

アナウンサーが発声した「怒り」の韻律的特徴と 聴覚的印象の主成分分析を用いた比較

轟 真紀子^{*1}・武 田 昌 一^{*2}

PCA-based comparison between prosodic features and
auditory impressions of “anger” speech according to
the degree of emotion uttered by announcers

TSURU Makiko^{*1} and TAKEDA Shoichi^{*2}

^{*1}Kurume Shin-Ai Women's college

^{*2}Kinki University

This paper describes PCA-based analysis of the prosodic features of “anger” expressions in Japanese speech. Aural impressions of speech were compared with the physical parameters of the prosodic features. We used 240 speech samples uttered by four announcers for a listening test. The speech samples consisted of “neutral” speech as well as speech with three types of emotions (“anger,” “joy,” and “sadness”) of three degrees (“light,” “middle”, and “high”). Prosodic-feature parameters were speech rate and F0 parameters. Analysis results were summarized as follows: (1) The differences of features depending on the degrees of emotions which the speaker intended to express and those which the subjects identified as were found in the variations of eigenvalues of the second principal component for the maximum fundamental frequency and the magnitude of accent command, (2) the differences of features depending on the degrees of emotions which the subjects identified as were found in the variations of the same eigenvalues, conspicuously in those for the magnitude of accent command, and (3) the subjects tended to identify the emotions as either of two groups: “neutral” or “anger” regarding any degrees of anger as in one group.

Key words : PCA, Prosody, Emotional Speech, Listening Test

キーワード : 主成分分析, 韻律, 感情音声, 聴取実験

はじめに

感情音声の研究は、コンピュータを有効なコミュニケーションツールにするために重要である。それらのアプローチは様々な形がとられている。

筆者らのグループでは、「怒り」「喜び」「悲しみ」「感謝」感情の違いと、それぞれの度合

(軽い感情・普通の感情・激しい感情)までも対象として、多数の音声数の韻律的特徴の解析を行っている¹⁾⁻³⁾。また、感情を伝える側、つまり発話者の発声した音声の特徴だけに重点を置いてしまうと、よりよいコミュニケーションツールになることはできない。感情を伝えられた、つまり聞き手の受容の特徴にも焦点をあてる必要がある。筆者らは発話者の感情音声の解

析にとどまらず、発話者の意図が聞き手にどの程度伝達されているのか、感情の度合の差異までも考慮に入れた聴取実験を続けている^{4),5)}。聴取実験には合成音声を使った報告もされているが、著者らは自然音声を使い、発話者と聞き手の感情の差異を韻律的特徴から分析するアプローチをとっている。本報告は、主成分分析を用いて発話者の意図する感情音声と被験者が受容した音声について、感情を表現している韻律の特徴を導き出すことを目的とする。

実験方法

音声試料および韻律パラメータ

- ・話者：元NHKアナウンサー4名（男性2名、女性2名、全員60歳代）。
- ・単語：4モーラ。アクセントは平板型の「やまごや」、中高型の「なみなみ」、頭高型の「いまにも」と各アクセント共通に「まなまな」。感情のない「平常」および「軽い怒り」「怒り」「激怒」、「軽い喜び」「喜び」「大喜び」、「軽い悲しみ」「悲しみ」「深い悲しみ」の感情を込めた音声を各5回発声。聴取実験には、そのうち1人のアナウンサーにつき、すべての感情音声のそれぞれ1発声の計60音声を抽出し、他の3人のアナウンサーも同様に抽出し、合計240音声を使用した。本実験では、韻律的特徴のパラメータとしては「平均発話速度」、「最高基本周波数」、基本周波数パターン生成過程モデル（藤崎モデル）にもとづいて「フレーズ指令の大きさ」、「アクセント指令の大きさ」、「最低基本周波数」などを用いる。

聴取実験方法

- ・240の音声をランダムに並べ、1音声の後に3秒間の「間」をとった。さらに、心理学的測定法の手法として、被験者が回答の練習をすることができるように各セッションの最初に10音声のダミー音声を加えた。ダミー音声には、3種類のアクセント型と4種類の単語（まなまな・やまごや・なみなみ・いまにも）

