 原 著

音声による感情表現の発達の検討

櫻庭 京子¹⁾ 今泉 敏²⁾ 笥 一彦¹⁾

要 約：就学前後の健常な幼児・児童（N=46）を対象に、「ピカチュウ」という発話を用いて、音声による意図的な感情表現の音響的特性を解析した。その結果、音節長、母音無声化、基本周波数の変化範囲やピークは感情に応じて変化することがわかった。F0が低く、変化範囲の狭い「怒り」、「平静」と、F0が高く、変化範囲の広い「喜び」、「驚き」に2極分化する傾向があった。「悲しみ」のF0の変化範囲は就学前児で高く、就学児で低くなった。母音無声化率は「平静」発話では高く、感情発話で低下した。発達に伴い、東京方言の環境に順応した無声化率の上昇がみられた。音節長の制御には年齢による差はなかった。本研究で得られた健常児の音声による感情表現の知見は、感情表現に問題を抱える児童の音声を評価する場合、有用と考えられる。

索引用語：幼児，児童，感情表現，感情，音響分析

Developmental Changes in Vocal Expression of Emotions

Kyoko Sakuraba¹⁾, Satoshi Imaizumi²⁾, Kazuhiko Kakehi¹⁾

Abstract : This study examined developmental changes in the acoustic characteristics of utterances of /pikachu/, the animated cartoon character, generated by 46 normal preschool and school-age children to intentionally express basic emotions. The results showed that syllable length, vowel-devoicing rate, dynamic range of fundamental frequency (F0), and peak timing significantly changed depending on the intended emotions. F0 was low with a narrow range for “anger” and “neutral,” while it was high with a wide range for “surprise” and “happiness.” For “sadness,” F0 was low in elder children and high in younger children. The vowel-devoicing rate was high for “neutral” but decreased during emotional expressions. It increased with age while adapting to the phonetic requirements of the Tokyo dialect. No significant difference was found in syllable length due to age. The developmental changes in acoustic properties revealed here can serve as a standard to evaluate disorders in vocal expression of emotions in children.

Key words : preschooler, school-age children, vocal expression, emotion, acoustic analysis

名古屋大学大学院人間情報学研究科統合音響情報拠点¹⁾：〒464-8601 名古屋市中種区不老町

東京大学大学院医学系研究科²⁾：〒113-0033 東京都文京区本郷7-3-1

¹⁾Graduate School of Human Informatics, Nagoya University, CIAIR : Furo-cho, Chikusa-ku, Nagoya 464-8601

²⁾Graduate School of Medicine, University of Tokyo : 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033

原稿受理：2000年9月19日

はじめに

感情表現に問題を持つ自閉症児などの¹⁻³⁾ 音声解析・評価する場合、比較対象となる健常者データは必ずしも多くない。特に幼児期の音声による感情表現を解析した研究は少ない。そこで本研究では、就学前後の健常な幼児・児童が意図的に感情表現をした音声の音響的特性を検討した。

感情は自然に表出される場合もあるものの、日常生活では他者との円滑なコミュニケーションを計るため、意図的に感情を抑制したり、本意とは異なる感情を表出する場合が多々ある⁴⁾。換言すれば、感情を意図的に表現する能力は他者とのコミュニケーションにおいて不可欠といえる。本研究では意図的に感情を表出した音声を録音して解析した。

音声による感情は「嬉しい」などの感情を表現する語を使って言語的に表出し得るものの、基本周波数(F0)や発話長など音声の韻律を介しても表出される。しかし、日本語のアクセントや中国語の四声のように韻律は語彙の意味弁別にも関係するので、必ずしも感情情報だけを伝達するわけではない。言語的情報と感情情報が重畳して表出される音声では、何らかの形で両者の折り合いをつける必要が起る。両者の関係は発達的に変化することも考えられる。

生後2ヵ月の乳児の音声を成人が判断すると感情情報は一定程度認知可能であるという報告があり⁵⁾、音声による感情表出は比較的早期から発達し始めていると考えられる。一方、言語情報の表出に関わる音声の諸特性の発達は母語の獲得と深く関わる。平静発話での母音無声化⁶⁾やピッチアクセント⁷⁾に関する発達的研究では、4〜5歳児でも習得途上にあるという報告があり、感情表出に比べて発達が遅い可能性がある。もしそうだとすると、ピッチアクセントや母音無声化、モーラ等時性といった日本語(東京方言)の音声的特徴ないし制約は感情表出によって変化するのだろうか。感情表出も年齢に応じた発達の変化を示すとすれば、音声の特徴と感情表出の関係も発達に依存して変化するのだろうか。本論文の目的は幼児の音声を解析して、以上の疑問に答えを見いだすことにある。

