

知覚された音声特性を9線分上に記述する試み

伊藤 精英*・福島 脩美・西脇 朗子*

福田ひかり*・中川 恵*

教育心理学**

(1990年10月30日受理)

I 問題と目的

我々の日常生活において、音声 (speech) によるコミュニケーションの果たす役割が大きいことは見逃せない事実である。音声を媒介として行なわれるコミュニケーションは、言語的意味内容の授受としての言語コミュニケーションと言語以外の要素 (準言語) の授受としての非言語コミュニケーション (nonverbal communication) とに大別することが出来る。これらのコミュニケーションの機能の主なものとして以下の2点が指摘されている (宇津木1982)。

- ① 相手の姿が視覚的に知覚されない状況 (暗闇とか電話など) において有効である。
- ② それゆえ危急を知らせる有効な手段となる。

言語コミュニケーションにおいては、言語的意味内容の表出・認知がその過程で行なわれるのにたいし、ノンバーバルコミュニケーションにおいては話者の態度・印象・情緒状態が表出され、認知される。

ところで音声コミュニケーションはどのような過程なのだろうか。エンコーディング過程は話手が意志や考えを言語形式に記号化し具体的な音声波を発生する過程である。言語音の生成過程には音声器官の生理的運動である呼気、発声、調音 (articulation) が含まれる。生成された音声波には、言語の意味内容とその意味内容を強調し修飾したり話手の意志や感情を暗示する情報 (準言語) とが含まれる。表情や動作が非言語的情報を伝えるのに対し、音声は言語と非言語との情報を同時に伝える機能を持つ。この意味で他の視覚的チャンネルとは異なった特徴を有しており、この事が準言語についての実験的研究を困難にしているのである。

一方デコーディング過程は聴覚系における音声波の分析中枢神経系における情報処理 (音声知覚、認知) である。言語内容のエンコーディングやデコーディングの問題は、言語学や認知心理学などの学問分野で主として扱われる。準言語的内容のエンコーディング・デコーディングの問題は主として心理学のノンバーバルコミュニケーションの問題として扱われる。

では、声の調子 (tone of voice) を手掛かりとするノンバーバル コミュニケーションの研

* 東京学芸大学教育学研究科

** 東京学芸大学 (184 小金井市貫井北町4-1-1)

究はどのような分野に別れるのであろうか。ここで我々は、宇津木 (1982) を参考に以下の4領域に分けて考えたい。

①特定の行動を固体の特定の属性の指標とする研究

具体的には特定の音声の音響学的特性を生理過程や身体の形質、あるいは心理的過程と関連させる研究がこれに属する。

②特定の行動が果たす機能の研究

会話行動の研究がこれに属する。

③音声の調子 (tone of voice) を用いた対人知覚研究

④情動表出・認知に関する研究

①の領域では主として音響学的分析の手法を用いて半世紀ほど以前から研究が行なわれている。③、④の分野では数多くの研究が積み重ねられているが、②の分野は相対的に見てあまり研究数が多いとは言えない。ここでは音声による情動表出における音声の音響学的分析研究について幾つかを挙げておくことにする。

Streeter, Kraus, Olson and Apple (1977) は、嘘をつく時と真実を話す時との2条件下で基本周波数 (ピッチ周波数) を比較した。その結果、真実を話す時より嘘を発言する時の方がピッチ周波数が高くなることを見出された。ここで言う基本周波数とは、向かい合う声帯によって構成されている声門の開閉によって生ずる周波数のことである。基本周波数が変化する要因として声門における呼気圧の増加と声帯筋の緊張の両者が係わっている。

Williams and Stevens (1972) は、3名の俳優に書下ろしの短い台本に基づき演技をさせ、怒り、恐れ、悲しみの各場面と感情的に中性な場面の音声をスペクトログラムによって分析した。パラメータとしては基本周波数の波形、平均スペクトル、一時的な特徴、調音である。俳優により感情場面の特性は、かならずしも一貫してはいないものの、感情場面によりこれらの音響学的特徴に違いがあることが認められた。杉藤、稲田 (1977) も俳優 (男性1名) の舞台台詞の内、平静時、喜び、悲しみ、怒りの感情場面の音響パラメータの分析を行なっている。その結果は、感情場面それぞれに音響学的特徴が表れた。杉藤らの結果は、Williamsらの結果と一致するものであった。すなわち、感情表出に関連の深い音響学的特性は基本周波数の値とその範囲である。怒りの基本周波数は、平均値が高く範囲も広い。悲しみにおいては平均値、変化範囲ともに小さい。

これらの研究により感情のエンコーディングの指標としての音響学的特性を取り上げることが妥当であるといえよう。しかしながら、音響学的分析を行なうためには録音条件を整え高品質の音質が要求される。現実生活に表れる感情の表出をそのまま周波数分析することは容易ではないのも事実である。そして感情の種類を指標としてピッチ周波数を利用出来るからといって感情のデコーディング過程でピッチ周波数とその手掛かりとして使用されているとはかぎらないことに注意するべきである。ピッチ知覚については以下の指摘がある (中田1977)。

①知覚されたピッチが上がり調子であるにもかかわらず、ピッチ周波数はフラットであるかむしろじゃっかん下降傾向にあることもある。

②ピッチが下降していると知覚されたにもかかわらずピッチ周波数はフラットあるいは上昇していることがある。

したがってピッチ周波数の変化波形を感情のデコーディングの指標とするにはかならずしも妥当であるとは言えない。

ピッチは準言語を構成する1要素である。感情伝達手段としての準言語の重要性は、Zaidel and Mehrabian (1969), Mehrabian (1970) などの言語と音調のメッセージの矛盾に関する研

