

# Perceptual judgments of emotion for bimodal expressions : interrelating facial and vocal expressions in two dimensional space

メタデータ	言語: jpn 出版者: 公開日: 2017-10-02 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	<a href="https://doi.org/10.24517/00000097">https://doi.org/10.24517/00000097</a>

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.



## Bimodal な表情の感情判断 — 顔と音声による曖昧表情の2次元布置化 —

工藤 治奈<sup>†</sup> 小島 治幸<sup>‡</sup>

<sup>†</sup> 金沢大学文学研究科 〒920-1192 石川県金沢市角間町

<sup>‡</sup> 金沢大学文学部 〒920-1192 石川県金沢市角間町

E-mail: <sup>†</sup> kharu@stu.kanazawa-u.ac.jp, <sup>‡</sup> hkojima@kenroku.kanazawa-u.ac.jp

あらまし 顔と音声による bimodal な表情を判断する際には、主に顔表情によって判断されるといわれている。しかし、本研究では音声表情の影響の大きさに着目し、顔優位説を再検討した。本研究では、次元説の立場に立ち、顔、音声の2次元の各表情要素がどのように影響し合っているのかを検討した。実験の結果、むしろ音声表情が bimodal 表情判断に大きく貢献する場合があります、感情判断は各感情によって様々で、一概に顔優位とも音声優位ともいえないということがわかった。

キーワード bimodal, 表情, 顔, 音声, 曖昧表情

## Perceptual judgments of emotion for bimodal expressions — interrelating facial and vocal expressions in two dimensional space —

Haruna Kudo<sup>†</sup> Haruyuki Kojima<sup>‡</sup>

<sup>†</sup> Graduate School of letters, Kanazawa University Kakuma-machi, Kanazawa-shi, Ishikawa, 920-1192 Japan

<sup>‡</sup> Faculty of letters, Kanazawa University Kakuma-machi, Kanazawa-shi, Ishikawa, 920-1192 Japan

E-mail: <sup>†</sup> kharu@stu.kanazawa-u.ac.jp, <sup>‡</sup> hkojima@kenroku.kanazawa-u.ac.jp

**Abstract** It has been said that facial expressions dominate vocal expressions in recognizing one's emotion. In this study, we re-examined the influence and the importance of vocal expressions on facial expressions, and vice versa, for bimodal expressions. The present study, then, tried to analyze the distances of emotional relationships in two dimensional space in terms of Multi-Dimensional Scaling. The results showed that vocal expressions sometimes rather contributes to assessing emotional states, so that facial expressions were not always dominant to vocal expressions.

**Keyword** bimodal, expression, face, voice, ambiguous expression

### 1. はじめに

#### 1.1. 本研究の意義

コミュニケーションにおける感情判断では、顔表情が主体で、音声表情は顔表情に基づく判断を「促進する文脈」であるとする研究報告がある(重野, 1998; 山田, 2000)。しかし、音声表情は単に顔表情の補佐的な働きしか持たないのであろうか? 本研究では先ず音声表情の影響の大きさを再検討したい。また、視覚表情と聴覚表情という bimodal 表情に対する感情評価の相互関係を検討することで、これまでの顔表情、音声表情単独のモデルを統合し、より包括的な表情判断モデルを作ることを目指している。

#### 1.2. 単独表情の判断過程について

これまで、カテゴリー説と次元説という、2つの大きな立場からの顔表情研究が、数多く行われてきた。Russell(1980)は次元説の立場から多次元尺度構成

法により、感情の類似性を2次元空間の中にプロットし、「快-不快」「覚醒-睡眠」の軸から成ることを示した。この2次元解釈はその後の顔表情研究においてしばしば用いられることとなった。しかし、「快-不快」や「覚醒度」次元に対応する顔表情の知覚的判断がどのような物理的次元、視覚的手がかりに基づいて行われているのかは、未だ明確とはいえない。

