

行為情報が乳児の視聴覚間協応に及ぼす影響

—視線計測からのアプローチ—

東京大学大学院教育学研究科¹⁾ 丸 山 慎
東京大学大学院教育学研究科 渡 辺 は ま

What do infants hear from seeing motion?
-an eye-tracking study on an impact of motional information on
the development of audiovisual congruence-

Graduate School of Education, The University of Tokyo MARUYAMA, Shin
Graduate School of Education, The University of Tokyo WATANABE, Hama

要 約

音の知覚とは、その音を生成している事象についての視覚性、運動性の知覚を伴ったもの、すなわち異種の感覚モダリティを横断するものである。これをさらに敷衍するならば、例えば音を発生させている行為者の行為を見ること（＝視覚）、その行為から聴こえてくるはずの音の特徴（＝聴覚）を予測させるような視聴覚間での協応を仮定することができるのではないだろうか。そこで本研究では、音の知覚における視聴覚間協応の発達の基盤を明らかにすることを目的として、4ヵ月齢および10ヵ月齢の乳児を対象にした視線計測を行った。視覚的刺激（コンピュータで描いた人型キャラクター）の行為（＝動き）の特徴と周期性が、同時に再生される聴覚的刺激の特徴と一致する／一致しない映像をランダムに呈示し、その際の乳児の注視領域および注視時間を計測することによって、行為と音の特徴との一致／不一致と乳児の視覚的探索のパターンとの関係を検討したのである。分析の結果、人型キャラクターの「顔」に対する注視時間は月齢が上がるほど長くなる傾向がみられ、またキャラクターの動きに対しては、どちらの月齢群の乳児も鋭敏に反応していたことが示された。乳児は「顔」を長く注視することによって、動きや音を変化させるキャラクターの意図のようなものを探ろうとしていた可能性も考えられるが、現段階では、視覚的な探索パターンが映像刺激の行為と音の組み合わせに応じて変化していたという明確な結果を得るには至らなかった。以上の結果について、映像刺激の構造や分析単位の拡張といった点から議論し、本研究の一応の総括とした。

【キー・ワード】乳児, 音知覚, 行為情報, 視聴覚間協応, アイトラッキング

¹⁾ 現所属：駒沢女子大学人文学科

Abstract

Sound perception in everyday life includes *event* perception of what is producing the sound. This means that sound perception is a product of the union of multiple perceptual sources, including visual and/or motoric information that structures the sound producing event. In the present study, we aimed to investigate what element of an event, such as the agent's action, was crucial as a visual cue to specify the auditory feature of the expected sound, and “when-and-how” the perceiver developmentally learns to perceive such an audiovisual link. We ran an eye-tracking study, using computer generated movies as stimuli: an original character, “the smiling clown” appeared at the center of the monitor and manipulated an object that produced either a drumming sound or a smooth sound by generating one of two types of action, beating the object or smoothly pulling/pushing it. Forty-nine 8-month-old, and thirty 10-month old infants participated in the study. We randomly presented the four possible types of the audio-visual couplings of the stimuli, such as the drumming sound with the beating action, and analyzed the infants' total looking time to the target areas of the stimuli. Our logic was that, if infants exhibited sensitivity to the audiovisual link in the context of the sound perception, they would show a perceptual preference for the congruent couplings of the audio-visual features of the stimuli (i.e., “drumming-sound with the beating-action” and “smooth-sound with the continuous-action”) to the incongruent couplings (i.e., “drumming-sound with the continuous-action” and “smooth-sound with the beating-action”). Although, at present, no bias for the audio-visual congruence has been clearly shown at any age, we have found that the total looking time to the clown's face was much longer than to any other area of the stimulus. We speculate on the implication of these results with reference to the relevant studies, especially on the human ability to infer the intentions of others, and discuss possible directions for future developments of our research.