究を通じて強調されて来ている。感情のディコーディング過程を研究するにあたっては、ピッチ変化などの準言語的の知覚に注目する必要がある。

そこで本研究では知覚された準言語的要素を量的質的な分析の可能な様式で記述する方法を開発し、その信頼性と音響学的物理量との相互関連性を検討することを目的とする。

II 方 法

1 評定方式

(1) 準言語的 (paralinguistic) 特性の選択

準言語的要素の内、主要なもの三つを記述の対照とした。すなわち、音の高低(ピッチ)、音量、速さの変化、である。

(2) 記述方式

ピッチ… ピッチ感覚の変化は、9本の線分を描いた用紙に直線あるいは曲線で表現し、線分間の中間値の表現はその中央値として読みとることとした。したがって段階値は17となる。

音量… 音量は7段階評価(とても大きい(+3)からとても小さい(-3))とした。

速さの変化… 音量と同様に7段階で評定することとした。

(3) 評定用紙

ピッチの曲線を描くために、9本の直線を等間隔に引き、その上方に刺激文を記したものを評定用紙とした。(付図参照)

(4) 評定手続き

評定は、1対比較法を用いた。基準となる刺激を提示しそのピッチ曲線を評定用紙に記述する。次に刺激語を提示し、基準のピッチ曲線、音量、速さと比較して、ピッチ曲線を描き音量と速さを評定し、その数値を用紙に記入する。

(5) 評定者

評定者は、ピッチ感覚の正確さなどを考慮し学芸大学で音楽教育研究科に在学する大学院生女子3名とした。評定者は、幼児期よりピアノ演奏の経験があり、ソルフエージュ(聴音)の訓練を受けてきている。

本研究で扱われる音響学的特性の記述は、相対的知覚に基づくものであり、聴取実験に入るに当たり3名の評定者のピッチ音量、速さの基準を一致させる必要があった。そこで1週間に1回から2回程度の訓練をほぼ4週間にわたって行なった。その結果基準の一致が見られたと判断し聴取実験を行なった。

2 信頼性の検討

(1) 聴取材料

本研究では、感情表出に伴う音声に含まれる準言語的特性の変化を記述することを試みた。

22名の表出者に「こんにちは」、「あれ、どうなったの」、「はい」と言いながら喜び、悲しみ、驚き、恐怖の5種類の感情を表出することを求めた。実験の冒頭で3短文をなにも感情を込めずに話すことを求めた。これらは評定の基準パターンとされた。表出者の発話内容は録音され基準となる刺激と感情表出刺激が対提示されるように編集された。基準刺激提示後約5秒で感情表出刺激が提示され刺激の対提示を1試行とする。試行間隔は10秒とした。

(2) 手続き

まず基準刺激を提示し評定者はそれを聴取しピッチ曲線を用紙に記入する。刺激は3名が記

述し終るまで繰返し（1回～3回）提示される。基準曲線は間隔尺度上で0を示す直線上に記述される。次に感情表出刺激が提示され評定者は基準パターンと比較して相対的位置にピッチ曲線を描く。音量、速さにも変化があると判断されればその変化を用紙に7段階で評定値を記入する。刺激の提示順序はランダムとし、且つ評定者の期待効果を除去する為に表出された感情の種類、その他表出者に関する情報は一切評定者には與えられなかった。

(3) 信頼性の分析

信頼性は、3側面から検討された。

①ピッチの一致率…ピッチ感覚が一致しているか否かを視察によって判定を行なった。視察の方法は3者のピッチ波形を重ね併せ各音節ごとに一致、不一致を17段階（9線分の中間の値まで読みとって17段階値）で確認し、「一致した音節数」の「全音節」に対する割合を一致率とした。

②波形の一致率…ピッチ感覚がずれていても波形が一致する場合が考えられる。そこで波形が一致している度合いを求める事とした。17段階尺度上で各音節のとり値を求め、後続の音節の取る値（例…「はい」の「い」）先行する音節の取る量（「はい」の「は」の値）との差の評定者間の一致度をもって一致率とした。

③再検査信頼性

ほぼ1箇月をおいて再検査を行い評定者内一致率を求めた。

3 相互関連性の検討

聴取実験で使用された表出者の内、女子1名を選び、その音声についてパーソナルコンピュータによるピッチ分析(AD変換を行なうソフトウェアを使用)を行い基本周波数とスペクトログラムを求めた。「こんにちは」、「はい」の音声刺激計12個が分析対象とされた。スペクトログラムは、音圧強度により4段階に分け数値化し、それを図示したものと及び基本周波数の波形とピッチ曲線とが比較された。

III 結 果

本研究はピッチ曲線を記述することの有効性を検証することを主たる目的としているため、その結果のみについて報告し分析を行なうこととする。

1 信頼性

(1) ピッチ

ピッチの評定に関する一致度についてまず詳述する。以下、評定者A、評定者B評定者Cの3名のピッチ曲線が一致した場合（以下3者一致）と評定者のうち任意の2名のピッチ曲線が一致した場合（2者一致）とに便宜上分けることにする。

Table 1に3者一致の人数、不一致人数及び一致率を、Table 2に2者一致の人数不一致人数及び一致率を示す。

3者一致については「こんにちは」の基準刺激に関する一致率は81.0%であり、感情刺激合計の一致率は10.5%であった。刺激語「はい」については基準刺激で81.0%であり、感情刺激合計の一致率は6.7%であった。刺激語「あれ どうなったの」では基準刺激及び感情刺激合計の一致率はそれぞれ0%であった。一方2者一致の成績を見ると、刺激文「こんにちは」での基準刺激一致率は100%であり、感情刺激合計の一致率は58.1%であった。刺激文「はい」では

基準刺激及び感情刺激合計の一致率はそれぞれ100%, 47.6%であった。刺激文「あれどうなっ

Table 1 ピッチに関する3者一致の人数と一致率

刺激語	一致	基準	幸福	悲しみ	驚き	恐れ	怒り	感情表出刺激計
こんにちは (音節5)	一致数	17	1	5	0	1	4	11
	不一致数	4	20	16	21	20	17	94
	一致率	81.0						10.5
あれどっ な ったの (音節8)	一致数	0	0	0	0	0	0	
	不一致数	22	22	22	22	22	22	
	一致率	0						0
はい (音節2)	一致数	17	1	2	1	3	0	7
	不一致数	4	20	19	20	18	21	98
	一致率	81.0						6.7

