社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2007－ＳＬＰ－69 (16)
2007／12／20

# 書き起こしへの付与を目指した発話印象の表現法に関する分析

小川 純平† 西田 昌史‡ 堀内 靖雄‡ 黒岩 眞吾‡

†‡千葉大学 大学院自然科学研究科
〒263-8522 千葉市稲毛区弥生町 1-33
E-mail: †j_ogawa@graduate.chiba-u.jp, ‡{nishida,hory,kuroiwa}@ faculty.chiba-u.jp

あらまし これまで我々は，討論や会議における書き起こしに発話印象を付与することを目指して，韻律情報をもとに発話印象を推定する手法について検討を行ってきた．本研究では，新たに韻律情報として F0 モデルから抽出したアクセント成分とフレーズ成分を用いて分析を行った．また，音声から推定された発話印象をどのように書き起こしに付与するかについて分析を行った．今回は，討論や会議といった音声の書き起こしを対象としているため，文字の太さ，大きさといった文字の装飾や感嘆符，疑問符などの記号の付与に着目した．対話音声の書き起こしにこれらのテキスト表現を行い，発話印象の主観評価実験を行った結果，音声から感じる発話印象とテキストから感じる発話印象の違いが明らかとなった．
キーワード 書き起こし，発話印象，韻律，F0 モデル，対話音声

# A Study on Indexing Method of Utterance Impression to Transcription

Junpei OGAWA† Masafumi NISHIDA‡ Yasuo HORIUCHI‡ and Shingo KUROIWA‡

†‡Chiba University Graduate School of Science and Technology
1-33 Yayoicho, Inage-ku, Chiba, 263-8522 Japan
E-mail: †j_ogawa@graduate.chiba-u.jp, ‡{nishida,hory,kuroiwa}@ faculty.chiba-u.jp

**Abstract** We have studied on estimation of utterance impression using prosody in order to index the utterance impression to transcription of debates and meetings. In this study, it estimated the utterance impression using accent and phrase elements extracted by F0 model. Moreover, it analyzed how to index the utterance impression to the transcription. We focused on thickness and size of character and sign of question and exclamation marks. We conducted subjective evaluation of the utterance impression using speech and text in dialogue speech. As a result, it demonstrated that the utterance impressions by speech and text are different.
**Keyword** Transcription, Utterance impression, Prosody, F0 model, Dialogue speech

## 1. はじめに

近年，音声認識技術を用いた音声の書き起こしの自動作成についての研究が盛んに行われてきている．これまでのシステムは音声を正確に書き起こすことに重点をおいているが，発話の内容をより正確に伝えるためには言語情報以外に議論の場面や話者の状態，感情といった情報も重要である．

我々はこれまで書き起こしに付与することを目的として，上記のような話者の感性情報（以下，発話印象）のうち言語情報から推定しづらいものを，対話音声をもとに韻律情報から推定するという研究を行ってきた[1]，[2]，[3]．しかし実際に発話印象を書き起こしに付与する場合，発話印象をどのようにテキストで表現するのかも大変重要である．

書き起こしと感性情報との関係に関する研究は，文字の大きさや色を変化させたときに興奮度合いの感じ方がどのように変化するのかを調べたもの[4]や，音声から受ける印象とフェイスマーク（顔文字）との関係を調査したもの[5]，テレビの字幕の色や大きさ，表示位置を変化させて感情の伝わりやすさと字幕の見やすさについて調べたもの[6]などがある．しかし感性情報と書き起こしとの関係は未だ明確にされていないのが現状である．

そこで我々は，討論や会議などの書き起こしに発話印象を付与することを目指し，テキストを変化させることによる発話印象の感じ方への影響について分析を行った．テキストの変化は，テキストが見づらくならないものを対象とし，文字を太くしたり，！，？などの記号を付与するものを扱う．また同時に今まで行ってきている音声による発話印象の推定の精度向上のた

