

基本 6 感情を表出した表情と音声に対する感情判断 -表情と音声を入れ替えた刺激に対する判断の日蘭比較-

Emotion perception by face and voice representing six basic emotions - In case of incongruent emotional stimuli -

田部井 賢一¹, 高木 幸子², Huis In't Veld, Lisanne³, de Gelder, Beatrice³, 田中 章浩⁴
Ken-ichi Tabei¹, Sachiko Takagi², Lisanne Huis In't Veld³,
Beatrice de Gelder³, Akihiro Tanaka⁴

¹三重大学大学院医学系研究科, ²東京女子大学大学院人間科学研究科,

³Department of Medical Psychology and Neuropsychology, Tilburg Univ., ⁴東京女子大学現代教養学部

¹Graduate School of Medicine, Mie Univ., ²Graduate School of Humanities and Sciences, Tokyo Woman's Christian Univ., ³Department of Medical Psychology and Neuropsychology, Tilburg Univ.,

⁴Department of Psychology, Tokyo Woman's Christian Univ.

tabei@clin.medic.mie-u.ac.jp

Abstract

Information from face and voice plays an important role in social communication. As shown in the study of speech perception, facial and vocal signals are integrated even in the perception of emotion. Our study investigated cultural differences between Japanese and Dutch participants in the multisensory perception of emotion. A face and a voice, expressing incongruent six basic emotions, were presented on each trial. Participants were instructed to judge the expressed emotion in audiovisual, visual-only and audio-only conditions. When participants judged the emotion of incongruent audiovisual stimuli, response based on vocal information was more frequently observed in Japanese than in Dutch participants, whereas response based on facial information was less frequent in Japanese than in Dutch participants. This result indicates that Japanese people are more attuned than Dutch people to vocal processing in the multisensory perception of six basic emotions in general.

Keywords — Facial expression, Vocal expression, Incongruent stimuli, Cultural differences

1. はじめに

人間同士のコミュニケーションでは、言語情報のみならず表情、声の調子、身振りなどといった非言語情報が表出され、聞き手に伝達される。日常生活では単一の感覚モダリティだけではなく、複数の感覚モダリティを用いたコミュニケーションがおこなわれている。表情と音声を用いたコミュニケーションでは、ある感情を表出する際に話し手の表情と音声は同一の感情を表出していると感じられ、異なった感情を表出していると感じら

れることはほとんどない。しかし、高木・平松・田中[1]は日本人の発話を収録して表情だけ、または音声だけというように分けて日本人に呈示すると、それぞれ異なった感情として知覚される刺激が多くみられることを見出した。具体的には、発話を収録して、表情のみあるいは音声のみを呈示した際に最頻回答がいずれも 50%以上である刺激、つまり回答がある程度特定の感情価に集中した刺激に限定した集計では、全体の 1 割程度の割合で表情と音声に対する最頻回答が異なっていた。さらに Takagi, Hiramatsu, Huis in't Veld, de Gelder, and Tanaka[2]はオランダ人の発話を収録して表情のみ、または音声のみをオランダ人に呈示すると、同様の刺激が全体の 2 割程度の割合で存在することを見出した。

表情と音声を用いた視聴覚刺激の知覚には文化差があることが示唆されている。Tanaka, Koizumi, Imai, Hiramatsu, Hiramoto, and de Gelder[3]は、怒りと喜びを表出した表情と音声による視聴覚情動認知の日蘭比較実験をおこない、多感覚情動情報の統合様式における文化差を検討した。研究では表情と音声の感情が一致（喜び顔 + 喜び声, 怒り顔 + 怒り声）または不一致（喜び顔 + 怒り声, 怒り顔 + 喜び声）の動画刺激を作成した。多感覚セッションでは表情と音声の両方を含む動画を呈示し、顔注意課題と声注意課題を実施した。顔注意課題では音声を無視して表情か

ら知覚される感情について判断するように、声注意課題では表情を無視して音声から知覚される感情について判断するように教示した。判断は「喜び」と「怒り」の二肢強制選択とした。その結果、日本人は顔の情動判断時には声からの影響が大きく、声の情動判断時には顔からの影響は小さかった。つまり、一貫して日本人はオランダ人よりも声への依存性が高かった。この結果は、感情の知覚において、表情と音声という多感覚情報の重みづけに文化差があることを示唆している。

