

短い文節から成る演技音声の「怒り」と「喜び」の程度を表す 発話の開始・終了の局所的特徴の研究

Research on Acoustic Features Appeared in both the Beginning and the Ending Parts of Spontaneous “Anger” and “Joy” Short Utterances Performers Speak

窪田 恭紘†
Takahiro Kubota

飯田 仁‡
Hitoshi Iida

1. はじめに

近年様々な分野で合成音声が使われているが、感情を含んだ音声を合成する技術はまだ不十分である。より伝わりやすい感情がこもっている音声を合成する為には、感情が強くなる毎にどのような特徴が見られるのかを解明する必要がある。そこで本論では、河津らの研究[1]で発話の文頭・文末(4文節の発話における第1、第4文節)に音響的特徴が見られたという結果を受け、文頭・文末をより細かく分析していくことで、感情の程度が強くなる毎に変化をする音響的特徴を探っていく。今回、各感情のデータが十分に得られなかった為、比較的データ数が得られた「怒り」及び「喜び」について分析を行っている。

2. 音声資料

有本ら[2]はオンラインゲームにおけるプレイヤーの対話を収録し、そこに現れる感情音声の音響的特徴の分析を行った。そこでは、発話の長さに関係なく分析を行っているが、局所的特徴を捉える必要があるため、本論でその研究で使用された音声から1文節の発話 B_1 、2文節の発話 B_2 の計36発話を扱う。 B_1 は文頭・文末を兼ねた音声として扱う。

表1. 発話内容の例

| 感情 | 様々な発話例 |
|----|---------------------------------|
| 怒り | 水/いて/むかつく/解除、解除ってさー |
| 喜び | ウハウハだよ/おらー/経験値をくれよ/ あいよー、倒した |

表2. 文節毎の発話数

| 感情の種類 | 文節数 | 発話数 |
|-------|-------|-------|
| 怒り | B_1 | 12 発話 |
| | B_2 | 4 発話 |
| 喜び | B_1 | 12 発話 |
| | B_2 | 8 発話 |

文節に関しては、より局所的な特徴を捉える為に、短い音声を分析する、という観点から動詞・名詞等に助詞・助動詞が2つまで繋がっている、もしくは繋がっていないものと定義した。品詞は形態素解析システム「茶筌」[3]に準拠した。

† 東京工科大学大学院バイオ・情報メディア研究科

‡ 東京工科大学

3. 音声分析

河津らの研究で怒りと喜び両方に見られた特徴を基に、音声分析ソフトウェア「praat」[4]を用いて $F0$ 最大値 $F0_{max}$ 、 $F0$ 最小値 $F0_{min}$ 、数値解析ソフトウェア「MATLAB」[5]を用いてダイナミックレンジ $F0_{dyna}$ 、短時間パワー最大値 P_{max} 、短時間パワー最小値 P_{min} 、短時間パワー変動量 P_{mag} を算出した。また局所的なパラメータとして、大野らの研究[6]で提案されている局所的発話速度の考えを基に、局所的発話速度の上昇が最大となるときを R_{max} 、最小となるときを R_{min} とし、その値の算出を独自に行った。尚、局所的発話速度は、平静を伴った発話と同一発話内容で感情のこめられた発話の2つについて DP マッチングを用いて比較し、そこから速度比を求めている為、平静を伴った発話についてパラメータ抽出を行えない。

上記で算出した値が、感情の程度が強くなる毎にどのように変化しているか観察し、パラメータ毎に分散分析を行った。また、局所的な特徴を捉える為、平均値の算出は行わなかった。

4. 分析結果

怒りの B_1 において $F0_{max}$ は、6発話で程度が強くなる毎に最も大きくなり、 P_{max} は9発話において程度が強くなる毎に上昇した。また R_{max} は8発話で上昇傾向が見られた。次に怒りの B_2 だが、特に上昇傾向等は見られなかった。続いて喜びの B_1 では8発話で $F0_{max}$ 、5発話で P_{max} 、10発話で R_{max} が上昇した。最後に喜びの B_2 だが、文頭・文末毎に上昇傾向のあったものを以下の図に示す。

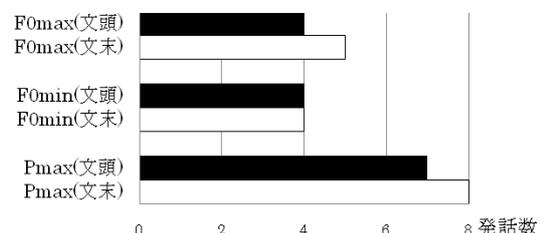


図1. 喜びの B_2 において傾向が見られた発話の数

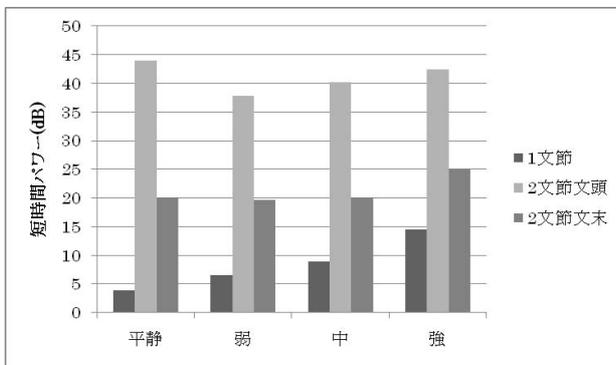
$F0_{max}$ と $F0_{min}$ 及び P_{max} に上昇傾向があり、特に文末で多く見られた。そして R_{max} については4発話で文頭に上昇傾向が見られた。また感情や文節数に関わらず R_{min} の値はばらつきがあり、上昇傾向等は見られなかった。

