中国人と日本人の感情認知における 表情と声の相互干渉性

曹 蓮 杉森 伸吉 東京学芸大学 高 史明 神奈川大学

The mutual interference of facial and vocal information in Chinese and Japanese people's perception of emotions

Lian Cao, Shinkichi Sugimori (Tokyo Gakugei University), and Fumiaki Taka (Kanagawa University)

In this study, we investigated cultural differences in multisensory perception of emotion between Chinese and Japanese participants, focusing on mutual interference of visual and auditory emotional information. In this experiment, the face-voice pairs were consisted of congruent or incongruent emotions (e.g., a happy (an angry) face with a happy (an angry) voice in congruent pairs, and a happy (an angry) face with an angry (a happy) voice in incongruent pairs). Participants were asked to judge the emotion of targets focusing on either face or voice while ignoring the other modality's information. In the voice-focus condition, the effect of to-be-ignored facial information was smaller in Japanese than Chinese participants, only when the participant and the target belonged to the same cultures (in-group). This indicated that Japanese people were more likely to be based on the voice information in multisensory perception of emotion of in-group. Our study illuminated that although both Japanese and Chinese people belonged to the Eastern culture, there were cultural differences in perceiving emotion from visual and auditory cues.

Key words: multisensory perception of emotion, audio-visual information, mutual interference, cross-cultural comparison, facial and vocal expression.

The Japanese Journal of Psychology

2017, Vol. 88, No. 1, pp. 1–10

J-STAGE Advanced published date: January 14, 2017, doi.org/10.4992/jjpsy.88.15032

中国と日本は地理的・文化的に近いが、感情の表示規則や解釈規則の文化固有性が異文化理解の妨げになることもあるのではないだろうか。例えば、家族を事故で亡くした場合、中国人は泣き声や表情で強く悲しみを表出するが、日本人は泣き声や表情を抑えることが多い。この場合、中国人にとっては、日本人はあまり悲しんでいないように見えるし、日本人にとっては、中国人は大げさに悲しみの演技をしているように見える。また、日本人からは中国人同士の会話を聞くと怒っているように感じるという声をよく耳にするが、中国人からは日本人女性の声は感情が過剰で違和感を覚えるという声をよく耳にする。こうした異文化間の感情

認知においては、それぞれの文化に固有の表示規則や解釈規則が正確な異文化理解を阻害する壁となることが少なくない（大坊, 2007; 李・松本, 2011）。

これまでの異文化間感情認知の研究は、東洋と西洋という大きな枠組みにおける文化圏間の比較が中心であったうに、対象とするモダリティは表情または声のみを用いた単一モダリティがほとんどであった。しかし、中国と日本の間には感情表出や感情認知における文化差があり（王・米谷, 2008; 趙, 2002）、また、日常生活における感情認知はマルチモーダルな情報によってなされる。したがって、本研究では同じ東洋文化圏内に属する地域間の感情比較研究として、中国人と日本人の表情と声によるマルチモーダルな感情認知について比較を行う。さらに、近年、モダリティ間の感情情報の相互影響についても文化差が報告されている（Liu, Rigoulot, & Pell, 2015; Tanaka et al., 2010）。そこで、本研究ではモダリティ間の相互影響について検

Correspondence concerning this article should be sent to: Lian Cao, United Graduate School of Education, Tokyo Gakugei University, Nukuikitamachi, Koganei 184-8501, Japan. (E-mail: sourenn09@yahoo.co.jp)

討するために、表情あるいは声の一方のみに基づいて感情を判断する際に、無視すべき他方のモダリティから生じる干渉の程度を測定する研究手法（以下、干渉パラダイムとする）を用いて実験的に検討する。それによって、視聴覚情報による自動的かつ抑制困難な感情認知における文化差を検討する。

感情認知の文化差

感情認知および表出の文化間比較は、研究の初期には主に表情からの感情情報について行われ、それらが文化によらず普遍的であることが報告されてきた (Ekman, 1972; Ekman & Friesen, 1971; Ekman & Friesen, 1986; Izard, 1971; Matsumoto & Ekman, 1989)。しかし、研究が進むにつれ、感情の表示規則や解釈規則には文化固有性もあることが示されてきた (Ekman & Friesen, 1975 工藤訳 1987; Gendron, Roberson, van der Vyver, & Barrett, 2014; Ishii, Reyes, & Kitayama, 2003; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Yuki, Maddux, & Masuda, 2007)。すなわち、どのような基本的な感情の種類を有するのか（例えば、喜びや怒り、嫌悪など）、また、それらがどのような表情表出に対応するのか（例えば、喜びの表情は目の周りの筋肉が収縮され、口角が引き上げられる）はかなりの程度で文化普遍的であるが、それらをいつ、どのように表出・認知するのか（表示・解釈規則）は、文化固有的である (Ekman & Friesen, 1975 工藤訳 1987)。なぜなら、人々は文化の価値基準や社会的な意味で適切か否かという基軸に基づいて、社会的場面に応じて感情表出を管理したり、感情認知を調整したりするからである (マツモト・工藤, 1996)。

このような文化固有の表示規則や解釈規則により、感情を認知する際に依存する情報源にも文化差があることが報告されている (Ishii et al., 2003; Masuda et al., 2008; Yuki et al., 2007)。Yuki et al. (2007) は、表情から感情を認知する際に、日本人はアメリカ人より、相手の目から得られる情報に依存すると報告している。そして、Ishii et al. (2003) は、声から感情を認知する際に、日本人はアメリカ人に比べ、イントネーションやリズムなどのパラ言語（ノンバーバル）情報に、アメリカ人は日本人より言語の意味（バーバル）情報に依存する傾向があると報告している。

こうした感情認知の文化差に関する研究は、日本とアメリカのような東西文化圏間の比較が中心であったうえに、視覚のみ、または聴覚のみの単一モダリティ内からの情報を用いた比較研究がほとんどであった。しかし、人間の感情表出や認知はマルチモーダルな情報チャンネルを通して行われる。したがって、文化比較研究においても、モダリティ間の感情情報、およびその相互作用を検討することが有益であると考えられる。そこで次に、複数の情報チャンネルから感情情報

