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

Title	歌声に含まれる基本周波数の微細変動成分の知覚に関する研究
Author(s)	北風, 裕教
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1335
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

歌声に含まれる基本周波数の微細変動成分の知覚に関する研究

指導教官　赤木正人 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

北風 裕教

2000 年 2 月 15 日

Copyright © 2000 by Hironori Kitakaze

要 旨

歌声に含まれる基本周波数の微細変動成分の分析及びその聴覚的効果を検討する。微細変動成分の分析は変調周波数と偏移幅に着目して行なう。分析の結果から「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という仮説を立てる。この仮説を基に合成音声を作成し順を追って知覚実験を行なうことで仮説を検証する。その結果、基本周波数が知覚に影響を与えていること、そして影響の大きさは変調周波数と偏移幅に依存することが明らかとなった。また、変調の変化成分も知覚に影響を及ぼすことが明らかとなった。更に微細変動成分の知覚は、規則的変動成分にはあまり関与されないことが明らかとなった。これらの結果から仮説を立証することができた。

目 次

1 序論 **1**

1.1 はじめに 1

1.2 背景 1

1.3 目的 2

1.4 本論文の構成 3

2 基本周波数の微細変動成分の分析 **4**

2.1 目的 4

2.2 分析方法 4

2.3 TEMPO2(STRAIGHT) による基本周波数の抽出 4

2.3.1 TEMPO2 の概略 6

2.3.2 合成音声を用いた TEMPO2 の精度評価 6

2.4 微細変動成分の抽出 14

2.4.1 歌声データ 14

2.4.2 抽出方法 15

2.5 微細変動成分の分析 17

2.5.1 変調周波数の分析と結果 17

2.5.2 偏移幅の分析と結果 18

2.6 考察（仮説の設定） 25

3 微細変動成分の知覚への影響 **26**

3.1 目的 26

3.2 予備実験：歌声に含まれる微細変動成分の知覚 27

3.2.1 目的 27

3.2.2 合成歌声の作成 27

3.2.3 予備実験 . . . 30
3.2.4 予備実験の考察 . . . 30
3.3 実験 1：変調の有無に対する実験 . . . 31
3.3.1 目的 . . . 31
3.3.2 実験 $1-a$ . . . 32
3.3.3 実験 $1-b$ . . . 34
3.3.4 実験 $1-c$ . . . 36
3.3.5 実験 $1-d$ . . . 38
3.3.6 実験 1 の考察 . . . 42
3.4 実験 2：変調が共に存在する場合に対する対比較実験 . . . 44
3.4.1 目的 . . . 44
3.4.2 実験 $2-a$ . . . 44
3.4.3 実験 $2-b$ . . . 47
3.4.4 実験 $2-c$ . . . 50
3.4.5 実験 2 の考察 . . . 53
3.5 変調が 2 つ存在する場合の実験 . . . 55
3.5.1 目的 . . . 55
3.5.2 音声データ . . . 55
3.5.3 実験 $3-a$ . . . 55
3.5.4 実験 $3-b$ . . . 58
3.5.5 実験 3 の考察 . . . 62

4 結論 **63**
4.1 本論文で明らかになったことの要約 . . . 63
4.2 今後の課題 . . . 64

謝辞 **65**

参考文献 **65**

学会発表リスト **68**

図 目 次

2.1 分析モデルの構成図 5
2.2 STRAIGHT(TEMPO2 の分析例) 7
2.3 $Rosenberg$波の作成 8
2.4 インパルス応答の作成 9
2.5 インパルス列の作成 9
2.6 $Klatt$ のシンセサイザーによる合成音声 10
2.7 f_0の変調周波数:左から$10Hz, 30Hz, 50Hz$、f_0の偏移幅:上から$0.5\%, 1.0\%, 1.5\%$ 分析する音声の基本周波数（破線）,TEMPO2 から抽出された基本周波数（実線） 12
2.8 $fm(t)$ と $fa(t)$ の相互相関の最大値の結果（$f_0 = 125Hz$） 12
2.9 f_0の変調周波数：左から$10, 30, 50, 70Hz$、f_0の偏移幅：上から$0.5, 1.0, 1.5\%$ 分析する音声の基本周波数（破線）, TEMPO2 から抽出された基本周波数（実線） 13
2.10 $fm(t)$ と $fa(t)$ の相互相関の最大値の結果（$f_0 = 250Hz$） 13
2.11 分析波形の例 (男性) 15
2.12 基本周波数と旋律概形成分 16
2.13 抽出した基本周波数の微細変動成分 16
2.14 微細変動成のスペクトル 18
2.15 微細変動成分のヒストグラム (サンプリング周期:$1ms$, bin 幅: 1cent) 19
2.16 基本周波数と偏移幅の関係 19

3.1 歌声の合成手法 28
3.2 採取した歌声の基本周波数 (TEMPO2 使用) 29
3.3 再合成歌声の基本周波数 (TEMPO2 使用) 29
3.4 実験 1-a：波形の例 33
3.5 変調周波数に対する分解能：基本周波数 $125Hz$、偏移幅 1% 33

3.6 実験 1-b：波形の例 35
3.7 変調周波数に対する分解能：基本周波数 $250Hz$、偏移幅 1% 35
3.8 実験 1-c：波形の例 37
3.9 変調周波数に対する分解能：基本周波数 $125Hz$、偏移幅 2% 37
3.10 実験 1-d：波形の例 39
3.11 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $3Hz$ 39
3.12 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $6Hz$ 40
3.13 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $12Hz$ 40
3.14 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $24Hz$ 41
3.15 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $48Hz$ 41
3.16 変調周波数差 $1Hz$のペア：基本周波数 $125Hz$、偏移幅 1% 45
3.17 変調周波数差 $1Hz$のペア：基本周波数 $125Hz$、偏移幅 2% 46
3.18 変調周波数差 $1Hz$のペア：基本周波数 $250Hz$、偏移幅 1% 46
3.19 変調周波数差 $1Hz$のペア：基本周波数 $250Hz$、偏移幅 2% 47
3.20 変調周波数差 $5Hz$のペア：基本周波数 $125Hz$、偏移幅 1% 48
3.21 変調周波数差 $5Hz$のペア：基本周波数 $125Hz$、偏移幅 2% 49
3.22 変調周波数差 $5Hz$のペア：基本周波数 $250Hz$、偏移幅 1% 49
3.23 変調周波数差 $5Hz$のペア：基本周波数 $250Hz$、偏移幅 2% 50
3.24 変調周波数差 $10Hz$のペア：基本周波数 $125Hz$、偏移幅 1% 51
3.25 変調周波数差 $10Hz$のペア：基本周波数 $125Hz$、偏移幅 2% 52
3.26 変調周波数差 $10Hz$のペア：基本周波数 $250Hz$、偏移幅 1% 52
3.27 変調周波数差 $10Hz$のペア：基本周波数 $250Hz$、偏移幅 2% 53
3.28 音声データの例 (上：変調周波数:$6Hz$, 下：変調周波数:$6Hz + 30Hz$) . . . 56
3.29 結果 3-a：$f_0 = 125Hz$ 57
3.30 結果 3-a：$f_0 = 250Hz$ 57
3.31 結果 3-b：$f_0 = 125Hz$, 変調周波数 f_{s2}の差が $1Hz$の音声ペア 59
3.32 結果 3-b：$f_0 = 125Hz$, 変調周波数 f_{s2}の差が $5Hz$の音声ペア 59
3.33 結果 3-b：$f_0 = 125Hz$, 変調周波数 f_{s2}の差が $10Hz$の音声ペア 60
3.34 結果 3-b：$f_0 = 250Hz$, 変調周波数 f_{s2}の差が $1Hz$の音声ペア 60
3.35 結果 3-b：$f_0 = 250Hz$, 変調周波数 f_{s2}の差が $5Hz$の音声ペア 61
3.36 結果 3-b：$f_0 = 250Hz$, 変調周波数 f_{s2}の差が $10Hz$の音声ペア 61

表 目 次

2.1 Klatt のシンセサイザーのフォルマント設定 8
2.2 FFT 分析パラメータ 17
2.3 L.M. 氏による偏移幅分析の結果 1 20
2.4 L.M. 氏による偏移幅分析の結果 2 20
2.5 S.K. 氏による偏移幅分析の結果 1 21
2.6 S.K. 氏による偏移幅分析の結果 2 21
2.7 M.S. 氏による偏移幅分析の結果 1 22
2.8 M.S. 氏による偏移幅分析の結果 2 22
2.9 T.Y. 氏による偏移幅分析の結果 1 23
2.10 T.Y. 氏による偏移幅分析の結果 2 23
2.11 T.H. 氏による偏移幅分析の結果 1 24
2.12 T.H. 氏による偏移幅分析の結果 2 24

3.1 予備実験の結果 31
3.2 実験 $1-d$ の結果 38
3.3 実験 2 の結果 53
3.4 実験 $3-a$ の結果 56
3.5 実験 $3-b$ の結果 58

第 1 章

序論

1.1　はじめに

近年、感情表現などの多様な合成音声を実現する研究が進められ、その１つとして歌声の特徴を最適なパラメータ制御より行ない、声楽家の「歌声らしい」音声合成の実現を目指す研究が行なわれている。ここで高品質な歌声を実現するために、歌声の特徴パラメータをどのように付加し、制御するかが問題となる。このためにはまず歌声に含まれる情報を明らかにすることが重要である。歌声情報の中から主要となる成分を抽出することが可能となれば、多様な合成歌声を生成できる歌声合成システムの構築や、プロ歌手を目指すための歌唱練習支援システム、歌声モーフィングや自動採譜などマンマシンインターフェイス技術にとって大変有益である。さらに、声質制御機能を備え、様々な合成歌声を生成できる音声合成システム自体の普及のために非常に重要である。

現在、合成音声の多様化の研究は、声質変換や様々な歌唱様式に対する歌声分析・合成などを中心に盛んになってきている。特に歌声特有の音程（基本周波数）の動特性の制御を精度良く行なうためには、歌声の持つ情報を把握することが非常に重要であり、そこで得られた知見は合成歌声の多様化だけでなく歌声知覚などの分野にも貢献をもたらす。

1.2　背景

品質の良い歌声合成を行うためには声帯や声道等に関するパラメータ制御が必要であり、特に基本周波数の動特性の制御が重要となる。歌声における基本周波数の動特性は、３つの変動成分に分けて考えることができる。１つは楽譜に記されている曲の旋律概形成分、そして歌声特有のビブラート成分、更に上述の成分を取り除いた後に残る微細変動成

分（不規則的変動成分）である。

この変動成分の定量的な評価と聴覚的効果そして合成への試みの研究として、矢田部・遠藤・粕谷は楽譜に記されている曲の旋律概形を構成する滑らかに変動するステップ状の成分および音程変化時のオーバーシュート・アンダーシュートに着目し特徴抽出を行ない、それらが付加された場合の合成音が聴覚的に歌声らしさの効果をもたらすことを示している [3]。また榊原・小坂は歌声特有の４から７ Hz で周期的に変化するビブラート成分に着目し、単純変調によるビブラート合成の低品質な問題を改善するために、正弦波重畳モデルを用いた定常楽音へのビブラート付加の方法を提案し音質の向上を目指している [5]。しかしこれまでの研究は、音声信号から平均化された基本周波数を抽出し分析を行なっているため、微細な時間変動を要する不規則的変動成分に対しては定量的な評価や歌声合成に及ぼす聴覚的効果を解明することは困難とされそれほど行なわれてはいない。

最近になってようやく微細な変動成分が着目されるようになり、幾つかの研究報告がされるようになった。小田切・粕谷はビブラートの揺らぎ成分に含まれる基本周波数の微細な不規則性が知覚に与える影響を含むことを分析し、これが合成音の自然性に影響を与えることを確認している [4]。この結果から、あきらかに不規則的な変動成分が知覚に何らか影響を与えていることは間違いないとされるが、その要因となる物理量は明らかにされてはおらず、早急の分析が必要とされている。また、この微細な不規則性はビブラートの揺らぎであり、旋律概形成分と規則的変動成分 (ビブラート) を取り除いた後に残る不規則的変動成分については検討が行なわれてはいないのが現状である。

1.3　目的

本研究は、歌声に含まれる基本周波数の微細変動成分を分析し、微細変動成分が知覚に与える影響について仮説を立て、知覚実験を行ない仮説の検証を行なうことを目的とする。歌声中に含まれる基本周波数の微細変動成分の特徴を明らかにすることができれば、歌手の「歌声らしい」音声合成の技術に応用できる。

そこで本研究は大衆声楽曲である日本童謡の歌唱から基本周波数の微細変動成分を抽出し、変調周波数と偏移幅に着目して定量的な評価を行なう。ここで明らかになった結果から「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という仮説を立る。この仮説を基に合成音声を作成し順を追って知覚実験を行なうことで仮説の検証を行なう。詳しい手法は次節に示す。

1.4 本論文の構成

本論文は 4 章で構成される。

1 章 本論文が対象としている研究分野の背景と問題点を指摘し、本論文の位置付けと目的を示す。

2 章 「歌声に含まれる基本周波数の微細変動成分が知覚に与える影響」を調査するために仮説を立てる。

はじめに本論文で取り扱う歌声分析の構成及び枠組を示す。特に基本周波数抽出アルゴリズム (TEMPO2) の概念を述べる。この基本周波数抽出アルゴリズムの精度の評価を合成音声を用いて行ない、本論文に用いるための使用可能性を検討する。次に基本周波数から微細変動成分 (不規則的変動成分) の抽出方法を提案する。最後に実際に採取した歌声に対して微細変動成分の分析を変調周波数と偏移幅の両面から行なう。これにより「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という仮説を立て、実際に歌声に含まれていた変調周波数と偏移幅の物理量を示す。

3 章 2 章で立てた仮説を基に合成音声を作成し順を追って知覚実験を行なうことで仮説の検証を行なう。

はじめに基本周波数の微細変動成分が知覚可能であるか合成歌声を用いて確認する。次に仮説の検証の基盤となる変調周波数と偏移幅についての検知限について様々な条件を足して確認を重ねる。最後に仮説の検証を行ない、考察する。

4 章 本論文で得られた結果を要約し、今後の展望を述べる。

第 2 章

基本周波数の微細変動成分の分析

2.1 目的

歌声に含まれる基本周波数の微細変動成分が知覚に与える影響を次章で調査するために、本章では実際に採取した歌声から基本周波数の微細変動成分を抽出し変調周波数と偏移幅に着目して分析を行なうことで、知覚に関与すると思われる幾つかの要因を考察し、それを基にして微細変動成分の知覚に関する仮説を立てることを目的とする。

2.2 分析方法

実際の歌声に微細変動成分 (不規則的変動成分) の変調周波数と偏移幅がどの程度含まれているかについて分析を行なった。図 2.1に本研究で用いる分析手法の概略を示す。分析は基本周波数抽出部、微細変動成分抽出部、偏移幅・変調周波数分析部の三段階で行なう。

2.3 TEMPO2(STRAIGHT) による基本周波数の抽出

歌声に含まれる基本周波数の微細な変動成分（不規則的変動成分）を分析するために、高精度の基本周波数抽出が可能でなければならない。したがって、本研究では河原らが提案した音声分析変換合成方法 $STRAIGHT$の基本周波数抽出アルゴリズム $TEMPO2$ を利用して基本周波数を推定する。

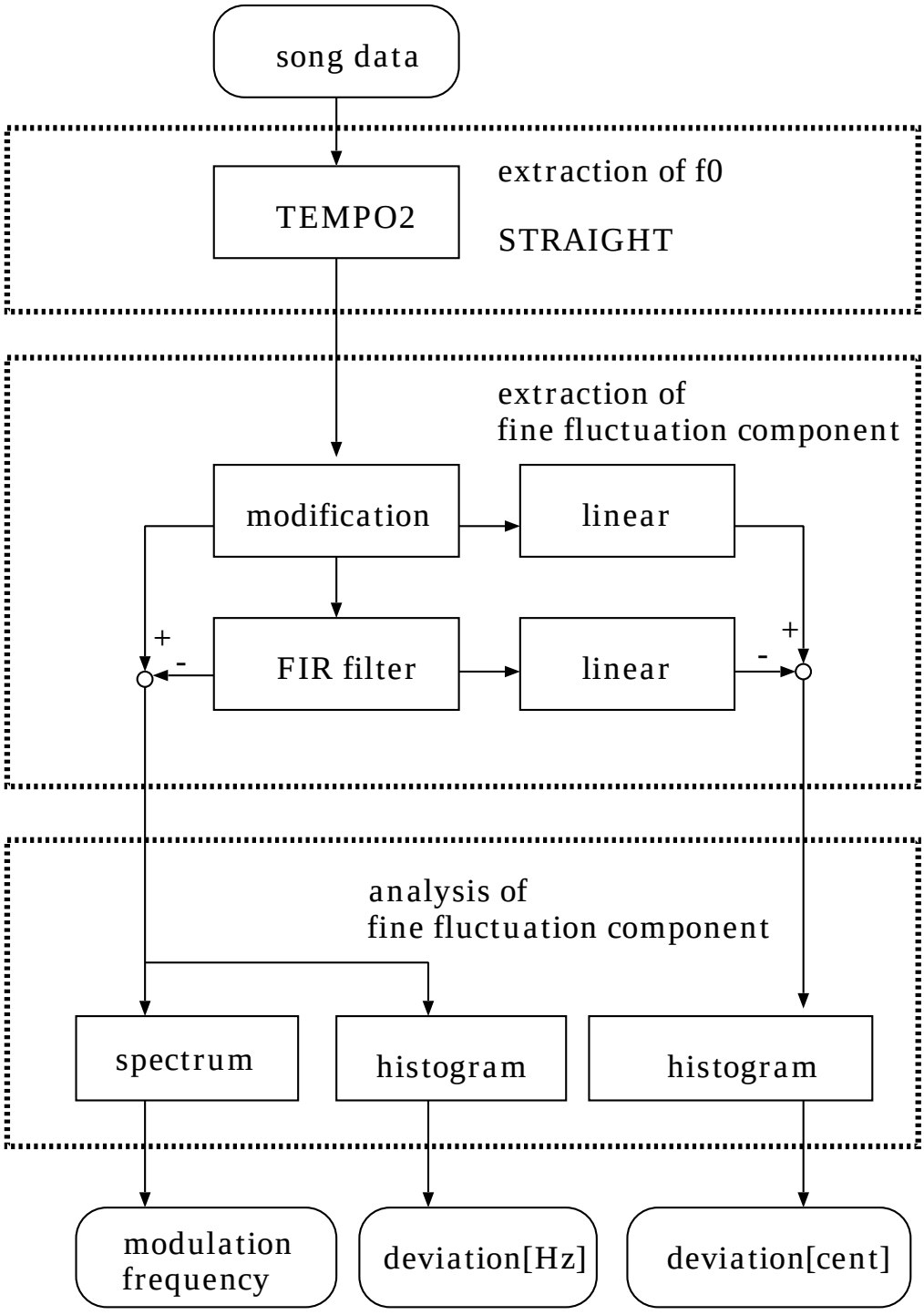


図 2.1: 分析モデルの構成図

2.3.1 TEMPO2 の概略

TEMPO2 は２つの処理段階から構成される。まず最初の段階では、log周波数軸で間隔が等しい同形の帯域通過フィルタが、フィルタの中心周波数からフィルタ出力の瞬時周波数へのマッピングにおいて浮動点を抽出するために使用される。これらの不動点は、基本周波数に対応する不動点を選択するために、推定されたC/N比によって評価される。この最初の基本周波数推定は適度に正確ではあるが、第２段階の手続きによって更に改善されることになる。

第２段階では基本周波数情報と基本周波数の微分係数を使う放射状の時間軸の伸縮が、基本周波数の適応$STFT$を行う前に導入される。この時間歪め$STFT$を基にした不動点分析が調波成分に対応する不動点を与える。そのとき、最小推定誤差をもつ基本周波数推定を与えるためのC/N情報を使うことで不動点の瞬時周波数は統合される。調波成分の推定されたC/N比はまた、音声再合成に適切な源信号の周期性を制御するための情報を与える。$TEMPO2$ の詳細は論文 [6][7][8][9] を参照して頂きたい。$TEMPO2$ の分析結果の例を 2.2に示す。

2.3.2 合成音声を用いた TEMPO2 の精度評価

基本周波数の動特性を分析するには微細な変動成分を高精度に抽出する必要がある。STRAIGHT の基本周波数抽出 (TEMPO2) がどのレベルまで精度良く微細変動を抽出できるかについて、基本周波数に変調をかけて作成した合成音声を用いて比較実験を行なった。

