心前回)，AnG (角回)，Cerebellum (小脳) など有意な脳活動が見られた．一方，Sing よりも Speak で有意な脳活動を示す脳部位はなかった．

表 2.3 に，Sing-Speak と Speak-Sing のコントラストでの有意な脳活動部位名と脳座標，クラスタ数を示す．図 2.8 に，Sing-Speak の脳活動差を脳表面にマッピングして表した図を示す．また，図 2.9 には，Sing-Speak の脳活動差の透し図を示す．

表 2.3: Sing と Speak のコントラストの脳活動結果

コントラスト	脳活動部位と脳座標	クラスタ数
Sing - Speak	LOrG [-27, 33, 0]	16
	SPL [-30, -45, 32]	21
	PrG [-24, -12, 40]	28
	MTG [-33, -9, 36]	28
	Pu [24, -3, 12]	3
	AnG [27, -57, 36]	27
	Cerebellum [6, -33, -32]	5
Speak - Sing	nothing	-

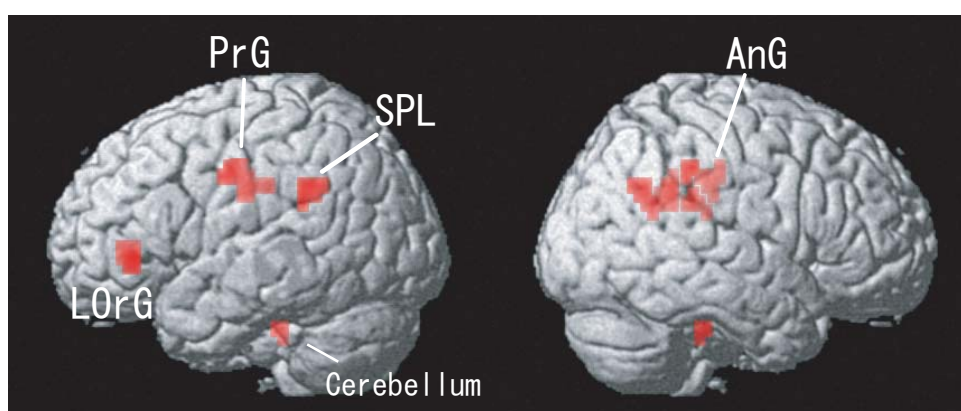


図 2.8: Sing - Speak の脳活動のレンダリング図 (左図：左脳，右図：右脳)

歌声特有のスペクトル形状の違いによる脳活動差

Base は，Speak に対して 3 kHz 付近のホルマントピークを 18dB 強調し，歌声特有のスペクトル形状に変化させた合成音である．Base を聞いたときの脳活動から，Speak を聞いたときの脳活動の差分を取れば，スペクトル形状の違いによる脳活動の違いが判明する．そこで，Base-Speak のコントラストを取ったところ，表 2.4 に示すような部位で脳活動の違いが見られた．

ヴィブラートの違いによる脳活動差

VR1，VR2，VR3 は Base にそれぞれの周波数を加えて作成した合成音である．ヴィブラートを加えた合成音から，Base を聞いたときの脳活動の差分を採ることにより，ヴィブラートの違いによる脳活動の違いを解析する．

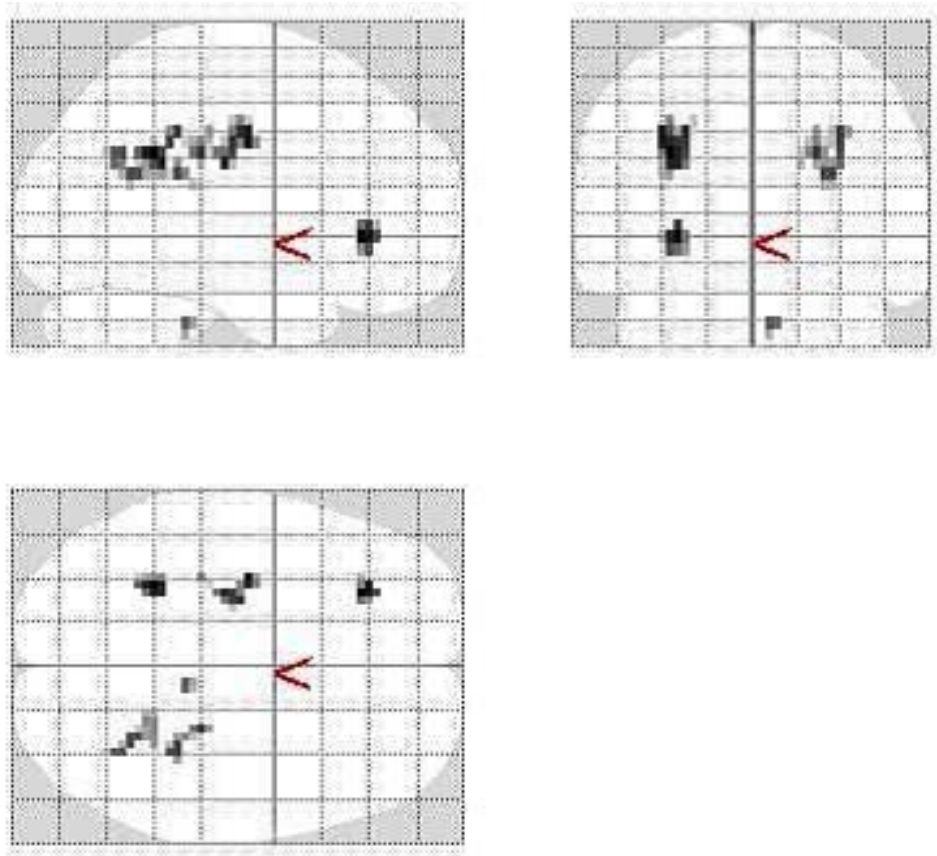


図 2.9: Sing - Speak の脳活動の透し図 (左上：側面図，左下：上面図，右上：前面図) 図中の黒い部分が有意な脳活動を示している

表 2.4: Base - Speak の脳活動差

コントラスト	脳活動部位と脳座標	クラスタ数
Base - Speak	IFGOp [60, 9, 20]	3
	CG [-18, -33, 32]	3
	PrG [-30, -18, 36]	4

VR1-Base , VR2-Base , VR3-Base のコントラストを取ったところ , 表 2.5 に示すような部位で活動の違いが見られた .

2.5.4 脳活動測定実験結果の考察

実験結果より , ヒトが発した歌声 (Sing) と話声 (Speak) をそれぞれ聞いた際の脳活動は異なることが判明した . 歌声では LOrG , SPL , PrG , MFG , AnG , Pu , Cerebellum

表 2.5: VR1 - Base , VR2 - Base , VR3 - Base の脳活動差

コントラスト	脳活動部位と脳座標	クラスタ数
VR1 - Base	IG [-33, 15, 4]	6
	PPo [-51, -6, 0]	4
VR2 - Base	nothing	-
VR3 - Base	POp [-60, -18, 16]	4
	ITG [-42, -48, -12]	3
	MTG [-51, -18, -8]	3

等の脳部位において話声よりも活動が強いことが判明した．同じく歌声と話声の脳活動に関する Callan らが行った実験結果と比較すると，歌声で活性化した共通部位は，PrG，SPL，Cerebellum であった [30]．LORg は，情動系の神経回路の一部と考えられている脳部位であり [2][11][36][39]，歌声を聞いた際の脳活動は，脳表面でなく脳内部の部位での活動が目立つことがわかる．一方，話声では歌声よりも活性化した脳部位はないことから，歌声を聞いた際の脳活動の方が活動が強く，広範囲であると考えられる．

Speak と合成音 Base のコントラストの解析結果から，スペクトル形状変化に影響する脳部位として CG，PrG 等が見られたものの，話声と歌声のコントラストほど大きな活動差は見られない．また，VR1 - Base，VR2 - Base，VR3 - Base の 3 組のコントラストにおいては，Base の脳活動が大きく，ヴィブラートを加えたことにより脳活動が共通して強くなる部位はなかった．

聴取実験と脳活動結果の関係については，「自然性」の評価が高い刺激音は，全体的に脳活動が強い傾向が見られた．しかし，ヴィブラートをもつ VR の中では，VR3 の「自然性」の評価が最も低いものの，脳活動は最も大きいことがわかった．このことは，VR3 が音声として不自然すぎることに注意を引き，他の VR よりも脳活動を強く引き起こしていると考えられる．また，「歌声らしさ」において高い評価を得た Sing と VR2 のコントラストにおいて脳活動の違いがかなり大きいことがわかり，「歌声らしさ」が共に評価が高くても，合成音の歌声はまだ実際の歌声には及ばないことが判明した．

実験 I では，Base，VR1，VR2 などの合成音を聞いたときよりも，実際の音声である Sing と Speak を聞いたときの脳活動が大きいことがわかった．また，合成音である Base，VR1，VR2，VR3 を聞いた際の脳活動は弱く，スペクトル形状やヴィブラートの違いが脳活動に及ぼす影響は，脳活動の大きな差異がないために判明しなかった．原因として，合成音の「自然性」が問題であったと思われる．

2.6 まとめ

本章では、脳活動測定実験Ⅰで用いる刺激音の作成方法、刺激音の聴取実験、脳活動測定実験Ⅰについて述べた。

脳活動測定実験Ⅰの目的は、実際の歌声と話声の脳活動の違いを調べることであり、また、「歌声」と知覚される具体的な音響的特徴の物理量を調べるため、歌声特有の音響的特徴を加えた合成音の脳活動に与える影響を調べることであった。そこで、刺激音は実際の歌声と話声 (Sing と Speak) と、話声に歌声の音響的特徴を適宜加えて作成した合成音 (Base, VR1, VR2, VR3) を用いた。

それら刺激音に対して「歌声らしさ」・「自然性」の聴取実験を行った結果、歌声の音響的特徴を加えた合成音 VR2 は、本物の歌声 Sing と同様の「歌声らしさ」の評価が得られた。合成音 Base は、歌声のスペクトル形状を加えることで、Speak よりも「歌声らしさ」の高い評価を得ており、また、「自然性」の評価においても、Speak と Sing などの本物の音声と同程度の評価を得られた。他のヴィブラートを加えた合成音 VR1, VR3 は、他の刺激音に比べて「歌声らしさ」・「自然性」ともに低い評価であるものの、刺激音作成の目的である歌声と話声を補間する合成音を作成することが出来た。

脳活動測定実験Ⅰでは、それらの刺激音を聞いた際の脳活動を測定し、それぞれの脳活動の違いを解析した。ヒトが発した歌声 (Sing) と話声 (Speak) において、それぞれ活動する脳部位が異なることが判明した。しかし、合成音を聞いた際の脳活動は弱く、音響的特徴の違いによる脳活動の違いは明らかにならなかった。

聴取実験と脳活動結果の関係については、「自然性」の評価が高い刺激音は、全体的に脳活動が強い傾向が見られたが、「歌声らしさ」に関する脳活動との関係は明らかにならなかった。

次章では、刺激音セットの異なる脳活動測定実験Ⅱについて記述する。

第3章 歌声と話声の音声知覚に関する 脳活動測定実験Ⅱ

3.1 はじめに

本章では、非言語情報のみ異なる刺激音作成と、刺激音が意識上でどのように知覚されているのか評価を行った聴取実験、刺激音を聞いた際の脳活動を測定する実験Ⅱについて記述する

3.2 実験Ⅱの目的

実験Ⅱは、歌声の重要な要素である基本周波数 (F0)、スペクトル形状、振幅エンベロープが、脳において「歌声」と知覚される上でどれほど寄与するのかを調べるために行う。

3.3 実験Ⅱで用いる刺激音

3.3.1 歌声に関する非言語情報のみ異なる刺激音の作成方法

実験Ⅱは、歌声の重要な要素である基本周波数 (F0)、スペクトル形状、振幅エンベロープが、脳において「歌声」と知覚される上でどれほど寄与するのかを調べるために行う。そこで、実験Ⅱで用いる刺激音は、実際にヒトが発した歌声と話声の基本周波数 (F0)、スペクトル形状、振幅エンベロープをそれぞれ取り出し、入れ替えて再合成した刺激音を作成する。

非言語情報が与える脳活動を調べるために、歌声に関する非言語情報のみ異なる刺激音を作成する。刺激音作成には、高品質な分析合成系 STRAIGHT [5] を用いて作成する。実験Ⅰでは、話声に歌声特有のスペクトル形状やヴィブラートなどの音響的特徴を加え、歌声らしく知覚される合成音を作成した。しかし、そのように作成した合成音では「自然性」が低く、それら合成音を聞いたときの脳活動は全体的に弱かった。そこで、実験Ⅰで問題であったと思われる「自然性」を考慮して、実験Ⅰの刺激音とは異なり、実際にヒトが発した話声 (Speak) と歌声 (Sing) から各 F0、スペクトル形状、振幅エンベロープを抽出し、組み替えて作成することで、より自然な合成音を作成する。実際にヒトが発した話声と歌声は、歌声データベース「日本語を歌・唄・謡う」の中から選定した男性テノ

ル歌手の音声を用いている [16] . 実験 I で使用した Speak と Sing と同じものである . ただ , Sing を 1.79 sec から 1.74 sec に継続時間を制御し , すべて同じ継続時間を持つものとした . 図 3.1 において , 入力音声からの F0 , スペクトル形状 , 振幅エンベロープの抽出手順を示し , 図 3.2 において , 抽出した F0 , スペクトル形状 , 振幅エンベロープからの合成手順を示す .

