

パラ言語情報の知覚とマガーク効果生起の関連

山本 寿子^{1,2} 河原 美彩子³ 田中 章浩¹

¹ 東京女子大学現代教養学部 〒167-8585 東京都杉並区善福寺 2-6-1

² 東京大学大学院教育学研究科 〒113-0033 東京都文京区本郷 7-3-1

³ 東京女子大学大学院人間科学研究科 〒167-8585 東京都杉並区善福寺 2-6-1

E-mail: ¹hisako_wy@lab.twcu.ac.jp

あらまし 他者の情動を読み取る手がかりとして顔と声、また声の中でも言語情報とパラ言語情報があるが、これらが異なる情動を表出することもある。先行研究では、複数の手がかりがあるときにどの情報を重視して情動判断を行うかについて、文化差や発達的变化が見られることが示唆されてきた。このような文化差および発達的变化は、視聴覚を利用した音韻知覚においても報告されている。これらの関連を調べるため、本研究では日本人の成人と5歳から12歳の子どもを対象に、複数情報を利用する情動判断課題と視聴覚音韻知覚課題を行い、その発達的变化と課題間の相関を調べた。児童期を通して、情動判断において声の情報およびパラ言語情報を利用する割合が増えること、視聴覚音韻知覚においては発達的变化が見られないこと、また11歳以上で課題間に相関が見られることが明らかとなった。複数情報を利用する課題間には共通の要因がある可能性、しかしその発達の様相は課題によって異なる可能性について議論する。

キーワード パラ言語情報, マガーク効果, 多感覚知覚, 認知発達

Developmental Relations between Emotion Perception from Speech and McGurk effect

Hisako W. YAMAMOTO^{1,2} Misako KAWAHARA¹ and Akihiro TANAKA¹

¹Tokyo Woman's Christian University 2-6-1 Zempukuji, Suginami-ku, Tokyo, 167-8585 Japan

²The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033 Japan

E-mail: ¹hisako_wy@lab.twcu.ac.jp

Abstract Facial expression and voice are available for judgement of others' emotion. Each cue sometimes expresses different emotional valence. Moreover, as for voice, emotions expressed by linguistic and paralinguistic information can conflict with each other. Previous research has revealed that humans put weight on multiple cues differently depending on their culture and age. In addition to emotional judgement, cultural difference has been observed in audio-visual phonetic perception (McGurk effect). In order to examine relations between these tasks, we investigated developmental changes in multimodal emotional judgement, audio-visual phonetic perception, and the correlations between these measures in Japanese adults and children from ages five to twelve. The utilization of voice and paralinguistic information in emotional judgement increased over age, while the use of audio information in phonetic perception remained the same during childhood. The correlations between tasks were observed in only adults and children from ages eleven to twelve. These results suggest the possibility that the common factors may exist between these multimodal perception tasks, but that the developmental path of each task differs.

Keywords Paralinguistic information, McGurk effect, Multimodal perception, Cognitive development

1. 問題と目的

他者とコミュニケーションを図る際、我々は、相手の表情や口の動き、声といった複数の手がかりを同時に用いて、情報を読み取っている。その際、数ある手

がかりのどこに重きを置いて判断を下すか、という情報間の重みづけの仕方には、文化差、発達的变化といった違いが見られることが、これまでの研究から明らかにされてきた。

たとえば、相手がどのような情動状態にあるかを読

み取る手がかりとして、相手の表情と話し声がある[1]。これらの手がかりに対して、オランダ人は表情を、日本人は話し声を重視して情動判断を行う、という文化差が見られている[2]。また、同じ日本人であっても、5歳の時点では表情を重視しているが、児童期を通じて、次第に話し声に重点をシフトしてゆくという発達的变化も報告されている[3]。