Table 2 ピッチに関する2者一致の人数と一致率

刺激語	一致	基準	幸福	悲しみ	驚き	恐れ	怒り	感情表出刺激計
こんにちは (音節5)	一致数	21	13	5	0	1	4	61
	不一致数	0	8	16	21	20	17	44
	一致率	100	61.9	57.1	38.1	66.7	66.7	58.1
あれどっ な ったの (音節8)	一致数	9	2	2	0	2	2	8
	不一致数	13	20	20	22	20	20	102
	一致率	40.9	9.1	9.1	0	9.1	9.1	7.3
はい (音節2)	一致数	21	11	9	7	14	9	50
	不一致数	0	10	12	14	7	55	
	一致率	100	52.4	42.9	33.3	66.7	66.7	47.6

たの」では基準刺激の一致率が40.9%であり、感情刺激合計の一致率は7.3%であった。

これらの結果を特徴付けていることは、以下のことである。

①刺激文についてみると3者一致、2者一致ともに「こんにちは」がもっとも一致度が高く、ついで「はい」であり、もっとも低い一致率を示したのが「あれ どうなっ」であった。

②3者一致、2者一致ともに基準刺激の方が感情刺激よりも一致率が高かった。3者一致における感情刺激合計の一致率が全般に低い(10.5%, 6.7%, 0%)ため、2者一致の場合のみ詳述する。刺激文「こんにちは」の場合「恐れ」と「怒り」が最も高い一致率で66.7%を示した。以下、「喜び」(61.9%), 「悲しみ」(57.1%), 「驚き」(38.1%)の順であった。刺激文は

「はい」について見ると、「恐れ」(66.7%)、「喜び」(52.4%)、「悲しみ」・「怒り」(42.9%)、「驚き」(33.3%)の順で一致率が高かった。刺激文「あれ どうなったの」では、「怒り」・「恐れ」・「悲しみ」・「喜び」の4感情の一致率は9.1%であるが、「驚き」の一致率は0%であった。

(2) 波形

ピッチの場合同様に3者一致と3者一致のそれぞれについて記述することとする。Table 3に波形に関する3者一致の不一致人数と一致率を示し、Table 4に2者一致の人数と不一致人数及び一致率を示す。ピッチの一致とは、波形が一致したことに加え、音程が一致することが条件となる。それに対し波形一致とは、多少の音程の不一致が見られたとしても記述された曲

Table 3 波形に関する3者一致の人数と一致率

刺激語	一致	基準	幸福	悲しみ	驚き	恐れ	怒り	感情表出刺激計
こんにちは (D=4)	一致数	17	15	16	7	17	17	72
	不一致数	4	6	5	14	4	4	33
	一致率	81.0	71.4	76.2	33.3	81.0	81.0	68.6
あれどっ なったの (D=7)	一致数	0	0	1	0	0	0	1
	不一致数	22	22	21	22	22	22	109
	一致率	0	0	4.8	0	0	0	0.9
はい (D=1)	一致数	18	11	9	7	14	9	80
	不一致数	3	10	12	14	7	12	25
	一致率	85.7	85.7	81.0	57.1	85.7	71.4	76.2

Table 4 波形に関する2者一致の人数と一致率

刺激語	一致	基準	幸福	悲しみ	驚き	恐れ	怒り	感情表出刺激計
こんにちは (D=4)	一致数	21	20	21	17	20	21	99
	不一致数	0	1	0	4	1	1	6
	一致率	100	95.2	100	81.0	95.2	100	94.3
あれどっ なったの (D=7)	一致数	11	9	7	9	8	8	41
	不一致数	11	13	15	13	14	14	6.9
	一致率	50.0	40.9	31.8	40.9	36.4	36.4	37.3
はい (D=1)	一致数	21	20	21	19	21	21	102
	不一致数	0	1	0	2	0	0	3
	一致率	100	95.2	100	90.5	100	100	97.1

線の波形 (contour) が一致していることを意味する。

まず3者一致について見ることにする。刺激文「こんにちは」の場合、基準刺激の一致率は81.0%、感情刺激合計の一致率は68.6%であった。刺激文「はい」については基準刺激で85.7%の一致率を示し、感情刺激合計で76.9%の一致率であった。刺激文「あれ どうなったの」では、基準刺激の一致率は、0%であり、感情刺激合計の一致率は0.9%にすぎなかった。2者一致の場合を見ると、刺激文「こんにちは」では基準刺激100%、感情刺激合計94.3%の一致率が認められた。刺激文「はい」においては基準刺激では100%、感情刺激合計の一致率は97.1%を示した。刺激文「あれ どうなったの」では、基準刺激で50.0%、感情刺激合計で37.3%の一致率が認められた。

これらの結果に見られる一貫した傾向は、3者一致より2者一致の一致率が高いこと、3者一致、2者一致共に刺激文「はい」が最も一致率が高く、ついで「こんにちは」であり、「あれ どうなったの」が最も低いこと、基準刺激の一致率が感情刺激の一致率より高いことなどである。刺激文「はい」と「こんにちは」では基準刺激で80%以上の一致率であり、感情刺激合計で70%前後の一致率を示しており、2者一致にいたっては基準刺激で100%、感情刺激合計で90%以上の高い一致率を示している。

これら2刺激文に関するかぎり、波形の一致度は極めて高いといえる。しかし、刺激文「あれ どうなったの」では2者一致において基準刺激の一致率が50%、感情刺激合計の一致率が40%弱と他に比べ極端に低い値にとどまっていることも事実である。

次に感情の性質による一致率の違いを述べることにする。3者一致の場合、刺激文「こんにちは」では、「恐れ」と「怒り」の一致率が81.0%と最も高く、次いで「悲しみ」(76.2%)、「喜び」(71.4%)、そして「驚き」は最も低く33.3%であった。刺激文「はい」については、「恐れ」と「喜び」が最も高く(85.7%)、次いで「悲しみ」(81.0%)、「怒り」(71.4%)、「驚き」(57.1%)であった。刺激文「あれ どうなったの」では等しく0%あるいはそれに近い値を示した。