を処理する際の文化差に関する研究を概観する。

マルチモーダルな感情認知における文化差

近年、マルチモーダルな感情処理も文化によって異なることが、正答率を用いた行動的指標と、脳波の変化を測定した生理的指標の両側面で検討されている (Liu et al., 2015; Tanaka et al., 2010)。Tanaka et al. (2010) は視聴覚感情認知におけるモダリティ間の相互干渉性に関する文化差を検討するために、日本人とオランダ人を対象に、表情と声の感情が一致する場合と不一致である場合の動画刺激を用いて、干渉パラダイム実験を行った。その結果、日本人はオランダ人に比べ、声による感情判断時には表情からの干渉が弱く、表情による感情判断時には、声からの干渉が強いことから、一貫して声への依存性が高いという文化差が示された。Liu et al. (2015) は中国人と北米人を対象に、同様の実験パラダイムを用いて、正答率と脳波の変化を測定し比較検討を行った。その結果、中国人は北米人に比べ、声による感情判断に与える表情からの影響が有意に弱かった。これらの先行研究によって、感情認知における表情と声による多感覚感情情報の重み付けは文化によって異なることが示された。

上記のような視聴覚感情認知における文化間の比較研究では、東洋人である中国人と日本人は、西洋人である北米人やオランダ人と比べ、声による感情情報に依存しやすいことが示されている (Liu et al., 2015; Tanaka et al., 2010)。しかし、このような文化差は東洋と西洋という文化圏の間だけでなく、同じ東洋文化圏内に属する地域間にも存在している可能性がある。そして、視聴覚感情情報の相互干渉性に焦点を当てた中日間の実証的比較研究はまだ行われていない。そこで、本研究では視聴覚感情認知におけるモダリティ間情報の相互干渉性について、干渉パラダイムを用いて中日間の比較検討を行う。

まず、中国人と日本人の表情と声による感情認知にはどのような文化差があるのかについて仮説を導出することを試みるために、両者の感情の表示規則について表情表出と言語的特徴の両側面に沿って数少ない研究を概観する。

中国人と日本人の感情表出の違い

マツモト・工藤 (1996) は、実際に感じている感情と顔で表出する感情が食い違うのは、日本人の大きな特徴とも言えると述べている。Ekman (1972) の研究では、日本人は自分の感じている極めて否定的な感情を隠すために、地位の高い人（実験者）がいるところでは微笑みを表出しやすいことが観察された。日本人の微笑み表情は幸福や友好的な感情を表すだけでなく、より複雑な意味や機能を担っている。そこで、日本人は表情表出の二面性を学習することで、表情によ

る感情表出をあまり信頼せず、表情に比べて恣意的操作がより困難な声からの感情情報に重みをおき、相手の本当の感情を認知しようとする傾向があるのではないかと考えられる。

一方、趙 (2002) は質問紙調査法を用いて感情表出の自己認識 (例えば、「あなたは公的な場面で、あるいは仕事に、怒りを感じている時、その感情がどの程度顔に現れますか」といった質問項目) を尋ね、中日間の比較を行った。6つの基本感情 (喜び、悲しみ、怒り、恐れ、驚き、嫌悪) について私的状況と公的状況に分けて分析した結果、私的状況では中日間の違いは見られなかったが、公的状況においては、日本人は中国人より喜びの感情を、中国人は日本人より怒りの感情を表出しやすいという自己認識を持っていることが示された。これらはあくまでも回答者の主観的な自己報告に基づくものではあるが、日常生活において中国人と日本人は異なる感情表出を経験していることを反映しているかもしれない。このことから、日本人に比べ、中国人は肯定的な感情表出に偏るのではなく、怒りなどのネガティブな感情も用いて様々な感情を伝達すると考えられる。

また、視聴覚感情認知は表情による表示規則だけでなく、言語的要因による影響も受けると考えられる。日本とアメリカの児童による感情発話をういて音響分析を行った櫻庭・今泉・笈 (2004) は、アメリカ人児童の発話では感情によって各音節の時間構成比にバラツキが見られるのに対し、日本人児童の発話ではそうした変化が見られずほぼ一定である (例えば、「ピカチュウ」に感情を込めても平静発話の際の1:1:2の時間構成比はほとんど変化しない) ことを示した。これをもとに、「各モーラ間に等時性を保つ」という日本語の言語固有の特性は感情発話にも反映されていると議論している。つまり、感情発話には、感情情報と言語固有の特徴が加算的に現れる (櫻庭他, 2004)。同様に、中国人の声による感情表現にも、中国語の言語固有の特性が含まれていると考えられる。

中国語の言語固有の特性の代表的なものとしては、「四声」が挙げられる。中国語は「声調言語」とも呼ばれ、「四声」という音節ごとの固有のアクセントによって語の意味が異なる音韻的規則がある。もちろん日本語にも「雨」と「鉛」のようにアクセント変化によって意味変化をもたらす単語はあるものの、中国語ほど意味認知に声のアクセント情報を用いない。柴田・柴田 (1990) の研究によると、同音語の中で、アクセントで区別できる単語の割合は、中国語は71.00%であるのに対し、日本語はわずか13.57%であった。このように日本語に比べ、中国語のアクセントは意味認知において担う役割が多いため、感情認知において担う役割が相対的に少なく、その役割を果たすのも比較的に困難になるのではないかと考えられる。例えば、

Chang (1958) は、中国語と英語の「軽蔑」の感情音声の音響的特徴が異なるのは、「四声」のような言語の音韻的制約によると示唆している。したがって、中国語の声による感情認知は、日本語の声に比べ、困難ではないかと考えられる。