また、一口に話し声と言っても、そのことばの内容である言語情報と、それをどのように話すかというパラ言語情報のそれぞれから情動を表出することが可能である。皮肉のような発言を考えると、「上手だね」というポジティブな言語内容を暗くネガティブな口調で言う、ということもあるだろう。つまり、同じ発話の中であっても、これらが示す情動価値が一貫しないこともありうるのだ。そのような話し声の中での二つの手がかりについても、幼児は言語情報に基づいて情動を判断し、年齢を追うごとに、次第にパラ言語情報を優先するようになってゆく、という発達的变化が報告されている[4-6]。また、アメリカ人に比べ、日本人はパラ言語情報による影響を受けやすいという文化差の存在も指摘されている[7]。

複数の手がかりのうち、どれを重視するか、という点に文化差と発達的变化が見られるのは、情動判断についてだけではない。マガーク効果の生起という現象を通して、視聴覚音韻知覚においても報告されている。マガーク効果とは、/ga/の口の動きに/ba/の音声を組み合わせると、そのどちらでもない/da/の音韻が知覚されるという現象であり[8]、音韻の知覚が聴覚情報と視覚情報という複数の手がかりによって成立しうること示すものである。本研究ではこの現象について、多感覚知覚という観点から、視聴覚音韻知覚の際に視覚情報を利用するほどマガーク効果が生起しやすく、聴覚情報を重視しているほど生起しづらい、と捉える。この視覚情報と聴覚情報への重みづけの仕方について、文化差および発達的变化の存在が示唆されている。まず、マガーク効果は、英語話者に比べて日本人では生起しづらいことが指摘されている[9]。また、英語話者においても、子どもは成人に比べてマガーク効果が生起しづらく[8]、児童期に生起率が増えていくことも報告されている[10]。

このように、複数の手がかりがあるときにどの情報を重視するか、という重みづけにおける文化差と発達的变化は、情動の判断、音韻の知覚のいずれにおいても報告されている。これらの課題における手がかりには、視覚情報、聴覚情報、パラ言語情報といった共通性があるが、その課題間の関連は明らかになっていない。そこで本研究は、日本人の成人と5歳から12歳の子どもの対象に、これらの課題を実施する。先行研究

での文化差・発達的变化が報告されている、顔と声を手がかりとする情動判断課題(顔声情動判断課題)、言語情報とパラ言語情報を手がかりとする情動判断課題(言語パラ言語情動判断課題)、映像と音声による音韻知覚課題(視聴覚音韻知覚課題)を実施し、まず、課題ごとに、複数の手がかりに対する重みづけの発達的变化を調べる。そして、課題間の重みづけの相関を求めることで、これらの背後に存在する共通の要因を検討する。

2. 方法

2.1. 実験参加者

子どもの参加者として、日本語を母語とする5-6歳児44名、7-8歳児92名、9-10歳児95名、11-12歳児37名が参加した。また、成人の参加者として、日本語を母語とする30歳から73歳までの173名が参加し、このうち、30歳から39歳までの63名(平均36.1歳)を分析対象とした。

2.2. 刺激と手順

【顔声情動判断課題】

日本語話者2名(女性)が発話している動画を用いた。動画刺激は、(発話者(女性2名)×情動の種類(顔が喜び・声が喜び/顔が怒り・声が怒り/顔が喜び・声が怒り/顔が怒り・声が喜びの4種類)×セリフ(「はいもしもし」「さようなら」「これなに」「そうなんですか」の4種類)の計32個であった。このうち、顔と声で表出している情動が一致する16個の動画を顔声一致動画、これらが一致しない16個の動画を顔声不一致動画とした。この他、練習用動画として、顔、声ともに喜び(怒り)の情動を「あれどうなったの」というセリフとともに表している2個の動画を用いた。これらの動画刺激は、黒い背景の15.6インチディスプレイの中央に640×480ピクセルの大きさに呈示した。音声はSENNHEISER製ヘッドホン(HDA300)を通じて快適レベルの音圧で呈示した。