また , Speak と Sing から抽出した音響的特徴をそれぞれ図 3.3 と図 3.4 に示す .

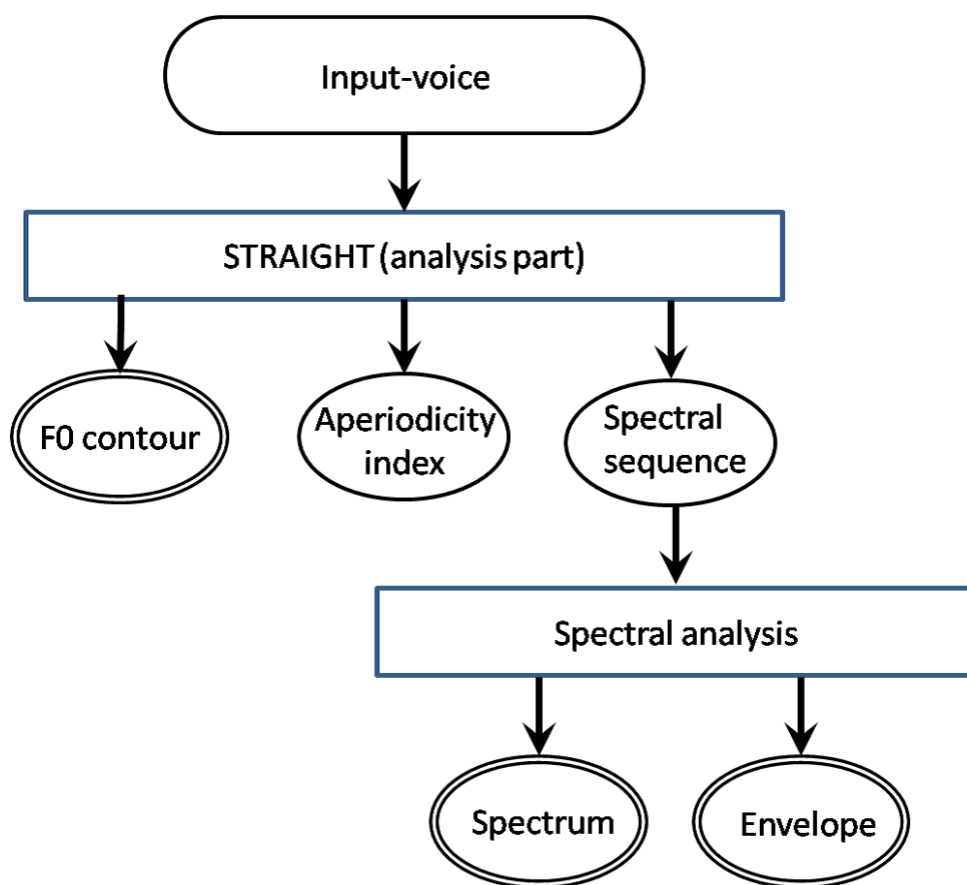


図 3.1: 各音響的特徴抽出の概要

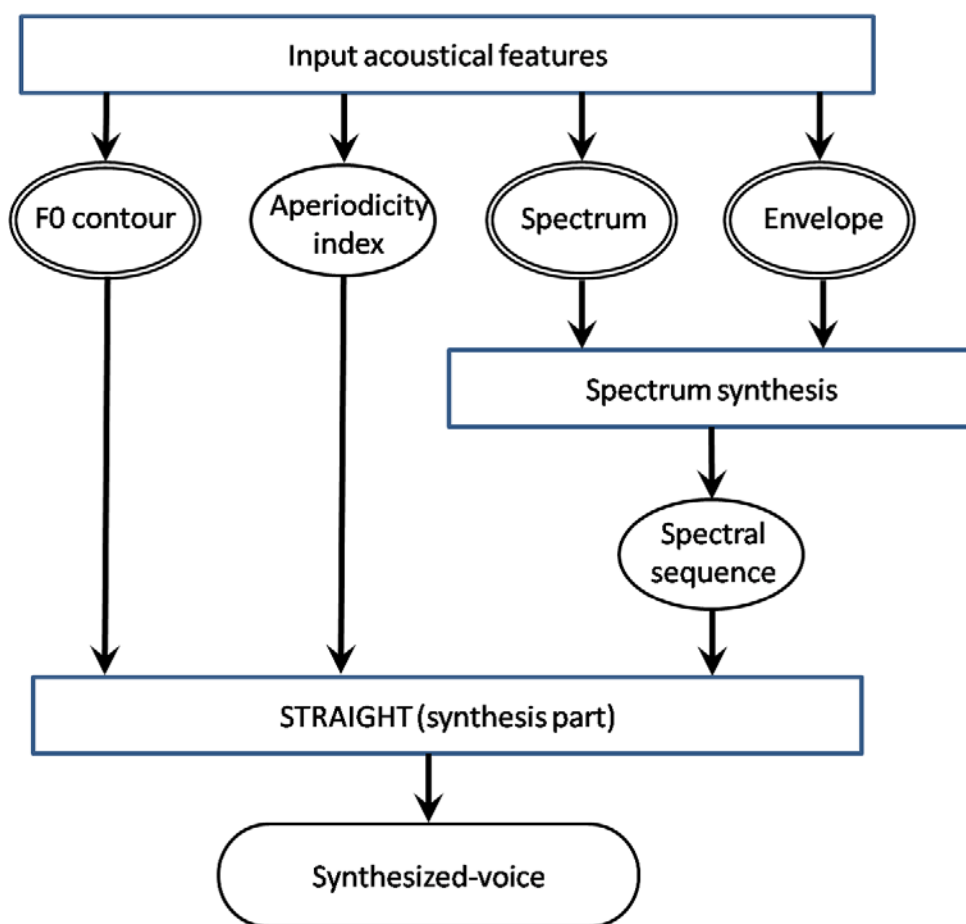


図 3.2: 各音響的特徴合成の概要

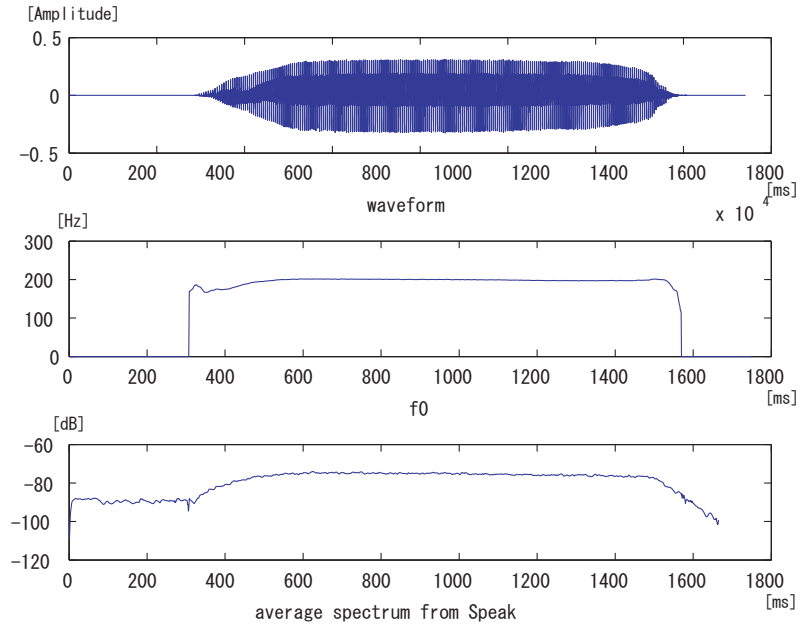


図 3.3: Speak の波形 (上図) と F_0 (中央図) と振幅エンベロープ (下図) .

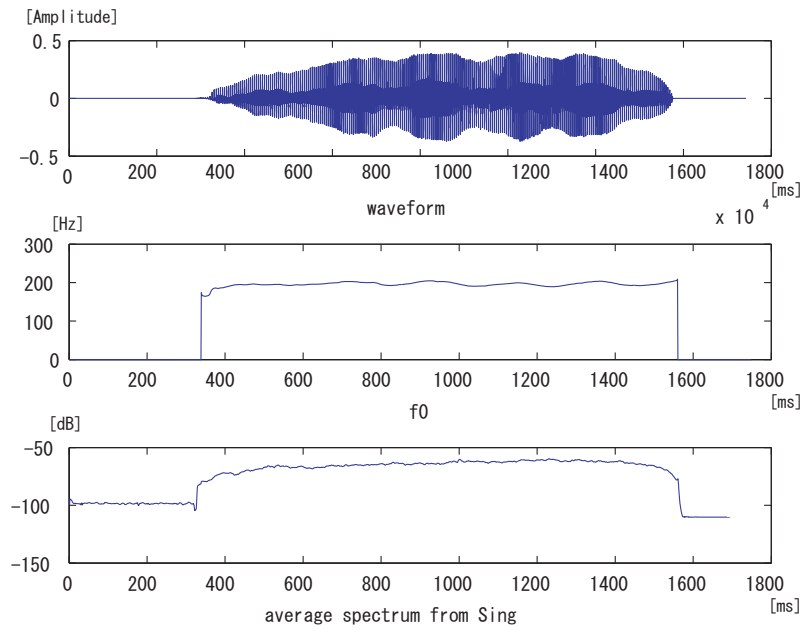


図 3.4: Sing の波形 (上図) と F_0 (中央図) と振幅エンベロープ (下図) .