一方、音声表情について、越川・向後(1994)は感情を込めて読んだ文章(「おはようございます」「今日はいい天気ですね」「頼みたいことがあるんです」等)に対して「快-不快」「強い-弱い」「速い-遅い」の3次元について評定させ、3次元における音声表情の布置を検討している。それによると、音声表情においてもこれまで顔表情の分野で報告されてきたものと同様の布置が存在することが示唆された。またそのなかで、他の感情領域の近くに布置する刺激はその感情を表現していると誤って判断されやすいことも考察している。

しかし、視覚の手がかり以外の、他の感覚による表情研究については未だ少ないと云わざるをえない。

### 1.3. bimodal 表情の判断過程について

日常、怒りを隠して笑い顔を作ってみせるなど、感情を100%表出しないことはよくある。このようなとき、表出者の顔は笑顔であっても声は怒っているというように、複数の表情が食い違って表出されていても、我々は相手の感情を読み取らなくてはならない。その場合手がかりとして、どの感覚様相 (modality) がより利用されるのだろうか。

重野 (1998) は、顔表情と音声表情の矛盾する場合について、顔画像に音声を同期させて提示し、選択肢の中から話者の情動をひとつ選択させた。その結果、顔画像の情動への同定率が音声表情への同定率を大きく上回っていた。これにより顔表情と音声表情では、顔表情のほうが優先されて利用されると結論づけた。

しかし、顔表情への同定率が高いからといって、音声表情はあまり情動判断に利用されていないといえるだろうか。この方法では被験者に顔表情か音声表情の二者択一を迫っており、情動判断における音声表情の貢献度を測定できない。

本研究では、統合的〔对人的〕情動判断における音声表情の有効性を再確認するとともに、音声表情と顔表情の相互関係について調べた。

## 2. bimodal 表情の評価実験

顔表情と音声表情が矛盾した場合のずれを検討し、顔表情と音声表情が、どのように統合、判断されるのかを明らかにすることを目的とした。顔表情と音声表情が組み合わされる場合、一方の表情がもう一方へどのように影響し、どちらが優位に働くのだろうか。

### 2.1. 刺激の作成、選定

#### 2.1.1. 顔表情刺激

まず、多数撮影した感情顔写真の中から、実験で用いる刺激材料を選定した。

大学生および大学院生5名が被写体となり、被写体の正面からデジタルカメラで撮影した。7表情 (喜び、楽しみ、驚き、怒り、悲しみ、恐れ、嫌悪) を指定し、鏡を渡して訓練させながら計35枚撮影した〔感情顔画像〕。評定には18~23歳の男女60名が参加した。撮影した35枚の顔画像をパーソナルコンピュータによりランダムに提示し、強制選択による7表情へのカテゴリー判断をしてもらった。

喜びと楽しみの混同率が高かったが、それらを同一カテゴリーとしてまとめたところ、全感情顔画像とカテゴリー判断との一致率が平均97.3%と非常に高くなったため、本実験では以下、同一の感情「喜び」として扱うことにした。恐れ表情に関しては一致率が

1.7%から36.7%、平均で15.3%と低いため除外した。

複数の感情顔画像のうち一致率が75%以上と高く、表情間での混同率の低い画像を各表情1枚ずつ計5枚選び、本実験での顔表情刺激とした。

#### 2.1.2. 音声表情刺激

顔表情と同様に、大学生および大学院生5名が話者となった。7表情 (喜び、楽しみ、驚き、怒り、悲しみ、恐れ、嫌悪) を指定し、本来感情に無関係な単語を感情様に発声させたものを、サウンドレコーダーにより録音した。単語は感情とは無関係な意味語 (つくえ・ねずみ・テレビ)、無意味語 (にりと・ままま・ねねね) 計6語を用い、各人 (5人) について7表情×6語=42音声、計210音声録音した。

評定には18~23歳の男女60名が参加した。録音した210の音声刺激を、パーソナルコンピュータにより、単語ごとにランダムに提示し、評定用紙での選択により7表情カテゴリーへ強制選択判断させた。

顔表情同様に、喜びと楽しみを同一カテゴリーとしてまとめることにより一致率が100%に近づいたため、1つの感情として扱うこととした。恐れは31.3%と正答率が低いので除外し、また認識率が75%以上であり、表情間での混同率の低い刺激を各表情1つ選んだ結果、本実験での音声刺激としたのは、「ままま」喜び、「ままま」驚き、「にりと」怒り、「にりと」悲しみ、「ねねね」嫌悪でとなった。