【Key words】 infant, sound perception, motional information, audiovisual congruence, eye-tracking

目 的

音は、単に聴覚的な刺激として受容されるだけではなく、その音を生成する人の行為やモノの運動についての知覚を伴う。すなわち音の知覚とは、音を生成する事象に関する視覚性、運動性の知覚を伴っているという点で、聴覚を含む複数の感覚モダリティを横断するものである。こうした音の知覚における多感覚性については、例えば事象と音の知覚との対応 (Spelke, 1976; 1979) をはじめ、人の発声と発声時の顔や口唇の形とのマッチング (Lewkowicz, 1996; Mugitani et al. 2008)、そして音高の知覚と空間性の知覚との一致 (例えば上行する音系列を聴取させると空間の上部が意識される

という聴覚的刺激としての音高と視覚的刺激としての空間の高低の一致：Rusconi, et al. 2005；Evans, 2010。乳児を対象にしたものはWagner, et al.1981; Walker, et al. 2010) に関する研究等、さまざまな知見が報告されている。

これらの先行研究が示唆するように、複数の感覚モダリティ間での協応が人の知覚の頑強な事実であるならば、ある知覚内容が、別の感覚の知覚内容を“予測”させるような関係も成立すると仮定することができる。例えば筆者らは、乳児の音高の知覚（聴覚）と空間性の知覚（視覚）との視聴覚間協応について視線計測装置を用いた検証を行い、装置の画面上を移動するターゲットが突然消失してしまった場合、つまり視覚刺激の入力が遮断された場合でさえも、聴覚刺激としての音高の情報にもとづいた視聴覚間協応が維持されることを明らかにした（Maruyama, et al. 2011, 丸山他, 2011）。消失したターゲットの行方を探索する際、音高が上行していく音階を提示した場合には、乳児の視線が画面の上部領域に向けられる傾向がみられたのである。つまり上行していく音の“高さ”が、不在となった視覚刺激の行方を予測的に誘導し、乳児の視線を空間のより“高い”位置に向けさせるという協応が起こったのである。

音の知覚をめぐる多感覚性、とりわけ音の知覚とその音を生起させている事象や行為の知覚との結び付きについては、知覚・認知に関する「生態学的アプローチ」（Gibson, 1966; 1979）およびその知見を音響学に応用した「生態学的音響学」と呼ばれる研究領域からも示唆に富んだ成果が報告されており、音を聴くことは、事象の運動を聴くことであるという主張が展開されている（Gaver, 1993; Shove & Repp, 1995）。これらの知見を踏まえると、ひとつの仮説を導くことができる。それは、行為の視覚的な見えには（その行為から）聴こえてくるはずの音の特徴を予測させる聴覚情報が含まれ、一方で、音はそれを生成している行為の特徴を予測させる視覚および運動情報を含んでいるのではないか、ということである。このような仮説が成り立つとすれば、果たして音を生成する行為とそこから予測される音との関係を特定する情報は、行為のどのような部分に内在しているのだろうか。そして聴取者は、どのような探索行動や発達のプロセスを経て、行為と音とを結び付ける情報を知覚できるようになっていくのだろうか。

人の発達初期における「行為の情報」と「音の知覚」との対応関係を検証した研究はきわめて少なく、例えば幼児を対象にして、特定の楽器の音とその楽器の演奏シーンの知覚との対応関係を検証した研究（Pick et al. 1994）がみられる程度である。そこで本研究では、視聴覚間協応をもたらす行為情報に対する探索能力とその発達という問題にアプローチし、上述の仮説を実験的に検証する。異なる行為の映像刺激を音の生成事象として呈示し、それらの行為に対応した（あるいは対応していない）音響刺激を聴取させた際の乳児の反応を複数の月齢群間で比較するのである。

本研究では、単に視聴覚間でのマッチングを測るだけではなく、視線計測装置を用いて映像刺激に対する乳児の視覚的な探索行動の詳細を記録する。視線の変化は、知覚内容を特定するような情報が集中する部分に収束していく可能性が高く、そこから感覚間での協応を生じさせている刺激の構造の特徴やその刺激を知覚的に処理する過程について推測することができると考えられる。例えば映像刺激として提示される行為とそこから予測される音の特徴が対応する場合、乳児の注視は、映像刺激の特定の部分に集中していきだろうか。逆に行為と音の間に対応がない場合には、視線が向けられる領