Klatt のシンセサイザーによる合成音声の作成

対象となる合成音声波形を考える。声門開大期をτ_1とし、声門閉小期をτ_2とすると、次式より $Rosenberg$波（三角波近似）による音源波形が与えられる。

$$f(t) = \begin{cases} a\{3(\frac{t}{\tau_1})^2 - 2(\frac{t}{\tau_1})^3\} & 0 \leq t \leq \tau_1 \\ a\{1 - (\frac{t-\tau_1}{\tau_2})^2\} & \tau_1 \leq t \leq \tau_1 + \tau_2 \end{cases} \tag{2.1}$$

なお、声門閉鎖期をτ_3とすると、周期Tは$T = \tau_1 + \tau_2 + \tau_3$ となる。図 2.3に$T = 4ms$の例を示す (サンプリング周波数 $48kHz$)。

ここで日本語母音/a/となるように共振の中心周波数及びバンド幅を設定し (表 2.1)、$Klatt$ のシンセサイザー (ホルマント型合成器) を用いて $48kHz$サンプリングで音源波形からインパルス応答を作成する (図 2.4)。後に、20 倍の $960kHz$(予備実験では $480kHz$)

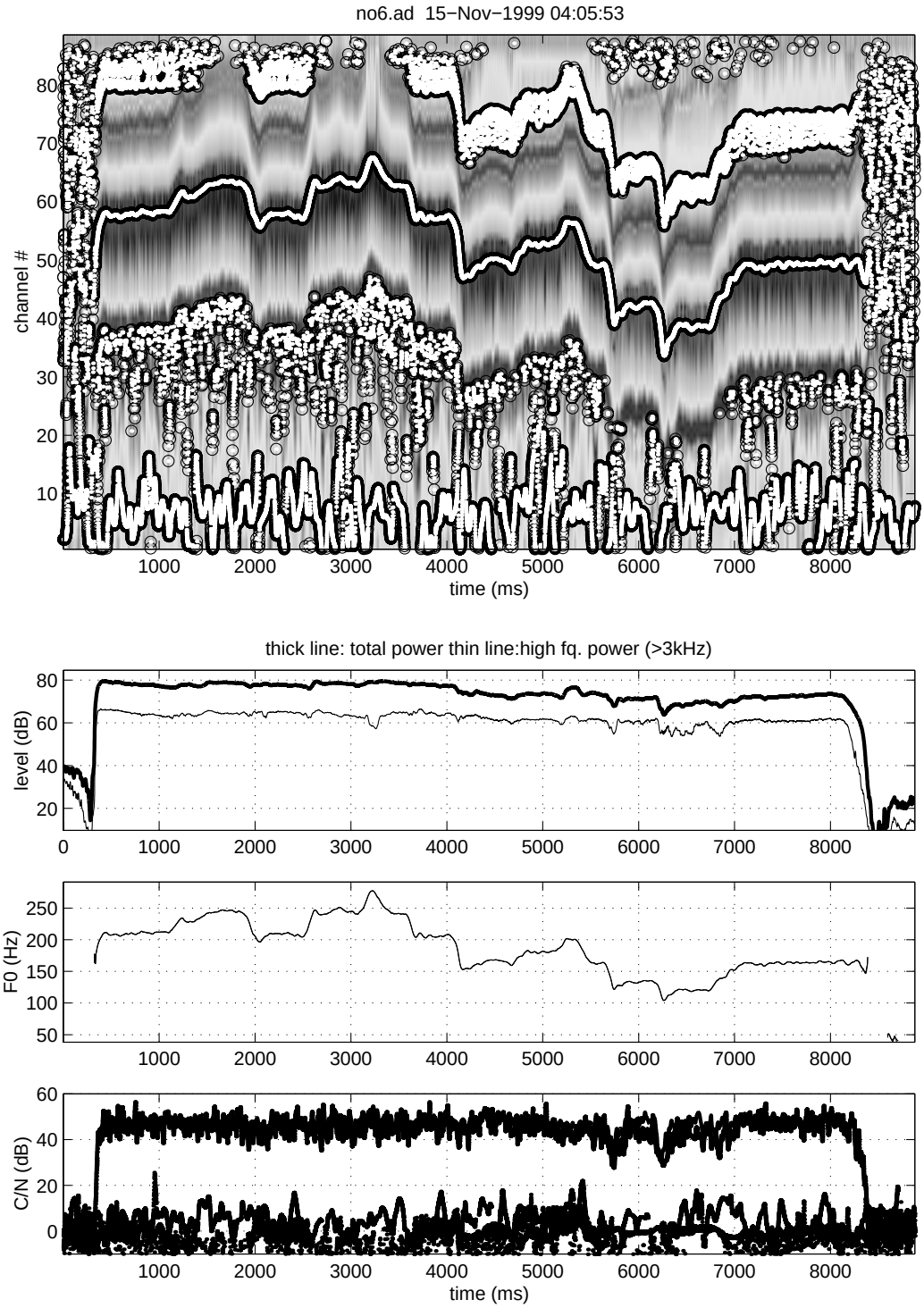


図 2.2: STRAIGHT(TEMPO2 の分析例)

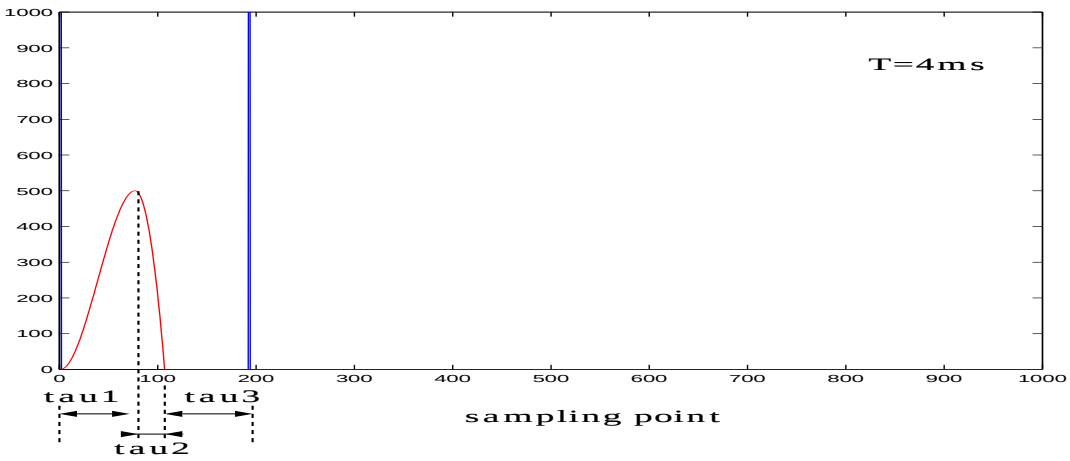


図 2.3: *Rosenberg*波の作成

にアップサンプリングを行なう。

表 2.1: Klatt のシンセサイザーのフォルマント設定

	f1	**f2**	**f3**	**f4**	**f5**	**f6**
フォルマント (Hz)	800	1200	2500	3500	4500	5500
バンド幅 (Hz)	80	120	250	350	450	550

本研究では、基本周波数の変調周波数 (f_s)・偏移幅 (α) を変化させた時の知覚分解能の検知限を調査するために、次式のように変調する基本周波数 $f_m(t)$ をモデル化する。

$$f_m(t) = (1 + \alpha \cdot sin(2\pi f_s t)) \cdot f_0 \tag{2.2}$$

これにより基本周波数 $f_m(t)$ の逆数 $1/f_m(t)$ で基本周期を決定することができる (図 2.5)。

図 2.4のインパルス応答と図 2.5のインパルス列とを畳み込む。畳み込んだ後、$48kHz$ までダウンサンプリングを行なうことで合成音声を作成することができる (図 2.6)。

ここで、アップサンプリングの処理を行なう理由は、サンプリング周波数が $48kHz$の場合、時間分解能が約 $20\mu s$ 間隔のサンプリング点しかとらない再合成波形となり、微細変動の情報が欠落してしまう原因となるからである。

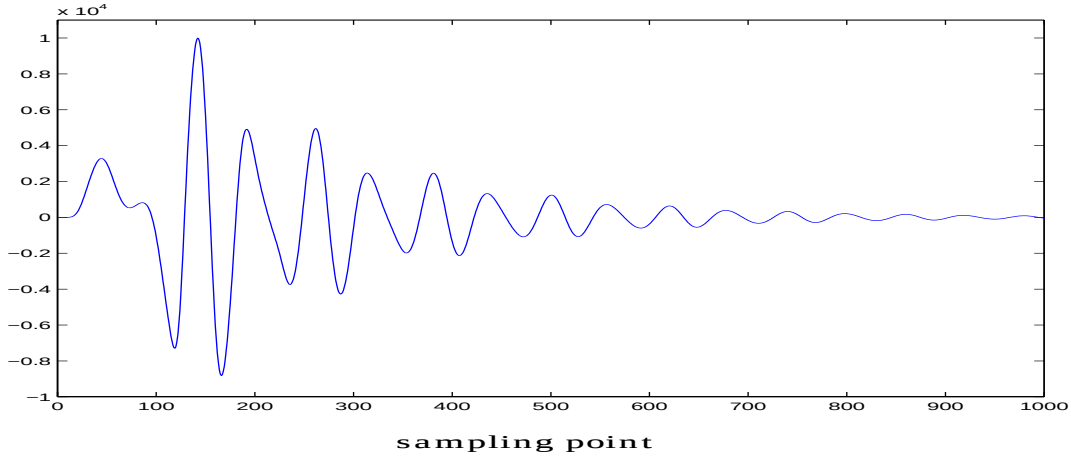
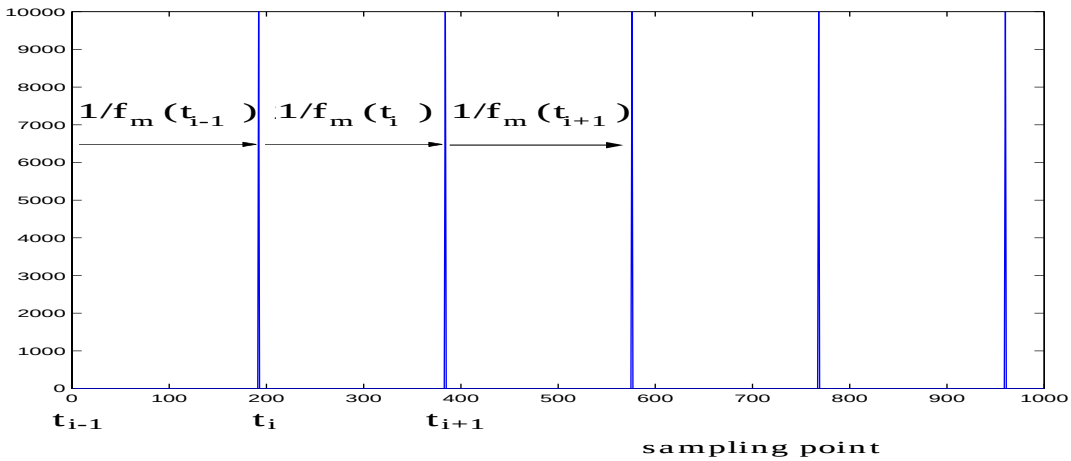


図 2.4: インパルス応答の作成



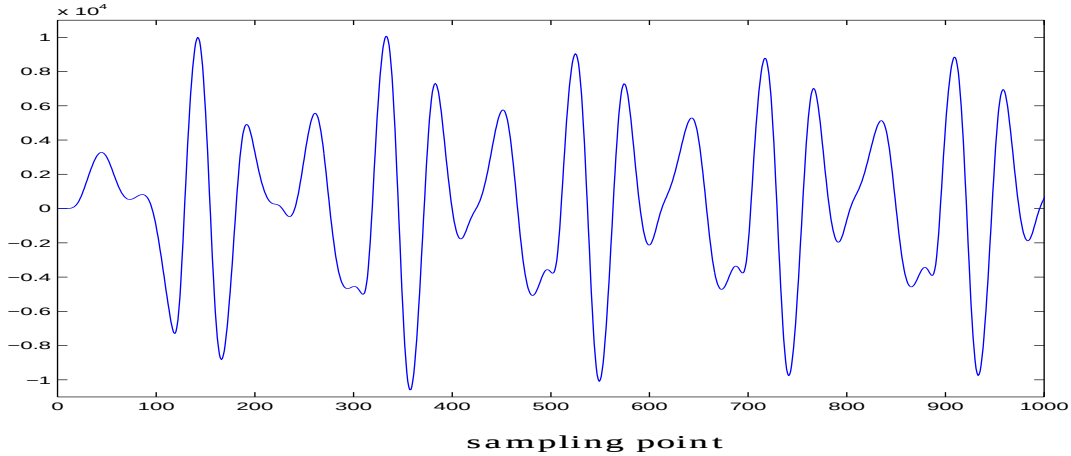


図 2.6: $Klatt$ のシンセサイザーによる合成音声

基本周波数の比較

このように作成された音声波形から基本周波数を再び STRAIGHT の $TEMPO2$ を用いて抽出する。基礎となる基本周波数 $fm(t)$ と STRAIGHT によって再抽出された基本周波数 $fa(t)$ の信号の類似性は、次のような相互相関係数の最大値により決定される。

$$\begin{aligned} \hat{R}_{fmfa} &= (\sum_{n=0}^{N-|m|-1} fm_n fa^*_{n+m}) / \sqrt{\hat{R}_{fm0fm0}} \cdot \sqrt{\hat{R}_{fa0}\hat{R}_{fa0}} \\ R_{cross} &= max[\hat{R}_{fmfa}] \end{aligned} \tag{2.3}$$

抽出精度の評価結果

f_0 125Hz

- 図 2.7は、Klatt のシンセサイザーで作成した合成音声の基本周波数（破線）と $TEMPO2$ の分析により再抽出した基本周波数（実線）との波形比較を示す。基準となる基本周波数は $125Hz$である。図は上から下へ基本周波数の偏移幅が $0.5\%(\pm 0.625Hz)$、$1.0\%(\pm 1.250Hz)$、$1.5\%(\pm 1.875Hz)$ を与えたものに相当し、左から右へ基本周波数の偏移幅が $10Hz$、$30Hz$、$50Hz$を与えたものに相当する。この結果から、変調周波数が上昇するに従い偏移幅はそれに伴い減少すること、そして位相遅れが生じることが言える。
- 図2.8は、Klatt のシンセサイザーで作成した合成音声の基本周波数と $TEMPO2$ の分析により再抽出した基本周波数とを、相互相関の最大値で比較 [式 (2.16)] し

た結果である。基準となる基本周波数は $125Hz$ である。基本周波数の偏移幅は 0.0% から 2.0% まで 0.1% 刻み、基本周波数の変調周波数は $0.0Hz$ から $50.0Hz$ まで $1.0Hz$ 刻みである。この結果、変調周波数が上昇すると相関値が低下することが言える。図から基本周波数 $125Hz$ に対しては、変調周波数 $25Hz$ までは安定して分析が可能であると考えられる。また、偏移幅による相関値の影響はなかった。

f_0 250Hz

- 図 2.9は、Klatt のシンセサイザーで作成した合成音声の基本周波数（破線）と $TEMPO2$ の分析により再抽出した基本周波数（実線）との波形比較を示す。基準となる基本周波数は $250Hz$ である。図は上から下へ基本周波数の偏移幅が $0.5\%(\pm 1.25Hz)$、$1.0\%(\pm 2.50Hz)$、$1.5\%(\pm 3.75Hz)$ を与えたものに相当し、左から右へ基本周波数の偏移幅が $10Hz$、$30Hz$、$50Hz$、$70Hz$ を与えたものに相当する。この結果、基本周波数が $125Hz$ の場合と同様変調周波数が上昇するに従い偏移幅はそれに伴い減少すること、そして位相遅れが生じることが言える。ただし、基本周波数が $125Hz$ よりも偏移幅の減少量が小さいことが伺える。
- 図2.10は、Klatt のシンセサイザーで作成した合成音声の基本周波数と $TEMPO2$ の分析により再抽出した基本周波数とを、相互相関の最大値で比較 [式 (2.16)] した結果である。基準となる基本周波数は $250Hz$ である。基本周波数の偏移幅は 0.0% から 2.0% まで 0.1% 刻み、基本周波数の変調周波数は $0.0Hz$ から $80.0Hz$ まで $1.0Hz$ 刻みである。この結果、変調周波数が上昇すると相関値が低下することが言える。図から基本周波数 $250Hz$ に対しては、変調周波数 $50Hz$ までは安定して分析が可能であると考えられる。また、偏移幅による相関値の影響はなかった。

これらの結果から基本周波数 f_0 に対して変調周波数が、$f_0/5Hz$ 以上では急激に相関が悪くなるといえる。これは、$TEMPO2$ の処理で用いられる時間窓が基本周期の 2 倍の窓であり、2 周期から 1 つの基本周波数を決定するために 2 周期の平均となることと、サンプリング定理から、サンプリング周波数 $f_0(125Hz or 250Hz)$ に対して $f_0/2$ までしか使用できないことからと予想される。これらのことから、基本周波数の 1/5 以内の変調周波数であれば、分析が可能であるといえる。

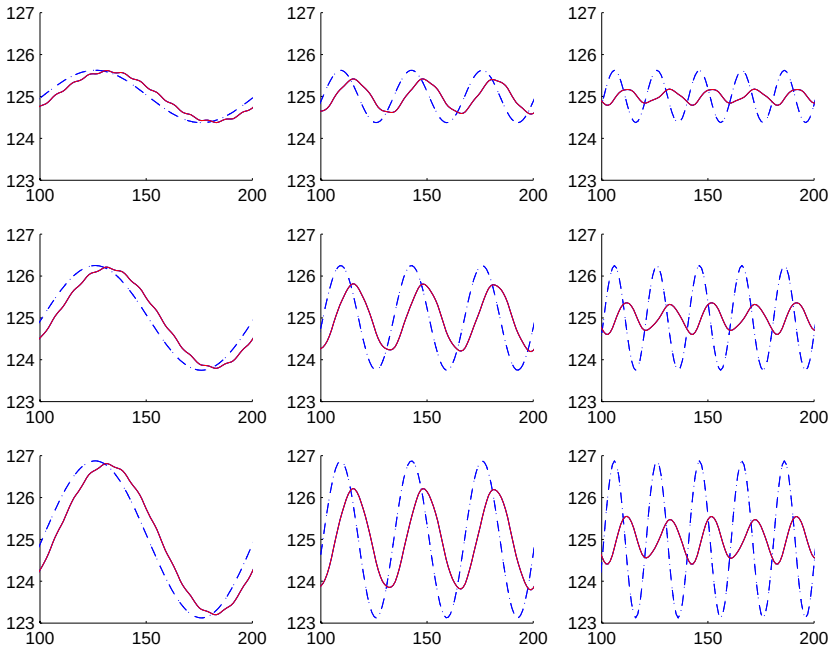
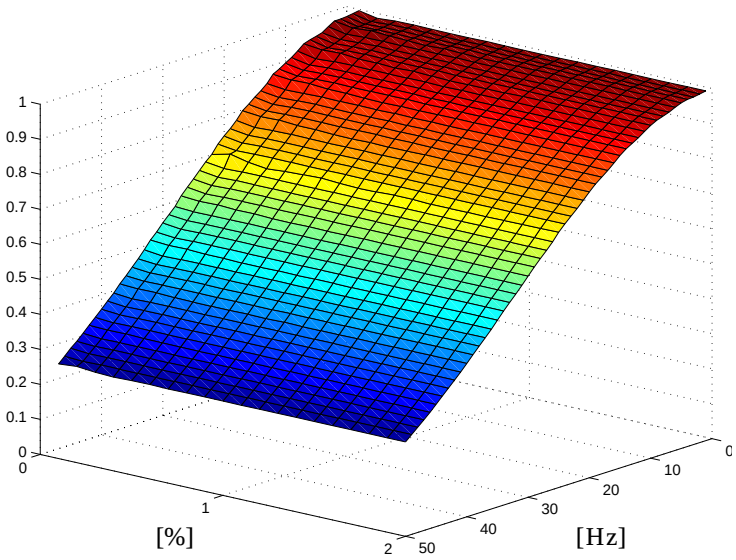