### 2.2. 実験協力者

19~23歳の男女18名が実験に参加した。

### 2.3. 装置

心理学実験用ソフトウェア Super Lab または e-Prime 1.1 を用いて、顔表情と音声表情の単独刺激、複合刺激を提示した。

### 2.4. 手続き

#### 2.4.1. 顔表情単独の感情評定

選定した各画像刺激を単独提示し、観察者にはそれぞれの感情の度合いを0~10の11段階で評定させた。ひとつの表情刺激に対して、「喜び」「驚き」「怒り」「悲しみ」「嫌悪」の5感情についてすべて評定させ、bimodal 表情の評定値と比較できるようにした。

#### 2.4.2. 音声表情単独の感情評定

選定した各音声刺激を単独提示し、それぞれの感情度を0~10の11段階で評定させた。ひとつの表情刺激に対して、「喜び」「驚き」「怒り」「悲しみ」「嫌悪」の5感情についてすべて評定させ、bimodal 表情の評定値と比較できるようにした。

#### 2.4.3. 顔、音声の bimodal 表情の感情評定

各画像刺激と音声刺激を同時に提示し、両刺激が完全に終了した後、それぞれの感情度を0~10の11段階で評定させた。ひとつの組み合わせ刺激に対して「喜

び」「驚き」「怒り」「悲しみ」「嫌悪」の5感情についてすべて評定させ、被験者毎に顔表情 5×音声表情 5×評定感情5の計 125 回試行した。

## 2.5. 結果と考察

bimodal 表情について、評定させた感情ごとに、顔表情 (5) × 音声表情 (5) の2要因 (被験者内) で分散分析を行った。すべての評定感情で、顔表情の主効果と共に音声表情の主効果があつた ( $p < 0.05$ )。また、感情「喜び」「驚き」「悲しみ」について顔表情と音声表情の交互作用がみられた。

顔表情と音声表情ともに主効果があつたことから、われわれが顔と音声で表情の矛盾した他者の感情を認知するときには、顔表情と音声表情のそれぞれを十分に読み取って判断しているということがいえる。よって、顔表情と音声表情は独立した情報として処理されていると考えられ、この段階において必ずしも顔表情が優先するとはいえないだろう。

ただし、bimodal 表情の感情判断をするためには、並列に処理された後に2つの情報を統合する必要がある。本評定の結果、顔表情と音声表情の評価の度合いに交互作用がみられたことは、われわれが顔と音声の各表情を独立に評価しているのではないことを示している。

## 2.6. 2次元的分析

次に、顔表情と音声表情が、次元空間においてどのように統合されているのかを検討するために、多次元尺度法による2次元布置化を試みた。ここでは、顔、音声の単独表情および bimodal 表情刺激 (顔単独 5、音声単独 5、bimodal 表情 25 刺激) の全刺激対において、ユークリッド距離を求めた。距離は、評定された感情ごとに評定値の差をとり、5つの差全ての2乗和から算出した。その関係を図1に示した。

これまでの多数の顔表情研究に照らし合わせて、次元1は「快-不快」次元、次元2は「覚醒度」次元と解釈することができる (Takehara and Suzuki, 1997; Ogawa, et al. 1999)。Takehara and Suzuki (1997) の、顔表情の合成刺激を用いた場合と同様に、顔表情と音声表情を合成した本研究でも、顔-音声合成表情は、顔あるいは音声単独の表情を結んだ直線付近に位置された。つまり、顔表情と音声表情という bimodal 表情でも、顔表情の合成と同じように、線形的な関係によって合成表情が認知されていることが示唆される。

