

音声と表情による感情の認知 一日米間の比較

重野 純

Recognition of emotion transmitted by vocal and facial expression: Comparison between the Japanese and the American

SHIGENO, Sumi

Abstract

Vocal as well as facial expression is considered to play an important role in the recognition of emotion. Furthermore, when we recognize a speaker's emotion, the display rules of the culture to which the speaker belongs give a great influence on the recognition of emotion. Two experiments – visual and auditory – were conducted to investigate the influence of cultural differences between the Japanese and the American on the recognition of facial and vocal expression. Japanese and American professional actors expressed six basic emotions using their faces and voices. Eleven Japanese and 11 American participants identified the actors' emotions. Results showed that the Japanese can identify the Japanese vocal as well as facial expression of emotion more accurately than the American can, and *vice versa*. It was suggested that the cultural differences between a speaker and a listener might be one of the critical factors that more or less decline the recognition of emotion.

Key words: recognition of emotion, vocal expression, facial expression, Japanese and American, display rules

日常のコミュニケーション行動においては、相手の顔に表れる表情だけではなく音声やジェスチャーなどの音声以外の手段も頻繁に用いられる。また、それらの間の相互作用も重要な手がかりとなる。さらに、話者の真の感情がストレートに表出されるとは限らない場合も多い。なぜなら話者が置かれている状況によって本当の感情を表出できない場合もあるだろうし、相手に自分の感情を知られたくない場合もあるだろうからである。たとえば、怒りの感情を感じたとしても社会的な場面においてはそのような感情をストレートに表出することがはばかられ、怒りを隠して笑顔で話さなくてはならないといった場面も少なくない。このような場合、私達の認知行動は顔の表情以外の手段を利用する。そうしないと相手の真の感情を察することが難しいからである。

ところで、感情の表出方法は、話者の属する

文化の中で用いられている顔の表情の統御に関する文化的規則すなわち表示規則 (display rules) によって異なる。したがって、感情についての認知行動は話者の属する文化的要因による影響を受けることになるが、さらに情報を受ける側の文化的要因 (表示規則) によっても大きく影響されるものと考えられる。たとえば、感情を明瞭に顔面に表出しない文化に属する人は、相手の感情を認知する能力も低いと言われている。それでは感情表出が控えめな日本人は、相手の感情を認知する能力も欧米人に劣るのであろうか。確かに日本人は欧米人の様には感情を顔にストレートに表出することは少ないが、他の手段を用いて欧米人よりもっと繊細に相手の感情を認知しているとも考えられる (重野, 1998)。一般に、表情と音声の間で矛盾した感情が表出される場合、西欧文化では音声より表情の方を手がかりにするとされている

(Ekman and Friesen, 1971). 日本人の場合はどうであろうか。

松見ら (1978) によれば、顔面表情が文化に固有なものなのか (文化固有説) または文化の相違を超えて同一なものなのか (普遍説) については見解の分かれるところであるが、基本的な感情に関する限り、文化に共通した表情があるらしい。そして、幸福、驚き、怒り、嫌悪、恐れ、悲しみの6種類の基本感情の表出には文化を超えた共通性のあることが確かめられている (Ekman & Friesen, 1971)。また、米谷ら (1994) の日米のTVドラマを用いた表情識別実験によれば、日本人の表情認識構造は、幸福、驚き、怒り、嫌悪、恐れについてはほぼ欧米人と同様になされているが、驚きと恐れ、怒りと嫌悪については混同されやすく、怒りと嫌悪の位置関係を除き、Schlosberg (1952, 1954) の円環モデルとよく類似している。しかし、悲しみの感情認知については差が認められている。

これまで感情認知について多くの研究が行われてきたが、それらの多くが顔面表情の認知に関するものであった。しかし上述したように実際のコミュニケーション行動においては、顔面表情と同等か場合によってはそれ以上に音声などの他の手段が重要な役割を果たすことがある。ところが音声を用いた研究は少ない (Ekman, 1985)。本研究では日本人とアメリカ人をそれぞれ情報を送り手 (話者) と受け手 (被験者) として用い、異なる文化的背景をもつ両者の認知行動を比較し、認知行動における社会・文化的環境の差異と感情認知に果たすそれらの役割を調べることを目的とする。