図 2.7: f_0の変調周波数：左から $10Hz, 30Hz, 50Hz$、f_0の偏移幅：上から $0.5\%, 1.0\%, 1.5\%$
分析する音声の基本周波数（破線）,TEMPO2 から抽出された基本周波数（実線）



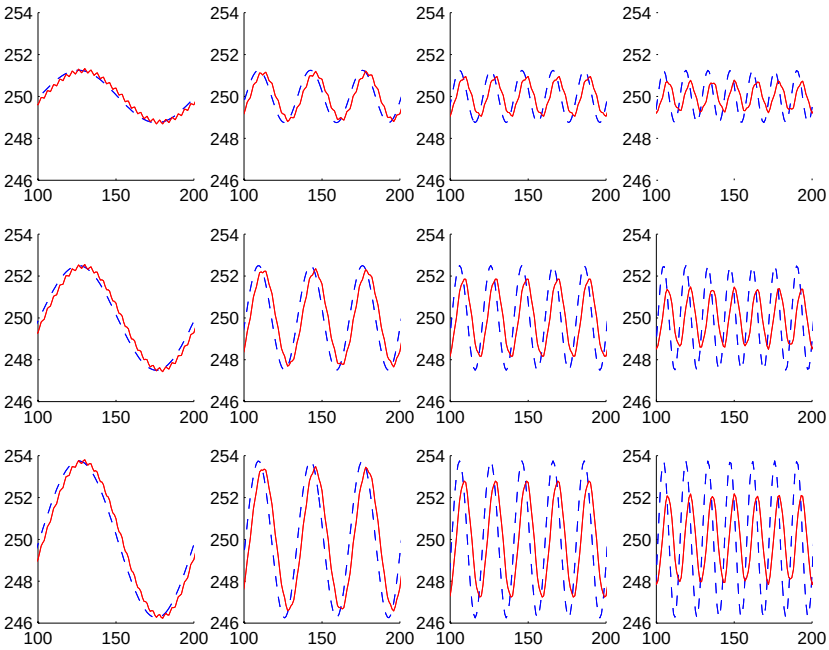
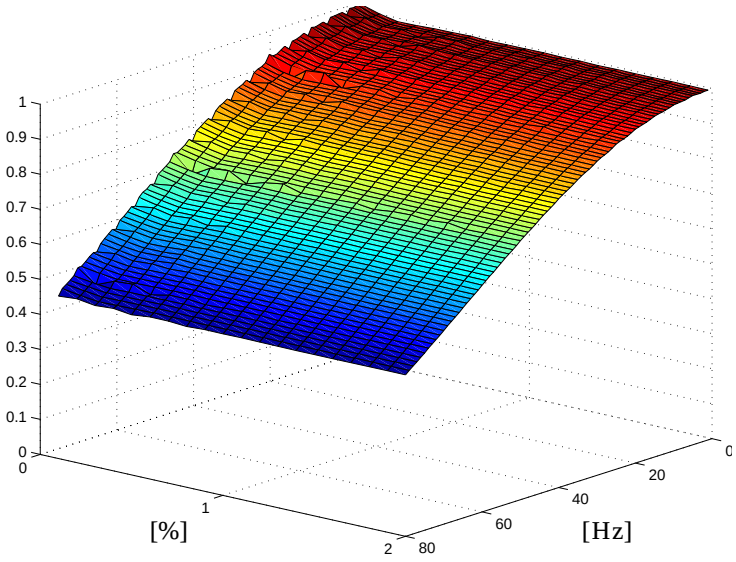


図 2.9: f_0の変調周波数：左から $10, 30, 50, 70Hz$、f_0の偏移幅：上から $0.5, 1.0, 1.5\%$
分析する音声の基本周波数（破線）, TEMPO2 から抽出された基本周波数（実線）



2.4 微細変動成分の抽出

2.4.1 歌声データ

歌声は、ボーカルスクールまたは声優養成所に通う学生 (男性 2 名、女性 3 名) から実際に採取したものを用いる。実験の簡略化のために対象となる歌は大衆声楽曲でもある日本童謡「七つの子」を日本語母音/a/で歌唱したものを用いる。被験者には歌唱法に関して特別な指示を与えることなく、また譜面に記される音階で歌唱するという制限も与えてはいない。その理由として、被験者の発声可能な音域を考慮に入れたためである。録音は防音室内部で行なわれ、サンプリング周波数 $48kHz$、量子化数 $16bit$ で DAT にデジタル録音する。このように録音された歌声を 4 小節ごとに切りだし、微細変動成分の抽出を行なった。

2.4.2 抽出方法

$TEMPO2$ で抽出した基本周波数に 201 段の FIR型ローパスフィルタ (カットオフ周波数 $5Hz$) を用いることで曲の旋律概形成分を推定し、抽出した基本周波数との差を求めることで、微細変動成分の抽出を行なう。

ここで、変動成分は次式により $f1$ と $f2$ の 2 つの音の音高の違いを表現する。

$$音高の差 = 1200log_2(f1/f2)(セント)$$

これにより半音は 100 セント、全音は 200 セント、オクターブは 1200 セントで表すことができるので、基本周波数の高さに関係なく音階間の距離を等しくでき、音楽的にも分析が可能となる。

図 2.11に歌声波形の例を示す。また図 2.12が $TEMPO2$ によって抽出した基本周波数とローパスフィルタ通過後の基本周波数である。図 2.13にこの両者の周波数の差から抽出した基本周波数の微細変動成分を示す。

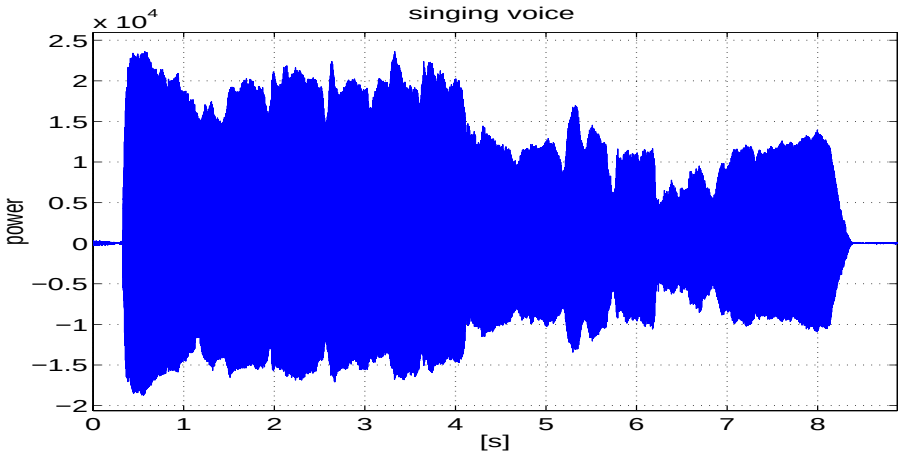


図 2.11: 分析波形の例 (男性)

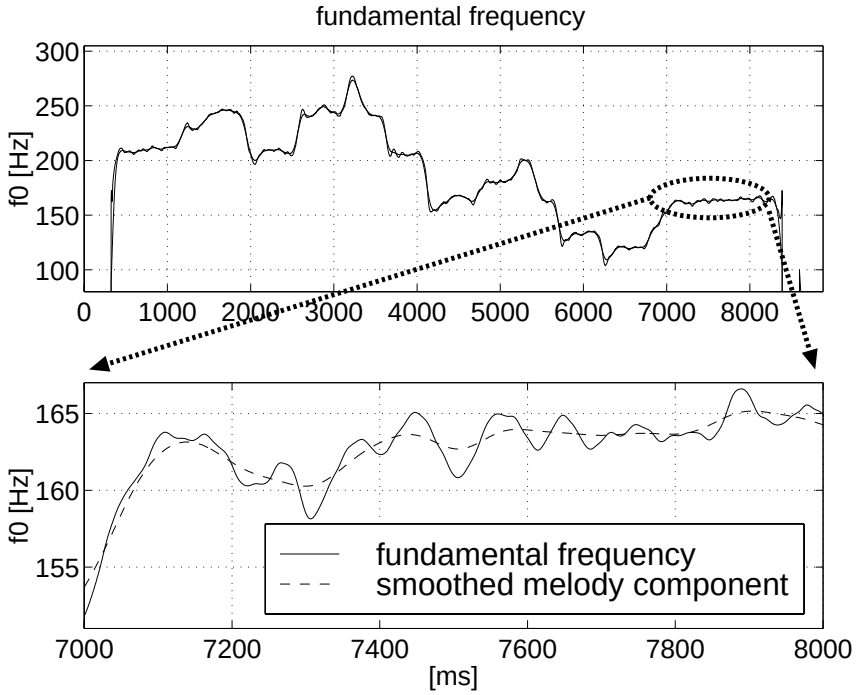


図 2.12: 基本周波数と旋律概形成分

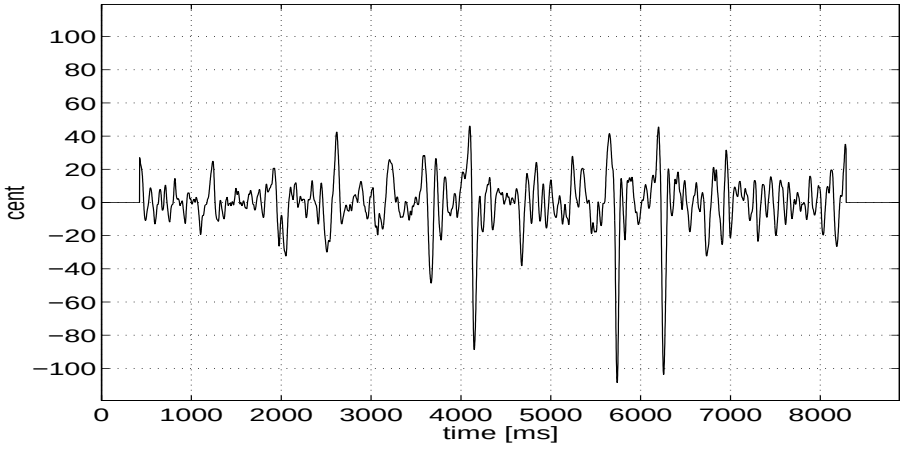


図 2.13: 抽出した基本周波数の微細変動成分

2.5 微細変動成分の分析

分析は、抽出した基本周波数の微細変動成分に対して、変調周波数と偏移幅に着目して行なう。変調周波数をランニングスペクトルから分析し、また偏移幅をヒストグラムを用いて分析する。また、基本周波数と偏移幅との関係についても同時に分析を行なう。

2.5.1 変調周波数の分析と結果

2.4節で抽出した基本周波数の微細変動成分に対するランニングスペクトルを調べた。FFTのパラメータは表2.2ようなものである。

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)} \qquad (FFT) \tag{2.4}$$

$$x(j) = \frac{1}{N}\sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)} \qquad (IFFT) \tag{2.5}$$

$$\omega_N = e^{-2\pi(j/N)} \tag{2.6}$$

表 2.2: FFT 分析パラメータ

分析パラメータ	設定（値）
sampling frequency	1.0kHz
window length	512msec
window type	hamming
fft length	2048(2^{11})
frame period	1msec

図2.13の微細変動成分のランニングスペクトルが図2.14(上) である。またスペクトルの全時間平均を行なった結果が図2.14(下) である。このような手法を用いて、全データに対して分析を行なった。

これらの分析から次のような結果が得られた。

- 音程変化時において変調周波数は$20Hz$程度の周波数成分を含む。
- 音程安定時において変調周波数は$10-15Hz$程度の周波数成分を含む。
- 音程安定時において変調周波数は急激な変化を起こすことはない。

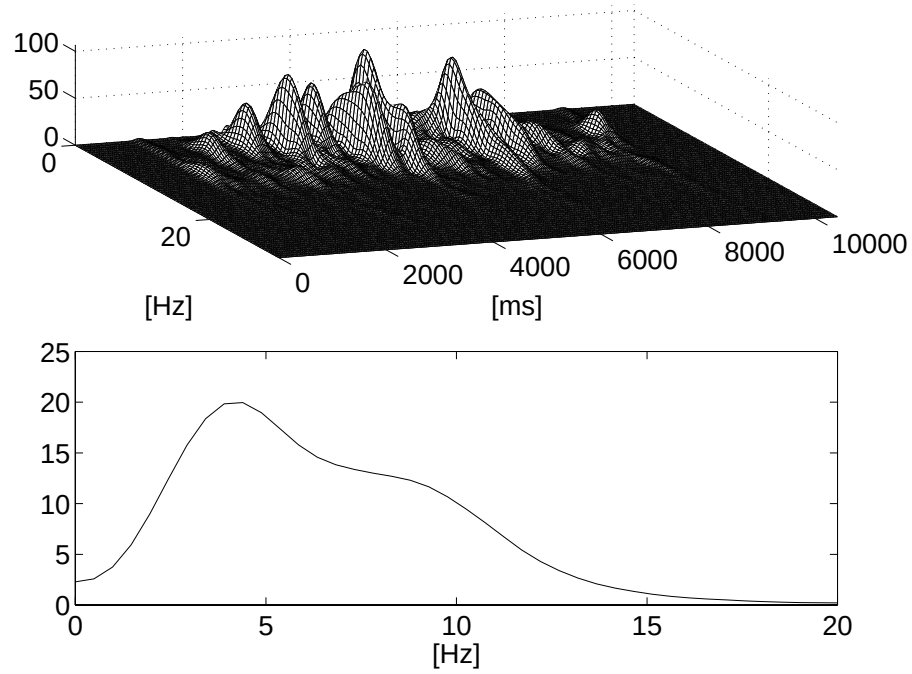


図 2.14: 微細変動成のスペクトル

2.5.2 偏移幅の分析と結果

2.4節で抽出した基本周波数の微細変動成分の波形からヒストグラムを計算することで偏移幅を分析する。図 2.15は図 2.13のヒストグラムの結果である。ただし無音区間は無視するものとする。男性２名（L.M. , S.K.）、女性３名（M.S. , T.Y. , T.H.）において分析を行ない、この結果の平均と標準偏差を求めたものが、表 2.3から表2.12である。また、基本周波数と偏移幅との関係についても分析を行なった (図 2.16)。なお、図 2.16(下) はその結果の絶対値である。

これらの結果から次のような考察が行なえる。被験者に対して偏移幅の大きさに差はあるが、偏移幅の標準偏差の平均は約$\pm 20cent$ であった。半音階が $100cent$ であることから、音階間の 1/5 は揺れていることになる。

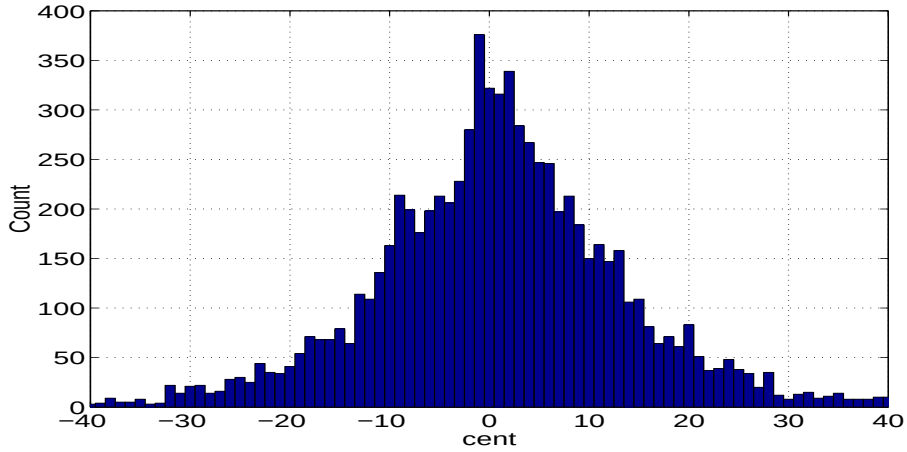


図 2.15: 微細変動成分のヒストグラム (サンプリング周期:1ms, bin 幅: 1cent)

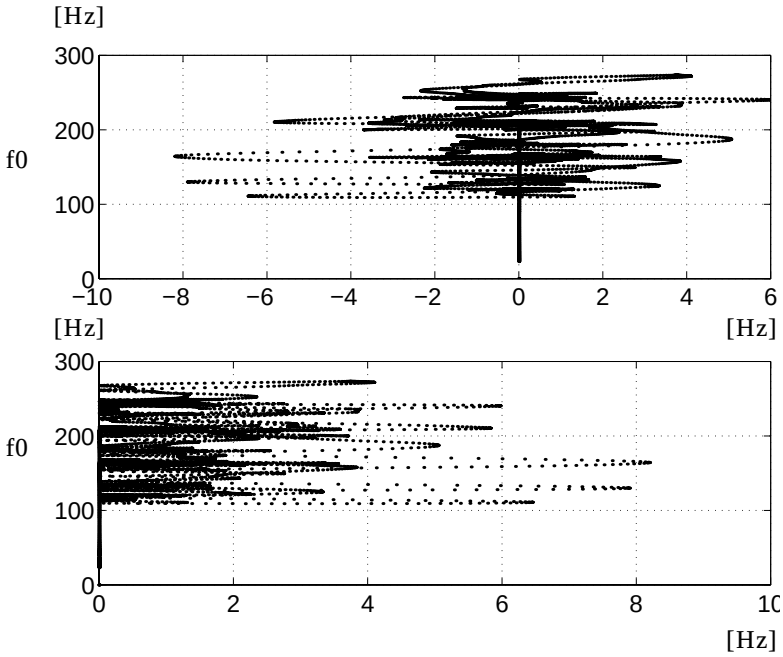


表 2.3: L.M. 氏による偏移幅分析の結果 1

L.M. （１回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	-0.0027	0.2654	-0.0351	2.8084
第 1 小節目〜第 4 小節目	0.0087	1.8859	-0.3616	17.8933
第 5 小節目〜第 8 小節目	0.0371	4.8991	-0.5559	37.1102
第 9 小節目〜第 12 小節目	0.0316	3.7302	-0.3454	33.4370
第 13 小節目〜第 16 小節目	0.0391	3.4839	-0.4748	40.6812
第 1 小節目〜第 4 小節目 (2 回目)	-0.0019	1.8798	-0.5682	17.4642
第 5 小節目〜第 8 小節目 (2 回目)	0.0471	2.5260	0.0328	18.7534

表 2.4: L.M. 氏による偏移幅分析の結果 2

L.M. （２回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0284	0.4853	0.7304	13.9965
第 1 小節目〜第 4 小節目	0.0023	1.2588	-0.1828	12.8892
第 5 小節目〜第 8 小節目	0.0072	3.6291	-0.7253	29.3570
第 9 小節目〜第 12 小節目	0.0110	2.1397	-0.0730	16.7575
第 13 小節目〜第 16 小節目	-0.0171	0.5276	-0.1673	47.7811
第 1 小節目〜第 4 小節目 (2 回目)	-0.0224	1.4644	-0.6418	14.0058
第 5 小節目〜第 8 小節目 (2 回目)	0.0011	2.5043	-0.4342	19.2371

表 2.5: S.K. 氏による偏移幅分析の結果 1

S.K. （１回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0098	0.8605	0.0590	8.5757
第 1 小節目〜第 4 小節目	0.0092	1.7062	-0.4789	19.6016
第 5 小節目〜第 8 小節目	0.0223	1.5307	-0.1605	13.9900
第 9 小節目〜第 12 小節目	0.0238	1.4047	-0.0945	15.4855
第 13 小節目〜第 16 小節目	0.0254	2.4146	-0.3976	26.7743
第 1 小節目〜第 4 小節目 (2 回目)	0.0353	1.4337	-0.2406	17.1392
第 5 小節目〜第 8 小節目 (2 回目)	0.0137	1.6170	-0.3439	16.6996

表 2.6: S.K. 氏による偏移幅分析の結果 2

S.K. （２回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0361	0.8986	0.3066	8.1771
第 1 小節目〜第 4 小節目	0.0056	1.2641	-0.3811	14.1235
第 5 小節目〜第 8 小節目	0.0111	1.4478	-0.3001	13.6884
第 9 小節目〜第 12 小節目	-0.0980	1.3708	-0.4800	14.6094
第 13 小節目〜第 16 小節目	0.1227	2.0745	0.8766	24.4535
第 1 小節目〜第 4 小節目 (2 回目)	0.0360	1.3051	-0.1807	14.3545
第 5 小節目〜第 8 小節目 (2 回目)	0.0259	1.7142	-0.3335	16.9193