3.3.2 作成した刺激音

上述した刺激音の作成方法を使って，作成した刺激音の名前と構成要素を表3.1に示す．表3.1に示すとおり，非言語情報のみ異なる刺激音である．

表 3.1: 刺激音の名前と構成要素

Number	Name	F0	Spectrum	Envelope
No.1	sp-sp-sp	Speak	Speak	Speak
No.2	sp-si-sp	Speak	Sing	Speak
No.3	sp-si-si	Speak	Sing	Sing
No.4	si-sp-sp	Sing	Speak	Speak
No.5	si-sp-si	Sing	Speak	Sing
No.6	si-si-si	Sing	Sing	Sing

以下に6つの刺激音の名称と特徴を記述する．

- **sp-sp-sp**
F0，スペクトル，振幅エンベロープが全て Speak から成る
- **sp-si-sp**
スペクトルのみ Sing から成る
- **sp-si-si**
F0 のみ Speak から成る
- **si-sp-sp**
F0 のみ Sing から成る
- **si-sp-si**
スペクトルのみ Speak から成る
- **si-si-si**
F0，スペクトル，振幅エンベロープが全て Sing から成る

すべての刺激音の音圧は同一，継続時間は 1.74 sec である．

3.4 聴取実験

本研究は，合成音を用いた歌声に関する脳活動の研究であるため，刺激音の「歌声らしさ」をあらかじめ評価する必要がある．また，刺激音がヒトの音声として聞こえるかを評

価するため、「自然性」もあらかじめ評価する必要がある．また，聴取実験の評価結果と脳活動測定実験の結果を比較し議論する．

3.4.1 聴取実験の方法

実験Ⅰで行った聴取実験の方法と同じく，上述の6種類の刺激音にたいして，「歌声らしさ」と「自然性」についてそれぞれ評価してもらう．評価方法は，2つ連続して呈示される刺激音に対して，どちらの刺激音がより「歌声らしい」か，もしくは「自然性」が高いかを5段階評価でそれぞれ採点してもらい，シェッフェの一対比較法を用いて評価する[8]．心理物理実験において，一対比較法は，数個の刺激を2つずつ対にして判断を求める方法である．

刺激条件

実験で用いる聴覚刺激は，先に示した6種類の刺激音を2つずつ対にしたものである．刺激音が6種類あるので，刺激対の数は，順序効果も考慮した $6 \times 5 = 30$ 対である．

実験参加者

実験Ⅰの聴取実験と同じく「歌声らしさ」「自然性」の評価を同じ方法で行う．被験者は正常な聴力をもつ9人（男性8人，女性1人）．被験者の年齢は20代（男性7人女性1人）である．

実験環境

実験Ⅰで行った聴取実験と同じ実験環境である．

実験手続き

実験Ⅰで行った聴取実験と同じく「歌声らしさ」と「自然性」に関して評価してもらった．なお，各実験参加者につき，30対の1セットを3回行い評価してもらう．

表 3.2: 母数の推定 (歌声らしさ) .

刺激音	母数
No.1	-1.00
No.2	-0.19
No.3	-0.10
No.4	0.14
No.5	0.25
No.6	0.9

表 3.3: 母数の推定 (自然性) .

刺激音	母数
No.1	0.7
No.2	-0.83
No.3	-0.76
No.4	0.62
No.5	0.24
No.6	0.04

3.4.2 聴取実験結果と考察

上記の実験方法で得られた「歌声らしさ」と「自然性」のデータの結果を、それぞれ表 3.2, 表 3.3 に示す。母数の値は、刺激音がどれだけ「歌声らしい」もしくは「自然性」が高く聴こえたかを表す値であり、正の大きな値であるほど評価が高い。

また、表に示した母数の値に従って、5 つの刺激の距離関係を直線上で示したものが、「歌声らしさ」においては図 3.5、「自然性」では図 3.6 になる。また、「歌声らしさ」と「自然性」の評価結果をまとめた結果を Fig.3.7 に示す。

尚、F 検定法による刺激間の有意差検定を行った結果「歌声らしさ」においては、sp-si-sp と sp-si-si、sp-si-sp と si-sp-sp、sp-si-si と si-sp-si、si-sp-sp と si-sp-si の刺激音間以外で、他の各刺激音間に 5% の水準で有意な差が確認された。「自然性」においては、sp-si-sp と sp-si-si、si-sp-si と si-si-si、sp-sp-sp と si-sp-sp の刺激音間以外で、他の各刺激音間に 5% の水準で有意な差が確認された。

図 3.5 より、「歌声らしさ」において、すべて Speak の要素である合成音 sp-sp-sp が最も歌声らしくなく、すべてが Sing の要素の合成音である si-si-si が最も歌声らしいことが確認された。他の合成音は sp-sp-sp の評価と si-si-si の評価の間に位置することもわかる。また、「歌声らしさ」という心理量に寄与する音響的特徴は、F0、スペクトル、振幅エン

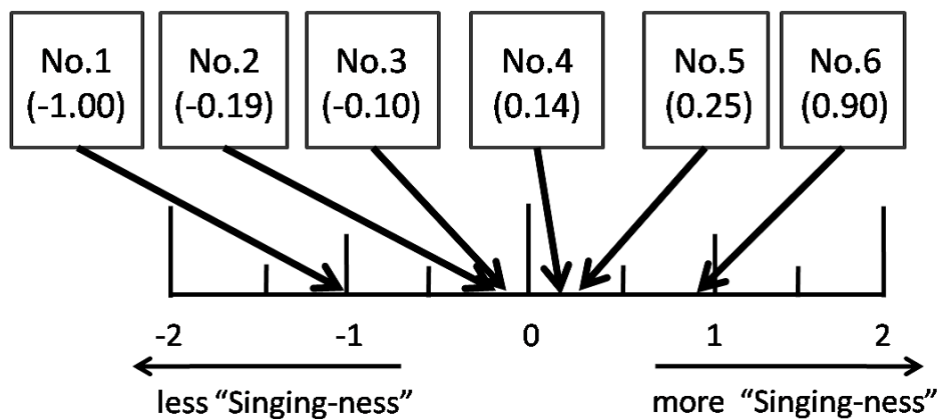


図 3.5: 「歌声らしさ」の関係 .

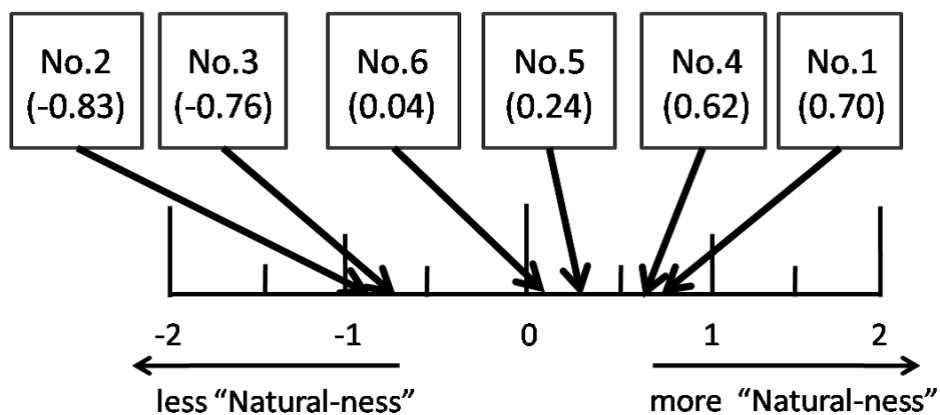


図 3.6: 「自然性」の関係 .

ベロープの順に強いことがわかる．なお，有意差検定の結果から振幅エンベロープの違いによる影響は少ないということが明らかになった．

また，「自然性」に関しては，特にスペクトル形状が Sing の要素である場合に音声として不自然になる傾向が見られるが，それ以外の合成音は「自然性」が高い結果が出ており，品質の良い刺激音が作成できたと考えられる．

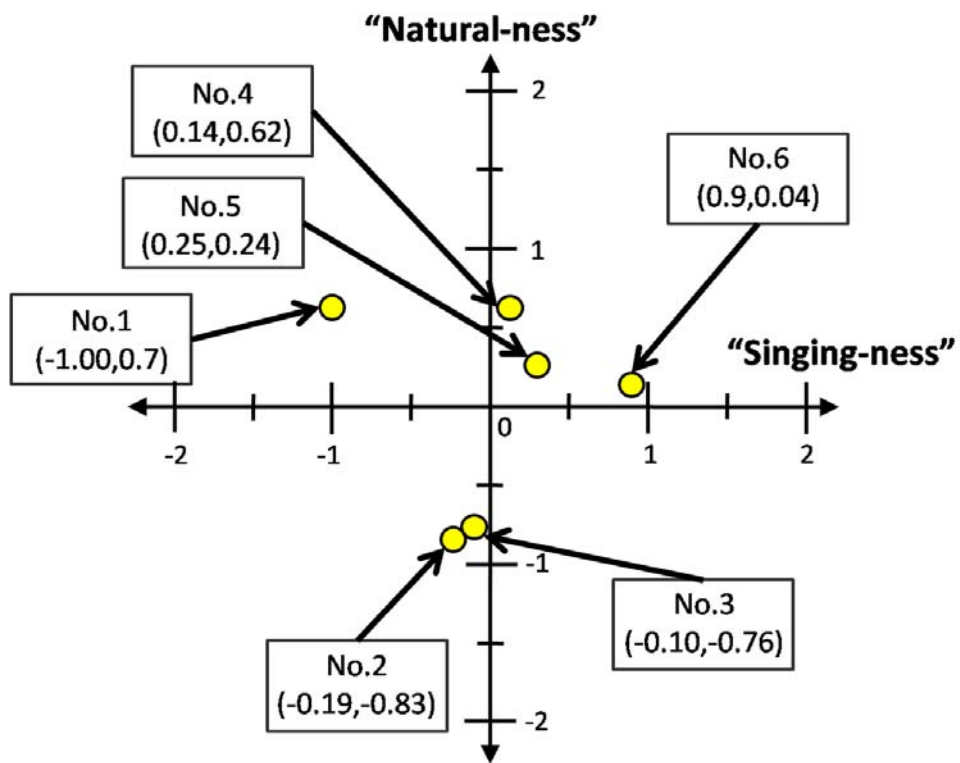


図 3.7: 「歌声らしさ」と「自然性」の関係 .

3.5 脳活動測定実験Ⅱ

上述した各刺激音を聞いた際の脳活動を測定し，歌声と話声，F0 やスペクトル形状，振幅エンベロープの違いが脳活動に及ぼす影響を調べるために，脳活動測定実験Ⅱを行う．そこで，歌声と話声の脳活動は異なるのか，音響的特徴の違いで脳活動は異なるのか，聴取実験の結果と脳活動結果に関係があるのかどうかを調べる．

3.5.1 脳活動測定実験Ⅱの方法

呈示方法

第3章で記述した脳活動測定実験Ⅰと同じ方法で行う．

実験参加者

被験者は正常な聴力をもつ16人（男性11人，女性5人）であり，年齢は50代の男性1人，40代の女性2人，30代男性2人，20代の男性8人，女性3人である．

実験装置

第3章の脳活動測定実験Ⅰと同じ実験装置である．

3.5.2 解析手法

第3章の脳活動測定実験Ⅰと同じ解析手法で解析を行う．なお，有意な脳活動部位は，実験Ⅰと同じくP値が0.001以上で，活動のクラスタ数が3以上の部位とし，活動部位を脳アトラス（『ATLAS OF THE HUMAN BRAIN』[28]）で調べて，脳部位名と脳座標（Talairach座標系）を記述する．

3.5.3 解析結果

各刺激音を聞いた際の脳活動の差を解析した結果を記述する．各6種類の刺激音を聞いたときの脳活動の違いを解析する．歌声と話声の脳活動の違い，また歌声に重要な音響的特徴が脳活動に与える影響を調べるために，以下の4つの脳活動の違いを調べる．

- 歌声（si-si-si）と話声（sp-sp-sp）の違い
- F0 のみの違い

- スペクトル形状のみの違い : 「響き」の違い
- F0 と振幅エンベロープの違い : 「揺れ」の違い

「響き」とは、歌声特有の singer's formant や高調波成分などを含むスペクトルの聴覚印象であり、「揺れ」とは、F0 と振幅エンベロープの変動の聴覚印象である。これらは、歌声らしさの知覚を構成する基本的な心理的特徴である [7]。

歌声 (si-si-si) と話声 (sp-sp-sp) の違い

刺激音 si-si-si と sp-sp-sp をそれぞれ聞いた際の脳活動の違いを解析したところ、表 3.4 に示す脳部位において違いがみられた。話声の sp-sp-sp よりも、歌声の si-si-si を聞いた際に強く活動した脳部位は、MOrG (眼窩回中央) や、SPL (上頭頂小葉)、Cerebellum (小脳) であった (図 3.8, 図 3.9)。一方、si-si-si よりも sp-sp-sp で強く活動した脳部位は、MTG (中側頭回) であった (図 3.10, 図 3.11)。