解析対象としてアニメ「ポケットモンスター」⁸⁾のキャラクター、ピカチュウの/pikachuu/という発話を活用した。ピカチュウは自身の名前/pikachuu/とその派生語/pikapika/などの発話だけで、自己の意図や感情を表現している。本研究では、ピカチュウの「喜怒哀驚」を示す表情図をみせて、表情図から「嬉しい」などの感情情報の抽出が可能であった幼児・児童に対

象にして、それらの感情を込めて/pikachuu/と発話させた。

東京方言の/pikachuu/は/ka/にアクセント核がある4モーラ3音節語であり、語頭の/pi/の母音は無声化する傾向がある。また、日本語音声の時間構造に関して、各モーラが等時性を示すというモーラ・タイミング仮説がある⁹⁾。さらに、ピッチアクセントに関わるF0は感情表出でも重要な役割を果たしており¹⁰⁾感情によって変化範囲が異なるという報告もある¹¹⁾。

そこで本研究では、(1)音声の時間構造、(2)語頭/pi/の母音無声化率、(3)F0のピークや範囲、が感情によってどのように変化するか、それは発達とどのように関係するかを調べた。

方 法

1. 発話被験者

就学前児は3歳5ヵ月〜6歳11ヵ月の16名(平均月齢68ヵ月SD11.11ヵ月)、就学児は7歳1ヵ月〜10歳10ヵ月30名(平均月齢105ヵ月SD13.73ヵ月)を被験者とした。どの被験者も週1回のテレビアニメを視聴していた。

2. 対象とした感情

「喜び」「悲しみ」「怒り」「驚き」と、特に感情のこもらない「平静」の5感情とした。

3. 録音手続き

最初に「平静」発話を録音した。識字可能な幼児・児童には「ぴかちゅう」という文字を読ませ、不可能な場合は特定の感情を示していないピカチュウの絵をみせて、「これは何?」と訊ね、/pikachuu/という発話を録音した。感情表現では、各感情を示すピカチュウの表情図を一枚づつ呈示し、ピカチュウの気持ちを訊ねた。ピカチュウの表情の背後にある感情語が表出可能であることを確認後、その感情を込めて/pikachuu/と発話するように指示した。表情図は感情ごとに2種類用意し、雑音の混入を考慮して2回づつ発話させた。マイクはピンマイクを使用し、被験者の口元から約10〜15cm離して上衣に取り付けた。録音はDATを用い、サンプリング周波数48KHzで行った。

4. 音声データ

課題途中の中断や/pikachuu/以外の発話をしたために、5感情に欠損値のあった5人(3歳1人、4歳1人、5歳1人、6歳2人)は今回の分析から除外した。その結果、就学前児の人数は11名となり、課題完了率は69%になった。発話は全体で709発話(就学前児179発話、就学児530発話)となった。音声データはサ

ンプリング周波数 12500 Hz, 16 bit で量子化し, コンピューターに格納した。

5. 聴取実験

話者の意図した感情が聴者に正確に伝わるかを調べるために, テレビアニメを視聴したことがない成人 10 名 (20~40 代の男女各 5 名) を被験者にして聴取実験を行った。手続きは, 音声 を 1 音声ずつヘッドホンにより聴取させ, 音声 が表わす感情を前述の 5 感情カテゴリーに「わからない」を加えた 6 選択肢の中から, いずれかを選択する強制選択法とした。話者感情と聴者が判断した感情が一致した場合, 聴者 1 人につき 1 点として得点化した一致得点に対して 2 要因の分散分析 (年齢群 2 水準×感情 5 水準) を行った。

6. 音響分析

本研究では成人が 70 % 以上の一致率で感情を同定できた音声 (就学前児 73 発話・就学児 238 発話) を対象として (表 1), それらの感情と年齢ごとの平均的な特性を検討した。70 % 以上に限定したため, 一部の感情, 特に「驚き」ではサンプル数が少なくなったものの, 感情や年齢に応じた平均的傾向を知るためには有用と考えた。これらのデータに関して, 時間構造, /pi/ の母音無声化率および F0 構造を以下の方法で検討した。