および「平常」を含めた全種類の感情音声を4人の声のいずれかで発声したものを取り入れた。

- ・被験者は1つの音声を聞いて、直感的に感じた感情とその度合（段階）を記入用紙（Figure 1）に記入してもらった。
- ・被験者は、静かな環境で一人1台のパソコンに向かい、同じメーカーのヘッドホンで被験者が聞きやすい十分に大きい一定音量に設定して実験を行った。
- ・被験者数は19歳～20歳の女子学生29人である。聴覚は正常である。
- ・同一被験者に250音声（ダミー10音声を含む）の聴取実験を2回行った。
- ・1回目と2回目の音声を逆順とした。

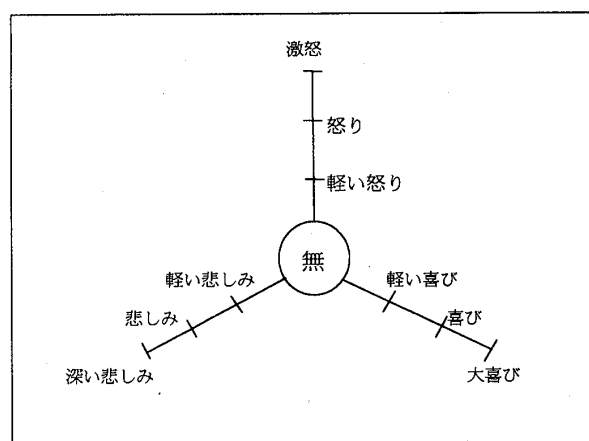


Figure 1 回答の記入様式

実験結果

聴取実験で得られたデータのうち「怒り」感情を3つの視点から主成分分析を行った。1つ目は発話者の意図による「怒り」音声、2つ目は発話者の意図と被験者の感情が一致した「怒り」音声、そして3つ目は被験者が受容した「怒り」音声である。比較のために「平常」の主成分分析を行った。使用する主成分は固有値で1以上を示す第2主成分までとする。

発話者が発声した「怒り」音声の主成分分析

アナウンサーが「怒り（軽い怒り・怒り・激

怒)」の感情で発声した72音声に対して度合別に主成分分析を行った。その結果、「軽い怒り」の第1主成分には「最低基本周波数」と「最高基本周波数」がそれぞれに0.58, 0.54の値を示し、声の高低の程度を表していると考えられる。第2主成分には「フレーズ指令の大きさ」が0.89の値を示し、発声時の声の荒々しさの程度

を表していると考えられる (Table 1)。この傾向は「怒り」にも見られた。次に、「激怒」の第1主成分には「軽い怒り」および「怒り」とは異なり、「アクセント指令の大きさ」と「最低基本周波数」がそれぞれ0.60と0.63の値を示している (Table 2)。

Table 1 アナウンサーが発声した「軽い怒り」音声の相関行列の固有値および固有ベクトル

主成分 (固有ベクトル)	第1主成分	第2主成分	第3主成分
平均発話速度 [sk_eve]	-0.53	0.18	0.34
フレーズ指令の大きさ [phrase]	0.06	0.90	-0.24
アクセント指令の大きさ [accent]	0.31	0.27	0.86
最低基本周波数 [F0_min]	0.59	-0.24	0.13
最高基本周波数 [F0_max]	0.54	0.17	-0.27
固有値	2.79	1.15	0.88
寄与率	55.7%	23.1%	17.6%
累積寄与率	55.7%	78.8%	96.4%

Table 2 アナウンサーが発声した「激怒」音声の相関行列の固有値および固有ベクトル

主成分 (固有ベクトル)	第1主成分	第2主成分	第3主成分
平均発話速度 [sk_eve]	-0.32	-0.20	0.84
フレーズ指令の大きさ [phrase]	-0.18	0.69	0.22
アクセント指令の大きさ [accent]	0.60	0.18	0.49
最低基本周波数 [F0_min]	0.63	-0.36	0.06
最高基本周波数 [F0_max]	0.32	0.56	-0.09
固有値	1.81	1.31	0.97
寄与率	36.3%	26.2%	19.3%
累積寄与率	36.3%	62.5%	81.8%

