

博士論文
音声と心理との関連についての
音声物理量の解析

論文指導教官
田村直良 教授
指導教官
影井清一郎 教授

横浜国立大学大学院 工学研究科
人工環境システム学専攻
00JE008 鈴木 朋子

2004年3月

横浜国立大学附属図書館



11472474

概要

現代社会では、感情のコミュニケーションが難しいと言われている。コミュニケーションとは、送り手の意識の状態（目標、意図、態度、感情、知識、欲求等）をある種の信号を介して受け手に伝達する過程であり、コミュニケーションの問題とは、送り手が送信したメッセージを受け手が正確に受信できない状況である。

このようなコミュニケーションの問題の背景には、2つの原因が考えられる。1つの原因は、送り手から受け手へと伝わる信号の量の問題であり、受け手は多量に累積した信号のうち、どこの箇所の信号を重要な信号として受け取れば良いのか判断に誤りが生じることである。もう一方の原因は、信号の種類の問題であり、目標、態度、感情等の多くの種類の信号が同時に存在するため、受け手はどの種の信号を受け取れば良いのか判断に誤りが生じることである。2つの原因によるコミュニケーションの誤りを防ぎ、正確なメッセージを送り手から受け手へ伝えるためには、重要な部分に焦点を当てて信号を扱うことが必要である。

そこで本研究では、正確な感情のコミュニケーションの促進を目標として、重要な箇所の信号と、感情に関する信号について重点的に扱うことで音声と心理状態（感情）の関連を検討し問題を解決することを目的とした。

第2章「音声に関する基本事項及び音声と心理状態の関連研究について」では、人間が音声を作り出す仕組みと、計算機での音声の処理の方法を概説し、音声と心理に関する関連研究を紹介した上で本研究における感情の扱いについて述べた。

第3章「感情に関する研究」では、心理学分野を中心とした感情に関する研究を概観した。本研究で用いる感情の定義として、独立のカテゴリにより構成される基本感情の立場から研究を行うことを述べた。

第4章「感情を表現した音声の分析」では、感情が表れた音声の分析と聴取実験を行った。声優が発声した感情音声と、複数の一般者が発声した感情音声から

音声物理量を抽出し、一文での代表値の統計的分析と図から目視で得た変化の分類を行うことで、感情と音声物理量との関連を検討した。感情を表現する音声物理量について感情表現の訓練を受けた人と訓練を受けたことのない一般の人に共通して表れる特徴があること、重要とされるパラメータが文全体で一つの代表値に注目する場合と文の中のダイナミックな変化に注目する場合とで異なることが分かった。感情音声の聴取実験からは、発声者の表現と聴取者の認知における感情の不一致が認められ、その背景として、発声者の個人性、聴取者の個人性、文脈の問題が考えられた。

第5章「音声物理量からの抑うつ傾向判定」では、感情の一種である抑うつに焦点を当て、抑うつ傾向判定システム構築を目標に抑うつ傾向と文読み上げの音声における音声物理量の関連を検討した。文読み上げの課題文として用いたBDI (Beck Depression Inventory) のうち6文は、音声物理量と抑うつ傾向の関連が見出された。抑うつが表れやすい箇所としての指標単語の概念を提案し、その選択法について検討を行った。3つの指標単語選択戦略にて選択した指標単語について、ロジスティック重回帰分析にて検討を行った。実験と分析の結果、統計指標単語選択戦略が他の2つの指標単語選択戦略よりも判定に対して有効な指標単語であることが分かった。

第6章「感情豊かな読み上げのための音声パラメータ」では、感情豊かな読み上げ音声の合成を目標として、感情が表れている単語としての指標単語に重点を置き、有効な音声物理量の検討を行った。音声物理量を操作したモデル音声を作成し音声聴取実験を行った。感情が表れていると選定されたモデル音声について、音声物理量を抽出し、基本周波数の変化を表現する方法を3点から検討した。得られた音声物理量を「かちかち山」物語読み上げの会話文の指標単語に付与し、聴取実験を行うことで有効性を検討した。結果、合成音声の指標単語の音声物理量を修正すると、音声から感じる落ち着いた印象が弱まり、文の意味よりも複雑な感情を認知させることが分かった。

以上の結果を踏まえて、第7章「結論」では、コミュニケーションの問題に対して、累積する信号の中で重要な箇所の信号、多種の信号の中で感情に関する信号を重点的に扱うことで問題の解決を目標とした本研究について全体的な討論を行った。

第一に、累積する信号の中で重要な箇所を扱う手法として提案した、指標単語について検討を行った。第5章においては抑うつが強調される箇所を指標単語として、指標単語の音声物理量を重要なパラメータとしての判定器の試作と評価を行った。第6章においては感情が強調される箇所を指標単語とし、指標単語に特化した感情情報の付与を行い合成音声にて評価を行った。共に指標単語により音声物理量と感情の関連が確認されたことから、多量の信号の中で箇所を限定する方法は、感情のコミュニケーションにとって有益であると示唆された。

第二に、多種の信号の中で焦点を当てた感情については音声物理量との関係について3つのアプローチから検討を行った。

意図的に感情を表現した音声を用いた第一のアプローチ（第4章）では、調査で選定された6感情と音声物理量の関連を検討した。感情表現の訓練の有無にかかわらず感情を表現した音声に共通の特徴が認められた。文全体の代表値については、振幅と発話時間が感情を表現する上で重要な音声物理量パラメータと考えられた。また、音声物理量の文の中の変化については、基本周波数と振幅が感情を表現する上で重要な音声物理量パラメータと考えられた。

自然な状態での発声者の感情と音声物理量の関連を検討した第二のアプローチ（第5章）では、特定の感情である抑うつを対象として音声物理量との関連を検討した。ここでは、高抑うつ傾向の者の発声した音声において、不安や悲しみ等の影響が音声物理量の特徴として表れることが分かった。

音声から聴取者が認知した感情と、音声物理量との関連を検討した第三のアプローチ（第6章）では、調査で選定された3つの感情と音声物理量の関連を検討した。感情と関連が認められたモデル刺激の音声物理量では一定の特徴が見られ、特に「悲しさ」に関しては、合成音声にモデル刺激より得られた音声物理量を付与した聴取実験においてもこの関連の情報の有効性が確認された。

以上の3つのアプローチによる相違を検討すると、怒りに関係した音声は、発声者が意図的に怒りを表現した音声と聴取者が音声より認知した感情とで一致が見られ、また、発声者が意図的に怒りを表現した音声と聴取者が怒りを表現していると選択した音声とで、音声物理量に共通点が見られた。すなわち、声が大きく発話時間が短い音声は怒りの音声と考えられた。

楽しみに関係した音声は、発声者が意図的に表現した音声と聴取者が音声より

認知した感情とでは一致が見られなかったが、発声者が意図的に楽しみを表現した音声と聴取者が喜びを表現していると選択した音声とでは、音声物理量に一部共通点が見られた。すなわち、発話時間が長いという点のみ楽しみの音声では一致していた。

悲しみに関係した音声は、発声者が意図的に表現した音声と聴取者が音声より認知した感情とでは一部一致が見られなかったが、発声者が意図的に表現した音声の音声物理量と聴取者が悲しみを表現していると選択した音声の音声物理量においては声が小さく発話時間が短いという差が見られた。しかし、悲しみと関連のある抑うつが自然に表現されていると考えられる音声の音声物理量は異なっており、声が高く声が大きいという特徴が認められた。すなわち、悲しみの音声では、発声者の持つ感情の音声への表れや意図的な表現、受け手の認知が異なっていた。

本研究における3つのアプローチからは、感情の種類によって発声者が意図的に表現した音声で伝わりやすいものと、伝わりにくいものが存在することが分かった。また、自然な状態での発声者の音声に表れた感情と、受け手が認知した感情とは異なることが考えられた。

以上の二点からの検討により、感情のコミュニケーションにおける問題については、指標単語に特化した音声物理量の分析と合成が有用であることが考えられた。また、感情の種類によっては、伝わりやすいものと伝わりにくいものがあり、音声物理量を用いて送り手の感情を分析する聴取の役割と、受け手が認知しやすい音声物理量を用いて合成を行う発声の役割とを計算機が担うことが有効であることが示唆された。

指標単語に関して言語学的な知識を用いて計算機によって自動的に選択をすること、感情の構造を独立したカテゴリではなく互いに関連しあったものと捉える見地からの研究が今後の課題とされる。

目次

1 序論	1
1.1 研究の背景	1
1.2 コミュニケーションの問題の整理	1
1.3 研究の目的	3
1.4 論文の構成	4
2 音声に関する基本事項及び音声と心理状態の関連研究について	6
2.1 生理学における音声	7
2.2 音声学・言語学における音声	8
2.2.1 音声の単位	8
2.2.2 音声の高さの変化のまとめ	9
2.3 工学における音声	10
2.3.1 音声波形のデジタル化	10
2.3.2 音声分析	11
2.3.3 基本周波数の抽出	13
2.3.4 音声合成	14
2.4 音声と心理状態の研究	15
2.4.1 音声からの心理状態の認知研究	15
2.4.2 音声に表れた心理状態の音声分析研究	16
2.4.3 心理状態が関連した音声合成研究	19
2.5 本研究での音声の扱いについて	20
2.5.1 本研究で用いた音声物理量	20
2.5.2 音声研究のアプローチ	21

3	感情に関する研究	22
3.1	感情とは	22
3.1.1	感情研究の4つの立場	23
3.1.2	基本感情に関する研究	24
3.2	本研究の感情への視点	27
4	感情を表現した音声の分析	29
4.1	感情と音声の関連検討のための研究の概略	29
4.2	調査	30
4.2.1	目的	30
4.2.2	方法	30
4.2.3	結果	31
4.3	声優の感情音声の音声物理量抽出	31
4.3.1	目的	31
4.3.2	方法	31
4.3.3	音声物理量の抽出	35
4.3.4	結果	35
4.3.5	考察	35
4.4	感情音声の聴取実験	36
4.4.1	目的	36
4.4.2	方法	36
4.4.3	結果	36
4.4.4	考察	42
4.5	一般者の感情音声の採取と統計的分析	43
4.5.1	目的	43
4.5.2	方法	44
4.5.3	結果	45
4.5.4	考察	46
4.6	考察	50
4.7	まとめ	52

5	音声物理量からの抑うつ傾向判定	53
5.1	背景と目的	53
5.2	抑うつ傾向判定のための研究の概略	55
5.3	BDI を用いた音声採取実験被験者, 課題文の選択	56
5.3.1	BDI について	56
5.3.2	調査の実施と被験者群の選定	57
5.3.3	課題文の選択	57
5.4	音声データの採取実験と分析	58
5.4.1	音声採取実験	58
5.4.2	音声物理量のパラメータについて	59
5.4.3	パラメータの差の検定と結果	59
5.4.4	重回帰分析による判別に使用する音声パラメータの選定	60
5.4.5	結果と考察	62
5.5	測定範囲の細分化	63
5.5.1	文節レベルでの分析の結果	63
5.6	指標単語選択戦略	64
5.6.1	統計による指標単語選択戦略	65
5.6.2	聴取による指標単語選択戦略	65
5.6.3	専門家による指標単語選択戦略	66
5.6.4	比較のためのランダム選択	66
5.7	それぞれの戦略による分析結果	66
5.7.1	統計による指標単語選択戦略による実験結果	66
5.7.2	聴取による指標単語選択戦略による実験結果	67
5.7.3	専門家による指標単語選択戦略での実験結果	68
5.7.4	比較のためのランダム選択での実験結果	68
5.7.5	指標単語選択戦略の結果について	69
5.8	結果と考察	70
5.8.1	指標単語について	70
5.8.2	抑うつ判定について	71
5.9	まとめ	72

6	感情豊かな読み上げのための音声パラメータ	74
6.1	背景と目的	74
6.2	音声パラメータ検討のための研究の概略	76
6.3	基本感情の選定及び感情尺度作り	76
6.3.1	方法	76
6.3.2	結果	77
6.4	モデル刺激の作成	79
6.4.1	音声採取実験	80
6.4.2	刺激音声の作成	81
6.5	モデル刺激を用いた音声聴取実験	82
6.5.1	方法	82
6.5.2	結果	83
6.6	モデル刺激からの音声物理量の抽出	84
6.6.1	指標単語調査	85
6.6.2	感情を付与した物語読み上げ	85
6.7	結果と考察	88
6.7.1	指標単語について	88
6.7.2	文から感じられる感情と感情因子の関係	88
6.7.3	感情と指標単語の位置の関係	89
6.8	まとめ	90
7	結論	91
7.1	研究のまとめ	91
7.2	指標単語について	92
7.3	音声物理量と感情の関連について	93
7.4	今後の課題	95
	謝辞	96

図一覧

1.1	Shannon&Weaver による通信（コミュニケーション）のモデル . . .	2
1.2	受信者の信号の誤選択	2
2.1	発声器官	8
4.1	感情状態尺度	34
4.2	「怒大」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	37
4.3	「怒小」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	37
4.4	「喜大」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	38
4.5	「無」の感情が表れた音声を刺激とした聴取実験の感情尺度	38
4.6	「安大」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	39
4.7	「安小」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	39
4.8	「鬱大」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	40
4.9	「鬱小」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	40
4.10	「驚大」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	41
4.11	「怖大」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	41
4.12	「怖小」の感情が表れた音声を刺激とした聴取実験の感情尺度 . . .	42
4.13	Wilcoxon の符合つき順位検定結果一覧	43
4.14	感情状態尺度因子ごとの得点平均（振幅）	45
4.15	感情状態尺度因子ごとの得点平均（sec）	45
5.1	ロジスティック重回帰分析にて選択された指標単語	69
6.1	文 1(怒り) の感情尺度平均得点	86
6.2	文 2(楽しさ) の感情尺度平均得点	87
6.3	文 3(悲しさ) の感情尺度平均得点	87

表一覽

2.1	藤崎による音声情報の分類	6
2.2	Davitz による情緒表現に結び付いた音声の手がかり	15
2.3	Scherer による感情と音声物理量のパラメータの対応表	18
4.1	感情状態尺度の因子分析結果	32
4.2	感情状態尺度 6 因子の因子相関行列	33
4.3	声優の感情音声の音声物理量	35
4.4	基本周波数の変化の分類表	47
4.5	振幅の変化の分類表	48
4.6	発話時間の変化の分類表	49
5.1	音声採取実験で採取した課題文	58
6.1	基本感情調査の因子分析結果 (promax 回転によるパターン行列)	78
6.2	基本感情調査 4 因子の因子相関行列	78
6.3	音声から感じられる感情調査結果	79
6.4	実験 A で用いた発声者への課題の指示	80
6.5	選定されたモデル刺激	83
6.6	「かちかち山」における指標単語と感情	85

第 1 章

序論

1.1 研究の背景

情報化社会の到来と共に、我々は必要な情報を迅速かつ簡便に入手できるようになった。しかし、情報技術の発展は快適な生活を提供すると同時に、感情のコミュニケーションにおける問題をもたらしたと言われている [40, 61]。

感情のコミュニケーションでの問題背景には、機械と人間とのコミュニケーションの問題が存在する。機械とのコミュニケーションでは、人間を相手にした場合のように、暖かく感情豊かなやりとりは不可能である。機械の無機質な印象に適応できず、ストレスを感じる人は多い。反対に、情報化社会への過適応のために、人間相手のコミュニケーションにおいて感情豊かな生身の交流が困難となる場合もある [56]。このように現代社会では感情のコミュニケーションにおいて問題が生じている。

コミュニケーションにおける感情の重要性は、精神医学や心理学の分野で繰り返し述べられてきた [47, 49, 44, 13]。コミュニケーションを円滑に行うため、人間の感情について様々な側面を解明し、機械での応用を考えることが必要である。

1.2 コミュニケーションの問題の整理

コミュニケーションとは、送り手の意識の状態（目標、意図、態度、感情、知識、欲求等）をある種の信号を介して受け手に伝達する過程である [24]。

Shannon&Weaver [75] は、電子通信理論を応用してコミュニケーション（通信）の過程を図 1.1 で図示している。このモデルでは、情報源（送り手）から送信機を通じてメッセージが発信される。送信信号は、通信路を通り受信信号となって受信機で受け取られ、メッセージとして受信地（受け手）へと伝わる。

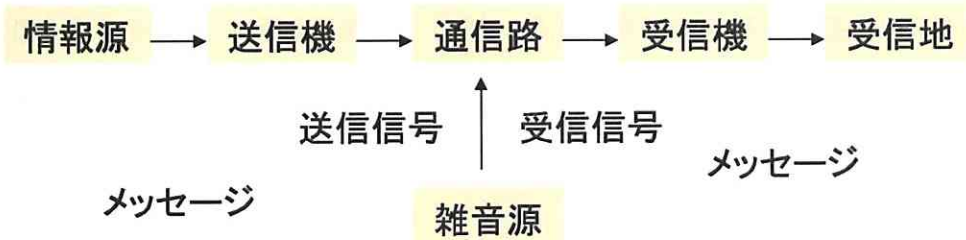


図 1.1: Shannon&Weaver による通信（コミュニケーション）のモデル

Shannon&Weaver のモデルから考えると、コミュニケーションの問題とは、送り手が送信したメッセージを受け手が正確に受信できない状況と定義することができる。このようなコミュニケーションの問題の背景には、2つの原因が考えられる（図 1.2）。

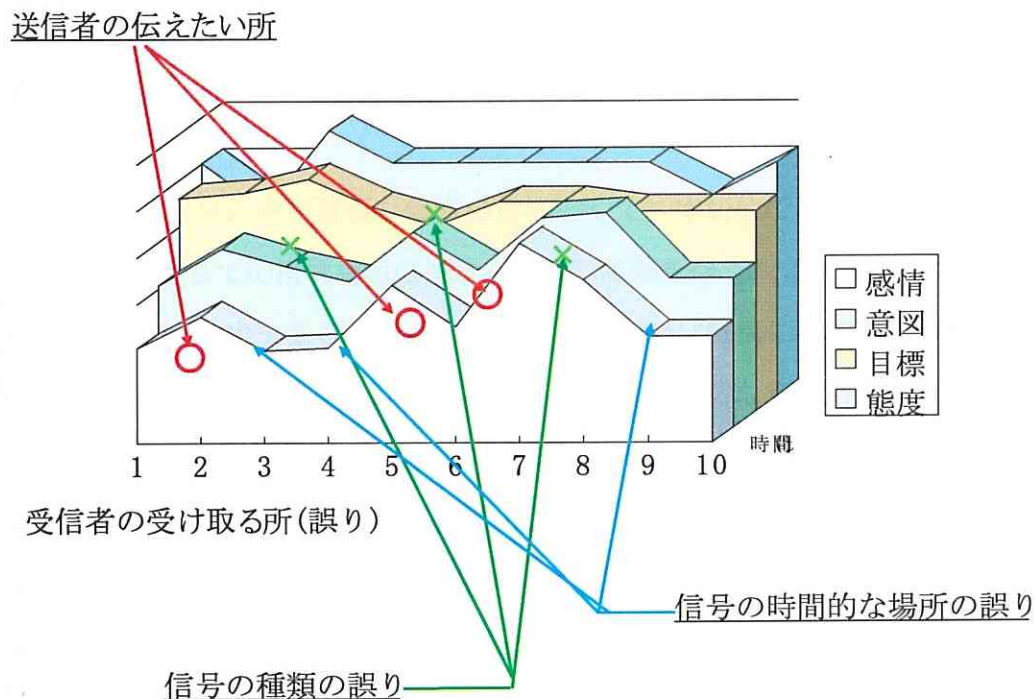


図 1.2: 受信者の信号の誤選択

1つの原因は、送り手から受け手へと伝わる信号の量の問題である。受け手が受信する信号は、コミュニケーションに要する時間が長くなるにつれて累積する。受け手は多量に累積した信号のうち、どこの箇所の信号を重要な信号として受け取れば良いのか判断に誤りが生じると考えられる。

もう一方の原因は、信号の種類の問題である。コミュニケーションの信号として送信される送り手の意識状態は、目標、態度、感情等の多くの種類が同時に存在する [24]。さらに、感情においても「喜」、「怒」、「哀」、「楽」のように質的に異なった種の信号が存在する。受け手はどの種の信号を受け取れば良いのか判断に誤りが生じると考えられる。

信号の量、信号の種類の問題の2側面の問題へ対応することが、コミュニケーションの問題解決への一つの方法となると考えられる。

1.3 研究の目的

コミュニケーションの問題として整理した原因の中で、コミュニケーションの誤りを防ぎ、正確なメッセージを送り手から受け手へ伝えるためには、全ての信号を一様に扱うのではなく、重要な部分に焦点を当てて扱うことが必要と言える。

そこで本研究では、累積する信号の中で重要な箇所の信号、多種の信号の中で重要な信号を重点的に扱うことで、より正確なコミュニケーションを促進することを目標とする。

累積する信号の中で重要な箇所の信号として、本研究では文の中で感情が強調される箇所としての指標単語の概念を提案し評価を行う。

多種の信号の中で重要な信号として、本研究では感情に関する信号を扱う。その際、人間のコミュニケーション手段として最も自然な形である音声言語を対象とする。音声言語には言葉の意味の持つ側面と、言葉の音を持つ側面があり、両面が信号として送り手から受け手へと伝わっているが、ここでは感情との関係が取り上げられてきた言葉の音の側面としての音声物理量に重点を置いて、音声と心理状態（感情）との関連を検討する。

なお、音声物理量と感情との関連については、具体的には次の3つのアプローチから検討を行う。

第一のアプローチ（第4章）は、意図的に感情を表現した音声における音声物理量と感情の検討である。計算機は、意図的に感情を表現した音声に対し聴取側の役割を取る事となる。本アプローチでは、音声による感情表現の訓練を受けた声優と、感情表現の訓練を受けたことのない一般の人による感情を表現した音声を分析し、音声物理量と特徴について検討する。また、音声聴取実験によって、発声者が表現した感情と、聴取者が認知した感情との関係を検討する。

第二のアプローチ（第5章）では、意図的な感情表現ではなく、自然な状態での発声者の感情と音声物理量の関連を検討する。計算機は、自然に感情が表れた音声に対し聴取・判断の役割を持つ事となる。ここでは音声に表れる感情として抑うつ傾向に焦点を当て、音声物理量との関連を検討する。また、文の中でも抑うつ傾向が表れやすい箇所を指標単語と定義し、抑うつ傾向判定のための指標単語を検討し、音声物理量と指標単語からの抑うつ判定器の作成を試みる。

第三のアプローチ（第6章）では、音声から聴取者が認知した感情と、音声物理量との関連を検討する。計算機は、感情を表現した音声の発声・合成の役割を持つ事となる。ここでは音声物理量を操作した音声に対して、聴取実験で感情ラベルを貼り、音声物理量と感情との関連を検討する。さらに、得られた音声物理量と感情との関連の情報は、感情が強調される箇所としての指標単語に特化して合成音声の音声物理量に修正を加えることで有効性を検討する。

1.4 論文の構成

第2章「音声に関する基本事項及び音声と心理状態の関連研究について」では、人間が音声を作り出す仕組みと、計算機での音声の処理の方法を概説し、音声と心理に関する関連研究を紹介する。

第3章「感情に関する研究」では、心理学分野を中心とした感情に関する研究を概観し、本研究で用いる感情の定義と、感情への視点を明らかにする。

第4章「感情を表現した音声の分析」では、声優と、複数の一般人が感情を表現して発声した音声について詳しい分析と考察を行う。感情の音声物理量への表れと、音声からの感情の認知について分析を行い、心理と音声の関連について議論する。

第5章「音声物理量からの抑うつ傾向判定」では、抑うつ傾向と音声物理量の関連について述べる。抑うつ傾向の高い人と低い人を対象に抑うつ文の読み上げ音声を採取し、抑うつ傾向の判定に重要な音声物理量パラメータについて検討する。さらに、抑うつ傾向が表れやすい部分である指標単語について議論する。

第6章「感情豊かな読み上げのための音声パラメータ」では、物語読み上げの合成音声に感情を付与することを目的として、感情と音声物理量の関連を議論する。聴取実験を行い、感情が表れていると判断された音声について音声物理量を抽出し、合成音声に付与する。合成音声の全体ではなく、部分について感情の付与を行うことで、感情を強調する単語（指標単語）に特化した感情豊かな音声を提案する。

第7章「結論」では、上記の章により検討した音声と心理の関連についての総合的考察を行う。

第 2 章

音声に関する基本事項及び音声と心理状態の関連研究について

本章では、音声に関する基本事項の紹介と、音声と心理状態の関連研究を概観する。

音声とは、人間がコミュニケーションのために人間の音声器官を使って発する音のことである [79]。音声により表現され、伝達される情報は、言語的情報、パラ言語的情報、非言語的情報の 3 種類に大別される [23] (表 2.1)。

表 2.1: 藤崎による音声情報の分類

	意図的表出	例
言語的情報	可能	音韻・統語・談話情報
パラ言語的情報	可能	意図・心的態度・強調
非言語的情報	不可能	性別・身体状態・情動

言語的情報とは辞書、統語、意味、談話のレベルで、文字言語に直接含まれるものである。パラ言語的情報は、直接的に文字表現に含まれないが音声言語の韻律的特徴によって話者の発話意図や感情状態などが表現されるもので、話者が意識的に、あるいは無意識的にその表現を制御できる。非言語的情報は、発話内容に関係せず、話者が意識的に制御し得ない個人的特徴や身体状態などの情報である。このように音声には、複雑な情報が含まれていることから人間のコミュニケー

ション手段としても特に重要なものと考えられ [25], 工学, 心理学, 生理学, 言語学といった様々な学問分野において研究が行なわれてきた. 本章では, 音声に関する基本事項を確認するため, 音声の生理学的側面, 音声学の側面からの説明を行った上で, 本研究の領域となる音声と心理状態の研究について概観し, 本研究での音声の扱いについて述べる.

2.1 生理学における音声

生理学では, 人間がどのように音声を作り出すかを扱っている. 音声を送り出す仕組み (図 2.1) は, 音源を発生する部分 (音源生成) と, 共鳴を発生する部分 (調音) から構成されている [26, 54].

音源生成

音声のエネルギー源は, 肺から押し出される空気 (呼気) である. 呼気は, 肺から気管支を通り喉頭, 咽頭, 口あるいは鼻という経路を経て体外へ出る. この際, 喉頭の声門の位置にある声帯が振動して, 音声の音源となる音波を発生する働きをしている.

調音

さまざまな言語音を作り出すために声道の形を調整することを調音と言う. 声道とは, 声門から口または鼻孔までの部分であり, 一種の音響管に例えられる. 舌, 唇, 歯, 顎は声道の形を決定するため調音器官と呼ばれ, これらの器官の位置により共振周波数が決定される.

音声波形の圧力変化の様子は, パルス的な三角波で近似することができる. この波形は, 正弦波に分解すると, 基本周期の逆数に対応する基本周波数 (fundamental frequency: F_0) の正弦波と, その整数倍の周波数の正弦波の組み合わせで表現できる.