め，使用する韻律パラメータについても検討した．

## 2. テキストの変化と発話印象の関係の分析

文字を太くするといった文字の装飾や，！などの記号を付ける，といったように，テキストに変化をつけることにより書き起こしを読んだ人に言語情報だけでは伝えることができない発話印象を伝えることができると考えられる．そこで，さまざまなテキストの変化によりどのような発話印象が感じられるのかを調査するために，テキスト変化の主観評価実験を行った．

### 2.1. 実験方法

変化させていないテキストと変化させたテキストの両方を提示し，「強調，疑問，驚き，自信，迷い，怒り，喜び，嫌悪，悲しみ，怖れ」の10個の発話印象について，0（感じられない），1（やや感じられる），2（感じられる），3（とても感じられる）の4段階で評定した．被験者は10名で，テキスト変化の種類は8種類である．具体的な実験の例を図1に，テキスト変化の種類の詳細を表1に示す．

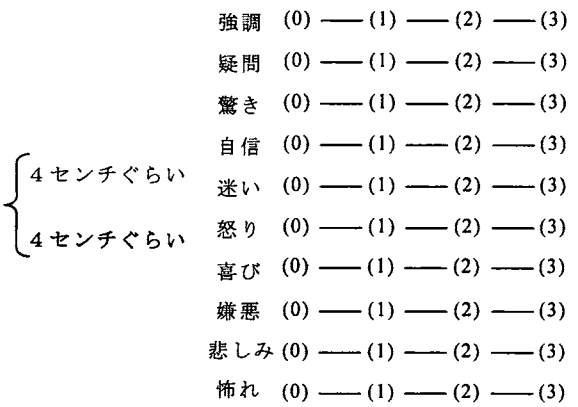


図1　テキスト変化の主観評価実験例

表1　テキストの変化の種類

| | 変化の種類 | 変化させたテキスト |
|---|---|---|
| 1 | 太字 | **4センチぐらい** |
| 2 | 文字縮小 | 4センチぐらい |
| 3 | 文字拡大 | 4センチぐらい |
| 4 | ？付与 | 4センチぐらい？ |
| 5 | ！付与 | 4センチぐらい！ |
| 6 | 下線 | 4センチぐらい |
| 7 | …付与 | 4センチぐらい… |
| 8 | ！？付与 | 4センチぐらい！？ |

図1のように，上段に通常のテキスト，下段に変化させたテキストを印刷した紙面上で提示し，通常のテキストに対して変化させたテキストから各発話印象がどの程度感じられるかを評定させた．なお今回は純粋にテキストの変化により受ける発話印象への影響のみを調べるため，以前実験により言語情報からは何も印象が感じられないとされた，「4センチぐらい」，「左上」，「右上にそのまま真っ直ぐ上に」，「Sの字を描きながら」の4つのフレーズを用いた．この4つのフレーズは3章で述べる音声の実験で用いたデータの1部である．したがってデータ数は10（発話印象の数）×8（テキスト変化の種類）×4（フレーズの種類）の340個となる．なお，表1に示したテキスト変化以外にも，組み合わせなどにより様々なテキスト変化のパターンが考えられるが，見やすさなども考慮して今回はなるべく単純に表現することのできる表1の8つを使用した．

### 2.2. 実験結果

本研究では，被験者10名の評定結果の平均値をその発話印象の評定値とした．また，評定値が1.0以上となったものはその発話印象が現われているもの（印象あり）とし，評定値が1.0より小さいものはその発話印象が感じられないもの（印象なし）として扱う．次章で音声の主観評価実験についても述べるが，それについても同様に評定値を扱う．

ここで，テキストにおける発話印象間の相関係数を表2に示す．なお，全10個の発話印象のうち，「怒り，喜び，嫌悪，悲しみ，怖れ」の5つについてはすべてのテキスト変化において印象なしとなった．これらの発話印象は今回用いたテキストの変化では表現できないと考えられる．そこで，今回はそれらを分析の対象外とし，表1のテキスト変化で表現できると考えられる「強調，疑問，驚き，自信，迷い」の5つに注目して分析を行う．