域が拡散し、より活発な探索活動が生じるかもしれない。また乳児の注視行動を異なる月齢群間で比較することによって、視覚的な探索の発達のな変化を捉えることができるはずである。以上の計測をもとに、行為情報に対する視覚的な探索が、音の特徴に対する予測的な知覚（聴覚的な知覚）に与える影響を検証し、視聴覚間での協応が生じる発達の基盤について明らかにすることを本研究の目的とする。

方 法

研究協力者 健常な 4 ヶ月齢児および 10 ヶ月齢児合計 79 名（4 ヶ月齢児 49 名 [男児 25 名, 女児 24 名 : 平均日齢 127.8 日, SD =5.458], 10 ヶ月齢児 30 名 [男児 12 名, 女児 18 名 : 平均日齢 306.4 日, SD =9.24]

使用機器 視線計測装置 Tobii-T120 (トビー・テクノロジー・ジャパン), コンピューター, スピーカー

視聴覚刺激 音を生成する行為の差異が明瞭であると思われる 2 種類の動作（垂直方向に動く“打動作”と水平方向に動く“伸縮動作”）を素材とし、それらをフラッシュ・アニメーションで描いたヒト型キャラクターに再現させたものを視覚的刺激として使用した。打動作と伸縮動作に関しては、移動に要するピクセル数を一致させることによって、動きの大きさを同一とみなせるように設定した（2 つの運動の違いが、動きの「方向」のみになるようにした）。一方、音響刺激は打動作から引き出される音および伸縮動作から引き出される音として「打撃音（打楽器音）」と「平滑音（グリッサンド音）」の 2 種類を使用した。打撃音は、約 1 秒間に 1 回の割合で C3 と D3 の音が順に各 3 回ずつ計 6 回再生され、平滑音（グリッサンド）は C3 と D3 の間で切れ目なく音高が変化し、約 2 秒で音高がずり上がる（ずり下がる）ように再生された。以上の映像および音響の刺激の組み合わせを操作し、動作の特徴とそこから予測される音響刺激との対応が高いと思われるもの（例えば打動作と打撃音）と対応が低いと思われるもの（例えば打動作と平滑音）とをそれぞれの動作と音について組み合わせ、計 4 種類の刺激を作成した（図 1）。これら 4 種類の映像刺激を各 6 回ずつランダムに再生できるように配置し、映像刺激の間にそれらとは無関係のアテンションゲッターを差し挟んで刺激を呈示していった。1 つの映像刺激の長さは約 17 秒で、全ての映像刺激を視聴した場合にかかる時間は約 9 分であった。

分析の視点と仮説 ①ヒト型キャラクターの行為情報と予測される音の特徴との対応が高いと思われる刺激に対しては、乳児の視線の方向はキャラクターの行為（手腕の動き）の中心的な部分に集中し、視線の遷移パターンが安定する。②行為情報と音の特徴との対応が低いと思われる刺激に対しては、視線の方向が拡散し（手腕だけではなくキャラクターの顔など様々な部位）、探索行動が増加する。