実験1

これまでの感情認知の研究では表情写真が用いられることが多かった。表情写真は特定の感情が最もよく表出された瞬間の静止画である。しかし実際のコミュニケーション場面では相手の表情が静止していることは少なく、表情運動のような動的な表情を認知する場合の方が多。また、音声による感情の認知と比較する場合には、時間とともに変化するという点で動的

な刺激の方がより適切であると考えられる。そこで本実験では、ビデオ映像を用いて動的な表情についての認知実験を行い、日米文化間の比較を行うこととする。

1. 方法

被験者 日本人大学生11名とアメリカ人大学生11名が実験に参加した。いずれも1年以上にわたる外国滞在経験はなかった。またこの種の実験には参加した経験がなかった。全員が日本人話者条件とアメリカ人話者条件の両方に参加した。

刺激 5つの語や短文を、日本語および米語からそれぞれ選んだ (表1)。これらの刺激材料は、地名、人名、時刻、曜日、あいさつ、相槌というように、有意味ではあるがそれ自体が特定の感情を表わすものではなく、かつ日常生活の種々の場面において感情を伴って発話することがあると考えられるものである。これらの刺激材料は6つの基本的な感情である「幸福」(happiness)、「驚き」(surprise)、「怒り」(anger)、「嫌悪」(disgust)、「恐れ」(fear)、「悲しみ」(sadness)を表すように表情によって表現された。話者は、顔の表情によって感情を明瞭に表現できる日本人とアメリカ人のプロの男性俳優各1名であった。両者とも30代で、演劇経験が8年以上あった。俳優は刺激材料をなるべく自然にかつ感情が明瞭に表れるような表情を作るように演技することを求められた。

各刺激材料は数回演技したものの中から、実験者とプロデューサーがともに最もよく感情を表出していると判断したものを一つ選んだ。刺激の収録と編集は、専門の業者に委託した。収録と編集はβカムSP (SONY, BVW-35)を用い

表1 実験に用いた刺激語と短文

日本語	米語
東京	New York
河原崎さん	Rio de Janeiro
11時半	Margaret
さようなら	Saturday
そうですか	Is that so?

て行い、実験用にS-VHSテープにダビングした。

手続き 5個の刺激材料のそれぞれを6種類の感情を表すように表現したものを、各刺激が2回ずつ提示されるようにランダムに提示した。したがって、試行数は60試行（5刺激材料×6感情×2回）であり、一つの感情につき被験者一人当たり10回の判断を得た。各刺激の提示前1.5sに合図音として200msの純音を提示した。反応時間はすべて3sとした。刺激の提示時間は、日本語では1693-2254 ms、米語では1858-2650msであった。

被験者は、各刺激について話者の感情を6つの基本感情の中から一つだけ選ぶように求められた。この際、2つ以上の感情を同時に感じた場合には、最も強く感じたものを答えるように指示した。刺激はすべてS-VHSビデオデッキ（Victor, HR-Z1）により再生し、29インチモニタ（Victor, AV-29F2）を用いて提示した。モニタから被験者までの距離は約1m73cmであった。実験中は実験者が被験者の隣に着席し、被験者がモニタを常に確実に見ていることを確認した。

2. 結果と考察

反応は、話者が意図した感情と同じ感情に被験者が同定した場合を正答とした。日本人話者、アメリカ人話者ごとの平均正答率を図1に示す。どちらの話者条件においても正答率は全体的に高かった。以下、話者条件別に結果を検討する。

日本人話者条件（図1(a), 図2(a)）: 日本人被験者では正答率は全体的には高かったが、恐れのみ非常に低かった。しかし、アメリカ人被験者では正答率が全体的に日本人被験者よりも低く、恐れが特に低いという傾向も認められなかった。正答数について被験者(2)×感情(6)の2元配置の分散分析を行ったところ、交互作用が有意であり [$F(5, 100) = 5.624, p < .0005$], 日本人とアメリカ人では認知しやすい感情とにくい感情の種類が異なることが分かった。そこで被験者条件ごとに感情の効果について調べ