活動の違いが見られた脳部位の MOrG (眼窩回中央) は、情動系の神経回路の一部として考えられている眼窩回に属する [11][36]。SPL (上頭頂小葉) は、体性感覚野と頭頂連合野に関係する [43]。Cerebellum (小脳) は、運動調節機能を司る部位であり [4]、心的イメージング、情動の調節、言語処理といった多くの認知課題を行っているときにも小脳が働いていることが報告されている [21]。MTG (中側頭回) は、側頭葉に属し、音韻の処理に関係すると考えられている [26][32][40][43]。

表 3.4: si-si-si minus sp-sp-sp: MOrG = medial orbital gyrus; SPL = superior parietal lobule; MTG = middle temporal gyrus

コントラスト	脳活動部位	クラスタ数
(si-si-si) - (sp-sp-sp)	MOrG [-18, 42, -8]	3
	SPL [18, -51, 56]	4
	Cerebellum [-3, -48, -28]	11
	Cerebellum [12, -39, -28]	10
(sp-sp-sp) - (si-si-si)	MTG [57, -42, 4]	4

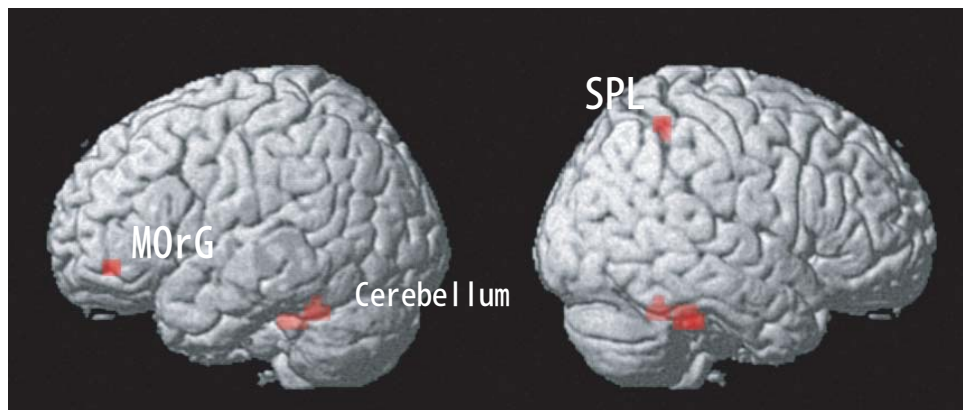


図 3.8: (Listening to si-si-si) minus (Listening to sp-sp-sp) の脳活動差 (左図: 左脳) (右図: 右脳) 赤点が活動の異なる脳部位であり, 脳活動を表面にマッピングして表示している

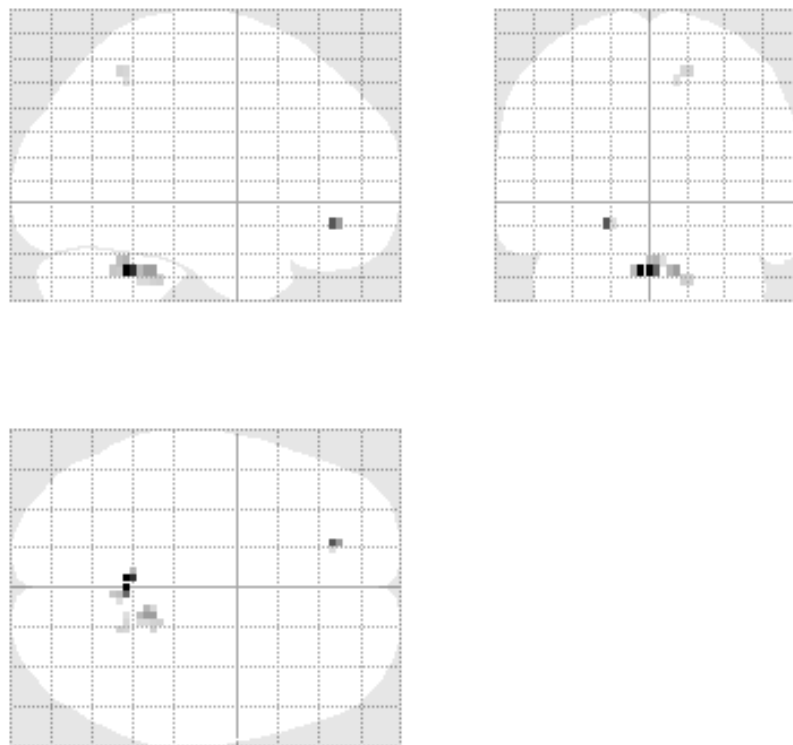


図 3.9: (Listening to si-si-si) minus (Listening to sp-sp-sp) の脳活動差 (左上: 側面図, 左下: 上面図, 右上: 前面図)

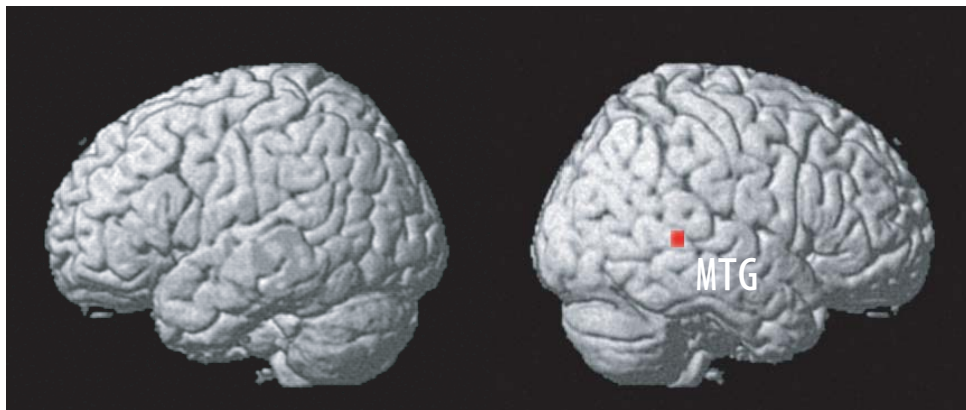


図 3.10: (Listening to sp-sp-sp) minus (Listening to si-si-si) の脳活動差 (左図：左脳) (右図：右脳) 赤点が活動の異なる脳部位であり，脳活動を表面にマッピングして表示している

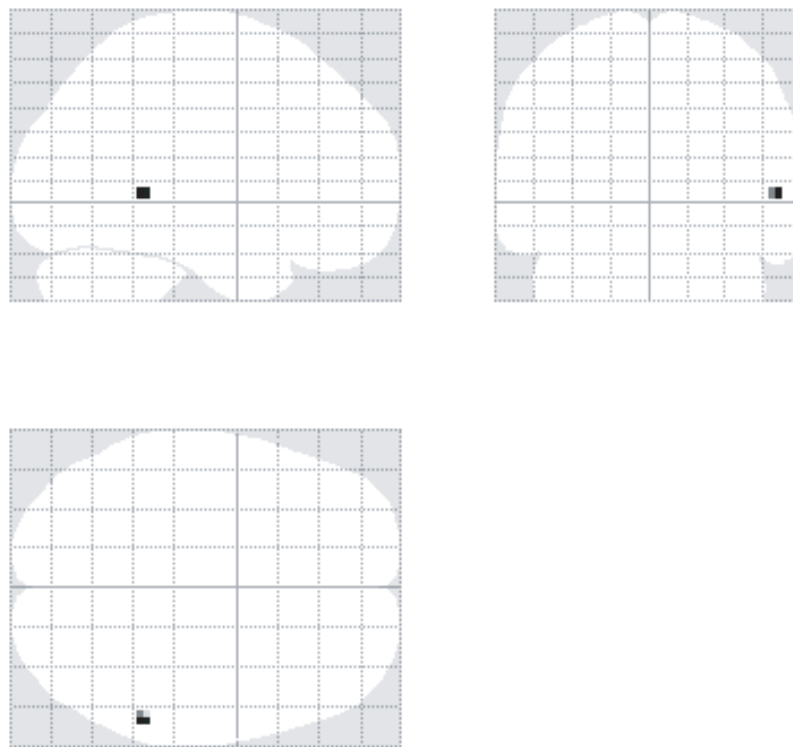


図 3.11: (Listening to sp-sp-sp) minus (Listening to si-si-si) の脳活動差 (左上：側面図，左下：上面図，右上：前面図)

F0 のみの違い

sp-sp-sp と si-sp-sp におけるコントラストなどの F0 のみ異なる刺激音の脳活動差を解析したところ，歌声の F0 を要素に持つことによって，ある脳部位を活性化させる結果となった (図 3.12，図 3.13)．表 3.5 に示されているように，CG (帯状回)，Ins (島)，SMG (縁上回)，STG (上側頭回)，MTG (中側頭回)，PCun (楔前部) などで有意な活動差を示した．一方，話声の F0 を要素に持つことによって，活性化する有意な脳部位はなかった．

表 3.5: F0 の違いによる脳活動の違い: CG = cingulate gyrus; MFPG = middle frontopolar gyrus; POp = parietal operculum; Ins = insula; TTG = transverse temporal gyrus/gyri; POTZ = parietooccipital transition zone; SMG = supramarginal gyrus; STG = superior temporal gyrus; OcG = occipital gyrus; ITG = inferior temporal gyrus; PCun = precuneus

コントラスト	脳活動部位	クラス数
(si-si-si) - (sp-si-si)	CG [-3, 6, 40]	71
	CG [-6, 30, 28]	
	CG [12, -54, 20]	7
	MFPG [-27, 42, 8]	12
	POp [48, -24, 16]	17
	Ins [42, -18, 4]	
	TTG [42, -27, 8]	
	POTZ [12, -87, 40]	11
	POTZ [36, -48, 20]	20
	SMG [45, -48, 16]	19
	STG [45, -45, 16]	
	STG [57, -39, 16]	7
	OcG [-21, -81, -4]	26
	ITG [42, -78, -4]	5
	PCun [-3, 60, 12]	7
(sp-si-si) - (si-si-si)	nothing	-

CG (帯状回) は，大脳辺縁系に属する脳部位である．大脳辺縁系は食欲，性欲などの本能，快・不快などの情動に関係する部位を含んでおり，その一部が帯状回である [15]．また，他の聴覚系の脳活動測定実験において，リズムより音色に注目することで活性化するという報告がある [9][10]．

Ins (島) は大脳辺縁系へ感情や感覚などを送る通路と考えられてる部位であり [23]，とくにネガティブな身体的反応に関連が深いことが知られており，情動に関係する扁桃体とも関係が深い [19][20]．