各試行では、まず、注視点(+記号)を500ミリ秒、同時に合図音(440Hzの純音)を100ミリ秒呈示した。注視点の位置は動画の中央に設定した。注視点の呈示開始から500ミリ秒後に動画刺激、続いてブランク画像を呈示した。実験参加者は動画を視聴し発話者が表している情動が喜びであるか、怒りであるかの強制選択を行った。回答はブランク画像時に、キー押しによって行った。実験参加者がキーを押してから500ミリ秒後に、次の試行を開始した。練習用動画を用いた練習試行2試行に続いて、本試行32試行(顔声一致試行、顔声不一致試行16試行ずつ)を実施した。本試行の順

序はランダムであった。

【言語パラ言語情動判断課題】

男女の日本語話者が発話した音声を用いた。音声刺激は、言語内容（ポジティブ / ネガティブの2種類）×セリフ（4種類：ポジティブな言語内容のセリフは「いいよ」「上手だね」「優しいね」「楽しいね」、ネガティブな言語内容のセリフは「だめだよ」「下手だね」「がっかりだよ」「つまらないね」）×パラ言語情報（ポジティブ / ネガティブの2種類）の計16個であった。このうち、言語内容とパラ言語情報が一致する8つの音声を言語パラ言語一致（以下、言パラ一致）音声、これらが一致しない8つの音声を言語パラ言語不一致（以下、言パラ不一致）音声とした。これらの音声刺激を実験参加者に対し、SENNHEISER製ヘッドホン（HDA300）を通じて快適レベルの音圧で呈示した。

各試行では、合図音（440Hzの純音）を100ミリ秒間呈示し、その1400ミリ秒後に1つの音声刺激を再生した。実験参加者は音声を聞き、発話者が楽しそうに話しているか、あるいは嫌そうに話しているかの強制選択を行った。回答はキー押しによって行った。実験参加者がキーを押してから500ミリ秒後に、次の試行を開始した。1つの音声刺激につき2試行、合計32試行（言パラ一致試行、言パラ不一致試行ともに16試行ずつ）を実施した。試行の順序はランダムであった。

【視聴覚音韻知覚課題】

日本語話者6名（男女3人ずつ）が単音節（/ka/, /ta/, /pa/のいずれか）を発話している動画を用いた。動画刺激として、映像と音声で表している音韻が一致する音韻一致動画12個（それぞれの音韻について4人ずつ）、/ka/の映像と/pa/の音声を組み合わせて作成したマガーク動画6個を動画刺激として用いた。これらの動画刺激は、黒い背景の15.6インチディスプレイの中央に640×480ピクセルの大きさで呈示した。音声はSENNHEISER製ヘッドホン（HDA300）を通じて快適レベルの音圧で呈示した。

各試行では、まず、注視点（+記号）を800ミリ秒、同時に合図音（440Hzの純音）を100ミリ秒呈示した。注視点の位置は、動画刺激の話者の鼻付近に設定した。注視点の呈示から1300ミリ秒後に動画刺激、続いてブランク画像を呈示した。実験参加者は、発話者が発した音が何であったか、/ka/, /ta/, /pa/の中から強制選択を行った。回答はブランク画像時に、キー押しによって行った。実験参加者がキーを押してから500ミリ秒後に、次の試行を開始した。視聴覚一致動画は各1試行ずつ（視聴覚一致試行）、マガーク動画は各2試行ずつ（マガーク試行）呈示した。したがって、視聴覚一