上述した観点から視聴覚感情認知における中国人と日本人の文化差をまとめると、以下の仮説が考えられる。(a) 視聴覚感情認知において、中国人は日本人より表情からの、日本人は中国人より声からの感情情報に、より敏感に反応する可能性がある。(b) 中国語は日本語より、声に基づいた感情認知が困難である可能性がある。

本研究では上記の仮説を検討するために、表情と声による視聴覚感情情報の相互干渉性に焦点を当て、視聴覚情報による自動的かつ抑制困難な感情認知について、中日間の比較実験を行うこととした。

方 法

参加者

日本での滞在経験や日本語の学習経験がない中国在住の中国人大学生21名、うち男性6名 (年齢: $M = 18.2$, $SD = 0.41$), 女性15名 (年齢: $M = 20.1$, $SD = 1.25$), 中国での滞在経験がなく、中国語の学習経験がほとんどない日本在住の日本人大学生27名、うち男性11名 (年齢: $M = 20.5$, $SD = 1.69$), 女性16名 (年齢: $M = 19.1$, $SD = 0.70$) が実験に参加した。プログラムがエラーで終了した日本人1名のデータは分析から除外された。

刺 激

まず、刺激収集を行った。中国人と日本人の表出者各10名 (男女各5名)、計20名の「喜び」あるいは「怒り」の感情が込められた表情写真と声を収集した。声刺激は重野 (2004) と Tanaka et al. (2010) を参考に、感情的に中立な意味の単語および短文を用い、表出者が各自の母語で喜びもしくは怒りの感情を込めて読み上げたものを録音した。中国人表出者には“bei jing (北京)” “shi yi dian (11時)” “xing qi yi (月曜日)” “zhe shi shen me (これはなに)” “shi ma (そうですか)” を、日本人表出者には「北京」を「東京」にした以外は同様の単語および短文を日本語で読み上げてもらった。

刺激収集はまず表情、次に声の順に実施した。

まず、表情刺激の撮影は、喜びの表情は嬉しかった出来事を、怒りの表情は腹立たしかった出来事を思い出して表情を作るように教示した上で行った。怒りの感情表出は Ekman, Friesen, & Hager (2002) を基に作成された「基本表情とそれに対応する顔面筋動作のリスト」(渡邊他, 2007) を参考にし、「両眉を中央に寄せ、目を見開き、瞼に力を入れる」などの参考情報を与え

た。撮影の前に、全ての表出者は鏡の前で練習をし、準備ができたと本人が判断してから撮影を実施した。撮影は、目的とする感情が十分に表情に込められていると表出者と実験者が共に判断するまで繰り返し行われた。撮影後、表情刺激に対する感情判断が表情以外の要素（髪型や服装の一部、顔の輪郭など）によって影響されないようにするために、表情画像は Adobe Photoshop Elements 12 (Adobe Systems Inc., 2013) を用いて、顔の輪郭線に沿って約 295×443 pixel の大きさに切り出し、背景色を RGB128 の灰色に設定した。

次に、声刺激の録音は、全表出者に発話の背景文脈を教示して感情的誘導を行い、目的とする感情を込めた言葉を実験者に向かって発話する方法で実施した。声刺激の録音も表情刺激の撮影と同様に、目的とする感情が声に十分に込められていると表出者と実験者が共に判断するまで繰り返し行われた。録音した声刺激は Audacity 2.0.5.0 (Audacity Team, 2013) を用いて単語の長さに合わせた長さに切り出した。声刺激はサンプリング周波数 44,100 Hz, 16 bit, モノラルであった。

上記の方法で収集した表情および声刺激の中から、表出者と実験者が共同して、「喜び」あるいは「怒り」の感情を最もよく表出していると感じられるものを選出した。さらに、選出した刺激の中から感情が十分に表出されている刺激を精選するために、中国人大学生 11 名、日本人大学生 16 名を対象に予備実験を行った。表情 (1,000 ms)、声 (600 ms – 1,400 ms) をそれぞれ単独に呈示し、「喜び」または「怒り」の感情がどのくらい含まれているのかについて「0: 全く感じられない」–「10: 非常に感じられる」の 11 件法で判断してもらい、各刺激に対する感情強度の評価得点を測定した。感情強度の評価得点が 5 点以上の刺激の中から、中国人参加者と日本人参加者の評価得点が類似する刺激を選出した。

その結果、中国人と日本人それぞれについて 4 名ずつ（男女各 2 名ずつ）、計 8 名の表出者について、1 名につき表情刺激と声刺激をそれぞれ 2 つずつ（喜びと怒りを 1 つずつ）選出した。すなわち、表情刺激と声刺激をそれぞれ 16 個ずつ選出した。

言語や感情の種類によって声刺激の長さに差があるのかについて t 検定を用いて検討した結果、有意差はなかった（中国語の声 ($M = 823.00, SD = 230.97$) と日本語の声 ($M = 955.00, SD = 94.39$) では $t(14) = 1.50, ns, r = .44$ 、喜び声 ($M = 876.38, SD = 192.30$) と怒り声 ($M = 901.63, SD = 186.74$) では $t(13.99) = 0.27, ns, r = .07$ ）。

以上の刺激を用いて、表情と声を同時に呈示する視聴覚刺激を作成した。視聴覚刺激は表情と声からの感情が相互一致する「一致条件」刺激（喜び表情 + 喜び声；怒り表情 + 怒り声）と、相互矛盾する「不一致条件」刺激（喜び表情 + 怒り声；怒り表情 + 喜び声）

であった。表情と声の組み合わせは同一人物の表情と声に限定せず、同国籍内の同性間でランダムに組み合わせ、計 64 個の視聴覚刺激を作成した。各視聴覚刺激の持続時間は組み合わせた声の持続時間と同様であった。

装置

実験は全て同一のノートパソコン (NEC, 1280×800 pixels, 32 bit, 12.1 inch) を利用して行われた。刺激呈示ソフトは SuperLab4.0 (Cedrus, 2009) を用いた。声刺激の呈示にはイヤホン (Apple Earphones with Remote and Mic MB770G/B) を利用し、音量は各参加者が大きすぎないと報告した程度に設定した。