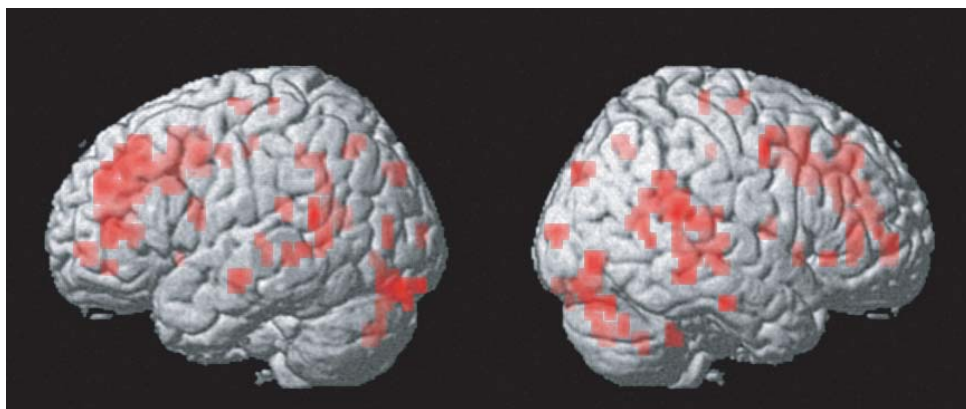


図 3.12: F0 の違いによる脳活動差 (左図：左脳) (右図：右脳) 赤点が活動の異なる脳部位であり，脳活動を表面にマッピングして表示している

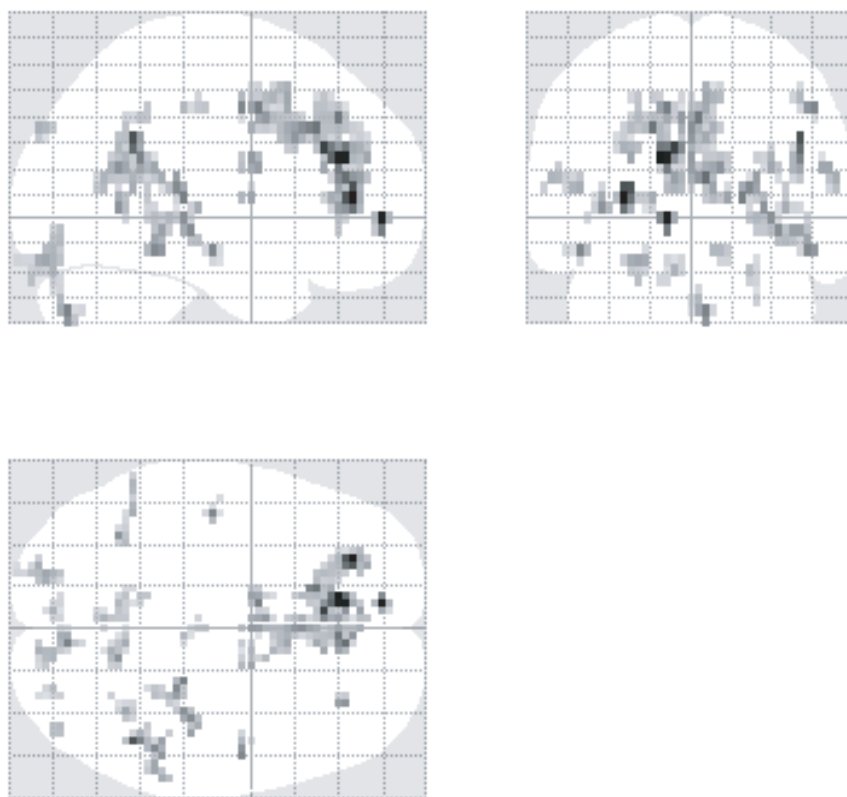


図 3.13: F0 の違いによる脳活動差 (左上：側面図，左下：上面図，右上：前面図)

SMG (縁上回) は音韻の系列処理に関する部位と考えられている [34] .

MTG (中側頭回) は音韻の処理に関する部位と考えられている [26][32][40][43] .

PCun (楔前部) は、詳しい機能は分かっていないが、Blood らの行った実験において協和音が増加するにつれて活性化した報告があり [27] , JermeY らの行った実験では、話している声と顔を見聞きし聴覚と視覚で話を知覚するよりも、聴覚単独で話を知覚した際により強く活動した報告がある [43] .

STG(上側頭回) は、言語の理解を担うウェルニッケ野の一部であり、構文処理や発話処理にも関係する部位と考えられている [37][38][44] . また、ある文献では知覚処理に関するといわれている [24][32][41][43][46] . また、母国語より第二外国語においてより活動するとい報告もある [31] . さらに、PET を用いた研究でスペクトル変化に関する部位は STG 前部であるという報告もあるが [45] , 今回の解析ではスペクトル変化によって活動差は示されてない .

スペクトル形状のみの違い

sp-si-sp と sp-sp-sp におけるコントラストなどのスペクトル形状のみ異なる刺激音の脳活動差を解析したところ、スペクトル形状の違いによって、脳活動が異なる部位が判明した (図 3.15 , 3.14) . 表 3.6 に示すように、CG (帯状回) , Cd (尾状核) , Cerebellum (小脳) , IG(島皮質) , Ins(島) などで有意な活動差を示した .

CG (帯状回) は、大脳辺縁系に属する . 大脳辺縁系は食欲、性欲などの本能、快・不快などの情動に関係する部位を含んでおり、その一部が帯状回である [15] . Cd (尾状核) は大脳皮質と視床、脳幹を結び付けている神経核の集まりである大脳基底核の一部であり、多くのドーパミン受容体があり、運動調節、認知機能、感情、動機付けなど様々な機能を担っている [2][15] . Cerebellum (小脳) は運動調節機能を担う [4] . Ins (島) と IG (島皮質) は大脳辺縁系へ感情や感覚などを送る通路と考えられてる部位であり [23] , とくにネガティブな身体的反応に関連が深いことが知られており、情動に関係する扁桃体とも関係が深い [19][20] .

表 3.6: スペクトル形状による脳活動の違い: CG = cingulate gyrus; Cd = caudate nucleus; IG = insular gyrus; Ins = insula; APul = anterior rectus capitis muscle; LgG = lingual gyrus; SFGM = superior frontal gyrus, lateral part; MFPG = middle frontopolar gyrus; MD = medial dorsal thalamic nucleus

コントラスト	脳活動部位	クラスタ数
(sp-si-sp) - (sp-sp-sp)	CG[12, -36, 20]	25
	CG[3, -24, 20]	
	Cd [-18, -15, 28]	23
	Cd [24, -33, 12]	10
	Cerebellum [-33, -57, -40]	16
(sp-sp-sp) - (sp-si-sp)	IG [30, 9, 12]	3
	Ins [-33, -12, 20]	3
	APul [15, -24, 4]	5
	LgG [9, -60, 0]	7
(si-si-si) - (si-sp-si)	SFGM [12, 9, 48]	5
	MFPG [-27, 60, 8]	5
	MD [3, -18, 8]	5
(si-sp-si) - (si-si-si)	nothing	-

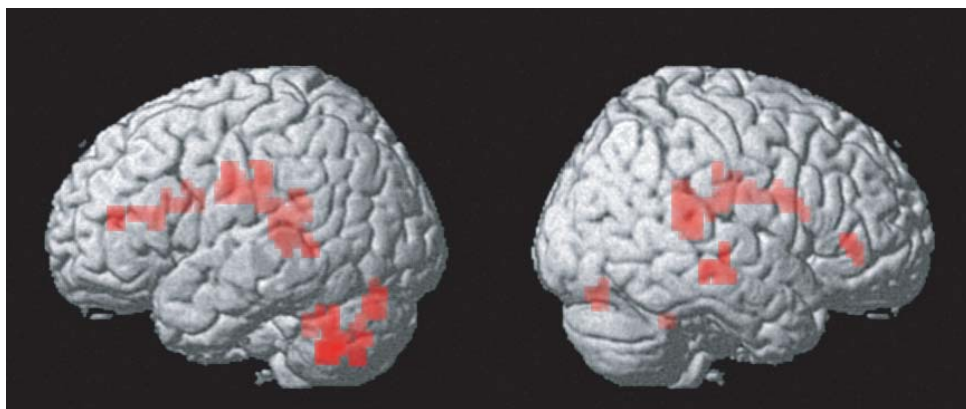


図 3.14: スペクトルの違いによる脳活動差 (左図：左脳，右図：右脳)

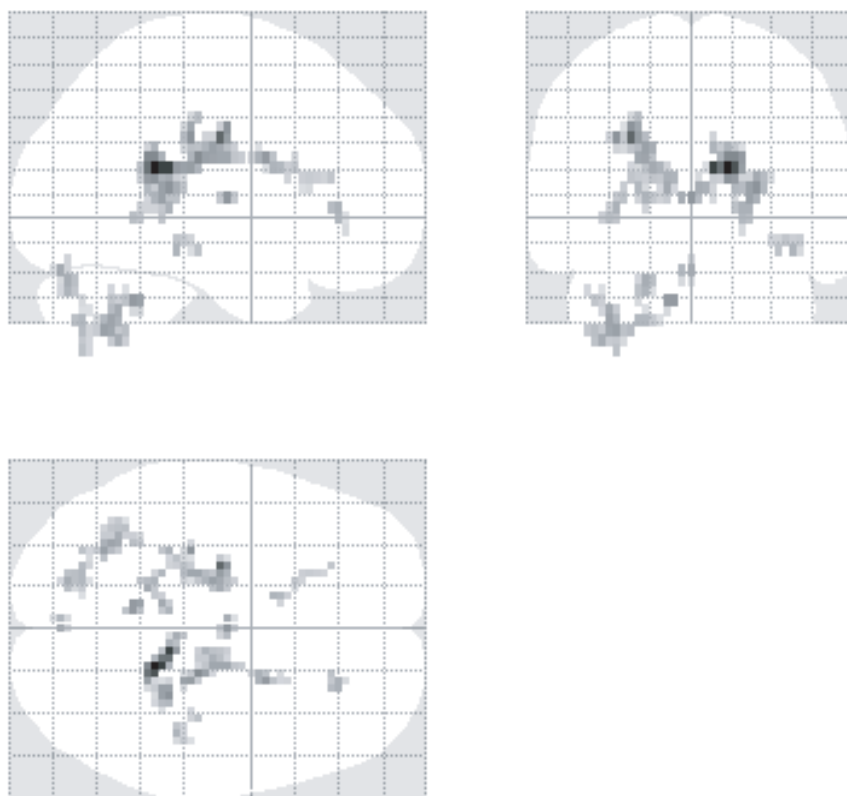


図 3.15: スペクトルの違いによる脳活動差 (左上：側面図，左下：上面図，右上：前面図)

F0 と振幅エンベロープの違い

si-si-si と sp-si-sp におけるコントラストなどの F0 と振幅エンベロープの異なる刺激音の脳活動差を解析したところ，表 3.7 に示すように有意な脳活動の差が見られた (図 3.15, 3.14) . 歌声の F0 と振幅エンベロープのものを聞いたときにより，IFGOr (下前頭回の眼窩部)，SG (直回)，MFPG において活動が強くなった．一方，話声の F0 と振幅エンベロープでは，STG (上側頭回)，IFGTr(下前頭回の三角部)，Cerebellum (小脳) などで活動の違いが見られた．

表 3.7: F0 と振幅エンベロープの違いによる脳活動の違い: IFGOr = inferior frontal gyrus, orbital part; SG = straight gyrus; MFPG = middle frontopolar gyrus; STG = superior temporal gyrus; IFGTr = inferior frontal gyrus, triangular part

コントラスト	脳活動部位	クラスタ数
(si-si-si) - (sp-si-sp)	IFGOr [-39, 30, -12]	9
	SG [-6, 36, -12]	29
	MFPG [-21, 60, 12]	3
(sp-si-sp) - (si-si-si)	STG [60, -45, 32]	13
	STG [66, -51, 24]	
	IFGTr [54, 21, 12]	7
	Cerebellum [-39, -57, -48]	10

IFGOr (下頭前回の弁蓋部)，IFGTr(下前頭回の三角部) は下頭前回の一部であり，ブローカ領域に属する [12] . SG は (直回) は前頭葉の眼窩回の内部にある脳回である [4] . MFPG は，前極回の中中部であり，どのような機能があるのか詳しく判明していない．STG(上側頭回) は，言語の理解を担うウェルニッケ野の一部であり，構文処理や発話処理にも関係する部位と考えられている [37][38][44] .