(1) 時間構造は, T1 を /pi/ から /ka/ のスパイクまでの区間, T2 を /ka/ から /chuu/ のスパイクまでの区間, T3 を /chuu/ のスパイクから母音の終了までの区間として, それらの長さを測定した。/pi/, /ka/, /chuu/ の各音節はすべて破裂部を持ち, スパイクはスペクトル上にほぼ明瞭に現れる。この測定方法であれば, 音声データ間および測定者間の測定誤差が最小限にとどまり, 頑健性がある。東京方言の特性として /pi/ で母音無声化が起これば T1 は T2 よりさらに短くなると予測される。一方, /chuu/ が 2 モーラとして発話されれば /ka/ より長いはずで, T2 < T3 と予測される。ただし, 先行研究^{12,13)}で指摘されているように語末長母音が短母音化されて発話されれば T2 = T3 となる可能性もある。T1, T2, T3 の長さを計測する理由はこれらの特性ないし制約が感情や発達に応

じて変化するかどうかを検討するためである。

各区間長が年齢と感情によってどのように変化するかを調べるために, 3 要因の分散分析 (年齢群 2 水準×感情 5 水準×区間位置 3 水準) を行った。また, 区間長の変化の程度を明確にするため, ① T1 を 1 とした時, T2, T3 の T1 に対する比率, ②「平静」の T1, T2, T3 を 1 とした時, 他の感情の T1, T2, T3 の「平静」に対する比率を求めた。

(2) /pi/ の母音が無声化したデータ数が全データ数に占める割合を /pi/ の母音無声化率とし, 母音無声化率が感情や年齢によって変化するかどうかを調べるために対数線型法¹⁴⁾による検定を行った。

(3) F0 構造を調べるために, /pi/, /ka/, /chuu/ の各有声区間の F0 の最高値を計測し, 最も高い値, すなわちピークがくる音節の割合を求めた。また, 各有声区間の F0 の最低値を計測し, 最低値から最高値までの範囲を F0 の変化範囲とした。F0 の最高値から最低値を引いた値が年齢や感情によって差があるかを調べるために, /pi/, /ka/, /chuu/ ごとに 2 要因の分散分析 (年齢群 2 水準×感情 5 水準) と, F0 の最高値や最低値が年齢や感情によって差があるかを調べるために, /pi/, /ka/, /chuu/ ごとに 2 要因の分散分析 (年齢群 2 水準×感情 5 水準) を行った。F0 の測定には STRAIGHT^{15,16)}を使用した。

結 果

1. 聴取実験

話者の意図した感情と聴者が判断した感情が一致した音声データの平均一致率は就学前児で 49 %, 就学児で 54 % となった。

感情別・年齢群別の一一致得点を図 1 に示した。エラーバーは標準誤差 (5 %) である。一致得点は両年齢群とも, 「平静」が最も高く, 「喜び」, 「驚き」の一一致得点が低いことは共通していた。就学前児では「平静」, 「怒り」, 「悲しみ」, 「喜び」, 「驚き」の順に, 就学児では「平静」, 「悲しみ」, 「怒り」, 「喜び」, 「驚き」の順に低くなった。

一致得点に対する 2 要因の分散分析 (年齢群 2 水準×感情 5 水準) の結果, 年齢間に有意差はなく, 上記の感情間差異 (F(4,699) = 27.290, $p < 0.0001$) と, 感情間差異の年齢による違い (F(4,699) = 3.411, $p < 0.01$) が有意であった。対比による検定を行ったところ, 「悲しみ」の感情において, 就学児の方が就学前児より有意に一致得点が高かった ($p < 0.01$)。有意差はなかったものの, 「怒り」においてのみ就学前児の方が就学児より一致得点が高かった。

表 1 音響分析を行う音声データの年齢・感情別内訳

	喜	驚	怒	悲	平静
就学前	11(6)	4(3)	23(9)	19(7)	16(11)
就学児	42(16)	17(11)	50(22)	84(25)	45(28)

単位: データ数 (発話者の人数)

2. 音響分析

T1, T2, T3 の実測値を年齢群別・感情別にしてグラフに示した(図2)。両年齢群とも T1, T2, T3 の順に区間長が延長した。しかし、どの感情でも一様に延長するわけではなく、延長の度合いは感情に依存し