表 2.7: M.S. 氏による偏移幅分析の結果 1

M.S. （１回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0878	1.8427	3.1063	61.8290
第 1 小節目〜第 4 小節目	-0.0151	2.4822	-1.1734	20.7736
第 5 小節目〜第 8 小節目	0.1342	3.1097	2.3167	42.0742
第 9 小節目〜第 12 小節目	-0.0080	2.1904	-0.2682	31.6317
第 13 小節目〜第 16 小節目	0.0056	2.0285	-0.2073	11.6257
第 1 小節目〜第 4 小節目 (2 回目)	0.0361	2.9298	-0.3907	19.9414
第 5 小節目〜第 8 小節目 (2 回目)	0.0776	3.1633	1.1842	60.7861

表 2.8: M.S. 氏による偏移幅分析の結果 2

M.S. （２回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	-0.0088	0.9212	0.1650	7.0932
第 1 小節目〜第 4 小節目	-0.0135	2.5060	-0.6110	18.3944
第 5 小節目〜第 8 小節目	0.1297	2.8741	2.2466	52.1874
第 9 小節目〜第 12 小節目	-0.0122	1.8678	-0.2293	9.0179
第 13 小節目〜第 16 小節目	0.0022	2.6635	-0.2824	29.6393
第 1 小節目〜第 4 小節目 (2 回目)	-0.0635	2.9207	-0.7310	17.5294
第 5 小節目〜第 8 小節目 (2 回目)	-0.0236	3.5246	1.1183	41.6493

表 2.9: T.Y. 氏による偏移幅分析の結果 1

T.Y.（１回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	-0.0190	0.8968	-0.1291	5.3435
第 1 小節目〜第 4 小節目	0.0230	3.3789	-0.7509	23.3840
第 5 小節目〜第 8 小節目	0.0414	3.1323	-0.3021	16.9174
第 9 小節目〜第 12 小節目	0.0506	2.7148	0.0055	14.0496
第 13 小節目〜第 16 小節目	0.0119	2.6621	-0.5222	17.4271
第 1 小節目〜第 4 小節目 (2 回目)	-0.0128	2.5902	-0.6284	16.6902
第 5 小節目〜第 8 小節目 (2 回目)	-0.0032	3.2312	-0.6037	18.4490

表 2.10: T.Y. 氏による偏移幅分析の結果 2

T.Y.（２回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0252	0.8811	0.1250	4.8686
第 1 小節目〜第 4 小節目	-0.0125	3.3739	-0.6951	20.3124
第 5 小節目〜第 8 小節目	-0.0215	3.8972	-0.5494	16.9104
第 9 小節目〜第 12 小節目	-0.0363	3.8064	0.6823	18.7653
第 13 小節目〜第 16 小節目	0.1144	3.3068	0.1594	18.5023
第 1 小節目〜第 4 小節目 (2 回目)	-0.0338	3.3744	-0.9168	19.8309
第 5 小節目〜第 8 小節目 (2 回目)	-0.0140	3.9621	-0.5326	17.1953

表 2.11: T.H. 氏による偏移幅分析の結果 1

T.H.（１回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0555	0.7999	0.2814	4.1094
第 1 小節目〜第 4 小節目	-0.0083	2.1360	-0.4215	12.4195
第 5 小節目〜第 8 小節目	-0.0017	3.5019	-0.5222	17.2589
第 9 小節目〜第 12 小節目	-0.0215	1.9856	-0.4025	11.1306
第 13 小節目〜第 16 小節目	0.0055	1.8377	-0.2068	10.4807
第 1 小節目〜第 4 小節目 (2 回目)	-0.0097	2.2356	-0.4409	13.8570
第 5 小節目〜第 8 小節目 (2 回目)	-0.0133	3.1108	-0.5388	16.1455

表 2.12: T.H. 氏による偏移幅分析の結果 2

T.H.（２回目）	周波数 (Hz)		cent 値 (cent)	
	平均	標準偏差	平均	標準偏差
日本語母音/a/自然発声	0.0191	0.5114	0.0988	2.7548
第 1 小節目〜第 4 小節目	-0.0025	2.1401	-0.3337	11.7757
第 5 小節目〜第 8 小節目	-0.0216	2.9413	-0.4868	13.6381
第 9 小節目〜第 12 小節目	0.0027	2.2427	-0.3236	12.1113
第 13 小節目〜第 16 小節目	0.0005	1.8653	-0.2637	10.3112
第 1 小節目〜第 4 小節目 (2 回目)	0.0027	2.2092	-0.3068	12.7579
第 5 小節目〜第 8 小節目 (2 回目)	0.0290	3.2850	-0.2960	15.1355

2.6　考察（仮説の設定）

歌声に含まれる基本周波数の微細変動成分から変調周波数と偏移幅を測定することで次のようなことが明らかとなった。

変調周波数 音程安定時には変調周波数が $15Hz$、音程変化時には変調周波数が $20Hz$程度の不規則的な微細変動成分を確認することができた。

偏移幅 偏移幅の標準偏差の平均は$\pm 20cent$ である。半音階が $100cent$ であることから、音階間の 1/5 は揺れている。ただし音程変化時では最大$\pm 60cent$ の変化が存在する。

このことから、音程安定時には変調周波数が低く偏移幅も小さいが音程変化により変調周波数が高く (低く) なるとそれにともない偏移幅も大きくなるということができる。したがって変調周波数と偏移幅には密接な関係があるといえる。

これらの結果と、今までの背景から次のような仮説を立てる。「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」次章ではこの仮説に対する検証を行なう。

第 3 章

微細変動成分の知覚への影響

3.1 目的

2 章では、知覚に与える影響を調べるために次のような仮説を立てた。

「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」

本章では 2 章で立てた仮説をもとに合成音声を作成し順を追って知覚実験を行なうことで仮説の検証を行なう。

以上のことをふまえて次の実験を行なう。

予備実験 人間が微細変動を知覚可能であるか確認する。これは上述で立てた仮説が的外れでないという事を確認するためである。実験は実際に採取した歌声から基本周波数を抽出し、微細変動成分を含む合成歌声と LPF により微細変動成分を取り除いた合成歌声の違いを正しく知覚できるか確認するものである。ここで確認ができれば、微細変動成分が歌声の知覚に何らかの影響を与えていることがいえる。

実験 1 基本周波数の変調が存在する場合と存在しない場合の合成音声の検知限を確認する。これは仮説の検証を行なうにあたり、幾つかの制約条件を与えた合成音声を用いて研究の基盤となる検知限を調べておく必要があるためである。また、変調周波数と偏移幅の依存関係も確認する。ここで確認が行なえれば、微細変動成分が歌声の知覚に与える影響を定量的に示すことができる。

実験 2 基本周波数の変調が共に存在する合成音声の検知限を確認する。実験 1 では変調の有無に対する実験であったが、実際には変調が存在しないものからの距離だけを知覚しているわけではなく、変調の変化成分も知覚に影響している可能性がある。

したがって実験 2 は変調周波数が時々刻々と変化する実際の歌声を考慮にいれ、対の音声において共に変調が存在する場合の検知限を調べるものである。この結果からは、変調周波数の幅がどのように知覚に影響を与えているかが考察でき、仮説の内容の「影響の大きさ」について検討ができる。

実験 3 低い変調周波数に高い変調周波数が存在する場合の合成音声の検知限を確認する。これは旋律概形成分や規則的変動成分 (ビブラート) などの要因によって微細変動成分の検知限が変化するかを考察するためである。この実験から、より歌声の条件に近い状態で、微細な変動成分が仮説どおりに知覚できるのかを検討することができる。

これらの段階を踏めば仮説を検証するための変調周波数と偏移幅の知覚への影響を考察できる。

3.2 予備実験：歌声に含まれる微細変動成分の知覚

歌声に含まれる基本周波数の微細変動成分が知覚に与える影響について合成音声を作成して聴取実験を行なった。

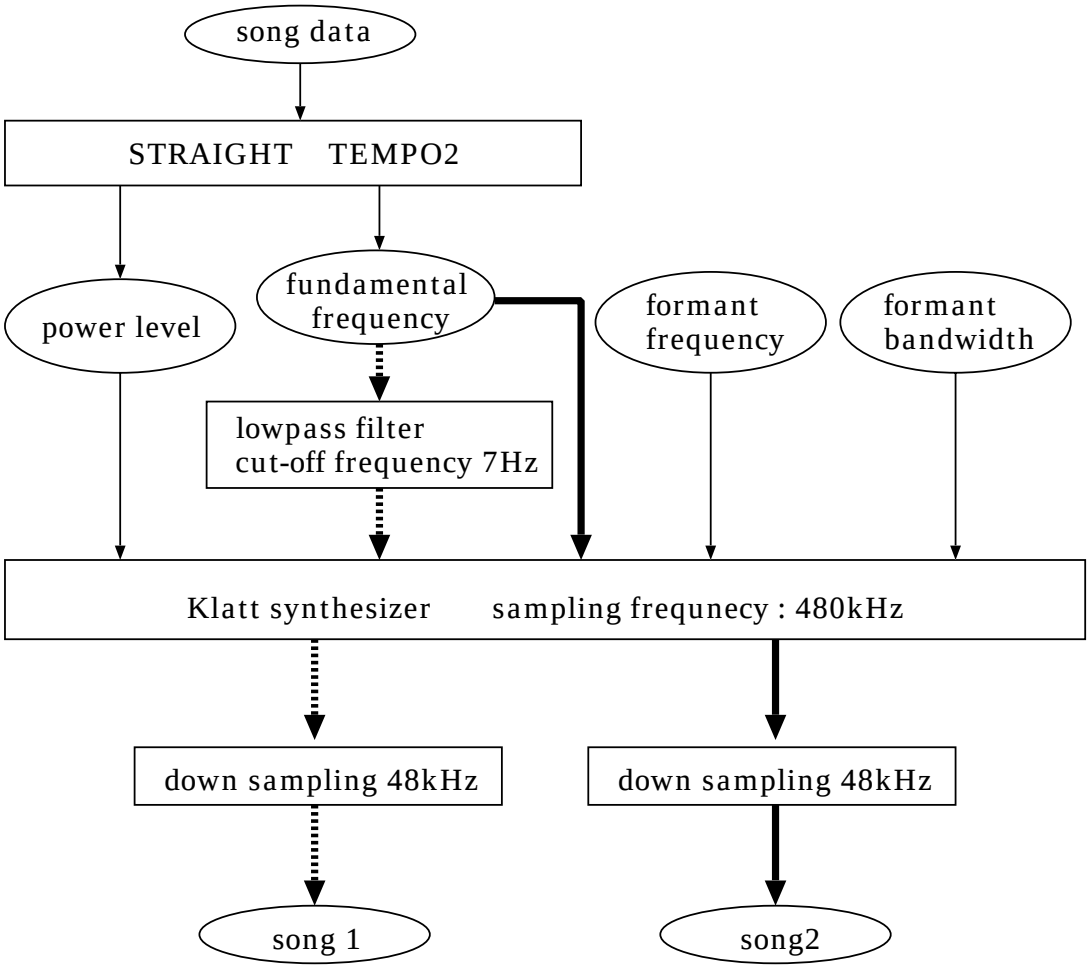
3.2.1 目的

2 章で立てた仮説の検証を行なう前に、実際に採取した歌声において、基本周波数の微細変動成分を知覚可能であるか調べる必要がある。これは「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という仮説が的外れでないという可能性を確認するためである。したがって本節では、TEMPO2(STRAIGHT) により抽出された基本周波数とカットオフ周波数 $7Hz$のローパスフィルタを通し微細変動成分を削除した基本周波数との違いを正しく知覚することが可能であるか合成音声を作成し対比較実験を行なう。ここでカットオフ周波数を $7Hz$に設定した理由は、曲の旋律概形成分 (5Hz 程度) 及びビブラート成分 ($4Hz$から $7Hz$で規則的に変動) を削除するためである。

3.2.2 合成歌声の作成

図 3.1に歌声の合成手法 (ブロックダイアグラム) を示す。このブロックダイアグラムは 2 段階から構成される。はじめの段階では、実際に採取した歌声 (song data) から$TEMPO2$

を用いて基本周波数 f_m とパワーを抽出する。次の段階で 2 章の表 2.1と同じ日本語母音 $/a/$ となるようにフォルマントの設定を行ない $Klatt$ のシンセサイザーを用いてインパルス応答を作成し、抽出した基本周波数 f_m の逆数 $1/f_m$ ごとにインパルスを持つインパルス列と畳み込みを行なう。ここでサンプリング周波数 480kHz の歌声を作成した後、48kHz までダウンサンプリングを行なうことで合成歌声 $(song1)$ を作成する。また、$TEMPO2$ から得られた基本周波数からカットオフ周波数 $7Hz$ のローパスフィルタを通して微細変動成分を削除した基本周波数を作成し、上述同様の処理により合成歌声 $(song2)$ を作成する。



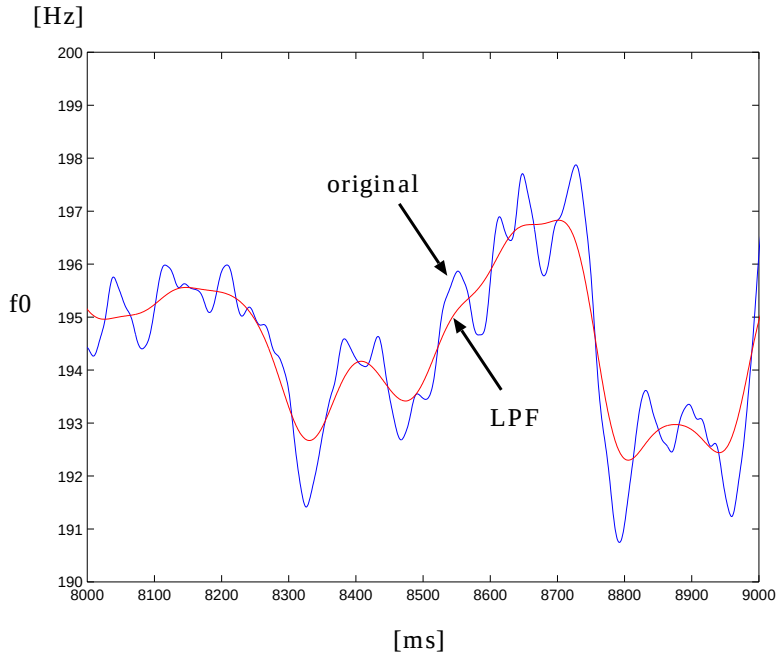
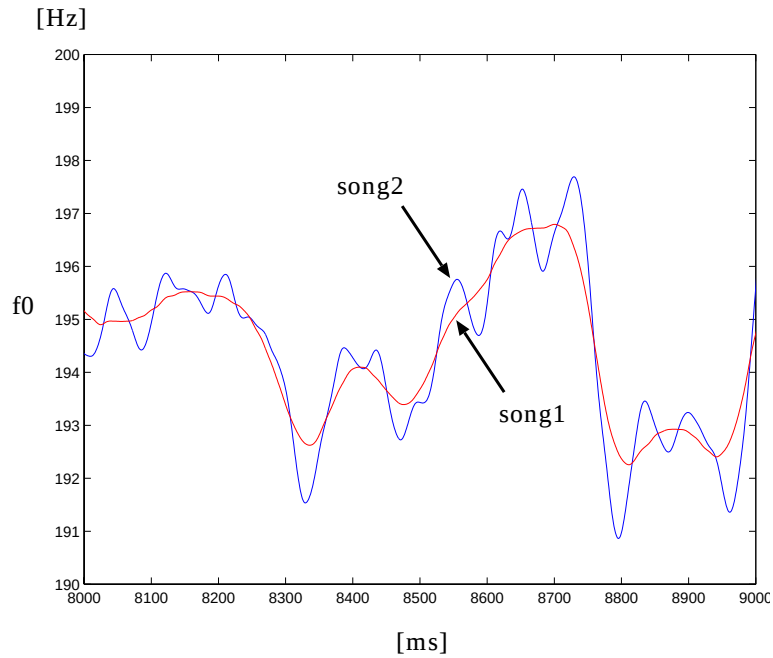


図 3.2: 採取した歌声の基本周波数 (TEMPO2 使用)



3.3 の結果である。この結果から分かるようにほぼ正しく合成音声を作成できていることが分かる。

3.2.3 予備実験

実験方法

歌声の基本周波数の微細変動成分が含まれている音声 (song1) とそうでない音声 (song2) の対比較実験をおこなう。対比較実験方法に基づき A_i, B_i の音声群を作成した。A_i は $TEMPO2$ によって抽出した基本周波数から作成した合成歌声 (song1) であり、B_i は $TEMPO2$ によって抽出した基本周波数にカットオフ周波数 $7Hz$ のローパスフィルタを通して微細変動成分を削除した基本周波数から作成した合成歌声 (song2) である。実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した歌声を提示する (両耳受聴)。被験者は提示音が同じ歌声と知覚した場合は○を解答用紙に記入し、異なる歌声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 13 名である。

実験結果

歌声の基本周波数の微細変動成分が含まれている音声とそうでない音声の対比較実験の結果を表 3.1に示す。基本周波数推定に利用した歌声 (song data) は、$S.K.$氏 (男性) が日本語母音/a/で歌唱した「七つの子」である。ここで $A_i - A_i(song1 - song1)$ または $B_i - B_i(song2 - song2)$ は同じ音声を提示したときに正しい音声であると知覚した認知率である。また $A_i - B_i(song1 - song2)$ または $B_i - A_i(song2 - song1)$ は違う音声を提示したときに違うものであると正しく知覚した認知率である。

この結果から、$A_i - B_i$ の第 5 小節目〜第 8 小節目 (2 回目) の認知率が 46% と低い認知率ではあるが、それ以外はほぼ高い認知率を示している。

3.2.4 予備実験の考察

この結果から、歌声に含まれる基本周波数の微細変動成分に対して、変調周波数と偏移幅の微細な変動の違いを知覚できることがいえた。換言すると、人間は歌声の微細変動成分を十分に知覚可能である。この予備実験から、以後の実験は意味があるといえる。

表 3.1: 予備実験の結果

	認知率 (%)		認知率 (%)	
	$A_i - A_i$	$B_i - B_i$	$A_i - B_i$	$B_i - A_i$
第 1 小節目〜第 4 小節目	77%	77%	92%	77%
第 5 小節目〜第 8 小節目	85%	85%	77%	92%
第 9 小節目〜第 12 小節目	77%	92%	62%	85%
第 13 小節目〜第 16 小節目	69%	69%	92%	100%
第 1 小節目〜第 4 小節目 (2 回目)	85%	69%	77%	69%
第 5 小節目〜第 8 小節目 (2 回目)	100%	85%	46%	77%

3.3 実験 1：変調の有無に対する実験

3.3.1 目的

予備実験では歌声に含まれる基本周波数の微細変動成分に対して、変調周波数と偏移幅の微細な変動を知覚することができた。しかしこの予備実験では変調周波数と偏移幅のどの要素に対して知覚することができたのかは明らかではない。また、仮説でも述べている「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という考え方に対して、「では変調周波数と偏移幅のどの変化に対して、どれくらいの変化から人間は知覚可能であるのか。」という質問に対しては結果は得られない。

したがって、仮説の検証を行なうにあたり、幾つかの制約条件を与えた合成音声を用いて本研究の基盤となる検知限を調べておく必要がある。まずはじめに本節では、変調が存在する場合と存在しない場合の対比較実験を行ない、変調周波数と偏移幅に対する検知限を調べる。ここで得られた結果を基にして実験 2(3.4節)、実験 3(3.5節) では歌声の分析に近付くように様々な条件を足して実験を重ねる。なお実験 1, 2 における刺激音は、2章 2.3.2節合成音声を用いた TEMPO2 の精度評価で用いた合成音声である。

3.3.2 実験 $1-a$

実験方法

基本周波数の変調がない音声と基本周波数に変調を与えた音声との対比較実験を行なう。基本周波数 $f_0 = 125Hz$、偏移幅 $1\%(\alpha = 0.01)$ 一定とし、対比較実験方法に基づき A, B_iの音声群を作成した。A は基本周波数の変調がない音声 (0Hz)、B_iは $f_s = 2Hz$から $62.5Hz$まで $2Hz$刻みで変調が存在する音声である。図 3.4に作成した基本周波数の波形の例を示す。この例は、上から基本周波数 $125Hz$で変調なし、基本周波数 $125Hz$で変調周波数 $6Hz$・偏移幅 1%、基本周波数 $125Hz$で変調周波数 $18Hz$・偏移幅 1%、基本周波数 $125Hz$で変調周波数 $60Hz$・偏移幅 1%である。このように作成した音声を用いて実験を行なう。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 15 名である。