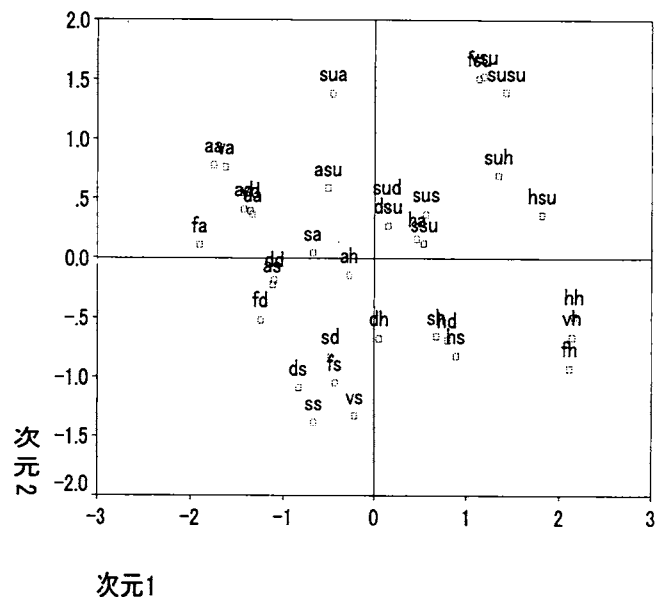
次に、顔表情が音声表情に、音声表情が顔表情に、どの程度優位に影響するのかを検討するために、各重心の移動に着目した。顔表情単独提示条件の5つの評価点の重心は、顔表情のニュートラル点と考えることができる。また、[ある顔表情と、5通りの音声表情と

の組み合わせ提示]の評価点の重心を顔表情5通りについて求めることができる (例、ah, asu, aa, as, ad の重心は、顔表情が怒りの bimodal 表情の重心であり、怒りの顔表情が bimodal 表情判断に与える影響をみることが出来る)。音声表情についても同様である。

例として、顔表情が嫌悪 (disgust: d) の場合を図2に示した。顔表情 [嫌悪] の単独提示時の評価点と、顔表情 [嫌悪] と5つの音声表情との bimodal 表情の重心、そして音声のニュートラル (VN) 点 (音声単独提示の5表情の評価点の重心) との関係を見ることが出来る。この場合、bimodal 表情「顔 d」重心が、VN と顔単独 d の中間点より顔単独 d に近ければ、音声表情よりも顔表情が優位に判断されているといえる。

また、評価点の移動は、x 軸 y 軸の表す感情因子方向への表情評価の変移を示すと考えられる。このため、x 軸、y 軸それぞれについて、重心移動指標 (Gravity Shift Index: GSI)

$$GSI = (\text{bimodal 重心} - N) / (\text{単独} - N)$$
 を計算し、その値を表1に示した。



顔単独、音声単独、混合表情すべての布置

図1.(f:顔単独, v:音声単独), (h:喜び, su:驚き, a:怒り, s:悲しみ, d:嫌悪), (1文字目:顔表情, 2文字目:音声表情 ex. ha:顔-喜び, 音声-怒り)

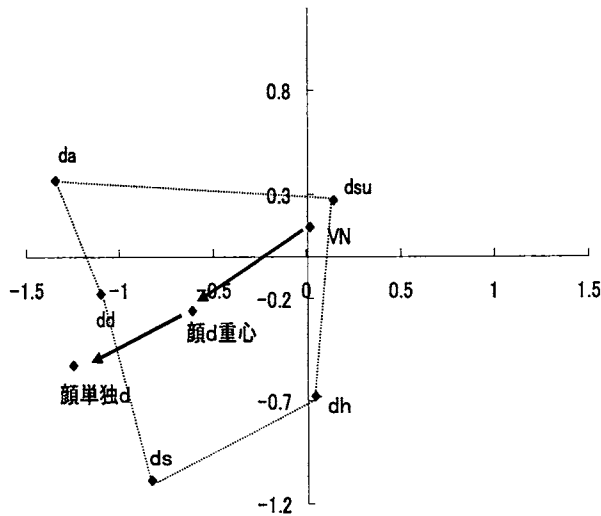


図 2. 顔表情が嫌悪(disgust)の場合の動き

表 1. 顔表情がそれぞれの場合に付加される音声重心 (VN) の移動 (GSI にみる評価点の動き), および音声表情がそれぞれの場合の顔表情重心 (FN) の移動