表2　発話印象間の相関係数（テキスト）

| | 強調 | 疑問 | 驚き | 自信 | 迷い |
|---|---|---|---|---|---|
| 強調 | 1.00 | -0.17 | 0.47 | 0.76 | -0.71 |
| 疑問 | -0.17 | 1.00 | 0.52 | -0.60 | 0.52 |
| 驚き | 0.47 | 0.52 | 1.00 | 0.00 | -0.08 |
| 自信 | 0.76 | -0.60 | 0.00 | 1.00 | -0.76 |
| 迷い | -0.71 | 0.52 | -0.08 | -0.76 | 1.00 |

表2により，自信と迷いのような反対の意味を持つと思われる発話印象間に強い負の相関がみられ，強調と自信のように似ていると思われる発話印象間には強い正の相関がみられることから，妥当な評定結果が得られていると考えられる．

次に，各テキスト変化におけるそれぞれの発話印象の評定値についてまとめた結果を表 3 に示す．表 3 の 1 列目の数字は表 1 の数字と対応している．

表 3　各テキスト変化における発話印象の評定値

| | 強調 | 疑問 | 驚き | 自信 | 迷い |
|---|---|---|---|---|---|
| 1 | 1.6 | 0.0 | 0.2 | 1.0 | 0.0 |
| 2 | 0.0 | 0.2 | 0.0 | 0.0 | 1.7 |
| 3 | 2.7 | 0.0 | 0.4 | 1.7 | 0.0 |
| 4 | 0.0 | 2.4 | 0.1 | 0.1 | 1.2 |
| 5 | 1.7 | 0.0 | 0.7 | 1.6 | 0.0 |
| 6 | 1.9 | 0.0 | 0.1 | 0.7 | 0.0 |
| 7 | 0.0 | 0.5 | 0.0 | 0.0 | 1.8 |
| 8 | 1.2 | 1.9 | 2.1 | 0.2 | 0.9 |

### 2.3. 考察

表 3 を見ると強調と自信では 3 の「文字拡大」，疑問では 4 の「？付与」，驚きでは 8 の「！？付与」，迷いでは 7 の「…付与」の評定値が最も高くなっている．ここで自信と驚きに注目する．自信と驚きは今回扱ったテキスト変化だけでは単独で表現できないことが分かる．つまり自信や驚きが感じられる（評定値 1.0 以上）テキスト変化では必ず別の発話印象も感じられてしまうということである．具体的には自信が感じられる 1, 3, 5 のテキスト変化には強調も同時に感じられ，驚きが感じられる 8 のテキスト変化では強調や疑問が同時に感じられるという結果が得られている．これは，表 2 において強調と自信の相関係数が 0.76，疑問と驚きの相関係数が 0.52 と発話印象間に相関関係が見られることとも整合がとれている．特に強調と自信については強い相関関係があるので，これらの発話印象をテキスト表記で区別することは難しいと考えられる．

また，表 3 において同じ発話印象であっても表現の違いによりその感じ方の度合いに変化が生じていることがわかる．これはテキスト変化の種類を使い分けることで発話印象の度合いも表現できることを示している．

## 3. 音声による発話印象の推定

我々は，テキストと発話印象の関係分析と平行して，音声から発話印象を推定する研究も行ってきている．これまでは F0（基本周波数）の平均値，最大値，最小値，レンジ，パワーの平均値，最大値，最小値，レンジ，そして平均モーラ長の 9 つの韻律パラメータを用いて発話印象の推定を行っていた．しかしこれらの韻律パラメータだけでは発話印象を推定するのに未だ不十分である．そこで今回は F0 モデルという韻律を数学的に説明するモデルに着目し，このモデルに使用されているパラメータを加えて推定を試みた．F0 モデルのパラメータを求めるためには非常に多くのパラメータを有した最適化問題を解くことが要求されるが，本研究室においてこの F0 モデルのパラメータ推定の自動化がなされており，また，話者交替の有無によりそのパラメータに有意差が生じる可能性を見出している [7].