図1 聴取実験結果。話者感情と聴者が判断した感情が一致した場合、聴者1人につき1点として得点化した。感情別・年齢群別に、一致得点の平均点を示した。

図2 各区間長の比較。T1を/pi/から/ka/のスパイクまでの区間、T2を/ka/から/chuu/のスパイクまでの区間、T3を/chuu/の開始から母音の終了までの区間として、感情別・年齢群別に示した。

ていた。感情による T1, T2, T3 の違いを示すために、区間長を比率にして検討した結果を表2に示す。T1=1として、感情ごとに T1 の T2, T3 に対する比を求めたものである。その結果、就学前児の T2 は T1 の 1.2~1.7 倍、就学児の T2 は 1.3~1.6 倍、就学前児の T3 は 1.9~3.1 倍、就学児の T3 は 1.8~2.8 倍となった。「平静」では T2, T3 がほぼ同じ比となった。一方、感情発話では T1, T2 がほぼ同じ比となり、T3 は T1, T2 のほぼ2倍に延長した。

次に、「平静」と感情発話の差を明らかにするために、「平静」の T1, T2, T3 をそれぞれ1として感情発話の「平静」に対する比をとり、表3に示した。T1, T2, T3 の延長は感情に依存した傾向がみられた。「悲しみ」は就学前児の T1 で「平静」の 1.9 倍、T2 で 1.5 倍、T3 で 2.5 倍、就学児の T1 で「平静」の 1.6 倍、T2 で 1.4 倍、T3 で 2.3 倍となり、どの区間でも、他の感情より長く延長した。一方、「驚き」は両年齢群とも T1, T2, は「平静」とほぼ同じ比率であり、T3 だけが就学前児で 1.9 倍、就学児で 1.7 倍と延長した。

T1, T2, T3 に対して年齢群×感情×区間位置の3要因の分散分析を行ったところ、区間位置 ($F(2,602)=196.989, p<0.0001$)、感情 ($F(4,301)=21.991, p<0.0001$) の主効果、および感情×区間位置 ($F(8,602)=10.341, p<0.0001$) の交互作用が有意だった。年齢の主効果や交互作用はみられなかった。

表4に示すように /pi/ の母音無声化率は「平静」では就学前児で 75 %、就学児で 84 %となった。しかし、感情を込めると、母音無声化率は減少した。両年齢群とも「悲しみ」では有声になることが多く、母音無声化率は就学前児で 21 %、就学児で 38 %だった。「喜び」では就学児で 57 %無声化しているのに対し、就学前児では 18 %しか起きなかった。「驚き」では就学前児

表2 各感情の T1 を 1 として、T2, T3 の T1 に対する比をとった

	T2					T3				
	喜	驚	怒	悲	平静	喜	驚	怒	悲	平静
就学前	1.5	1.2	1.2	1.3	1.7	2.3	3.1	2.1	2.5	1.9
就学児	1.3	1.4	1.4	1.4	1.6	2.4	2.8	2.3	2.5	1.8

表3 「平静」の T1, T2, T3 を 1 として、各感情の「平静」に対する比をとった

	T1				T2				T3			
	喜	驚	怒	悲	喜	驚	怒	悲	喜	驚	怒	悲
就学前	1.5	1.1	1.6	1.9	1.4	0.8	1.2	1.5	1.9	1.9	1.8	2.5
就学児	1.4	1.1	1.4	1.6	1.2	1.0	1.2	1.4	2.0	1.7	1.8	2.3

で 50 %, 就学児で 65 % と他の感情に比べて両年齢群とも母音無声化率が高かった。すべての感情において就学前児より就学児の方が母音無声化率は高かった。

対数線型法による検定を行った結果、年齢×無声化 ($\chi^2(1)=10.73$, $p<0.01$), 感情×無声化 ($\chi^2(4)=32.64$, $p<0.0001$) の交互作用が有意となり、就学児と就学前児では感情の母音無声化率への影響が異なることが示唆された。