発話者と被験者の感情が一致した音声の主成分分析

一致した音声とは、被験者が発話者の感情とその度合まで一致して回答した音声を指す。被験者のうち一人でも発話者との感情が一致した音声を取り出し、その韻律的特徴を主成分分析にかけた。結果は第1主成分および第2主成分ともに前項と同じ傾向であった。アナウンサーの「怒り」発声について、第1主成分から、発話者が怒り感情を表現するのに「最高および最低基本周波数」つまり「声の高低」に表れていると考えられる。第2主成分から「フレーズ指令の大きさ」つまり発声時における「声の荒々しさ」に表れていると考えられる。また、「激怒」においては第1主成分に別の韻律的特徴が示されており、感情の度合間に主成分となる韻律的特徴が異なっていることが分かった。これは発話者が「怒りの度合」に変化を与える際、それを特徴づける韻律を変化させて発声していることが考えられる。

被験者が「怒り」感情に同定した音声の主成分分析

同定した音声とは、発話者の意図とは無関係に被験者が同定した音声を指す。被験者が「怒り」（軽い怒り・怒り・激怒）と同定したすべての音声を主成分分析の対象とした。よって分析の対象には、発話者が「怒り」以外の感情で発声した音声も含まれ、また発話者が「怒り」感情で発声した音声でも被験者が「怒り」に同定しなければ含まれないことになる。その結果、「軽い怒り」、「怒り」、「激怒」のいずれも主成分の傾向は同じであった。「激怒」を例に見ると、第1主成分には「最高基本周波数」と「最低基本周波数」にそれぞれ0.56, 0.55の値を示し、「声の高低」の程度を表していると考えられる。これは発話者の意図で発声された音声と同様の傾向である。第2主成分には「フレーズ指令の大きさ」と「アクセント指令の大きさ」がそれぞれ0.64, 0.61の値を示し、発声時における「息の荒々しさ」と「アクセントの強弱」を表していると考えられる（Table 3）。被験者が同定した「怒り」感情音声において、その度合の違いによる主成分の変化はなかった。

Table 3. 被験者が「激怒」感情に同定した音声の相関行列の固有値および固有ベクトル

主成分（固有ベクトル）	第1主成分	第2主成分	第3主成分
平均発話速度 [sk_eve]	-0.52	0.18	0.01
フレーズ指令の大きさ [phrase]	-0.22	0.64	0.63
アクセント指令の大きさ [accent]	0.24	0.61	-0.67
最低基本周波数 [F0_min]	0.56	-0.21	0.33
最高基本周波数 [F0_max]	0.55	0.37	0.22
固有値	2.29	1.34	0.76
寄与率	45.9%	26.7%	15.2%
累積寄与率	45.9%	72.6%	87.9%

固有ベクトルの散布図による比較

次に、各度合の第1主成分と第2主成分の固有ベクトルの散布図を、発話者と被験者の感情および度合が一致した音声(Figure 2)と、被験者が感情および度合に同定した音声(Figure 3)のグラフを示す。