致試行とマガーク試行を各12試行ずつ、合計24試行を実施した。なお、すべての試行の順序はランダムであった。

3. 結果

3.1. 各課題における発達的变化

【顔声情動判断課題】

顔声一致試行、顔声不一致試行の各試行において、声に基づいて答えた割合を声反応率として算出した。各年齢群の平均値を図1に示す。声反応率について、年齢群（5-6歳・7-8歳・9-10歳・11-12歳・成人）×試行の種類（顔声一致・顔声不一致）の二要因の分散分析（年齢群は被験者間、試行の種類は被験者内）を行ったところ、年齢群の主効果（ $F(4, 326)=12.10, p<.001, \eta_p^2=.13$ ）、試行の種類的主効果（ $F(1, 326)=2795.40, p<.001, \eta_p^2=.90$ ）、これらの交互作用が有意であった（ $F(4, 326)=8.03, p<.001, \eta_p^2=.09$ ）。単純主効果の検定を行ったところ、顔声一致試行、顔声不一致試行のいずれにおいても、年齢群による差が有意であった（一致： $F(4, 326)=3.07, p=.017$ 、不一致： $F(4, 326)=10.48, p<.001$ ）。多重比較（Shaffer法）の結果、顔声一致試行では、5-6歳より7-8歳（ $p=.012$ ）、9-10歳（ $p=.016$ ）の声反応率が高かった。顔声不一致試行では、5-6歳より9-10歳（ $p=.010$ ）と11-12歳（ $p<.001$ ）で、7-8歳より9-10歳（ $p=.041$ ）と11-12歳（ $p<.001$ ）で、9-10歳より11-12歳（ $p=.003$ ）の声反応率が高かった。また11-12歳の声反応率は成人より高かった（ $p<.001$ ）。

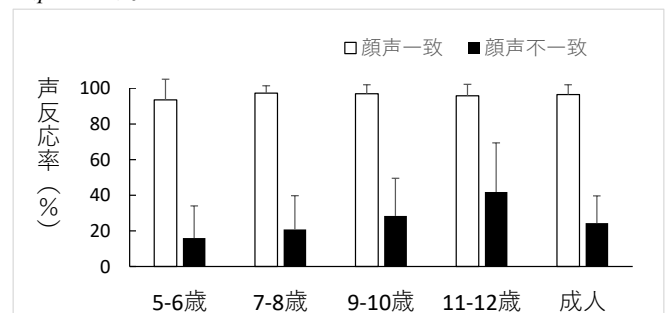


図1 顔声情動判断課題
年齢群ごとの声反応率

【言語パラ言語情動判断課題】

言パラ一致試行、言パラ不一致試行の各試行において、パラ言語情報に基づいて答えた割合をパラ言語反応率として算出した。各年齢群の平均値を図2に示す。パラ言語反応率について、年齢群×試行の種類（言パラ一致・言パラ不一致）の二要因の分散分析（年齢群は被験者間、試行の種類は被験者内）を行ったところ、年齢群の主効果（ $F(4, 326)=17.12, p<.001, \eta_p^2=.17$ ）、試

行の種類の主効果 ($F(1, 326)=222.11, p<.001, \eta^2=.41$), これらの交互作用が有意であった ($F(4, 326)=10.49, p<.001, \eta^2=.11$)。単純主効果の検定を行ったところ, 言パラ一致試行, 言パラ不一致試行のいずれにおいても, 年齢群による差が有意であった (一致: $F(4, 326)=4.54, p=.001$, 不一致: $F(4, 326)=15.32, p<.001$)。多重比較 (Shaffer 法) の結果, 言パラ一致試行では 5-6 歳より 7-8 歳 ($p=.012$) と 11-12 歳 ($p=.003$), 9-10 歳より 11-12 歳の ($p=.038$) の声反応率が高かった。言パラ不一致試行では, 5-6 歳はその他すべての年齢群より声反応率が低かった (7-8 歳: $p=.001$, その他: $p<.001$)。7-8 歳は, 9-10 歳 ($p=.046$), 11-12 歳 ($p<.001$) より声反応率が低かった。