表 5 に、各発話内で F0 のピークがどの音節にあるかを割合で示した。平静発話では就学前後ともに約 90 % の発話で /ka/ に F0 ピークがあり、/pi/, /chuu/ に F0 ピークがある割合は 10 % 以下だった。しかし、感情を込めると多くの発話で F0 ピークが /chuu/ に移行した。就学前児では、「怒り」で 69 %, 「驚き」で 67 %, 就学児では「怒り」で 61 %, 「驚き」で 60 % 発話で、/chuu/ に F0 ピークがあった。また、就学前児の「喜び」では 30 %, 「悲しみ」では 16 % の発話で /chuu/ に F0 ピークが移行したものの、「喜び」で 60 %, 「悲しみ」では 68 % の発話で /ka/ に F0 ピークがあり、/ka/ に F0 ピークのある割合の方が多かった。一方、就学児では「喜び」は /ka/ に 50 %, /chuu/ に 45 % の割合で、「悲しみ」は /pi/ に 24 %, /ka/ に 35 %, /chuu/ に 40 % の割合で F0 のピークがあり、平静発話のように /ka/ に F0 ピークが集中する傾向はみられなかった。このように感情を込めることによって /ka/ 以外に F0 ピークがある割合が高くなるものの、その割合は感情や年齢によって異なっていた。

図 3 に F0 の変化範囲、すなわち F0 の最高値の平均と最低値の平均を示した。就学児ではほぼ 400 Hz

を境界にして F0 が高く、変化範囲が広い「喜び」, 「驚き」と、F0 が低く、変化範囲が狭い「平静」, 「悲しみ」, 「怒り」に 2 群に分れた。就学前児の「悲しみ」は 400 Hz より高い範囲にあった。400 Hz を境界にして 2 分する傾向は /pi/ よりも /ka/, /chuu/ で顕著だった。

/pi/, /ka/, /chuu/ ごとに、F0 の最高値および最低値に対して、2 要因の分散分析(年齢群 2 水準×感情 5 水準)を行った結果、すべての音節で最高値、最低値とも感情によって有意な差があった。しかしながら、Fisher の下位検定の結果、/pi/ では最高・最低値ともに、最も F0 が低い「怒り」と「喜び」, 「驚き」, 「悲しみ」のそれぞれの感情間に、F0 が次いで低い「平静」と「喜び」, 「驚き」のそれぞれの感情間に有意差があるだけだった。一方、/ka/, /chuu/ の最高値では、F0 が高く、変化範囲が広い「喜び—驚き」間や、F0 が低く、変化範囲が狭い /ka/ の「平静—怒り」および /chuu/ の「悲しみ—怒り」を除くすべての感情間で有意差があった。最低値では、/ka/ の「平静—怒り」と /chuu/ の「喜び—悲しみ」を除いたすべての感情間に有意差があった。

「悲しみ」は就学前後で F0 の変化範囲が異なっており、就学前児の方が就学児よりも F0 が高かった。F0 の最高・最低値に対して、年齢群と感情の交互作用があり(最高値: $(F(4,283)=3,927, p<0.01)$ (最低値: $(F(4,283)=3,239, p<0.05)$), 対比による下位検定を行ったところ、/ka/ では有意差があった($p<0.001$)。

F0 の最高値から最低値をひいた差を変化幅として、変化幅の年齢や感情による変化をみた。/pi/ ではどの感情も 40~100 Hz の幅にほぼおさまった。しかしながら、/ka/ では変化幅が狭い「平静」, 「怒り」, 就学前児の「驚き」(50~70 Hz) と、変化幅が広い「喜び」, 就学児の「驚き」(160~220 Hz) に 2 分する傾向がみられた。/chuu/ ではさらに変化幅の差が広がり、変化幅の狭い感情「平静」, 「悲しみ」, 「怒り」(80~150 Hz) と、変化幅の広い感情「喜び」, 「驚き」(310~380 Hz) に 2 分した。

/pi/, /ka/, /chuu/ ごとに 2 要因の分散分析(年齢

表 4 /pi/の母音無声化率

	喜	驚	怒	悲	平静
就学前	18	50	35	21	75
就学児	57	65	58	38	84

単位: %

表 5 F0 ピークの位置。F0 の最高値が/pi/, /ka/, /chuu/のどこにくるかを調べ、割合で示した

	就学前					就学児				
	喜	驚	怒	悲	平静	喜	驚	怒	悲	平静
pi	10	0	6	16	7	5	0	5	24	7
ka	60	33	25	68	93	50	40	34	35	89
chuu	30	67	69	16	0	45	60	61	40	5

単位: %

図3 F0の変化範囲(上: /pi/ 中: /ka/ 下: /chuu/). グラフの両端は上がF0の最高値の平均, 下が最低値の平均である.