手続き

本研究では計 2 つの課題を一人ずつ実験室で実施した。

まず、課題 1 として「多感覚課題」を行った。表情と声による感情情報の相互干渉性を測定し、視聴覚感情認知における中国人と日本人の文化差を検討することが目的であった。視聴覚刺激を用いて、声は無視して表情のみに基づいて感情を認知する「表情セッション」と、表情は無視して声のみに基づいて感情を認知する「声セッション」を参加者間でカウンターバランスをとって実施した。各セッション内では視聴覚刺激を 1 回ずつランダムに呈示し、計 128 試行 (2 (セッション: 表情・声) $\times$ 64 個 (視聴覚刺激)) を行った。各セッションにおいて、視聴覚刺激が一致した場合に比べ、不一致である場合にエラー率が高まる程度を、表情と声による感情情報の相互干渉性を測定する指標とした。

各試行の流れとして、まず、初期画面として、表情刺激の残像や声刺激の残響による影響を除外するために、ランダムドット刺激の中心に注視点を呈示すると同時に、300 Hz の純音 (“pi-pi-pi-pi”) を 2,000 ms 呈示した。次の視聴覚刺激の呈示が終了すると、判断画面を呈示し、注目するように教示されたモダリティの刺激が表す感情が「喜び」か「怒り」かについて 2 肢強制選択法で判断を求めた。回答が終了すると、即座に ISI が呈示され、次の試行が開始された。判断はできる限り正確に、かつ速く行うように教示し、エラー率と反応時間を測定した。また、利き手やキーの押しやすさなどの影響を除外するため、反応キーの「喜び」と「怒り」への割り当ては参加者間でカウンターバランスを取った。

課題 2 として「単感覚課題」を実施した。表情刺激と声刺激を単独で呈示した場合の判断難易度について確認することが目的であった。表情 (1,000 ms) もしくは声をそれぞれ単独で呈示し、どのような感情が感じられるのかを「喜び」と「怒り」の 2 肢強制選択法

Table 1
視聴覚感情認知における中国人と日本人の正答率および標準偏差

	表情セッション							
	中国人の怒り顔		中国人の喜び顔		日本人の怒り顔		日本人の喜び顔	
	怒り声	喜び声	怒り声	喜び声	怒り声	喜び声	怒り声	喜び声
中国人参加者	100.00 (0.00)	98.21 (4.48)	99.40 (2.73)	97.62 (5.03)	99.40 (2.73)	98.21 (4.48)	100.00 (0.00)	99.40 (2.73)
日本人参加者	99.52 (2.45)	99.52 (2.45)	97.60 (6.14)	99.04 (3.40)	99.52 (2.45)	97.60 (6.14)	97.60 (5.02)	99.04 (3.40)
	声セッション							
	中国人の怒り声		中国人の喜び声		日本人の怒り声		日本人の喜び声	
	怒り顔	喜び顔	怒り顔	喜び顔	怒り顔	喜び顔	怒り顔	喜び顔
中国人参加者	97.62 (5.03)	89.29 (15.45)	87.50 (16.30)	98.21 (5.98)	95.83 (9.13)	86.31 (21.25)	88.69 (13.05)	96.43 (7.01)
日本人参加者	95.19 (9.41)	84.62 (17.07)	77.40 (22.64)	93.75 (9.52)	99.52 (2.45)	99.04 (3.40)	99.52 (2.45)	99.04 (3.40)

注) () 内の数字は標準偏差を示す。

で判断する「表情のみセッション」と「声のみセッション」を行い、正答率を測定した。それ以外の手続きは課題1と同様であった。

結 果

視聴覚感情情報の相互干渉性を検討する指標として、表情と声の感情が一致する場合に比べ、不一致である場合にエラー率が高まる程度を分析した。すなわち「不一致条件」のエラー率から「一致条件」のエラー率を引いて補正したエラー率（以下、補正後エラー率とする）を、無視すべきモダリティから生じた干渉量の指標とした。

視聴覚感情認知における正答率はTable 1に示している。

視聴覚感情情報の相互干渉性における中日間の違い

感情認知は文化ごとの表示規則に対する学習経験や習慣に基づいて行われる (Engelmann & Pogosyan, 2013)。したがって、感情認知について検討する際には、表出者と認知者が属する文化の異同も考慮すべきである。本研究では、表出者側と認知者側が属する文化が同じである場合を「内集団」、異なる場合を「外集団」とし、それぞれの表情セッションと声セッションにおける中国人と日本人の感情認知の文化差を、以下の分析を通して検討した。

2 (参加者国籍：中国、日本) × 2 (セッション：表情、声) × 2 (集団：内、外) を独立変数、補正後エラー率を従属変数とした参加者間1要因と参加者内2要因の3要因混合計画分散分析を行った (Figure 1)。その結果、セッションの主効果 ($F(1, 45) = 21.89, p < .01, \eta^2 = .33$) と集団の主効果 ($F(1, 45) = 8.47, p$

$< .01, \eta^2 = .16$) が有意であった。参加者国籍の主効果は有意ではなかった ($F(1, 45) = 0.21, ns, \eta^2 = .01$)。さらに、参加者国籍 × セッション × 集団の相互作用 ($F(1, 45) = 15.93, p < .01, \eta^2 = .26$) が有意であったため下位検定を行った。

集団ごとの参加者国籍とセッションの効果を検討するために2要因分散分析を行った。まず、外集団においては、セッションの主効果のみが有意であり ($F(1, 45) = 24.60, p < .01, \eta^2 = .35$)、参加者国籍の主効果 ($F(1, 45) = 1.50, ns, \eta^2 = .03$) および参加者国籍 × セッションの交互作用 ($F(1, 45) = 1.08, ns, \eta^2 = .02$) は有意ではなかった。この結果より、外集団に属する表出者の表情や声に対する感情認知には中国人と日本人の間で文化差がないことが分かった。