2者一致の場合「こんにちは」で「驚き」が81.0%で最も低く、他は95%以上の高い一致率を示した。「はい」の場合は全感情の一致率が90%を越えているものの、相対的には「驚き」の一致率が低かった(90.5%)。「あれ どうなったの」では30%から40%の範囲に一致率が集中していた。感情の性質により一致率が異なる事がピッチの場合と同様に明かとなり、「恐れ」が最も高く、「驚き」が最も低いという結果は注目に値する。

2 再検査による評価の一貫性

表出者のうち任意の7名を選び、再検査を行なった。Table 5は、ピッチ曲線の評定者別一致率を示している。またTable 6には、波形の評定者別一致率を示す。全音節が一致した場合を完全一致とし、1音節のみが不一致な場合を準一致として記述することとする。

なお結果が煩雑になることを避けるために、ここでは基準刺激についてのみ報告する。

完全にピッチ曲線が一致した場合を見ると、評定者Aでは刺激文「こんにちは」において100%、「あれ どうなったの」では28.6%、「はい」では100%であった。評定者Bの場合を見ると、刺激文「こんにちは」の一致率は57.1%、「あれ どうなったの」では28.6%、「はい」で71.4%であった。評定者Cの一致率は刺激文「こんにちは」、「あれ どうなったの」、「はい」それぞれ57.1%、42.9%、100%であった。最も高い内的一貫性を示したのは評定者Aであった。評定者Bは30%程度から70%の中程度の一致率を示し、評定者Cは40%程度から100%という中程度から非常に高い一致率を示していた。

準一致についてであるが、刺激文「こんにちは」の場合評定者BとCの準一致率は共に85.7%

の価である。「あれ どうなったの」では評定者Aは85.7%、評定者Bが57.1%、評定者Cが85.7%の一致率を示している。「はい」の場合は評定者Bで85.7%、評定者Cで100%であった。

完全一致の場合「はい」の評定者内一致率がかなり高く、次いで「こんにちは」であり最も評定者内一致率の低い文は「あれ どうなったの」であった。準一致になると三つの刺激文を通じて3名の評定者全て一致率の上昇が認められる。一致率は57%から100%の間に分布している。準一致の場合刺激文の種類にかかわらず3名の個人内一致率は高いことが明かとなった。しかしながら3名の評定者は、全て絶対的(1名は相対的)音感を習得しているにもかかわらず完全なピッチの一致にばらつきや個人差が認められた事は、ピッチ知覚の変動性を示唆するものと思われる。

次に音程を無視してピッチ曲線の波形(contour)の評定者内一貫性について見

ることとする。評定者Aについては完全一致の場合「こんにちは」と「はい」の一致率は100%であるが、「あれ どうなったの」では28.6%にすぎない。準一致になると「あれ どうなったの」で85.7%の一致率を示している。評定者Bについては完全一致のばあい「はい」で100%の一致、「こんにちは」で71.4%の一致、「あれ どうなったの」では42.9%の一致がそれぞれ認められた。準一致になると「こんにちは」で100%の一致、「あれ どうなったの」のばあいで57.1%の一致を示し共に上昇していることが分かる。評定者Cの場合についてであるが、完全一致の一致率は刺激文「こんにちは」、「はい」、「あれ どうなったの」それぞれ57.1%、100%、42.9%である。準一致になると「こんにちは」で100%、「あれ どうなったの」で85.7%と一致率の上昇が認められる。

3名の評定者のうち最も個人内一致率の高かった者は評定者Aであり、以下評定者B、評定者

Table 5 ピッチに関する評定者別評定の一致度 (基準刺激の再検査法)
評定者

刺激語	一致率	A	B	C
こんにちは	完全一致 (5 / 5)	100%	57.1%	57.1%
	準一致 (4 / 5 以上)	—	85.7%	85.7%
あれどうなったの	完全一致 (8 / 8)	28.6%	28.6%	42.9%
	準一致 (7 / 8 以上)	85.7%	57.1%	85.7%
はい	完全一致 (2 / 2)	100%	71.4%	100%
	準一致 (1 / 2 以上)	—	85.7%	—

(数値は話し手7人中の一致した人数の比率)

Table 6 波形に関する評定者別評定の一致度 (基準刺激の再検査法)
評定者

刺激語	一致率	A	B	C
こんにちは	完全一致 (4 / 4)	100%	71.41%	57.1%
	準一致 (3 / 4 以上)	—	100%	100%
あれどうなったの	完全一致 (7 / 7)	28.6%	42.9%	42.9%
	準一致 (6 / 7 以上)	85.7%	57.1%	85.7%
はい	完全一致 (1 / 1)	100%	100%	100%

(数値は話し手7人中の一致した人数の比率)

Cの順である。ピッチ曲線の場合に比べて波形の一致率に上昇が認められた者は、評定者Bであり他の2名はピッチ曲線と波形の一致率との間にほとんど差が表れなかった。

Table 7 ピッチに関する評定者別評定の一致度 (話し手7人中一致した人数)

刺激語	一致基準	評定者		
		A	B	C
こんにちは	完全一致 (5/5)	7/7	4/7	4/7
	準一致 (4/5以上)	—	6/7	6/7
あれどうな ったの	完全一致 (8/8)	2/7	2/7	3/7
	準一致 (7/8以上)	6/7	4/7	6/7
はい	完全一致 (2/2)	7/7	5/7	7/7
	準一致 (1/2以上)	—	6/7	—

(単位は人数)

Table 8 波形に関する評定者別評定の一致度 (話し手7人中一致した人数)

刺激語	一致率	評定者		
		A	B	C
こんにちは	完全一致 (4/4)	7/7	5/7	4/7
	準一致 (3/4以上)	—	7/7	7/7
あれどうな ったの	完全一致 (7/7)	2/7	3/7	3/7
	準一致 (6/7以上)	6/7	4/7	6/7
はい	完全一致 (1/1)	7/7	7/7	7/7