群2水準×感情5水準)を行った結果, /ka/ と /chuu/ の変化幅に対して, 感情による差があった (/ka/: (F(4,283)=14.383, $p<0.0001$) (/chuu/: (F(4,297)=31.428, $p<0.0001$). Fisherの下位検定を行ったところ, 変化幅の広い感情間(「喜び」「驚き」)および変化幅の狭い感情間(「悲しみ」「怒り」「平静」)では, /chuu/ の「怒り」「平静」間を除いて有意な差はなく, 変化幅の広い感情と狭い感情の間で有意差があった.

考 察

図2, 表2に示すように, どの感情においても, 年齢によらず, T1はT2より短いことが観測された. また, 平静発話ではT2とT3はほぼ同じ長さになり, 語末の長母音の短母音化を指摘した先行研究^{12,13)}を確認する結果となった. この関係は感情をのせると保存されず, 感情を含めた発話ではT3がT2の1.5~2倍程度に延長した. この結果は「短母音化した語末長母音が感情を含めると長母音として発話される」と解釈するよりも, 「成人の感情音声では語末音節が平静より約2倍程度延長される¹⁷⁾」という感情表現の典型性を就学前児がすでに獲得していると考えの方が自然であると解釈する. 感情が短母音が長母音かという音韻選択に関与すると考えるより, 後にも議論するように, 長音化した区間でF0変化幅の拡大やピークの強調など感情伝達に効果的な表現を行うと考える方が自然だからである.

表3に示すように, 感情を含めると, 平静発話より区間長が延長する傾向が観察された. しかし, どの感情においても一様に区間長が延長するわけではなく, 「悲しみ」はどの区間も長くなるとか, 「驚き」はT3のみが長くなるなど, 延長の仕方は感情に依存していた. この傾向は両年齢群でみられ, 時間長による感情表現の方略は就学前後で変わらないことが示唆された.

表4に示したように, /pi/ の母音無声化率はすべての感情において就学児の方が就学前児より上回っており, 特に「喜び」では差が大きかった. 東京方言話者の環境下では, 5歳までに母音無声化がほぼ定着するという報告があり⁶⁾, 就学前児で無声化率が低いことと合致する. /pika/ のように無声音, 有声音が交互に現れる単語では, 母音無声化率は発話速度が速いほど上昇することが知られている^{6,18)}. 本研究においても, 感情による母音無声化率の違いをみると, /pi/ の音節長が長い「悲しみ」「怒り」「喜び」では低く, 短い「驚き」「平静」では高い傾向が観測された. 感情による母音無声化率の差異は発話速度に依存した現象と考えることができる. これに対して, 図2に示したように就学前児と就学児のT1, T2, T3に有意差がみられないので, 加齢に伴う母音無声化率の上昇は発話速度の上昇によるものではなく, 東京方言の音声的制約を学習した結果と考えられる.

「平静」発話では, 両年齢群を通してF0のピークはほぼ /ka/ の音節にあり, /pikachuu/ のアクセント核が /ka/ にあることを反映する結果となった. しかし, 感情を含めるとピークは他の音節, 特に /chuu/ に移

行する傾向が強くなった。アクセント核はピッチの急激な下降によって表現される¹⁹⁾ことからF0ピークが平静発話で/ka/にあることは合理的である。感情をのせると音響的なF0ピークが語末の/chuu/に移動したという今回の結果は必ずしもアクセント核が語末音節に移動したことを意味しない。局所的なF0ピークとそれに続く下降がアクセント核を音響的にも保存している場合もあるし、また/ka/に局所的なピークがなくても聞き手が語末音節に出現したF0ピークを感情表現のための現象と知覚する可能性もある。予備的な検討では少数例を除いて少なくとも日本語を母語とする聞き手は/chuu/に出現したF0ピークをアクセント核と知覚することは少なく、むしろ感情表現の手段として余分に出現したピークと知覚する傾向が強かった。語末音節の延長とF0ピークの出現は米語発話でも観測されており²⁰⁾、アクセント核という言葉特異的な現象としてではなく、日米に共通な感情表現の方略として顕在化した傾向だと思われる。ただし、日本語を母語とする聞き手がアクセント核の移動と判断したサンプルは皆無ではなかったので、それが聞き手の認知特性に依存するのか、話者の感情表現上の方略や発達に依存するのか、さらに慎重に検討したいと考えている。