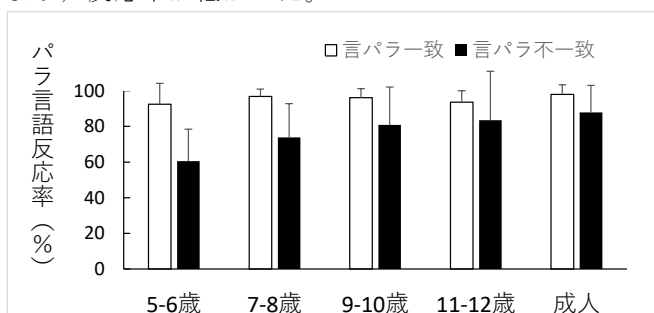


図2 言語パラ言語情動判断課題
 年齢群ごとのパラ言語反応率

【視聴覚音韻知覚課題】

視聴覚一致試行の正答率の平均値は, 5-6 歳が 94.1% ($SD=10.3$), 7-8 歳が 96.1% ($SD=6.1$), 9-10 歳が 95.9% ($SD=8.4$), 11-12 歳が 95.7% ($SD=6.7$), 成人が 97.4% ($SD=6.5$) であった。視聴覚一致試行の正答率について, 年齢群による一要因の分散分析 (被験者間) を行ったところ, 年齢群による差は見られなかった ($F(4, 326)=1.91, p=.315, \eta^2=.02$)。

マガーク試行については, 各試行における反応のうち, /ka/としたものを視覚反応, /ta/としたものを融合反応, /pa/としたものを聴覚反応とし, 視覚反応の割合と融合反応の割合を合わせたものを, 視覚影響率として算出した。各年齢群の平均値を図3に示す。視覚影響率について, 年齢群による一要因の分散分析 (被験者間) を行ったところ, 主効果が有意であった ($F(4, 326)=22.41, p<.001, \eta^2=.22$)。多重比較 (Shaffer 法) の結果, 成人の視覚影響率は, その他のすべての年齢群より高かった ($p<.001$)。その他の年齢群間の違いは見られなかった。

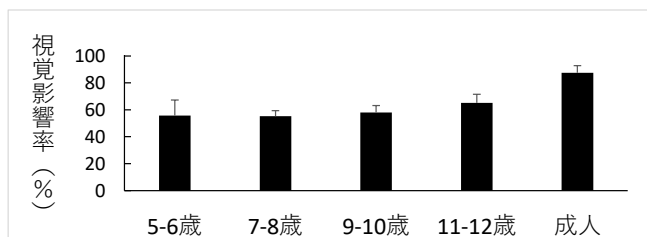


図3 視聴覚音韻知覚課題
 マガーク試行における視覚影響率

3.2. 課題間の相関

複数の手がかりのうちどの情報を重視するか, という重みづけにあたって, 課題間に関連があるかを調べるために, 顔声情動判断課題における顔声不一致試行の声反応率, 言語パラ言語情動判断課題における言パラ不一致試行のパラ言語反応率, 視聴覚音韻知覚課題におけるマガーク試行の視覚影響率間の相関係数を, 年齢群ごとに求めた。

5-6 歳, 7-8 歳, 9-10 歳の年齢群においてはいずれの課題間にも有意な相関は見られなかった。11-12 歳の年齢群においては, 声反応率と視覚影響率間に弱い負の相関が見られた。また, パラ言語反応率と視覚影響率の間に, 有意傾向ではあるが, 弱い正の相関が見られた (表1)。成人においては, 声反応率とパラ言語反応率の間に弱い負の相関が見られた。有意傾向ではあるが, 声反応率と視覚影響率間に弱い負の相関が見られた。また, パラ言語反応率と視覚影響率の間に, 弱い正の相関が見られた (表2)。