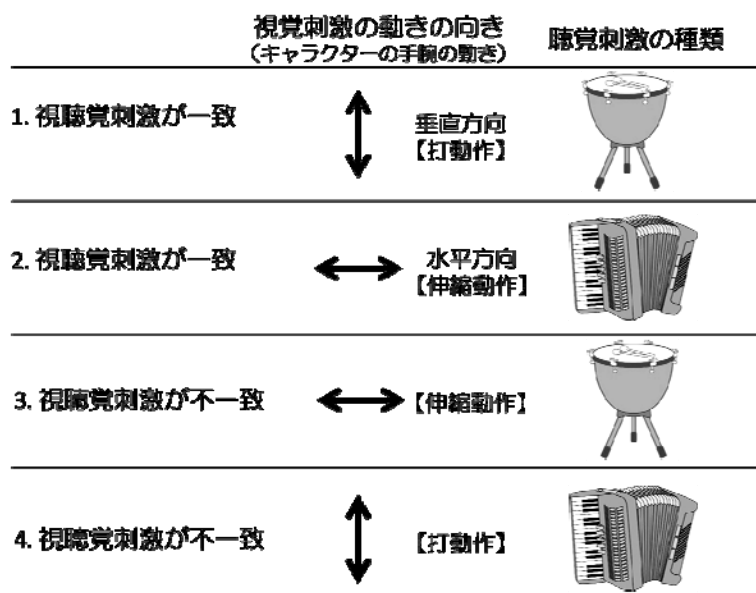


図1 視聴覚刺激の種類

手続き 保護者に研究内容を十分に説明し同意書の記入を確認後、乳児は保護者とともに防音スペースへ移動した。乳児は、刺激が呈示されるモニタ正面で保護者の膝の上に着座し、キャリブレーション（モニタの4隅と中央の計5点）を行った後、計測を開始した。計測中、保護者には乳児に声をかけたり、指差し行動などをしないように教示した。乳児がモニタを注視している限り映像刺激を呈示し続け、明らかに飽きてしまってモニタを見なくなったり、泣き出してしまった時点で計測を終了した。

結果

有効データの基準 映像刺激の呈示開始時から2秒間のうち80%以上（1.6秒以上）映像のいずれかの部分を注視していたことが確認され、かつその基準を満たした試行が4種類の視聴覚刺激の全てについて2試行以上取得できていた乳児のデータを有効とした。また4ヵ月齢児のうち1名の早産児のデータは今回の分析には含めなかった。以上の基準から最終的に分析対象としたのは4ヵ月齢児13名、10ヵ月齢児22名のデータであった。

注視領域と注視時間の算出 映像刺激に登場するキャラクターの「顔」および手腕の運動領域を「行為領域A」と「行為領域B」に分割し（図2）、それぞれの領域に対する乳児の注視時間を計測した。視聴覚刺激の組み合わせのうち同一種類の試行の注視時間を平均し、その長さを1回の映像刺激の再生時間に対する割合として表し、それらを月齢間で比較した。

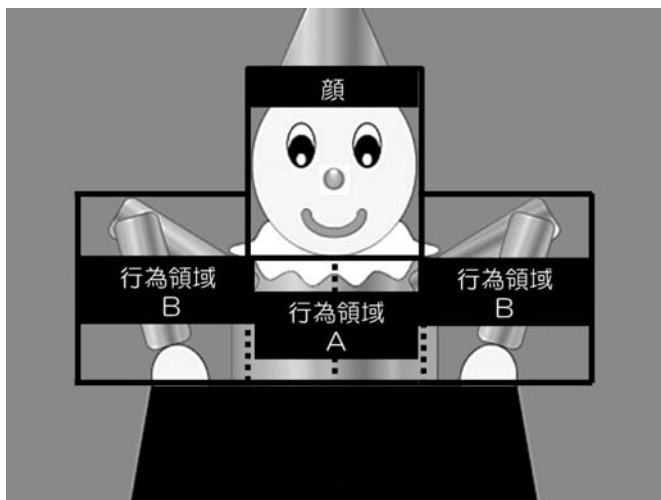


図 2 映像刺激の注視領域の区分

「顔」に対する注視時間 図 3 は、キャラクターの「顔」に対する各月齢の注視時間の割合を 4 種類の視聴覚刺激の種類別にまとめたものである。注視時間の割合について、月齢と視聴覚刺激の種類を要因とする 2 (4 ヶ月, 10 ヶ月) × 4 (1, 2, 3, 4) の分散分析を行ったところ、月齢の主効果に有意傾向 ($F[1, 33]=4.051, p=0.052$) がみられ、視聴覚刺激の組み合わせの主効果には有意差がみられた ($F[3, 99]=4.698, p=0.004$)。視聴覚刺激の種類について Scheffe の方法による多重比較を行った結果、10 ヶ月齢の「2」と「4」および「3」と「4」の間にそれぞれ有意差がみられた。