F0の変化範囲は、300~400 Hzを中心とした「怒り」、「平静」、および就学児の「悲しみ」と、400~800 Hzを中心とした「喜び」、「驚き」および就学前児の「悲しみ」に2分化された。感情を快・不快の2つのカテゴリーで分類すると、「喜び」は快感情に、「怒り」、「悲しみ」は不快感情に分類できる。「驚き」は瞬時の感情であり、表出後すぐに他の感情に移行する傾向があるため、快・不快への分類が難しい。Tomkins^{21,22)}のように、「驚き」を「快感情」に分類する立場が正しいとすれば、今回分析したF0の変化範囲は、「喜び」、「驚き」の快感情と、「怒り」、「悲しみ」の不快感情の2群に分化し、前者では高く広く、後者では逆になることになる。先行研究^{23,24)}の報告では「悲しみ」のF0変化範囲は低く狭いとするものが多い。本研究の「悲しみ」のF0は、就学児で297~413 Hzとほぼ410 Hz以下の115 Hz程度の範囲に、就学前児では398~583 Hzと400 Hz以上の185 Hzの範囲に分布した。就学児では成人を対象とした先行研究^{23,24)}と一致する傾向を示したのに対して、就学前児では異なる傾向を示した。

Fairbanks²⁵⁾は4ヵ月児の空腹時の「泣き」のF0を計測し、平常時よりF0が200 Hzも上昇したと報告している。また、幼児は「泣く」などの活動表出的語彙をほぼ同意味で使用しているという報告があ

る^{26,27)}。これらの研究を踏まえると、就学前児の「悲しみ」でF0が高いことは、就学前児が「悲しみ」の感情を「泣き」として表現した可能性が考えられる。このように考えると、就学児の「悲しみ」でF0が低いことは、「落胆」など「泣き」を伴わない「悲しみ」の表現を身に付けた結果とも解釈できる。

話者感情—聴者知覚感情の平均一致率が就学前児で49%、就学児で54%となり、成人話者を対象に聴取実験を実施したScherer²⁸⁾の51.1%、白澤ら²⁹⁾の51.5%とほぼ同様の聴取結果となった。これより、音声による感情の意図的な表現能力は就学前にある程度完成していることが示された。しかしながら、就学前児の課題完了率は69%であるため、年齢が幼いと意図的な感情表出能力には個人差が大きいたことが示唆された。

以上のように音声の特性は表現する感情に応じて平静発話とは異なる特性を示す。この特性の一部は発達的に変化する。感情表現に問題を抱える児童の音声を評価する場合、このような側面を考慮する必要があると考えられる。

結 論

就学前後の健常児の音声による感情表現を調べた結果、表現する感情に応じて時間構造、母音無声化、基本周波数(F0)の変化範囲やピーク位置が有意に変化した。「怒り」、「平静」では変化幅が狭く低いのにに対して、「喜び」、「驚き」では変化幅が広く、高いなど、感情による2極分化傾向も明らかになった。母音無声化、F0の変化範囲など一部の特性は年齢によっても変化した。感情表現に問題を抱える児童の音声を評価する場合、このような健常児の発達の側面も考慮する必要があると考えられる。

STRAIGHTの使用に関して、和歌山大学の河原英紀先生に感謝する。ピカチュウの使用に関して、任天堂、ゲームフリーク、ゲームクリエイチャーズ、小学館プロダクション(盛武源氏、福田剛士氏、伊藤忠久氏)に感謝する。

この研究の一部は科研費基盤B114800の支援を受けた。

文 献

- 1) Ricks, D.: Making sense of experience to make sensible sounds. Before Speech: The Beginning of Interpersonal Communication. (edited by Bullowa, M.), Cambridge University Press, New York, pp. 245-268, 1979.
- 2) Hobson, R.P.: The autistic child's appraisal of emotion. J. of Child Psychology and Psychiatry,