実験結果

変調の有無に対する微細変動成分の知覚分解能を図 3.5に示す。図の折れ線グラフは $A - B_i, B_i - A$ のペアを被験者に提示した時に、正しく違いを知覚することができた認知率に相当する。したがって、図の横軸は B_iの変調周波数である。また、棒グラフは $A - A, B_i - B_i$を被験者に提示した時に、共に等しい音声であるにも関わらず、異なる音声であると誤認識した率である。この結果、変調周波数が約 20Hz までは違いを聞き分けることが可能であるが、$20Hz$以上であると聞き分けることが困難であることが分かる。

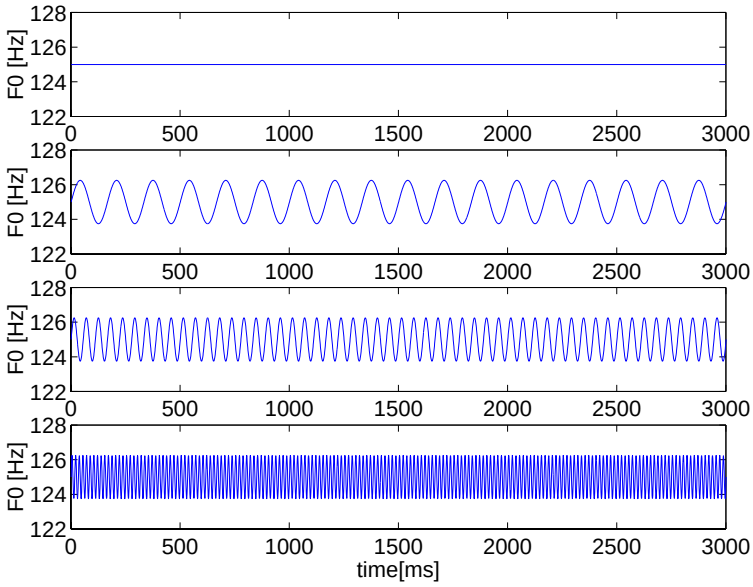


図 3.4: 実験 1-a：波形の例

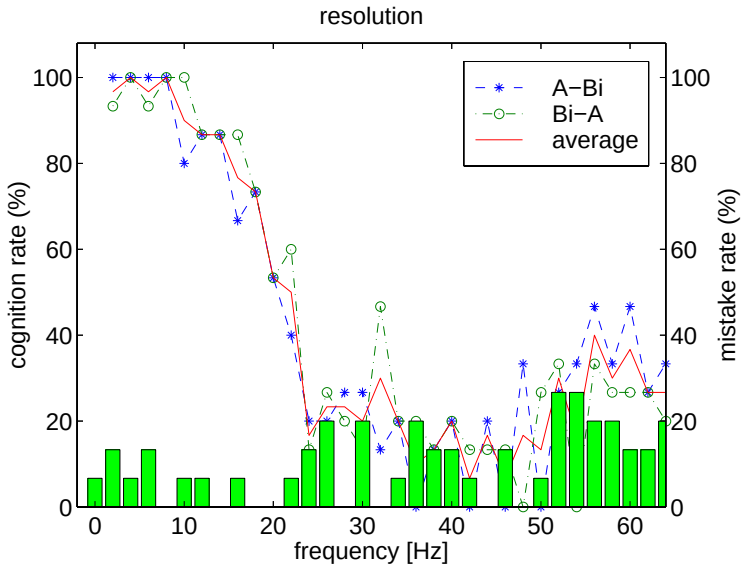


図 3.5: 変調周波数に対する分解能：基本周波数 $125Hz$、偏移幅 1%

3.3.3 実験 $1-b$

実験方法

3.3.2節 (実験 $1-a$) では基本周波数が $125Hz$であったが、歌唱の際に曲の旋律概形に音程を合わせるために基本周波数の上昇が必要となる。したがって基本周波数の上昇に伴ない微細変動成分の知覚の影響が存在する可能性があることから基本周波数が $250Hz$ の場合においても同様の実験を行なう。その他の実験環境は実験 $1-a$ と同じである。図3.6に作成した基本周波数の波形の例を示す。この例は、上から基本周波数 $250Hz$で変調なし、基本周波数 $125Hz$で変調周波数 $6Hz$・偏移幅 1%、基本周波数 $250Hz$で変調周波数 $6Hz$・偏移幅 1%である。同じ偏移幅 1%でも基本周波数によって距離が違うことに注意。このように作成した音声を用いて実験を行なう。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する(両耳受聴)。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 15 名である。

実験結果

基本周波数が $250Hz$の場合における微細変動成分の知覚分解能を図3.7に示す。この結果、変調周波数が約 60Hz までは違いを聞き分けることが可能であるが、$60Hz$以上であると聞き分けることが困難であることが分かる。実験 $1-a$ の結果と比較して基本周波数が上昇した場合には変調周波数に対する検知限の上昇も見られることが分かる。

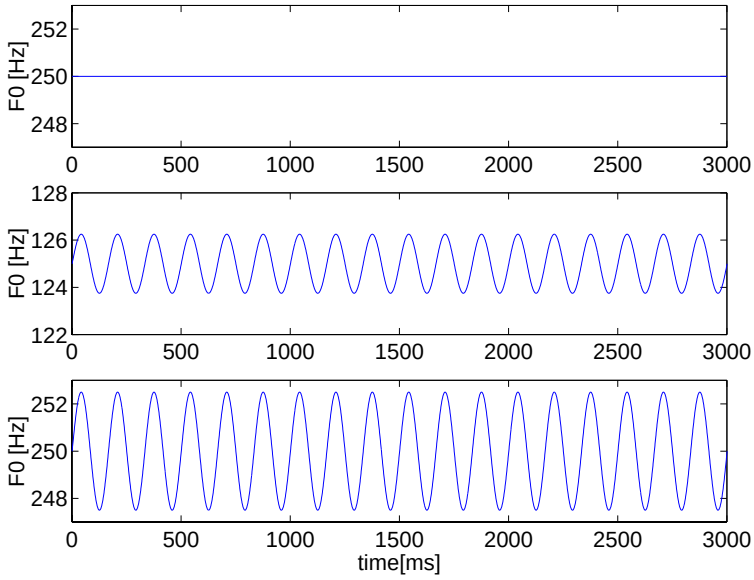


図 3.6: 実験 1-b：波形の例

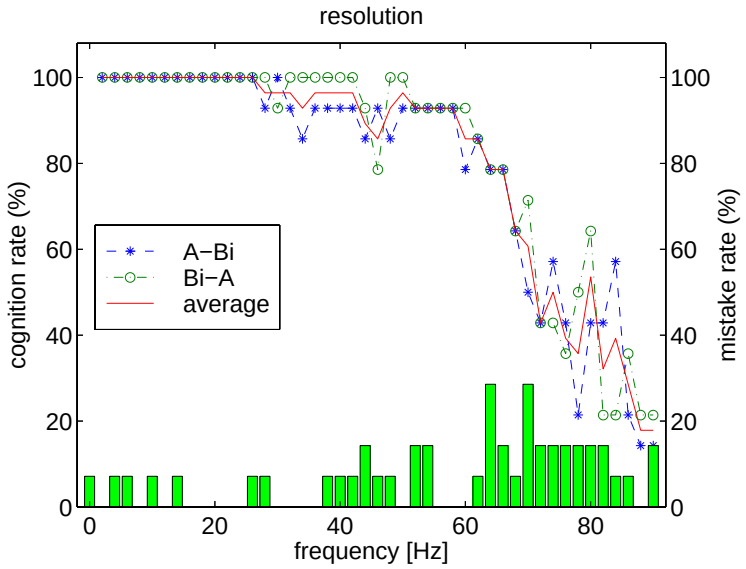


図 3.7: 変調周波数に対する分解能：基本周波数 $250Hz$、偏移幅 1%

3.3.4 実験 $1-c$

実験方法

3.3.2節 (実験 $1-a$) では偏移幅が 1%であったが、歌声の分析では微細変動成分（不規則的変動成分）の偏移幅が最大箇所で 2%存在すると言える結果が得られた。(平均では偏移幅は約 1.2%程度である。) したがって偏移幅の上昇に伴ない微細変動成分の知覚の影響が存在する可能性があることから偏移幅が 2%の場合においても同様の実験を行なう。その他の実験環境は実験 $1-a$ と同じである。図 3.8に作成した基本周波数の波形の例を示す。この例は、上から基本周波数 $125Hz$で変調なし、基本周波数 $125Hz$で変調周波数 $6Hz$・偏移幅 1%(実験 $1-a$ で用いた偏移幅)、基本周波数 $125Hz$で変調周波数 $6Hz$・偏移幅 2%である (実験 $1-c$ で用いる偏移る幅)。このように作成した音声を用いて実験を行なう。ここで、実験 $1-b$ の基本周波数 $250Hz$で偏移幅 1%と、基本周波数 $125Hz$で偏移幅 2%は、偏移幅が共に $2.5Hz$で等しいことに注意。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 15名である。

実験結果

偏移幅が 2%の場合における微細変動成分の知覚分解能を図 3.9に示す。この結果検知限は $40Hz$であるといえるが、変調周波数が 40Hz を越えても微細変動成分の知覚が可能であることが分かる。実験 $1-a$ の結果と比較して偏移幅が上昇した場合には変調周波数に対する検知限の上昇も見られることが分かる。

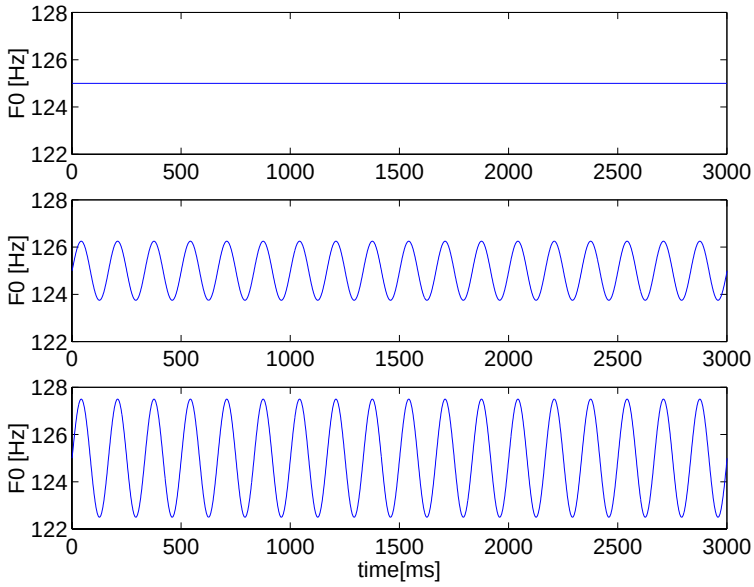


図 3.8: 実験 1-c：波形の例

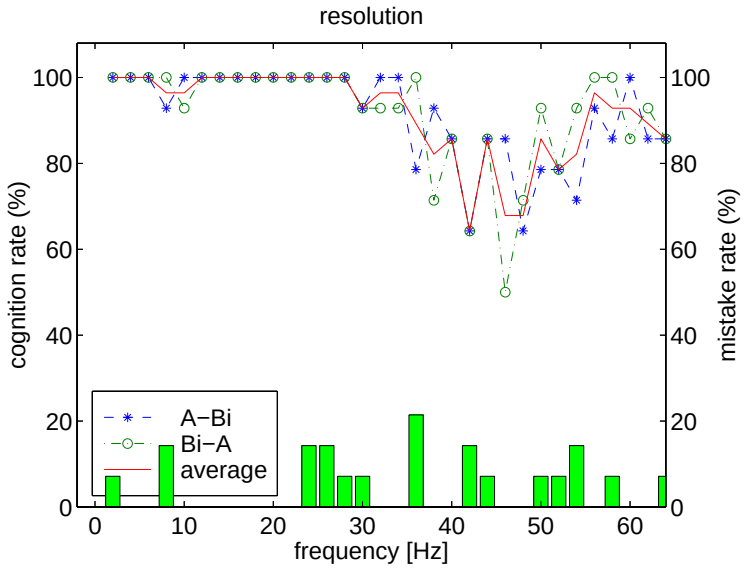


図 3.9: 変調周波数に対する分解能：基本周波数 $125Hz$、偏移幅 2%

3.3.5 実験 $1-d$

実験方法

これまでは基本周波数と偏移幅を一定にして変調周波数に対する検知限について実験を行なってきたが、次に基本周波数 $(125Hz)$ と変調周波数を一定にして偏移幅に対する検知限について検討を行なった。対比較実験方法に基づき A, B_i の音声を作成する。A は基本周波数の偏移幅がない音声 (0%)、B_i は 0.1% から 2.0% まで 0.1% きざみで偏移幅が存在する音声である。実験環境は上述の聴取実験と同じである。図 3.10に作成した基本周波数の波形の例を示す。この例は、上から基本周波数 $125Hz$ で変調なし、基本周波数 $125Hz$ で変調周波数 $6Hz$・偏移幅 0.5%、基本周波数 $125Hz$ で変調周波数 $6Hz$・偏移幅 1%、基本周波数 $125Hz$ で変調周波数 $6Hz$・偏移幅 2% である。このように作成した音声を用いて実験を行なう。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 15 名である。

実験結果

変調周波数が $3Hz, 6Hz, 12Hz, 24Hz, 48Hz$ において実験を行なった結果を図 3.11から図 3.15に示す。また各々の検知限を示したものを表 3.2に示す。上述同様折れ線グラフは認知率、棒グラフは誤り率である。この結果、変調周波数が $3Hz$ の時には基本周波数に対する偏移幅が 0.6%以上であれば、十分に違いを聞き分けることが可能であることが分かる。また変調周波数が $6Hz$ では 0.7%、$12Hz$ では 1.0%、$24Hz$ では 1.1%、$48Hz$ では 1.4% 以上の偏移幅であれば違いを聞き分けることができることが言える。

表 3.2: 実験 $1-d$ の結果

f_s	3Hz	6Hz	12Hz	24Hz	48 Hz
検知限 (偏移幅)	0.6%	0.7%	1.0%	1.1%	1.4%

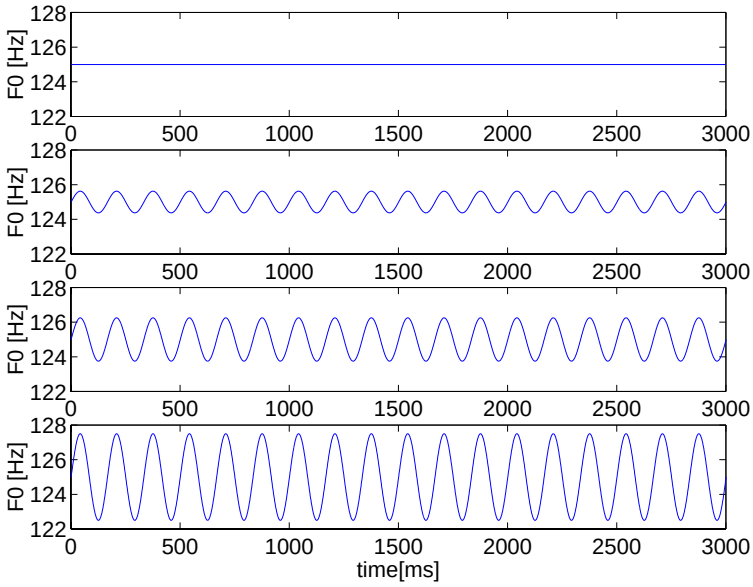


図 3.10: 実験 1-d：波形の例

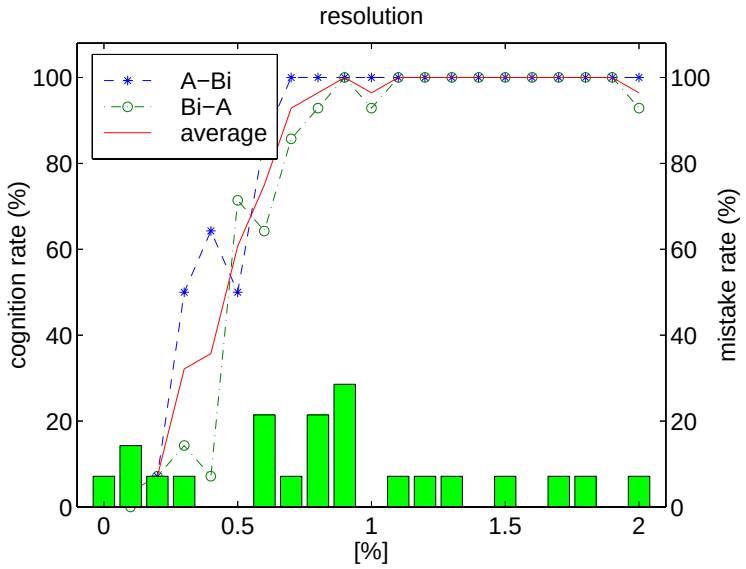


図 3.11: 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $3Hz$

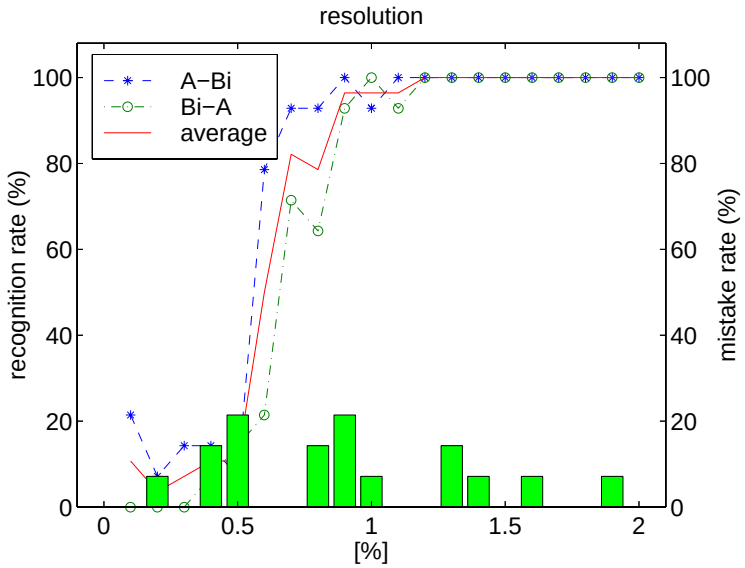


図 3.12: 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $6Hz$

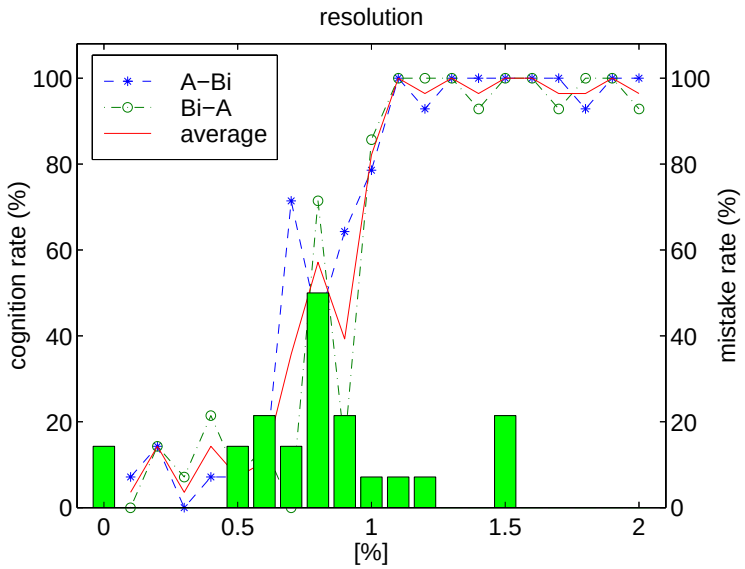


図 3.13: 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $12Hz$

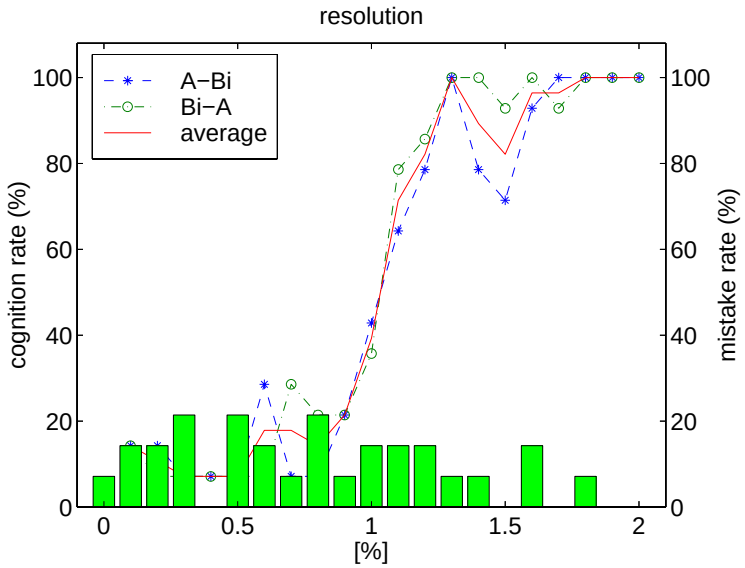


図 3.14: 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $24Hz$

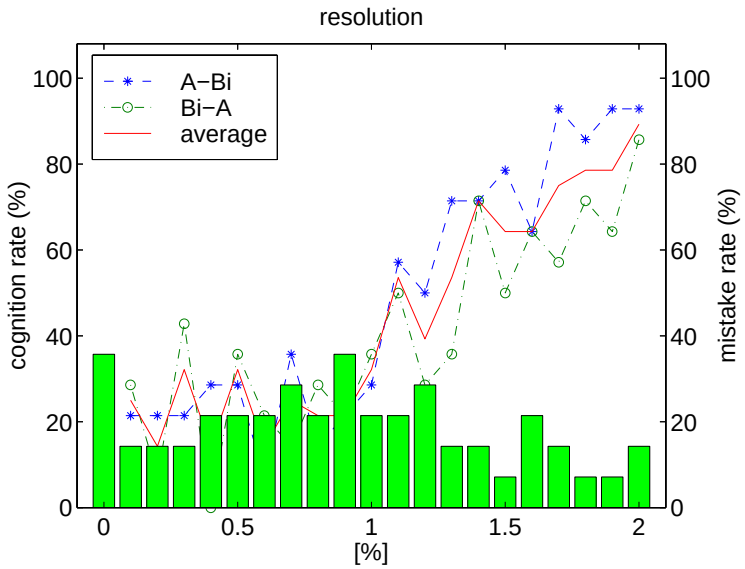


図 3.15: 偏移幅に対する分解能：基本周波数 $125Hz$、変調周波数 $48Hz$

3.3.6 実験 1 の考察

実験 1 の変調が存在する場合と存在しない場合の対比較実験の結果から次のようなことが考察できる。

- 基本周波数に対する検知限について実験 $1-a$ と実験 $1-b$ の比較から考察する。実験 $1-a$ より基本周波数 $125Hz$・偏移幅 1%、変調周波数が $0Hz$ から $62.5Hz$まで変化した時の検知限は約 $20Hz$である。また実験 $1-b$ から基本周波数 $250Hz$・偏移幅 1%、変調周波数が $0Hz$ から $90Hz$まで変化した時の検知限は約 $60Hz$である。この結果から基本周波数が単純に変化しただけで、微細変動成分の検知限に影響があることが分かる。基本周波数には検知限を変動させる要因がある可能性がある。