表1 11-12 歳における課題間の相関係数

	声反応率	パラ言語反応率	視覚影響率
声反応率			
パラ言語反応率	-0.19		
視覚影響率	-0.34*	0.28+	

* $p<.05$, + $p<.1$

表2 成人における課題間の相関係数

	声反応率	パラ言語反応率	視覚影響率
声反応率			
パラ言語反応率	-0.37**		
視覚影響率	-0.24+	0.26*	

** $p<.01$ * $p<.05$, + $p<.1$

4. 考察

4.1. 課題ごとの手がかりの重みづけの発達的变化

顔声情動判断課題において, 顔声不一致試行の声反応率は, 5-6 歳から 11-12 歳にかけて増加し, 成人では

それより減少することが示された。日本人の場合、児童期を通じて顔から声に手がかりの重点がシフトする、という点は、先行研究[3]の結果と一致している。一方、成人の参加者においては、11-12歳より声反応率が低いという特徴が見られた。これは先行研究[3]で見られていた、成人と11-12歳とで声反応率に差が見られなかった結果と対照的である。先行研究と本研究の成人の参加者の違いは、前者が大学生であったのに対し、後者は子どもの参加者の保護者であり、かつ、平均36歳と前者より年齢層が上であったことがあげられる。この違いが成人の声反応率の違いに繋がった理由として、第一に、日本人は成人後に加齢に伴い、声から再び顔に重きを置くようになっていくという可能性が考えられる。第二に、日常的に子どもと接するという相互作用により、声反応率が減少するという可能性も考えられる。

言語パラ言語情動判断課題において、言パラ不一致試行のパラ言語反応率は、5-6歳から成人にかけて増加することが示された。このように、パラ言語反応率が児童期を通して増加し、成人でピークを迎えるという発達的变化は、先行研究の知見[6]と一貫していると言える。

視聴覚音韻知覚課題では、子どもの参加者の中では、年齢による視覚影響率の違いは見られなかった。日本語母語児を対象とした先行研究でも、児童期ではマザー効果の生起に発達的变化は見られておらず[10]、今回の結果はそれと一貫したものだと言える。ただし、成人における視覚影響率は子どもより高かったことから、視聴覚音韻知覚において視覚情報も手がかりにする、という重みづけのシフトが児童期以後に生じる可能性が示唆された。

4.2. 課題間の重みづけの関連

11-12歳の年齢群と成人では、顔声情動判断課題における声反応率と、視聴覚音韻知覚課題における視覚影響率の間に負の相関が見られた。これらの課題はいずれも、顔を見るという視覚情報、音声を聞くという聴覚情報の二種類の手がかりを用いるという点で共通している。顔声情動判断課題における声反応率というのは、話者の顔という視覚情報を軽視し、聴覚情報を重視する度合い、と言い換えることができる。これと視聴覚音韻知覚課題における視覚情報を重視する割合に負の相関が見られたということから、情動判断において視覚情報を重視することと、視聴覚による音韻知覚において視覚情報を重視することの背後には、相手の「顔への注目」という共通の要因が存在すると考えられる(図4-(a))。相手の顔に注目する人ほど、顔と声の情動判断において顔を重視し、視聴覚音韻知覚課

題においても口元の動きを重視すると考えられる。

次に、11-12歳の年齢群と成人では、言語パラ言語による情動判断課題におけるパラ言語反応率と、視聴覚音韻知覚課題における視覚影響率の間に正の相関が見られた。言語パラ言語による情動判断課題の2つの手がかり、つまり言語情報とパラ言語情報は、いずれも聴覚情報である。その中のパラ言語を重視することと、視聴覚による音韻知覚において視覚情報を重視することの間に正の相関が見られたということになる。言語情報にとってはそれを構成する音韻情報が重要である一方、パラ言語情報は音韻以外の音の要素で規定される部分が多い。ここから、「音韻情報の重視」という要因が存在し、パラ言語反応と、視聴覚音韻知覚課題において音韻そのもの以外から影響を受ける視覚影響率に対して、それぞれ抑制的に働くと考えられる(図4-(b))。