	音声重心(VN)の移動: 顔表情の影響		顔重心(FN)の移動: 音声表情の影響	
	x	y	x	y
happy	0.574	0.407	0.389	0.160
surprise	0.518	0.522	0.598	0.425
anger	0.541	-4.524	0.440	0.771
sadness	0.321	0.572	1.110	0.393
disgust	0.502	0.606	0.273	0.009
AVE.	0.491	-0.483	0.562	0.351

GSI が 0 に近いことは顔表情 (音声表情) の評価の重心に、音声表情 (顔表情) 刺激がほとんど影響を与えなかったということである。また、1 に近ければ、顔表情 (音声表情) の評価の重心に、音声表情 (顔表情) 刺激が極めて大きな影響を与えたと考えられる。

表 1 の顔表情の影響においては、anger 表情を除いた 4 表情で、x 値も y 値も 0.5 に近くなっていることから、この場合、bimodal 表情の判断で顔表情はおよそ 2 分の 1 の影響力しかもたないことを示している。Anger 表情については、例外的に y 軸への影響がマイナスに大変大きくなっている。つまり、anger という顔表情は、単独の場合よりも、音声表情と組み合わせられた時に大きな影響をもつ顔表情であるといえる。また、音声表情の影響において、sadness は 1.110 と、顔重心を x 軸方向に強力に牽引している。一方 disgust の音声表情はあまり bimodal 表情の感情判断に貢献していないと考えられる。

このように、bimodal 表情における顔表情、音声表情の貢献度は、各感情によって様々であり、一概に顔優位とも音声優位ともいえないものである。特に anger 音声表情、sadness 音声表情などは顔重心 (FN) の移動つまり顔表情判断への影響が大きい (顔表情判断への強い牽引力をもっている)。このことは、顔-音声の bimodal 表情では単純に顔優位に判断されるのではなく、むしろ逆に音声表情が bimodal 表情判断に大きく貢献する場合があることを示している。怒りなど、覚醒度に大きな特徴のある音声表情は、感情判断に強く影響する。つまり、音声表情は顔表情の判断を付随的に「促進」するものではなく、それ自体の感情評価を引き出す可能性があるといえる。

### 3. おわりに

今回用いた多次元尺度法によって導出された x 軸は「快-不快」、y 軸は「覚醒度」次元と考えられるが、今後は、顔-音声の bimodal 表情の感情判断とこれらの次元(因子)との関連について明らかにする必要がある。さらに、モデルを精密化し、予言性を高め、ヒトの複雑で曖昧な表情の認知機構を明らかにしてゆきたい。

### 文 献

- [1] Ogawa, T. & Takehara, T. & Monchi, R. & Fukui, Y. & Suzuki, N. "Emotion space under conditions of perceptual ambiguity" *Perceptual and Motor Skills* 88, Pp.1379-1383 (1999)
- [2] Ogawa, T. & Suzuki, N. "Response differentiation to facial expression of emotion as increasing exposure duration" *Perceptual and Motor Skills* 89, Pp.557-563 (1999)
- [3] 小川時洋 "曖昧表情認識研究とその生態学的妥当性" *Japanese Psychological Review* 43, 2, Pp.295-306 (2000)
- [4] 越川房子・向後礼子 "音声の感情認知に関する予備的研究 (I)" *日本教育心理学会第 36 回総会発表論文集* 308 (1994)
- [5] 重野純 "感情認知に果たす視聴覚情報の役割 - 日本人とアメリカ人の比較 -" *日本音響学会講演論文集* Pp.379 - 380 (Sep.1998)
- [6] 重野純 "感情を表現した音声の認知と音響的性質" *The Japanese Journal of Psychology*, 74, 6, Pp.540-546 (2004)
- [7] Takehara, T. & Suzuki, N. "Morphed images of basic emotional expressions: ratings on Russel's bipolar field" *Perceptual and Motor Skills* 85, Pp.1003-1010 (1997)
- [8] Russell JA. "A circumplex model of affect" *Journal of Personality and Social Psychology*, 39 Pp.1161-1178 (1980)