たところ、日本人被験者もアメリカ人被験者も6つの感情の正答率には差があり（日本人被験者: $F(5, 50) = 18.231, p < .000001$; アメリカ人被験者: $F(5, 50) = 3.449, p < .01$), 日本人被験者の場合には恐れのみが他の5つの感情よりも有意に低かった（TukeyのHSD検定による多重比較で $p < .0005$ ）が、アメリカ人被験者の場合には嫌悪が幸福や怒りよりも低かった ($p < .05$)。また、感情ごとに日米被験者の結果を比較したところ、嫌悪の場合のみ日本人被験者の方がアメリカ人被験者よりも正答率が高かった ($p < .05$)。

次に、6つの基本感情が日米の感情認知構造の中でどのような関係にあるのかを明らかにするため、多次元尺度構成法（MDS）によって結果を分析した。図2(a)はKruskalのMDSで分析した結果を2次元空間の中にプロットしたものである。ストレスは日本人被験者では2次元解の時0.089、アメリカ人被験者では0.0045であり、よく適合していた。感情の配置は日米被験者間で類似していたが、悲しみの位置のみ異なっていた。日本人被験者では「幸福－驚き－恐れ－悲しみ－怒り－嫌悪」の順にほぼ円環状に並ぶことが認められた。この並び方は表情写真についての先行研究と類似していた。ただし、悲しみの位置は米谷ら（1994）の結果では、日本人被験者は怒りと幸福の中間の表情として認知する傾向があったとしたが、本実験の結果では恐れと怒りの中間に位置していた。一方、アメリカ人被験者では6つの基本感情は円環からだいぶ外れていた。

アメリカ人話者条件（図1(b), 図2(b)）: 被験者条件と感情の効果の間に交互作用が認められ [$F(5, 100) = 2.742, p < .05$], 日本人とアメリカ人では認知しやすい感情とにくい感情の種類が異なることが分かった。しかしどの感情においても日米被験者間に差は認められなかった。また、日本人被験者のみ感情の種類により正答率に差があった [$F(5, 50) = 4.368, p < .005$] が、アメリカ人被験者はどの感情に対しても同等の高い正答率を示した。