Tanaka ら[3]の研究では、表出と知覚が比較的容易であり、快-不快に明確に区分できる喜びと怒りの感情の組み合わせに絞ることで、表情と音声による感情認知に文化差があることを明らかにした。現実場面においても、表情と音声の感情価が食い違っていることは存在し[1]、喜びと怒りに限らず、さまざまな組み合わせが想定できるが、すべての組み合わせにおいて同様の文化差が生じるのか、それとも感情の組み合わせによってモダリティ優位性の文化差に違いが見られるのかどうかは明らかではない。

そこで本研究では、基本 6 感情（怒り、嫌悪、恐怖、喜び、悲しみ、驚き）[5]を表現した動画をもとに、表出された表情と音声の感情価の組み合わせを入れ替えた刺激を作成して、多感覚情報の重みづけにおける文化差を系統的に検討した。これらの 6 感情はすべて、アクションユニット[4]などによって感情表出の方法が系統立てられているため、共通のプロトコルを用いて統制された刺激を作成することが可能である。

2. 方法

実験参加者: 実験参加者は、日本在住の日本人大学生 26 名（男性 13 名、平均年齢 20.3 歳±1.4; 以下、日群）とオランダ在住のオランダ人大学生 25 名（男性 6 名、平均年齢 21.4 歳±3.2; 以下、蘭群）であった。

刺激: 刺激は、本研究室で開発した表情音声刺激セット[1]から、日本人とオランダ人の各 4 名のモデルが基本 6 感情（怒り、嫌悪、恐怖、喜び、悲しみ、驚き）を表出した表情と音声をそれぞれ抽

出した。さらにそれらを用いて編集加工し、表情と音声の示す感情が一致・不一致である人工的な刺激を作成した。前者を一致刺激、後者を不一致刺激とする。

手続き: 刺激は、表情と音声を同時に呈示して表情と音声の両方に注意するように教示した AV 全体条件、表情と音声を同時に呈示して表情のみに注意するよう教示した AV 顔注意条件、表情と音声を同時に呈示して音声のみに注意するよう教示した AV 声注意条件の 3 条件を設けた。なお、表情のみを呈示した V 条件、音声のみを呈示した A 条件も実施した。日・蘭両群ともに同一文化に属する人物の刺激のみを呈示した。

3. 結果

AV 全体条件:

AV 全体条件の不一致刺激呈示時の結果を図 1 から図 3 に示した。図の作成にあたってはまず混同行列を刺激の表情の感情価別に作成した (e.g., 表 1)。混同行列は、行方向がモデルの音声の感情価で、列方向が実験参加者によって回答された感情価である。各行列はモデルによって表現された感情価に対して、実験参加者が回答した感情価の回答率を示している。下線の値は正答率であり、モデルの表現した感情価と一致して回答された率を示している。さらに表情の示す感情と一致する回答である顔反応（背景色グレー）と、音声の示す感情と一致する回答である声反応（背景色黒）を刺激の表情の感情価ごとに求めた。

表 1: 表情の感情価が怒りのときの混同行列。下線の値は正答率であり、モデルの表現した感情価と一致して回答された率を示す。背景色グレーは表情の示す感情と一致する回答である顔反応、背景色黒は音声の示す感情と一致する回答である声反応である。

		表情=怒り					
		回答					
		怒り	嫌悪	恐怖	喜び	悲しみ	驚き
音 声	怒り	<u>74</u>	17	3	0	3	3
	嫌悪	63	<u>25</u>	4	0	5	4
	恐怖	28	12	<u>28</u>	0	15	17
	喜び	55	25	4	<u>0</u>	6	10
	悲しみ	33	16	17	0	<u>28</u>	6
	驚き	37	20	14	1	1	<u>27</u>