- 偏移幅に対する検知限について実験 $1-a$ と実験 $1-c$ の比較から考察する。実験 $1-c$ では実験 $1-a$ の偏移幅を 1%から 2%に変えて実験を行なった。この結果、検知限は $20Hz$から $40Hz$へと上昇する結果となった。このことから、偏移幅は微細変動の知覚に大きく影響を及ぼしていることがいえる。

- 実験 $1-b$ と実験 $1-c$ について考察する。実験 $1-b$ は基本周波数 $250Hz$に対して偏移幅 $1\% = 2.5Hz$、また実験 $1-c$ は基本周波数 $125Hz$に対して偏移幅 $2\% = 2.5Hz$と、偏移幅は同じ $2.5Hz$である。単純に偏移幅だけで検知限が決定するのであれば、検知限は等しくなるはずである。しかし、実験 $1-b$ では検知限が $60Hz$であり、実験 $1-c$ では検知限が $40Hz$であり、検知限は等しくならない。この結果から、微細変動成分は偏移幅のみで決定されるということはないように思われる。これは図 3.7と図 3.9を比較しても理解できる。更に、基本周波数と偏移幅には何らかの関係があることがいえる。

- 実験 $1-d$ では基本周波数と偏移幅を一定にして変調周波数に対する検知限を調べた。この結果から、変調周波数が低い周波数 (実験例では $3Hz$) に比べて変調周波数が高い周波数 (実験例では $48Hz$) は偏移幅の検知限が低いことがいえる。これは低い変調周波数のほうが知覚し易いことを意味している。

実験 1 の変調が存在する場合と存在しない場合の対比較実験の結果と 2章で得られた分析結果とてらし合わせて仮説の検証を行なう。2章では歌声には偏移幅が平均 $20cent$ 程度存在し変調周波数は $15Hz$程度存在することが分析できている。

ここで基本周波数が約 $125Hz$から $250Hz$までを変動している歌声を知覚可能であるかを検討する。歌声分析によって得られた偏移幅 $20cnet$ は基本周波数 f_0によらず、$\pm 1.2\%$

の偏移幅である。したがって、この場合基本周波数$125Hz$から$250Hz$で変動する歌声に対しては、偏移幅が約$\pm 1.2\%$で変調周波数は$15Hz$程度含まれていることになる。実験$1-a$の結果では、基本周波数が$125Hz$で偏移幅1%の場合には、変調周波数を$20Hz$まで知覚可能であるという結果が得られている。また実験$1-b$の結果から、基本周波数が$250Hz$で偏移幅1%の場合には、変調周波数を$60Hz$まで知覚可能であるという結果が得られている。実験$1-c$の結果から偏移幅が上昇した場合には検知限も上昇することがいえているので、偏移幅が1.2%である場合は、検知限も上昇して考えても良いといえる。したがって、歌声に含まれる基本周波数の微細変動成分は知覚可能であることがいえた。

逆に変調周波数が$15Hz$存在する歌声を知覚するには何%の偏移幅が必要かを考察する。この場合、実験$1-d$の実験結果(表3.2)から偏移幅が1.1%存在すれば良いといえる。歌声には偏移幅が平均で1.2%存在していたことから、微細変動成分は知覚可能であるといえる。

3.4　実験 2：変調が共に存在する場合に対する対比較実験

3.4.1　目的

実際の歌声を考えた場合、時々刻々と変調周波数と偏移幅が変化していると考えられる。実験 1 では変調が存在する場合と存在しない場合との検知限を調べたものであるが、実際には変調が存在しないものからの距離を知覚しているだけではなく、変調の変化成分も知覚に影響している可能性がある。したがって変調周波数が時々刻々と変化する実際の音声を考慮し、対の音声において共に変調が存在する場合の検知限についても同様の実験を行なう。実験 $2-a$ では変調周波数差が $1Hz$の違いを持つ音声ペアに焦点をあてた。また実験 $2-b$ では実験 $2-a$ と同様の実験で変調周波数差が $5Hz$、実験 $2-c$ では $10Hz$で、その音声対の差を広げる。これらの実験により知覚のメカニズムを考察する。実験 2 も実験 1 同様、2 章 2.3.4 節で作成した合成音声を用いる。

3.4.2　実験 $2-a$

実験方法

基本周波数 $125Hz, 250Hz$の各々において偏移幅 $1\%, 2\%$で変調周波数差が $1Hz$の違いを持つ音声のペアについて実験を行なう。このパラメータに対する検知限の値を測定する。実験環境は実験 1 と同様である。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 15 名である。

実験結果

図 3.16に基本周波数 $125Hz$、偏移幅 1%で変調周波数差が $1Hz$のペアにおける対比較実験の結果を示す。ここで A_iは $0Hz$から $19Hz$まで $1Hz$刻み、B_iは $1Hz$から $20Hz$まで $1Hz$刻みの変調周波数を持つ音声である。3.3節同様折れ線グラフは認知率を示し、横軸は A_i-B_i, B_i-A_iペアの変調周波数の小さい方に相当する。また棒グラフは、誤り率に相当し、横軸は A_i-A_i, B_i-B_iの変調周波数に相当する。この結果、変調周波数が $5Hz-6Hz$ペアまでは違いを聞き分けることが可能であるが、それ以上であると違いを聞き分けることが困難になると考えられる。したがって検知限は $5Hz$である。すなわち変調周波数が

$5Hz$と $6Hz$の違いまでは知覚可能であるが、$6Hz$と $7Hz$では違いを聞き分けることが困難になるといえる。

図 3.17に基本周波数 $125Hz$、偏移幅 2%における対比較実験の結果を示す。この結果から検知限は $7Hz$である。すなわち変調周波数が $7Hz$と $8Hz$の違いまでは知覚可能であるが、$8Hz$と $9Hz$では違いを聞き分けることが困難になるといえる。

図 3.18に基本周波数 $250Hz$、偏移幅 1%における対比較実験の結果を示す。この結果から検知限は $6Hz$である。すなわち変調周波数が $6Hz$と $7Hz$の違いまでは知覚可能であるが、$7Hz$と $8Hz$では違いを聞き分けることが困難になるといえる。

図 3.19に基本周波数 $250Hz$、偏移幅 2%における対比較実験の結果を示す。この結果から検知限は $8Hz$である。すなわち変調周波数が $8Hz$と $9Hz$の違いまでは知覚可能であるが、$9Hz$と $10Hz$では違いを聞き分けることが困難になるといえる。

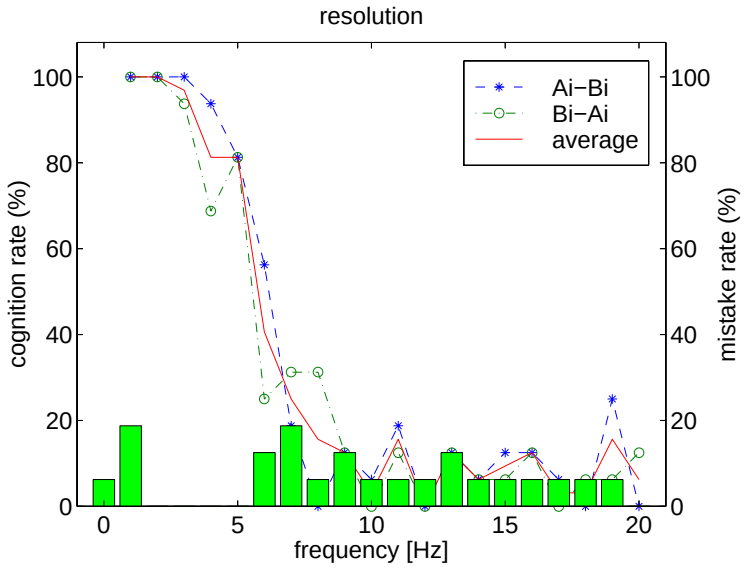


図 3.16: 変調周波数差 $1Hz$のペア：基本周波数 $125Hz$、偏移幅 1%

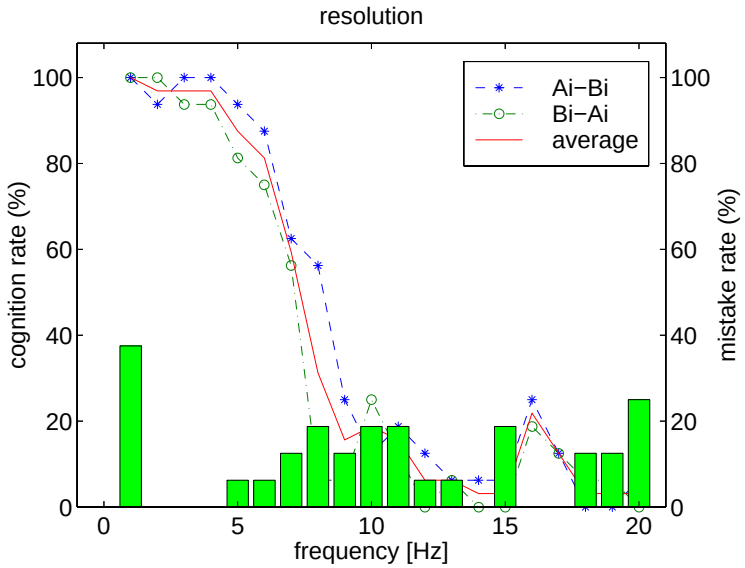


図 3.17: 変調周波数差 $1Hz$のペア：基本周波数 $125Hz$、偏移幅 2%

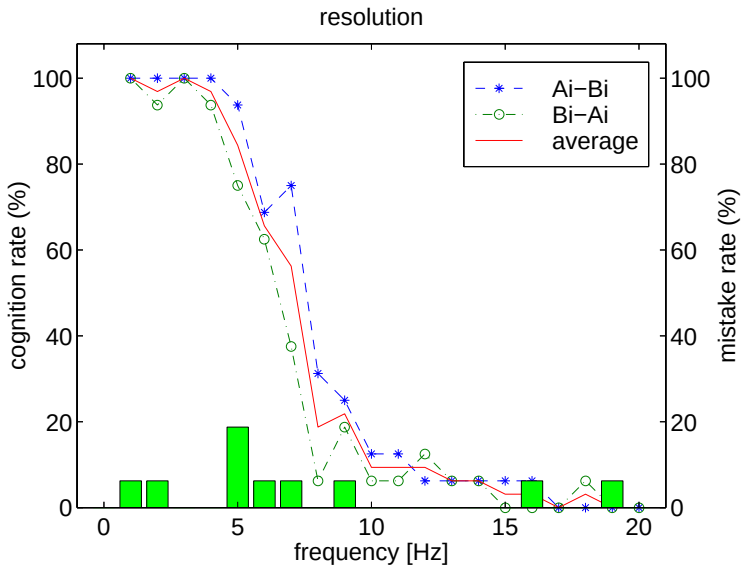


図 3.18: 変調周波数差 $1Hz$のペア：基本周波数 $250Hz$、偏移幅 1%

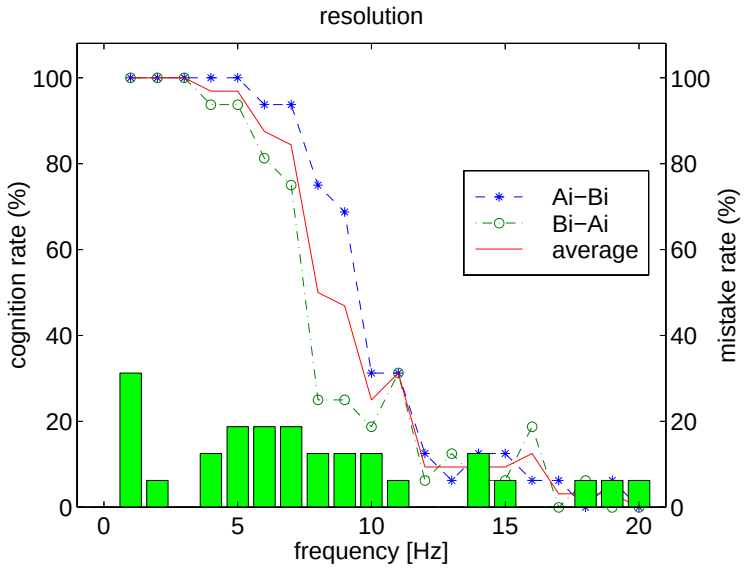


図 3.19: 変調周波数差 $1Hz$のペア：基本周波数 $250Hz$、偏移幅 2%

3.4.3 実験 $2-b$

実験方法

基本周波数 $125Hz, 250Hz$の各々において偏移幅 $1\%, 2\%$で変調周波数差が $5Hz$ の違いを持つ音声のペアについて実験を行なう。このパラメータに対する検知限の値を測定する。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 14 名である。

実験結果

変調周波数差が $5Hz$における結果は次のようなものであった。図 3.20に基本周波数 $125Hz$、偏移幅 1%における対比較実験の結果を示す。この結果から検知限は $9Hz$である。すなわち変調周波数が $9Hz$と $14Hz$の違いまでは知覚可能であるが、$10Hz$と $15Hz$では違いを聞き分けることが困難になるといえる。

図 3.21に基本周波数 $125Hz$、偏移幅 2%における対比較実験の結果を示す。この結果から検知限は $11Hz$である。すなわち変調周波数が $11Hz$と $16Hz$の違いまでは知覚可能であるが、$12Hz$と $17Hz$では違いを聞き分けることが困難になるといえる。

図 3.22に基本周波数 $250Hz$、偏移幅 1%における対比較実験の結果を示す。この結果から検知限は $11Hz$である。すなわち変調周波数が $11Hz$と $16Hz$の違いまでは知覚可能であるが、$12Hz$と $17Hz$では違いを聞き分けることが困難になるといえる。

図 3.23に基本周波数 $250Hz$、偏移幅 2%における対比較実験の結果を示す。この結果から検知限は $13Hz$である。すなわち変調周波数が $13Hz$と $18Hz$の違いまでは知覚可能であるが、$14Hz$と $19Hz$では違いを聞き分けることが困難になるといえる。

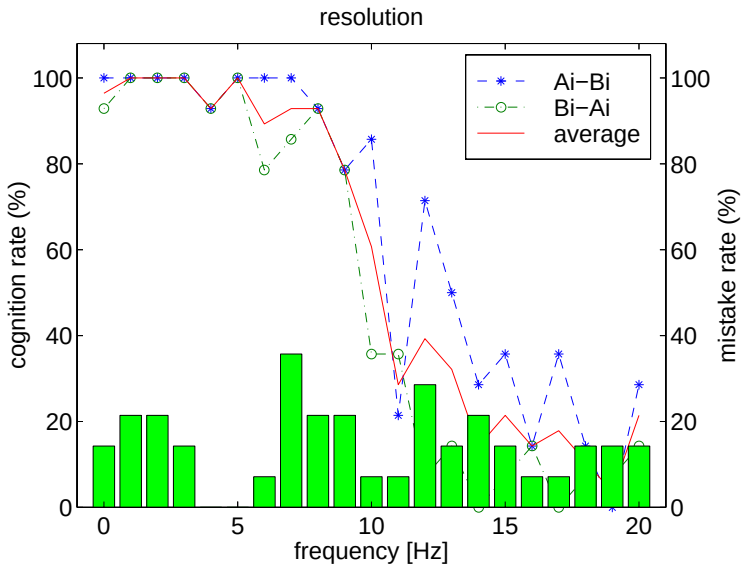


図 3.20: 変調周波数差 $5Hz$のペア：基本周波数 $125Hz$、偏移幅 1%

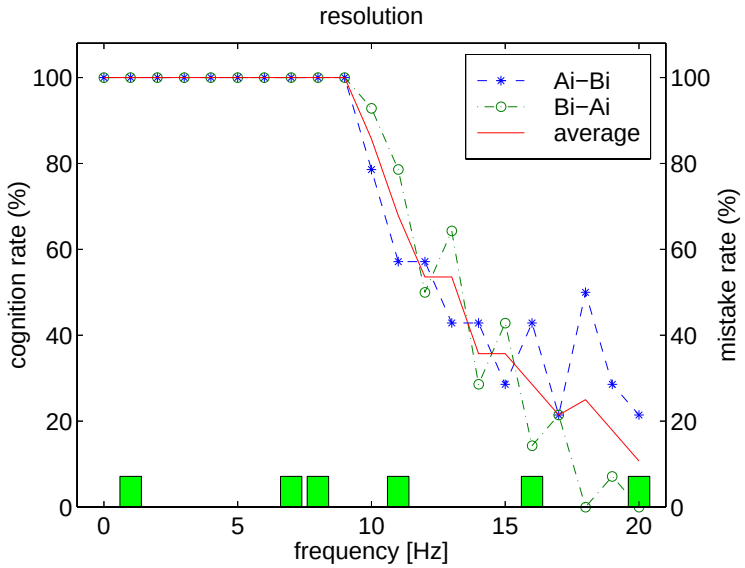


図 3.21: 変調周波数差 $5Hz$のペア：基本周波数 $125Hz$、偏移幅 2%

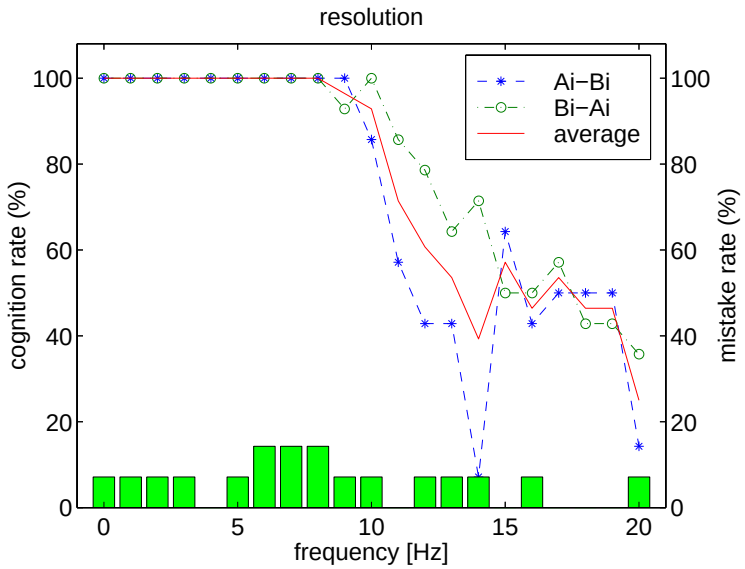


図 3.22: 変調周波数差 $5Hz$のペア：基本周波数 $250Hz$、偏移幅 1%

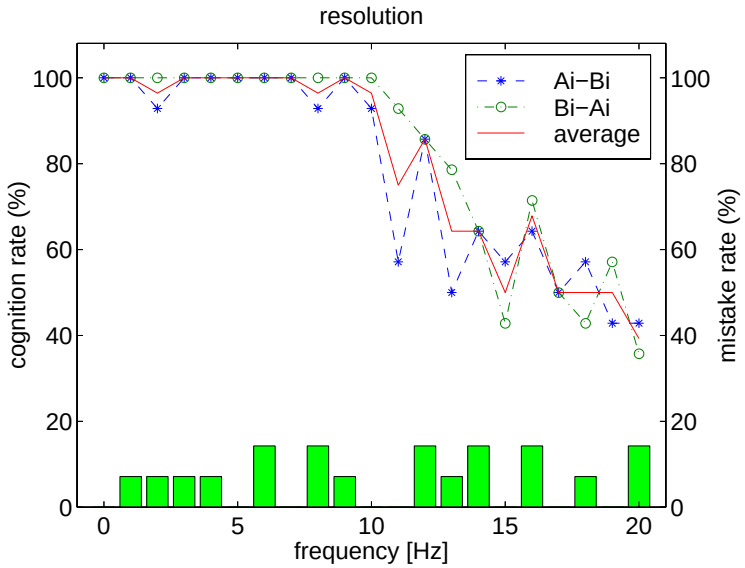


図 3.23: 変調周波数差 $5Hz$のペア：基本周波数 $250Hz$、偏移幅 2%

3.4.4 実験 $2-c$

実験方法

基本周波数 $125Hz, 250Hz$の各々において偏移幅 $1\%, 2\%$で変調周波数差が $10Hz$ の違いを持つ音声のペアについて実験を行なう。このパラメータに対する検知限の値を測定する。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 14 名である。