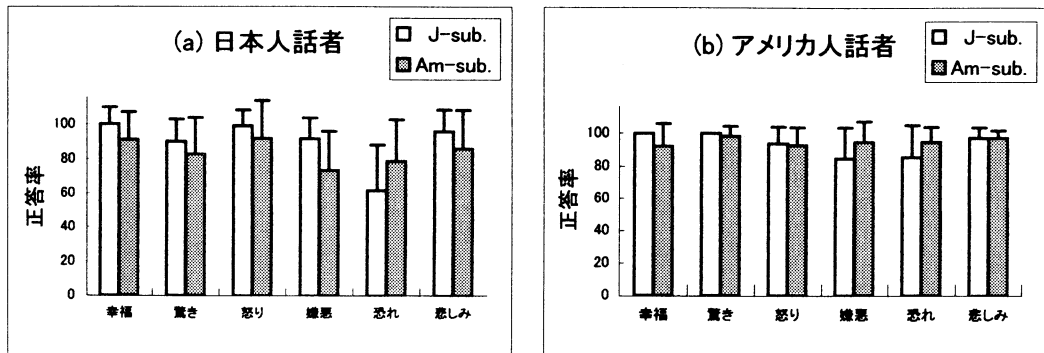


図1 表情認知の実験結果

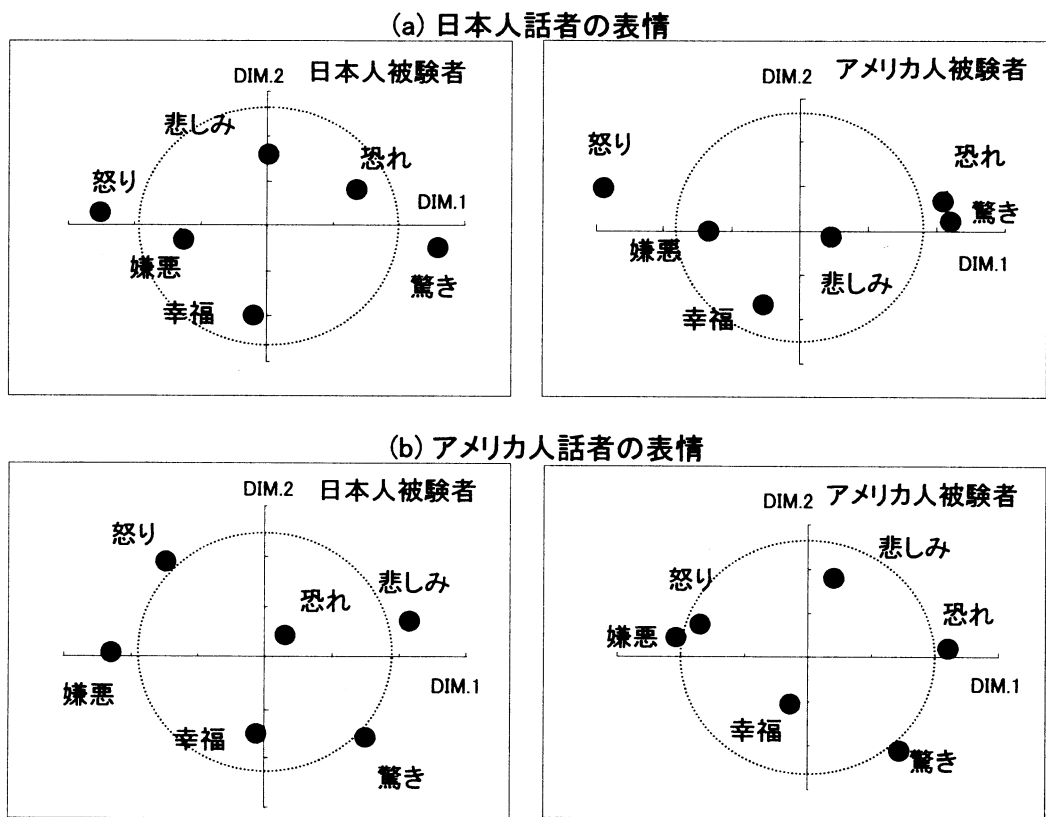


図2 表情認知の実験から求めた2次元心理空間における6つの基本感情

MDSにより結果を分析した(図2(b)). ストレスは日本人被験者では2次元解の時0.075, アメリカ人被験者では0.019であり, よく適合していた. 日米被験者間で類似した感情の配置が認められたが, 日本人話者の場合と同様に悲しみの位置が異なっていた. アメリカ人被験者の方が日本人被験者よりもより怒りに近く配置されていた.

以上の結果から, 日本人話者の場合もアメリカ人話者の場合も, 顔面表情の認知においては日米被験者間には悲しみを除いてほとんど差異が認められず, 動的な表情の認知においても文化を超えた普遍性のあることが考えられる.

実験2

音声による感情表現の場合について, 話者と聞き手の間の文化的差異によって感情認知が影響を受けるのかどうか, また受ける場合はどのような影響なのかについて検討する.

1. 方法

被験者 実験1に参加した日本人大学生11名とアメリカ人大学生11名が参加した.

刺激 刺激材料は実験1で用いた5つの語や短文を, 日米のプロの男性俳優が6つの基本感情を表現するようになるべく自然にかつ感情が明瞭に表れるように発話した音声であった. 俳優は2名とも実験1と同一であった. 各音声刺激は数回発話したものの中から, 実験者とプロデューサーがともに最も明瞭に感情が表出していると判断したものを一つ選んだ. 刺激の収録と編集は, 実験1と同様であった.

手続き 5個の刺激材料のそれぞれを6種類の感情を表すように音声によって表現したものを, 各刺激が2回ずつ提示されるようにランダムに提示した. 試行数は60試行であり, 一つの感情につき被験者一人当たり10回の判断を得た. 各音声の提示前1.5sに合図音として200msの純音を提示した. 反応時間はすべて3sとした.

刺激の持続時間は693-1254ms(日本語), 858-1650ms(米語)であった. 音声の提示レベルは, 68-75 dB Aであった.

被験者は, 各刺激について話者の感情を6つの基本感情の中から一つだけ選ぶように求められた. この際, 2つ以上の感情を同時に感じた場合には, 最も強く感じたものを答えるように指示した. その他の方法は, 実験1と同様であった.