音声における音声波形の物理的特徴と, 知覚 (聴覚) との関連からは次のように考えられている. 基本周波数の変化は, 聴覚的には声の高さ (pitch) の感覚を生じる. また, 圧力の変化は音声波形の強さ (intensity) に対応し, 声の大きさ

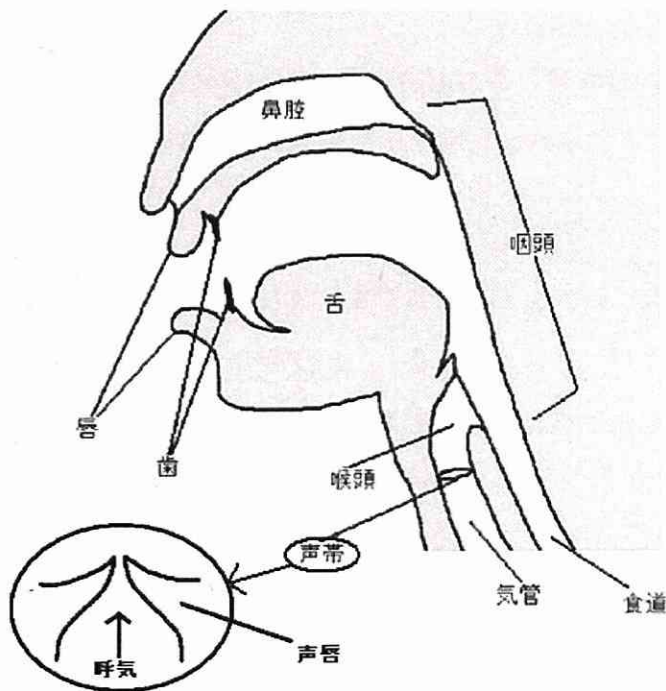


図 2.1: 発声器官

(loudness) の感覚を引き起こす。声帯の振動は、声門の開閉の開始、停止や、繰り返し回数や、呼気の送出量や喉頭の緊張度を調節することで、発声者が意識的に制御することができるが、この制御により、音波の持続時間、基本周波数、強度が変化し、言葉の区切りや、発話の全体的な速さ、リズムや、単語のアクセントや文の抑揚、句の強調が伝達される。また、子音母音などの音の変化は調音によって生成される共振周波数 (formant) によって決定される。

2.2 音声学・言語学における音声

音声学や言語学において、音声は分割された単位として扱われる。本節ではその分類について述べる。

2.2.1 音声の単位

音声で用いられる単位は、小さいものから音素、音節、モーラがある。

音素

音声の最小単位は音素または単音と呼ばれる。音素は、音源の種類、調音器官の形状により規定されるもので、母音と子音に分けられる。

- 母音

母音とは、呼気の流れが妨害されることなくスムーズに口から出る基本的な音のことを言う。日本語では、「あ」、「い」、「う」、「え」、「お」が母音にあたる。母音は、声帯が振動して声を伴うため、「有声音」とも呼ばれる。

- 子音

子音とは、声道の中で息が一時的に閉鎖されたり、狭めによる摩擦が生じたりしてつくられる音のことを言う。「k」、「p」、「t」のように母音を破裂によって区切る子音を「破裂音」、「s」のように呼気が通るときの摩擦によってできる子音を「摩擦音」と言う。また、「k」、「s」、「t」、「p」は「無声子音」と呼ばれ、「g」、「z」、「d」、「b」は「有声子音」と呼ばれる。

音節

音節とは、切目が感じられない音の連続のことをいう。母音だけ、あるいは母音が核となって前後どちらかもしくは両方に子音が従った形になることが多い。例えば「鳥」の場合は、to/ri の2つとなる。

モーラ

一定のリズムをなす音節のことをモーラ（拍）という。日本語では、仮名1文字が1モーラになり、他言語と異なり一対一で文字と対応する音声単位であるため、日本語にとってはモーラが自然な区切りと考えられることがある。

2.2.2 音声の高さの変化のまとめ

音声の高さの変化には、単語ごとの局所的な上がり下がりを示すアクセントと、句などのより長い区間での変化のイントネーションがある。

アクセント

それぞれの言語で語ごとに定まっているピッチやストレスに関する決まりをアクセントという。アクセントは、地方によって異なる。

イントネーション

イントネーションとは、話し言葉の中で生じる声の高さの時間的変化のことである。イントネーションは、アクセントの影響を受けるが、それ以外にも

- 意味上のまたは文法上のくぎりと連続の表示
- 話の焦点 (話者の意図, 新しい情報)
- 疑問, 反問, 勧誘等の区別

の働きが加わる。

これらの単位は言語学的に意味・文法の区別をもたらす上で重要である。音声の高さ (基本周波数) とその変化には、アクセント・イントネーションの双方の働きが重畳されている。

2.3 工学における音声

工学では、計算機を用いて音声を離散的なデータとして扱う。本節では、音声の音声波形のデジタル化について述べた後、音声分析と音声合成の方法について概説する。

2.3.1 音声波形のデジタル化

音声は、空気中では連続的な波形の形で存在する。計算機を用いて音声を分析する際には、連続した音声波形 (アナログ波形) をデジタル信号に変換する作業、AD変換が必要とされる。AD変換は、時間的に連続な波形を離散的な値の系列で表現する標本化 (sampling)、波形の値を有限個の値の中の一つで近似する量子化 (quantization)、具体的な表現 (通常は2進数表現) を行う符合化 (coding) によって行われる [26]。

2.3.2 音声分析

音声进行分析し音声を特徴づける物理的統計的性質を明らかにすることを音声分析という [52].

音声分析には、音声波の発声に関して特定のモデルを設定せずに行うフーリエ解析、帯域フィルタ分析、自己相関分析、ケプストラム分析と、音声発声過程のモデルを用いて計算を行う線形予測分析がある。ここでは、基本となるフーリエ解析について説明し、基本周波数の抽出法を述べる。

フーリエ分析

周期を持った波は、複雑な波でも単純な波の足しあわせとして表現できる。これを式で表したものがフーリエ級数であり式 (2.1) で表される。

$$F(t) = A_0 + \sum_{n=1}^{\infty} \{A_n \cos(n\omega t) + B_n \sin(n\omega t)\} \quad (2.1)$$

音声信号 $x(t)$ とそのフーリエ変換であるフーリエスペクトル $X(\omega)$ の関係は式 (2.2) で表される。

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad (2.2)$$

このようにして算出したフーリエ変換によって、周期のある波、周期のない波どちらにも、基本周期 T において、正弦波成分、および余弦波成分の含有量を解析することができる。

なお、計算機で信号処理を行う場合、サンプリング信号を有限の長さで打ち切った離散フーリエ変換 (DFT:discrete Fourier transform) として式 (2.3) を用いて計算を行う。

$$X(\omega) = \sum_{n=0}^{N-1} x(n)e^{-j\omega n} \quad (2.3)$$

N が大きい場合には膨大な計算量が必要になる。そのため計算手順を工夫することにより計算量を減少させたのが、高速フーリエ変換 (FFT; fast Fourier transform) と呼ばれるアルゴリズムである。

窓関数

離散フーリエ変換では、ある区間の音声信号を切り出した上で信号を連続した周期関数と見なし解析を行う。そのため、信号を切り出す際には、周期の開始時点及び終了時点を正確に切り出す必要があり、正確に切り出せない場合は切り出し部分において誤差を生じる。

一般的な信号においては、この誤差は必然的に生じる。そのため、信号の切り出し区間の両端に急激な変化が起こらないよう徐々に0にすることで誤差を最小限に抑える方法として窓関数が用いられる。

以下に、代表的な窓関数を説明する。

- 方形窓 (Rectangular Window)

入力信号を単純に切り出し使用する方法である。

$$W_R(n) = 1 \quad (2.4)$$

- ハニング窓 (Hanning Window)

$$W_N(n) = 0.5 - 0.5\cos\left(\frac{2\pi n}{N-1}\right) \quad (2.5)$$

$x(n)(0 \leq n \leq N) =$ 標本化された時系列

- ハミング窓 (Hamming Window)

$$W_H(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \quad (2.6)$$

$x(n)(0 \leq n \leq N) =$ 標本化された時系列

なお、音声分析ではハミング窓を用いることが多い。

2.3.3 基本周波数の抽出

声の高さに対応する基本周波数を抽出するには、音声波形の周期性を調べる必要がある。音声波形を生成する声帯の駆動は周期に例えられるが、完全な周期を持っているわけではない。また、音声波形には駆動音源と声道の影響が重なりあっているため、声道の影響を取り除いて声帯波のみを取り出すことは難しいと言われる。

このような議論の中で、基本周波数抽出方法が考案されてきた。以下に代表的なものを挙げる。

1. 波形処理

ピッチパルスに該当する波形の周期的ピークを検出する方法であり、並列処理法、データ減少法、零交差計数法等がある。

2. 相関処理

音声のデジタル信号処理において最も広く用いられている方法であり、相関処理が波形の位相ひずみに強く、また、ハード化においても比較的簡単な構造で実現できる。自己相関法、変形相関法等がある。

3. スペクトル処理

ケプストラム分析を用いる方法である。ケプストラム分析により、スペクトル包絡構造とスペクトル微細構造が分離され、有声音の場合のケプストラムは、ピッチ抽出に相当するケフレンシーにおいて強いピークを持ち、ピッチ周期が抽出される。

自己相関法

基本周波数の抽出法の中で良く用いられるのは自己相関法である。

音声波形 $x(n)$ の自己相関関数 $\phi(m)$ は、基本周波数周期 (T) と一致する点において顕著なピークを持つ。自己相関法ではこのピーク値を検出して基本周波数抽出を行う。つまり、信号を時間的に少しずつずらしてもとの信号とかけあわせたものの和をとることで、自己相関の最大値としての基本周波数が求められる。式は以下を用いる。

$$\phi(m) = \frac{1}{N} \sum_{n=0}^{N-1-|m|} x(n)x(n+|m|) \quad (|m| = 0, 1, \dots, N-1) \quad (2.7)$$

N : 波形を切り出して分析に用いる短区間の標本数

2.3.4 音声合成

テキストを音声に変換して出力することを、テキストからの音声合成と呼ぶ。音声合成の方式は大きく分けて2つあげられる。一つはフォルマント合成であり、もう一方はコーパスベース合成である。

フォルマント合成

フォルマント合成とは、調音器官の形状を制御する調音モデルを用いる合成方式である。自由な内容の音声合成では、音声合成の制御ソフトウェアとして、人が音声言語習得の際に獲得した知識を必要とする [32]。そのため、フォルマント合成では、実際の音声言語データの分析に基づいて、人間の発声機構や韻律的特徴を近似したさまざまな規則を用いて音声を作成する。この合成法では人間の声は一切用いられないため、出来上がった合成音声は不自然に聞こえるとの評価を受けることが多い。しかし、音声物理量パラメータの操作性が高いという利点があり、音声における感情表現の研究に用いられることがある [72]。

コーパスベース合成

前もって作成された音声データベースから最も適切な単位を選択しスペクトルを生成する合成方式をコーパスベース合成と言う。この手法による合成音声は、適切なデータがデータベースから検索されると非常に自然な音声合成される。データベースへの音声データの蓄積には規定されたパラメータを用いることで容量を小さくしているものの、データベースの作成には多大な労力が伴う。また、検索の際のパラメータ重み付けに結果が大きく影響される。

2.4 音声と心理状態の研究

音声と心理状態の関連にはじめに注目したのは Darwin [15] とされている。 Darwin 後は、方法的な問題等から音声と心理状態に関する研究が減少した時期もあったが、計算機技術や音声処理技術の目覚ましい発展と共に、近年には盛んに研究が行われている。ここでは、音声と心理状態の関連研究として、音声からの心理状態の認知研究、音声に表れた心理状態の音声分析研究、心理状態が関連した音声合成研究を取り上げる。

2.4.1 音声からの心理状態の認知研究

聴取実験を行い、発声者の心理状態を音声から推測する方法による研究である。

Davitz&Davitz [16] は、話し手にアルファベットの一部を諳んじながら 10 の異なる感情を表現させたものを刺激として、聞き手に音声に含まれた感情を当てさせる実験を行った。結果は、話し手によって音声だけで 54 % の感情が伝わった人もいたが、23 % しか伝わらなかった人もいたという。しかし全体的には、話し手と聞き手の情緒の伝達度は偶然では説明が付かないほど高いと結論している。 Davitz [17] は、同手法による一連の実験結果をまとめ、非言語的な声の特徴から推測できる感情の一覧表（表 2.2）を作成している。

表 2.2: Davitz による情緒表現に結び付いた音声の手がかり

感情	大きさ	高さ	音色	速さ	変化	リズム	発音
愛情	静か	低い	ひびく	ゆっくり	同じ調子だが少し上昇気味	規則的	不明瞭
怒り	大きい	高い	かなり立てる	速い	不規則に上下する	不規則	ボキボキ折れる
退屈	普通～小さい	普通～低い	中程度のひびき	適度にゆっくり	単調か次第に下降する		やや不明瞭
快活	適度に大きい	適度に高い	適度にやかましい	適度に速い	上下し全体としては上昇気味	規則的	
いらだたしさ	普通～小さい	普通か適度に高い	適度にやかましい	適度に速い	やや上昇気味		ややボキボキ折れる
喜び	大きい	高い	適度にやかましい	速い	上昇気味	規則的	
悲しみ	静か	低い	ひびく	ゆっくり	下降気味	不規則な休止	不明瞭
満足	普通	普通	ややひびく	普通	やや上昇気味	規則的	やや不明瞭

音声から発声者の感情を推測する方法は、精神障害を持つ患者の感情の認知に関する研究としても応用されている。丸山 [47] は、統合失調症患者の情動疎通性の異常に着目して、言葉の音調の受け取り方からコミュニケーションに必要な共感能力の有無を調べる音調テストを考案した。「こっちへいらっしゃい」というセリフの緩急、強弱、抑揚を変えたもの 8 種類他を刺激として用いた聴取実験を行った結果、統合失調症患者は対照健常者よりも話者の感情をずれて受け取り、質的

にも違う意味づけをすることが明らかになった。三須 [49] は、統合失調症患者とその家族に対して音調テストを実施した。結果は類型に整理されたが、父、母、患者と同胞の順に音調テストの結果が悪いということが明らかになった。その後、丹野、町山、荒尾 [89] が、統合失調症患者の距離弁別と感情認知との関連に着目し音調テストを実施している他、安西、丹羽、斎藤、増井、亀山、平松、岡崎 [5] も同様の手法を用いて統合失調症患者の感情認知について検討を行っている。

以上の研究を通じて、人間は音声から感情を認知することが確認されてきた。同手法の研究は感情音声の評価としての聴取実験で頻繁に用いられている。

2.4.2 音声に表れた心理状態の音声分析研究

計算機の普及と音声処理技術の発展と共に、感情が表現された音声を分析し、音声物理量の特徴と心理状態との関連を検討する研究が行われてきた。

Williams と Stevens [97] は、話者の感情状態を反映する音声のパラメータを同定することを目的として、感情を表現した音声と、日常生活で感情が表れた状況での音声とを比較している。その結果、「怒り」、「恐れ」、「悲しみ」の状況では、音声基本周波数、スペクトラム、テンポ、波形の形態などに相違がみられたと報告している。

水木 [50] は、15名の被験者が発声した驚き、喜び、嫌悪、怒り、恐怖の5つの感情が表現された「え」の音声を録音し、分析した。音長、波形の概形、ピッチ周期、スペクトル包絡線の傾斜の観点からその音声を分析したところ、音長は無感情よりも感情語の方が長く、ピッチ周期は無感情では感情語より短く、スペクトル包絡線は無感情より感情語の方が急で、波形の概形は3つに分類できることが分かった。また、伊東 [34] は、音声のダイナミックな性質に着目し、同じ課題を7名の被験者に実施し、音声の波形の概形、ピッチ周期の時間変化、スペクトル包絡線の傾斜の時間変化を分析している。

広瀬、高橋、藤崎、大野 [29] は、感情的に中立的な短文に「喜び」、「怒り」、「悲しみ」、「おそれ」、「驚き」、「同情」の感情を表現し発声した音声と、文章を朗読した音声を比較している。その結果、感情を表現した音声では、文頭から文末へ向かう基本周波数の下降が不明確となり、喜びでは基本周波数が数十%上昇し、怒りでは後半の基本周波数が上昇し、悲しみでは基本周波数の起伏が小

さくレベルが低くなるということを報告している。

平賀, 斎藤, 森島, 原島 [28] は, 演劇経験者が「怒り」, 「喜び」, 「悲しみ」, 「嫌悪」の感情を表現して発声した単語音声と短文音声を録音し, 「平静」音声との比較を行った。また, 自然な感情音声としてFM ラジオからの音声を採取し, 同様に分析, 比較を行った。ピッチ周波数, ピッチ振幅の変化パターンを中心に分析を行ったところ, 感情を表現した音声, および自然な感情音声に共通した特性が見られたと報告している。

また, 重永 [77, 76] が, 「喜び」, 「怒り」, 「悲しみ」, 「嫌々」, 「平静」, 「疑問」の単語音声を「平静」での音声の物理量と比較し, 基本周波数の変化についてパターンへの分類を行っている他, 感情が表現された音声を分析した研究は多数あげられる [94, 95, 39, 37, 67, 66]。

このような感情が表現された音声の研究を幅広く概観し, 感情と音声物理量のパラメータの対応表の作成を試みたものが, Scherer による表 2.3 である [69]。

以上述べてきたように, 音声物理量と心理状態の関連を検討する研究は, 感情を心理状態の対象としている場合が主流である。だが, ごく少数ながら感情以外の心理状態を対象としている場合もある。丹波 [87, 88] は, 航空事故機のボイスレコーダから緊急時の音声における基本周波数の増加を報告している。また, 宇津木と岡村 [93] は, 文章を読ませた時と雑談中との音声基本周波数を比較し, 後者の標準偏差と平均が増加することを報告している。他にも音声を用いた虚偽検出も試みられている [101]。

音声は心理状態を表す指標の一つであるとの見地から, 精神障害と音声との関連を検討する研究もある。安西ら [5, 4, 6] は, 統合失調症患者 10 名および対照健常者 10 名の音声を録音し, 基本周波数の最小周波数, 最大周波数, 中央値, 全体の持続時間, ポーズ, 無子音時間の側面から分析を行った。結果, 周波数変動率は, 「悲哀」, 「喜び」で統合失調症患者が低いことが分かった。坂口, 郭, 吉益, 百蹊, 東 [64] は, 統合失調症患者解体型 20 名, 統合失調症患者妄想型 20 名と対照健常者 40 名の物語読み上げ音声を比較した結果, 妄想型と解体型とで音声物理量の特徴が異なっていると述べている。

Andreasen, Alpert & Martz [3] は, 統合失調症患者と感情障害の患者 61 名を対象として, 半構造化面接の際の音声を分析した。対象を感情平板化群, 非感情平

表 2.3: Scherer による感情と音声物理量のパラメータの対応表

		Enjoy	Elation	Displeasure	Contempt	Sadness	Grief
		-Happy	-joy	-Disgust	-Scorn	-Dejection	-Desperation
F0	Perturbation	<=	>			>	>
	Mean	<	≥	>	<>	<>	>
	Range	<=	≥			≤	>
	Variability	<	≥			≤	>
	Contour	<	>			≤	>
	Shift regularity	=	<				
	F1 mean	<	<	>	>	>	>
	F2 mean			<	<	<	<
	F1 bandwidth	>	<>	<<	<	<>	<<
	Formant precision		>	>	>	≤	>
Intensity	Mean	<	≥	>	>>	≤≤	>
	Range	<=	>			<	
	Variability	<	>			<	
	Frequency range	>	>	>	>>	>	>>
	High-frequency energy	<	<>	>	>	<>	>>
	Spectral noise					>	
	Speech rate	<	≥			≤	>
	Transition time	>	<			>	<
		Anxiety	Fear	Irritation	Rage	Boredom	Shame
		-Worry	-Terror	-Cold	-Hot anger	-Indifference	-Guilt
F0	Perturbation		≥		>		
	Mean	>	≥≥	<>	<>	≤	>
	Range		≥≥	<	≥≥		
	Variability		≥≥	<	≥≥		
	Contour	>	>>	<	=		>
	Shift regularity		<		<	>	
	F1 mean	>	>	>	>	>	>
	F2 mean	<	<	<	<	<	<
	F1 bandwidth	<	<<	<<	<<	<	<
	Formant precision	>	>	>	>		>
Intensity	Mean		>	≥	≥≥	<>	
	Range		>	>	>		
	Variability		>		>		
	Frequency range		>>	>	>	>	
	High-frequency energy	>	≥≥	>>	>>	<>	>
	Spectral noise						
	Speech rate		≥≥		≥		
	Transition time		<		<		

> は増加, < は減少,

>> 等二重になっているものは増加 (減少) が著しいもの, <> は異なる2つのタイプの結果があるもの

板化群に分け検討した結果、感情平板化群では、音声の振幅と基本周波数の変動性が小さいと報告している。

うつ病患者の音声特徴について、音声物理量を解析した研究もある。Alpert, Pouget & Silva [1] は、うつ病患者と健常者の音声の比較を行った。結果、うつ病患者の音声は、声が大き、発話時間が短い、基本周波数が低いという差がみられたと報告されている。

Ellgring & Scherer [20] は、うつ病患者 16 名の治療前と治療後の音声を、発話時間と基本周波数について比較している。その結果、治療前に比べて治療後の患者の発話率（時間あたりの発話量）が上昇し、間が減り、基本周波数が下がると報告している。

2.4.3 心理状態が関連した音声合成研究

音声合成技術が発展するに従い、心理状態が関連した音声合成、特に感情音声の合成が検討されるようになった。

コーパスベース合成音声の手法を感情音声合成に応用した研究として飯田、キャンベル、安村 [30] の研究があげられる。飯田らは感情が表現された音声を集め、大規模な感情コーパスを作り、自然音声直接波形接続型音声合成システム CHATR を使って合成を行っている。評価の結果、合成された音声は自然性が低いものの感情は伝わると報告している。

フォルマント合成方式を用いた感情音声合成も行われている。Cahn [12] は、17 種の音声パラメータを操作できる Affect Editor を作成し、文献を参考に適切に操作したパラメータをフォルマント型合成装置である DEC-Talk に与え、6 種の感情音声を合成している。また、Murray & Arnott [51] は、感情表現時の音声パラメータの変化を整理し、DEC-Talk を用いた感情音声合成器 HAMLET を作成している。

木村、富田、本多 [41] は、「驚き」、「喜び」、「怒り」、「嫌悪」の 4 つの感情を「え」に表現して発声した音声を聴取実験し、結果について声帯特性と声道特性とに分離し、4 つの感情と音声の音長、初期ピッチ周波数、ピッチ周波数レンジ、ピッチ周波数の時間的变化、ホルマント周波数におけるスペクトルの大きさとの関係を調べた。声帯特性中のパラメータのとり値が「驚き」と「喜び」

とあまり異なっておらず、声道特性のパラメータが感情の判断に寄与していると考え、声道特性を操作することで「喜び」、「驚き」の感情を持つ合成音声を作成し、一定の結果を得ている。

フォルマント合成方式では、感情音声合成という実用的な目的のほかに、パラメータの操作による音声物理量と、聴取者の音声からの感情認知との対応を検討しやすいため、音声と感情の関連の検討手法としても有効とされる。合成音声ではないもののこの手法を参考に、Breitensrtein, Lancker & Daum は聴取実験に重点を置き、音声の知覚に関連する物理的な手がかりを系統的に調べている [9]。Breitensrtein によると、発話速度については、発話率を低くした場合「悲しみ」にラベル付けされ、発話率を高くしたところ「怒り」、「恐れ」、「中性」に分類される。また、ピッチの変動を減少させた場合は「悲しみ」や「中性」に分類され、変動を増加させると「恐れ」、「怒り」、「悲しみ」に分類されたと報告されている。

2.5 本研究での音声の扱いについて

本研究で用いた音声物理量と、音声研究のアプローチについて述べる。

2.5.1 本研究で用いた音声物理量

本研究で用いた音声物理量は基本周波数、振幅、発話時間、発話速度である。

基本周波数

基本周波数は自己相関法にて抽出し、窓関数はハミング法を用いる。測定区内で最も高い値を最高基本周波数、低い値を最低基本周波数、算術平均を平均基本周波数、最高基本周波数と最低基本周波数との差を基本周波数レンジとする。抽出のアルゴリズムを以下に示す。

1. 音声波形データを入力
2. 窓関数で波形をフレーム化
3. 各波形の自己相関関数値を計算

4. 各フレーム毎に自己相関値に基づく基本周波数上位 5 候補を抽出し，連続性を保つ基本周波数値を正解として抽出

なお，基本周波数抽出にはプログラム `pitcher` [57] を一部改変したものを用いた。

振幅

音声データ波形の振幅を算出する。測定区間内で最も高い値を**最大振幅**，低い値を**最小振幅**，区間の算術平均を**振幅平均**，最高振幅と最低振幅との差を**振幅レンジ**とする。

発話時間

発話に要した時間を**発話時間**とする。

発話速度

1 モーラ読み上げに要した時間を**発話速度**とする。

2.5.2 音声研究のアプローチ

本研究での音声研究のアプローチは，節 2.4.2 で説明を行った先行研究と対応している。すなわち，第 4 章では音声からの心理状態の認知研究と音声に表れた心理状態の音声分析研究を用い，第 5 章では，音声に表れた心理状態の音声分析研究を用い，第 6 章では心理状態が関連した音声合成研究からのアプローチである。

第 3 章

感情に関する研究

本章では、本論文で扱う感情について述べる。現在の感情研究の動向を概観することを目的に、感情における4つの代表的立場と基本感情について述べた上で、本研究における感情への視点を定義する。なお、感情は、心理学、哲学、社会学ほか様々な学問分野において研究の対象となってきたが、ここでは感情研究の中核を担う心理学の分野における研究を中心に述べる。

3.1 感情とは

今日、「感情」(emotion)と呼ばれているものは、元来パッション (passions) (情念) という単語を用い表現されていた [60]。現在、感情と類似の意味で用いられる単語には、気分、気持、情緒、情動がある。これら単語の表す概念の差異に関する明確な定義はないが、「大辞林」 [48] によると、「感情のうち、急速にひき起こされ、その過程が一時的で急激なもの」を「情緒」や「情動」と呼び、「比較的弱く長時間持続する感情の状態」を「気分」や「気持」と呼び、一方「感情」は、一時的なものである情緒や情動と、持続的なものである気分や気持との両方を包括した概念とされている。本論文では、混乱を避けるために、全てを包括した「感情」という言葉を用いる。

感情は、心理学の分野でも「動機付け」、「対人関係（特に対人魅力）」、「臨床」、「発達」といった様々な立場から研究が行われている。そのため、研究者により定義が異なり、全ての人が満足するような感情の定義や分類はないと考えられている [13]。そこで、Randolph [60] の感情への4つの視点による分類を参

考に、心理学における感情研究を概観する。特に、基本感情の立場について詳細に述べた上で、本論文における感情への視点を提示する。

3.1.1 感情研究の4つの立場

心理学における感情研究には4つの立場があると言われている [60].