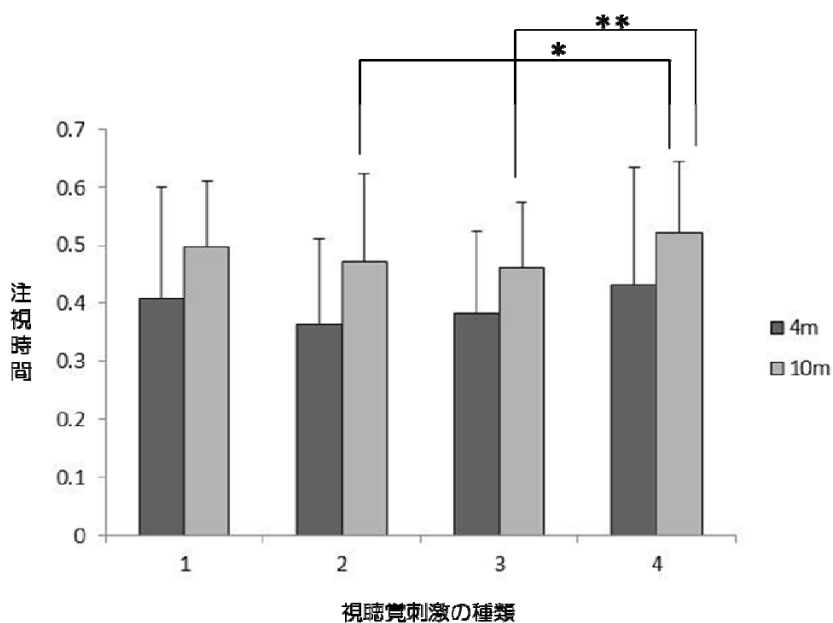


図 3 キャラクターの「顔」に対する注視時間

「行為領域」に対する注視時間 図4は、キャラクターの「行為領域（図2の行為領域AとBを合わせたもの）」に対する各月齢群の注視時間の割合を4種類の視聴覚刺激別にまとめたものである。注視時間の割合について、月齢と視聴覚刺激の種類を要因とする2（4ヵ月、10ヵ月）×4（1, 2, 3, 4）の分散分析を行ったところ、月齢に有意差はみられなかったが、各月齢内での視聴覚刺激の種類の主効果に有意差がみられた（4ヵ月齢： $F[3, 99]=3.839, p=0.012$, 10ヵ月齢： $F[3, 99]=9.162, p<0.01$ ）。Scheffeの方法による多重比較を行った結果、4ヵ月齢では「1」と「2」, 「1」と「3」, および「2」と「4」の間に有意差がみられ、10ヵ月齢では「1」と「2」, 「1」と「3」, 「2」と「4」, および「3」と「4」の間に有意差がみられた。