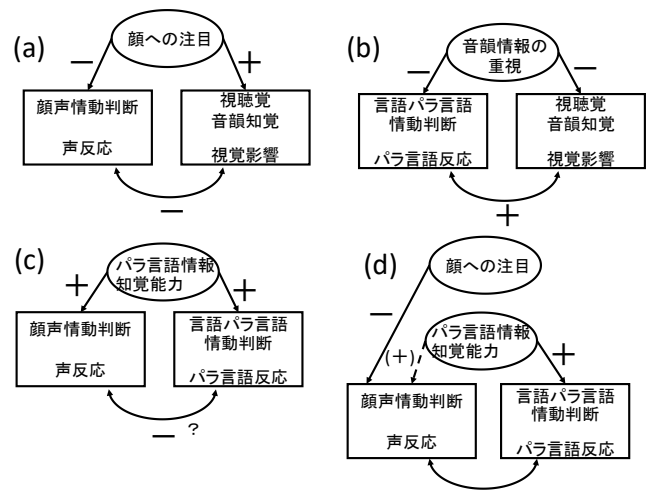


図4 課題間の相関に関わる各要因

(四角は課題のパフォーマンス、円は背後の要因、+は促進(正の相関)、-は抑制(負の相関)を表す)

また、パラ言語情報を重視するためには、そもそもパラ言語情報を理解しなければならないため、「パラ言語情報知覚能力」要因の存在も考えられる(図4-(c))。しかしここで、成人で見られた、顔声情動判断課題における声反応率と、言語パラ言語情動判断課題におけるパラ言語反応率の間に負の相関が見られたことを考えねばならない(図4-(c)、?部分)。顔声情動判断課題で使った音声は、言語情報はニュートラルなものであり、話し方、つまりパラ言語情報のみで情動を表すものであった。したがって、「パラ言語情報知覚能力」は、顔声情動判断課題における声反応率にも促進的な影響をもたらすと考えられる。それならば、顔声情動判断課題における声反応率と言語パラ言語情動判断課題におけるパラ言語反応率との間に負の相関が見られるの

はなぜだろうか。この理由として、「顔への注目」要因が声反応率を抑制する影響力が、「パラ言語情報知覚能力」が声反応率を促進する影響力を上回っているため、という可能性が考えられる（図 4-(d)。「顔への注目」をする人が情動判断において顔を重視するのは、彼らの「パラ言語情報知覚能力」の低さゆえに声から情動を読み取ることができず、顔に頼らざるを得なかったから、ではなく、あくまでもパラ言語情報を読み取った上で、それよりも顔の情報を優先したから、ということになる。ただし、今回実施した言語パラ言語情動判断課題では、成人は不一致課題においてもパラ反応率が極めて高く、天井効果によって、個人差が見えづらくなっている、という可能性については注意すべきだろう。