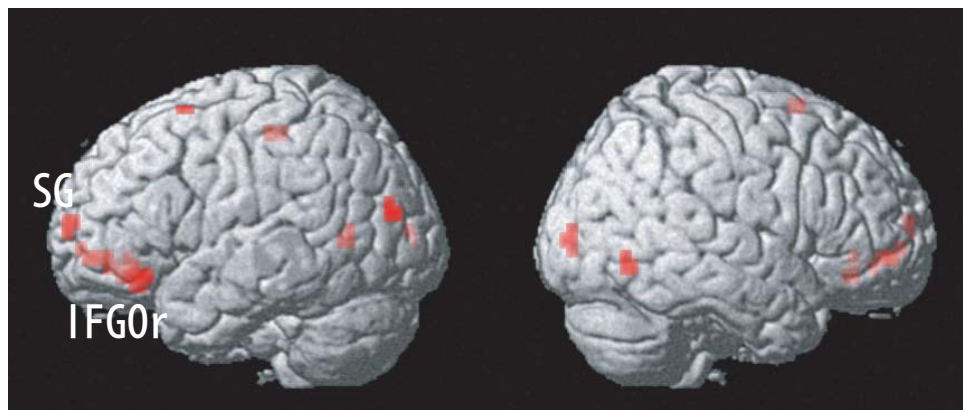


図 3.16: (Listening to si-si-si) minus (Listening to sp-si-sp) の脳活動差 (左図: 左脳) (右図: 右脳) 赤点が活動の異なる脳部位であり, 脳活動を表面にマッピングして表示している

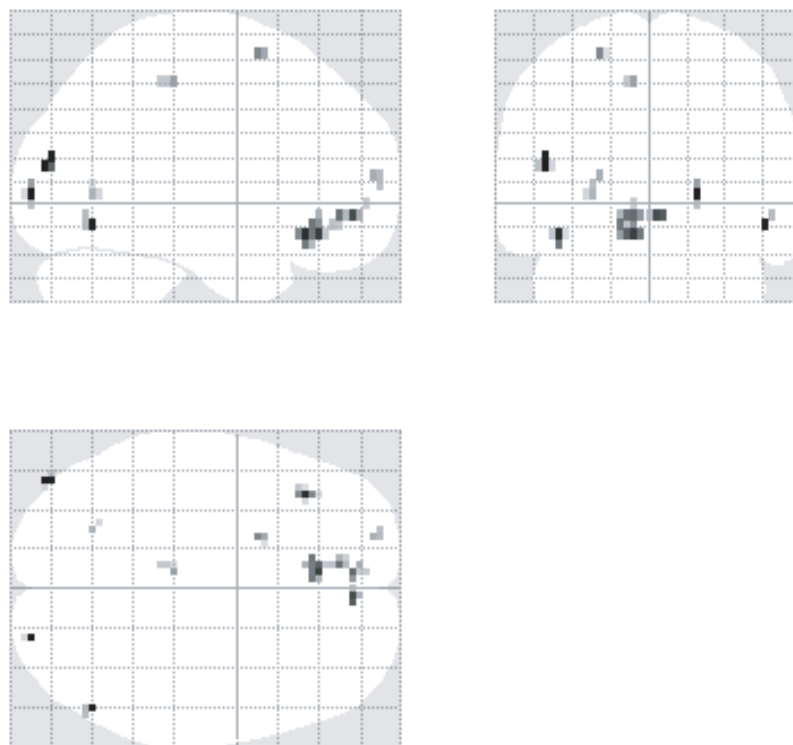


図 3.17: (Listening to si-si-si) minus (Listening to sp-si-sp) の脳活動差 (左上: 側面図, 左下: 上面図, 右上: 前面図)

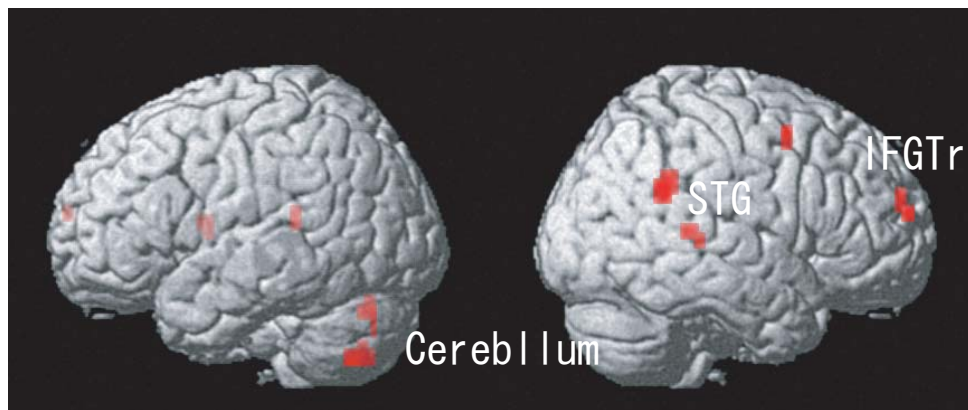


図 3.18: (Listening to sp-sp-sp) minus (Listening to si-si-si) の脳活動差 (左図: 左脳) (右図: 右脳) 赤点が活動の異なる脳部位であり, 脳活動を表面にマッピングして表示している

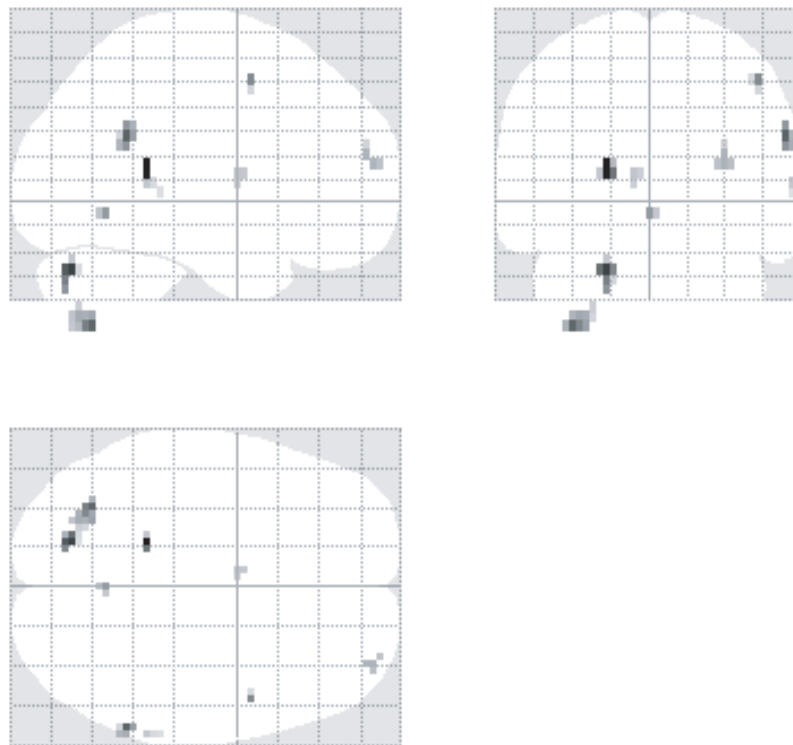


図 3.19: (Listening to sp-si-sp) minus (Listening to si-si-si) の脳活動差 (左上: 側面図, 左下: 上面図, 右上: 前面図)

3.5.4 脳活動測定実験結果の考察

本章では、歌声に関する刺激音を聞いた際の脳活動を測定し、それぞれの脳活動の違いを解析した。全て歌声の要素である刺激音 si-si-si と全て話声の要素である sp-sp-sp の脳活動の違いを解析した結果、眼窩回や小脳の部位に活動の違いが見られた。眼窩回の脳部位は、情動にも深く関係のある部位である。また、F0 やスペクトル形状、振幅エンベロープなどの音響的特徴それぞれが異なることにより、脳活動が異なることも判明した。主に、CG (帯状回) や Cd (尾状核)、Ins (島) などの脳内部での脳部位に活動差が見られた。これらの部位は、大脳辺縁系や大脳基底核などの脳内部に属する部位であり、情動に関係する部位である。

また、歌声の音響的特徴を含んだ刺激音の方が、話声の音響的特徴を含んだ刺激音よりも脳活動が強い傾向にあった。つまり、聴取実験と脳活動結果の関係については「歌声らしさ」が高いほど、全体の脳活動は強い傾向が見られた。しかし、「自然性」に関しては関係が見られなかった。

3.6 まとめ

本章では、脳活動測定実験Ⅱで用いる刺激音の作成方法、刺激音の聴取実験、脳活動測定実験Ⅱについて述べた。

実験Ⅱの目的は、実際のヒトの歌声と話声の基本周波数 (F0)、スペクトル形状、振幅エンベロープが脳活動に与える影響を調べることであった。そこで、実験Ⅱで用いる刺激音は、実際にヒトが発した歌声と話声の基本周波数 (F0)、スペクトル形状、振幅エンベロープをそれぞれ取り出し、入れ替えて刺激音を作成した。

それら刺激音に対して聴取実験を行った結果「歌声らしさ」において、すべて Speak の要素である合成音 sp-sp-sp が最も歌声らしくなく、すべてが Sing の要素の合成音である si-si-si が最も歌声らしく、他の合成音は sp-sp-sp の評価と si-si-si の評価の間に位置する結果となった。また、「自然性」に関しては、特にスペクトル形状が Sing の要素である場合に音声として不自然になる傾向が見られたものの、それ以外の合成音は「自然性」が高い結果が出ており、品質の良い刺激音が作成できた。

脳活動測定実験Ⅱでは、それらの刺激音を聞いた際の脳活動を測定し、それぞれの脳活動の違いを解析した。歌声 (si-si-si) と話声 (sp-sp-sp) において、それぞれ活動する脳部位が異なることが判明した。また、音響的特徴の違いによる脳活動の違いも解析することが出来た。聴取実験と脳活動結果の関係については「歌声らしさ」が高いほど、全体の脳活動は強い傾向が見られたが、「自然性」に関しては関係が見られなかった。

次章では、実験Ⅰと実験Ⅱの結果から全体考察を行う。

第4章 歌声と話声に関する非言語情報の違いが引き起こす脳活動の違い

4.1 はじめに

脳活動測定実験Ⅰと実験Ⅱの結果から、歌声と話声にそれぞれ特有の脳活動部位と、音響的特徴が異なることで違う脳活動部位をまとめ、歌声知覚や非言語情報知覚に関して活性化する脳活動の全体考察を行う。

4.2 実際のヒトの歌声と話声を聞いた際の脳活動の違い

実験Ⅰと実験Ⅱの結果より、ヒトが発した歌声と話声において、それぞれ知覚した際の脳活動が異なることが判明した。歌声を聞いたときに、話声を聞いたときよりも強く活動し、実験Ⅰと実験Ⅱに共通する脳部位は、LOrG (側部眼窩回) や MOrG (眼窩回中央) などの眼窩回の一部であった (図 4.1.4.2)。この部位は、前頭葉眼窩皮質の一部であり、扁桃体と密接な繊維連絡を有し、ヒトおよび動物では、電気刺激により呼吸停止、血圧低下、瞳孔拡大などの自律反応や注意反応が起こり、上位中枢として快・不快に反応する扁桃体や、自律神経とホルモンの調節を行う視床下部を含む脳幹の情動行動に關与する領域を制御すると考えられていおり、情動系の神経回路の一部と考えられている [2][11][36][39]。Stefan らが行った研究では、呈示される音楽の曲の構成が正しいか判断するタスクにおいて前頭葉眼窩皮質の部位が活動を示した例もある [37]。

この結果より、ヒトが歌声として知覚する際には、話声を知覚した際よりも、なんらかの感情を引き起こし情動に關する脳部位が活動することで、「歌声」と知覚する可能性があると考えられる。言い換えれば、感情を揺さぶる声を「歌声」として知覚しているとも考えられる。

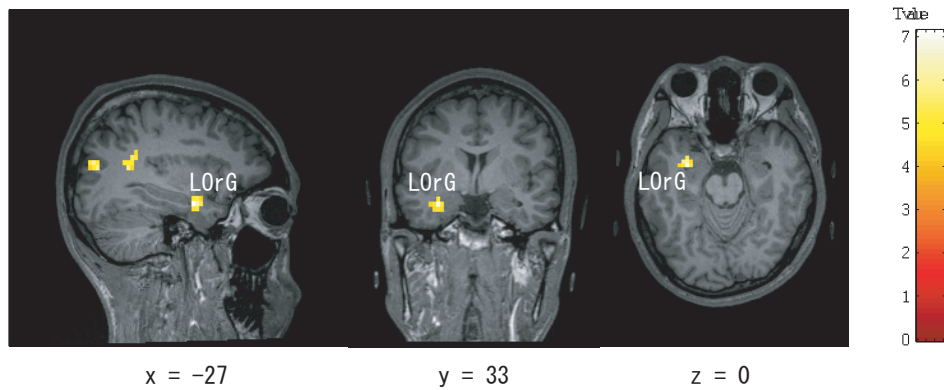


図 4.1: LOrG (側部眼窩回) の位置 (左図：矢状断面図，中央図：正中断面図，右図：水平断面図)