第一の立場は、Darwinの研究 [15] に端を発する立場であり、感情を進化の過程で表れた必然的なものと捉える立場である。1872年、Darwinは”The Expression of the Emotions in Man and Animals [15]”において、人間と動物の感情表現として表情、音声、体の動きを解説した。Darwinによると、感情表出は動物から人間へと連続した形を持つもので、感情が存在する理由も進化論の枠組みから捉えることが可能である。例えば、人間が嫌悪感を表す時の表情は、元来はまずい食べ物を吐き出す動物の行動と関連した「有用な連合性習慣」であったと考察されている。進化論を基盤にした感情研究の立場は、現在では普遍的な少数の感情が様々な感情の基底にあるとする基本感情論の立場として発展している。

第二の立場は、感情を身体的反応の結果と捉える立場である [36]. 「ジェームス＝ランゲ説」を代表とするこの立場では、例えば「怖ろしい」という感情を例に取ったとき、「怖ろしい」から心臓が動悸するのではなく、動悸がして口が渴くために「怖ろしい」という感情が発生すると考える。身体反応を重視するこの立場は、今日では、交感神経系¹の活動と感情との関連を論じる研究が主流である。応用として、人の顔の表情を操ることにより感情経験をコントロールすることが可能だという「顔面フィードバック仮説」(facial feedback hypothesis) [91, 11]へと発展している。

第三の立場は、感情は特定の評価パターンに基づいて生起すると考える認知説である。この立場において感情とは、知覚する環境に対し、人間が有害あるいは有益と評価したことを適切に処理するための「覚悟と準備をさせる」ための反応と考えられている [43]. 第二の立場では、身体の変化が感情を生起させると捉えたのに対して、認知説では、身体の変化へ人間が評価を与えることによって感情

¹副交感神経とともに自律神経系を形成する。自律神経系は、分泌線や心臓といった内臓の機能を制御しており、発汗、唾液の分泌、消化液の分泌の制御、胃の運動の変化、涙の分泌を司っている。

が生起し、その評価の意味付けによって体験される感情が異なると考える。

最後の立場は、感情の存在そのものを社会的目的に寄与する社会的構造体と考える社会的構築主義の立場である。この立場では、感情の経験と表現は学習された慣習あるいは規則に左右されるものと捉えており、環境、文化によって生じる感情やその形態が異なると考えている。また、感情には重要な社会的機能があると想定しており、例えば、「怖れ」という感情を例に取ったとき、「怖れ」の感情を持つことで人間は社会的な規範からの逸脱を防止する効果があると考えられる [22]。

3.1.2 基本感情に関する研究

基本感情とは

人間は様々な感情を経験する。膨大な種の感情に対して、すべての感情の中核を成す「一次的感情」、あるいは「基本感情」があると想定する立場がある。基本感情論では、感情を生存、進化の過程で必要な存在と捉え、少数の基本となる純粋な感情の混合や合成から複雑な感情が派生するという前提に立っている [60]。

はじめて基本感情を提唱したのは Ekman と言われている [19, 18]。Ekman と Friesen は表情活動の記号化システム FACS(Facial Action Coding System) の比較文化的な研究において、幸福感、悲しみ、恐れ、嫌悪、怒り、驚き、興味、軽蔑が人類に共通に存在する基本感情であることを見出した。

同様に表情識別の実験を行った Izard [35] は、分離情動理論の立場から 10 の基本感情（興味・興奮、喜び、驚き、苦悩・不安、怒り、嫌悪、軽蔑、恐れ、恥、罪悪感）を報告している。Izard [35] は、基本感情について、

1. 個別で特定の神経的基盤がある。
2. 個別で特定の表情がある。
3. 意識にのぼる個別で特定の感情を所有する。
4. 進化 - 生物学的過程を通じて生じたものである。
5. 適応的機能をもつ体制化と動機づけの特徴がある。

の 5 つを基本感情の条件として挙げている。

基本感情論では、感情の存在理由を進化の過程で必要であったためと説明するが、Izard [35] は、さらに発展させた2つの観点から感情の機能を論じている。第一の観点は、養育者を引き付けるための感情の機能である。人間の子供は他の動物に比べて未熟な状態で生まれるため、長い期間に渡って他者から世話をされることが生存のためには必要とされる [59]。このため、感情の一種であるアタッチメント（愛着）を母子間で築くこと [8] は生存のための必要条件となる。第二の観点は、社会的コミュニケーションとしての感情の機能である。言葉を持たない乳児は、社会的世界とかかわることが難しい。そのため、表情や音声、類型化された行動を用いて感情の表出を行うことが大人社会とかかわる手段とされている。このような情緒的なかわり（コミットメント）の力は、性や空腹といった動因よりも強いと考えられている。Izard が感情の機能を新たに検討したことにより、基本感情は単なる外界に対する反応のパターンではなく、対人関係での社会的な意味を持つものと考えられるようになった。

基本感情論と異なる立場

基本感情論は、少数の代表的な感情で人間の持つ様々な感情体験を要約し、説明しようとする立場であるが、このような感情の捉え方に対して異論を唱える立場もある。

基本感情論に対する意見の一つは、感情次元論である。感情の構造について論じる中から台頭したこの立場は、基本感情論の多くがいくつかの独立したカテゴリ構造を持つと仮定しているのに対して、感情とは少数の次元的な構造を持つものとする。

感情次元論は、主観的感情経験を意識の次元として概念化した Spencer [78] から始まったとされている。Wundt [99] は、情動や情動経験によって記述される意識の領域が、快-不快、弛緩-緊張、沈着-興奮の3次元によって説明できると主張した。Schlosberg [70] は、個別の情動の顔面表出を分類し、多くの表情が、「愛、幸福、嬉しさ」、「驚き」、「恐れ、苦悩」、「怒り、決意」、「嫌悪」、「軽蔑」のカテゴリによって分類できることを示したが、後に快-不快、注目-拒否、睡眠-緊張の3次元でこれらの感情を集約した [71]。Woodworth [98] は、表情を、快-不快、受容-拒否、睡眠-緊張の3つの次元で表現できると提唱し

た。

Plutchik [68] は、感情語の調査を通して、感情を強度、類似度、両極性からなる立体モデルで表している。Plutchik は感情について「基本となる一時的な感情があり、その他はそこから派生した二次的な感情である。ちょうど色に原色と、混合されたその他の色があるのと意味は同じだ」と述べて8つの一時的な原型感情（恐れ、怒り、喜び、悲しみ、受容、嫌悪、期待、驚き）を定義し、それぞれの間色にあたる情動が存在するという意味で次元を論じている。Polivy [58] は、多次元感情尺度を用いて測定したところ、ターゲットとしている感情以外の感情が喚起されていることを指摘した。Polivy は被験者に感情を区別する能力がないか、感情そのものが独立したカテゴリーで分けられるものではなく被験者は実際に多種の感情が喚起するように感じているかのいずれかが原因として考察されている。また、Russell [63] は、統計手法である主成分分析を用いて感情が2次元であることを述べているが、最近では感情はカテゴリ構造を持つという意見と、次元構造を持つという意見が統合されて行く傾向にある [96]。

基本感情への他方の立場からの意見は感情の発達論である [45]。Bridges [10] が乳児の観察により感情の発達と分化を述べたことにはじまった感情の発達論は、その後感情がどのように生まれるかについての議論に発展した。基本感情が普遍的で生得的なものと仮定している生物学的モデルであることに対して、社会化モデルでは、感情行動に感情のラベルづけと適切な社会的反応を与えることで乳児は後天的に感情を学習すると考える。このモデルでは、感情状態はある文脈における社会的な環境の反応として学習されると考える。

精神分析学の流れを汲む Klein は、乳幼児の精神発達を論じる中で、怒りの後に抑うつが、羨望の後に嫉妬が生まれるとして、感情には早期から存在するものと後に発生するものがあると述べている。さらに、感情の種類によって、二者関係で発生する感情と発達の進んだ三者関係に至ってはじめて発生する感情とを区別することで、感情は自然発生的に分化してゆくのではなく、対象関係の発達に伴って複雑に発達して行くことが述べられている [73]。一方 Emde [21] は、乳幼児の感情を観察する中で、生理学的に、あるいは状況相関的に感情を分類することは無意味で、感情は世話へのメッセージと社会的相互作用に通じるものと強調している。

このように、社会化モデルでは、感情を環境や養育者との相互作用から学習するものと考えるが、Lewis & Rosenblum [45] は、どちらか一方のモデルにて人間の感情を説明することは不十分であると警鐘を鳴らしている。Lewis らによるとこの2つのモデルは感情の種類によって適合が異なり、例えば一次的感情 (primary affects) は、生物学的なモデルが説明に適し、後天的感情 (learned affects) は、社会化されたモデルを説明するのに適しているという。さらに発達の初期では生物学的なモデルが適切だが、より後期の発達過程では学習モデルが適切であるため、人間の持つ様々な感情についての一つのモデルのみを適用することは不十分であると結論している。

3.2 本研究の感情への視点

本研究では、感情をいくつかの基本的なカテゴリで分類できるものとする基本感情論の立場より感情を扱う。

第一の理由は、感情そのものが基本的なカテゴリを持つと仮定することで、研究の目的や方略が明確に整理されたと考えられたためである。本研究で対象とする音声は、それ自体が複数の情報を持つものであり、次元や機能面を重視した立場より感情を扱うことは研究の主旨の混乱を招く恐れがある。Cowie [14] は、音声研究の自然な出発地点は基本感情を用いることだと述べているが、本研究においても感情次元論のように関連するの連続的な感情を定義するよりも、必要最小限のカテゴリを明快に定義することが有益だと考えられた。

第二の理由は研究の実用性を考慮したことである。感情と音声の関連を検討するという本研究での目的は、計算機による感情の判断と、合成音声に感情情報を付与するという方法で評価される。ここでは、複雑な感情や、発達後期に生まれる感情があったとしてもそれを対象とするのは時期焦燥であり、少数の基本となる感情を扱う必要があると考えられる。これらの実用面を考慮したとき、生理学的根拠、認知や社会的相互作用を重視した感情論よりも、むしろできるだけ少数の、それぞれの目的に必要なとされた感情のみを基本感情の中から用いることが有益と考えられた。

第三の理由は、本研究が感情そのものの構造についての心理学的研究ではない

ことである。このため、感情の複雑な構造を討論することを前提にした方法よりも、むしろ感情を独立したカテゴリに分けられるものであり、人間の発達早期より発生する普遍的なものと仮定するほうが有益と考えられた。

以上の理由から、本研究では感情の基本的なカテゴリを、第4章「感情を込めた音声の分析」では6次元からとらえ、第5章「抑うつ傾向と音声に関する研究」では基本的カテゴリの中の一つの次元を重点的にとらえ、第6章「感情豊かな読み上げのための音声パラメータ」ではさらに基本的カテゴリを4次元に集約して研究の目的に必要な部分のみを扱うこととする。

第 4 章

感情を表現した音声の分析

本章では、「精神分裂病患者の感情について」 [80] において得られたデータを再分析し、音声と感情の関連を論じる。ここでは、音声と感情の関連についてのみ焦点をあて、論文での主要テーマであった精神分裂病患者（統合失調症患者）の感情についての分析と考察は行わない。

はじめに、本章での研究に関する概要を述べ、研究の流れについて紹介する。次に、音声聴取実験で用いる尺度を作成する。さらに、感情が表現された音声（以下、感情音声と呼ぶ）の聴取実験の報告と、刺激として用いた感情音声の分析について述べる [82]。最後に、22名の発声者が発声した感情を表現した音声の音声物理量についての統計的分析を行い [84]、感情と音声の関連について考察を行う。

4.1 感情と音声の関連検討のための研究の概略

感情と音声の関連検討のための研究は以下の流れで行う。

1. 調査

実験で使用する簡便な感情状態尺度を作成する。

2. 声優の感情音声の音声物理量抽出

声優学校へ通う女性が発声する感情を表現した音声について、音声物理量の抽出を行い、特徴について検討する。

3. 感情音声聴取実験

感情を表現した音声を刺激として、感情音声の聴取実験を行う。

4. 一般者の感情音声の採取と統計的分析

感情表現について特別な訓練を受けない一般者 22 名が発声した感情音声の音声物理量を抽出し、統計的解析を行う。さらに、音声物理量の変化の特徴を把握することを目的に、音声物理量のデータの目視による分類の手法を提案し、音声と感情の関わりについて検討を行う。

4.2 調査

4.2.1 目的

感情音声の聴取実験では、強制選択法にて音声に表れている感情を測定することが多いが、正確に感情を測定するためには、感情状態を多側面より測定可能な簡便な尺度が必要だと考えられる。信頼性¹や、妥当性²が共に優れている尺度としては、多面的感情状態尺度 [90] があげられるが、実験で繰り返し使用するには項目数が多く、調査対象者の負担が大きいと思われる。そこで、実験で使用できるような簡便な感情状態尺度を作成することを目的に調査を行う。

4.2.2 方法

対象

大学生を調査対象者とする。学生 227 名に調査を実施し、分析に用いた有効回答数は 206 部であった。

調査項目

調査項目は、多面的感情状態尺度 [90] に感情表現辞典 [53] から選定した感情状態を表す形容詞 70 語を加えた、150 項目である。教示として「あなたは今、どのような感情を持っていますか。最も近いものに○をつけて下さい。」という文章と、例を用意し、その後に質問項目を記載する。評定は「まったく感じていない」、

¹再テスト時の値の安定性を示す

²測定する概念の包括程度を示す

「いづらか感じている」，「かなり感じている」，「はっきり感じている」の4件法で求める。

4.2.3 結果

調査回収後，感情状態を表す形容詞項目の回答は，「まったく感じていない」を1点，「いづらか感じている」を2点，「かなり感じている」を3点，「はっきり感じている」を4点と得点化し，全150項目について，因子分析(主因子解，promax回転)を行った。いづれの因子にも負荷量が低い項目，複数の因子に負荷量が高い項目を除き，因子分析を繰り返し行った。

因子分析の結果105項目6因子が抽出された。結果のパターン行列及び因子相関行列を表4.1，表4.2に示す。なお，音声聴取実験では，調査での男女差が大きい項目，主成分分析の負荷量が少ない項目を抜き，6因子12項目を感情状態尺度として使用することとした。尺度項目の信頼性検討のため，Cronbachの α 係数を算出したところ，「怖」0.829，「喜」0.813，「安」0.750，「鬱」0.740，「怒」0.779，「驚」0.888であり，高い信頼性が得られた。実際に使われた感情状態尺度は図4.1であった。

4.3 声優の感情音声の音声物理量抽出

4.3.1 目的

音声と感情の関連を捉えることを目的として，感情が表現された音声から音声物理量を抽出する。

4.3.2 方法

感情音声の採取

静かな部屋で，声優養成学校へ通う女性が感情を表現して「こっちへいらっしゃい」という台詞の発声を行う。発声者には，感情状態尺度の因子となる6感情，すなわち「喜」，「鬱」，「怖」，「驚」，「安」，「怒」をそれぞれ大小，それに「無」を加えた13の感情を読み分けることを課題として提示する。音声の録音にはMD(sony MZ-R50)を用いる。

表 4.1: 感情状態尺度 (105 項目) の因子分析結果 (promax 回転によるパターン行列)

*は尺度で用いた項目

質問項目 / 因子		1	2	3	4	5	6
怖	びくついた	0.880	-0.150	0.051	-0.134	-0.072	0.036
	震え上がった	0.862	-0.124	-0.093	-0.098	-0.082	0.079
	* 恐ろしい	0.839	-0.070	-0.100	-0.071	0.032	-0.057
	* おびえている	0.821	-0.057	-0.035	0.144	-0.112	-0.042
	がたがたとした	0.782	-0.036	-0.115	-0.032	-0.039	0.032
	冷や冷やした	0.764	0.002	0.006	0.011	-0.036	0.191
	縮み上がった	0.753	-0.083	-0.052	-0.031	-0.023	0.150
	泣ける	0.752	0.038	-0.005	-0.066	0.062	-0.094
	怖い	0.735	0.057	-0.110	0.005	0.065	-0.157
	どうてんした	0.718	0.032	-0.153	0.104	-0.155	0.224
	臆病な	0.715	0.031	0.041	0.315	-0.208	-0.053
	おろおろした	0.712	0.075	-0.080	0.295	-0.139	0.008
	しんみりした	0.696	0.067	0.148	-0.061	0.017	0.000
	不気味な	0.688	-0.038	0.039	-0.109	0.126	-0.067
	ほろっとした	0.680	0.175	0.000	-0.191	0.056	-0.092
	ぞっとした	0.671	-0.036	-0.014	-0.054	0.046	0.147
	びくりとした	0.649	0.025	-0.142	0.091	-0.012	-0.028
	ぎょっとした	0.631	-0.041	0.000	0.124	0.002	0.234
	驚いた	0.614	0.109	0.021	-0.145	0.018	0.278
	びっくりした	0.597	0.169	-0.087	-0.055	-0.009	0.341
	目が潤む	0.596	0.072	-0.085	-0.173	0.203	0.024
	うろたえた	0.584	-0.012	0.004	0.259	-0.171	0.172
	緊張した	0.581	0.214	-0.055	0.144	0.045	0.034
	物悲しい	0.558	-0.027	0.094	0.203	0.069	-0.130
	ひやっとした	0.558	-0.062	0.020	-0.108	0.120	0.216
	うらんだ	0.556	-0.011	-0.162	-0.079	0.365	-0.112
	胸騒ぎがする	0.533	0.147	-0.024	0.116	0.119	0.028
	寒気がする	0.456	0.015	-0.029	0.114	0.165	0.196
	寂寥とした	0.454	-0.095	0.207	0.120	0.092	-0.227
	哀れな	0.446	0.130	0.002	0.276	0.180	0.016
沈痛な	0.425	-0.057	0.087	0.135	0.268	0.114	
悲しい	0.384	0.038	0.052	0.272	0.277	-0.075	
喜	* 喜んでいる	-0.013	0.948	-0.234	-0.006	0.017	-0.147
	* 楽しい	-0.079	0.930	-0.087	0.114	-0.069	-0.028
	嬉しい	0.010	0.891	-0.124	0.089	-0.122	-0.088
	活気のある	0.001	0.870	-0.212	0.066	0.051	0.055
	面白い	-0.141	0.853	-0.017	0.258	0.064	-0.051
	うきうきした	0.006	0.814	-0.073	0.080	0.085	-0.018
	はつらつとした	-0.098	0.809	-0.224	0.060	0.043	0.019
	愉快的な	-0.148	0.804	-0.078	0.173	0.067	0.069
	元気いっぱい	0.101	0.798	-0.174	-0.075	-0.109	0.123
	充実した	0.022	0.734	-0.116	-0.130	0.013	-0.051
	機嫌の良い	-0.002	0.724	0.125	-0.050	-0.026	0.014
	弾んだ	-0.039	0.717	0.011	0.053	0.175	0.086
	満足した	-0.050	0.716	-0.025	-0.014	-0.039	-0.079
	明るい	-0.104	0.712	0.046	0.063	-0.009	0.078
	気力に満ちた	0.034	0.710	-0.135	-0.161	0.062	0.007
	晴れ晴れとした	-0.050	0.710	0.067	-0.127	0.044	0.071
	陽気な	0.033	0.706	0.084	-0.038	0.015	0.043
	わくわくした	0.071	0.705	0.044	-0.030	0.049	0.056
	満ち足りた	0.078	0.704	0.009	0.070	-0.095	0.003
	幸せな	-0.030	0.693	0.071	-0.001	-0.044	-0.045
	快調な	-0.024	0.678	0.148	-0.112	-0.035	0.031
	快適な	-0.021	0.677	0.177	0.018	0.044	-0.072
	晴れやかな	0.079	0.675	0.059	-0.136	0.115	0.037
	爽快な	0.021	0.674	0.162	0.006	-0.133	0.066
	さわやかな	-0.029	0.614	0.242	-0.065	0.107	0.035
	心地良い	-0.034	0.595	0.237	-0.024	-0.050	-0.055
	のびのびとした	-0.105	0.591	0.346	0.002	0.006	0.007
	心安い	0.124	0.562	0.194	0.143	-0.157	-0.050
	安心した	0.174	0.541	0.251	-0.179	-0.024	-0.064
	救われた	0.288	0.514	-0.075	-0.023	0.114	-0.078
気持ちの良い	0.200	0.477	0.284	-0.177	0.035	-0.017	
ほくほくする	0.112	0.456	0.092	0.013	-0.039	0.017	

安	のどかな	-0.166	-0.034	0.850	0.054	0.052	-0.005
	ゆっくりした	0.022	-0.127	0.843	0.067	-0.099	0.013
	のんびりした	-0.103	0.013	0.819	0.138	-0.101	0.015
	* くつろいだ	-0.035	0.166	0.639	-0.022	0.066	0.036
	* 落ち着いた	-0.162	0.105	0.638	-0.012	0.132	0.164
	平静な	-0.099	0.100	0.604	0.050	0.144	0.075
	のんきな	-0.081	0.108	0.581	0.196	-0.085	-0.030
	ゆうゆうとした	-0.028	0.284	0.580	-0.075	0.139	0.056
	気長な	-0.132	0.060	0.565	0.202	0.069	0.166
	安堵した	0.119	0.245	0.550	-0.250	-0.040	-0.116
	気分がなごんだ	0.130	0.282	0.549	-0.222	-0.034	-0.117
	おっとりした	0.152	0.129	0.548	0.015	-0.106	0.125
	ゆったりした	-0.312	0.173	0.544	0.227	-0.015	0.003
	ぼんやりした	0.392	-0.226	0.539	0.129	-0.050	-0.084
	平安な	-0.084	0.334	0.484	0.027	-0.107	-0.052
やわらいだ	0.270	0.256	0.454	-0.183	-0.109	-0.087	
鬱	自信がない	0.187	0.053	0.017	0.717	-0.163	0.066
	悩んでいる	0.062	0.043	-0.008	0.716	0.010	0.062
	ばやばや	-0.082	0.016	0.402	0.631	0.001	0.137
	引け目を感じている	0.200	0.044	0.078	0.619	-0.090	-0.011
	疲れた	-0.062	-0.012	0.029	0.584	0.044	0.093
	くよくよした	0.383	-0.002	0.023	0.555	0.009	-0.056
	無気力な	0.002	-0.222	0.142	0.522	0.143	0.051
	だるい	-0.062	-0.138	-0.068	0.501	0.127	0.107
	* 沈んだ	0.283	-0.112	0.117	0.460	0.177	-0.212
	* 心淋しい	0.314	0.097	0.070	0.458	0.092	-0.160
悲観した	0.133	-0.072	0.146	0.389	0.267	-0.059	
怒	敵意のある	0.099	0.008	0.009	-0.012	0.770	-0.128
	むっとした	0.178	0.031	-0.083	0.114	0.692	-0.052
	憎らしい	0.316	0.037	-0.122	-0.037	0.611	-0.123
	ばからしい	0.205	-0.228	0.122	0.067	0.601	-0.017
	* かつとした	0.359	-0.013	0.012	-0.100	0.600	0.025
	攻撃的な	0.099	0.193	-0.176	0.006	0.594	0.137
	* おこった	0.307	0.054	-0.112	-0.014	0.540	0.001
	呆れた	0.228	-0.162	0.210	0.021	0.509	0.148
	不機嫌な	0.145	0.028	-0.059	0.331	0.491	0.013
	鋭敏な	0.079	0.144	0.106	-0.012	0.481	0.153
気分を害した	0.240	-0.078	-0.012	0.073	0.462	-0.064	
驚	* 面食らった	0.254	-0.013	0.133	0.163	-0.038	0.760
	* 度肝を抜かれた	0.350	0.030	-0.004	0.084	-0.068	0.658
	たまげた	0.419	-0.122	0.132	-0.110	0.122	0.556

表 4.2: 感情状態尺度 6 因子の因子相関行列

		1	2	3	4	5	6
1	怖	1.000					
2	喜	0.156	1.000				
3	安	0.132	0.567	1.000			
4	鬱	0.396	-0.364	-0.166	1.000		
5	怒	0.561	0.065	0.051	0.432	1.000	
6	驚	0.207	0.122	-0.076	-0.140	0.169	1.000

図 4.1: 感情状態尺度

あなたは、今聞いた声に、どのような気持ちが含まれていると思いますか？
最も近い番号に○をつけて下さい。

		全く感じていない	いくらか感じている	かなり感じている	はっきり感じている
1 喜んでいる	1	2	3	4
2 心淋しい	1	2	3	4
3 楽しい	1	2	3	4
4 度肝を抜かれた	1	2	3	4
5 面食らった	1	2	3	4
6 落ち着いた	1	2	3	4
7 おびえている	1	2	3	4
8 恐ろしい	1	2	3	4
9 おこった	1	2	3	4
10 かつとした	1	2	3	4
11 くつろいだ	1	2	3	4
12 沈んだ	1	2	3	4

4.3.3 音声物理量の抽出

録音した音声は、計算機に入力し WAVE 形式ファイルに変換する。音声物理量として、それぞれの課題において、声の大きさに対応する一文の波形の平均振幅（パワー値）、声の高さに対応する一文の平均基本周波数 (Hz)、話すテンポに対応する発声時間を抽出する。なお、「喜小」については録音状態が悪いため分析から除外する。

4.3.4 結果

感情音声から抽出した音声物理量の結果の一覧を表 4.3に掲載する。

表 4.3: 声優の感情音声の音声物理量

	怒大	怒小	喜大	無	安大	安小	鬱大	鬱小	驚大	驚小	怖大	怖小
平均周波数 (Hz)	161	212	271	196	204	255	171	283	174	259	171	155
振幅 (パワー値)	235	209	235	219	218	195	219	203	222	214	229	206
時間 (sec)	1.3	1.1	1.7	1.3	1.6	1.4	1.5	1.5	1.5	1.3	1.2	1.0

4.3.5 考察

採取、分析を行った音声は各条件につき一データであるため、代表値から音声物理量の特徴を述べる。

平均基本周波数は感情を大きく表現したものではありません。小さく表現したものは高い傾向にある。振幅は、感情を大きく表現したものは大きく、小さく表現したものは小さい傾向にあり、「怒大」、「喜大」では特に大きい傾向にある。発話時間は「怒」、「怖」では発話時間が短く、「喜」、「安」、「鬱」では長い傾向にある。

全体としてまとめると、感情の大小については、感情が大きい場合は声が低く大きく、感情が小さい場合は声が高く小さい傾向にある。また、声の大きさや、発話のテンポといった音声物理量は感情のカテゴリによって異なることが示唆されている。

4.4 感情音声の聴取実験

4.4.1 目的

感情音声にどのような感情が表れていると感じられるか推定を行う聴取実験を行い、発声者の表現した感情と聴取者が認知した感情を比較検討する。

4.4.2 方法

音声刺激

声優養成学校へ通う女性が発声する感情音声（前述の音声）を聴取実験の刺激として用いる。

聴取対象者

学生 10 名を対象とする。

手続き

実験者は教示を述べた後、テープレコーダーで音声の刺激を 2 回続けて提示する。聴取対象者は感情音声刺激にどのような感情が表れているかを感情状態尺度で評定する。なお、尺度の項目と刺激の提示順序は、順番効果を考慮して調査対象者によりランダムに配置する。

4.4.3 結果

感情状態尺度の回答は、「まったく感じていない」を 1 点、「いくらか感じている」を 2 点、「かなり感じている」を 3 点、「はっきり感じている」を 4 点と得点化し、各因子ごとの合計得点を算出した。全ての刺激に関して、6 感情因子（喜因子、鬱因子、驚因子、怖因子、安因子、怒因子）を要因とした Friedman 検定、多重比較として Wilcoxon の符号つき順位検定を行った。

「怒大」

「怒大」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 37.722$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、怒が他

の因子よりも、怖が喜、安よりも得点が有意に高かった。結果を図 4.2に示す。

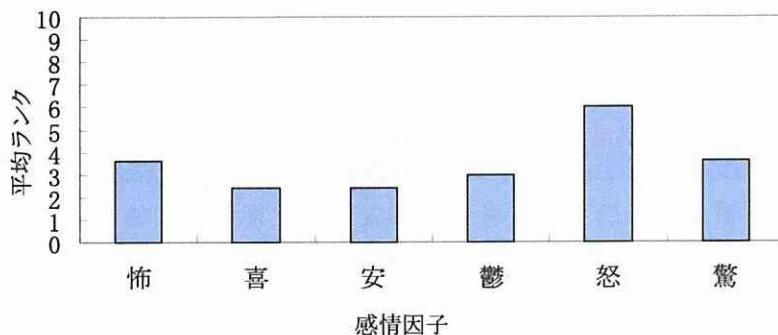


図 4.2: 「怒大」の感情が表れた音声を刺激とした聴取実験の感情尺度

「怒小」

「怒小」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 32.042$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、怒が他の因子よりも、怖が喜、安よりも、鬱が喜よりも得点が有意に高かった。結果を図 4.3に示す。

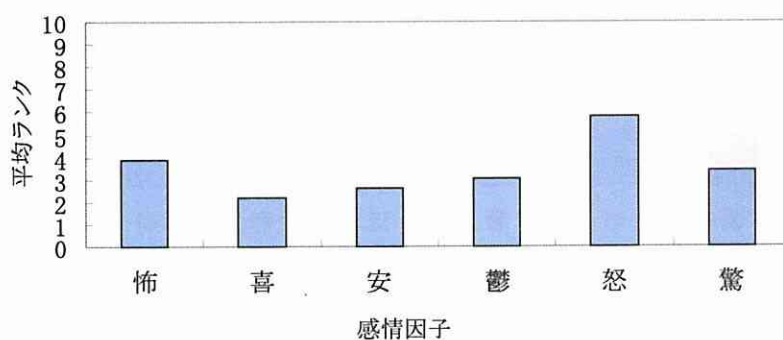


図 4.3: 「怒小」の感情が表れた音声を刺激とした聴取実験の感情尺度

「喜大」

「喜大」は、Friedman 検定では差が認められなかったが、Wilcoxon の符号つき順位検定を行ったところ、喜が安、鬱、怒が鬱、怖よりも得点が有意

に高いと傾向が認められた。結果を図 4.4に示す。

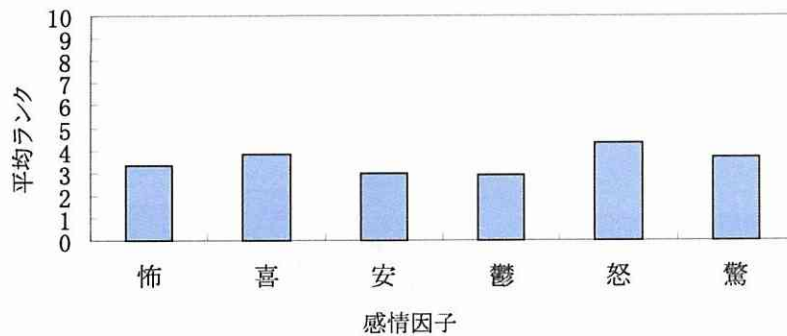


図 4.4: 「喜大」の感情が表れた音声を刺激とした聴取実験の感情尺度

「無」

「無」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 35.842$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、怒が他の因子よりも、鬱が喜、怖、驚よりも、安が怖、喜、驚よりも得点が有意に高かった。結果を図 4.5に示す。

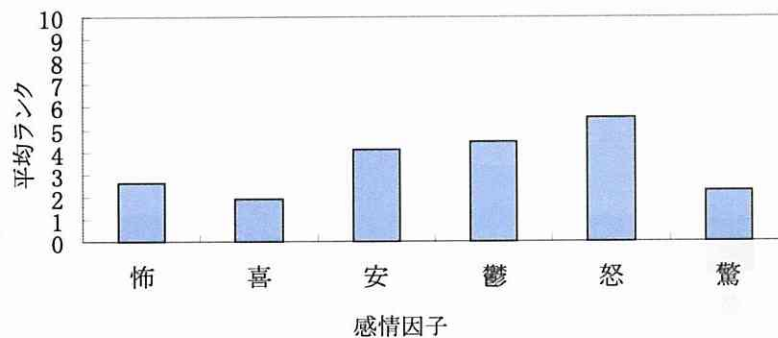


図 4.5: 「無」の感情が表れた音声を刺激とした聴取実験の感情尺度

「安大」

「安大」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 21.636$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、安が驚、怒、喜、怖よりも、鬱が怒、驚よりも得点が有意に高かった。結果を図

4.6に示す.