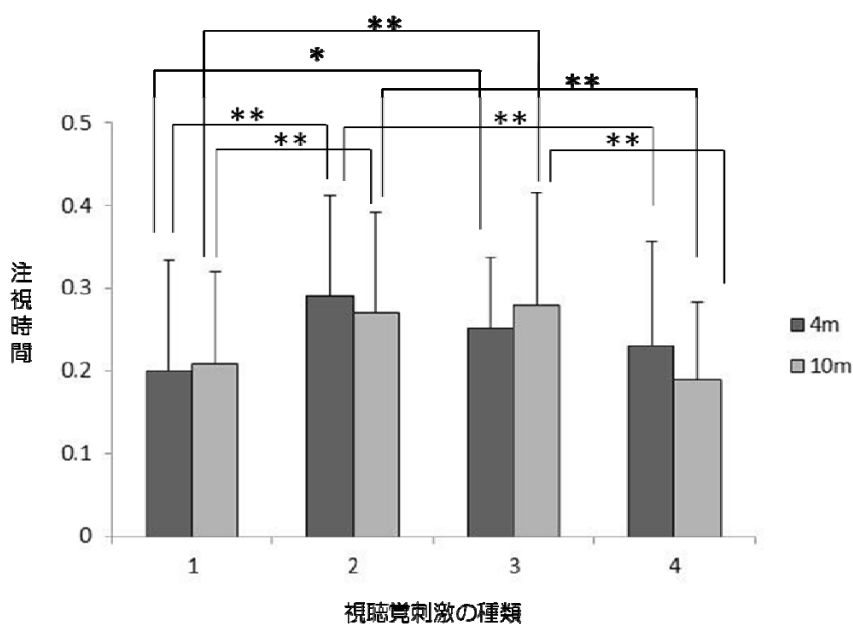


図4 キャラクターの「行為領域（A + B）」に対する注視時間

考 察

注視時間の分析が示すように、本研究に参加した乳児たちは、概して視聴覚刺激に登場するヒト型キャラクターの「顔」に相当する部分に高い関心を示しており、その傾向は、月齢が上がるとより顕著になっていた（10ヵ月齢群の方がより長く「顔」を注視する傾向があった）。また10ヵ月齢群に関しては、視聴覚刺激の違いによる「顔」への注視時間の長さにも変化がみられた。一方、「行為領域」に対する注視時間には、月齢の違いによる差はみられなかったが、各月齢群内では視聴覚刺激の違いによる注視時間の長さの変化が観察された。

ここで、どの視聴覚刺激の間で注視時間に差があったのかを改めて確認してみると、「顔」の場合には10ヵ月齢の「2」と「4」および「3」と「4」の間、そして「行為領域」の場合には、どちらの

月齢群でも共通して「1」と「2」, 「1」と「3」および「2」と「4」の間に差があり, これに10ヵ月齢群では「3」と「4」の間が追加されるというものであった。しかし, ここで挙げた刺激の組み合わせは, 先述した本研究の仮説を積極的に支持するものではなかった。すなわち, 「キャラクターの行為とそこから予測される音の特徴との対応が高い刺激（「1」と「2」）については, 乳児の視線はキャラクターの行為領域（手腕の動き）に集中し, 逆にキャラクターの行為と音の特徴との対応が低い刺激（「3」と「4」）については, （探索行動の増加に関連して）手腕だけではなく, 顔などの他の領域にも視線の方向が拡散する」という仮説が当てはまるとすれば, キャラクターの行為が同一であるにもかかわらず, 聴こえてくる音響的特徴が異なる「1」と「4」および「2」と「3」の間で注視時間の長さに差が生じ, かつ「顔」と「行為領域」への注視時間の長さについても, 一定の変化のパターンがみられる必要があったからである。

本研究において, 映像刺激の“視-聴”覚的特徴の一致/不一致が, 注視時間の変化に反映されなかったのは何故だろうか。例えば視聴覚間協応の発達に関する代表的な先行研究である Spelke (1976) および Spelke (1979) からも明らかのように, 視覚的事象と音響的刺激との統合的な知覚は4ヵ月齢児でも現れているはずであり, 本研究に参加した乳児の月齢に起因する問題であったとは考えにくい。そこで可能な理由として考えられるのは, 本研究で使用した映像刺激の構造が, 視聴覚間で冗長に処理されていたのではなく, 視覚的な処理により強く影響を与えるものであったのではないかということである。

その根拠のひとつは, 乳児がキャラクターの行為の変化をかなり精緻に追視していたということである。図5は, キャラクターの行為領域を「A」と「B」とに分け（図2参照）, それぞれの領域に対する乳児の注視時間の割合を視聴覚刺激の種類ごとにまとめたものである。どちらの月齢でも, 折れ線のプロフィールが対称的な形となっている。行為領域「A」は, キャラクターの胴体部分に位置しており, 主に打動作（垂直方向）を繰り返して行う際に手の部分がこの領域内に入っていた。打動作が呈示された「1」および「4」で, 領域「A」に対する注視時間の割合が増加していたということは, つまり乳児が打動作をするキャラクターの手の動きを注視していたことを示唆している。一方, 行為領域「B」は, キャラクターの左右両端に拡がっており, グリッサンドの動作（水平方向）が呈示される際にキャラクターの両腕がこの領域内に入っていた。グリッサンドの動作が呈示された「2」と「3」において, 領域「B」への注視時間の割合が増加していたということは, 領域「A」の場合と同様に, 乳児がグリッサンドの動作をするキャラクターの両腕の動きを注視していたことを示唆している。本研究に参加した乳児は, 音の特徴がどのようなものであれ, キャラクターの行為を追視し続けており, その結果, 視聴覚間での協応が仮説通りには生じなかったのではないだろうか。