感情をプールした結果では、2（日・蘭群）×2（顔・声反応）の2要因分散分析を実施した結果、交互作用（ $F(1,48)=105.48, p<.001$ ）がみられた。単純主効果検定の結果、日群の方が蘭群よりも声反応が多く、蘭群の方が日群よりも顔反応が多かった（ $p<.001$ ）（図1）。つまり、表情と音声の感情価が不一致であるとき、全体としては、日本人のほうがオランダ人よりも声に基づいた判断をおこなう傾向が確認できた。

感情別の結果では、日群では怒り・嫌悪・喜び・驚きの刺激に対しては表情の示す感情と一致する回答である顔反応が多く、恐怖・悲しみの刺激に対しては音声の示す感情と一致する回答である声反応が多かった（図2）。一方、蘭群ではすべての感情価において顔反応が多く、本実験においても蘭群の表情知覚優位が示唆された（図3）。

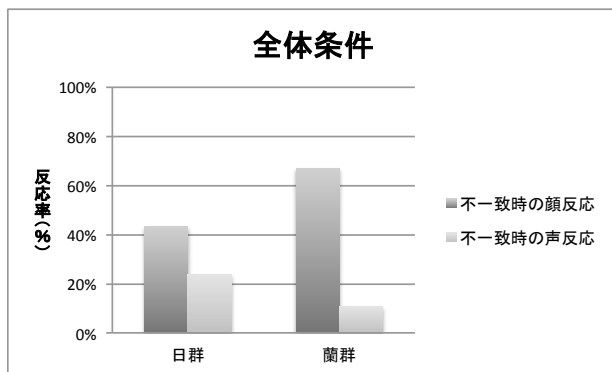


図1：表情と音声不一致時の顔反応と声反応（AV全体条件，日・蘭群）

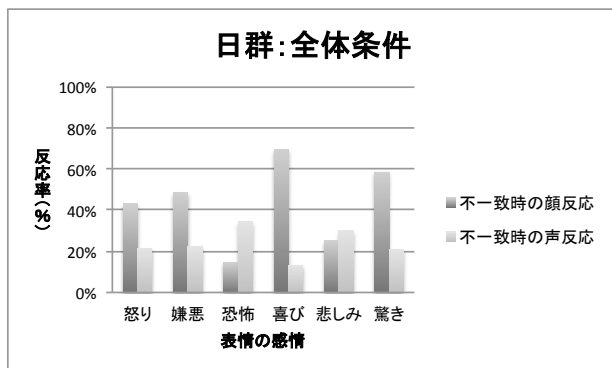


図2：表情と音声不一致時の顔反応と声反応（AV全体条件，日群）

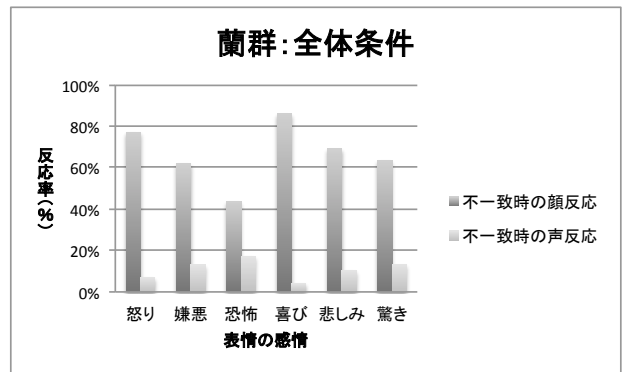


図3：表情と音声不一致時の顔反応と声反応（AV全体条件，蘭群）

AV顔注意，AV声注意条件：

AV顔注意，AV声注意条件の結果を図4から図6に示した。

感情をプールした結果では、日群では両条件ともに一致刺激の方が不一致刺激よりも正答率が高かったが（ $p<.05$ ），蘭群では声注意条件のみ一致刺激の方が不一致刺激よりも正答率が高かった（ $p<.05$ ）（図4）。蘭群の顔注意条件では一致刺激と不一致刺激の正答率に差がなかったことから、蘭群は表情知覚において音声からの影響をほとんど受けないことが示唆される。