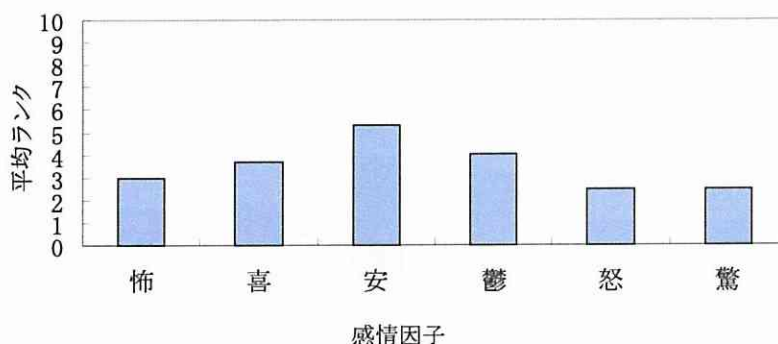


図 4.6: 「安大」の感情が表れた音声を刺激とした聴取実験の感情尺度

「安小」

「安小」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 33.899$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、安が怖、怒、驚、喜よりも、鬱が怒、驚、怖よりも得点が有意に高かった。結果を図 4.7に示す。

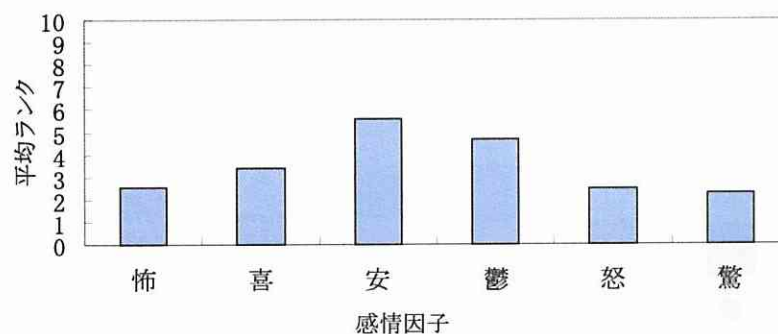


図 4.7: 「安小」の感情が表れた音声を刺激とした聴取実験の感情尺度

「鬱大」

「鬱大」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 25.573$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、安が

喜, 怒, 驚, 怖よりも, 鬱が驚, 怒, 怖よりも得点が有意に高かった. 結果を図 4.8に示す.

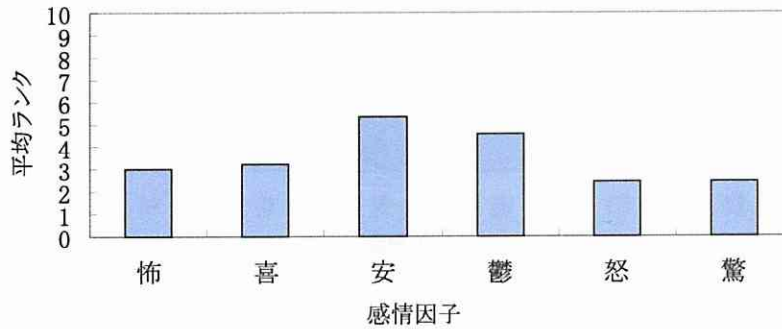


図 4.8: 「鬱大」の感情が表れた音声を刺激とした聴取実験の感情尺度

「鬱小」

「鬱小」は, Friedman 検定を行った結果, 有意の差が確認された ($\chi^2 = 25.137, df = 5, p < .01$). Wilcoxon の符号つき順位検定の結果, 鬱が驚, 怒, 喜, 怖よりも, 安が喜, 驚, 怒よりも, 怖が驚よりも得点が有意に高かった. 結果を図 4.9に示す.

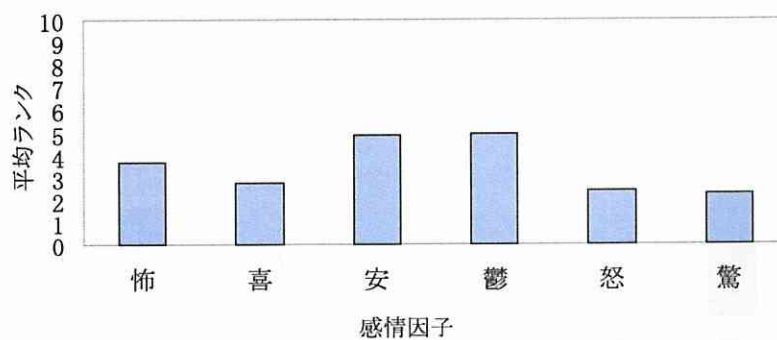


図 4.9: 「鬱小」の感情が表れた音声を刺激とした聴取実験の感情尺度

「驚大」

「驚大」は, Friedman 検定を行った結果, 有意の差が確認された ($\chi^2 = 12.961, df = 5, p < .01$). Wilcoxon の符号つき順位検定の結果, 怒が

驚、鬱、怖よりも得点が有意に高かった。結果を図 4.10に示す。

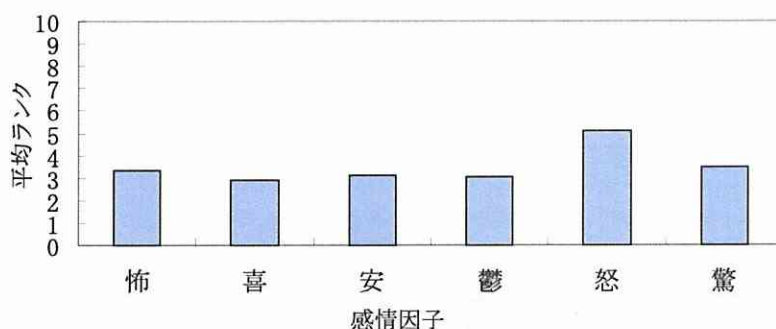


図 4.10: 「驚大」の感情が表れた音声を刺激とした聴取実験の感情尺度

「驚小」

「驚小」は、Friedman 検定で有意の差は認められなかった。

「怖大」

「怖大」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 = 28.393$, $df = 5$, $p < .01$)。Wilcoxon の符号つき順位検定の結果、驚が鬱、安、怒、喜よりも、怖が安、鬱よりも得点が有意に高かった。結果を図 4.11に示す。

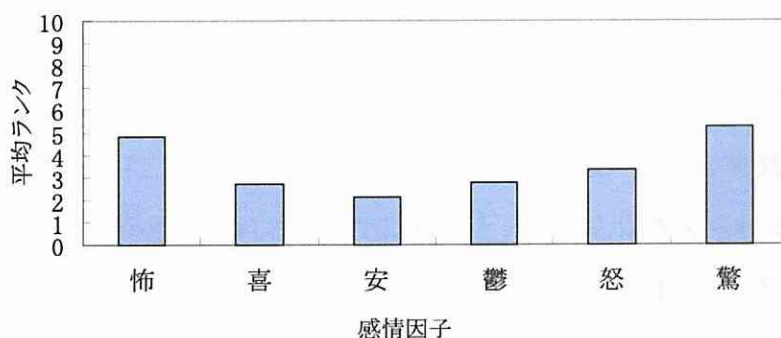


図 4.11: 「怖大」の感情が表れた音声を刺激とした聴取実験の感情尺度

「怖小」

「怖小」は、Friedman 検定を行った結果、有意の差が確認された ($\chi^2 =$

26.449, $df = 5$, $p < .01$) . Wilcoxon の符号つき順位検定の結果、怖が喜, 安, 鬱, 驚よりも, 驚が喜よりも, 鬱が喜よりも得点が有意に高かった. 結果を図 4.12に示す.

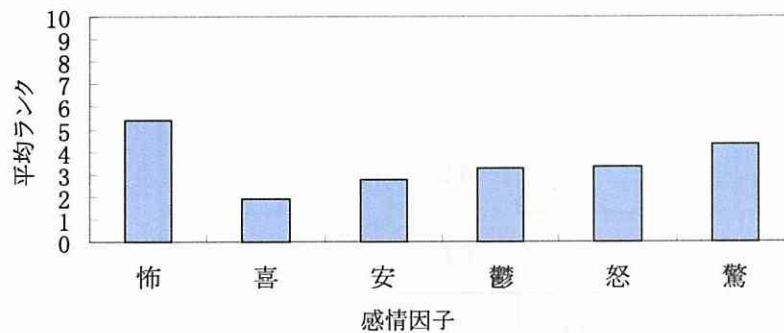


図 4.12: 「怖小」の感情が表れた音声を刺激とした聴取実験の感情尺度

4.4.4 考察

聴取実験の尺度を統計解析した結果、12の音声刺激のうち、10の刺激において差の検定で統計的に有意な差が認められた. 多重比較を行った結果、他の感情状態因子に比べて平均値に差が認められた感情状態因子について、得点が高いものからまとめたものが次の図 4.13である.

発声者が表現した感情と聴取者が推定した感情とが一致しやすかったものとして、「怒」, 「安」があげられるが、一致しなかったものとして「喜」, 「驚」があげられる. ここでは、音声によって伝わりやすい感情と、伝わりにくい感情があるということと共に、聴取者による音声の感情推定は、必ずしも発声者の表現した感情と同じものが推定されるわけではないということが考えられる.

また、聴取者が推定した感情を整理すると、「安」と「鬱」, 「怒」と「怖」が同じ音声刺激から感じ取られる感情として順位を並べていることが多い. 音声から推定される感情としては、近いものとして感じられる感情があると考えられる.

図 4.13: Wilcoxon の符合つき順位検定結果一覧 (数字は感情因子の得点平均, 枠内は他の因子と比べ得点に有意な差があるもの)

		聴取者が音声に表れていると推定した感情					
		怒	喜	安	鬱	驚	怖
発声者が音声に込めた感情	怒大	6.0	2.4	2.4	3.0	3.6	3.6
	怒小	5.9	2.2	2.6	3.1	3.4	3.9
	喜大	4.3	3.9	3.0	5.9	3.7	3.3
	無	5.6	2.0	4.1	4.5	2.3	2.6
	安大	2.5	3.7	5.3	4.1	2.5	3.0
	安小	2.5	3.4	5.6	4.7	2.3	2.6
	鬱大	2.4	3.3	5.4	4.6	2.4	3.0
	鬱小	2.4	2.8	4.9	5.0	2.3	3.7
	驚大	5.1	2.9	3.1	3.1	3.5	3.4
	怖大	3.3	2.7	2.2	2.8	5.3	4.9
	怖小	3.4	2.0	2.8	3.3	4.3	5.4

4.5 一般者の感情音声の採取と統計的分析

4.5.1 目的

音声の物理的側面と感情の関連を捉えることを目的として、感情が表現された音声について、分析による音声物理量の抽出及び統計解析と、人手による音声物理量の変化特徴の分類を行う。

4.5.2 方法

発声者

方言を持たない者 22 名を感情音声発声の対象者とする。

方法

発声者には、「喜」、「鬱」、「驚」、「怖」、「安」、「怒」の感情を表現して「こっちへいらっしゃい」というセリフを読み分けることを課題として提示する。音声の録音には MD(sony MZ-R50) を用いる。

音声物理量の抽出

録音した音声は、計算機に入力し WAVE 形式ファイルに変換する。音声の物理量として、課題の各条件における、一文での基本周波数（基本周波数平均、最小基本周波数、最大基本周波数、基本周波数レンジ）、振幅、発話時間を抽出する。

音声物理量の変化の分類

今までの分析において、音声物理量のパラメータは一文での代表値として算出してきた。しかし、人間が音声を認知する過程を考えると、数個の単語から成る文の中での代表値を感じ取るよりも、むしろ音声物理量のダイナミックな変化を特徴として感じ取っている可能性がある。そこで、ここでは、一文を通した音声物理量パラメータの変化の様子を捉えることを目標に、人による変化の分類を行う。

分類の方法としては、基本周波数、振幅、発話時間の 3 つのパラメータについてそれぞれ値を算出し、グラフで図示する。分類を行う者は、グラフを見ながら「こっちへいらっしゃい」のそれぞれ「こ」、「ちへ」、「いら」、「しゃい」にあたる部分について、どのように変化をしているかを目で見て分類する。分類は、基本周波数は高低、振幅は大小、発話時間は長短をそれぞれの箇所で行う。分類は、各パラメータについて 2 名が別々に行う。

4.5.3 結果

音声物理量の統計的分析

得られた音声物理量の各パラメータは一元配置分散分析（6感情因子）を行うことで整理した。振幅では「鬱」と「安」が、「怒」よりも振幅が小さいという差が認められた ($F(289.822, 5) = 3.587p < .01$)。結果を図 4.14に示す。発話時間では「驚」、「怒」が「喜」、「鬱」、「怖」、「安」よりも時間が短いという差が認められた ($F(1.285, 5) = 8.537p < .01$)。結果を図 4.15に示す。基本周波数については統計的に有意な差は認められなかった。

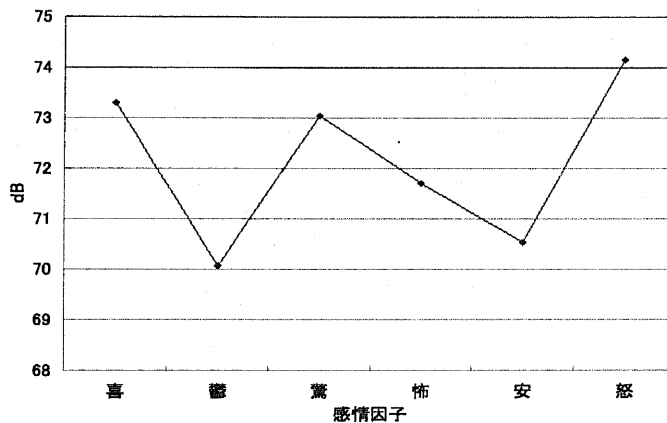


図 4.14: 感情状態尺度因子ごとの得点平均 (振幅)

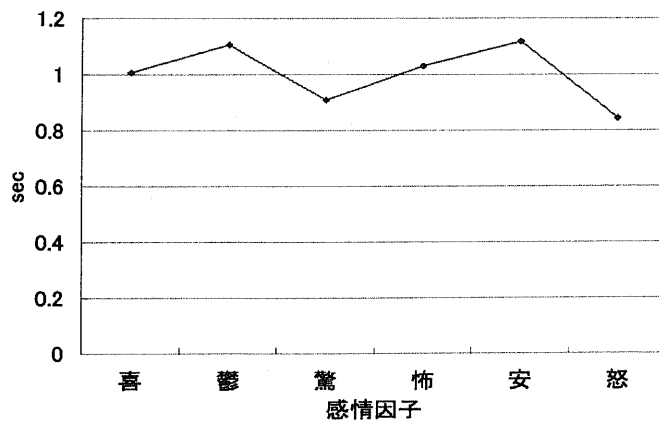


図 4.15: 感情状態尺度因子ごとの得点平均 (sec)

音声物理量の変化の分類

基本周波数、振幅、発話時間の音声物理量変化の分類については、各分類のパターンごとに集計を行った。結果を表 4.4、表 4.5、表 4.6に示す。

4.5.4 考察

音声物理量の統計的分析

複数の者が発声した感情の表れた音声に関して、音声物理量を抽出し、平均値の差の検定を行ったところ、振幅及び発話時間で統計的に有意な差が認められた。この二つの音声物理量パラメータに関しては、22名の発声者によって感情ごとに同様の傾向を持って感情の表現がなされているということが示唆されている。すなわち、「怒」の感情の音声は声が大きく発話時間が短くなり、「鬱」、「安」は声小さく発話時間が長い。「驚」は発話時間が短く、「喜」は発話時間が長い。

それに対して、基本周波数では、22名の発声者によって同様の傾向が認められないことから、基本周波数は個人の表現のばらつきが大きい領域である可能性があると考えられる。

声優による感情音声の音声物理量（4.3節）と比較すると、振幅、発話時間については、特徴が一致したものが多い。すなわち、「怒」では声が大きく発話時間が短く、「鬱」、「安」は声小さく、「喜」、「鬱」、「安」は発話時間が長い。基本周波数については、声優による感情音声においても感情の大小において関連が見られたのみで、感情のカテゴリとの関連で特徴として捉えることが不可能であった。以上のことから、発声者が感情を表現した音声における、文全体の音声物理量については、特に振幅と発話時間において特徴が表れると考えられる。

音声物理量の変化の分類

基本周波数、振幅、発話時間の音声物理量パラメータの変化を分類したところ、発話時間では結果に傾向が見られなかったが、基本周波数、振幅では感情によって傾向があった。基本周波数では、「喜」、「驚」では、「こ」が低く、「ちへ」が高く、「いら」が低く、「しゃい」が高いという変化に分類された音声が多かつ

表 4.4: 基本周波数の変化の分類表 (は基本周波数の低いもの, は基本周波数の高いもの, 太字は最も度数の大きなもの)

こっ	ちへ	いらっ	しゃい	喜	鬱	驚	怖	安	怒
-	-	-	-	3	3	0.5	0.5	5	3
-	-	-	-	0	0	0.5	0	0	0
-	-	-	-	0	0	0	0.5	0	0
-	-	-	-	3	1	1.5	2	0.5	5
-	-	-	-	1.5	0	0.5	0	0	0.5
-	-	-	-	6	2	10	1.5	4	4
-	-	-	-	3.5	1	0	1	2.5	3
-	-	-	-	0	0	0	0	0	0
-	-	-	-	0.5	0	0	2	0	0
-	-	-	-	0	0	0	0	0	0
-	-	-	-	0	0.5	0	0	0	0
-	-	-	-	0	0	0	1	0	0
-	-	-	-	0	0.5	0	0	0	0.5
-	-	-	-	0	0	0	0	0	0
-	-	-	-	4	4	6.5	11	9	5.5

表 4.5: 振幅の変化の分類表 (_ は振幅の小さいもの, $\bar{\quad}$ は振幅の大きいもの, 太字は最も度数の大きなもの)

こっ	ちへ	いらっ	しゃい	喜	鬱	驚	怖	安	怒
-	-	-	$\bar{\quad}$	0.5	2	0	1	2	0
-	-	-	$\bar{\quad}$	0	0.5	0	0	0	0
-	-	-	-	0	0.5	0	0	0.5	0
-	-	-	$\bar{\quad}$	3.5	7	6	4	4	5.5
-	-	-	-	0	0.5	1	0	0	0
-	-	-	$\bar{\quad}$	2.5	2.5	3.5	2	2.5	0
-	-	-	-	0	0	0	1.5	0.5	0
-	-	-	-	0	0	1	1.5	0	0.5
-	-	-	$\bar{\quad}$	5	2.5	1	3	5	0
-	-	-	-	0	0	0	0	0	0
-	-	-	-	0	0	0	0	0	0
-	-	-	-	0	0	0	0	0	0
-	-	-	$\bar{\quad}$	0	0	0	0.5	0	0
-	-	-	$\bar{\quad}$	0	0	0	0.5	0	0
-	-	-	$\bar{\quad}$	11	8.5	8	11	9.5	12

表 4.6: 発話時間の変化の分類表 (_ は発話時間の短いもの, - は発話時間の長いもの, 太字は最も度数の大きなもの)

こっ	ちへ	いらっ	しゃい	喜	鬱	驚	怖	安	怒
-	-	-	-	1.5	0.5	5	2.5	0.5	3
-	-	-	-	0.5	0	1.5	1	1	0
-	-	-	-	0	1	0	0	1	0.5
-	-	-	-	0	0	0	0	0	0.5
-	-	-	-	0	0	0.5	0	0	0
-	-	-	-	6.5	1	3	1	3	2
-	-	-	-	3.5	9	4.5	5	3	5
-	-	-	-	0	0	0	0	0	0.5
-	-	-	-	5.5	2.5	1.5	3	4	5
-	-	-	-	1.5	3	3.5	2.5	6	1
-	-	-	-	0.5	0.5	0	0.5	0	0
-	-	-	-	0	0	0	0.5	1.5	0
-	-	-	-	0.5	0	0	0	0	0
-	-	-	-	0	0	0	0	0	0
-	-	-	-	2	4	3.5	5.5	2	4.5

た。また、「鬱」、「怖」、「安」、「怒」では、「こ」、「ちへ」、「いら」、「しゃい」共に同じ高さで変化がないと分類された音声が多かった。

振幅は「喜」、「安」では、「こ」が大きく、「ちへ」が大きく、「いら」が小さく、「しゃい」が大きいという変化に分類された音声が多かった。また、「鬱」、「驚」、「怖」、「怒」では、「こ」が小さく、「ちへ」が大きく、「いら」が大きく、「しゃい」が大きいという変化に分類された音声が多かった。

以上のことから、音声物理量の一文での統計的解析で複数が同じ傾向を持つ音声物理量と、人手で分類した変化で複数が同じ傾向を持つ音声物理量とは、パラメータが異なることが考えられた。すなわち、一文での代表値となる大ざっぱなパラメータと、文の中のダイナミックな変化を示すパラメータとで、感情との関連で重要とされるパラメータが異なるということが考えられた。

4.6 考察

本章で行った声優の感情音声の音声物理量抽出、感情音声聴取実験、一般者の感情音声の採取と統計的分析からは、感情を表現した音声について、発声者と聴取者の感情の不一致という問題と、感情を表現する音声物理量パラメータの特徴が示唆された。

発声者と聴取者の感情の不一致

感情音声聴取実験からは、発声者が音声に表現した感情と、聴取者が音声から認知する感情とは必ずしも一致しないということが考えられた。この問題の背景には、次の3つの原因が推測できるであろう。

考えられる原因の一つは、発声者の個人性の問題である。感情と音声物理量との関連を検討した数多くの研究では結果の不一致が報告されている[69]。発声者の感情表現の技量によって、あるいは発声者の表現の癖によって、発声者が表現した感情が正確に聴取者に伝わらないことが考えられる。本章の感情音声聴取実験では、声優養成学校に通う女性、すなわち音声による感情表現の訓練を受けた者を刺激音声の発声者としたが、発声者の個人性が結果へ影響を与えた可能性は完全には否定できない問題である。

考えられる二点目の問題は、聴取者の個人性の問題である。心理検査の一手法として、曖昧な刺激を提示することで被験者の内界や認知の偏りを投影させる投影法と呼ばれる手法がある [38]。本章の感情音声聴取実験では、文の意味が感情の認知に影響することを避けるため、「こっちへいらっしやい」という意味的には中立的な音声刺激を用いたが、中立的であるためにかえって聴取者の受け取り方が投影法の心理検査と同様に結果に反映されている可能性がある。つまり、聴取者自身の主観的な意味付けを通して音声から感情を認知していることが原因として考えられる。

発声者の表現と聴取者の認知における感情の不一致に関する三点目の原因には、文脈の問題があげられる。本章で行った実験と同様に感情の表現を検討している表情に関する一連の研究に対して、Russell [62] は表情の正確な解釈は表現の瞬間そのものよりも、表現の文脈によって決定されると批判を行っているが、ここでも同様の批判が考えられる。すなわち、「こっちへいらっしやい」という中立的な文を様々に感情を表現して読み分けたものを突然提示されたとしても、聴取者は、その音声に至る文脈が与えられないために正確な感情認知ができないという可能性である。

感情と音声物理量の関連を検討する上では、発声者の表現と聴取者の認知における感情の不一致と、その原因を考慮することが必要と考えられる。

感情を表現する音声物理量について

本章では、声優の感情音声の音声物理量抽出と一般者の感情音声の採取と統計的分析から感情と音声物理量の関連を検討した。

声優による感情音声と一般者の感情音声から抽出した文全体の音声物理量代表値の比較を行ったところ特徴の一致が見られた。すなわち、「怒」では声が大きく発話時間が短く、「鬱」、「安」は声が小さく、「喜」、「鬱」、「安」は発話時間が長くなることから、感情表現の訓練を受けた者と訓練を受けたことのない一般者に共通して見られる傾向であった。文全体の代表値については、振幅と発話時間が感情を表現する上で重要な音声物理量パラメータと考えられる。

また、一般者の感情音声に対して行った音声物理量の変化の分類では、基本周波数は、「喜」、「驚」が同様の変化を示すと分類され、振幅は「喜」、「安」及

び「鬱」，「驚」，「怖」，「怒」が同様の変化を示すと分類された。音声物理量の文の中の変化については，基本周波数と振幅が感情を表現する上で重要な音声物理量パラメータと考えられる。

感情を表現する音声物理量については，一文で一つの代表値に注目した場合と，文中でのダイナミックな変化に注目した場合で，重要な音声物理量パラメータが異なることが分かった。このことから，一文での代表値だけではなく時間的変化を考慮に入れた音声物理量パラメータの抽出が必要であることが示唆されたと考えられる。

4.7 まとめ

調査によって実験で使用する簡便な感情状態尺度を作成した上で，声優が発声した感情音声の音声物理量を抽出し，特徴について検討した。また，感情音声を刺激として聴取実験を行い，発声者の表現した感情と聴取者が認知した感情について検討を行った。さらに，22名の一般者によって発声された感情音声から音声物理量を抽出し，一文での代表値の統計的分析と図から目視で得た変化の分類を行うことで，複数の発声者に共通する感情と音声物理量との関連を検討した。

感情音声聴取実験からは，発声者の表現と聴取者の認知における感情の不一致の現象が認められた。不一致の背景として，発声者の個人性，聴取者の個人性，文脈の問題が考えられた。声優による感情音声と一般者の感情音声の音声物理量における検討からは，感情を表現する音声物理量について感情表現の訓練を受けた者と訓練を受けたことのない一般の人に共通して表れる特徴があることと，重要とされるパラメータが文全体で一つの代表値に注目する場合と文の中のダイナミックな変化に注目する場合とで異なることが分かった。

第 5 章

音声物理量からの抑うつ傾向判定

本章では、感情の一種として悲しみや憂鬱さに関連する抑うつの感情について取り上げ、音声との関連を検討する [86, 81, 83]. 第 4 章では、基本感情を 6 カテゴリーに分け、その全般について感情と音声物理量との関連を検討したが、ここではうつ病など気分障害と関連している抑うつの感情にのみ焦点をあて、自然に音声上に表れる感情について詳しい分析を行うと共に、抑うつ感情が強調されやすい単語を提案し、その選択について検討する.