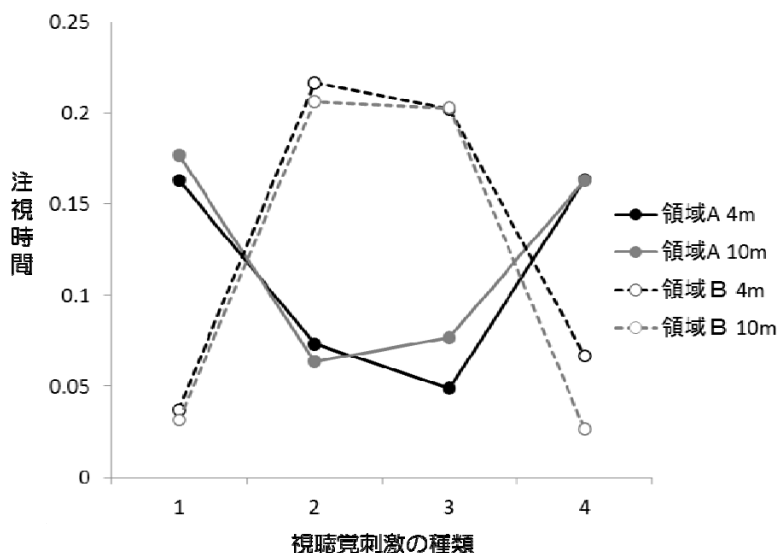


図5 行為領域「A」および「B」を分割して算出した注視時間の割合

これらの考察は、あくまで現段階における注視時間の分析のみをもとに述べたものであり、本研究にはまだ検証すべき課題と発展の可能性がいくつも残されている。例えば「顔」への注視時間が、月齢が上の群で長くなる傾向があり、かつ視聴覚刺激の組み合わせによって、その長さに差が生じていたという結果は（それが必ずしも仮説を支持する組み合わせではなかったとしても）、最近の「他者理解（目的的な行為を行う他者の意図の推測）の発達の基盤」に関する研究成果との関連をうかがわせるものである。Myowa et al. (2012) は、例えば「ヒトがコップに飲み物を注ぐ」といった他者の行為の意図を推測するような場面を見ているときの注視行動の特徴を、ヒトの成人と乳児（8および10 ヶ月齢児）そしてチンパンジーを対象にして比較した。そして、チンパンジーは「顔」にはあまり視線を向けずに、一貫してモノを見ていたのに対し、ヒトの乳児は、チンパンジーよりも他者の「顔」を長く見ていたということを報告している（成人は、他者による目的的な行為のゴールが達成された後 [例えばジュースを注ぎ終わった後] に「顔」への注意が減少したという）。こうした結果から、Myowa et al. (2012) は、ヒトが何らかの行為を行っている際の他者の「顔」をみるのは、他者の心の状態、すなわちその行為の意図に関する情報を抽出しようとするからであると指摘したのである。

例えば本研究の場合、10 ヶ月齢児の「顔」への注視時間に差があったケースには、視聴覚刺激の「2」と「4」の間、すなわち「音響的な特徴（グリッサンド音）は同一で、キャラクターの行為の向きが異なる」という組み合わせがひとつ含まれていた。Myowa et al. (2012) の知見を踏まえて、この点を解釈してみると、10 ヶ月齢児は“同一の音響刺激”が“異なる行為（視覚刺激）”から発せられているというある種のコンフリクトが生じた理由を探るために（キャラクターの意図を読み取ろうとするために）、行為と音の特徴との一致性が低いと考えられる「4」の刺激が呈示された際に、キャラクターの「顔」をより長くみていたということができないのではないだろうか。ただし本研究の場合、「顔」といっても、それはコンピュータで描いたアニメーションのキャラクターの「顔のような

特徴を備えた部分」というべきであり、厳密に言えば Myowa et al. (2012) のいう「他者の顔」ではなかったのかもしれない。しかし、こうした映像刺激の「顔」的な部分が、乳児に「他者性」を感じさせたり、行為の意図の手掛かりを探索するための参照先になっていたのだとすれば、それは乳児の他者理解の枠組みを人工物にまで拡張できる可能性を示唆する成果であったということができるとはならないだろうか。

