

発話と映像的身振りの統合的理解における聞き手の視線

— 幼児と成人の比較 —

Where is Listener Watching at the Speaker in Understanding the Message Integrated of Speech and Iconic Gesture? :

Comparison of Young Children and Adults.

三宅 英典¹⁾

Hidenori MIYAKE

関根 和生²⁾

Kazuki SEKINE

問題と目的

私たちはコミュニケーション場面において、話し手の発話や言語情報だけに注意しているわけではない。話し手は、しばしば発話とともに身振りをを用いて、メッセージを伝達している。このような発話に伴う身振りは、自発的身振りと呼ばれる。自発的身振りとは発話は、時間的にも意味的にも同期して産出されるが、自発的身振り（以下では単に「身振り」と表記）には、しばしば発話には表されていない情報が表されている（McNeill, 1992）。例えば、「昨日魚釣りに出かけて、その時に大きな魚が釣れてね…」と言いながら魚の大きさを身振りで示したとしよう。この時、発話は、釣った魚が大きかったことを伝達しているが、具体的な魚の大きさは示されていない。一方、身振りは、その魚の大きさを具体的に表現しており、聞き手はこのような発話と身振りの情報をもとに「そんなに大きな魚を釣ったのか」と解釈することが可能になる。このように、身振りとそれに同期する発話は、

同じ指示対象に関する情報を伝達しつつも、指示対象の異なる側面を表すことがある。

上記のように、指示対象についてそれぞれ異なる側面を伝達する発話と身振りをもとに、聞き手が統合的にメッセージを理解することを統合的理解と呼ぶ。先の例で挙げた魚の大きさを表すような身振りは映像的身振りと呼ばれ、事物や動作のような視覚的な特徴を表現する。この発話と映像的身振りの統合的理解は、幼児期に発達していくことが報告されている（Sekine, Sowden, & Kita, 2015; Stanfield, Williamson, & Ozçalışkan, 2015）。

Sekine et al. (2015) は、3歳児・5歳児・成人を対象に、日常的な動作を表すメッセージを発話と身振りで提示し、その後、メッセージと一致する選択肢を4枚の写真から選ばせた。例えば、「投げる」という動作を表すメッセージでは、参加者に「投げています」と言いながらチェストパスの身振りを行っているモデルの動画を提示した。その動画の後に、4枚の写真を提示し、そこから、モデルの伝達するメッセージと最も一致するものを選択させた。参加者が4つの選択肢から、モデルが表す「投げる」という発話情報と「動作（投

1) 現代子ども教育学科

2) 早稲田大学人間科学学術院

げ方)」という身振り情報の両方と一致する選択肢を選んだ時に、発話と身振りを統合したと判断した。その結果、5歳児は、発話と身振りを成人と同程度に統合することができた。つまり、一つの指示対象の異なる側面を意味する身振りと発話の情報を統合させることができた。一方、3歳児は両者を統合することが難しく、発話情報に依存することが明らかになった。ただし、3歳児は発話と身振りをビデオではなく実演で提示すると、両者を統合することができた。

Sekine et al. (2015) では、発話と身振りの統合能力が幼児期に発達しているだけでなく、実際にこの能力に基づく統合的理解が、実演もしくはビデオのように状況の影響を受けることが明らかにされている。これらをもとに、Miyake & Sugimura (2018) は、発話と身振りの提示時に、発話に身振りを参照する指示詞を加える条件を新たに設定して、両者の統合的理解の検討を行なった(例:「こうやって投げます」と言いながらチェストパスの身振りをする)。その結果、指示詞が加わることで、幼児の発話と身振りの統合的理解が促進されることを明らかにした。

指示詞が発話と身振りの統合的理解を促進するメカニズムとして、聞き手の身振りに対する注意の促進が考えられる。Hamasaki (2002) によると、指示詞の理解は2歳半から3歳にかけて可能になる。そのため、3歳児以降は、指示詞によって身振りに対する注意を高められていたと考えられる。もし、身振り情報に関する指示詞の作用が、身振りへ注意を向けさせることであるならば、指示詞は、語用論的な作用によって身振りが発話情報と同一の指示対象の別側面を指していることを聞き手に明示していると考えられる。また聞き手の注意は、身振りに対する視覚的注意にも反映されている可能性がある。

Kelly, Özyürek, & Maris. (2010) は、統合的理解の処理過程における発話と身振りの関係性を検討している。彼らは2種類の動画を用いて、実験参加者に両者の情報が一致するかどうかを判断させた。具体的には、1つ目の動画で、ある動作を行っている動画を提示した(例:野菜を包丁で切っている)。2つ目の動画では、1つ目の動画に一致する(あるいは一致しない)情報を、発話と身振りで表した動画で提示した(例:一致条件では、「切る」と発話しながら包丁を使う身振りをする。不一致条件では「切る」と発話しながらハサミを使う身振りをする)。実験参加者は、これらの動画を順に視聴し、2つ目の動画における発話と身振り情報(あるいは発話情報のみ)が、1つ目の動画に対して一致するかどうかを判断するよう求められた。その結果、統合的理解において身振りと発話が、相互の理解に影響を与えあっていることが明らかになった。そのため、身振りに対する視覚的注意においても、統合的理解における発話の理解に影響を与えている可能性がある。