5.1 背景と目的

抑うつ傾向を持つ人間の音声、及び抑うつの文を読む際の人間の音声についての物理量の特徴について検討し、抑うつ傾向判定を行う上で重要なパラメータの抽出を行う.

精神科の診察やカウンセリングのように人間の心を扱う臨床場面において、患者の多くが訴える心的状態の一つに抑うつ状態がある. 世界で幅広く用いられている精神疾患の診断手引き DSM-IV [2] には、うつ病や抑うつの状態を示す「大うつ病」の診断基準として、抑うつ気分や、活動における興味や喜びの喪失、無価値観、罪悪感といった話の内容から認められるものと共に、食欲、体重、睡眠の変化や精神運動の制止のように行動の諸特徴における症状があげられている. ここでいう精神運動の制止とは、具体的には「会話、思考、体動の遅いこと、応答の前の時間が長くなる、声量、抑揚、会話量、内容の豊かさの減少や無口」と定

義されており、行動や音声の諸特徴に関する症状である。

症状を客観的に捉えるという見地から、うつ病患者の音声に焦点をあてた研究が行われている。Alpert, Pouget & Silva [1] は、うつ病患者と健常者の音声を録音し、流暢さ (fluency) と韻律 (prosody) の側面から比較検討を行っている。ここでは、うつ病患者の音声は、健常者の音声に比べて、声が大き、発話時間が短い、基本周波数が低いという差がみられたと報告されている。また、Ellgring & Scherer [20] は、うつ病患者 16 名の治療前と治療後の音声を、発話時間と基本周波数について比較し、治療前に比べて治療後の患者の発話率 (時間あたりの発話量) が上昇し、間が減り、基本周波数が下がると報告している。さらに、精神運動制止仮説、認知障害仮説、社会感情仮説の 3 つの仮説から検証した結果、音声の特徴は社会感情 (特に悲しみ、不安) の影響が大きいと結論している。

このように抑うつ状態における音声物理量の分析は、徐々に行われるようになってきたが、今までに報告されてきた結果は一定ではない。その理由として Ellgring ら [20] は、音声パラメータが発話率や空白時間、振幅など研究者間によって一致していないことや、対象とする音声の課題が複雑な認知操作を必要とするものとしにくいものが同列に扱われていることをあげている。また、これらの問題のほかに、先行研究では、うつ病患者と対照健常者の音声を録音解析して比較するにとどまっており、たとえば抑うつ状態の診断補助となりえるような、音声からの診断指標の提案や抑うつ判定システムの構築といった実用化を目指した研究はされていないことがあげられる。

これらの問題に対して本研究では、感情を含む音声の音声物理量パラメータとして有用と考えられているものの、Ellgring & Scherer [20] では用いられなかった高さ、大きさ、速さをもとにした音声物理量のパラメータを用いて、音声の分析を行う。また、課題については、Alpert ら [1] が行った数を数える課題よりもより現実状況に近く、Ellgring ら [20] が行った自由発話よりも音声物理量の他者間比較に信頼性がおけると考えられる抑うつ状態測定のための尺度で使われている抑うつ文を用いる。

なお、抑うつ状態の音声特徴を把握するには、実際に医療機関でうつ病と診断されている患者を対象として被験者群を選定することが望ましい。しかし、医療行為を目的とする医療機関では、音声採取に適した実験環境を整えることは難し

く、また、医療機関以外の場所で実験環境を整えたとしても治療に訪れる患者を研究のために移動することも現実的には難しい。そのため、大学生を対象に調査を行い、高抑うつ傾向を持つ群と低抑うつ傾向を持つ群を選定し、音声の採取、分析を行うことで、うつ傾向が高い学生グループと低いグループにおける比較を行う。

以上の方法を用いて、抑うつ傾向判定システムの構築を目標に、抑うつ傾向と音声物理量の関連を見だし、判定システムを試作することを目的とする。具体的には、うつ状態測定尺度であるBDI(Beck Depression Inventory) [27]を用いて高抑うつ群と低抑うつ群を選定し、文読み上げの実験を行う。録音した音声は物理量を抽出し、抑うつ傾向判定に重要な音声物理量パラメータの選定と発声時の抑うつ傾向を良く示す単語としての指標単語の提案を行う。作成された判定システムは情報検索で用いられる再現率適合率をもとに評価される。

5.2 抑うつ傾向判定のための研究の概略

抑うつ傾向判定のための研究の音声の採取、物理量抽出は以下の手順で行う。

1. **高抑うつ群と低抑うつ群の選定** 既存のうつ状態を測定する尺度 (BDI) を用いて調査を行う。調査対象者の中から、BDIの結果(得点)に応じて、抑うつ傾向の高い「高抑うつ群」と傾向が低い「低抑うつ群」を選定し、音声採取実験の実験被験者とする。
2. **音声採取実験課題文の作成** 高抑うつ群、低抑うつ群の2群の間で統計的に有意な差がみられたBDIの項目を、音声採取実験で用いる抑うつ課題文とする。
また、抑うつ課題文に対して、抑うつ的な意味合いが含まれない文を非抑うつ課題文とする。
3. **音声採取実験** 高抑うつ群、低抑うつ群に対して課題文読み上げの音声採取実験を行う。
4. **音声の分析** 採取した音声を計算機で処理し、音声物理量を算出する。得られた音声物理量について統計的処理を行う。

5. 指標単語選定 以下の3つの戦略によりそれぞれ指標単語を選定する。

統計による指標単語選択戦略

抑うつ課題文を文節に分割し、文節ごとに音声物理量を算出し、統計的検定により両群間で有意の差が見られる単語を指標単語とするもの。

聴取による指標単語選択戦略

第三者が両群の音声を聞き比べ、差があると指摘した単語を指標単語とするもの。

専門家による指標単語選択戦略

臨床心理士及び臨床心理学専修の大学院生に対して課題文を書面にて提示し、抑うつ傾向が表れるとされた単語を指標単語とするもの。

6. 指標単語の抽出 それぞれの指標単語選択戦略を用いて重回帰分析を行い、抑うつ傾向判定に重要な指標単語を選択する。

5.3 BDI を用いた音声採取実験被験者、課題文の選択

音声採取実験での被験者となる抑うつ傾向の高い者、低い者を選出し、さらに実験で用いる課題文を選定することを目的として質問紙調査を行う。

5.3.1 BDI について

BDI(Beck Depression Inventory) [7]¹は、抑うつ状態測定のために臨床現場や研究で幅広く用いられている尺度である。日本では、林 [27] が 1988 年に BDI 日本語版を作成している。

この尺度では、質問項目は全 16 項目であるが、それぞれの項目が番号 0 から 3 までの 4 つの文から構成されており、番号が大きくなるに従って抑うつ傾向の高い内容の文となる。たとえば第 5 問は、

[0] 私は自分自身に失望していない。

[1] 私は自分自身に失望している。

¹実際に用いた BDI を付録 (1) に示す

[2] 私は自分自身にいや気がさしている。

[3] 私は自分自身が嫌いだ。

の4文からなっている。

実施の際には、回答者は項目の4つの文を読み、最も自分の気持ちに当てはまる文に○をつける。結果は、項目番号を足すことで算出し、全体の合計で得点が低ければ抑うつ傾向が低い、得点が高ければ抑うつ傾向が高いとされる。なお、合計得点の最低点は0点、最高点は48点である。

5.3.2 調査の実施と被験者群の選定

BDI尺度は、大学生を対象に理系教養科目の授業終了後に連続して実施した。調査に協力した男子大学生104名のうち、回答に記入漏れのない98名について、回答者ごとに選択された項目番号の合計を算出した。調査回答者の中で、得点の高い者から25名、低い者から24名を選出し、それぞれ調査回答者の中で抑うつ傾向が高いことから「高抑うつ群」、抑うつ傾向が低いことから「低抑うつ群」とした。

高抑うつ群の平均得点は16.0(標準偏差3.0)点、低抑うつ群の平均得点は2.1(標準偏差1.5)点であった。なお、高抑うつ群において高く得点される傾向が強かった項目は、5(自己嫌悪) 11(外観への悲観) 12(仕事の遅滞)であった。

Beck [7] は、BDIとうつ病の関連について、BDIの合計平均点が正常は10.9、軽症うつ病は18.7、中等症うつ病は25.4、重症うつ病は30.0と述べている。BDIの結果から、音声採取実験の被験者は、高抑うつ群は低抑うつ群に比べうつ病圏に近い状態にあると考えられるが、本研究では医療機関への通院を必要とするほどの重篤な抑うつ状態の者はいなかった。

5.3.3 課題文の選択

高抑うつ群、低抑うつ群の両群の違いを明確に表す課題を選定するため、BDIの項目ごとに2群のBDI得点のt検定を行った。

t検定の結果、5%以下の水準でBDI得点における群間の有意な差が認められた項目の中で、群間の平均値の差が大きく、かつ群内での標準偏差が小さな3つの項目の文(計12文)を、安定して両群の違いを表している項目と考え、実験で

用いる文読み上げのための抑うつ課題文とした。

また、抑うつのニュアンスを含まない文を非抑うつ課題文とした。最終的に音声採取実験で用いた課題文を表 5.1 に示す。

表 5.1: 音声採取実験で採取した課題文

非抑うつ課題文	
	私は横浜国立大学の学生です
抑うつ課題文 (文頭の数字は課題文番号)	
5-0.	私は自分自身に失望していない
5-1.	私は自分自身に失望している
5-2.	私は自分自身にいや気がさしている
5-3.	自分自身が嫌いだ
11-0.	私は自分の現実よりは悪く見えない
11-1.	私は年をとり魅力を失って見えるのではないか気になる
11-2.	私はだんだん魅力がなくなったように思う
11-3.	私は自分の見かけが見苦しくなってきたと信じている
12-0.	私は以前と同様に仕事ができる
12-1.	何かをしようとする時、前よりも余分な努力をしなければならない
12-2.	何かを始める時、うんと頑張らなくてはならない
12-3.	私は何もしたくない

5.4 音声データの採取実験と分析

5.4.1 音声採取実験

「高抑うつ群」, 「低抑うつ群」両群の音声を採取するための実験を行う。

音声採取実験の被験者は, BDI 調査により選定された「高抑うつ群」, 「低抑うつ群」のうち実験に協力した 30 名 (各群 15 名) である。

実験は, 個室で被験者が計算機のモニターに表示された課題文を読み上げる形式で実施される。抑うつ課題文は被験者ごとにランダムに提示され, 文読み上げの音声は文の提示のタイミングを示すピープ音とともに DAT で録音される²

²実験で使用した機器類は以下である。

録音機材: DAT(Sony TCD D-100), 48KHz サンプリング, 16bit 量子化

マイク: Sony ECM-M957 指向性マイク

計算機では, 録音した音声を WAVE 形式ファイルとして処理する。なお, WAVE 形式への変換には Canopus DA-Port USB を用いる。

なお、実験室は研究棟の静かな個室をあて、実験に必要な機材以外は反響を避けるため置かないようにした。被験者は固定されたイスに座って発声を行い、マイクとの距離は30cmとした。実験の所要時間は1人あたり約10分であった³。

5.4.2 音声物理量のパラメータについて

音声の特徴量としては、高さ、速さ、大きさが基本的である。本研究では、感情を含む音声からの物理特徴量抽出の先行文献 [77] を参考に、以下に示す音声物理量のパラメータを算出する。DAT から計算機への音声データの取り込み作業及び文と文節の切り分け、モーラ数のカウントは手動で行うが、音声物理量のパラメータについてはプログラムを作成し、自動的に算出する。なお、基本周波数の算出には、自己相関関数による方法を用い、窓関数はハミング法を用いる。

フレーム周期は5ms、基本周波数の抽出範囲は50から400Hzとする。また、パワー値(15dB)を閾値とし、閾値に至らないものを無音区間と判断し、該当箇所に関しては音声物理量パラメータを算出しない。

振幅 音声データ波形の振幅を算出する。測定区間内で最も高い値を**最大振幅**、低い値を**最小振幅**、区間の算術平均を**振幅平均**、最高振幅と最低振幅との差を**振幅レンジ**とする。

基本周波数 基本周波数の値を算出する。測定区間内で最も高い値を**最高基本周波数**、低い値を**最低基本周波数**、算術平均を**平均基本周波数**、最高基本周波数と最低基本周波数との差を**基本周波数レンジ**とする。

発話速度 課題文読み上げの際に、1モーラ読み上げに要した時間
(課題文読み上げ総時間(ms) / 課題文総モーラ数) を**発話速度**とする。

5.4.3 パラメータの差の検定と結果

抑うつ課題文と非抑うつ課題文との違いを明らかにするため、課題文発声の音声物理量に被験者群間で相違が認められるか統計的に検討する。

³音声採取実験の際に、再びBDI尺度を被験者に対して実施する方法も考えられたが、調査での得点結果を元に文読み上げの課題文を選択している関係上、再度の実施は不相当と考えられた。本研究では、調査の時点のBDI得点と同一であると考えて音声採取実験を行っている。

全ての音声パラメータに関して、高抑うつ群と低抑うつ群の群間の差の検定としてノンパラメトリック検定（Mann-Whitney 検定）を行う⁴。

各パラメータでの検定の結果を以下に示す。

振幅 課題全てにおいて両群に有意な差は認められなかった。

基本周波数 最大基本周波数では、課題文 5-1 ($U_0 = 56.5, p < .01$)、課題文 11-3 ($U_0 = 55.5, p < .01$)、課題文 12-2 ($U_0 = 60.5, p < .01$) において高抑うつ群の方が低抑うつ群よりも周波数が高いという差が見られた。基本周波数レンジでは、課題 5-3 において高抑うつ群の方が低抑うつ群よりもレンジが大きいという差が見られた ($U_0 = 61, p < .01$)。最低基本周波数、平均基本周波数では有意な差は認められなかった。

発話速度 発話速度では、課題全てにおいて両群に有意な差は認められなかった。

ノンパラメトリック検定の結果で、非抑うつ課題文 0 については被験者群間で差が見られず、他の抑うつ課題文では被験者群間での差が認められたことから、抑うつ課題文を読む際に、抑うつ傾向が高いか低いかという発声者の心理状態が影響を与えていることが考えられる。

5.4.4 重回帰分析による判別に使用する音声パラメータの選定

本研究での音声パラメータのなかで抑うつ傾向を統計的に有意に説明できるパラメータを明らかにするために、二項ロジスティック回帰分析⁵を行う。

⁴ 2 群の母集団 G_1, G_2 に差があるかどうかの検定は以下の手順で行う（Mann-Whitney 検定）。

検定を行う 2 群のケース数をそれぞれ n_1, n_2 とする。2 群を結合して一つの集合とみなして順位づける。各ケースの順位について、 G_1, G_2 に属するものの順位の番号をそれぞれ合計して R_1 と R_2 とし、次の値を求める。

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_1 + 2)}{2} - R_2$$

$$U_0 = \min(U_1, U_2)$$

次に Mann-Whitney 検定表より n_1, n_2, P （有意差）にあたる u を求め、 $U_0 > u$ であれば、両群の値には差がないとされる。

⁵ ロジスティック回帰分析は、従属変数をロジスティック変換することで、カテゴリデータを扱うことが可能となった分析手法である [31]。ロジスティック変換は次式による。

本研究では、群属性（高抑うつ群，低抑うつ群）を従属変数としていたため通常の重回帰分析ではなくロジスティック回帰分析を用う。音声パラメータを独立変数，群属性を従属変数として， p について解き，閾値 0.5 で判断する。なお，分析のパラメータ投入にはステップワイズ法を用いる。

結果，群属性については，課題文 11-1，課題文 11-3 の最高基本周波数と，課題文 12-1 の平均基本周波数，課題文 12-2 の基本周波数レンジ，課題文 11-2，課題文 11-3，課題文 12-3 の振幅レンジのパラメータが重要であることが明らかになった。これらのパラメータを用いた回帰式は次のようになる。

$$\begin{aligned} \log \frac{p}{1-p} = & 7.11 \times 11-1 \text{最高基本周波数} \\ & - 22.6 \times \text{課題文}11-2 \text{最高振幅} \\ & - 7.08 \times \text{課題文}11-3 \text{最高基本周波数} \\ & + 11.29 \times \text{課題文}11-3 \text{振幅レンジ} \\ & + 3.84 \times \text{課題文}12-1 \text{平均基本周波数} \\ & - 1.34 \times \text{課題文}12-2 \text{基本周波数レンジ} \\ & - 6.27 \times \text{課題文}12-3 \text{振幅レンジ} \\ & + 21.16 \end{aligned} \quad (5.1)$$

5.1式によって判定器を作成し，その判定の精度について評価を行う。評価には，回帰式を用いて，高抑うつと判定される際の再現率，適合率，F 値，低抑うつと判定される際の再現率，適合率，F 値を算出する。算出式は次式である。

$$\text{再現率 (R)} = \frac{\text{正しく判定した事例数}}{\text{実際の事例数}}$$

$$\text{適合率 (P)} = \frac{\text{正しく判定した事例数}}{\text{判定した事例数}}$$

$$\text{F 値} = \frac{2PR}{R+P}$$

$$p \mapsto \log \frac{p}{1-p}$$

ロジスティック変換された値を従属変数として次式重回帰分析を行う。

$$\log \frac{p}{1-p} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \beta_0$$

ただし、ここでは判定器により低抑うつと判定される場合と高抑うつと判定される場合の2つの場合があり、研究の目的からも両者ともに重要であると考えられるため、次式を用いる。

$$\text{低抑うつ再現率} = \frac{\text{正しく低抑うつ群と判定した事例数}}{\text{実際の低抑うつ群の事例数}}$$

$$\text{低抑うつ適合率} = \frac{\text{正しく低抑うつ群と判定した事例数}}{\text{低抑うつ群と判定した事例数}}$$

$$\text{低抑うつ F 値} = \frac{2 \times \text{低抑うつ再現率} \times \text{低抑うつ適合率}}{(\text{低抑うつ適合率} + \text{低抑うつ再現率})}$$

$$\text{高抑うつ再現率} = \frac{\text{正しく高抑うつ群と判定した事例数}}{\text{実際の高抑うつ群の事例数}}$$

$$\text{高抑うつ適合率} = \frac{\text{正しく高抑うつ群と判定した事例数}}{\text{高抑うつ群と判定した事例数}}$$

$$\text{高抑うつ F 値} = \frac{2 \times \text{高抑うつ再現率} \times \text{高抑うつ適合率}}{(\text{高抑うつ適合率} + \text{高抑うつ再現率})}$$

これらの式を用いて、訓練する事例とテストする事例に同じ事例を用いたクローズドテストと、事例を分割し訓練する事例とテストする事例に異なる事例を用いたオープンテストを行うことで評価を行う。なお、オープンテストでは、2分割の交差検定を用いる。

5.4.5 結果と考察

ロジスティック回帰分析では、重要なパラメータとして、課題文 11-1、課題文 11-3 の最高基本周波数と、課題文 12-1 の平均基本周波数、課題文 12-2 の基本周波数レンジ、課題文 11-2、課題文 11-3、課題文 12-3 の振幅レンジが抽出された。この結果からは、文全体の代表値を用いた分析では発話速度には差が見られないものの、基本周波数と振幅において抑うつ傾向の高低と音声物理量に間連があることが示唆されている。

クローズドテストでは、両判定における再現率、適合率、F 値はそれぞれ 100% であった。オープンテストとして二分割交差検定を行ったところ、高抑うつ群再現率は 54%、適合率は 58%、F 値は 56%、低抑うつ群再現率は 60%、適合率は 55%

%, F 値は 57% であった.

ロジスティック回帰分析の結果としては, クローズドテストによる評価が非常に良い. しかし評価が 100% という値, つまりミスがっさいないということは実際の判定ではまずあり得ない. ここでは判定式が与えられた事例に対して機械学習でいう過学習を行ってしまったと考えられる. オープンテストの評価がクローズドテストの評価に比べて極端に低いということも, 判定式が過学習を起こしていることを裏付けているのではないかと考えられる.

5.5 測定範囲の細分化

今までの分析では, 文全体の発声における物理量の代表値をそれぞれのパラメータの値として分析を行ってきた. しかし, 一文で一つの代表値を用いると, 例えば特定の単語に抑うつ傾向が強く表れていたとしても, 文中の他の個所の音声物理量の影響によって特徴が分かりにくくなる可能性がある. また, 将来的な応用の一方法として, 自由発話の中での判定を考えた時に, 特定の文を判定に用いるよりも, 抑うつ傾向が表れやすい単語を用いた方が実用的と考えられる. このため, 音声の中の箇所を特定して細分化した分析が必要と考えられる.

そこで, 文という大きな単位よりもより小さな単位として, 文節のレベルで今までと同様の統計的分析を行う. なお, 文全体でのパラメータの分析において, 非抑うつ課題文ではいずれのパラメータでも被験者群間の差が見られなかったため, ここでは抑うつ課題文のみを分析の対象としている.

5.5.1 文節レベルでの分析の結果

すべての文節におけるすべての音声パラメータを独立変数として分析を行った結果, 課題文 5-1 第 3 文節の発話速度, 課題文 5-2 第 3 文節の振幅レンジ, 課題文 11-3 第 6 文節の基本周波数レンジ, 課題文 12-2 第 2 文節の振幅レンジが重要なパラメータとして抑うつ傾向の高低を説明できることがわかった. これらのパラメータを用いた回帰式は次のようになる.

$$\log \frac{p}{1-p}$$

$$\begin{aligned}
&= -3.18 \times \text{課題文5 - 1第3文節発話速度} \\
&\quad + 3.21 \times \text{課題文5 - 2第3文節振幅レンジ} \\
&\quad - 3.69 \times \text{課題文11 - 3第6文節基本周波数レンジ} \\
&\quad + 11.19 \times \text{課題文12 - 2第2文節振幅レンジ} \\
&\quad + 64.27 \tag{5.2}
\end{aligned}$$

5.2式により判定器を作成し、再現率と適合率を算出した。クローズドテストでは、両判定における再現率、適合率はそれぞれ100%であった。オープンテストとして二分割交差検定を行ったところ、高抑うつ群再現率は54%、適合率は53%、F値は54%、低抑うつ群再現率は48%、適合率は48%、F値は48%であった。

5.6 指標単語選択戦略

上述した分析ではすべての抑うつ課題文について文節ごとに分析を行った。

しかし全ての文節についてのパラメータを分析に用いることは、機械学習での訓練を行う関係上、パラメータの数に比べて訓練データの数が相対的に少なくなってしまう学習の精度が下がる可能性がある。

また、実際に分析を行った結果をみると、判定式を用いたクローズドテストでの再現率、適合率が100%であり、過学習を起こしている可能性がある。

そこで、ここではより効果的に判定を行うために、現象を代表できる指標単語を導入する。指標単語を選定するために、次に述べる3つの戦略を検討する。

3つの戦略は、比較を行うことを目的として、それぞれ性質の異なるものを選ぶ。一つは理論的、統計的に裏付けを行い指標単語を選定する**統計による指標単語選択戦略**であり、一つは計算機で抽出した音声物理量に対しての聴覚判断に基づく**聴取による指標単語選択戦略**（参考）であり、もう一つは専門家（臨床心理士等）による文からの抑うつ単語選定の**専門家による指標単語選択戦略**である。

これら3つの戦略による指標単語選択との比較のために、ランダムに選択したパラメータを用いた分析の結果もあわせて示す。

5.6.1 統計による指標単語選択戦略

統計的手法により単語を選択する。前述の音声物理量パラメータを文節ごとに算出して、高抑うつ群、低抑うつ群の両群間ノンパラメトリック検定 (Mann-Whitney 検定) を行う。検定の結果、音声パラメータで被験者群間で差がみられた 28 文節を用いる。

5.6.2 聴取による指標単語選択戦略

第三者が音声データを聴取し判定を行う実験により文節を選択する。

統計による指標単語選択戦略、及び専門家による指標単語選択戦略との比較のために、一般人が聴覚によって単語に表れた抑うつ傾向を判断する方法を参考として取り入れる。ここでは、音声採取実験に参加していない第三者 3 名によって選択を行う⁶。聴取者は、両群の音声データを聴き比べ、抑うつ感情が表れていると音声から判断された順に文節に順位をつける。順位をつける際の判断は、聴取者が、抑うつ傾向を聴覚から経験的に感じ取る、いわば一般人の直感を基準としている。なお、聞き比べる音声データは、両群で同一の文を読み上げたデータとするが、高抑うつ群と低抑うつ群の音声データの組み合わせは、聴取を行う者によってランダムになっている。音声データが高抑うつ群と低抑うつ群どちらのものかについては、文節の順番判断に影響はないと考えられたため、聴取者にも分かるようになっている。このようにして得られた 3 人の結果を整理し、聴取によって最も抑うつが表れていると判断された文節をそれぞれ 1 文につき 1 つ選択する。

⁶聴取による指標単語選択戦略では、臨床心理士等のうつ状態に関する専門的な知識を持つ者ではない、一般人を対象としている。ここでは、専門家との判定の正確さを比較するよりも、一般人の聴覚の判定との比較を目的としたからである。ただし、3 名と対象者が少なく、実験対象者の主観性が結果へ影響する可能性もあるため、参考として掲載している。

5.6.3 専門家による指標単語選択戦略

臨床心理士として医療関係などの現場で仕事を行っている者、および臨床心理学を専修している大学院生 14 名を対象に調査を行い単語を選択する⁷。調査では、12 の抑うつ課題文を文節ごとに区切りをいれた用紙が配られ、もっとも抑うつがよく表れると考えられる箇所に線を引くように求められる。14 名の結果を整理し、もっとも抑うつが表れるとされた文節をそれぞれ 1 文につき 1 つ選択する。

5.6.4 比較のためのランダム選択

上記の 3 つの指標単語選択戦略との比較のためにランダムに文節のレベルでのパラメータを選択する。ほかの選択戦略とパラメータ数などの条件が近くなることを考慮して、14 のパラメータを選択する。

5.7 それぞれの戦略による分析結果

それぞれの戦略によって選択された指標単語について、二項ロジスティック回帰分析を行う。

5.7.1 統計による指標単語選択戦略による実験結果

統計による指標単語候補についてロジスティック回帰分析を行った結果、課題文 5-1 第 3 文節「失望して」の発話速度、課題文 11-3 第 1 文節「私は」の最低振幅、課題文 12-2 第 2 文節「始める時」の振幅レンジが重要なパラメータとして抑うつ傾向の高低を説明できることがわかった。これらのパラメータを用いた回帰式は次のようになる。

$$\begin{aligned} \log \frac{p}{1-p} \\ = -0.11 \times \text{課題文5-1第3文節発話速度} \end{aligned}$$