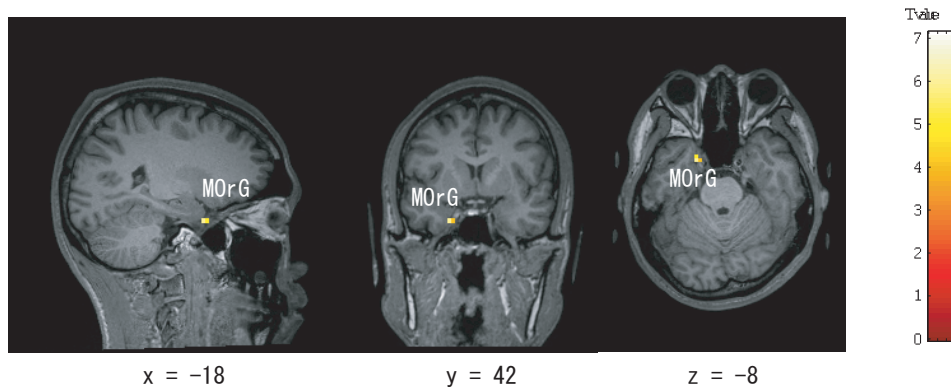


図 4.2: MOrG (眼窩回中央) の位置 (左図：矢状断面図，中央図：正中断面図，右図：水平断面図)

4.3 歌声特有の音響的特徴が与える脳活動の違い

脳活動測定実験で用いた刺激音は，歌声特有の音響的特徴を適宜変化させた合成音も呈示している．実験Ⅰでは，合成音を聞いた際の脳活動が弱く，音響的特徴の違いによる脳活動差は明らかにならなかった．しかし，合成音の作成方法を改良した結果，実験Ⅱでは，F0 やスペクトル形状の違いによる脳活動差を解析することが出来た．そこで，F0 とスペクトル形状の違いによる脳活動部位の違いから，非言語情報の知覚に関する考察を記述する．

4.3.1 F0 の違いが与える脳活動の影響

F0 の違いによる脳活動差は，特に CG (帯状回)，Ins (島) など有意な活動差を示した．

CG (帯状回) は、大脳辺縁系に属し、食欲、性欲などの本能、快・不快などの情動に関係する部位を含んでおり、その一部が帯状回である [15]。特に、活動を示した CG (帯状回) の前部は、感覚刺激の意味に関する情報を、扁桃体、視床背内側核、および前頭葉眼窩皮質などから受け、感情情報を出力情報に転換する過程で重要な役割を果たしていると考えられている部位である [2]。また、PET を用いた楽器の音色の認知についての脳活動測定実験において、リズムに注目して聞くよりも音色に注目して聞いた時の方が、CG (帯状回)、STG 前部 (上側頭回)、扁桃体、海馬傍回、右側の島皮質、下頭頂小葉で有意な活性化が見られるという結果が出ている [9][10]。

また、Ins (島) は大脳辺縁系へ感情や感覚などの信号を送る通路と考えられてる部位であり [23]、好奇心や意欲に関係するとの報告がある [33]、また、とくにネガティブな身体的反応に関連が深いと考えられており、扁桃体や視床下部とも関係が深い部位である [2]。なお、左右の島皮質で発声および歌声の発声のコントロールを分けもって司っているという報告もある [25]。このように、歌声に関する非言語情報の処理は、大脳表面の聴覚野以外に脳内部に属する脳部位を活性化させているのではないかと考えられる。

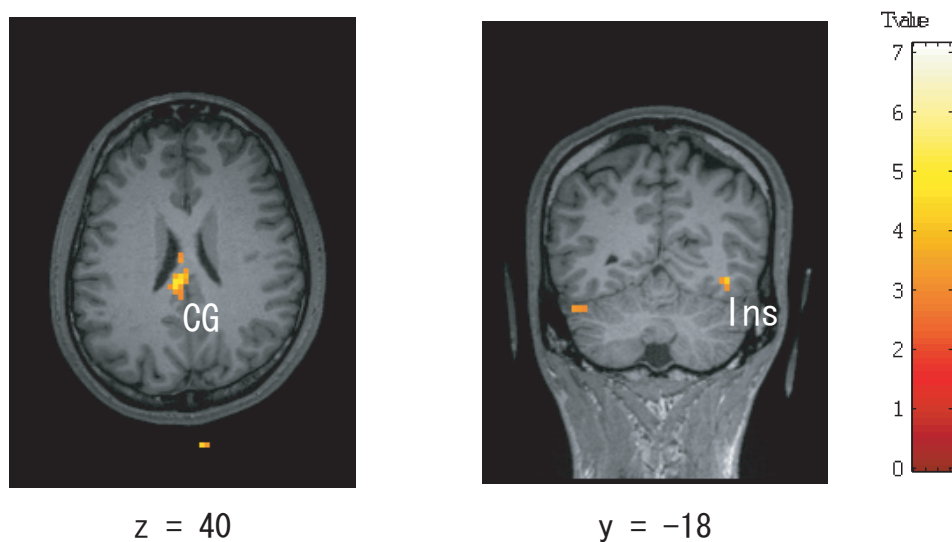


図 4.3: CG (帯状回), Ins (島) の位置 (左図：水平断面図, 右図：正中断面図)

4.3.2 スペクトルの違いが与える脳活動の影響

スペクトル形状による脳活動差は、CG (帯状回)、Cd (尾状核)、MFPG (前極回中部)、Cerebellum (小脳) などでは有意な活動を示した。また、表 3.6 からわかるように、歌声のスペクトル形状である方が、話声のスペクトル形状であるよりも、脳活動の差が大きいことがわかる。

CG (帯状回) は、F0 の音響的特徴の違いにおいても、脳活動の差が起きた脳部位である。図 4.4 に示す Cd (尾状核) は、大脳皮質と視床、脳幹を結び付けている神経核の集まりである大脳基底核の一部であり、多くのドーパミン受容体があり、運動調節、認知機能、感情、動機付けなど様々な機能を担っている [2][15]。また、この部位はバイリンガルの言語を切り替える際に活動を示すことが確認されている [15]。Cerebellum (小脳) は、運動調節機能を司る部位であり [4]、心的イメージング、情動の調節、言語処理といった多くの認知課題を行っているときにも小脳が働いていることが報告されている [21]。

このように、F0 の違いと同様に、スペクトル形状の違いにおいても、大脳辺縁系や大脳基底核などの脳内部の部位に、有意な脳活動の違いが見られた。

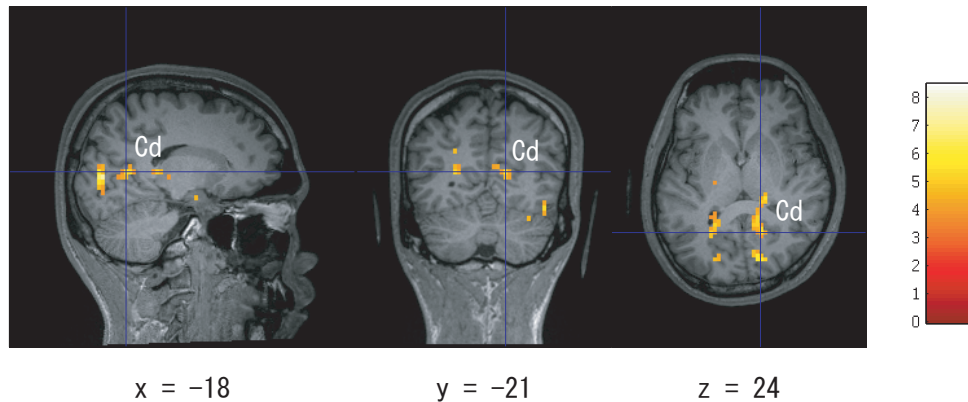


図 4.4: Cd (尾状核) の位置 (左図：矢状断面図，中央図：正中断面図，右図：水平断面図)

4.4 他の実験結果との比較

Callan らの行った脳活動測定実験においても歌声と話声で脳活動が異なっている [30]。彼らの用いた刺激音は 20 秒間の日本語の童謡であり、言語情報が含まれているが、我々の歌声と話声の脳活動の結果との共通部位として一次運動野に属する PrG (中心前回) [-46, 1, 28]、体性感覚野と頭頂連合野に属する SPL (上頭頂小葉) [-20, -60, 44]、運動調節機能を司る Cerebellum (小脳) [-40, -48, -28] が一致しており、また、眼窩回近くの OFC (眼窩前頭皮質) [-2, 46, -16] においても活動が見られている [30]。運動に関する部位に活動が見られることから、歌声を聞いた際には発声などに関わる部位も活動する可能性が高いと考えられる。また、F0 やスペクトルの違いで活動した CG [18, 13, 23]、Cd [-8, 5, 18] も活動を示している。このように、Callan らの結果と多く一致する。これらの共通部位は、言語情報に関係なく非言語情報の違いで活性化する部位であると考えられる。一方で、Callan らの実験でのみ活動が見られる部位は、聴覚連合野のヘシュル回や記憶に関する海馬などで、言語情報に関する部位だと考えられる。

4.5 聴取実験結果と脳活動の関係

聴取実験の評価結果と脳活動を比較すると、実験Ⅰでは、「自然性」の高い刺激音の方が、全体の脳活動が強い傾向があることが判明した。このことから、実際のヒトの音声に比べ、「自然性」の低い刺激音を聞いた際には、脳活動の領域が狭く弱いと考えられる。しかし、「自然性」が高くなるにつれて、ある脳部位の活性化が強くなるということは判明しなかった。また、実験Ⅰにおいて、「歌声らしさ」が高いSingとVR2は、評価は共に高くとも脳活動が全く異なっており、「自然性」の違いにより脳活動の違いが生じたと考えられる。また、「歌声らしさ」の評価結果と脳活動の関係を調べたところ、ある脳部位が徐々に強くなるということとはなかったが、全体的な脳活動は「歌声らしさ」が高い刺激音を聞いた際の方が強い傾向が見られた。

第5章 結論

5.1 本論文で明らかになったことの要約

本稿の目的は、歌声と話声に関する非言語情報を含む刺激音を聞いた際の脳活動を測定し、活動している脳部位の違いを検討し、歌声と話声それぞれに特有の脳部位を調べ、非言語情報の知覚の解明を試みることである。

脳活動測定実験の結果より、歌声と話声、音響的特徴の異なる刺激音を聞いた際の脳活動が異なることが判明した。歌声が話声よりも活性化する脳活動部位は、LOrG (眼窩回側部) や MOrG (眼窩回中央) などの情動系に関する脳部位での活動が見られた。これらの脳部位は、情動系に関係する部位であると考えられており、歌声を聞いた際には、なんらかの感情を引き起こし情動に関する脳部位が活動することで「歌声」と知覚する可能性があると考えられる。また、歌声の基本周波数変化やスペクトル形状などの非言語情報の音響的特徴の違いは、CG (帯状回) などの大脳辺縁系に属する脳内部の部位において活動差が見られることが判明した。CG は、大脳辺縁系に属し、扁桃体や前頭葉眼窩皮質などに関係し、情動にも関係する脳部位である。これらのことから、ヒトが歌声に関する基本周波数変化やスペクトル形状を知覚することによって、大脳表面の聴覚野ではなく、脳内部における大脳辺縁系や大脳基底核などの原始的な脳部位で脳活動を示すと考えられる。