以上のような本研究の可能性を精査し、データをより詳細に解釈するために、今後は映像刺激の各領域に対する注視時間の算出だけではなく、乳児の視線遷移のパターンを時系列に沿って明らかにし、刺激呈示の直後にはどの領域に視線が向けられていたのか、それが時間の経過とともに（あるいは複数回にわたる刺激呈示とともに）どのように変化していくのかといった、視覚的探索の動的変化なども指標として取り入れた分析も試みる予定である。また本研究で対象にした乳児の 2 つの月齢群（4 ヶ月齢と 10 ヶ月齢）の間を埋める月齢群（例えば 7 ヶ月齢）を設定した同様の検討を行い、「行為領域」に対する注視時間に 4 ヶ月齢と 10 ヶ月齢の間では差が明確に出なかったという結果が、「2 つの月齢群が同じだった」ということなのか、あるいはその間に何らかの質的な変化が生じていたのか、といった点についても明らかにしていきたい（7 ヶ月齢での注視行動が、本研究で取り上げた 2 つの月齢群のどちらも異なっていたとすれば、4 ヶ月齢から 10 ヶ月齢に至るまでの間に何らかの質的変化が生じていたことが推測できる）。このような計測や分析を追加していくことによって、現段階では十分に検証するに至らなかった当初の仮説と実際のデータとの整合性について改めて議論し、視聴覚間協応の発達の基盤についての理解を深めていきたいと考えている。

謝 辞

本研究の実施にあたり、ご協力いただいたお子様と保護者の皆様に深く感謝申し上げます。本研究の実施にあたり有益なご示唆をいただいた多賀厳太郎先生（東京大学）、そして映像刺激の作成およびデータの分析作業をお手伝いいただいた滝山聖士氏（公立はこだて未来大学）にもこの場を借りて御礼を申し上げます。最後になりましたが、本研究の実施を可能にいただいた「公益財団法人 発達科学研究教育センター」の学術研究助成事業に対しまして心より御礼を申し上げます。

引用文献

- Evans, K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10 (1): 6, 1-12.
- Gaver, W. W. (1993). What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, 5, 1, 1-29.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin Company.
- 佐々木正人・古山宣洋・三嶋博之 [監訳] 生態学的知覚システム論 感性をとらえなおす 東京大学出版会 2011 年

- Gibson, J.J. (1979). *The ecological approach to visual perception*. Lawrence Erlbaum Associates, Inc. 古崎 敬・古崎愛子・辻 敬一郎・村瀬 旻 [訳] 生態学的視覚論 ヒトの知覚世界を探る サイエンス社 1985年
- Lewkowicz, D. (1996). Infants' response to the audible and visible properties of the human face: 1. role of lexical-syntactic content, temporal synchrony, gender, and manner of speech. *Developmental Psychology*, 32, (2), 347-366
- Maruyama, S., Watanabe H., & Taga, G. (2011) Spatial awareness induced by pitch direction: A study on the developmental origin of audio-visual congruence in target detection task. *The 15th European Conference on Developmental Psychology*. Bergen, Norway 23-27 August.
- 丸山 慎・渡辺はま・多賀巖太郎 (2011) 高い音は上? : 乳児の視線計測からのアプローチ 日本認知科学会第 28 回大会発表論文集 130-133
- Mugitani, R., Kobayashi, T., & Hiraki, K. (2011). Robust right-side bias of infants' audiovisual matching in lips and voice for rapidly changing sound. *認知科学*, 18, 402-415
- Myowa-Yamakoshi, M., Scola, C., & Hirata, S. (2012) Humans and chimpanzees attend differently to goal-directed actions. *Nature Communications*, 10.1038/ncomms1695
- Pick, A. D., Gross, D., Heinrichs, M., Love, M., & Palmer, C. (1994). Development of perception of the unity of musical events. *Cognitive Development*, 9, (3), 355-375
- Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, 99, 113-129
- Shove, P. & Repp. H. B. (1995). Musical motion and performance: Theoretical and empirical perspectives. In Rink, J. (Ed.), *The Practice and Performance: Studies in Musical Interpretation*, 55-83. Cambridge University Press.
- Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, 8, 553-560
- Spelke, E. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, 15, 626-636
- Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development*, 52, 728-731
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, 21 (1), 21-25