⁷なお、専門家はうつ病患者を想定して文節を選択している可能性があるが、選択を行う文節が一文につき数個と少ないこと、「指標単語」という代表値を得るための方法であることから、選択された結果は専門家がどのような対象を想定して選択を行っているにしろ、結果には影響ないと考えられる。

$$\begin{aligned}
& + 0.32 \times \text{課題文11 - 3第1文節最低振幅} \\
& + 0.50 \times \text{課題文12 - 2第2文節振幅レンジ} \\
& - 14.33 \qquad (5.3)
\end{aligned}$$

5.3式により判定式を作成し、再現率と適合率を算出した。クローズドテストでは、両判定における再現率適合率はそれぞれ87%であった。オープンテストとして二分割交差検定を行ったところ、高抑うつ再現率は57%、適合率は63%、F値は60%、低抑うつ再現率は64%、適合率は59%、F値は60%であった。

5.7.2 聴取による指標単語選択戦略による実験結果

聴取による指標単語選択戦略についてロジスティック回帰分析を行った結果、課題文11-3第1文節「私は」の最低振幅、課題文12-0の第1文節「私は」の基本周波数レンジ、課題文12-1の第1文節「私は」の基本周波数レンジと最高振幅、課題文12-2の第3文節「うんと」の振幅レンジと発話速度が重要なパラメータとして抑うつ傾向の高低を説明できることがわかった。これらのパラメータを用いた回帰式は次のようになる。

$$\begin{aligned}
& \log \frac{p}{1-p} \\
& = 57.82 \times \text{課題文11 - 3第1文節最低振幅} \\
& \quad - 1.55 \times \text{課題文12 - 0第1文節基本周波数レンジ} \\
& \quad + 7.62 \times \text{課題文12 - 1第1文節基本周波数レンジ} \\
& \quad - 36.25 \times \text{課題文12 - 1第1文節最高振幅} \\
& \quad + 56.59 \times \text{課題文12 - 2第3文節振幅レンジ} \\
& \quad - 11.09 \times \text{課題文12 - 2第3文節発話速度} \\
& \quad - 162.97 \qquad (5.4)
\end{aligned}$$

5.4式により判定器を作成し、再現率と適合率を算出した。クローズドテストでは、両判定における再現率適合率はそれぞれ100%であった。オープンテストとして二分割交差検定を行ったところ、高抑うつ再現率は54%、適合率は45%、F値は49%、低抑うつ再現率は32%、適合率は39%、F値は35%であった。

5.7.3 専門家による指標単語選択戦略での実験結果

専門家による指標単語選択戦略についてロジスティック回帰分析を行った結果、課題文 5-2 第 3 文節「いや気が」の最低振幅、課題文 5-3 第 3 文節「嫌いだ」の振幅レンジ、課題文 11-0 第 5 文節「見えない」の最低振幅、課題文 11-1 第 8 文節「気になる」の平均基本周波数、課題文 11-2 第 4 文節「なくなったように」の基本周波数レンジ、課題文 12-0 第 5 文節「できる」の最高基本周波数が重要なパラメータとして抑うつ傾向の高低を説明できることがわかった。これらのパラメータを用いた回帰式は次のようになる。

$$\begin{aligned} \log \frac{p}{1-p} &= -12.32 \times \text{課題文5-2第3文節最低振幅} \\ &\quad - 11.4 \times \text{課題文5-3第3文節振幅レンジ} \\ &\quad + 23.50 \times \text{課題文11-0第5文節最低振幅} \\ &\quad + 7.56 \times \text{課題文11-1第8文節平均基本周波数} \\ &\quad - 3.48 \times \text{課題文11-2第4文節基本周波数レンジ} \\ &\quad + 2.14 \times \text{課題文12-0第5文節最高基本周波数} \\ &\quad - 1119.87 \end{aligned} \tag{5.5}$$

5.5式により判定器を作成し、再現率と適合率を算出した。クローズドテストでは、両判定における再現率適合率はそれぞれ 100% であった。オープンテストとして二分割交差検定を行ったところ、高抑うつ再現率は 66%、適合率は 52%、F 値は 58%、低抑うつ再現率は 39%、適合率は 54%、F 値は 45% であった。

5.7.4 比較のためのランダム選択での実験結果

ランダムに選ばれたパラメータについてロジスティック回帰分析を行った結果、課題文 11-0 第 2 文節「自分の」の最高基本周波数が重要なパラメータとして抑うつ傾向の高低を説明できることがわかった。このパラメータを用いた回帰式は次のようになる。

$$\log \frac{p}{1-p}$$

$$= 0.05 \times \text{課題文11 - 0第2文節最高基本周波数}$$

$$- 6.80 \quad (6)$$

(6) 式により判定器を作成し、再現率と適合率を算出した。クローズドテストでは、両判定における再現率適合率はそれぞれ67%であった。オープンテストとして二分割交差検定を行ったところ、高抑うつ再現率は39%、適合率は45%、F値は42%、低抑うつ再現率は53%、適合率は47%、F値は49%であった。

5.7.5 指標単語選択戦略の結果について

3つの戦略によって指標単語候補を選択し、判定器を作成した。3つの戦略による結果の一覧を図5.1に示す。なお、複数の戦略によって指標単語として選択された文節およびパラメータは課題文11-3第1文節「私は」の最低振幅のみである。

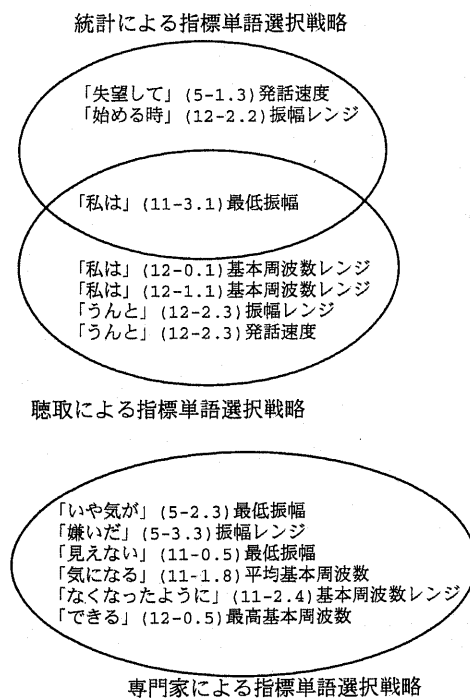


図 5.1: ロジスティック重回帰分析にて選択された指標単語
(括弧内の数字は前が課題文番号, 後が文節である)

5.8 結果と考察

5.8.1 指標単語について

3つの戦略によって指標単語候補を選択し、判定器を作成し、評価を行った。

ランダムに選んだパラメータを用いて作成した判定器と、それぞれの指標単語選択戦略を比べると、統計による指標単語選択戦略については低抑うつ判定高抑うつ判定ともに、統計による指標単語選択戦略の方が評価が良い。これは統計による指標単語がランダムに選んだよりも有用に抑うつ傾向を判定することを示唆している。なお、専門家による指標単語選択戦略と聴取による指標単語選択戦略は低抑うつ判定についてはランダムに選んだパラメータを使用した方の評価が良いが、高抑うつ判定に関しては、専門家による指標単語選択戦略と聴取による指標単語選択戦略の方が評価が良い。これは、専門家による指標単語選択戦略および聴取による指標単語選択戦略が、ともに、抑うつが低いことよりも抑うつが高いことについての判別に有用であることを示していると考えられる。

指標単語選択戦略ごとに選ばれたパラメータの相違を考えてゆくと、専門家が課題文を読むことから抑うつ傾向を判断した専門家による指標単語選択戦略での重要なパラメータと、物理量に差が見られた統計による指標単語選択戦略での重要なパラメータとは異なっていた。さらに、専門家による指標単語選択戦略は、統計による指標単語選択戦略よりも再現率、適合率の両方において評価が低かった。

また、音声を直接聞きながら聴取者が音声の箇所を選んだ聴取による指標単語と、統計による指標単語とは、重なっている部分もあるが、大部分は異なっていた。聴取による指標単語選択戦略は、統計による指標単語選択戦略よりも再現率、適合率の両方において評価が低かった。

専門家による指標単語選択戦略と統計による指標単語選択戦略との重要なパラメータの相違からは、専門家が文面から感じ取る抑うつ傾向と、音声に現れる抑うつ傾向とはやや違いが見られることが考えられる。聴取による指標単語選択戦略は、音声を聞きながら抑うつが表れている箇所を判断するという意味で、音声だけではなく文の意味が関連している課題であり、そのために統計による指標単

語選択戦略で抽出された重要なパラメータと関連している箇所と、全く関連のない箇所とが現れたのではないかと考えられる。

本研究で最も判定に有効な評価が現れた単語は、統計による指標単語選択戦略から選ばれたいくつかのパラメータであったが、文と音声に関する知見についてはさらなる研究と検討が必要とされる問題と考えられる。

5.8.2 抑うつ判定について

すでに述べたように、うつ病の症状の一つである精神運動制止では、会話の遅さ、応答の前の時間の延長、声量、抑揚の減少が挙げられている。このことからうつ状態が音声上の物理的な特徴として表れることが示唆されている。高抑うつ群と低抑うつ群の文読み上げの音声を検討した本研究で、高抑うつ傾向を持つ者の音声上の重要なパラメータとされたのは、ある種の文節における基本周波数と抑揚の上昇であった。

Ellgring et al. [20] は、うつ病患者 16 名の治療前と治療後の音声を時間と基本周波数の観点から比較し、治療後に比べ治療前の音声の基本周波数が高いという結果に着目した。精神運動制止の仮説だけでは音声に現れる特徴を説明するのには不十分と考え、うつ病患者の音声へ影響を与えるメカニズムとしての社会感情の重要性を主張した。つまり、抑うつ状態の音声は、精神運動制止の現象だけではなく、悲しみや不安の心理が反映されたものと捉えている。

うつ状態と関連のある感情としては、不安、悲しみ、寂しさ、焦燥感が論じられてきた [92]。Izard [35] は、感情を完全に独立なものというよりも他の感情と結び付き関連して行くものと捉え、特にうつ状態は背景に複数の感情が喚起されている複雑な感情であると述べている。うつ状態の背景にある不安感情の音声特徴としては基本周波数の上昇と振幅の増大が述べられており、悲しみ感情の音声特徴としては基本周波数の低下と振幅の減少が報告されている [69, 74]。

実験では抑うつを表す者として BDI で選定した抑うつ傾向の高い者を音声採取の対象とした。対象がうつ病患者ではなく、一般の抑うつ傾向の高い者であったため、精神運動制止の現象よりも、うつ状態やうつ傾向の背景にある不安や悲しみが、音声特徴量として強く反映された可能性がある。さらに、課題となった文の性質が影響を与えた可能性も考えられる。つまり、抑うつ傾向が高い者が、抑

うつ文や抑うつを連想させる単語を発声する際に、抑うつの連想を回避しようと緊張をしたため、発声器官での筋緊張が起こり、基本周波数と抑揚の上昇が起きたことも考えられるであろう。

結果からは、抑うつ傾向が高い状態にある人間が発声する音声は、うつ状態と関連する複数の感情や、抑うつの文を読む際の緊張等が音声物理量に反映される可能性があると考えられた。なお、今回は音声採取実験の対象が一般の人であるため、即座にうつ病や抑うつ状態の判定として結果を応用することは難しい。実際にうつ病の患者や、抑うつ状態の心理判定に対して用いるには、抑うつ傾向と音声物理量との関連が見られたことを土台として、更なる研究と改良が必要とされると考えられる。

5.9 まとめ

抑うつ傾向判定システム構築を目標に、高抑うつ群、低抑うつ群を対象とした文読み上げの実験を行った。文読み上げの課題文として用いたBDI(Beck Depression Inventory)のうち6文は、音声における最高基本周波数、平均基本周波数、基本周波数レンジ、振幅レンジのパラメータで抑うつ傾向との関連が見いだされた。

さらに、抑うつ傾向判定のための文読み上げの指標単語選定を目的に、3つの方法で選んだ指標単語選択戦略について、ロジスティック重回帰分析による音声物理量の重要なパラメータの抽出を行った。実験と分析の結果、統計指標単語選択戦略が他の2つの指標単語選択戦略よりも判定に対して有効な指標単語であることが分かった。

しかし、学生を対象にBDIを用いて協力者を選定し、抑うつ傾向と音声物理量の関連と、抑うつ傾向の指標単語を提案しているため、即座にうつ状態での臨床で応用することは難しい。抑うつ状態を訴える患者からの実験協力者確保は現実的に困難ではあるが、今後は、医療機関等との連携をもとに患者を対象に研究を進展させてゆくことが課題である。

また、3つの指標単語選択戦略の結果の相違から、専門家は低抑うつ傾向よりも高抑うつ傾向を持つ者に対する判別精度が良いと考えられたが、今後判定器

を改良してゆく過程で、専門家等人間の判断に近づくよう高抑うつ傾向の者に対してのみ精度をあげて行くか、あるいは専門家が判別をしなかった低抑うつ傾向の人間に対しても精度をあげて行くか、その方針についても考えることを課題としたい。

第 6 章

感情豊かな読み上げのための音声パラメータ

本章では、聴取実験に重点を置いて感情と音声物理量の関連を論じる。第 4 章では基本感情を 6 カテゴリに分け感情が表現された音声の音声物理量抽出と聴取実験での関連を検討し、第 5 章では抑うつ感情にのみ焦点をあて音声物理量の抽出と文の性質との関連を論じたが、第 6 章では、計算機による感情音声の発声を目指し、様々に変化する音声から感情がどのように聴取されるかという視点によって感情と音声物理量の関係を論じる。得られた感情と音声物理量に関する情報は、合成音声に付与し、聴取実験により評価を行う。

6.1 背景と目的

感情豊かな読み上げ音声の作成を目標として、有効な音声パラメータの検討を行う [85].

音声と感情の関連は、音声に人間の感情がどのように表れるか、また、どのようにして音声上へ人工的に感情を付与できるかという見地より、古くから論議されている問題である。

音声に表れた人間の感情を捉えようとする研究は、主に発声者が感情を表現した音声から音声物理量を抽出し、その関連を検討する方法で行われてきた [77, 50, 33, 34]. この方法からは一定の知見が得られているが、分析の対象となる音声を発声する、発声者の表現力等の個人性が結果へ影響を与える可能性が高い。

一方、音声合成の応用分野においては、合成音声の自然性向上や、音声に障害

を持つ人への補助を目的として、音声上へ人工的に感情を表現する韻律を付与する研究が行われてきた [72]. 感情音声合成の手法の一種として、感情を表現した音声の大規模なデータベースを構築し、そこからターゲットとなる音声と感情を合成するコーパスベースの合成法があり、有効性が確認されている [46, 65, 30]. しかしこのアプローチでは、感情認知の精度はデータベースの出来に大きく影響される上、データベース構築には大きな負担が強いられるという問題がある.

感情音声合成のもう一方の手法は、人間の発声機構や韻律的特徴を近似したさまざまな規則を用いて音声を作成するフォルマント合成を用いた手法である. フォルマント合成による感情音声合成では、音声物理量を様々に変化させた音声を刺激として聴取実験を行い、得られた音声物理量と聴取者が認知する感情との関連情報を用いて合成を行う方法が用いられる [12, 51]. フォルマント合成による合成音声は人工的な不自然なものとして聞こえるとの反論はあるものの、聴取実験に重点を置くこの方法は、音声と感情との対応を把握する上でも発声者の個性や、研究方法の効率といった問題への解決法と考えられている.

人間同士のコミュニケーションを考えたとき、感情的情報が内包される場合はミス・コミュニケーションが生じやすいと言われている [100]. 感情を伝えるために、文字言語では、絵文字、色彩やフォントを用いて、文の中でも箇所を限定して感情をより明確に表現することが日常的であり効果を得ているが、音声言語を作成する感情音声の合成では、箇所を限定しての変化の試みはなく、文全体の音声物理量を変化させることで感情を付与している. 感情が適切に伝わるということの主眼におくと、文の中でも特に感情が表れている箇所に焦点を当て、その箇所に感情を表す音声物理量を修正することで感情がより正確に伝わる可能性がある.

そこで、読み上げの音声に豊かな感情を付与することを目的として、パラメータを様々に変化させた刺激音声を作成して聴取実験を行い、被験者が認知した感情という観点から感情が感じられる音声を選定する. 選定された音声は、物語（「かちかち山」）の読み上げ音声を対象に音声物理量を修正し、聴取実験を行うことで有効性を評価する.

6.2 音声パラメータ検討のための研究の概略

音声パラメータ検討のための研究は以下の手順で行う。

1. **基本感情の選定及び感情尺度作り** 基本感情の選定及び感情尺度作りを目的に、感情に関する言葉についての調査を行う。因子分析によって結果を整理し、本研究でターゲットとする基本感情、及び、音声聴取実験で用いる感情尺度を作成し、音声から感じられる感情についても評定を行う。
2. **モデル刺激の作成** 演劇経験者を対象に音声採取実験を行う。録音音声を用いて刺激を作成する。
3. **モデル刺激を用いた音声聴取実験** 作成した刺激を用いて音声聴取実験を行う。被験者は、感情尺度にて音声から感じられた感情の評価を行う。感情尺度の回答を整理し、感情が感じられる刺激を選定する。
4. **モデル刺激からの音声物理量の抽出** モデル刺激から音声物理量を抽出する。また、基本周波数の変化を表現する方法として、線形関数による波形近似を導入する。
5. **指標単語の音声物理量を修正した物語読み上げ** 物語（「かちかち山」）読み上げのテキスト合成音声の指標単語に、選定された刺激の音声物理量を付与する。音声聴取実験を行い、結果を検討する。

6.3 基本感情の選定及び感情尺度作り

本研究では、実験に先駆けて調査を行うことにより、人間の持つ感情の因子構造を捉え、その感情を実験で対象とする基本感情と定義し、ここで捉えた感情因子を評定する尺度も同時に作成する。さらに、どのような感情が音声に表れると思うかそのイメージについても調査を行う。

6.3.1 方法

調査対象者 大学生を対象として調査を行う。学生 227 名に調査を実施し、分析に用いた有効回答数は 88 部（男性 48 名、女性 40 名、平均年齢 21.8 歳）で

あった。

調査内容 基本感情調査では、感情に関する言葉として、「日本語語彙大系」 [55] の中で属性が「感情」である 53 語を項目に用いる。「あなたは今、どのような感情をもっていますか」という質問に続けて感情に関する言葉が並べられる。調査対象者は 5 件法（「あてはまらない」、「あまりあてはまらない」、「どちらでもない」、「少しあてはまる」、「あてはまる」）で回答を行う。音声に表れる感情表現については、「あなたは、人の声にどのような感情が含まれていると思いますか」という質問に続けて同様の項目が並べられる¹。

6.3.2 結果

調査回収後、「あてはまらない」を 1 点、「あまりあてはまらない」を 2 点、「どちらでもない」を 3 点、「少しあてはまる」を 4 点、「あてはまる」を 5 点と得点化し、基本感情調査全項目について因子分析（主因子法，promax 回転）を行うことで結果を整理する。いずれの因子にも負荷量が低い項目、複数の因子に負荷量が高い項目を除き繰り返し分析を行った結果、4 因子 12 項目にて安定した因子が抽出された。結果を表 6.1、表 6.2 に示す。

第 1 因子に負荷量の高い項目は「楽しみ」、「喜び」、「満足」から構成され「楽しさ」に関する因子と解釈された。第 2 因子に負荷量の高い項目は「悪意」、「嫌悪」、「憎悪」から構成され「怒り」に関する因子と解釈された。第 3 因子に負荷量の高い項目は「狼狽」、「悲しみ」、「心配」から構成され「悲しさ」に関する因子と解釈された。第 4 因子に負荷量の高い項目は「鎮静」、「落ち着き」、「信用」から構成され「落ち着き」に関する因子と解釈された。

音声聴取実験では、4 因子 12 項目を感情尺度として用い、それぞれ 5 件法で回答を求める²。尺度の信頼性検討のため、各因子における Cronbach の α 係数を算出する（表 6.1）。全ての因子にてほぼ高い信頼性が確認された。

第 1 因子「楽しさ」は、第 4 因子「落ち着き」と相関がかなり高く、また、第 2 因子「怒り」は第 3 因子「悲しさ」と相関がかなり高い。対して第 4 因子「落

¹実際に用いた調査用紙を付録 (2) に示す。

²実際に用いた感情尺度を付録 (3) に示す。

表 6.1: 基本感情調査の因子分析結果 (promax 回転によるパターン行列)

因子	質問項目	1	2	3	4
楽しさ	楽しみ	0.806	0.102	-0.137	-0.078
	喜び	0.783	0.010	0.156	-0.228
	満足	0.654	-0.066	-0.122	0.205
怒り	悪意	0.063	0.861	-0.115	-0.018
	嫌悪	0.130	0.687	0.167	0.044
	憎悪	-0.093	0.647	0.018	0.069
悲しさ	狼狽	-0.026	-0.042	0.843	-0.027
	悲しみ	-0.101	0.193	0.566	0.036
	心配	0.053	-0.043	0.420	-0.069
落ち着き	鎮静	-0.220	0.111	-0.044	0.725
	落ち着き	0.052	-0.037	-0.158	0.628
	信用	0.348	-0.076	0.248	0.525
Cronbach の α 係数		0.65	0.77	0.65	0.65

表 6.2: 基本感情調査 4 因子の因子相関行列

因子	1(楽しさ)	2(怒り)	3(悲しさ)	4(落ち着き)
1	1.000	-0.187	0.006	0.431
2	-0.187	1.000	0.559	-0.287
3	0.006	0.559	1.000	-0.264
4	0.431	-0.287	-0.264	1.000

「落ち着き」は第2因子「怒り」や第3因子「悲しさ」とは弱い負の相関の関係にある。このことから、「楽しさ」は「落ち着き」と共に生起し、「怒り」と「悲しさ」は「落ち着き」と共存しない感情と考えられる。

なお、因子抽出後の負荷量平方和は4因子の累積で52.6%であり、調査項目である「日本語語彙大系」の「感情」に関する属性のうち半数がこの4つの因子として抽出された感情に収束されると考えられる。本研究では「楽しさ」、「怒り」、「悲しさ」、「落ち着き」の4因子を基本感情と定義する。

音声から感じられる感情調査については、同様の手順で回答を得点化し、調査対象者の平均点を算出する。前述の調査により基本感情として抽出された4因子

表 6.3: 音声から感じられる感情調査結果

因子	項目	平均得点
楽しさ	楽しみ	4.18
	喜び	4.32
	満足	3.85
怒り	悪意	3.62
	嫌悪	4.05
	憎悪	3.84
悲しさ	狼狽	3.98
	悲しみ	4.07
	心配	3.85
落ち着き	鎮静	3.47
	落ち着き	3.22
	信用	3.15

平均得点：
 「あてはまらない」 1点
 「あまりあてはまらない」 2点
 「どちらでもない」 3点
 「少しあてはまる」 4点
 「あてはまる」 5点
 回答者全員の平均点

における結果を表 6.3に示す。

「楽しさ」, 「怒り」, 「悲しさ」の因子については, 因子全体の平均が 3.5 を超えており, これらの感情が音声から感じられると考えられていることが分かる。

「落ち着き」の因子は平均が 3 に近く他の 3 因子に比べ, 比較的音声からは感じられないと考えられていることが分かる。

6.4 モデル刺激の作成

合成音声への音声物理量修正の見本とする刺激をモデル刺激と呼ぶ。また, 文の中でも特に感情が表れている箇所を指標単語と呼ぶ。モデル刺激は, 音声合成段階で指標単語における音声物理量を修正するモデルとなる。

音声合成には Visual Speech CreatorII³の韻律修正機能を用いるため, 聴取実験で用いるモデル刺激は, 合成段階での操作と関連している必要がある。また, 感情と音声の関係を捉える上では聴取実験における音声の自然性が重要である。そこで, モデル刺激作成には, 音声物理量のパラメータ変化を限定して発声した音声を採用する必要がある。

音声聴取実験の段階では, 指標単語が出現する箇所を想定して, モデル刺激を文の一部に組み込む必要がある。また, 様々な単語が指標単語となった場合の一

³NTT-IT の音声合成編集ツールである。韻律修正機能では, いったん合成した音声について, 基本周波数, 抑揚, 音量, 速度等を調整可能である。特に, 基本周波数と音量については, 一音素につき数箇所の時点での物理量を設定することにより, 変化を自由に操作できる。

表 6.4: 実験 A で用いた発声者への課題の指示

No	声の高さ	セリフのテンポ
1	高い声で	素早く
2	低い声で	素早く
3	高いから低い声に	素早く
4	低いから高い声に	素早く
5	高い声で	ゆっくりと
6	低い声で	ゆっくりと
7	高いから低い声に	ゆっくりと
8	低いから高い声に	ゆっくりと

般化を考慮し、文やモデル刺激は言葉が意味を持たないものを選定する必要がある。以上の点を考慮して、モデル刺激と音声聴取実験の刺激音声を作成する。

6.4.1 音声採取実験

発声者 演劇経験がある男性 3 名を対象とする。

課題 発声の課題は以下に述べる実験 A, B の 2 つである。なお、実験 A の発声課題については、発声者によって課題の順番をランダムに並び替える。

1. 実験 A の課題

「あ」、「い」、「う」、「え」、「お」の 5 つの母音について、発声者は声の高さ、声の長さを変化させ発声を行う。声の高さは高い声、低い声の組み合わせ 4 パターンであり、セリフのテンポは素早く、ゆっくりとの 2 パターン、計 8 パターンである (表 6.4)。

2. 実験 B の課題

物語「かちかち山」の数カ所を朗読の課題とする。感情を交えずに読み上げた後に、感情を表現して読み上げる。

手続き 実験は静かな個室で行う。発声者は実験の手続きを十分理解し、適度に練習を行った上で実験に参加する。発声者は発声しやすいよう録音マイクから 40 センチ程の距離に立ち発声を行う。録音レベルは発声者の声の大きさによって調整を行う。

実験に使用した機器 実験では次の機器を使用する。

録音機材には DAT(Sony TCD D-100) を用いる。48kHz サンプリング, 16bit 量子化の設定にて録音を行う。マイクは, Sony ECM-M957, 指向性マイクを用いる。

計算機では, 録音した音声を WAVE 形式ファイルとして処理する。なお, WAVE 形式への変換には Canopus DA-Port USB を用いる。

6.4.2 刺激音声の作成

モデル刺激

指標単語のモデル刺激は, 音声パラメータが様々に変化し, その変化が自然である必要がある。ここでは, 音声採取実験で採取した実験 A の音声のうち, 3 人の発声者の中で, 最も基本周波数が安定していた発声者 2 の音声「あ」の音声の 8 パターンを使用する。

無意味文

1. 無意味文の作成

モデル刺激と, 無意味文の音声は, 音声接続部分の違和感が少ない方がよい。そのため, 音声採取実験 B 「かちかち山」の感情を交えない朗読を元に, 次の 3 文の無意味文を作成する。一文はある部分の音声を時間的に逆に再生させたものである。残る二文の音声は, 朗読音声を不自然にならない程度に細かく分割しランダムにつなげたものであり, 分割の細かさが 2 文で異なっている。

2. 無意味文の選定

作成した無意味文 3 文について, 予備の音声聴取実験を行う。被験者は 9 人で, それぞれの音声を聞き, 音声に表れた感情について尺度で回答し, 音声が不自然に聞こえないか自由に記述する。このようにして得られた回答結果は, 感情尺度の 4 因子の基本感情から整理し, 基本感情に片寄りが最も少ないと考えられた無意味文 1 文を選択した。

刺激音声の作成

指標単語が位置する箇所を文中と文末と想定し、無意味文におけるモデル刺激の位置 2 箇所、声の高さの変化 4 パターン、声の長さの変化 2 パターン、声の大きさの変化 2 パターン、モデル刺激が入らない無意味文のみの計 33 パターンの音声を刺激音声とする。声の高さの変化及び声の長さの変化は音声採取実験 A で発声者が意図した変化であり、声の大きさは、発声者が発声したままの大きさの音声と、その音声から 12dB 減じた音声を作成する。なお、いずれの刺激音声についても、モデル刺激が無意味文と不連続に聞こえないよう、間が少なく、継続した一つの文として自然に聞こえる同一箇所を挿入位置に選び、両音声の振幅の最大値を揃えた上で接続する。