2. 結果と考察

話者が意図した感情と同じ感情に被験者が同定した場合を正答とした. 日本人話者, アメリカ人話者ごとの平均正答率を図3に示す. 以下, 話者条件別に結果を検討する.

日本人話者条件(図3(a), 図4(a)): 日本人被験者では顔の表情の結果(実験1)ほどではなかったが, 正答率は全体的に高かった. しかし, やはり恐れの場合のみ非常に低かった. 一方, アメリカ人被験者では顔面表情の場合と同様に, 正答率が全体的に日本人被験者よりも低く, その程度は表情の場合よりも大きかった. しかし, 恐れが特に低いということはなかった. 被験者(2)×感情(6)の2元配置の分散分析を行ったところ, 交互作用が有意であり [$F(5, 100) = 11.558, p < .000001$], 音声による感情認知においても表情認知と同様に, 日本人とアメリカ人では認知しやすい感情としにくい感情があることが分かった. 被験者条件ごとに感情の効果を調べたところ, 日本人被験者もアメリカ人被験者も6つの感情の正答率には差があった(日本人被験者: $F(5, 50) = 17.210, p < .000001$; アメリカ人被験者: $F(5, 50) = 3.439, p < .01$). 日本人被験者の場合には恐れのみが幸福を除く他の4つの感情よりも有意に低かったが ($p < .0005$), アメリカ人被験者の場合には感情間で正答率には有意な差異はなかった. また, 感情ごとに日米被験者の結果を比較したところ, 驚きと恐れを除く他の4つの感情において日本人被験者の方がアメリカ人被験者よりも正答率が高かった ($p < .01$).