次に、内集団においては、参加者国籍の主効果 ($F(1,$

Figure 1. 視聴覚感情認知における中国人と日本人の補正後エラー率。

注) 補正後エラー率は不一致条件のエラー率から一致条件のエラー率を引いたもの (エラーバーは標準誤差)。

45) = 9.75, $p < .01$, $\eta^2 = .18$), セッションの主効果 ($F(1, 45) = 7.07$, $p < .05$, $\eta^2 = .14$) が有意であった。そして、参加者国籍 × セッションの交互作用 ($F(1, 45) = 14.45$, $p < .01$, $\eta^2 = .24$) も有意であったため下位検定を行った。

まず、内集団の各セッションにおける参加者国籍の効果を検討するために t 検定を実施した。その結果、内集団の表情セッションにおいて、中国人参加者 ($M = 0.00$, $SD = 4.42$) と日本人参加者 ($M = 1.68$, $SD = 3.77$) の間に有意差はなかった ($t(45) = 1.41$, ns , $r = .21$)。

一方、内集団の声セッションにおいて中国人参加者 ($M = 9.52$, $SD = 12.12$) と日本人参加者 ($M = 0.00$, $SD = 3.06$) の間には有意差が見られた ($t(22.07) = 3.51$, $p < .01$, $r = .60$)。この結果より、内集団に属する表出者の声に対する中国人と日本人の感情認知には文化差があることが分かった。具体的には、日本人が日本人の声に基づいて感情を認知する場合に比べ、中国人が中国人の声に基づいて感情を認知する場合の方が、表情からの干渉を受けやすく、声に基づいて感情を正確に認知するのが困難であることが分かった。

以上の結果から、表出者側と認知者側の文化が同じである場合 (内集団) に、日本人は中国人に比べ、声による感情認知に与える表情からの干渉が顕著に低く、声による感情情報に依存しやすいことが分かった。

他の要因による影響の確認

上記のように、内集団の声による感情認知に与える表情からの干渉において、中日間の文化差が示された。しかし、この結果は、他の要因によるものではなかったのかを確認する必要がある。したがって、以下は反応時間、表情と声の刺激の判断難易度において中国人と日本人の間に違いがなかったのかを確認した。

反応時間と正答率のトレード・オフ 上記は中国人と日本人の視聴覚感情情報の相互干渉性を検討するために、補正後エラー率を指標として用いた。しかし、感情認知を正確に、かつ速く行うように教示をしたため、反応時間と正答率のトレード・オフが生じうる。本研究では表情と声と同時に終了するまで判断できないように設定されているため、中国人と日本人のどちらの方がより反応が早かったのかについては検討できないが、判断画面が呈示されてから判断終了までの時間的長さに大幅な遅れはなかったのかを確かめる必要がある。例えば、(多感覚課題の) 内集団の声課題において、日本人が中国人より補正後エラー率が有意に低かったのは、中国人と比べて反応時間が遅かったからではないのかを確認する必要がある。したがって、反応時間も考慮した指標として反応時間を正答率で割った値 (Inverse Efficiency: 以下、IE とする) を用いた。IE は反応時間と正答率を合わせたパフォーマンス

に関する指標であり、値が大きいほどパフォーマンスが悪いことを示す (Spence, Pavani, & Driver, 2004)。視聴覚相互干渉性によって反応時間と正答率を合わせたパフォーマンスがどのように変化するかを検討するために、表情と声の感情が一致する場合の IE と比べ、不一致である場合の IE が高まる程度を分析した。すなわち、「不一致条件」の IE から「一致条件」の IE を引いて補正した IE (以下、補正後 IE とする) を、無視すべきモダリティから生じた干渉量の指標とした。

反応時間の外れ値処理については条件ごとの「平均値 $\pm 3SD$ 」を基準値とし、これを超えた値は全て外れ値とみなし、基準値に置き換えた。

2 (参加者国籍: 中国, 日本) $\times$ 2 (セッション: 表情, 声) $\times$ 2 (集団: 内, 外) を独立変数、補正後 IE を従属変数にした参加者間 1 要因、参加者内 2 要因の 3 要因混合計画分散分析を行った結果、セッションの主効果が有意であった ($F(1, 45) = 19.94$, $p < .01$, $\eta^2 = .31$)。そして、参加者国籍 $\times$ セッション $\times$ 集団の交互作用が有意であった ($F(1, 45) = 7.22$, $p < .01$, $\eta^2 = .14$) ため、下位検定を行った (Figure 2)。主に、内集団の各課題における中国人と日本人の文化差を検討するために t 検定を実施した結果、補正後エラー率を指標とした分析結果と同様であった。具体的には、内集団の表情セッションにおいては、中国人参加者 ($M = 0.20$, $SD = 0.53$) と日本人参加者 ($M = 0.12$, $SD = 0.56$) の間に有意差はなかった ($t(45) = 0.48$, ns , $r = .07$)。一方、内集団の声セッションにおいては、中国人参加者 ($M = 0.88$, $SD = 1.06$) が日本人参加者 ($M = -0.03$, $SD = 0.36$) より補正後 IE が有意に高かった ($t(23.72) = 3.77$, $p < .01$, $r = .61$)。つまり、内集団の声による感情認知において、日本人は中国人に比べ、パフォーマンスが顕著に高かった。

以上の分析結果より、内集団の声による感情認知に

Figure 2. 視聴覚感情認知における中国人と日本人の補正後 Inverse Efficiency (IE)。

注) 補正後 IE は不一致条件の IE から一致条件の IE を引いたもの (エラーバーは標準誤差)。

与える表情からの干渉において見られた中日間の違いは反応時間と正答率のトレード・オフでは説明できないことが確認できた。

刺激の判断難易度による影響の確認 刺激自体の判断難易度に違いがあり、それが「多感覚課題」の結果に影響したのではないのかを検討するために、課題2の「単感覚課題」のデータを用いて以下の分析を行った。