6.5 モデル刺激を用いた音声聴取実験

モデル刺激を用いて作成した刺激音声の聴取実験を行うことで、感情が表れていると感じられるモデル刺激を選定する。

6.5.1 方法

1. 被験者

聴取実験被験者は学生 20 名（男性 9 名，女性 11 名）である。

2. 手続き

被験者は、実験者より実験の説明を受けた後、一人で実験を行う。被験者は音声刺激を聞き、音声に表れていると感じた感情について、感情尺度を用いて回答を行う。感情尺度は「あなたは今の声を聞いて、どのような感情が声に含まれていると思いましたか？」の教示の後に、4 因子 12 項目が並べられ、被験者は 1 「あてはまらない」～5 「あてはまる」を選択する。33 パターンの刺激音声、12 項目の感情尺度は順番効果を考慮してランダムに提示される。刺激音声の提示には計算機を使用し、被験者は自分のペースで、必要であれば繰り返し同じ刺激音声を聞くことができる。

6.5.2 結果

感情尺度

感情尺度を得点化し、感情の合計点を各感情因子で算出し、基本感情「楽しさ」、
「怒り」、「悲しさ」、「落ち着き」の4水準間の平均値の差の検定を行う。

1. 無意味文の感情尺度

無意味文のみを刺激音声とした場合における感情尺度での平均値の差の検定の結果、5%水準での有意な差は認められなかった。従って、無意味文はいずれか特定の感情が含まれていると認知される刺激ではないことが確認された。

2. モデル刺激が接続された刺激音声の感情尺度

32パターンの刺激音声について、それぞれ感情尺度での平均値の差の検定を行った。5%水準で統計的に有意な差が認められた刺激は13パターンであった。13パターンの刺激の中で、音声から感じられる感情調査で音声に表れやすいとされた感情因子、すなわち「楽しさ」、「怒り」、「悲しさ」について、最も当該の感情が表れていると聴取された刺激について「文中」、「文末」にモデル刺激が位置するものそれぞれ一刺激の計6刺激を代表として選択した。感情が表れている基準については、他の3因子に比べて当該の感情因子が最も高い得点を取っているものとした。

表 6.5: 選定されたモデル刺激

感情	位置	大きさ	テンポ	F0(レンジ)
楽しさ	文中	大きさ小	ゆっくりと	167Hz(20Hz)
楽しさ	文末	大きさ大	ゆっくりと	167Hz(20Hz)
怒り	文中	大きさ大	素早く	150Hz(65Hz)
怒り	文末	大きさ大	素早く	150Hz(65Hz)
悲しさ	文中	大きさ小	ゆっくりと	124Hz(87Hz)
悲しさ	文末	大きさ小	ゆっくりと	124Hz(87Hz)

6.6 モデル刺激からの音声物理量の抽出

モデル刺激の音声は、音声採取実験にて発声者が音声物理量を意図的に変化させているが、合成音声に修正を加える場合には、実際の物理量を測定する必要がある。そのため、刺激音声から以下の音声物理量を抽出した。なお、音声物理量抽出については実験に使用した無意味文を含む文全体ではなくモデル刺激であった「あ」についてのみ対象として抽出を行い、パワー値 (30dB) を閾値とし、閾値に至らないものを無音区間と判断し、該当箇所に関しては音声物理量を抽出しなかった。なお、感情尺度の分析により代表とされた6音声についての音声物理量及び発声者への指示を表 6.5に示す。

1. 基本周波数

基本周波数の平均、レンジを抽出した。基本周波数の算出には、自己相関関数による方法を用い、窓関数はハミング法を用いた。フレーム周期は 5ms、基本周波数の抽出範囲は 30 から 500Hz とした。

2. 基本周波数の変化を表現する方法

感情が表れる音声物理量としては、音声全体の音声物理量（基本周波数、大きさ、発話時間）の平均の他、基本周波数の全体、あるいは特定箇所での変化が重視されることが多い [77, 76]。そこで、基本周波数の変化の表現として、線形関数による波形近似を用いる。基本周波数の変化の表現は、7つの時点にて値を抽出し、合成音声の音声物理量修正の見本とする。ここで7点を選択した基準は、予備の検討を通して、抽出する点が少ないと変化の表現を合成音声に反映することが難しく、多いと単音から構成されるような短い言葉の修正が困難であると考えられたからである。

モデル刺激の目視による傾き同定では目視を行う基準が人によって曖昧になる可能性がある。そこで、基本周波数の変化を線形 n 次関数で近似する。

予測関数を $F(x)$ とし、近似させる線形 n 次関数のモデルを $G(x)$ とおく。

$$G(x) = b_n x^n + b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \dots + b_1 x + b_0 \text{ である。}$$

観測値との誤差を最小限にする方法として、最小二乗法を用いる。

表 6.6: 「かちかち山」における指標単語と感情

文	感情	文と <u>指標単語</u>
文 1	怒り	「よーし、悪い <u>タヌキめ</u> . <u>しかえし</u> してやる」
文 2	楽しさ	「山にたくさん <u>おいしいキノコ</u> が出ているよ. いっしょに行ってごちそう作って <u>食べないか</u> 」
文 3	悲しさ	「 <u>ごめんなさい</u> , <u>ごめんなさい</u> . もう二度と悪さはしません」

方程式を解くことで得られた近似曲線上で等間隔に7点の値を抽出し、合成音声に7点の値を付与する。近似の次元数については、3次、5次、7次、10次で試行し、ここでは基本周波数の変化を少ない計算コストで近似できた7次を用いる。

6.6.1 指標単語調査

「かちかち山」の文の中で特に感情が感じられる部分（指標単語）を選択するために調査を行う。調査対象者は8名の学生で、「かちかち山」の一部を読み、「楽しさ」、「怒り」、「悲しさ」の感情が表れていると考えられる単語についてできるだけ多く選択する。回答から、最も多く感情が表れていると選択された文中と文末の単語を指標単語とする。選定された指標単語と、指標単語から感じられると推測された感情について表 6.6に示す。

6.6.2 感情を付与した物語読み上げ

調査で選択された指標単語に対して、モデル刺激から抽出された音声物理量及び線形曲線による基本周波数の変化近似の物理量を、合成された指標単語部分に以下のように付与した。

線形曲線による基本周波数変化近似の情報の付与については、合成音声の基本周波数の平均とレンジをモデル刺激の発声者の音声と合わせた上で行った。また、発話時間については、発声者が意図した音声物理量の変化を参考に、ゆっくりと発声したモデル刺激については通常の合成音声の1.5倍、素早く発声したモデル

刺激については0.9倍に修正した。音声の大きさについては、刺激音声作成の際を参考に、音声採取実験の音量そのままであった条件については通常の合成音声の200%、12dB減じた条件については50%に修正を行った。「楽しさ」、「怒り」、「悲しさ」について、それぞれ文中の指標単語を修正した合成音声と、文末の指標単語を修正した合成音声を用意し、比較のために指標単語に修正を加えない合成音声の計9刺激を作成した。

修正した合成音声を用いて聴取実験を行った。被験者は10名であり、「かちかち山」会話部分の単純な合成音声と、指標単語について音声物理量の修正を行った合成音声を聞き、どのような感情が表れているかについて感情尺度（1「あてはまらない」～5「あてはまる」）を用いて回答を行った。結果は図6.1、6.2、6.3に示す。

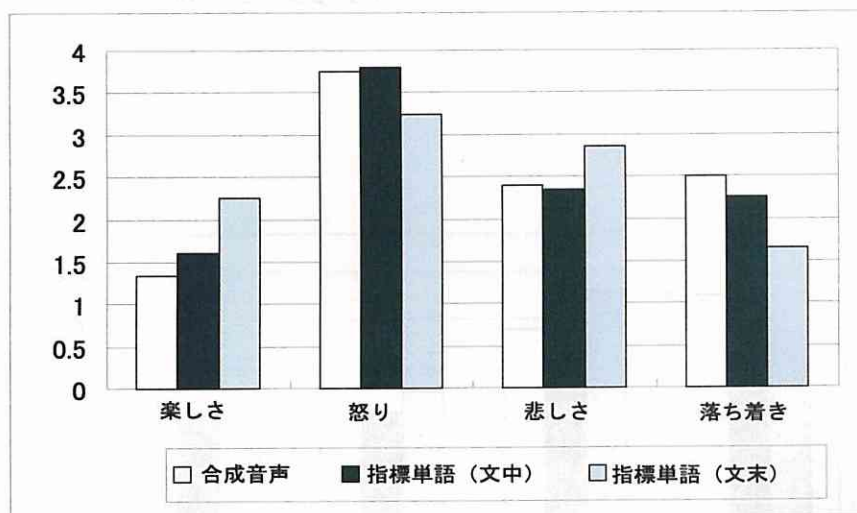


図 6.1: 文1(怒り)の感情尺度平均得点

なお、合成音声についてそれぞれ自然であるかどうかを自由に回答してもらったところ、文1、文3、文2の順に指標単語に修正を加えた合成音声の方が自然であるとの結果を得ている。

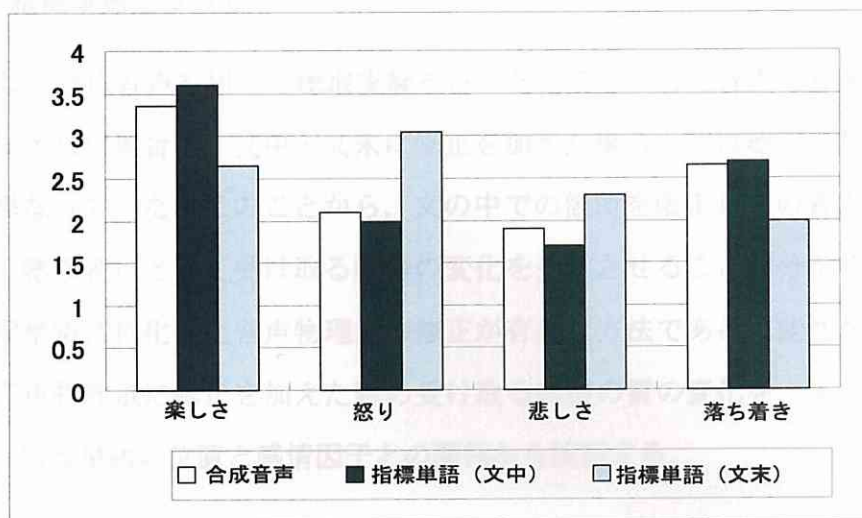


図 6.2: 文 2(楽しさ) の感情尺度平均得点

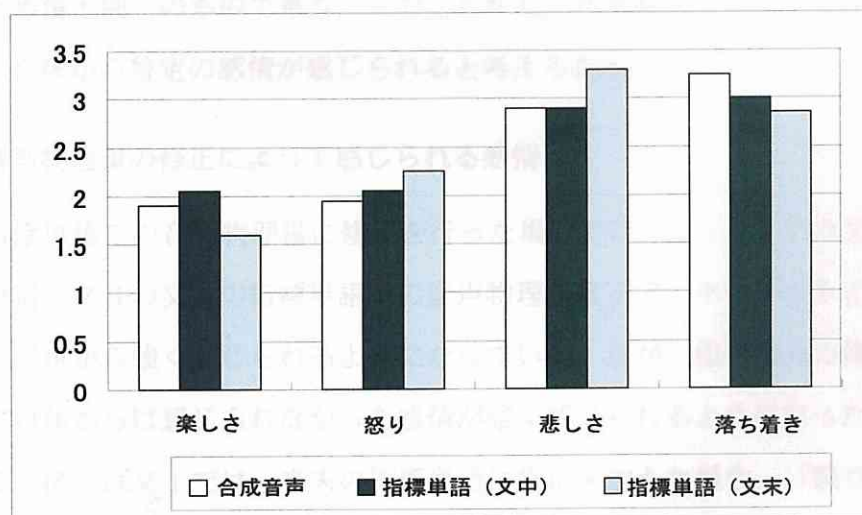


図 6.3: 文 3(悲しさ) の感情尺度平均得点

6.7 結果と考察

6.7.1 指標単語について

作成した合成音声を用いて聴取実験を行ったところ、合成音声に音声物理量の修正を加えない場合と、文中・文末に修正を加えた場合とでは聴取される感情の得点が異なっていた。このことから、文の中での箇所を限定しての音声物理量の修正は、聴取者にとって受け取る感情の変化を生じさせることが分かる。すなわち、指標単語に特化した音声物理量の修正が有用な方法である可能性がある。そこで、音声物理量に修正を加えた際の受け取る感情の質の変化を、文と感情因子の関係、指標単語の位置と感情因子との関係から検討する。

6.7.2 文から感じられる感情と感情因子の関係

文自体から感じられる感情

音声物理量に修正を加えない合成音声では、特定の感情因子の得点が高い。得点の高い感情因子は、指標単語調査で指標単語から感じられると推定された感情と同一のものであり、このことから、合成音声を聴取する際、文の意味自体から特定の感情が感じられると考えられる。

音声物理量の修正によって感じられる感情

指標単語での音声物理量に修正を行った場合では、文1、文2の文中の指標単語、文3の文末の指標単語での音声物理量修正で、ターゲットとした感情が音声から強く感じられるようになっている。だが、他の部分の修正では、文自体からは感じられなかった感情が強く感じられるようになったものもある。例えば文1では、文末の指標単語に修正を加えた場合、「怒り」と共に「悲しさ」の感情因子得点が上昇しており、文2では「楽しさ」と共に「怒り」の感情因子得点が上昇している。

これらの異なった感情の生起という現象を考えた場合、背景に二つの要因が推測できる。

1つの要因は、感情構造そのものとして共起しやすい感情の存在である。例えば「怒り」と「悲しさ」の関係を考えた場合、感情因子間の関係を表6.2

から見ると、この二因子は相関が高い。このことから、文1で見られるように、「怒り」と「悲しさ」は感情の構造として共に上昇する可能性が高いと考えられる。

もう一方の要因は、音声物理量の共通性である。例えば「楽しさ」と「怒り」の特徴を検討すると、本研究でのモデル刺激では共に声の大きさが大きく、基本周波数が高いといった共通点がある。感情が感じられる音声の手がかりについては声の高さ等が述べられてきたが [50, 33, 34], 複合した手がかりを用いることで「喜び」の音声「怒り」と聴取されたという報告も見受けられる [76]. 本研究でも、聴取者は音声物理量の様々な要素を手がかりに判断を行うことで、特徴が共通した2つの感情因子の得点が共に上昇したと考えられる。

6.7.3 感情と指標単語の位置の関係

声の高さの変化は、アクセント、イントネーション等と関連して述べられることがあり、特に文末での声の高さの変化は、疑問、反問、勧誘などの情報を伝えるとされる [79]. 本研究では、文末にある指標単語の音声物理量を修正した場合、「落ち着き」の感情得点が低下すると共に「悲しさ」の感情得点が上昇する傾向が見られた。特に文末において音声物理量が増加すると、断定というよりも疑問として感じられ、音声から聴取される落ち着いた印象が低下する可能性がある。

合成音声の指標単語の音声物理量を修正すると、感情の構造や音声物理量の手がかりから、文の意味自体から判断するよりも複雑な感情を感じ取るようになり、落ち着いた印象が低くなると考えられる。必ずしもターゲットとした感情のみが強く感じられる結果とはならなかったが、合成音声の自然性を評価した自由回答からは、複雑な感情が受け取られている文1が最も自然に聞こえるとの結果を得ており、ターゲットの感情だけが伝わるのが自然な音声とは考えにくい。本研究では、音声合成での分かりやすさを重視した関係上、基本感情を前提とした立場で研究を行ってきたが、表 6.2で見られた感情因子間の相関関係や、感情次元論で述べられている感情の空間布置を考慮すると [62], 感情そのものの性質として異なる感情の共起関係が音声の自然性、感情そのものの自然性に影響しているとも考えられよう。

なお、実験では、指標単語に特化した音声物理量修正の効果を検討するために、あえて文全体での音声物理量は無修正にとどめた。聴取者にとって、音声におけるパラ言語的特徴と韻律特徴は独立のものとして受け取られていると言われるものの [42]、合成音声の自然性を向上するためには、アクセントやイントネーションといった情報を重畳していく仕事も必要だと考えられる。今後の検討課題としていきたい。

6.8 まとめ

感情豊かな読み上げ音声の作成を目標として、感情が表れている箇所（指標単語）における、有効な音声パラメータの検討を行った。第一に、研究で用いる基本感情を抽出するために調査を行った。調査で得られた4つの感情因子「楽しさ」、 「怒り」、 「悲しさ」、 「落ち着き」を本研究での基本感情と定義し、4因子12項目を感情尺度として刺激音声の評価に用いた。

様々な文への一般化を考え、無意味文とモデル刺激を用いた刺激音声を作成し、聴取実験を行った。感情が表れていると選定されたモデル音声について、音声物理量（基本周波数）を抽出し、基本周波数の変化を表現する方法を導入した。

得られた音声物理量を「かちかち山」物語読み上げの会話文の指標単語に付与し、聴取実験を行った。結果、合成音声の指標単語に音声物理量を付与すると、音声から感じる落ち着いた印象が低くなり、文の意味だけからよりも複雑な感情を、音声物理量の手がかりと、感情そのものの構造から感じ取るようになると考えられた。

第 7 章

結論

7.1 研究のまとめ

本研究では、正確な感情のコミュニケーションの促進を目標として、音声と心理状態（感情）の関連について検討を行った。

第 2 章、第 3 章では、音声に関する研究と感情に関する研究を概観し、本研究での感情への視点を定義した。ここでは、独立のカテゴリにより構成される基本感情の立場から研究を行うことを述べた。

第 4 章「感情を表現した音声の分析」では、感情が表れた音声の分析と聴取実験を行った。声優が発声した感情音声と、複数の一般者が発声した感情音声から音声物理量を抽出し、一文での代表値の統計的分析と図から目視で得た変化の分類を行うことで、感情と音声物理量との関連を検討した。感情を表現する音声物理量について感情表現の訓練を受けた人と訓練を受けたことのない一般の人に共通して表れる特徴があること、重要とされるパラメータが文全体で一つの代表値に注目する場合と文の中のダイナミックな変化に注目する場合とで異なることが分かった。感情音声の聴取実験からは、発声者の表現と聴取者の認知における感情の不一致が認められ、その背景として、発声者の個人性、聴取者の個人性、文脈の問題が考えられた。

第 5 章「音声物理量からの抑うつ傾向判定」では、感情の一種である抑うつに焦点を当て、抑うつ傾向判定システム構築を目標に抑うつ傾向と文読み上げの音

声における音声物理量の関連を検討した。文読み上げの課題文として用いた BDI (Beck Depression Inventory) のうち 6 文は、音声物理量と抑うつ傾向の関連が見出された。抑うつが表れやすい箇所としての指標単語の概念を提案し、その選択法について検討を行った。3つの指標単語選択戦略にて選択した指標単語について、ロジスティック重回帰分析にて検討を行った。実験と分析の結果、統計指標単語選択戦略が他の2つの指標単語選択戦略よりも判定に対して有効な指標単語であることが分かった。

第6章「感情豊かな読み上げのための音声パラメータ」では、感情豊かな読み上げ音声の合成を目標として、感情が表れる箇所としての指標単語に重点を置き、有効な音声物理量の検討を行った。音声物理量を操作したモデル音声を作成し音声聴取実験を行った。感情が感じられると選定されたモデル音声について、音声物理量を抽出し、基本周波数の変化を表現する方法を導入した。得られた音声物理量を「かちかち山」物語読み上げの会話文の指標単語に付与し、聴取実験を行うことで有効性を検討した。結果、合成音声の指標単語での音声物理量を修正すると、音声から感じる落ち着いた印象が低くなり、文の意味よりも複雑な感情を認知させることが分かった。

7.2 指標単語について

本研究では、はじめに感情のコミュニケーションの問題を取り上げ、問題が起きる背景として、コミュニケーションの信号の多量さ、信号の多種さによる受け手の混乱を仮定した。この問題に対しては、重要な部分、すなわち累積する信号の中で重要な箇所の信号、多種の信号の中で感情に関する信号を重点的に扱うことで、より正確な感情のコミュニケーションの促進を検討した。

累積する信号の中で重要な箇所を扱う手法として、本研究では指標単語の概念を提案した。第5章においては抑うつが強調される箇所を指標単語として、指標単語の音声物理量を用いた判定器の試作と評価を行った。第6章においては感情が強調される箇所を指標単語とし、指標単語に特化した音声物理量の修正を行い合成音声にて評価を行った。共に指標単語により音声物理量と感情の関連が確認されたことから、多量の信号の中で箇所を限定する方法は、感情のコミュニケー

ションにとって有益であると示唆された。

7.3 音声物理量と感情の関連について

また、音声物理量と感情との関連については3つのアプローチから検討を行った。

意図的に感情を表現した音声を用いた第一のアプローチ（第4章）では、調査で選定された6感情と音声物理量の関連を検討した。ここでは、「怒」では声が大きく発話時間が短く、「鬱」、「安」は声小さく、「喜」、「鬱」、「安」は発話時間が長くなるのが、感情表現の訓練を受けた者と訓練を受けたことのない一般者に共通して見られる傾向であり、文全体の代表値については、振幅と発話時間が感情を表現する上で重要な音声物理量パラメータと考えられた。また、音声物理量の文の中の変化の分類では、基本周波数は、「喜」、「驚」が同様の変化を示すと分類され、振幅は「喜」、「安」及び「鬱」、「驚」、「怖」、「怒」が同様の変化を示すと分類され、音声物理量の文の中の変化については、基本周波数と振幅が感情を表現する上で重要な音声物理量パラメータと考えられた。感情音声の聴取実験では、発声者が表現した感情と聴取者が推定した感情とが一致しやすかったものとして、「怒」、「安」が、一致しなかったものとして「喜」、「驚」があった。

自然な状態での発声者の感情と音声物理量の関連を検討した第二のアプローチ（第5章）では、特定の感情である抑うつを対象として音声物理量との関連を検討した。ここでは、高抑うつ傾向の者の発声した音声において、声が高く大きいことが認められ、不安や悲しみ等の影響が音声物理量の特徴として表れることが考えられた。

音声から聴取者が認知した感情と、音声物理量との関連を検討した第三のアプローチ（第6章）では、調査で選定された3つの感情と音声物理量の関連を検討した。感情と関連が認められたモデル刺激の音声物理量は、「楽しさ」では基本周波数が高く発話時間が長く、「怒り」では基本周波数が低くから高くへ変化し、声が大きく発話時間が短く、「悲しさ」では基本周波数が高くから低くへ変化し、声小さく発話時間が短かった。特に「悲しさ」に関しては、合成音声の音声

物理量を修正した上での聴取実験でもこの情報の有効性が確認された。

3つのアプローチによる相違は、怒り、楽しみ、悲しみの感情において検討する。

怒りに関係した音声は、発声者が意図的に怒りを表現した音声と聴取者が音声より認知した感情とで一致が見られ、また、発声者が意図的に怒りを表現した音声と聴取者が怒りを表現していると選択した音声とで、音声物理量に共通点が見られた。すなわち、声が大きく発話時間が短い音声は怒りの音声となる。

楽しみに関係した音声は、発声者が意図的に表現した音声と聴取者が音声より認知した感情とでは一致が見られなかったが、発声者が意図的に楽しみを表現した音声と聴取者が喜びを表現していると選択した音声とでは、音声物理量に一部共通点が見られた。すなわち、発話時間が長いという点のみ楽しみの音声では一致している。

悲しみに関係した音声は、発声者が意図的に表現した音声と聴取者が音声より認知した感情とでは一部一致が見られなかったが、発声者が意図的に表現した音声の音声物理量と聴取者が悲しみを表現していると選択した音声の音声物理量においては、声小さく発話時間が短いという一致が見られた。しかし、悲しみと関連のある抑うつが自然に表現されていると考えられる音声の音声物理量は異なっており、声が高く声が大きいという特徴が認められた。すなわち、悲しみの音声では、発声者の持つ感情の音声への表れや意図的な表現、受け手の認知が異なる。

自然な状態での発声者の感情については悲しさに関連した感情の検討を行ったのだが、本研究における3つのアプローチからは、感情の種類によって発声者が意図的に表現した音声で伝わりやすいものと、伝わりにくいものが存在することが分かる。また、自然な状態での発声者の音声に表れた感情と、受け手が認知した感情とは異なることが考えられる。このような3つのアプローチの結果の不一致に関しては、音声物理量を用いて送り手の感情を分析する聴取の役割と、受け手が認知しやすい音声物理量を用いて合成を行う発声の役割とを計算機が担うことが有効であることが示唆されている。

7.4 今後の課題

本研究では、累積する信号の中で重要な箇所を扱う手法としての指標単語を提案した。指標単語については、分析、合成という二つのアプローチにてある程度の有効性が確認されたが、選択方法については更なる検討が必要と考えられる。本研究での抑うつが表れる箇所としての指標単語は、統計による選択戦略、聴取による選択戦略、専門家による選択戦略、すなわち、計算機による音声物理量の利用と一般人の聴覚と専門家が文の意味から読み取る方法により選択を行い、合成で用いた指標単語は一般人が文の意味から読み取る方法により選択を行った。感情が表れる箇所としての指標単語を考えたとき、単語の持つ単語特有の意味の情報と、文脈から派生する全体的な流れの中での意味の情報との両側面を含む必要が考えられるであろう。計算機による言葉の意味の解析は困難な課題であるが、最小限のコストで正確なコミュニケーションを可能とする指標単語を用いるために、言語学的な知識を利用した計算機による自動的な選択は必要不可欠の課題と考えられる。

本研究では感情に対して独立したカテゴリに分類されるという基本感情の立場より感情を扱った。しかし聴取実験の結果を見ると、音声との関連で複数の基本感情を採択した第4章では一つの感情を表現した音声刺激から複数の感情が認知され、音声物理量を変化させた音声刺激を用いた第6章でも複数の感情が認知されていた。この結果は、感情の構造自体がそれぞれ独立したカテゴリではなく、互いに関連しあったものであるということを示唆している可能性が高い。感情が互いに関連しあったものであるなら、むしろ計算機では複数の感情を複合的に扱ってゆく分析・合成を行うことが人間にとって自然なものとして受け入れられるのではないか。感情の構造については今後の研究が必要と考えられる。

謝辞

本研究を進めるにあたり、終始親身な御指導、御助言をいただきました横浜国立大学大学院 環境情報研究院 影井清一郎教授、論文を御指導いただきました横浜国立大学大学院 環境情報研究院 田村直良教授に心から感謝の意を捧げます。

また、検討会における議論や、意見交換によって多大なる援助を下された大学院大学院工学研究科電子情報学専攻博士課程前期在籍の山野辺啓雄氏、大学院環境情報学府情報メディア学専攻博士課程前期在籍の阿部亮介氏、浅山敦史氏、狩野剛氏、北條孝氏、川本大輔氏、松尾宣彦氏、山本隼祐氏には多大なる御協力を頂き感謝致します。

参考文献

- [1] M.Alpert, E.R.Pouget, and P.R.Silve. Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders*, Vol. 66, pp. 59-69, 2001.
- [2] American Psychiatric Association, editor. *American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. American Psychiatric Association, 1994.
- [3] N.C.Andreasen, M.Alpert, and M.J.Martz. Acoustic analysis. *Arch Gen Psychiatry*, Vol. 38, pp. 281-285, 1981.
- [4] 安西信雄, 丹羽真一, 斎藤治, 増井寛治, 亀山知道, 平松謙一, 岡崎祐士, 伊藤憲治. 精神分裂病患者の感情認知, 表出障害の研究 (2) . 臨床精神医学, Vol. 12, No. 3, pp. 341-351, 1983.
- [5] 安西信雄, 丹羽真一, 斎藤治, 増井寛治, 亀山知道, 平松謙一, 岡崎祐士, 伊藤憲治. 精神分裂病患者の感情認知, 表出障害の研究 (1) . 臨床精神医学, Vol. 12, No. 1, pp. 85-95, 1983.
- [6] 安西信雄, 丹羽真一, 斎藤治, 増井寛治, 亀山知道, 平松謙一, 岡崎祐士, 伊藤憲治. 精神分裂病患者の感情認知, 表出障害の研究 (3) . 臨床精神医学, Vol. 14, No. 3, pp. 333-342, 1985.
- [7] A.T.Beck, C.H.Ward, M.Mendelohn, J.Mock, and J.Erbaugh. An inventory for measuring depression. *Archives of General Psychiatry*, Vol. 4, pp. 561-571, 1961.