(単位は人数)

3 音響学的特質との相互関連性

三つの刺激文のうち、「こんにちは」と「はい」について周波数分析による結果とピッチ曲線の記述との対応関係について述べる。

まず周波数分析の結果をFig. 1とFig. 2に示す。また, Fig. 3とFig. 4に基本周波数の波形を示す。刺激文「こんにちは」の周波数スペクトルに表れた特徴は以下のとおりである。刺激文「こんにちは」での基準刺激のフォルマントは概してフラットな状態である。「悲しみ」と「恐れ」についても基準刺激に近い波形であり, フラットになっている。「驚き」では概してやや上昇した後にフラットな状態になる。「喜び」では前半部に上昇が見られ, 後にフラットな状態を維持する。「怒り」の場合概してフォルマントの波形は急な上昇の後, 多少の下降が見ら

Fig. 1 スペクトrogram (刺激文「こんにちは」の場合)

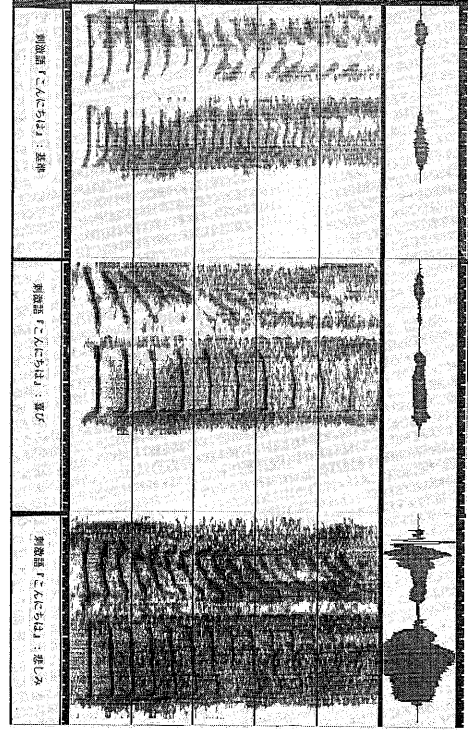
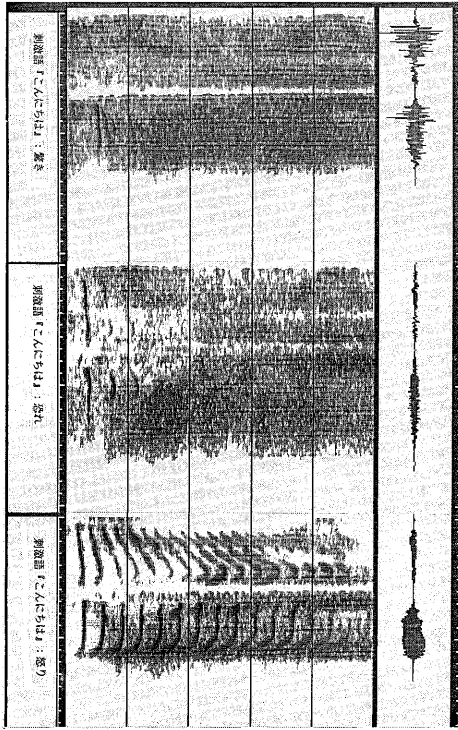
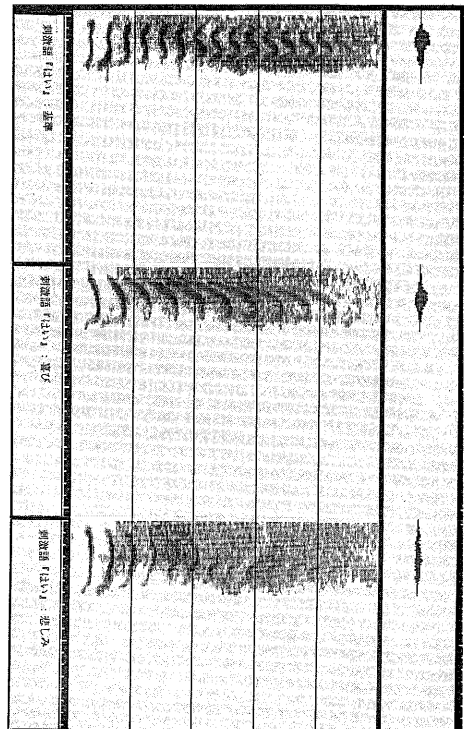
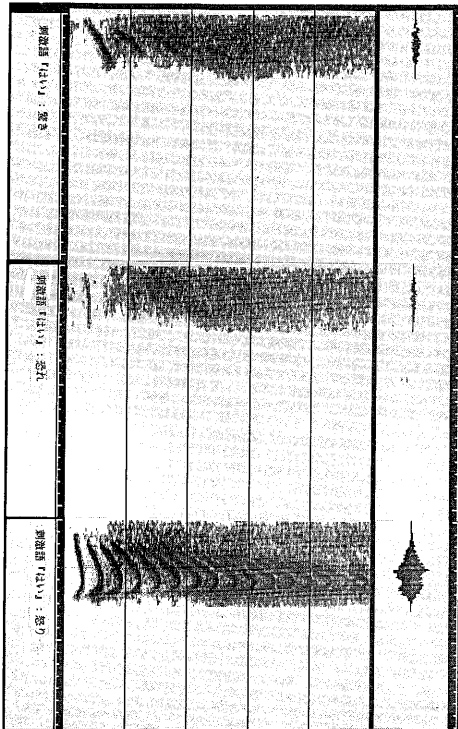


Fig. 2 スペクトrogram (刺激文「はい」の場合)