2 (参加者国籍：中国、日本) × 2 (セッション：表情のみ、声のみ) × 2 (集団：内、外) を独立変数、正答率を従属変数とした参加者間1要因と参加者内2要因の3要因混合計画分散分析を行った (Figure 3)。この結果、セッションの主効果 ($F(1, 46) = 28.51, p < .01, \eta^2 = .38$) と集団の主効果 ($F(1, 46) = 8.70, p < .01, \eta^2 = .16$) が有意であった。参加者国籍の主効果 ($F(1, 46) = 0.30, ns, \eta^2 = .01$) は有意ではなかった。そして、セッション × 集団の交互作用 ($F(1, 46) = 8.70, p < .01, \eta^2 = .16$) が有意であったため下位検定を行った結果、内・外集団にかかわらず、声のみ刺激 ($M = 98.44, SD = 4.18; M = 94.01, SD = 8.15$) より表情のみ刺激 ($M = 99.74, SD = 1.80; M = 99.74, SD = 1.80$) による感情認知のほうが正しく判断しやすいことが分かった ($t(47) = 2.34, p < .05, r = .32; t(47) = 4.65, p < .01, r = .56$)。そして、外集団の声より、内集団の声に基づいて感情を判断する場合に正答率が高く、感情を判断しやすいことがわかった ($t(47) = 3.13, p < .01, r = .42$)。

これ以外の参加者国籍を含んだいずれの効果も、有意ではなかった ($F(1, 46) < 2.81, ns, \eta^2 < .06$)。

これらの分析結果から、特に内集団では表情刺激においても声刺激においても、中国人参加者と日本人参加者の間の違いは見られなかった。したがって、「多感覚課題」の内集団の声セッションにおいて、中国人参加者と日本人参加者は判断難易度が同程度に高い声刺激に基づいて感情を判断し、その際に提示された中国人と日本人の表情刺激も同程度の判断難易度であっ

Figure 3. 表情のみ・声のみの感情認知における中国人と日本人の正答率 (エラーバーは標準誤差)。

たことが確認できた。

以上の結果より、「多感覚課題」で内集団の声による感情認知に与える表情からの干渉において見られた中国人と日本人の文化差は、刺激自体の判断難易度 (例えば、日本人の表情刺激自体が中国人の表情刺激より判断難易度が高く、声に対する干渉が生じにくかったなど) による影響では説明できないと言える。

声からの感情認知に与える言語的違いによる影響

次に、声からの感情認知に言語の違いによる影響があるのかについて検討する。感情発話には感情情報の特徴だけでなく、言語固有の特徴が加算的に現れる (櫻庭他, 2004)。そのため、声による感情認知において、中国語と日本語の言語固有の特徴による効果があったのかについても検討する必要がある。したがって、以下は言語ごとの中国人と日本人の感情認知について検討した。

視聴覚感情認知における言語的違いによる影響 言語的特徴の違いが声による感情認知に影響を与えるのかを検討するために、中国語と日本語の声に基づいて感情を認知する際の表情からの干渉程度に焦点を当て、中国人と日本人の違いを検討する。そのために多感覚課題における声セッションのみのデータを用いて以下の分析を実施した。

2 (参加者国籍：中国、日本) × 2 (言語：中国語、日本語) を独立変数、補正後エラー率を従属変数にした2要因分散分析を行った。その結果、参加者国籍の主効果は有意ではなかった ($F(1, 45) = 0.68, ns, \eta^2 = .02$) が、言語の主効果 ($F(1, 45) = 13.79, p < .01, \eta^2 = .24$) は有意であった。そして、参加者国籍 × 言語の交互作用 ($F(1, 45) = 10.57, p < .01, \eta^2 = .19$) が有意であったため、下位検定を行った (Figure 4)。言語ごとの中国人と日本人の違いについて検討した結果、日本語の声による感情認知においては、日本人参加者 ($M = 0.00, SD = 3.06$) は中国人参加者 ($M = 8.63, SD = 15.37$) より、表情からの干渉が有意に低かった ($t(21.29) = 2.53, p < .01, r = .48$)。一方、中国語の声に

Figure 4. 中国語と日本語の声による感情認知に与える表情からの干渉性における中国人と日本人の補正後エラー率。注) 補正後エラー率は不一致条件のエラー率から一致条件のエラー率を引いたもの (エラーバーは標準誤差)。

よる感情認知においては、中国人参加者 ($M = 9.52$, $SD = 12.12$) と日本人参加者 ($M = 13.46$, $SD = 13.43$) の間に有意差はなく ($t(45) = 1.04$, ns , $r = .15$), 中国人も日本人と同じく表情による干渉を受けやすいことが示された。

上記のように、中国語の声による感情認知は、日本人だけでなく、中国語を母国語とする中国人においても、表情からの干渉を受けやすいことから、中国語の声による感情認知は日本語に比べ、より困難であると考えられる。

聴覚のみの感情認知における言語的違いによる影響

声のみによる感情認知においても言語的違いによる影響は生じうるのかを検討するために、単感覚課題における声のみセッションのデータを用いて以下の分析を行った。

2 (参加者国籍: 中国, 日本) $\times$ 2 (言語: 中国語, 日本語) を独立変数, 正答率を従属変数にした 2 要因分散分析を実施した。その結果, 参加者国籍の主効果 ($F(1, 46) = 0.02$, ns , $\eta_p^2 = .00$), 言語の主効果 ($F(1, 46) = 2.81$, ns , $\eta_p^2 = .06$) は有意ではなかった。参加者国籍 $\times$ 言語の交互作用 ($F(1, 46) = 8.70$, $p < .01$, $\eta_p^2 = .16$) が有意であったため下位検定を行った (Figure 5)。言語ごとの参加者国籍の効果を検討するために 1 要因分散分析を行った。その結果, 日本語の声において, 日本人参加者 ($M = 99.54$, $SD = 2.41$) は, 中国人参加者 ($M = 95.24$, $SD = 6.22$) より, 正答率が有意に高かった ($F(1, 46) = 10.87$, $p < .01$, $\eta_p^2 = .19$)。一方, 中国語の声においては, 中国人参加者 ($M = 97.02$, $SD = 5.46$) と日本人参加者 ($M = 93.06$, $SD = 9.39$) の間には有意差がなかった ($F(1, 46) = 2.96$, ns , $\eta_p^2 = .06$)。