実験結果

変調周波数差が $10Hz$における結果は次のようなものであった。図 3.24に基本周波数 $125Hz$、偏移幅 1%における対比較実験の結果を示す。この結果から検知限は $12Hz$である。すなわち変調周波数が $12Hz$と $22Hz$の違いまでは知覚可能であるが、$13Hz$と $23Hz$では違いを聞き分けることが困難になるといえる。

図 3.25に基本周波数 $125Hz$、偏移幅 2%における対比較実験の結果を示す。この結果から検知限は $14Hz$である。すなわち変調周波数が $14Hz$と $24Hz$の違いまでは知覚可能で

あるが、15Hzと 25Hzでは違いを聞き分けることが困難になるといえる。

図 3.26に基本周波数 250Hz、偏移幅 1%における対比較実験の結果を示す。この結果から検知限は 17Hzである。すなわち変調周波数が 17Hzと 27Hzの違いまでは知覚可能であるが、18Hzと 28Hzでは違いを聞き分けることが困難になるといえる。

図 3.27に基本周波数 250Hz、偏移幅 2%における対比較実験の結果を示す。この結果から検知限は 20Hz以上であるといえる。

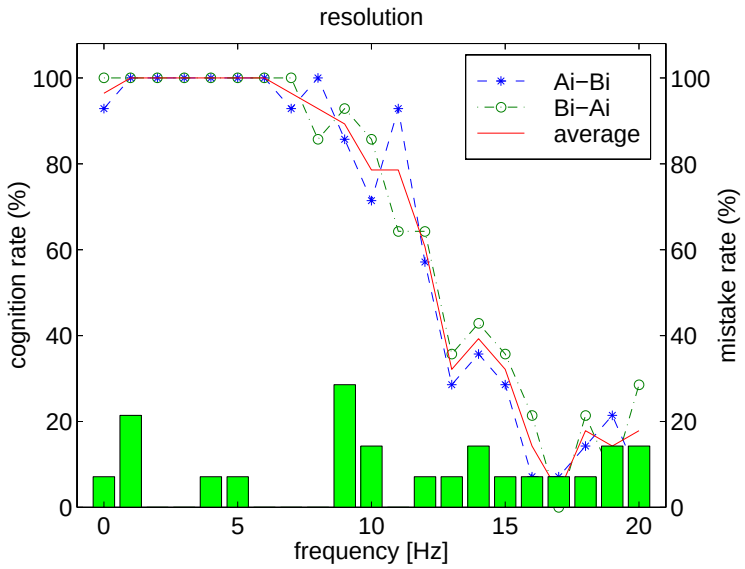


図 3.24: 変調周波数差 10Hzのペア：基本周波数 125Hz、偏移幅 1%

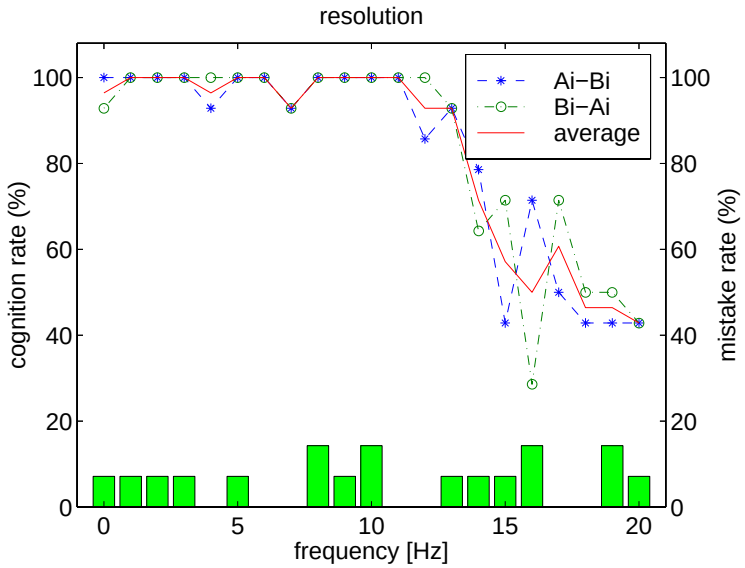


図 3.25: 変調周波数差 $10Hz$のペア：基本周波数 $125Hz$、偏移幅 2%

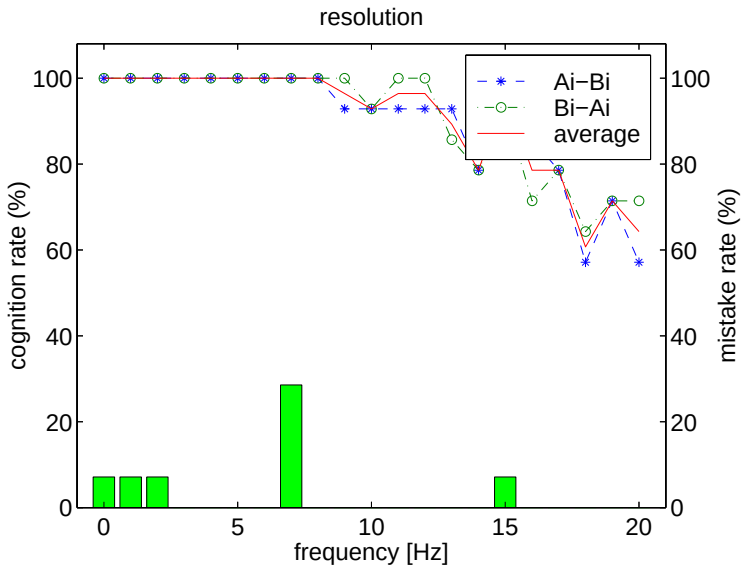


図 3.26: 変調周波数差 $10Hz$のペア：基本周波数 $250Hz$、偏移幅 1%

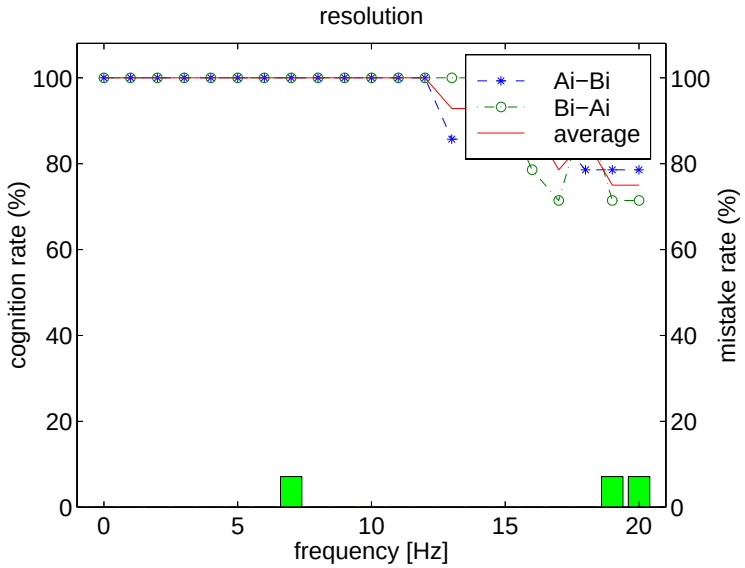


図 3.27: 変調周波数差 $10Hz$のペア：基本周波数 $250Hz$、偏移幅 2%

3.4.5 実験 2 の考察

表 3.3: 実験 2 の結果

変調周波数差	$f_0 : 125Hz$		$f_0 : 250Hz$	
	$\alpha : 1\%$	$\alpha : 2\%$	$\alpha : 1\%$	$\alpha : 2\%$
1Hz（実験 $2-a$）	5Hz	7Hz	6Hz	8Hz
5Hz（実験 $2-b$）	9Hz	11Hz	11Hz	13Hz
10Hz（実験 $2-c$）	12Hz	14Hz	17Hz	20Hz 以上

実際の歌声を考えた場合、時々刻々と変調周波数と偏移幅が変化していると考えられる。実験 1 では変調が存在する場合と存在しない場合との検知限を調べたものであるが、実際には変調が存在しないものからの距離を知覚しているだけではなく、変調の変化成分も知覚に影響している可能性がある。実験 2 の結果を用いて考察を行なう。

実験 $2-a, 2-b, 2-c$ の検知限の結果をまとめたものを表 3.3に示す。この結果から次のことがいえる。

- 対比較実験に用いる音声ペアの変調周波数差が $1Hz$ の対比較実験において基本周波数が $125Hz, 250Hz$の両者とも偏移幅が大きい方が高い変調周波数まで認知できる

ことが分かる。これは変調周波数差が $1Hz$ に限ったことではなく、実験 $2-b, 2-c$ の変調周波数差 $5Hz, 10Hz$ に対してもいえる。この結果は、偏移幅が微細変動の知覚に大きく影響を及ぼしているという実験 1 の考察と類似する。

- 音声対の変調周波数差 $1Hz, 5Hz, 10Hz$ を比較した場合、音声対の変調周波数差が大きい方が高い周波数まで認知可能であるという結果が得られた。これから基本周波数の微細変動成分の急激な変化に対して非常に敏感に認知できる可能性があるといえる。

- 基本周波数 $125Hz$ と $250Hz$ を比較してみると単純に基本周波数が高い方が、高い変調周波数まで認知できるという結果がえられた。これは実験 1 において基本周波数の変動に対する変調周波数の検知限の影響を示しており、実験 1 と同様の結果であると考察できる。

実験 2 の変調が共に存在する場合に対する実験と 2章で得られた分析結果とをてらし合わせて仮説の検証を行なう。ここで仮説は次のようなものであった。

「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」

2章の考察で音程が安定している時には変調周波数が低く ($15Hz$程度)、偏移幅も小さい ($20cent$) が、音程変化により基本周波数が変化すると変調周波数が高くなり ($20Hz$)、それに伴い偏移幅も大きくなる (60 cent) という変調の変化成分が増大する分析結果が得られている。この分析結果は上述の考察から音程の変化時において、基本周波数の変動が見られると変調周波数が増加する。変調周波数が増加すると、前の変調周波数からの差が必然的に大きくなる。変調周波数の差が大きくなれば、より知覚しやすい状態になるという結果にたどり着く。また変調周波数が増加すると、偏移幅の増加も生じる (分析結果より)。このような連鎖反応で譜面にあわせて音程変化を行なう歌声の場合は、より微細な変動を知覚し易くなる可能性がある。したがって仮説に示した「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という可能性をより強く示すことができる。

3.5　変調が 2 つ存在する場合の実験

3.5.1　目的

低い変調周波数に高い変調周波数が存在する場合の合成音声の検知限を確認する。これは規則的変動成分 (ビブラート) などの要因によって微細変動成分の検知限が変化するかを考察するためである。また 3.4節よりも更に変調周波数が時々刻々と変化する実際の音声に近似した合成音であると考えることができ、知覚に与える影響を持つパラメータをより詳しく決定することができる。

実験 $3-a$ では変調周波数 f_{s2}(式 3.1) を変化させた時の検知限について焦点をあてる。ここで変調周波数 f_{s1}はビブラートを考慮にいれて、$6Hz$一定とする。

また実験 $3-b$ では変調周波数差 f_{s2}(式 3.1) が $1Hz$、$5Hz$、$10Hz$を持つ音声ペアに対して焦点をあてる。実験 $3-a$ と同様、波数 f_{s1}はビブラートを考慮にいれて、$6Hz$一定とする。

3.5.2　音声データ

基本周波数の規則的変動成分に微細変動成分が含まれている場合、規則的変動成分 (ビブラート) などの要因によって微細変動成分の検知限が変化する可能性がある。そこで変調された基本周波数 $f_m(t)$ を次式で表し、2.3.2節で用いた手法により合成音声を作成し知覚に及ぼす影響の確認を行なう。ここでα_1, α_2は偏移幅であり、f_{s1}, f_{s2}は変調周波数である。これまでと同様 f_0は基本周波数である。

$$f_m(t) = f_0 + \alpha_1 \cdot f_0 \cdot sin(2\pi f_{s1} t) + \alpha_2 \cdot f_0 \cdot sin(2\pi f_{s2} t) \tag{3.1}$$

3.5.3　実験 $3-a$

実験方法

基本周波数 $f_0 = 125Hz, 250Hz$とし対比較実験法に基づき A, B_iの音声対を作成する。A は式 2.2の変調周波数 f_sを $6Hz$、偏移幅αを 1%一定とした音声である。B_iは式 3.1の変調周波数 f_{s1}が 6Hz で f_{s2}が $0Hz$から $62.5Hz(90Hz)$（$2Hz$刻み）で変調している音声である。また偏移幅はα_1, α_2共に 1%である。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異な

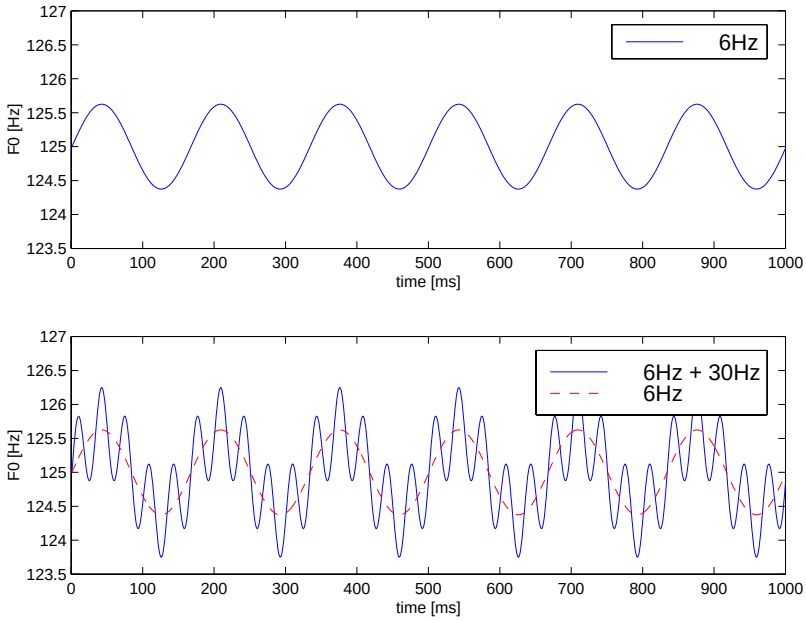


図 3.28: 音声データの例 (上：変調周波数:$6Hz$, 下：変調周波数:$6Hz + 30Hz$)

る音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 14 名である。

実験結果

基本周波数 $125Hz$に対しての結果を図 3.29に示す。また基本周波数 $250Hz$に対しての結果を図 3.30に示す。図の折れ線グラフは $A - B_i, B_i - A$ のペアを被験者に提示した時に、正しく違いを知覚することができた認知率に相当する。したがって、図の横軸は $A - B_i, B_i - A$ ペアの変調周波数 $B_i(f_{s2})$ である。また、棒グラフは $A - A, B_i - B_i$を被験者に提示した時に、共に等しい音声であるにも関わらず、異なる音声であると誤認識した率である。したがって、図の横軸は $A - A, B_i - B_i$の変調周波数 (f_{s2}) に相当する。この結果、基本周波数 $125Hz, 250Hz$に対して検知限が各々$20Hz, 60Hz$であった表 3.4。

表 3.4: 実験 $3 - a$ の結果

f_0	125Hz	250Hz
検知限 (変調周波数)	$20Hz$	$60Hz$

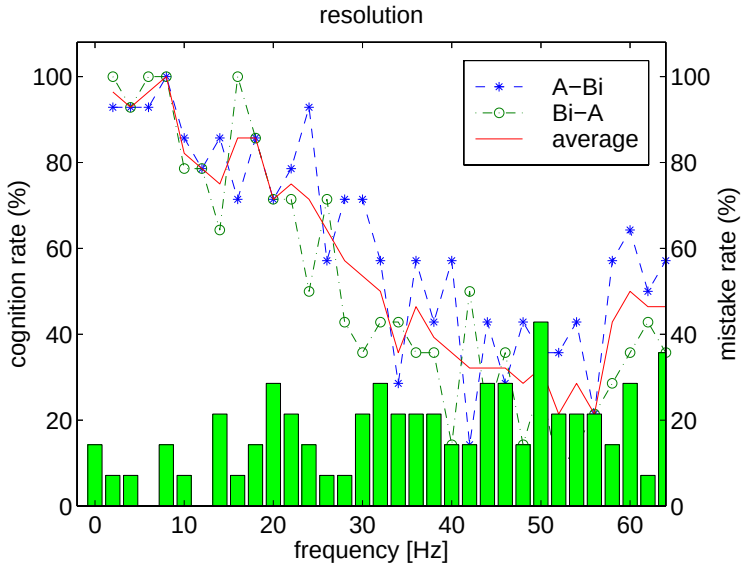


図 3.29: 結果 3-a：$f_0 = 125Hz$

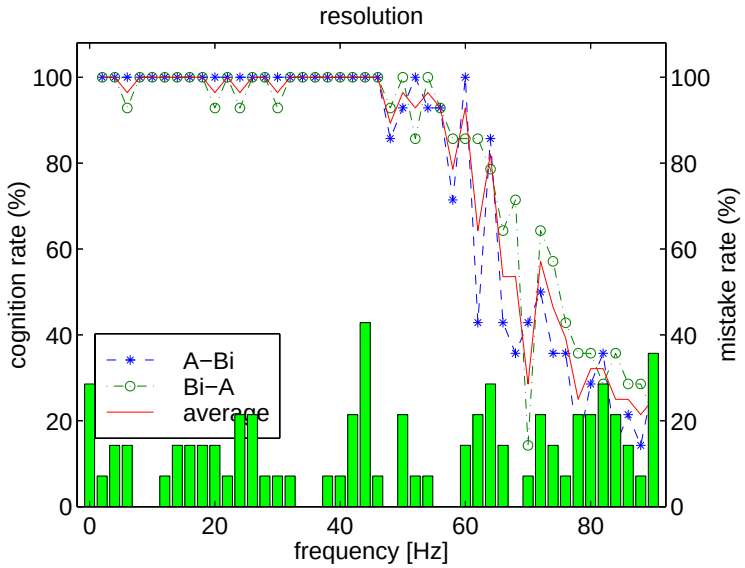


図 3.30: 結果 3-a：$f_0 = 250Hz$

3.5.4 実験 $3-b$

実験方法

緩やかな変調に微細な変調があるような 2 つの変調周波数を含む音声同士の聴取実験を行なう。対比較実験法に基づき A_i, B_iの音声対を式 3.1 より作成する。実験は基本周波数 $125Hz, 250Hz$の両者について行なう。各々のパラメータは変調周波数 f_{s1}が 6Hz で f_{s2}が $0Hz$から $62.5Hz(90Hz)$（$2Hz$刻み）で変調している音声を用いる。また偏移幅はα_1, α_2 共に 1%である。ここで音声対 A_i, B_iの違いは、変調周波数 f_{s2}の差として $1Hz$、$5Hz$、$10Hz$を与えたことにある。