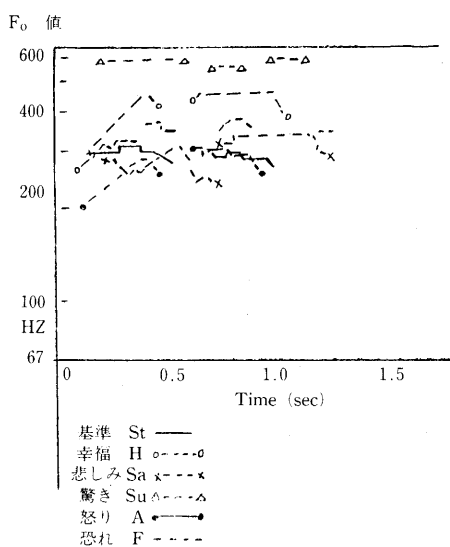


Fig.3 基本周波数波形
(刺激文「こんにちは」の場合)

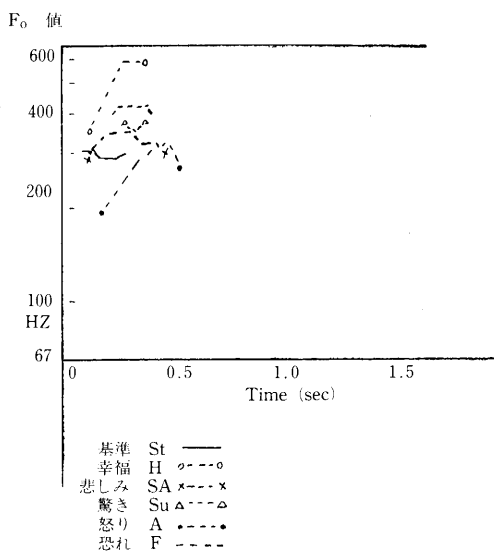


Fig.4 基本周波数波形
(刺激文「はい」の場合)

れる。フォルマントの数は「基準」では13, 「怒り」で15, 「悲しみ」で13ほど読みとれ, これらはフォルマント数が多く間隔が狭いことが特徴である。「恐れ」と「驚き」ではフォルマント数が少ない。「喜び」では前半6本のフォルマントが読みとれ, それらは上昇的であるが後半ではフォルマントは9本になりフラットとなる。

基本周波数の高い方から順に列挙するならば「驚き」, 「喜び」, 「恐れ」, 「悲しみ」, 「基準」, 「怒り」となる。

つぎに刺激文「はい」について見ると, フォルマントの形では「基準」が概してフラットであり, 「恐れ」がそれに順ずる。「悲しみ」はゆるやかな弓がたを示し, 「驚き」は直線の上昇を示す。「喜び」では概して上昇が急であり下降が少ない。「怒り」においては大きな上昇と大きな下降が見られる。フォルマント数では「基準」が17と最も多く, 次いで「怒り」が14, 「喜び」が10, 「悲しみ」が7, 「驚き」では4, 「恐れ」では2である。ただし「驚き」, 「恐れ」ではともにフォルマントが不明瞭である。基本周波数の平均値の高い順に列挙すると, 「喜び」, 「恐れ」, 「驚き」, 「悲しみ」, 「基準」, 「怒り」となる。

次に評定されたピッチ曲線について述べる。

①刺激文「こんにちは」

「基準」では3者とも一致した曲線を描いている。「喜び」では評定者Cを除き「こん」から「にちは」にかけて上昇が見られる。ただし評定者AはBより勾配が急で大きい。「悲しみ」では「こん」で基準値すなわち0あるいはじゃっかん低い値をとり「にちは」にかけてじゃっかんの上昇をする。「驚き」では「こん」で0.5から1.5という高い価であり後半に掛けて0.75から1.85の値まで3名の記述した曲線は上昇する。評定者

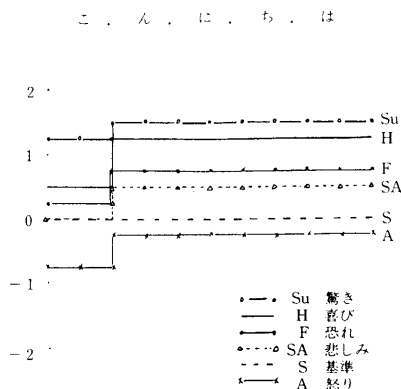


Fig.5 評定者Aによるピッチ曲線

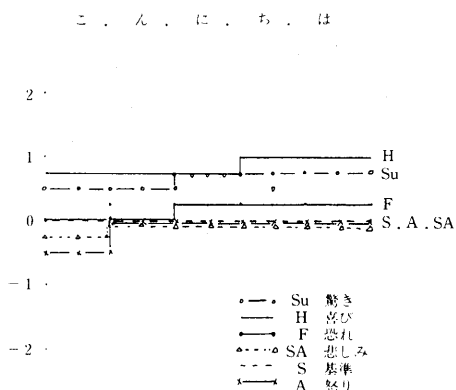


Fig.6 評定者Bによるピッチ曲線

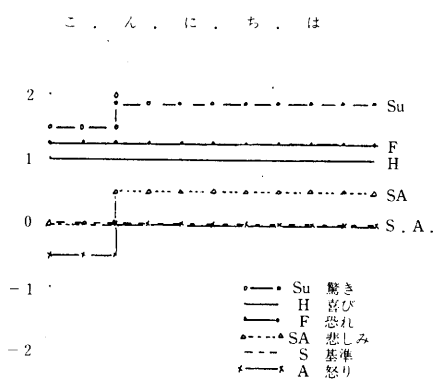


Fig.7 評定者Cによるピッチ曲線

A, Cでは急上昇する曲線である。「恐れ」については評定者A, Bの曲線では前半部で基準値に近い値(0, 0.25)から後半部にかけて上昇が見られる。ただし評定者Aの方が急勾配である。評定者Cは基準より高い値である1.25程度でフラットとなっている。「怒り」の場合3名の評定者とも基準より低い値から基準値あるいはその近似点へと上昇する曲線を記述している。勾配に関するかぎり3名の記述に差は認められない。ピッチの評定に関して高い方から順に列挙するならば、以下ようになる。評定者Aでは「驚き」、「喜び」、「恐れ」、「悲しみ」、「基準」、「怒り」, である。評定者Bにおいては「喜び」、「驚き」、「恐れ」、「悲しみ」、「基準」、「怒り」, である。評定者Cの場合は、「驚き」、「恐れ」、「喜び」、「悲しみ」、「基準」、「怒り」となる。