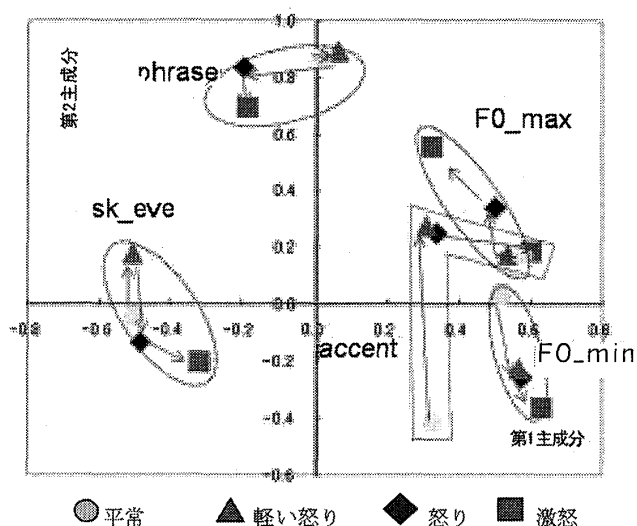


Figure 2 アナウンサーと被験者の感情が一致した音声の固有ベクトル散布図

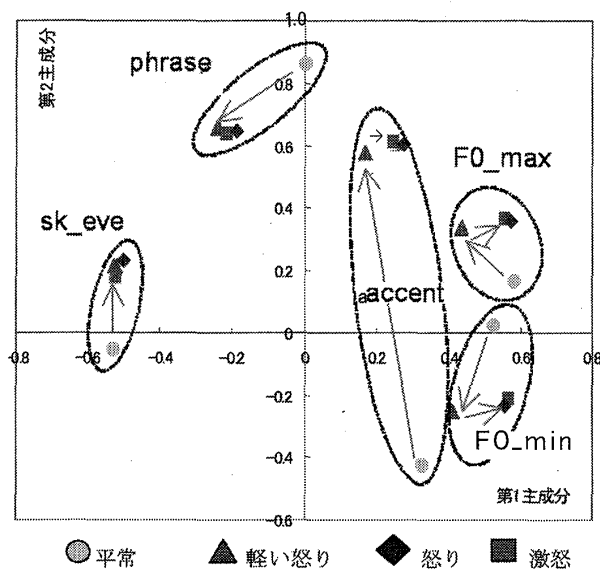


Figure 3 被験者が同定した音声の固有ベクトル散布図

その結果、Figure 2にあるように被験者が「平常」と異なる特徴付けをしている韻律は「アクセント指令の大きさ」と「最低基本周波

数」であることがわかる。それ以外の韻律では「平常」と「軽い怒り」あるいは「平常」と「怒り」がほぼ同じ特徴付けになっている。一方、被験者が同定した「怒り」音声では、Figure 3に示しているように、被験者は「平常」と「怒り」グループの間にある韻律的特徴を明確にその違いを捉えていることがわかる。このことは、発話者の韻律的な表現と被験者が捉えているそれとは隔たりがあるということが考えられる。

結 論

本研究では、発話者が発声する感情音声と聞き手の受けた感情の相違、またそれに起因する韻律的特徴を導き出すために韻律的特徴の主成分分析を試みた。その結果、アナウンサーの「怒り」感情において、それを表現する韻律パラメータは「最高基本周波数」と「最低基本周波数」であることが示唆された。よって、声の高低で「怒り」の表現をしているという結果が得られた。また発話者が「激怒」感情を表現するときに「軽い怒り」および「怒り」とは異なる韻律的特徴を与えていると考えられる結果を得た。さらに「怒り」を表現する発話者の韻律パラメータと、「怒り」を受容する被験者のそれとは隔たりがあることも示唆された。

今後は、「喜び」あるいは「悲しみ」音声についても分析を進める。

謝 辞

擬似感情音声を発声して下さったアナウンサーの方々に謝意を表す。また本研究に携わりたくさんの貴重な解析データを残してくださった武田研究グループの卒業生および聴取実験に快く協力してくださった皆様に対して深く感謝の意を表す。

なお、本報告は日本音響学会2007年春季研究発表会において報告したことを付記する。

参 考 文 献

- 1) 武田昌一, 村岡輝雄, 大山 玄, “感情・感性的表現をもつ音声の韻律的特徴の解析”, 平成12年度～16年度文部省科学研究費補助金特定領域研究 研究成果報告書, 2005
- 2) 橋澤保輝, 武田昌一, MUHAMMED DZ ULKHIFLEE HAMZAH, 大山 玄, “アナウンサーと声優が発声した感情を表現する音声の感情の程度に応じた韻律的特徴の話者グループ間の比較”, 音講論集, 2-1-2, 2005
- 3) 武田昌一, 大山 玄, 栃谷綾香, 西澤良博, “日本語音声における「怒り」を表現する韻律的特徴の解析”, 日本音響学会誌58巻9号, 561-568, 2002
- 4) 靄真紀子, 武田昌一, “感情表現音声の韻律的特徴 -聴取実験による検討-”, 久留米信愛女学院短期大学研究紀要第27号, 67-73, 2004
- 5) 靄真紀子, 武田昌一, “アナウンサーが発声した感情を表現する音声の感情の程度に応じた韻律的特長と聴覚的印象との比較”, 音講論集, pp.295-296, 2005