次にMDSによって分析し, 図4(a)のように

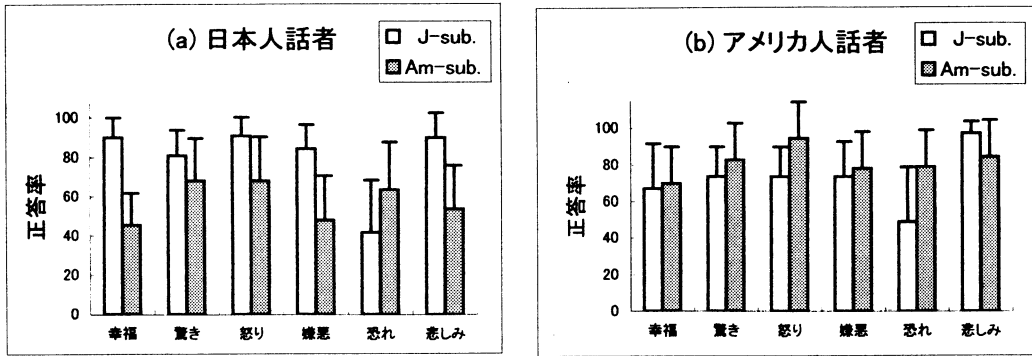


図3 音声による感情認知の実験結果

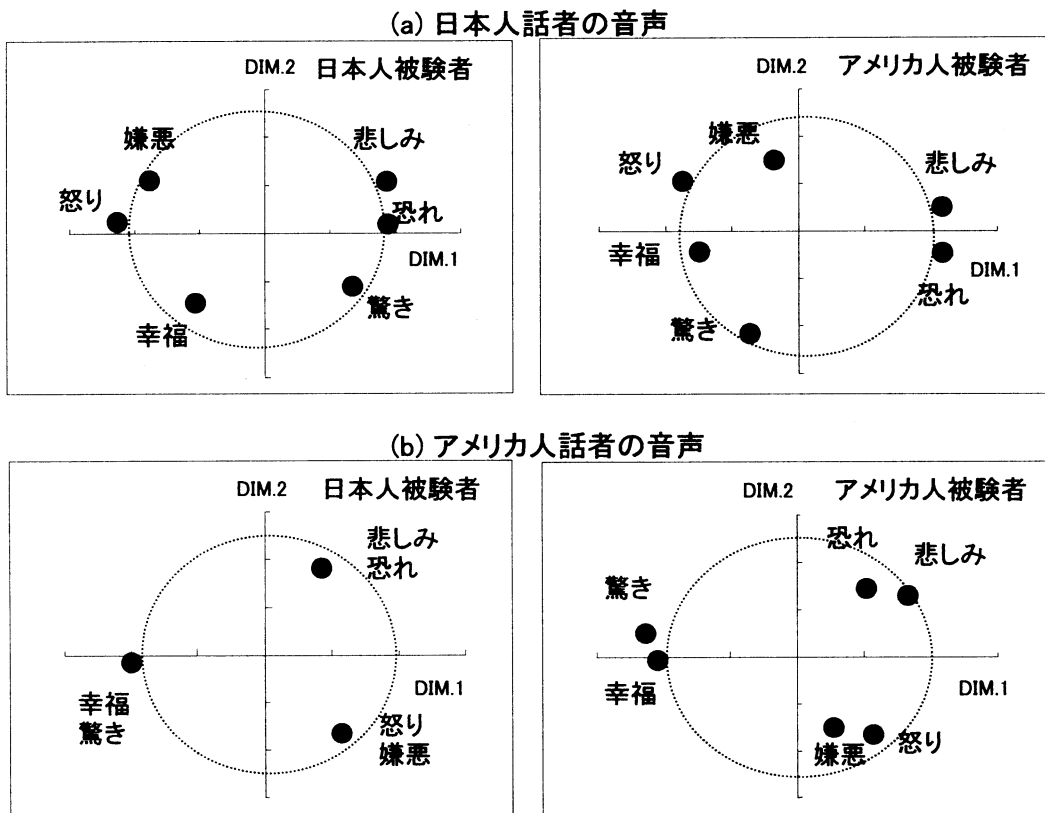


図4 音声の感情認知の実験から求めた2次元心理空間における6つの基本感情

6つの感情を2次元空間の中にプロットした。ストレスは日本人被験者では2次元解の時0.058, アメリカ人被験者では0.066であり, よく適合していた。感情の配置は日米被験者間でよく類似しており, 「幸福—驚き—恐れ—悲しみ—嫌悪—怒り」の順にほぼ円環状に並ぶことが認められた。これは, 表情写真についての先行研究と類似している。悲しみの位置は表情の場合とは異なり, 日米被験者とも恐れと嫌悪の中間で恐れに近いところにあった。

アメリカ人話者条件 (図3(b), 図4(b)): 日本人被験者では全体的に正答率は低く, 表情の結果 (実験1) と同様に恐れが特に低かった。一方, アメリカ人被験者では実験1ほどではなかったが, 正答率は日本人被験者よりも全体的に高かった。被験者条件と感情の効果の間には交互作用が有意に認められ [$F(5, 100) = 4.195, p < .005$], 日本人とアメリカ人では認知しやすい感情としにくい感情の種類が異なることが示された。日米被験者間に有意な差が認められたのは, 怒り [$F(1, 20) = 13.634, p < .005$], 恐れ [$F(1, 20) = 9.182, p < .01$], 悲しみ [$F(1, 20) = 5.414, p < .05$] の3つの感情だった。さらに, 日米被験者とも感情の種類によって正答率に差が認められた (日本人被験者: $F(5, 50) = 7.012, p < .00005$; アメリカ人被験者: $F(5, 50) = 3.456,$

$p < .01$)。日本人被験者では, 恐れが幸福を除く他の4つの感情よりも正答率が低かった ($p < .05$) のに対して, アメリカ人被験者では幸福が怒りよりも低かった ($p < .05$)。

MDSによる分析の結果 (図4(b)), 日米被験者間でよく類似した感情の配置が認められた。ストレスは日本人被験者では2次元解の時0.000004, アメリカ人被験者では0.003であり, よく適合していた。日米被験者とも6つの感情の配置は類似していた。表情の場合とは異なり, アメリカ人話者が音声により表現した6つの感情はほぼ3つのカテゴリーに分類された。それは, 幸福と驚き, 怒りと嫌悪, 恐れと悲しみの3つのカテゴリーであった。このことからアメリカ人話者が音声で表現する感情は日本人話者ほどには「きめ細かくない」ことが考えられる。本実験では話者は1名であったが, 同様の傾向は2名のアメリカ人話者について調べた実験結果においても認められている (重野, 2003)。