②刺激文「はい」

「基準」では3名全てフラットな線分を描いている。「喜び」の場合には3名が基準より高い一で線分が記述されフラットである。「悲しみ」では基準値より高い一でフラットな線分が記述されている。「驚き」の場合には評定者Bでは0.5から1.0への上昇直線である。他の2名の記述した直線は基準値よりかなり高い値でフラットである。「恐れ」の場合0.75から1.25の値でフラットな直線である。「怒り」では評定者A, Cは基準値より低い値でフラットな直線である。評定者Bは-0.25から0への上昇直線を描いている。ピッチについて高い順に列挙するならば、以下のごとくである。

評定者Aについて

は、「驚き」、「喜び」、「悲しみ」、「恐れ」、「基準」、「怒り」となる。評定者Bの場合は「喜び」、「驚き」、「恐れ」、「悲しみ」、「基準」、「怒り」の順である。評定者Cを見ると、「驚き」、「喜び」、「恐れ」、「悲しみ」、「基準」、「怒り」となる。

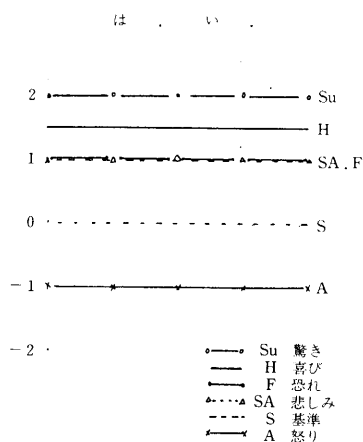


Fig.8 評定者Aによるピッチ曲線

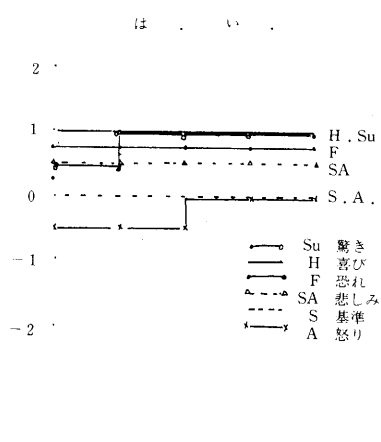


Fig.9 評定者Bによるピッチ曲線

次に周波数分析の結果と評定との関連をしてみることにする。6刺激について知覚されたピッチの順位と基本周波数による6刺激材料の順位とが完全に一致したのは「こんにちは」では評定者Aであり、「はい」では評定者Bであった。また3名全てが「悲しみ」、「基準」、「怒り」の下位3刺激については一致が見られた。波形の対応関係については「こんにちわ」の「喜び」において評定者Aの記述した曲線と F_2 及び F_0 の曲線とがほぼ対応している。同刺激文の「悲しみ」では、700以下のフォルマント波形に3名の記述した曲線が類似している。「驚き」では F_0 の波形と評定者A, Cが類似している。「恐れ」では対応は見られない。「怒り」においては3名の評定者の記述は、 F_0 の波形とほぼ対応している。「基準」の平均基本周波数値は294hzであり300hzに近似する。この値は間隔尺度上で0に相当するからである。以上の結果から知覚されたピッチ曲線とピッチ周波数の波形は、ある程度の対応関係があるといえる。しかしながら、複数の刺激間の知覚された相対的ピッチの順位とピッチ周波数の平均値による材料間の順位とはかならずしも一致しないという結果が示された。このことは、知覚されたピッチ感覚と物理的ピッチとはかならずしも関数的対応を示さないことを示唆するものである。

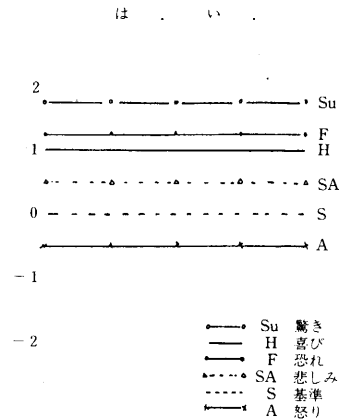


Fig.10 評定者Cによるピッチ曲線

IV 考 察

本研究の主たる目的は、感情のコミュニケーションにおけるディコーディング過程を研究するに当たり、知覚されたピッチ、音量、相対的速度を量的、質的に記述する方法の開発である。そこで本方式の有効性を評定者間一致度及び評定者内一貫性、物理量と感覚量との対応の観点から検討を行なうこととした。ここでは、以上の観点から考察を加える。

評定者間の一致度については、前項において刺激文による差はあるものの、3者間である程度の一致率が得られた。またピッチ曲線よりも波形の一致度が高いことが明かとなった。音声による感情表出時における音響学的特質の分析を行なった幾つかの研究 (Williams & Stevens 1972, 杉藤・稲田 1977) によると、感情のエンコーディングと関連の深い音響学的特質は基本周波数の平均値、波形、振幅である。知覚された波形を記述することは、これらの情報がディコーディングされる過程の研究には不可欠である。本研究において波形の一致度が一部を除き、満足のいくものであったことは、この点に関する限り本方式に対する信頼性が得られたことを意味する。

ところで評定者間一致率の変動に関与した要因として刺激文の長さを上げることができる。評定者間一致率は刺激文「こんにちは」、「はい」、「あれ どうなったの」の順であった。「はい」のように音節数が少なすぎると、調音による音声情報が乏しく、その結果曲線の記述が困難になる。また「あれ どうなったの」のような刺激文では音節数が多いため、調音による音声情報量も増大する。

本方式は継時的-対比較法を用いているため、評定者はピッチ曲線の記述にあたり記憶に頼らざるをえず過剰な情報量によりマスキングされ記憶が不正確となることが考えられる。その結果評定者間の記述にばらつきが表れると思われる。このことから記述された曲線の信頼性を保持するためには、刺激文「こんにちは」のような適当な音節数の音声刺激を選択する必要が