- 27 : 321-341, 1986 a.
- 3) Hobson, R.P. : The autistic child's appraisal of emotion : A further study. *J. of Child Psychology and Psychiatry*, 27 : 671-680, 1986 b.
 - 4) Ekman, P., Roper, G., Hager, J. C. : Deliberate facial movement. *Child Development*, 51 : 886-891, 1980.
 - 5) 志村洋子, 今泉 敏 : 生後2ヵ月の乳児の音声における非言語情報. *音声言語医学*, 36(3) : 365-371, 1995.
 - 6) Imaizumi, S., Fuwa, K., Hosoi, H. : Development of adaptive phonetic gestures in children : Evidence from vowel devoicing in two different dialects of Japanese. *J. of the Acoustical Society of America*, 106(2) : 1033-1044, 1999.
 - 7) 窪園晴夫, 白勢彩子 : 複合語アクセント規則の獲得. 理論とデータ. 特定領域研究(A)「心の発達 : 認知的成長の機構」ニューズレター8 : 19-25, 2000.
 - 8) 田尻 智 : ポケットモンスター. (石原恒和監修), 小学館, 東京, 1998.
 - 9) Bloch, B. : Studies in colloquial Japanese : Phonemics. *Language*, 26 : 86-125, 1942.
 - 10) Scherer, K.R. : Vocal affect expression : Review and a model for future research, *Psychological Bulletin*, 99(2) : 143-165, 1986.
 - 11) Murray, I.R., Arnott, J.L. : Toward the simulation of emotion in synthetic speech : A review of the literature on human vocal emotion. *J. of the Acoustic Society of America*, 93(2) : 1097-1108, 1993.
 - 12) 助川泰彦, 小河原義朗 : 長母音短縮現象と語中位置および発話速度との関係に関する考察, 第13回日本音声学会 : 131-136, 1999.
 - 13) Kubozono, H. : Visual perception of speech timing. 「心の発達 : 認知的成長の機構」平成12年度研究成果報告書 : 267-273, 2001.
 - 14) SAS/STAT User's Guide, SAS Institute, 6.17 : 1999.
 - 15) Kawahara, H., Katayose, H., de Cheveigné, A., et. al. : Fixed point analysis of frequency to instantaneous frequency mapping for accurate estimation of F0 and periodicity. *Proc. EUR-OSPEECH' 99*, 6 : 2781-2784, 1999.
 - 16) Kawahara, H., Katsuse-Masuda, I., de Cheveigné, A. : Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous frequency based F0 extraction : Possible role of a repetitive structure in sounds. *Speech Communication*, 27 : 187-207, 1999.
 - 17) 上床弘幸, 小林 豊, 新美康永 : 音声の感情表現の分析とモデル化. 電子情報通信学会技術研究報告, SP 92-131 : 65-72, 1993.
 - 18) Imaizumi, S., Hayashi, A., Deguchi, T. : Listener adaptive characteristics of vowel devoicing in Japanese dialogue. *J. of the Acoustical Society of America*, 98(2) : 768-778, 1995.
 - 19) 杉藤美代子 : 日本語音声の研究2—日本人の英語—. 和泉書店, 大阪, 1996.
 - 20) 櫻庭京子, 今泉 敏, 寛 一彦, 他 : 日米語の言語的制約が感情表現に及ぼす影響の音響的比較—ピカチュウ vs Pikachu—. 電子情報通信学会技術研究報告, SP 2000-166 : 47-54, 2001.
 - 21) Tomkins, S.S. : Affect, Imagery, Consciousness. Vol. I (Positive affect), Springer, New York, 1962.
 - 22) Tomkins, S.S. : Affect, Imagery, Consciousness. Vol. II (Negative affect), Springer, New York, 1963.
 - 23) Williams, C.E., Stevens, K.N. : Emotions and speech : Some acoustical correlates. *J. of Acoustical Society of America*, 52 : 1238-1250, 1972.
 - 24) Banse, R., Scherer, K.R. : Acoustic profiles in vocal emotion expression. *J. of Personality and Social Psychology*, 70(3) : 614-636, 1996.
 - 25) Fairbanks, G. : An acoustical study of the pitch of infant hunger wails. *Child Development*, 13 : 227-232, 1942.
 - 26) 星野喜久三 : 表情の感情的意味理解に関する発達的研究. *教育心理研究*, 26(2) : 129-132, 1969.
 - 27) 櫻庭京子, 今泉 敏 : 2~4歳児における情動語の理解力と表情認知能力の発達の比較. *発達心理研究*, 12(1) : 35-45, 2001.
 - 28) Scherer, K.R., Banse, R., Wallbott, H.G., et al. : Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, 15(2) : 123-148, 1991.
 - 29) 白澤敏行, 加藤義弘, 大西 昇 : 音声が入りに与える感性情報の分析. 電子情報通信学会技術研究報告, HIP 96-9 : 47-52, 1996.

別刷請求先 : 〒464-8601 名古屋市千種区不老町
名古屋大学人間情報学研究所
櫻庭京子