上記の結果より, 日本語の声による感情認知において, 日本人は中国人に比べて正答率が顕著に高かったが, 中国語の声による感情認知においては, 中国人にとっては母国語であるにもかかわらず, 日本人と比べて顕著に高い結果は示されなかった。この結果から, 中国語の声による感情認知は日本語に比べ, より困難である可能性があると考えられる。

Figure 5. 中国語と日本語の声による感情認知における中国人と日本人の正答率 (エラーバーは標準誤差)。

考 察

本研究は, 中国人と日本人の視聴覚感情認知における文化差を検討するために, 表情と声による感情情報の相互干渉性に焦点を当てて比較実験を行った。

まず, 表情による感情認知に与える声からの干渉においては, 中日間の文化差は示されなかった。その理由として, 表情に対する文化普遍的な感情認知機能が働いた可能性が考えられると同時に, 表情課題における正答率の天井効果による影響も考えられる (Table 1)。したがって, 視聴覚感情認知において, 中国人は日本人より表情からの感情情報により敏感に反応するという仮説 (a) は支持されなかった。今後の課題として, 表情による感情認知に与える声からの干渉程度について中日間の文化差を検討するには, 表情にノイズを付加し, 判断難易度を高めた (Tanaka et al., 2010) 刺激, あるいは感情強度が低く, 判断難易度が高い刺激を用いて, 再検討する必要がある。

次に, 声による感情認知に与える表情からの干渉について, 表出者側と認知者側の文化的異同を考慮して検討した結果, 内集団 (表出者と認知者が同じ文化に属する場合) においてのみ, 中国人と日本人の感情認知に違いが見られた。具体的には, 日本人は中国人に比べ, 内集団の声による感情認知に与える表情からの干渉が顕著に低かった。この結果から, 日本人は中国人より声からの感情情報に敏感に反応するという仮説 (a) は, 内集団の声による感情認知においてのみ支持された。このように日本人が他の文化と比べ, 声からの感情情報に敏感に反応する結果は, 日蘭の比較研究を行った Tanaka et al. (2010) の知見とも整合している。一方, 中国人は北米人と比較すると, 声による感情情報に依存しやすいことが示されている (Liu et al., 2015)。これらの先行研究では東洋人である中国人と日本人は, 西洋人である北米人やオランダ人よりも, 声からの感情情報に敏感であると述べているが, これはあくまでも集団主義社会と個人主義社会という大きな枠組みで区分された東西文化間における知見である。本研究では同じ東洋文化圏内に属する中国人と日本人の間においても文化差が存在し, 内集団の感情認知において, 日本人は中国人より声からの感情情報に依存しやすいことが示された。

それでは, 上記のような感情認知における文化差はどのような要因によって規定されているのだろうか。まず, 感情の表示規則の違いが考えられる。東洋人である中国人と日本人が, 西洋人である北米人やオランダ人と比べ, 声による感情情報に依存しやすい (Liu et al., 2015; Tanaka et al., 2010) のは, 中国人も日本人も表情からの感情表出を抑制する機能が働きやすいからであると考えられる。日本人は表情によるネガティブな感情表出を抑制し, 社会的場面にふさわしいと思

われる感情を表出するように表情をコントロールする傾向が、アメリカ人などの西洋文化と比べ顕著である (Ekman & Friesen, 1975 工藤訳 1987; マツモト・工藤, 1996)。このような社会的場面における感情表出の抑制機能は東アジア文化に固有の特徴である (Argyle, 1996) と指摘されているが、これはあくまでも個人主義社会とされる西洋文化と、集団主義社会とされる東洋文化という大きな枠組みにおける知見であり、必ずしも同じ東洋文化内に属する地域間に文化差がないことではない。李・松本 (2011) では韓国人も日本人のように社会的場面では感情表出を抑制すべきであるとみなしているが、その程度は日本人と比べ小さいと示されている。さらに、公的な状況において、日本人は中国人より喜びの感情を、中国人は日本人より怒りの感情を表出しやすいという自己認識を持っている (趙, 2002)。このように、同じ東洋文化圏内の地域間にも感情の表示規則の違いが報告されている。

東洋文化圏内に属する地域間における感情の表示規則の違いについては、より広範な実証的研究が必要であると考えられる。中日間の表情表出の文化差についてはまだ明らかではないが、以下のような可能性が想定される。

中国人も日本人も表情表出に抑制機能が働きやすいが、それぞれの抑制機能は異なる方法で働いていると考えられる。日本人は微笑みなどの表情表出を用いて、ネガティブな感情を隠したり、偽装する方法で抑制機能が働くのに対し、中国人は表情表出が少ない (無表情が多い) ことで抑制機能が働くのではないかと考えられる。つまり、日本人は本心とは異なる表情表出が多いことから、表情による感情表出への信頼性が低く、表情に比べて恣意的な操作がより困難な声からの感情情報に重みをおき、感情を認知する傾向があるのではないかと考えられる。一方、中国人は微笑みなどの表情に偏らず、怒りなどのネガティブな表情も表出しやすいことから、日本人よりは声からの感情情報への依存性が低いと考えられる。