### 3.1. F0 モデル

F0 モデルは藤崎らによって提案された韻律を数学的に説明するモデルであり [8]，喉頭の制御に基づいた声帯の振動の変化から生理的に説明されている．F0 モデルはフレーズ成分とアクセント成分という 2 つの成分の線形和によって構成されている．フレーズ成分は発話頭から発話末にかけて緩やかに減衰する成分であり，インパルス応答の形で記述される．またアクセント成分は局所的に上昇下降する成分であり，ステップ応答の形で記述されている．

F0 モデルは(1)式のように記述される．

$$\ln F_0(t) = \ln F_b + \sum_{i=1}^{I} A_{pi} G_{pi}(t - T_{0i}) + \sum_{j=1}^{J} A_{aj} \left\{ G_{aj}(t - T_{1j}) - G_{aj}(t - T_{2j}) \right\} \quad (1)$$

ここで，$F_b$ は基本周波数の基底値であり，話者ごとのベースとなる F0 値を示す．$A_p$ はフレーズ指令（インパルス）の大きさ，$A_a$ はアクセント指令（ステップ）の大きさであり，$T_{0i}$ は i 番目のフレーズ指令の生起時点，$T_{1j}$ は j 番目のアクセント指令の始点，$T_{2j}$ は j 番目のアクセント指令の終点を示す．

また，(1)式内の $G_p$, $G_a$ はそれぞれフレーズ制御機構，アクセント制御機構の関数であり(2)式，(3)式によって記述される．

$$G_{pi}(t) = \begin{cases} \alpha_i^{\,2} t e^{-\alpha_i t} & : t \geq 0 \\ 0 & : t < 0 \end{cases} \quad (2)$$

$$G_{aj}(t) = \begin{cases} \min\left[1 - (1 + \beta_j t) e^{-\beta_j t}, \gamma\right] & : t \geq 0 \\ 0 & : t < 0 \end{cases} \quad (3)$$

ここでαはフレーズ制御機構の固有角周波数でありフレーズ成分の減衰の速さを（図 2），βはアクセント制御機構の固有角周波数でありアクセントの上昇下降の早さを（図 3）決定するパラメータである．

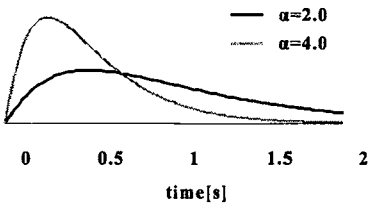

図 2　αの値によるフレーズ成分の減衰の差

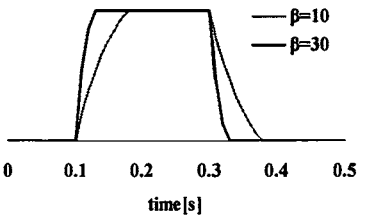

図 3　βの値によるアクセント成分の応答速度の差

これを実際の音声に適用すると図 4 のようになる．

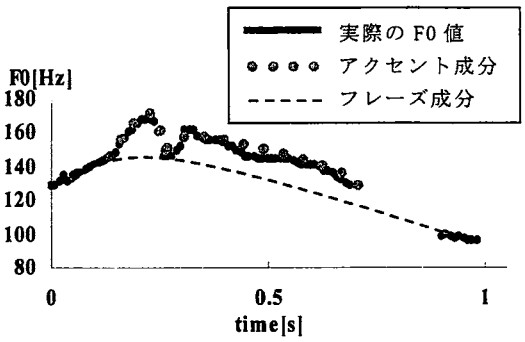

図 4　F0 モデルの適用例

このように，F0 モデルでは F0 をフレーズ成分とアクセント成分の線形和の形で表すことができる．

今回は $F_b$，$A_a$，$A_p$，αに着目し，$F_b$，文頭の $A_a$，文末の $A_a$，$A_p \times \alpha$ の平均値を発話印象推定のためのパラメータとして追加した．文頭，文末の $A_a$ はそれぞれ文頭，文末から 0.5 秒間のアクセント指令の大きさで，文頭，文末 0.5 秒間にアクセント指令がなかった場合は値を 0 とした．αはフレーズ成分の減衰の速さを表すパラメータであり，$A_p \times \alpha$ は F0 パターンの勾配を表すパラメータになっている．