感情別の結果では、日群では怒り・嫌悪・喜び・悲しみ・驚きの刺激では顔注意条件の方が声注意条件よりも正答率が高く（ $p<.05$ ），恐怖の刺激では声注意条件の方が顔注意条件よりも正答率が高かった（ $p<.05$ ）（図5）。一方、蘭群ではすべての感情価において顔注意条件の方が声注意条件よりも正答率が高かった（ $p<.01$ ）（図6）。これらの結果から、日群にとって恐怖感情は他の感情価と比べて、表情から知覚しにくいことがわかった。

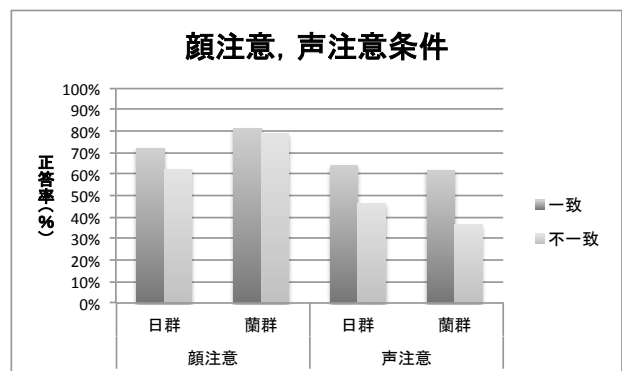


図4：表情・音声一致，不一致時の正答率（AV顔注意・AV声注意条件，日・蘭群）



図 5: 表情・音声不一致時の正答率 (AV 顔注意・AV 声注意条件, 日群)

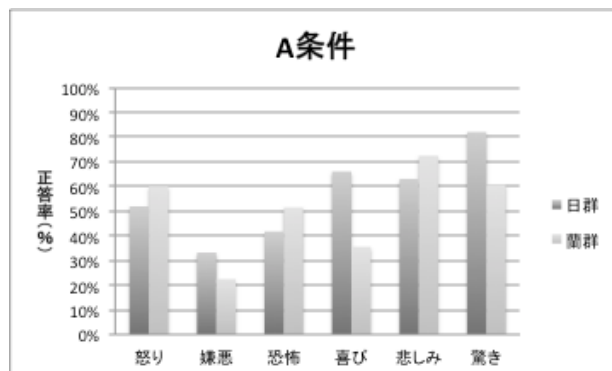


図 8: A 条件の正答率 (日群, 蘭群)

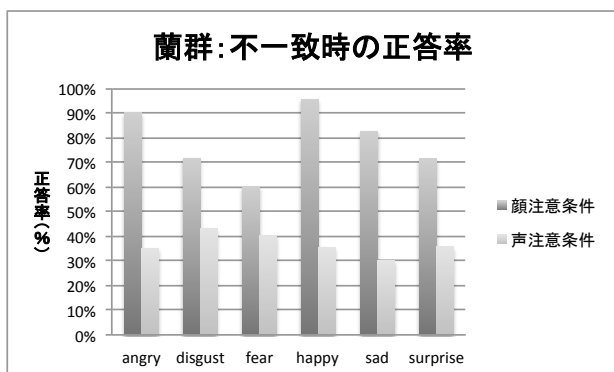


図 6: 表情・音声不一致時の正答率 (AV 顔注意・AV 声注意条件, 蘭群)

V 条件, A 条件:

V 条件と A 条件では呈示条件によらず, 嫌悪, 喜び, 驚きの感情価においては日群の方が蘭群よりも正答率が高く, 怒り, 恐怖, 悲しみの感情価では蘭群の方が日群よりも正答率が高いという結果が得られた ($p < .05$) (図 7, 8).