ここまで述べてきた課題間の相関をすべてふまえると、図 5 のようなモデルを想定することができる。

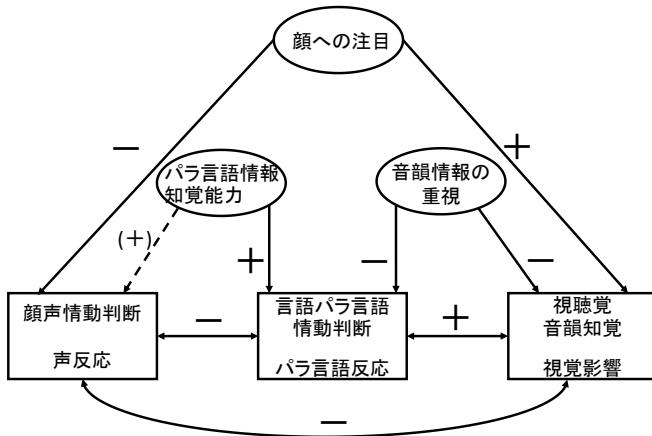


図 5 情動判断課題と視聴覚音韻知覚の
 多感覚知覚に関わる要因

(四角は課題のパフォーマンス，円は背後の要因，
 +は促進（正の相関），-は抑制（負の相関）を表す）

以上に示してきたように、11 歳以降の参加者において課題間の相関が見られたことから、これらの課題の背後に共通する要因が存在することが示唆された。その一方で、10 歳以下の子どもにおいては相関が見られなかった。この理由の可能性の 1 つとして、視聴覚音韻知覚と情動判断とで、手がかりへの重みづけの発達の变化的時期と方向に違いがあることが考えられる。本研究の結果では、情動の読みとりにおける重みづけの変化は、顔と声、言語情報とパラ言語情報のいずれにおいても児童期に生じていた。また、顔と声において、その変化は視覚情報から聴覚情報にシフトする、という方向で生じていた。一方、視聴覚による音韻知覚については、その重みづけの変化が児童期以降に生じる可能性が示唆されていた。また、成人に向けての発達の变化的方向は、聴覚情報から視覚情報、という

ものであった。発達過程の途上にある子どもにおいて、成人のような課題間の相関が見出されなかった理由の 1 つには、このような課題間の非対称性があるのかもしれない。

最後に本研究の展望として、多感覚知覚に関わる要因の普遍性を検討する必要性を述べる。欧米人に比べると、日本人は声から情動判断をしやすく [2]、またマガーク効果も起こりづらい [9][10]。図 5 のモデルによれば、その文化差は日本人の「顔への注目」が低いため、と説明することができる。しかし、それらの課題の結果をもとに言語パラ言語情動判断課題の結果を予測すると、日本人のパラ言語反応は欧米人より低くなるはずである。しかし実際には、日本人は欧米人よりもパラ言語情報に影響を受けやすいという知見が報告されている [7]。この齟齬をどう考えるべきかは大きな課題である。多感覚知覚における日本人と欧米人との文化差の背景には、今回の結果から想定されるもの以外の要因とメカニズムが介在している可能性も、今後は検討する必要があるだろう。

文 献

- [1] B. de Gelder and J. Vroomen, "The perception of emotions by ear and by eye," *Cognition & Emotion*, vol.14, no.3, pp.289-311, 2000.
- [2] A. Tanaka, A. Koizumi, H. Imai, S. Hiramatsu, E. Hiramoto, and B. de Gelder, "I feel your voice. Cultural differences in the multisensory perception of emotion," *Psychological Science*, vol. 21, no. 9, pp. 1259-1262, 2010.
- [3] M. Kawahara, D. A. Sauter, and A. Tanaka, "Development of cultural differences in emotion perception from face and voice," Poster presented at 31st International Congress of Psychology, 2016.
- [4] M. Friend, "Developmental changes in sensitivity to vocal paralinguistic," *Developmental Science*, vol. 3, no. 2, pp. 148-162, 2000.
- [5] 池田慎之介, 針生悦子, "幼児における話者の感情の推測", 日本発達心理学会第 27 回大会ポスター発表, 2016.
- [6] 野口由貴, 小澤由嗣, 山崎和子, 今泉敏, "音声から話者の心を理解する能力の発達," *音声言語医学*, vol.45, pp269-275, 2004.
- [7] K. Ishii, J. A. Reyes, and S. Kitayama, "Spontaneous attention to word content versus emotional tone: Differences Among Three Cultures," *Psychological Science*, vol. 14, no. 1, pp. 39-46, 2003.
- [8] H. McGurk, and J. McDonald, "Hearing lips and seeing," *Nature*, vol. 264, pp. 746-748, 1976.
- [9] K. Sekiyama and Y. Tohkura, "McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility," *Journal of Acoustic Society of America*, vol. 90, no.4, pp. 1797-1805, 1991.
- [10] K. Sekiyama and D. Burnham, "Impact of language on development of auditory-visual speech," *Developmental Science*, vol. 11, no. 2, pp. 306-320, 2008.