### 3.2. 実験方法

まずテキストの実験と同様に，音声から感じられる発話印象の主観評価実験を行う．実験データには千葉大学で収録された日本語地図課題コーパスを用いた．これは一人の話者が正解ルートの描かれた地図を見ながら，正解ルートの描かれていない地図を持っているもう一人の話者に指示を出し目的地へ導くという対話を収録したものである．ここから話者 10 名，それぞれ十数発話の計 130 発話を用いた．発話は言語情報による発話印象への影響を抑えるため，文脈情報の現れにくい 2，3 秒程度の短いものを選んだ．被験者 10 名にこの音声を聞かせ，2 章において表 1 のテキスト変化で表現できるとされた「強調，疑問，驚き，自信，迷い」の 5 つの発話印象がそれぞれどの程度感じられたかを-3 から 3 の 7 段階で評定させ，10 名の評定した結果の平均値をその音声の評定値とした．テキストの実験と同様に，評定値が 1.0 以上のものはその発話印象が感じられるものとし，1.0 より小さいものはその発話印象が感じられないものとする．

その評定値と韻律パラメータを用いて，重回帰分析により発話印象の推定実験を行う．重回帰分析とは複数の変数（説明変数）から 1 つの変数（目的変数）を推定するときなどに使われる手法で，今回は韻律パラメータから発話印象の評定値を推定する．また，重回帰分析を通して推定に最も有効なパラメータの組み合わせを知ることができる．なお，重回帰分析の結果から推定精度を求めているが，これは評定値と推定結果が，ともに 1.0 以上の場合を「印象あり」で正解，ともに 1.0 以下となった場合を「印象なし」で正解，それ以外を不正解として精度を求めたものである．

### 3.3. 実験結果と考察

主観評価実験の結果，各発話印象において印象ありとなったデータ数の内訳を表 4 に示す．なお，全 130 発話のうちどの発話印象も感じられない音声が 42 発話あり，本研究ではこれを平静音声とみなす．推定すべき発話印象以外の発話印象による韻律の変化による影響を考慮し，重回帰分析には各発話印象の印象ありの音声と平静音声を学習データとして用いた．たとえば強調ならば，印象ありの 52 発話と平静音声の 42 発話の計 94 発話で重回帰分析を行っている．

表 4　各発話印象の印象ありのデータ数

| | |
|---|---|
| 強調 | 52 |
| 疑問 | 40 |
| 驚き | 21 |
| 自信 | 31 |
| 迷い | 30 |

次に，音声における発話印象間の相関関係を表 5 に

示す.

表 5　発話印象間の相関係数（音声）

| | 強調 | 疑問 | 驚き | 自信 | 迷い |
|---|---|---|---|---|---|
| 強調 | 1.00 | -0.09 | 0.54 | 0.62 | -0.44 |
| 疑問 | -0.09 | 1.00 | 0.47 | -0.69 | 0.78 |
| 驚き | 0.54 | 0.47 | 1.00 | -0.04 | 0.29 |
| 自信 | 0.62 | -0.69 | -0.04 | 1.00 | -0.86 |
| 迷い | -0.44 | 0.78 | 0.29 | -0.86 | 1.00 |

表 5 において，自信と迷いのように逆の意味を持つと思われる発話印象には負の相関が，強調と自信のように似ていると思われる発話印象には正の相関がみられるので，妥当な評定結果が得られていると考えられる.

次に，F0 パラメータを追加する前の変数選択した結果を表 6 に，追加した後の変数選択の結果を表 7 にそれぞれ示す.