あるだろう。

次に評定者内の記述の安定性について考察を加える。本研究では任意の7名の表出した音声刺激計42個についてほぼ一箇月をおいて再検査を行い評定の安定性を検討した。その結果評定者内の一致率は中程度以上であり、曲線の記述に関するかぎり、評定者内の一貫性は確認された。しかしながら、評定者間の一致率の場合と同様に刺激文によって、一致率に差が認められた。このような差異は音声刺激の波形や音節数によると考えられる。波形の振幅が少なく、波形が単純で音節数が少ないほど評定者内の安定度は高いといえる。上述の結果を総合すると本方式を有効に活用するためには刺激材料となる音声の特性を充分に吟味し、使用する音声刺激によって記述されたピッチ曲線あるいは波形の信頼性が損なわれないようにすることが必要である。

ところで今回の研究に参加した評定者は絶対的あるいは相対的音感を有している。しかながら評定者間、評定者内の一致率を見ると、ピッチ曲線の一致率が100%に達しないことが少なからず認められた。このことから音声のピッチ知覚は変動しやすいと考えられる。

次にピッチの物理量と知覚されたピッチとの対応関係について考察する。前項で示したように幾つかの音声刺激に関して知覚された波形と基本周波数あるいはフォルマント波形との間に対応関係が認められた。ゆえに本方式によりある程度は物理量を心理尺度上に記述することが可能であることが示唆された。

次に知覚されたピッチと物理的ピッチとの関係を見ることにする。複数の刺激間における知覚された相対的ピッチの順位と同音声間の基本周波数の平均値による順位とはかならずしも一致しないことが示された。これは中田(1977)の指摘を支持しており、知覚されたピッチと物理的ピッチとは関数的関係にあるとは考えにくいことを意味している。聞き手のピッチ知覚は話者の発話努力(音の強さ、音色)によって多分に影響を受けていることが考えられる。

最後に今後の課題について述べることにする。

①評定者の訓練

本研究では回数にわたり評定者の聴取及び曲線記述の訓練を行なった。評定者の訓練によりピッチ曲線や波形の正確さが向上するのであろうか。評定の一致度に及ぼす学習の効果を検討する必要がある。

②絶対音感の有無の効果の検討

今回の評定者は音楽を専門とする大学院生であった。ソルフェージュなどの音楽的訓練を受けていない評定者がどのような曲線を記述するのであろうか。記述されたピッチ曲線・波形の信頼性に及ぼす聴音能力の効果を検討する必要がある。

③評定方式の改良

評定方式をさらに鋭敏でかつ簡便なものに改良することによって、この方式の信頼性と妥当性をさらに検討する必要がある。

④感情の性質が評定の一致度に及ぼす効果の検討

本研究では感情の種類により評定者間及び評定者内一致率に差異が見られた。評定の一致度に関与する要因を解明するなかで、音響学的特質の変化に関する詳細な分析が必要となろう。

付記：本研究を行うにあたり次の方々に多大なお力添えをいただきました。東京学芸大学障害児教育研究室の小川仁教授と出口利定助教授には音声の物理的特性の記録法に関してご指導を賜りました。また大学院生の小沢康司さんと董玲玲さんには実験の準備と図表の整理を手伝っていただきました。ここにお名前を記して感謝の意を表します。最後に音声表出や聞き取りに

積極的に協力して下さった東京学芸大学の院生と学部生の方々に心から感謝いたします。

引用文献

1. Mehrabian A. 1970 When Are Feelings Communicated Inconsistently? Journal Of Experimental Research in Personality 4, 198-212
2. 中田和男 1977 音声 日本音響学会編 音響工学講座7 コロナ社
杉藤美代子, 稲田裕子 1977 演劇言語における感情表現 - 基本周波数と持続時間に関して - 音声学会会報, 第155号, July Pp. 13-18
3. Streeter L. A., Krauss R.M, Olson G. C. and Apple W. 1977 Pitch Changes During Attempted Deception Journal of Personality and Social Psychology Vol. 35(5), 345-350
4. 宇津木成介 1982 音声におけるノンバーバル行動 サイコロジー10サイエンス社Pp. 54-59
5. Williams C. E., Stevens K. N. 1972 Emotions and Speech:Some Acoustical Correlates The Journal of the Acoustical Society of America Vol. 52, No. 4, 1238-1245
6. Zaidel S. F. and Mehrabian A. 1969 The Ability to Communicate and Infer Positive and Negative Attitudes Facially and vocally Journal of Experimental Research in Personality 3, 233-241

参考文献

1. 大泉充郎監修, 藤村靖編 1972 『音声科学』 東京大学出版会
2. 境久雄, 中山剛 1978 『聴覚と音響心理』 日本音響学会編 音響工学講座6 コロナ社

付図

番号	氏名
----	----

強さ

2 -----

1 -----

0 -----

-1 -----

-2 -----

速さ

強さ

2 -----

1 -----

0 -----

-1 -----

-2 -----

速さ

A Study on a Method of Describing Perceived Vocal Properties on 9 Parallel Lines

Kiyohide ITO, Osami FUKUSHIMA, Akiko NISHIWAKI,
Hikari FUKUDA and Megumi NAKAGAWA

Department of Educational Psychology

Abstract

It was designed to develop a method which would allow researchers to make further studies on decoding process of nonverbal communication of affect by tone of voice. 22 encoders were asked to express five emotional states (anger, fear, happiness, sadness and surprise) and neutral one, reciting three sentences which were standard content. 132 utterances tape-recorded were randomly compiled and re-recorded on one tape.

Three trained raters, who were all female graduate students majoring in music, were required to listen to utterances one by one and describe a contour (an intonation pattern) on nine lines, rating loudness and rate on seven-point scales, on basis of perceived pitch and intonation of utterances. Results reveal as follows: (1) moderate degree of interrater agreements and high degree of intrarater agreements were found; (2) described contours correspond with fundamental frequency contours, although degree of agreements varies from one emotional state to another. These findings represent that it is necessary to improve this method. In addition, they are going to be discussed from the viewpoint of perception of pitch of voice.