以上の実験結果は, 若干の差異はあるものの表情で表わされた感情表現に認められた結果と類似した傾向が音声による感情の認知にも認められることを示している。

感情認知における日米文化間の比較

図5は表情や音声によって表現された6つの

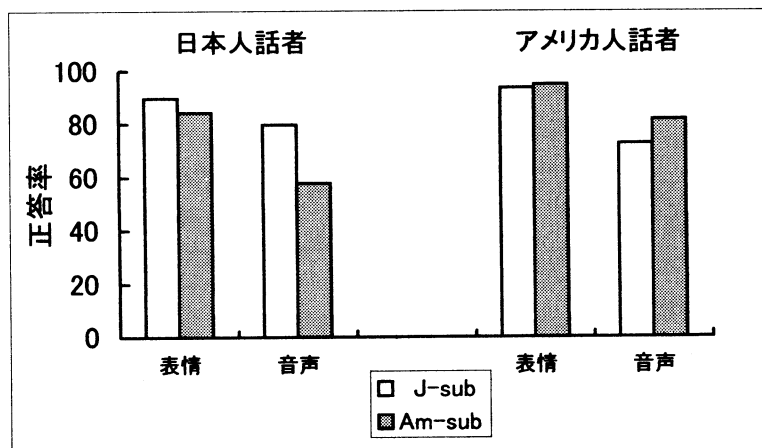


図5 表情と音声による感情認知の比較

基本的な感情を認知する場合、情報の送り手（話者）と受け手（被験者）が日米間で異なるときにどのような差異があるのかをグラフにまとめたものである。今回の実験では視覚の場合も動的な表情（ビデオによる表情運動）を用い、聴覚と同じように時間とともに変化する刺激を用いたが、提示方法により正答率には差があり、聴覚（音声）提示の場合よりも視覚（表情）提示の方が正答率は全般的に高かった。また、表情にせよ音声にせよ、送り手（話者）と受け手（認知する被験者）が同じ文化に属する場合（同文化条件）の方が異なる文化に属する場合（異文化条件）よりも、認知成績（正答率）は高かった。その差異は、表情の場合にはごくわずかであったが、音声の場合にはかなり開きがあった。そして開きの大きさは、話者が日本人である場合の方がアメリカ人である場合よりも大きかった。このことから日本人はアメリカ人よりも音声の感情表現により敏感であると考えられる。特に、日本人話者の音声に対しては、日本人被験者の正答率はかなり高く、日常生活において私たちがアメリカ人よりも音声を手がかりとして感情認知を行っていることを示唆している。

本研究から感情認知には情報の送り手や受け手の文化的条件だけではなく、両者の間の文化的差異（同異関係）がきわめて重要であることが示唆された。さらに、日米間を比較すると、日本人は感情表出が控えめではあるが、音声による感情表現にはより敏感であることが示唆された。

本研究は平成9年度サウンド技術振興財団研究助成および平成9-12年度科学研究費補助金特定領域研究(A)「心の発達：認知的成長の機構」の補助を受けた。

引用文献

- Ekman, P. 1985 *Telling lies*. New York: Norton.
Ekman, P., Friesen, W.V. 1971 Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, **17**, 124-129.
米谷 淳・瀧上凱令 1994 日米のTVドラマを

- 用いた表情識別実験 神戸大学国際文化学部 紀要「国際文化学研究」, **3**, 29-54.
松見淳子・Boucher, J.D. 1978 情動, 顔面表情および文化的差異について 心理学研究, **49**, 167-172.
Schlosberg, H. 1952 The description of facial expression in terms of two dimensions. *Journal of Experimental Psychology*, **44**, 229-237.
Schlosberg, H. 1954 Three dimensions of emotion. *Psychological Review*, **61**, 81-88.
重野 純 1998 音声の情動認知に及ぼす視覚情報の影響 日本音響学会聴覚研究会資料, H-98-8, 1-7.
重野 純 2003 米語話者が表現した感情音声の認知 日本心理学会第67回大会発表論文集, 認知, p. 703.