表 6　変数選択後の韻律パラメータ（追加前）

| | |
|---|---|
| 強調 | F0 最大値，パワー最大値，パワー平均値 |
| 疑問 | F0 平均値，F0 レンジ，パワー最大値<br>パワーレンジ |
| 驚き | F0 レンジ，F0 最小値，F0 最大値<br>パワーレンジ |
| 自信 | F0 レンジ，パワー平均値，平均モーラ長 |
| 迷い | F0 最大値，パワー最小値，平均モーラ長 |

表 7　変数選択後の韻律パラメータ（追加後）

| | |
|---|---|
| 強調 | F0 最大値，パワー最大値，$A_p \times \alpha$ 平均値 |
| 疑問 | F0 平均値，F0 レンジ，パワー最大値<br>パワーレンジ |
| 驚き | F0 レンジ，F0 最小値，F0 最大値<br>パワーレンジ |
| 自信 | F0 レンジ，パワー平均値，平均モーラ長 |
| 迷い | F0 最大値，パワー最小値，平均モーラ長<br>文頭 $A_a$，文末 $A_a$，$A_p \times \alpha$ 平均値 |

表 6 と表 7 を比較すると，強調と迷いにおいて追加したパラメータが選択されていることがわかる．逆に，疑問，驚き，自信では追加したパラメータは変数選択の結果残らなかった．次に，F0 パラメータを追加した後の重回帰式を(4)式から(8)式に示す．なお，式中の $w_{i1}$，$w_{i2}$，$w_{i3}$，$w_{i4}$ はそれぞれ F0 の平均値，最大値，最小値，レンジを，$w_{i5}$，$w_{i6}$，$w_{i7}$，$w_{i8}$ はそれぞれパワーの平均値，最大値，最小値，レンジを，$w_{i9}$ 平均モーラ長を，$w_{i10}$ は $F_b$ を，$w_{i11}$，$w_{i12}$ はそれぞれ文頭，文末の $A_a$ を，$w_{i13}$ は $A_p \times \alpha$ の平均値を表している.

強調　$$y_i = 0.61 + 0.44w_{i2} + 0.64w_{i6} + 0.23w_{i13} \quad (4)$$

疑問　$$y_i = 0.58 + 0.50w_{i1} + 0.57w_{i4} - 0.56w_{i6} + 0.72w_{i8} \quad (5)$$

驚き　$$y_i = -0.11 - 0.69w_{i2} + 1.44w_{i3} + 1.83w_{i4} + 0.36w_{i8} \quad (6)$$

自信　$$y_i = 0.50 + 0.24w_{i4} + 0.87w_{i5} - 0.23w_{i9} \quad (7)$$

迷い　$$y_i = 0.51 + 0.39w_{i2} - 0.29w_{i7} + 0.43w_{i9} + 0.23w_{i11} - 0.22w_{i12} + 0.26w_{i13} \quad (8)$$

重回帰式を見ると強調ではパワー最大値が，疑問ではパワーレンジが，驚きでは F0 レンジが，自信ではパワー平均値が，迷いでは平均モーラ長が最も有効なパラメータであることが分かる.

次に，表 8 と表 9 に F0 パラメータを追加する前と F0 パラメータを追加した後の重回帰分析による発話印象の推定精度を「印象あり」，「印象なし」，「全体」それぞれ別々に求めたものを示す.

表 8　発話印象の推定精度（F0 パラメータ追加前）

| | 強調 | 疑問 | 驚き | 自信 | 迷い | 平均 |
|---|---|---|---|---|---|---|
| 印象あり | 73% | 55% | 48% | 61% | 37% | 54% |
| 印象なし | 88% | 93% | 100% | 90% | 95% | 93% |
| 全体 | 80% | 74% | 83% | 78% | 71% | 77% |

表 9　発話印象の推定精度（F0 パラメータ追加後）

| | 強調 | 疑問 | 驚き | 自信 | 迷い | 平均 |
|---|---|---|---|---|---|---|
| 印象あり | 65% | 55% | 48% | 61% | 43% | 55% |
| 印象なし | 90% | 93% | 100% | 90% | 98% | 94% |
| 全体 | 77% | 74% | 83% | 78% | 75% | 77% |