続いて分散分析の結果だが、有意水準を0.05として分析を行った。結果は以下の表の通りである。

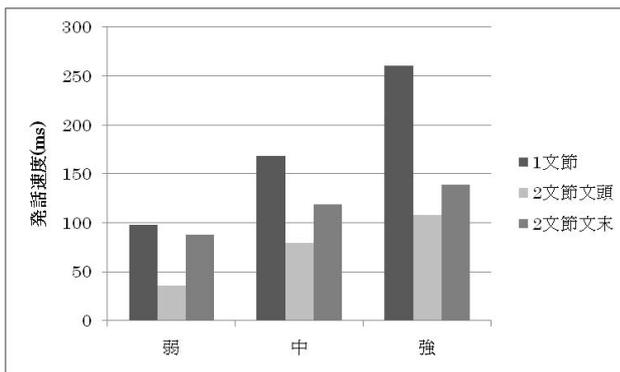
表3. 音響パラメータ毎のP値

| | 怒り | 喜び |
|-------------|-------------|----------|
| FO_{max} | 0.086162735 | 0.730189 |
| FO_{min} | 0.020793339 | 0.012216 |
| FO_{dyna} | 0.121165496 | 0.050397 |
| P_{max} | 0.015094684 | 0.100151 |
| P_{min} | 0.011335152 | 0.00372 |
| P_{mag} | 6.86122E-06 | 0.049196 |
| R_{max} | 0.023924175 | 0.036781 |
| R_{min} | 0.006128429 | 0.116024 |

怒り・喜び両方では FO_{min} , P_{min} , P_{mag} , R_{max} において、怒りだけでは FO_{min} , R_{min} において有意差が認められた。更に有意差が認められたパラメータについて、文節数・感情の程度毎の平均を算出した。

図2 怒りの P_{mag} における傾向

怒りの P_{mag} においては2文節文頭で平均値が高くなっている。ただし平均値に関して上昇傾向は見られておらず、1文節においては上昇傾向が見られている。

図3 喜びの R_{max} における傾向

喜びの R_{max} においては1文節で平均値が高くなっている。また全文節において平均値は上昇傾向にある。

5. 考察

5.1 怒り

怒りについては2文節で上昇傾向等は見られなかったが、1文節において程度が上がる毎に P_{max} , R_{max} の上昇が見られ、 FO_{max} は程度が弱で最大になるか上昇するという傾

向が見られた為、 P_{max} , R_{max} が怒りの程度に関わる最も重要な要素であり、次いで FO_{max} が重要だと思われる。また、継続時間長を分析した河津らの研究では、程度が上がる毎に発話は遅くなると述べられているが、本論の分析から局所的には発話速度が上昇するという結果が得られた。このことから発話速度が平均的に遅くなっても、局所的に発話を速くすることで、感情を強くしているのではないかと考えられる。

5.2 喜び

喜びについてもほぼ怒りと同様 FO_{max} , P_{max} , R_{max} に傾向が見られ、更には FO_{min} にも傾向が見られたことから声の高低で感情を表していると思われる。文頭で R_{max} に傾向が見られたが、喜びについても河津らが程度が上がる毎に継続時間長が遅くなったと述べている為、発話速度が平均的に遅くなっても、局所的に発話を速くして感情を強くしていると思われる。文末で FO_{max} と FO_{min} 及び P_{max} に上昇傾向が見られるが、これは先行研究で述べられている結果と大差がない為、それらの上昇傾向が喜びの特徴であると言える。また R_{min} にも特徴が見られるのではないかと考えたが、そのような結果が得られなかった。

6. むすび

本論で分析した怒りと喜びにおける B_1 , B_2 の文頭・文末に見られる韻律的特徴の中で、特に局所的発話速度に上昇傾向等があるのではないかと考えたが、 R_{max} については感情の程度が上がる毎に速くなっていくという結果が得られ、分散分析でも有意差が見られたことから、感情の程度を左右する傾向が局所的に表れていると考えられる。ただし発話中のどの部分で傾向が見られたかは不明である。そこで今後は発話毎に局所的発話速度変化の概形を求めることで、各発話に見られる共通の傾向を探っていく。また、今回1・2文節しか分析していないことやデータ数自体が少ないことなども問題として挙げられる。その為、今後は新たにデータを収録して分析を進めていく。そして各音響パラメータから得られた特徴を組み合わせ、感情を伴った音声を合成し、聴取実験によって妥当性を検証する予定である。

参考文献

- [1] 河津 宏美, 大野 澄雄, “感情の程度変化に対する各種韻律的特徴の役割”, 日本音響学会 2007 年秋季研究発表会講演論文集, pp.307-308, (2007).
- [2] 有本 泰子, 河津 宏美, 大野 澄雄, 飯田 仁, “感情音声のコーパス構築と音響的特徴の分析- MMORPG における音声チャットを利用した対話中に表れた感情の識別 -”, 情報処理学会第70回音声言語情報処理研究会 (SIG-SLP) 研究報告, pp.133-138, (2008).
- [3] 茶筌 <http://chasen.naist.jp/hiki/ChaSen/>
- [4] praat <http://www.fon.hum.uva.nl/praat/>
- [5] MATLAB <http://www.mathworks.co.jp/>
- [6] 大野 澄雄, 福宮 誠道, 藤崎 博也, “発話速度の局所的変動の定量的分析手法”, 1996 年電子情報通信学会総合大会, pp.242, (1996).