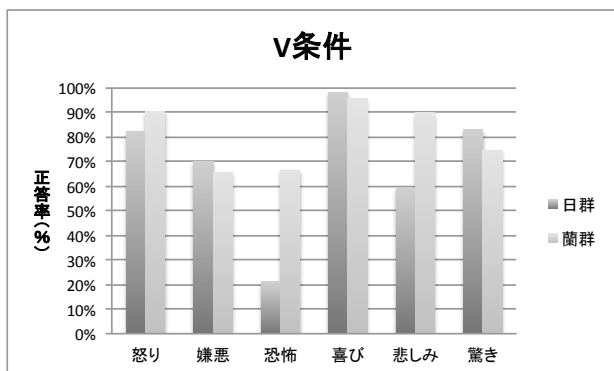


図 7: V 条件の正答率 (日群, 蘭群)

4. 考察

本研究では Tanaka ら [3] を発展させ, 現実場面で想定できるさまざまな組み合わせにおいても同様の文化差が生じるのか, それとも感情の組み合わせによってモダリティ優位性の文化差に違いが見られるのかを検討するために, 基本 6 感情を表現した動画をもとに, 表出された表情と音声の感情価の組み合わせを入れ替えた刺激を用いて, 多感覚情報の重みづけにおける文化差を系統的に検討した.

全ての感情の組み合わせの全体的傾向として, オランダ人と比較して日本人に音声知覚優位が見られ, 日本人と比較してオランダ人に表情知覚優位が示された. この結果は怒りと喜びの感情価を用いた先行研究 [3] の結果と整合しており, 基本 6 感情全体としても同様の傾向がみられることが新たに示唆された. 特にオランダ人の AV 顔注意条件において, 一致刺激と不一致刺激に対する正答率がほぼ同じであることは, 表情の判断をする際に, 表情と不一致な感情価の音声の影響をほとんど受けない, つまり表情知覚優位であることを示唆する興味深い結果である.

感情別の結果では感情価によって傾向の違いがみられた. AV 全体条件において, 日本人の恐怖と悲しみでは表情より音声に対する反応率が高く, 顔注意条件と声注意条件の比較においては日本人の恐怖の感情価において声注意条件の方が顔注意条件よりも正答率が高いというように, 感情の種類によって回答傾向が顕著に異なることが示された. 恐怖と悲しみの感情価において日本人に音声

知覚優位が示唆されたことに関しては、本実験の V 条件の結果（日本人は恐怖表情および悲しみ表情の正答率が低い）をふまえると、オランダ人と比較して日本人では表情の表出そのものが顕著ではないことが要因かもしれない。つまり、恐怖と悲しみの表情表出が顕著でないために適切に情報を読み取れず、表情よりは感情を読み取りやすい音声に依存する傾向が高くなるという可能性を挙げることができる。しかし、日本人が恐怖音声や悲しみ音声の感情認知に長けているかといえ、必ずしもそうとは言い切れない。A 条件における恐怖感情および悲しみ感情の正答率を日蘭で比較すると、日本人の正答率はオランダ人よりも低い。したがって、恐怖感情における日本人の音声知覚優位は感情表出に関わるのか、それとも感情知覚に関わるのかは、本実験だけでは結論づけることができない。今後は日本人がオランダ人刺激を視聴した場合やオランダ人が日本人刺激を視聴する実験を設けることで、さらに日本人の音声知覚優位とオランダ人の表情知覚優位に関して検討をおこなっていききたい。

参考文献

- [1] 高木幸子・平松沙織・田中章浩. (2011). 感情を含んだ表情と音声を組み合わせた日蘭刺激作成とその評価. 日本心理学会第 75 回大会発表論文集, 576.
- [2] S. Takagi, S. Hiramatsu, E.M.J. Huis in't Veld, B. de Gelder, and A. Tanaka, (2011) "Recording and validation of audiovisual expressions by faces and voices", 12th *International Multisensory Research Forum Program Abstracts*, pp.34.
- [3] A. Tanaka, A. Koizumi, H. Imai, S. Hiramatsu, E. Hiramoto & B. de Gelder. (2010). I feel your voice: Cultural differences in the multisensory perception of emotion. *Psychological Science*, 21, 1259-1262.
- [4] P. Ekman, and W. V. Friesen, (1978). *Facial action coding system*, Palo Alto: Consulting

Psychologist. Press.

- [5] P. Ekman, & W. Friesen ., (1976) "Pictures of facial affect" Palo Alto: Consulting Psychologist. Press.

謝辞

本研究は、総務省 SCOPE 若手 ICT 研究者育成型研究開発(10210311)、および JSPS 科研費若手研究(A) (24680030)の支援を受けました。