「歌」として判断する音響的特徴の具体的な物理量は明らかに出来なかったが、歌声の基本周波数変化とスペクトル形状が、ヒトが脳において「歌声」として知覚する音響的特徴として重要であることが判明した。

5.2 今後の課題

本研究での有意な脳活動部位を、閾値がP 値が0.001 以下でクラスタ数が3 以上の部位とした。そこで、今後の課題として、P 値を0.001 以下に厳しくしても活動差が表れるように被験者を増やして解析を行う必要があると考える。fMRI を用いて測定した脳活動データは、空間分解能は高いものの、時間分解能が高くなく、歌音を知覚する際にどのような順番で脳部位が活動していくのかということは明らかに出来なかった。そのためには、fMRI の装置改良や他の脳活動測定装置を併用して行うことにより、より詳しく歌声知覚の解明できるのではと期待する。また、今回の実験は歌声と話声に着目した非言語情報に関する知覚の一部の解明のみであるので、今後は、感情などの非言語情報の違いに関しても研究を行い、より非言語情報の知覚の解明を進めていくのを期待する。

謝辞

本研究を遂行するにあたり、終始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 赤木正人 教授と 甲南大学 知能情報学部 北村達也 准教授に深く感謝致します。

本研究を遂行するにあたり、日頃から有益な御助言をいただき、御指導いただいた、北陸先端科学技術大学院大学 情報科学研究科 鷗木祐史 准教授、李軍鋒 助教、国際電気通信基礎技術研究所 脳活動イメージングセンタ 能田 由紀子 研究員、河内山隆紀 研究員、同志社大学 大学院工学研究科 博士後期過程 橘亮輔さん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 教授と徳田功 准教授に心より感謝致します。

日頃の研究生生活において、多くの御助言ならびに激励を賜りました、北陸先端科学技術大学院大学 情報科学研究科 博士後期過程 羽二生篤さん、木谷俊介さん、並びに音情報処理学講座の皆様、知能情報処理学講座の皆様、及び諸先輩方に厚く御礼申し上げます。

なお、本研究は総務省戦略的情報通信研究開発推進制度（課題番号 071705001）および科学研究費補助金（課題番号 20300064）の援助を受けて行われたものであります。ここに感謝の意を表します。

最後に、石川県での2年間に研究生生活をいつも暖かく見守ってくれ、心の支えとなってくれた両親、姉、並びに多くの友人たち、諸先輩方、後輩たち、そして多くの知人の方々に心より感謝致します。

参考文献

- [1] 渥美 義賢, 渡辺 哲也, 小田 侯朗, 大内 進, “障害児教育と関連した脳科学的知見について-脳の可塑性と障害の機能補填-,” 国立特殊教育総合研究所紀要, **33**, 15-25, 2006.
- [2] 甘利 俊一, 外山 慶介 編集, “脳科学大事典,” 朝倉出版, 2000.
- [3] 小川 誠二, “Functional Magnetic Resonance Imaging,” 神経研究の進歩, **38(2)**, 309-318, 1994.
- [4] 河合 良訓, 原島 広至, “脳単,” *NTS*, 2005.
- [5] 河原 英紀, “聴覚の情景解析と高品質音声分析変換合成法 STRAIGHT,” 日本音響学会講演論文集, 1-2-1, 189-192, March 1985.
- [6] 斎藤 毅, “歌声知覚・生成機構の解明に向けた歌声合成システム構築に関する研究,” JAIST 情報科学研究科博士論文, 2006.
- [7] 斎藤 毅, 辻 直也, 鷗木 祐史, 赤木 正人, “歌声らしさの知覚モデルに基づいた歌声特有の音響特徴量の分析,” 音響学会誌, **64(5)**, 267-277, 2008.
- [8] 佐藤 信, “統計的官能検査法,” 日科技連, 225-244, March 1985.
- [9] 佐藤 正之, “音楽の脳内処理 -音色認知と扁桃体-,” 高次脳機能研究, **25(2)**, 139-144, 2005.
- [10] 佐藤 正之, “音の種類の脳内認知機構: 神経心理学的研究,” 日本音響学会講演論文集, 857-858, Sept. 2007.
- [11] , 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)**, 116-128, 2005.
- [12] 田川 皓一, “失語症の画像診断,” 高次脳研究, **27**, 1-9, 2007.
- [13] 月本 洋, 菊池 吉晃, 妹尾 淳史, 安保 雅博, 渡邊 修, 米本 恭三, “脳機能画像入門 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.

- [14] 辻 直也, 赤木 正人, “歌声らしさの要因とそれに関連する音響特徴料の検討,” 日本音響学会聴覚研究会資料, H-2004-8, 2002.
- [15] 富永 靖弘, “脳のしくみ,” 新星出版社, 2007.
- [16] 中山 一郎, “日本語を歌・唄・唄う,” 日本音響学会誌, **51(11)**, 688-693(2003).
- [17] 古井 貞熙, “音声情報処理,” 森北出版株式会社, 1998.
- [18] 正木 信夫, “脳イメージング技術の聴覚・音声・言語研究への貢献,” 電子情報通信学会, 17-23, 2002.
- [19] 理化学研究所 脳科学総合研究センター 編集, “脳研究の最前線 (上) 脳の認知と進化,” 講談社, B1571, 2007.
- [20] 理化学研究所 脳科学総合研究センター 編集, “脳研究の最前線 (下) 脳の疾患と数理,” 講談社, B1571, 2007.
- [21] フロイド E ブルーム, “新 脳の探検 (上) 脳・神経系の基本地図をたどる,” 講談社, B1431, 2004.
- [22] フロイド E ブルーム, “新 脳の探検 (下) 脳から「心」と「行動」を見る,” 講談社, B1432, 2004.
- [23] Hermann Ackermann, Axel Riecker, “The contribution of the insula to motor aspects of speech production: A review and a hypothesis,” *Brain and Language*, **89**, 320-328, 2004.
- [24] Christian F. Altmann, Christoph Bledowski, Michael Wibral, Jochen Kaiser, “Processing of location and pattern changes of natural sounds in the human auditory cortex,” *NeuroImage*, **35**, 1192-1200, 2007.
- [25] Riecker Axel, Ackermann Hermann, Wildgruber Dirk, Dogil Grzegorz, Grodd Wolfgang, “Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum,” *Neuroreport*, **11(9)**, 1997-2000, 2000.
- [26] J. R. Binder, J. A. Hammeke, S. M. Rao, R. W. Cox, “Function of the left planum temporale in auditory and linguistic processing,” *Oxford University Press*, **119**, 1239-1247, 1996.
- [27] Anne J. Blood, Robert J. Zatorre, Patrick Bermudez, Alan C. Evans, “Emotion responses to pleasant and unpleasant music correlate with activity in paralimbic brain,” *Nat. Neurosci.*, **2**, 382-387, 1999.

- [28] JURGEN K MAI , “ATLAS OF THE HUMAN BRAIN 3RD EDITION , ” *ACADEMIC PRESS* , 2008 .
- [29] Steven Brown , Michael J . Martinez , Donald A . Hodges , Peter T . Fox , Lawrence M . Parsons , “The song system of the human brain , ” *Cognitive Brain Research* , **20** , 365-375 , 2004 .
- [30] Daniel E . Callan , Vassiliy Tsytsarev , Takashi Hanakawa , Akiko M . Callan , Maya Katsuhara , Hidenao Fukuyama , Robert Turner , “Song and speech:Brain regions involved with perception and covert production , ” *NeuroImage* , **31** , 1327-1342 , 2006 .
- [31] Daniel E . Callan , Jeffery A . Jones , Akiko M . Callan , Reiko Akahane-Yamada , “Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models , ” *NeuroImage* , **22** , 1182-1194 , 2004 .
- [32] Daniel E . Callan , Keiichi Tajima , Akiko M . Callan , Rieko Kubo , Shinobu Masaki , Reiko Akahane-Yamada , “Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast , ” *NeuroImage* , **19** , 113-124 , 2003 .
- [33] Manes Facundo , Springer Jane , Jorge Ricardo , Robinson Robert G , “Verbal memory impairment after left insular cortex infarction , ” *J. Neurol. Neurosurg. Psychiatry* , **67**(4) , 532-534 , 1999 .
- [34] Frank H. GUENTHER , “Neural Modeling of Speech Production , ” *Proceedings of the 4th International Nijmegen Speech Motor Conference* , 13-16 , 2001 .
- [35] Atsuko Gunji , Ryouhei Ishii , Wilkin Chau , Ryusuke Kakigi , Christo Pantev , “Rhythmic brain activities related to singing in humans , ” *NeuroImage* , **34** , 426-434 , 2007 .
- [36] N.R . Horn , M . Dolan , R . Elliott , J.F.W . Deakin , P.W.R . Woodruff , “Response inhibition and impulsivity: an fMRI study , ” *Neuropsychologia* , **41** , 1959-1966 , 2003 .
- [37] Stefan Koelsch , Thomas Fritz , Katrin Schulze , David Alsop , Gottfried Schlang , “Adults and children processing music: An fMRI study , ” *NeuroImage* , **25** , 1068-1076 , 2005 .
- [38] Stefan Koelsch , “Neural substrates of processing syntax and semantics in music , ” *Neurobiology* , **15** , 1-6 , 2005 .

- [39] B .R. .Kaada ,K .H .Pribram ,J .A .Epstein , “Respiratory and vascular responses in monkeys from temporal pole, insula, orbital surface and cingulate gyrus: a preliminary report , ” *Neurophysiol* , **12** , 347-356 , 2006 .
- [40] Philippe P .Maeder ,Reto A .Meuli ,Michela Adriani ,Anne Bellmann ,Eleonora Fornari , Jean-Philippe Thiran , Antoine Pittet , Stephanie Clarke , “Distinct Pathways Involved in Sound Recognition and Localization: A Human fMRI Study , ” *NeuroImage* , **14** , 802-816 , 2001 .
- [41] Friedemann Pulvermuller , Martina Huss , Ferath Kherif , Fermin Moscoso del Prado Martin , Olaf Hauk , Yury Shtyrov , “Motor cortex maps articulatory features of speech sounds , ” *PNAS* , **103(20)** , 7865-7870 , 2006 .
- [42] Seashore , C.E. , “The vibrato University of Iowa Studies in the Psychology of Music , ” Vol.I , (1932) .
- [43] Jeremy I. Skipper , Howard C . Nusbaum , Steven L . Small , “Listening to talking faces: motor cortical activation during speech perception , ” *NeuroImage* , **25** , 76-89 , 2005 .
- [44] Stephen M Wilson , Ayse Pinar Saygin , Martin I Sereno , Marco Iacoboni , “Listening to speech activates motor areas involved in speech production , ” *NATURE NEUROSCIENCE* , **7(7)** , 701-702 , 2004 .
- [45] Robert J.Zatorre , Pascal Belin , “Spectral and Temporal Processing in Human Auditory Cortex , ” *Oxford University Press* , **11** , 946-953 , 2001 .
- [46] Robert J.Zatorre , Ernst Meyer , Albert Gjedde , Alan C . Evans , “PET studies of Phonetic Processing of Speech: Review,Replication,and Reanalysis , ” *Oxford University Press* , **6** , 21-30 , 1996 .

本研究に関する研究業績

学会発表リスト

中村 友彦，北村 達也，赤木 正人，“fMRI を用いた歌声と話声における脳活動の差異の検討，” 日本音響学会 聴覚研究会，H-2008-108，2008．

T.Nakamura，T.Kitamura，M.Akagi，“A study on nonlinguistic features in singing and speaking voices by brain activity measurement，” Proc．NCSP09，Hawai，USA，March 2009 (to be appear)．

中村 友彦，北村 達也，赤木 正人，“脳活動測定による歌声と話声に関する非言語特徴の検証，” 日本音響学会平成 21 年度春季研究発表会 講演論文集，2-Q-12，March 2009 (発表予定)．