疑問，驚き，自信では F0 モデルのパラメータが残らなかったため変化はない．強調では，F0 モデルパラメータを追加したことで印象なしの推定精度は上がっ

ているのに対し，印象なしの推定精度は下がっていることから，F0モデルパラメータ追加前に残っていたパワー平均値が強調ありの推定に有効であることがわかった．また，迷いでは，F0モデルパラメータを追加したことで全体の推定精度が向上していることから，文頭，文末の $A_a$ などのパラメータは迷いの推定に有効であることがわかった．

## 4. テキストと音声における相違点

今回の分析によりテキストでは単独で表現できない発話印象があることがわかった．2章でも述べたように，自信が感じられるテキスト変化では必ず強調も感じられてしまう．しかし音声においては自信のみが感じられる音声が存在することも確認されている．表2のテキストにおける発話印象間の相関係数と，表5の音声における発話印象間の相関係数を比較してみると，強調と自信の相関係数がテキストでは0.76，音声では0.62となっており，テキストの場合の方がより相関が強いことが分かる．この他にも迷いと強調，迷いと疑問などテキストと音声の場合で相関関係が大きく変化しているものが見られる．このように，テキストと音声では発話印象の感じ方に違いが生じることが明らかになった．この結果は，今後音声から推定した発話印象をテキストに付与することを考える上で大変興味深い．テキストと音声で発話印象の感じ方に違いが生じることにより，音声から推定することができてもテキストでは表現できない発話印象が存在する可能性がある．よって今後はそのようなことも考慮して，音声，テキスト両方で表現可能な発話印象を利用して分析をしていく必要がある．

## 5. おわりに

本稿ではテキストの変化と発話印象の関係について分析した．その結果，テキストでは単独で表現できない印象がある可能性が示された．同時に，F0モデルのパラメータを新たに追加して，音声による発話印象の推定を行った．また，テキストと音声の発話印象間の相関関係ついて分析したところ，テキストと音声では発話印象の感じ方に違いが生じることが明らかになった．

今回のテキストの評定実験では8種類のテキスト変化についてのみ分析を行ったが，さらに種類を増やして分析する予定である．また音声の発話印象の推定についても，有効と思われる韻律パラメータをさらに追加しての分析や，さまざまな長さの発話を利用しての分析を行っていく予定である．発話印象についても音声とテキストの両方でうまく表現できるようなものを今後検討していく必要がある．


## 謝辞

本研究を進めるにあたり，F0モデルの推定において協力をしていただいた同研究室所属の木村太郎氏に深く感謝いたします．


## 文　献


[1] 西田 昌史，小川 純平，堀内 靖雄，市川 熹，“議事録への付与を目指した発話印象の分析，”音講論，1-6-7，pp.235-236，Sep.2005

[2] 西田 昌史，小川 純平，堀内 靖雄，市川 熹，“対話音声を対象とした韻律情報による発話印象のモデル化，”信学技報，SP2005-103，pp.79-84，Dec.2005

[3] 西田 昌史，小川 純平，堀内 靖雄，市川 熹，“韻律特徴に基づく対話における発話印象の推定，”音講論，1-4-7，pp.225-226，Mar.2006

[4] 江尻 芳雄，金森 康和，“話者の興奮度合いを適用した字幕表現，”信学技報，SP2006-58，Sep.2006

[5] 醤野 和博，柏岡 秀紀，ニック キャンベル，“フェイスマークを用いた自然発話音声における感性情報の分析，”音講論，2-2-7，pp285-286，Sep.2004

[6] 片山 滋友，鈴木 久仁子，谷 史織，“テレビの字幕提示における感情伝達の方法とその効果，” 電子情報通信学会総合大会講演論文集，基礎・境界，424，Mar.2002

[7] 木村 太郎，西田 昌史，堀内 靖雄，市川 熹，“遺伝的アルゴリズムによるF0モデルパラメータ推定手法と話者交替分析への適用，”信学技報，SP2006-82，pp.37-42，Nov.2006

[8] H. Fujisaki and K. Hirose, “Analysis of voice fundamental frequency contours for declarative sentences of Japanese,” Jour. Acoust. Soc. Jpn. (E), Vol.5, No.4, pp.233-242, 1984.