実験は防音室内部で行ない、ヘッドホンを介して被験者に作成した合成音声を提示する（両耳受聴）。被験者は提示音が同じ音声と知覚した場合は○を解答用紙に記入し、異なる音声であると知覚した場合は×を記入する。ここで被験者は正常聴力を有する学生 14 名である。

実験結果

図 3.31に基本周波数 $125Hz$で変調周波数 f_{s2}の差が $1Hz$の結果を示す。ここで図の A_i は $0Hz$から $19Hz$まで $1Hz$刻み、B_iは $1Hz$から $20Hz$まで $1Hz$刻みの変調周波数を持つ音声である。3.3節、3.4節と同様、折れ線グラフは認知率を示し、横軸は $A_i - B_i, B_i - A_i$ ペアの変調周波数 f_{s2}の小さい方に相当する。また棒グラフは、誤り率に相当し、横軸は $A_i - A_i, B_i - B_i$の変調周波数 f_{s2}に相当する。同様に図 3.32、図 3.33は基本周波数 $125Hz$ で変調周波数 f_{s2}の差が $5Hz$、$10Hz$の結果である。また、図 3.34、図 3.35、図 3.36は基本周波数 $250Hz$で変調周波数 f_{s2}の差が $1Hz$、$5Hz$、$10Hz$の結果である。これらの結果から検知限を導いたものが表 3.5 である。

表 3.5: 実験 $3-b$ の結果

変調周波数の差	125Hz	250Hz
$1Hz$	$7Hz$	$9Hz$
$5Hz$	$9Hz$	$11Hz$
$10Hz$	$11Hz$	$14Hz$

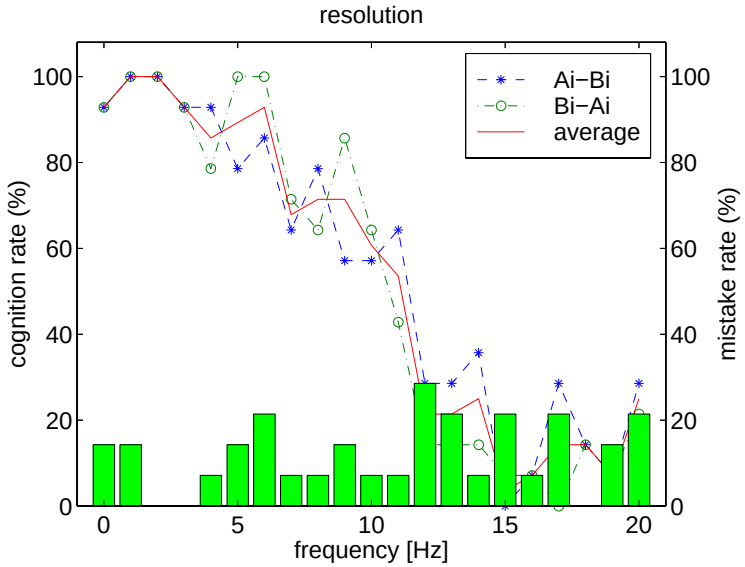


図 3.31: 結果 3-b：$f_0 = 125Hz$, 変調周波数 f_{s2}の差が $1Hz$の音声ペア

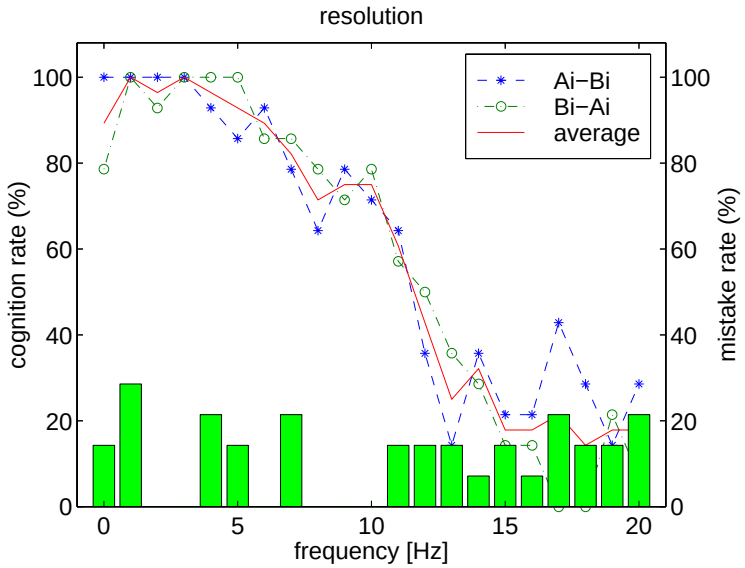


図 3.32: 結果 3-b：$f_0 = 125Hz$, 変調周波数 f_{s2}の差が $5Hz$の音声ペア

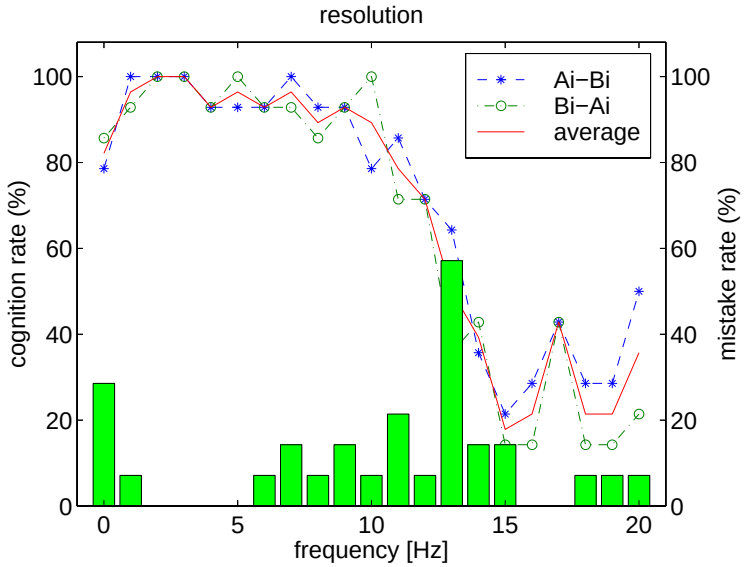


図 3.33: 結果 3-b：$f_0 = 125Hz$, 変調周波数 f_{s2}の差が $10Hz$の音声ペア

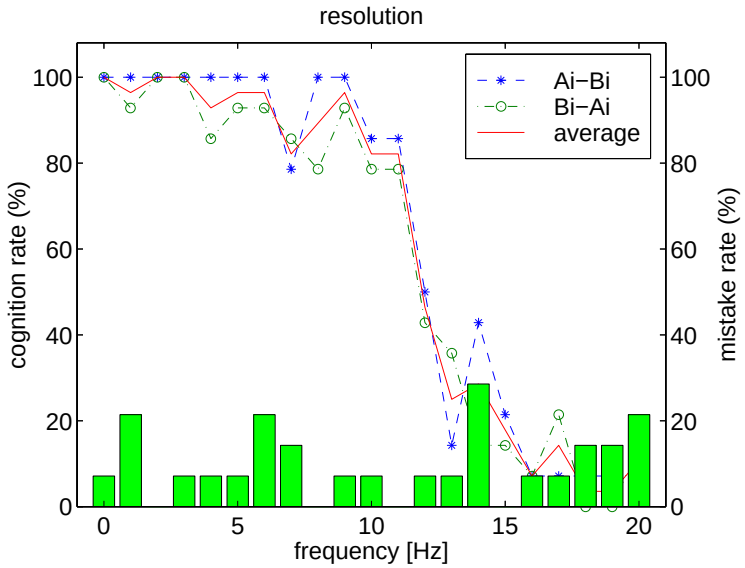


図 3.34: 結果 3-b：$f_0 = 250Hz$, 変調周波数 f_{s2}の差が $1Hz$の音声ペア

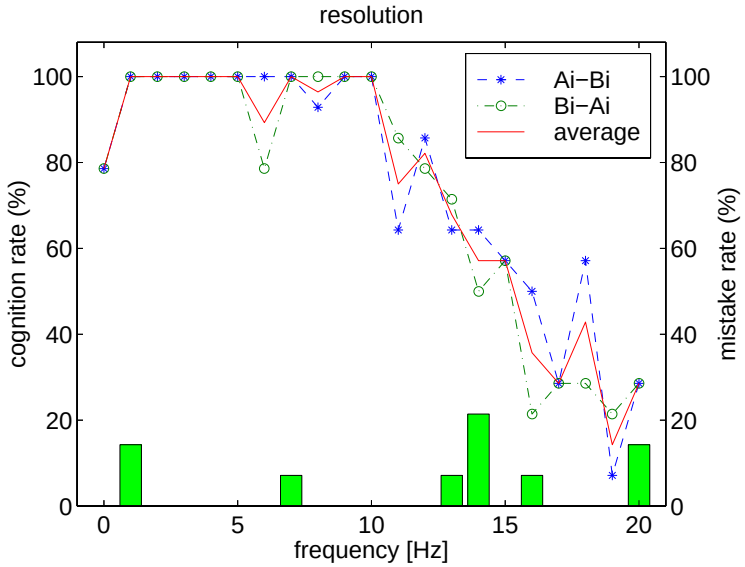


図 3.35: 結果 3-b：$f_0 = 250Hz$, 変調周波数 f_{s2}の差が $5Hz$の音声ペア

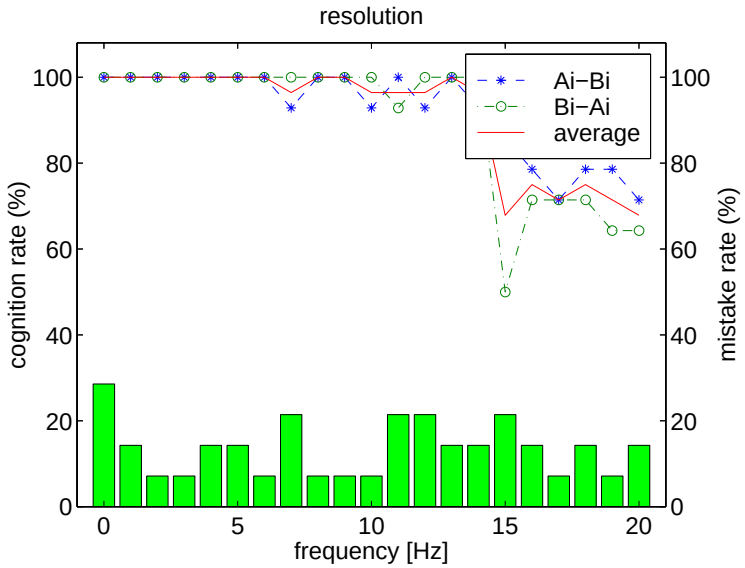


図 3.36: 結果 3-b：$f_0 = 250Hz$, 変調周波数 f_{s2}の差が $10Hz$の音声ペア

3.5.5 実験 3 の考察

歌声の規則的変動成分 (ビブラート) などの要因によって微細変動成分の検知限が変動する可能性がある。ここで検知限が大きく変動してしまい、認知が非常に難しくなれば、歌声の微細変動成分の知覚に影響を及ぼすことになる。3.5節で得られた結果をもとに上述の問題点を考察し、合わせて仮説の検証を行なう。

実験 3 から得られた結果から次のような考察ができる。

- 実験 $3-a$ の結果から基本周波数 $125Hz$での検知限は $20Hz$であり、基本周波数 $250Hz$では検知限が $60Hz$であるという結果がえられた。ここで、ここで 3.3節の実験 $1-a, 1-b$ の変調周波数が存在する場合と存在しない場合の実験結果と比較してみると、検知限は共に同じであるという結果が得られた。したがって規則的変動成分 (ビブラート) を考慮にいれて調周波数 f_{s1}を $6Hz$一定としたにもかかわらず、規則的変動成分の影響は受けることなく微細な変動成分を知覚することができたことになる。この考察から規則的変動成分の影響は受けないという可能性があげられる。

- 実験 $3-b$ の結果から考察を行なう。基本周波数が $125Hz, 250Hz$の両者において変調周波数 f_{s2}の音声ペアの差が大きくなれば、それに伴い検知限が上昇する結果が得られた。この結果は、3.4節実験 2 表 3.3の結果と同じである。検知限の値からみると実験 3.4と全く同じ結果は得られなかった。ここに実験 2 と実験 3 の提示音の違いが現れていると考えられる。しかし非常に類似した検知限となることがいえる。

これらの結果から仮説の検証を行なう。上述の考察から分かるように、規則的変動成分 (ビブラート) による微細変動の知覚への影響はほとんど見られなかった。この結果から歌声の微細変動成分には規則的変動成分の影響を受けることなく知覚可能であるといえる。ただし、今回の実験では式 3.1の偏移幅α_1, α_2を共に 1% としている。実際の歌声ではα_1の偏移幅がα_2の偏移幅よりも大きいと予想される。したがって偏移幅$\alpha_1 > \alpha_2$の条件を満たす合成音声を作成し、同様の実験を行なうことで、規則的変動成分が微細変動成分に与える影響について考察を行なう必要があると思われる。

第 4 章

結論

4.1 本論文で明らかになったことの要約

本研究では歌声に含まれる基本周波数の微細変動成分 (不規則的変動成分) が知覚に与える影響を検討した。そのために基本周波数抽出アルゴリズム (TEMPO2) の精度評価を合成音声を用いて行ない、本研究に利用するための使用可能性を検討した。これにより、基本周波数 f_0に対して $f_0/5$ 以下の変調周波数の解析が可能であることが確認できた。また基本周波数から微細変動成分 (不規則的変動成分) の抽出方法を提案し、実際に採取した歌声に対して微細変動成分の分析を変調周波数と偏移幅の両面から行なった。採取した歌声の分析より変調周波数では音程安定時には $15Hz$程度、音程変化時には $20Hz$程度の変調が確認できた。偏移幅に関しては、音程安定時に基本周波数に対して約 $20cent$、音程変化時には最大 $60cent$ の偏移幅を持つことが確認できた。これらから微細変動成分が知覚に与える影響を持つだろうという背景と、今回得られた分析結果をもとに「歌声に含まれる基本周波数の微細変動成分は、歌声の知覚に影響を与え、影響の大きさは基本周波数の変調周波数と偏移幅に依存する。」という仮説を立てて、合成音声を作成し仮説の検証を順時行なった。まず、基本周波数の微細変動成分が知覚可能であるか合成歌声を用いて確認した。この結果、非常に高い認知率で微細変動成分を知覚していることが明らかとなった。したがって以後の仮説の検証に意味を持つことがいえた。仮説の検証においては、

- 変調の有無に対する実験
- 変調が共に存在する場合に対する実験
- 変調が２つ存在する場合の実験

の 3 段階の実験を重ねることで行なった。まず「変調の有無に対する実験」において、歌声分析で得られた変調周波数 $15Hz$(音程安定時) と偏移幅 $20cent$(平均)、基本周波数 $125Hz - 250Hz$、という条件の下では共に知覚可能である揺らぎであり、仮説は正しいことを示した。次の「変調が共に存在する場合に対する実験」では、実際の歌声知覚において変調の有無による距離の違いだけではなく、変調の変化成分も知覚に影響していることを結果から考察した。これにより、譜面にあわせて音程変化を行なう歌声の場合は、より微細な変動を知覚し易いことを示した。最後の「変調が２つ存在する場合の実験」では規則的変動成分が微細変動成分の知覚に与える影響があるかを検討した。この結果、規則的変動成分が知覚にそれほど関与していないことを明らかにした。

4.2 今後の課題

以下に今後の課題を列挙する。

1. 歌声の分析手法の提案
   今回は微細変動成分の分析に対して変調周波数と偏移幅に着目して行なったが、他にも譜面に記されている基本周波数からの逸脱距離などの微細な変動を分析するためのパラメータを用いて分析を重ねる必要がある。

2. 実験 $3-1$ の課題
   今回の実験では式 3.1の偏移幅α_1, α_2を共に 1%としている。実際の歌声ではα_1の偏移幅がα_2の偏移幅よりも大きいと予想される。したがって偏移幅$\alpha_1 > \alpha_2$の条件を満たす合成音声を作成し、同様の実験を行なうことで、規則的変動成分が微細変動成分に与える影響についてさらなる考察を行なう必要があると思われる。

謝辞

常日頃から数多くの有益な御助言、御指導を頂きました赤木正人教授、ならびに赤木研究室の皆様に深く感謝致します。

また本研究を進めるにあたって、歌声を採取させていただいた YAMAHA PMS ボーカル科の政氏志津代さん、寺野波瑠奈さん、島崎健一さん、岩村美知宏さん、声優養成所の田中由実子さんに心から感謝致します。

また本研究を進めるにあたって、聴取実験に協力していただいた赤木研究室の皆様、並びに丸山晃生さん（小野研究室）、高瀬泰宏さん（落水研究室）に厚くお礼申し上げます。

最後に、3 年間の研究生活を支えて下さった家族、友人、私と関わりのあった方皆様に感謝致します。

参考文献

[1] Johan Sundberg, The Acoustics of the Singing Voice', Sci. Am 236 , 82-91 (1977))

[2] Ichiro Nakayama, Masuzo Yanagida, Yuki Kakita, Tsutomu Uehara, 'Intelligibility of Japanese sung vowels', J.Acoust. Soc Jpn.(E) 14,6 (1993).

[3] 矢田部 学, 遠藤 康男, 粕谷 英樹, 神戸 孝夫, ' 歌声の基本周波数の動特性', 音響論集,3-8-6, (1998).

[4] 小田切 わか菜, 粕谷 英樹, ' 歌声のビブラートの分析・合成・知覚に関する検討', 音響論集,1-4-6, (1999).

[5] 榊原 健一, 小坂 直敏, ' 信号モデルによるビブラートの制御について', 音響論集,1-7-5, (1999).

[6] 河原 英樹, 片寄 晴弘, Roy D.Patterson, Alain de Cheveigné, ' 瞬時周波数を用いた基本周波数の高精度の抽出について', 日本音響学会 聴覚研究会資料,MA-98-42, (1998).

[7] 河原 英樹, 片寄 晴弘, Roy D.Patterson, Alain de Cheveigné, ' 周波数-瞬時周波数の安定平衡点の性質に基づく音源情報の抽出について', 音響論集,1-2-4, (1999).

[8] Hideki Kawahara, Parham Zolfaghari, Alain de Cheveigné, Roy D.Patterson, 'Source Informatiuon Extraction using Fixed Points of a Frequency to Instantaneous Frequency Map', The Acoustical Society of Japan H-99-46, July 8 ,(1999).

[9] Hideki Kawahara, Haruhiro Katayose, Alain de Cheveigné, Roy D.Patterson, 'FIXED POINT ANALYSIS OF FREQUENCY TO INSTANTANEOUS FREQUENCY MAPPING FOR ACCURATE ESTIMATION OF F_0 AND PERIODICITY', EUROSPEECH-99, September 5-9, Budapest,Hungary,(1999).

[10] 中山一郎, 小林範子, ' 歌の声', 日本音響学会誌 52 巻 5 号, pp.383-388,(1996).

[11] 鈴木久喜, ’ピッチ抽出の今昔’, 日本音響学会誌 56 巻 2 号, pp.121-128,(2000).

学会発表リスト

北風、赤木:' 基本周波数の微細変動成分に対する知覚', 音声研究会 (1999.3 発表)