次に、感情認知における文化差を規定する要因として、言語的要因による影響が考えられる。視聴覚感情認知において、日本語に比べて中国語の声による感情認知の方が表情からの干渉を受けやすかった。そして、声のみを単独で提示した場合も、日本語の声による感情認知では中国人に比べて日本人の方が、正答率が高かったのに対し、中国語の声による感情認知においては、中国人と日本人の正答率に差が見られなかった。これらの結果より、中国語は日本語に比べ、声に基づいた感情認知が困難であるという仮説 (b) は支持された。このことから、中国語は日本語より声による感情認知を困難にする特徴をもっている可能性があると考えられる。

これを支持する近年の知見として、Shigeno (2013)

が挙げられる。Shigeno (2013) は、声のみによる感情認知の場合、中国人も日本人も内集団優位性が見られるが、中国人発話者の喜び感情の判断は中国人参加者にとっても日本人参加者にとっても困難であることを示している。また、MDS (multi-dimensional scaling) 分析の結果、中国語と日本語の喜び声異なる次元にプロットされていることから、中国語と日本語の声による感情表出・認知に異なる特徴や規則があると考えられるのである。これは、中国語の言語的特徴である声調の多様な変化と、それが意味認知に与える強い影響 (柴田・柴田, 1990) によると考えられるが、具体的にどのような言語的特徴がどのように影響を与えるのかについては今後検討すべき課題である。

また、声のみによる感情認知において、6種の基本感情を用いた Shigeno (2013) では中国人参加者でも日本人参加者でも内集団優位性が見られているが、怒りと喜びのみを用いた本研究では日本人参加者のみに見られている。これは、中国語での喜び感情への認知が中国人にとっても困難であるという上述の点に関わっている可能性が考えられる。したがって、今後の課題として、中国語と日本語の言語固有の特徴によって感情認知に促進的・阻害的效果が生じるのはそれぞれどのような感情であり、その際に言語的要素と感情的要素がどのように相互作用するのかについて6種の基本感情を用いて検討を重ねる必要がある。

最後に、本研究では中国と日本の地域や方言による影響を取り除くために、全国の各地域から集まり、かつ標準語の使用に慣れている大学生を対象に検討した。したがって、本稿の結果は地域や方言による影響では説明できないと考えられる。一方、本研究では中国の漢民族を対象として検討したため、民族間の感情認知の違いによる影響を受けた可能性は考えられる。中国は統一された多民族の国として、56の民族を持っているが、中国の漢民族は人口の約90%以上を占めている。このことから、本研究の結果には中国人人口の大多数を占める漢民族の感情認知の傾向が強く反映されていると考えられる。しかし、民族間の違いによる影響の有無については確認する必要があり、それは今後検討すべき課題とする。

引用文献

- Argyle, M. (1996). *Bodily communication*. London: Routledge.
- Chang, N. C. (1958). Tones and intonation in the Chengtu dialect (Szechuan, China). *Phonetica*, 2, 59-85.
- 大坊 郁夫 (2007). 社会的脈絡における顔コミュニケーションへの文化的視点 対人社会心理学研究, 7, 1-10.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. R. Cole (Ed.),

- Nebraska Symposium on Motivation, 1971* (pp. 207–283). Lincoln, NE: University of Nebraska Press.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*, 124–129.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial cues*. Englewood Cliffs, NJ: Prentice-Hall.
- (エクマン, P.・フリーゼン, W. V. 工藤 力 (訳編) (1987). 表情分析入門——表情に隠された意味をさぐる—— 誠信書房)
- Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion, 10*, 159–168.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System (FACS)*. Salt Lake City, UT: A Human Face.
- Engelmann, J. B., & Pogossyan, M. (2013). Emotion perception across cultures: The role of cognitive mechanisms. *Frontiers in Psychology, 4*, 118.
- Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Cultural relativity in perceiving emotion from vocalizations. *Psychological Science, 25*, 911–920.
- Ishii, K., Reyes, J. A., & Kitayama, S. (2003). Spontaneous attention to word content versus emotional tone: Differences among three cultures. *Psychological Science, 14*, 39–46.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*, 1543–1548.
- 李 礼真・松本 芳之 (2011). 日本人と韓国人における表示規則 心理学研究, 82, 415–423.
- Liu, P., Rigoulot, S., & Pell, M. D. (2015). Culture modulates the brain response to human expressions of emotion: Electrophysiological evidence. *Neuropsychologia, 67*, 1–13.
- Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94*, 365–381.
- Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13*, 143–157.
- マツモト デイビッド・工藤 力 (1996). 日本人の感情世界——ミステリアスな文化の謎を解く—— 誠信書房
- 櫻庭 京子・今泉 敏・寛 一彦 (2004). 「ピカチュウ」にこめられた感性情報 音声研究, 8, 77–84.
- 柴田 武・柴田 里程 (1990). アクセントは同音語をどの程度弁別しうるか——日本語・英語・中国語の場合—— 計量国語学, 17, 317–327.
- 重野 純 (2004). 感情を表現した音声の認知と音響的性質 心理学研究, 74, 540–546.
- Shigeno, S. (2013). Multimodal perception of emotion in Japanese and Chinese. *Psychology Research, 3*, 504–511.
- Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience, 4*, 148–169.
- Tanaka, A., Koizumi, A., Imai, H., Hiramatsu, S., Hiramoto, E., & de Gelder, B. (2010). I feel your voice: Cultural differences in the multisensory perception of emotion. *Psychological Science, 21*, 1259–1262.
- 王 愛東・米谷 淳 (2008). インタビューにおける表情表出——中国人と日本人の比較—— 対人社会心理学研究, 8, 97–102.
- 渡邊 伸行・鈴木 竜太・吉田 宏之・續木 大介・番場 あやの・Chandrasiri, N. P. …山田 寛 (2007). 顔情報データベース FIND——日本人の顔面像データベース構築の試み—— 感情心理学研究, 14, 39–53.
- Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology, 43*, 303–311.
- 趙 特雷 (2002). 表示規則の日中比較研究 国際文化学, 6, 77–89.

—— 2015. 7. 30 受稿, 2016. 9. 24 受理 ——