- [8] J.Bowlby. *Attachment and loss*. Hogarth Press, 1980. (『母子関係の理論』黒田実郎(訳), 岩崎学術出版社, 1976-1981).
- [9] C.Breitenstein, D.V.Lancker, and I.Daum. The contribution of speech rate and pitch variation to the perception of vocal emotions in a german and an american sample. *Cognition and Emotion*, Vol. 15, pp. 57-79, 2001.
- [10] K.M.B.Bridges. Emotional development in early infancy. *Child Development*, Vol. 3, pp. 324-341, 1932.
- [11] R.Buck. Nonverbal behavior and the theory of emotion: The facial feedback hypothesis. *Journal of Personality and Social Psychology*, Vol. 38, pp. 811-824, 1980.
- [12] J.E.Cahn. The generation of affect in synthesized speech. *Journal of the American Voice I/O Society*, Vol. 8, pp. 1-19, 1990.
- [13] C.E.Izard. *The psychology of emotions*. Plenum Press, 1991. (『感情心理学』比較発達研究会(訳), ナカニシヤ出版, 1996).
- [14] R.Cowie. Describing the emotional states expressed in speech. *Speech Emotion*, pp. 11-18, 2000.
- [15] C.Darwin. *The expression of the emotions in Man and Animals*. University of Chicago Press, 1872/1965. (『人及び動物の表情について』浜中浜太郎(訳), 岩波書店, 1931).
- [16] J.R.Davitz and L.Davitz. The speech. *Journal of communication*, Vol. 9, pp. 110-117, 1959.
- [17] J.R.Davitz. *The Communication of Emotional Meaning*. McGraw-Hill, 1964.
- [18] P.Ekman and W.V.Friesen. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, Vol. 17, pp. 124-129, 1971.

- [19] P.Ekman and W.V.Friesen. *Unmasking the face : a guide to recognizing emotions from facial clues*. Englewood Cliffs, 1975. (『表情分析入門：表情に隠された意味をさぐる』工藤力(訳編), 誠信書房, 1987).
- [20] H.Ellgring and K.R.Sherer. Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior*, Vol. 20, No. 2, pp. 83-110, 1999.
- [21] R.N.Emde, D.H.Kligman, J.H.Reich, and T.D.Wade. *Emotional expression in infancy:I.Initial studies of social signaling and an emergent model*. Plenum Press, 1978. M.Lewis & M.A.Rosenblum.The Development of affect.
- [22] N.H.Frijda and B.Mesquita. *The social roles and functions of emotions.*, pp. 51-87. 1994. (S.Kitayama and H.R.Markus(Eds.) *Emotion and culture. Empirical studies of mutual influence*).
- [23] 藤崎博也. 韻律研究の諸側面とその課題. 日本音響学会秋季研究発表会講演論文集, pp. 287-290, 1994.
- [24] 藤原武弘. 社会心理学, 現代心理学シリーズ, 第9巻. 培風館, 1997.
- [25] 深田博己. コミュニケーション心理学: 心理学的コミュニケーション論への招待. 北大路書房, 1999.
- [26] 古井貞熙. 音響・音声工学. 近代科学社, 1992.
- [27] 林潔. 学生の抑うつ傾向の検討. カウンセリング研究, Vol. 20, No. 2, pp. 76-83, 1988.
- [28] 平賀裕, 斎藤善行, 森島繁生, 原島博. 音声に含まれる感情情報抽出の一検討. 信学技報, Vol. HC93-66, pp. 1-8, 1994.
- [29] 広瀬啓吉, 高橋登, 藤崎博也, 大野澄雄. 音声の基本周波数パターンにおける話者の意図・感情の表現. 信学技報, Vol. HC94-41, pp. 33-40, 1994.
- [30] 飯田朱美, ニック・キャンベル, 安村通晃. 感情表現が可能な合成音声の作成と評価. 情報処学会論文誌, Vol. 40, No. 2, pp. 479-486, 1999.

- [31] 石村貞夫, デズモンドアレン. すぐわかる統計用語. 東京図書, 1997.
- [32] 板橋秀一, 粕谷英樹, 北脇信彦. 音声分野 (特集 音響学における 20 世紀の成果と 21 世紀に残された課題 - 第 1 部 分野別の流れ). 日本音響学会誌, Vol. 57, No. 1, pp. 11-20, 2001.
- [33] 伊東 (水木) 久美子. 感情を含む音声に関する基礎研究 (2). 人間工学, Vol. 21, No. 2, pp. 81-87, 1985.
- [34] 伊東 (水木) 久美子. 感情を含む音声に関する基礎研究 (3). 人間工学, Vol. 22, No. 4, pp. 211-217, 1986.
- [35] C.E.Izard. *Human emotions*. Plenum Press, 1977.
- [36] W. James. What is an emotion? *Mind*, Vol. 19, pp. 188-205, 1984.
- [37] 門谷信愛希, 阿曾弘具, 鈴木基之, 牧野正三. 音声に含まれる感情の判別に関する検討. 情報処理学会研究報告, Vol. 119, pp. 43-48, 2000.
- [38] 片口安史. (改訂) 新・心理診断法. 金子書房, 1987.
- [39] 川波弘道, 広瀬啓吾. 態度・感情音声における韻律的特徴の考察. 信学技報, Vol. SP97-67, pp. 73-80, 1997.
- [40] 河野友信. ハイテク社会の病. 現代の病シリーズ. 医学書院, 1992.
- [41] 木村広樹, 富田豊, 本多敏. 声帯および声道特性中のパラメーター変化による感情を持つ音声の合成. 人間工学, Vol. 32, No. 6, pp. 319-325, 1996.
- [42] 小林聡, 北澤茂良. 日本語の自然対話音声におけるパラ言語的特徴の検討. 日本音響学会誌, Vol. 56, pp. 467-476, 2000.
- [43] R.S.Lazarus. *Emotion and adaptation*. Oxford University Press, 1991.
- [44] J.P.Leff and C.Vaughn. Expressed emotion in families : its significance for mental illness. *Guilford Press*, 1985. (『分裂病と家族の感情表出』三野善央, 牛島定信 (訳), 金剛出版, 1991).

- [45] M.Lewis and M.A.Rosenblum. *The Development of affect*. Plenum Press, 1978.
- [46] 丸本徹, N. Campbell. 波形接続型音声合成方式による発話様式の制御. 日本音響学会講演論文集, Vol. 1-7-9, pp. 213-214, 2000.
- [47] 丸山芳也. 精神分裂病者における言語の音調理解に関する研究. 精神医学, Vol. 2, No. 10, pp. 667-678, 1960.
- [48] 大辞林. 三省堂, 1995.
- [49] 三須秀亮. 音調テストによる精神分裂病家族の研究. 精神医学, Vol. 12, No. 4, pp. 305-312, 1970.
- [50] 水木久美子. 感情を含む音声に関する基礎研究. 人間工学, Vol. 20, No. 4, pp. 225-230, 1984.
- [51] I.R.Murray and J.L.Arnott. Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Communication*, Vol. 16, pp. 369-390, 1995.
- [52] 中川聖一(章主任). 音声処理, pp. 1316-1331. オーム社, 1989. 情報処理学会(編), 『情報処理ハンドブック』.
- [53] 中村明. 感情表現辞典. 六興出版, 1979.
- [54] 日本音声言語医学会編. 声の検査法(第二版). 医歯薬出版, 1994.
- [55] 日本語語彙大系 CD-ROM版, 1999.
- [56] 小此木啓吾. 一・五の時代. 筑摩書房, 1987.
- [57] <http://vision.kuee.kyoto-u.ac.jp/lecture/dsp/pitch/pitchprog.html>.
pitcher.
- [58] H.G.Wallbott. Vocal behavior and psychopathology. *Pharmacopsychiatry*, Vol. 22, pp. 13-16, 1981.

- [59] A.Portmann. *Biologische fragmente zu einer lehre vom menschen*. Schwabe, 1944. (『人間はどこまで動物か：新しい人間像のために』高木正孝(訳), 岩波書店, 1961).
- [60] C.R. Randolph. *The science of emotion : research and tradition in the psychology of emotions*. Prentice Hall, 1996. (『感情の科学：心理学は感情をどこまで理解できたか』齊藤勇(監訳), 誠信書房, 1999).
- [61] G.rod. *Technostress : the human cost of the computer revolution*. Addison-Wesley, 1984. (『テクノストレス』池央耿, 高見浩(訳), 新潮社, 1984).
- [62] J.A.Russell. Pancultural aspects of the human conceptual organization of emotion. *Journal of Personality and Social Psychology*, pp. 1281–1288, 1983.
- [63] J.A.Russell and B.Fehr. Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology*, Vol. 117, pp. 89–90, 1987.
- [64] 坂口守男, 郭哲次, 吉益文夫, 百蹊陽三, 東雄司. 音響分析を用いた精神精神分裂病者の感情表現能力に関する研究. *精神医学*, Vol. 39, No. 6, pp. 583–591, 1997.
- [65] 坂本正治, 齊藤隆. 高品質な感情音声を合成する手法と顔型ロボットへの実装例. *情報処理学会研究報告*, Vol. SLP37, No. 5, pp. 25–30, 2001.
- [66] 櫻庭京子, 今泉敏, 箕一彦. 幼児・児童の感情表現における音響的分析 – 「ひかちゅう」にこめられた感性情報 –. *信学技報*, Vol. SP99-86, pp. 1–8, 1999.
- [67] 櫻庭京子, 今泉敏, 箕一彦. 感情表現が言語的制約に及ぼす影響の発達の検討. *信学技報*, Vol. SP2000-39, pp. 53–59, 2000.
- [68] K.R.Scherer and P.Ekman, editors. *Approaches to emotion*, chapter R.Plutchik. Emotions: A general psychoevolutionary theory. Erlbaum, 1984.

- [69] K.R.Scherer. Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, Vol. 99, No. 2, pp. 143–165, 1986.
- [70] H.Schlosberg. A scale for the judgement of facial expressions. *Journal of Experimental Psychology*, Vol. 44, pp. 229–237, 1941.
- [71] H.Schlosberg. Three dimensions of emotion. *Psychological Review*, Vol. 61, , 1958.
- [72] M. Schroder. Emotional speech synthesis—a review. *Proceedings of Eurospeech*, Vol. Aalborg, pp. 561–564, 2001.
- [73] H.Segal. *Introduction to the work of Melanie Klein*. The Hogarth Press, 1873. (岩崎徹也 (訳) 『メラニー・クライン入門』, 岩崎学術出版社, 1977).
- [74] A.W.Seigman and S.Boyle. Voices of fear and anxiety and sadness and depression:the effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology*, Vol. 102, No. 3, pp. 430–437, 1993.
- [75] C.E.Shannon and W.Weaver. *The mathematical theory of communication*. University of Illinois Press, 1949. (『コミュニケーションの数学的理論：情報理論の基礎』長谷川淳, 井上光洋 (訳) , 明治図書, 1969).
- [76] 重永実, 小川孝, 中尾光志. 単語音声による感情表現について. 信学技報, Vol. SP95-15, pp. 39–46, 1995.
- [77] 重永実. 感情の判別分析からみた感情音声の特性. 電子情報通信学会論文誌 A, Vol. 83, No. 6, pp. 726–735, 2000.
- [78] H.Spencer. *The principles of psychology*. appleton, 1890.
- [79] 杉藤美代子. 声にだして読もう! : 朗読を科学する. 明治書院, 2003.
- [80] 鈴木朋子. 精神分裂病患者の感情について : 音響分析と音調テストを用いて, 1997.

- [81] 鈴木朋子, 藤森正仁, 田村直良. 心理的うつ状態と文読み上げの音声物理量について. 信学技報, Vol. SP2000-43, pp. 9-16, 2000.
- [82] 鈴木朋子. 感情を込めた音声に関する研究 - 音声の物理的特徴・聴取による音声の感情推定 -. 日本心理学会 (第 65 回大会研究発表集), 2001.
- [83] 鈴木朋子, 田村直良. 心理状態の音声物理量への関連について. 信学技報, Vol. HCS2002-4(2002-04), pp. 19-24, 2002.
- [84] 鈴木朋子. 感情を込めた音声に関する研究 (3). 日本感情心理学会第 11 回大会, 2003.
- [85] 鈴木朋子, 山本隼佑, 田村直良. 感情豊かな読み上げのための音声パラメータの検討. 信学技報, Vol. HIP2003-86(2003-12), pp. 51-56, 2003.
- [86] 鈴木朋子, 田村直良. 音声物理量からの抑うつ傾向判定. 電子情報通信学会論文誌, 2004 掲載予定.
- [87] 丹波聖尚. 緊急時における音声変容の研究 (1). 航空医学実験隊報告, 1970.
- [88] 丹波聖尚. 緊急時における音声変容の研究 (2). 航空医学実験隊報告, 1971.
- [89] 丹野義彦, 町山幸輝, 荒尾裕子. 精神精神分裂病における距離弁別と感情認知との関連. 臨床精神医学, Vol. 15, No. 2, pp. 243-254, 1986.
- [90] 寺崎正治, 岸本陽一, 古賀愛人. 多面的感情状態尺度の作成. 心理学研究, Vol. 62, No. 6, pp. 350-356, 1992.
- [91] R.Tourangeau and P.C.Ellsworth. The role of facial response in the experience of emotion. *Journal of Personality and Social Psychology*, Vol. 37, pp. 1519-1531, 1979.
- [92] 内山喜久雄, 筒井未春, 上里一郎. うつ病. 同朋舎, 1990.
- [93] 宇津木成介, 岡村紀子. 感情を含む音声に関する基礎研究. 人間工学, Vol. 20, No. 4, pp. 225-230, 1976.

- [94] 宇津木成介. 音声による感情表出と認知. 第43回日本心理学会大会論文集 (抄録), 1979.
- [95] 宇津木成介. ノンバーバル行動の実験的研究, 音声による情動表出と非言語的な弁別手がかり. 川島書店, 1993. 荘巖舜哉 (編) .
- [96] D.Watson and A.Tellegen. Toward a consensual structure of mood. *Psychological bulletin*, Vol. 98, pp. 219–235, 1985.
- [97] C.E.Williams and K.N.Stevens. Emotions and speech : some acoustical correlates. *Journal of the acoustical society of America*, pp. 1238–1250, 1972.
- [98] R.S.Woodworth. *Experimental psychology*. Henry Holt and Company, 1838.
- [99] W.Wundt. *Outlines of psychology*. Stechert, 1896.
- [100] 八田武志, 唐澤かおり, 岡本真彦, 杉村智子, 川上綾子, 岩原昭彦. インターネット・パソコン通信における文字情報の伝達効率改善に関する認知心理学的研究. 電気通信普及財団研究調査報告書, Vol. 18, pp. 228–236, 2003.
- [101] 安木博臣, 岩宮真一郎. 音声分析に基づく虚偽検出の可能性. 信学技報, Vol. SP2000-152, pp. 73–77, 2001.

発表論文一覧

主要な発表論文

1. 鈴木朋子, 田村直良. 音声物理量からの抑うつ傾向判定. 電子情報通信学会論文誌 (採録)

その他の発表論文

1. 鈴木朋子, 藤森正仁, 田村直良. 心理的うつ状態と文読み上げの音声物理量について. 信学技報, SP2000-43, pp.9-16, 2000.
2. 馬場謙一, 鈴木朋子, 竹内理英, 松本京介, 長谷川麻衣子. 自我機能の発達と病態化の研究 (その1). 放送大学研究年報, vol.18, pp.1-10, 2000.
3. 鈴木朋子, 井上果子. 日本における精神分析学のはじまり (1): 久保良英の貢献. 横浜国立大学大学院教育学研究科 教育相談・支援総合センター紀要, vol.1, pp.101-114, 2002.
4. 鈴木朋子, 田村直良. 心理状態の音声物理量への関連について. 信学技報, HCS2002-4(2002-04), pp.19-24, 2002.
5. 鈴木朋子, 井上果子. 日本における精神分析学のはじまり (2): 大槻快尊の貢献. 横浜国立大学大学院教育学研究科 教育相談・支援総合センター紀要, vol.2, pp.127-139, 2002.
6. 鈴木朋子, 山本隼佑, 田村直良. 感情豊かな読み上げのための音声パラメータの検討. 信学技報. HIP2003-86(2003-12), pp.51-56, 2003.

7. 鈴木朋子. 久保良英によるビネー式知能検査の改訂. 心理学史・心理学論, vol.5, pp.1-13, 2003.

大会発表

1. 鈴木朋子, 門前進. 発声に与える干渉の強弱. 日本心理学会 (第 62 回大会発表論文集), 1998.
2. 芳川玲子, 春日井翠, 正田節子, 鈴木朋子, 竹内理英. 子供の遊びの変遷. 第 4 回精神衛生学会分科会「子どもと環境」: 最近の少年非行, 1998.
3. 鈴木朋子. 精神分裂病患者の感情の認知について - 音調テストを用いて - 日本心理学会 (第 64 回大会発表論文集). 2000.
4. 鈴木朋子. 文章音読時の音声物理量に関する研究. 日本心理臨床学会 (第 19 回大会研究発表集), 2000.
5. 鈴木朋子. 感情を込めた音声に関する研究 - 音声の物理的特徴・聴取による音声の感情推定 -. 日本心理学会 (第 65 回大会研究発表集), 2001.
6. 鈴木朋子. 感情を込めた音声に関する研究 (2) - フォルマントからの検討 -. 日本感情心理学会第 10 回大会, 2002.
7. 鈴木朋子. 精神分裂病患者の感情の喚起について - 音楽を用いて -. 日本心理学会 (第 66 回大会研究発表集), 2002.
8. 鈴木朋子, 井上果子. ボーダーライン水準の中年女性との心理療法 - 心の境界をめぐって -. 日本心理臨床学会 (第 21 回大会), 2002.
9. 鈴木朋子. 統合失調症患者の感情の表出について - 音声による表出 -. 日本心理学会 (第 67 回大会研究発表集), 2003.
10. 鈴木朋子. 感情を込めた音声に関する研究 (3). 日本感情心理学会第 11 回大会, 2003.

付録 1

BDI の質問紙

BDI—I

この質問紙のそれぞれの質問文をよく読んでください。そして最近のあなたの気持ちを最もよく表している質問文を、各質問の中からそれぞれ1つ選択して、その番号に○をつけてください。それぞれの質問に同じ程度の質問文（選択肢）が複数あれば、複数に○をつけてください。○をつける前に、各質問の質問文を全部読んでください。

第1問 0. 私は落ち込んでいない。

1. 私はいつも落ち込んでいる
2. 私はいつも落ち込んでいるから急に元気になれない。
3. 私はとてもがまんができないほど落ち込んでいるし不幸だ。

第2問 0. 私の将来について特に失望していない。

1. 私の将来について失望している。
2. 私の将来に期待するものはない。
3. 私の将来には希望がもてないし、物事はよくならないと思う。

第3問 0. 日常生活では大変満足している。

1. 日常生活の出来事を楽しんでいない。
2. 私は何にも本当に満足できない。
3. 私はどんなことにも満足できないし退屈だ。

第4問 0. 私は罰をうけている（いわば罰があたっている）とは思わない。

1. 私は罰せられるかも知れないと思う。
2. 私は罰せられるだろうと思う。
3. 私は罰せられていると思う。

第5問 0. 私は自分自身に失望していない。

1. 私は自分自身に失望している。
2. 私は自分自身にいや気がさしている。
3. 私は自分自身が嫌いだ。

第6問 0. 私は自分は他の人よりは良くないとは思わない。

1. 私は自分の弱さや失敗について自分自身を責めている。
2. 私は自分の欠点をいつも自分のせいにする。
3. いやなことが起こるとそれを自分自身のせいにする。

- 第7問 0. 自殺について考えたことはない。
1. 自殺について考えたことはあるが実行したことはない。
 2. 自殺したいと思う。
 3. もし機会があったら自殺するだろう。

- 第8問 0. 私はいつもより悲しい気持ちではない。
1. 私はいつも悲しい気持ちでいる。
 2. 私は今はいつも悲しい気持ちでいる。
 3. 私はいつも泣いていたが今では泣きたいと思っても泣けない。

- 第9問 0. 私はこれまでのようにいらいらしない。
1. 私は今までより簡単に悩んでしまうしいらいらする。
 2. 私はいつでもいらいらを感じる。
 3. 私はいらいらすらも感じなくなった。

- 第10問 0. 私は自分なりの判断力がある。
1. 私は今までのような判断力に乏しい。
 2. 私は依然よりも物事の判断に困難を感じる。
 3. 私は何も判断することができない。

- 第11問 0. 私は自分の現実よりは悪く見えない。
1. 私は年を取り魅力を失って見えるのではないか気になる。
 2. 私はだんだん魅力がなくなったように思う。
 3. 私は自分の見かけが見苦しくなって来たと感じている。

- 第12問 0. 私は以前と同様に仕事ができる。
1. 何かをしようとする時、前よりも余分な努力をしなければならない。
 2. 何かを始める時、うんと頑張らなくてはならない。
 3. 私は何もしたくない。

- 第13問 0. 私はいつものようによく眠れる。
1. 私はいつものようにはよく眠れない。
 2. 私は以前よりも1～2時間早く目がさめるし、それからはなかなか眠れない。
 3. 私は以前より数時間早く目がさめるし再び眠れない。

第14問 0. 私はいつもより特に疲れたりはしない。

1. 私は以前よりも簡単に疲れる。
2. 私は何をやっても疲れる。
3. 私は余り疲れるので何もできない。

第15問 0. 食欲は普通だ。

1. 食欲は以前よりは良くない。
2. 今は食欲がない。
3. 食欲が全くない。

第16問 0. 性についての興味は特に変わっていない。

1. 以前より性に対する興味が減少した。
2. 今では性に対する興味が大変減少した。
3. 性に対する興味が全くなくなった。

お疲れ様でした。

調査はこれで終了です。

ご協力ありがとうございました。

付録 2

基本感情及び音声に表れる感情表現の調査用質問紙

あなたは今、どのような感情をもっていますか。
最も近いものに○をつけて下さい。

		あてはまらない	あまりあてはまらない	どちらでもない	少しあてはまる	あてはまる
1	悪意	1	2	3	4	5
2	喜び	1	2	3	4	5
3	憤り	1	2	3	4	5
4	安心	1	2	3	4	5
5	反省	1	2	3	4	5
6	尊敬	1	2	3	4	5
7	恨み	1	2	3	4	5
8	ひがみ	1	2	3	4	5
9	興奮	1	2	3	4	5
10	くつろぎ	1	2	3	4	5
11	苦しさ	1	2	3	4	5
12	焦躁	1	2	3	4	5
13	疎隔	1	2	3	4	5
14	利己心	1	2	3	4	5
15	公共心	1	2	3	4	5
16	不信用	1	2	3	4	5
17	愛好	1	2	3	4	5
18	尊重	1	2	3	4	5
19	自棄	1	2	3	4	5
20	落ち着き	1	2	3	4	5
21	名誉	1	2	3	4	5
22	嫉妬	1	2	3	4	5
23	同情	1	2	3	4	5
24	憎悪	1	2	3	4	5
25	心配	1	2	3	4	5
26	恥	1	2	3	4	5
27	親近	1	2	3	4	5
28	忠孝	1	2	3	4	5
29	恐れ	1	2	3	4	5
30	愛惜	1	2	3	4	5
31	嫌悪	1	2	3	4	5
32	軽蔑	1	2	3	4	5
33	義侠	1	2	3	4	5
34	驚き	1	2	3	4	5
35	悲しみ	1	2	3	4	5

36	狼狽	1	2	3	4	5
37	無視	1	2	3	4	5
38	愛	1	2	3	4	5
39	誇り	1	2	3	4	5
40	鎮静	1	2	3	4	5
41	怒り	1	2	3	4	5
42	悔い	1	2	3	4	5
43	恩	1	2	3	4	5
44	楽しみ	1	2	3	4	5
45	驕り	1	2	3	4	5
46	卑下	1	2	3	4	5
47	信用	1	2	3	4	5
48	満足	1	2	3	4	5
49	善意	1	2	3	4	5
50	虚栄	1	2	3	4	5
51	人情	1	2	3	4	5
52	自信	1	2	3	4	5
53	感謝	1	2	3	4	5

あなたは、人の声(文の意味をのぞいて音自体)にどのような感情が含まれていると思いますか。
あなたの考えに最も近いものに○をつけて下さい。

		あ て は ま ら な い	あ ま り あ て は ま ら な い	ど ち ら で も な い	少 し あ て は ま る	あ て は ま る
54	驕り	1	2	3	4	5
55	狼狽	1	2	3	4	5
56	同情	1	2	3	4	5
57	善意	1	2	3	4	5
58	誇り	1	2	3	4	5
59	驚き	1	2	3	4	5
60	利己心	1	2	3	4	5
61	公共心	1	2	3	4	5
62	不信用	1	2	3	4	5
63	反省	1	2	3	4	5
64	悔い	1	2	3	4	5
65	苦しさ	1	2	3	4	5
66	親近	1	2	3	4	5
67	尊敬	1	2	3	4	5
68	鎮静	1	2	3	4	5
69	悲しみ	1	2	3	4	5
70	自信	1	2	3	4	5
71	嫌悪	1	2	3	4	5
72	愛惜	1	2	3	4	5
73	愛	1	2	3	4	5
74	尊重	1	2	3	4	5
75	疎隔	1	2	3	4	5
76	恥	1	2	3	4	5
77	悪意	1	2	3	4	5
78	忠孝	1	2	3	4	5
79	義侠	1	2	3	4	5
80	人情	1	2	3	4	5
81	くつろぎ	1	2	3	4	5
82	卑下	1	2	3	4	5
83	虚栄	1	2	3	4	5
84	憎悪	1	2	3	4	5
85	ひがみ	1	2	3	4	5
86	愛好	1	2	3	4	5
87	喜び	1	2	3	4	5
88	安心	1	2	3	4	5

89	自棄	1	2	3	4	5
90	嫉妬	1	2	3	4	5
91	恐れ	1	2	3	4	5
92	楽しみ	1	2	3	4	5
93	軽蔑	1	2	3	4	5
94	感謝	1	2	3	4	5
95	興奮	1	2	3	4	5
96	心配	1	2	3	4	5
97	恨み	1	2	3	4	5
98	無視	1	2	3	4	5
99	怒り	1	2	3	4	5
100	落ち着き	1	2	3	4	5
101	恩	1	2	3	4	5
102	慎み	1	2	3	4	5
103	信用	1	2	3	4	5
104	焦躁	1	2	3	4	5
105	満足	1	2	3	4	5
106	名誉	1	2	3	4	5

ご協力ありがとうございました。

付録 3

感情尺度

あなたは今の声を聞いて、どのような気持ちが声に含まれていると思いましたか？
当てはまる番号(1-5)を入力してください。

1…あてはまらない
2…あまりあてはまらない

3…どちらでもない

4…少しあてはまる
5…あてはまる

番号(1-5)

楽しみ	
喜び	
満足	
狼狽	
悲しみ	
鎮静	
憎悪	
信用	
心配	
悪意	
落ち着き	
嫌悪	