また、三宅 (2019) では、幼児の発話と身振りの統合は、成人よりも身振りを意図的に発話と統合させる必要性があることを示唆している。そのため、幼児にとっては、成人よりも身振りへ意図的に注意を向けなければ、両者の統合が難しいと考えられる。

しかしながら、発話と身振りの統合的理解において、聞き手の話者に対する視覚的注意の様相は明らかにされておらず、幼児と成人で視覚的注意に差があるかどうか不明である。そのため、本研究は、指示詞による統合的理解の促進メカニズムを明らかにするための予備調査として、Sekine et al. (2015) のパラダイムを用いて、発話と身振りの統合過程における聞き手の視覚的注意を測定し、話者に対する視覚的注意を幼児と成人で比較する

ことを目的とする。

方法

参加者

参加者は、6歳児21名（女児9名、平均6歳2ヶ月）と女子大学生21名（平均19名）であった。

課題の材料と刺激

実験には、Sekine et al. (2015) のパラダイムを使用した。このパラダイムでは、実験参加者に視覚刺激と聴覚刺激を提示し、その後、刺激に対する反応を4つの選択肢から選択させた (Figure 1)。本研究では、このパラダイムに沿って作成された Miyake & Sugimura (2019) の提示刺激と選択肢の写真を使用した。

まず、実験課題の材料には、日常的な動作を表す8つのメッセージを用いた（書く・投げる・乗る・読む・飲む・開ける・食べる・登る）。これらの動作は、3歳児の80%以上が理解している (Watamaki & Ogura, 2004) と判断される動詞から選出された。メッセージの視覚情報では、手腕を使って動作の仕草

を表現する身振りを提示した。聴覚情報では、動作を表す動詞を「○○ます」と表現した発話を提示した。メッセージは、女性のモデルが椅子に座った状態で表現した。その際、モデルには常にマスクを着用してもらった。これは、参加者が、モデルの口や唇の動きから、発話内容を推測させないようにするためであった。各動作のメッセージは、それぞれビデオカメラを用いて録画した。動画はそれぞれ約5秒間であった。

上記の材料をもとに、提示刺激には3つの条件を参加者内要因として設定した。1つ目はモデルの身振りと言話の両方を提示する発話一身振り条件 (Verbal-Gesture 条件；VG 条件)、2つ目はモデルの静止画 (Figure 1の①) と発話を提示する発話条件 (V 条件)、3つ目は発話をミュートにしてモデルの身振り動画を提示する身振り条件 (G 条件) であった。

次に、提示刺激に対する反応を確かめるために使用する写真 (選択肢) を作成した (Figure 1の右側)。4つの選択肢は、それぞれ次の特徴に分けて作成した (①モデルが発した発話情報のみと一致する発話選択肢、②身振り情報のみと一致する身振り選択肢、③

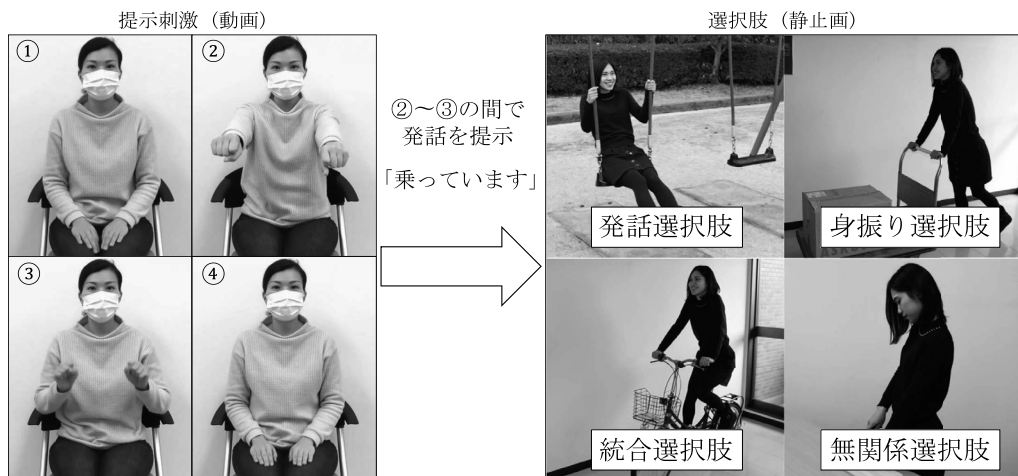


Figure 1 提示刺激と刺激に対する選択肢の例（乗る動作）

発話と身振りの情報を組み合わせた時に一致する統合選択肢, ④発話と身振りのどちらの情報とも一致しない無関係選択肢)。

実験手続き

実験課題について、大学生の参加者は静かな個室で実施した。幼児は自身が通う幼稚園のホール室の一部を仕切りで区切って実施した。実験刺激の提示には、14インチのラップトップを用いた。すべての参加者はヘッドフォンを着用して課題に取り組んだ。刺激を見ている最中の視線を測定するために、視線計測の機材にトビー・テクノロジー社製 Tobii nano を使用し、刺激の提示ならびに計測した視線の分析ソフトウェアに同社製 Tobii Pro ラボを使用した。

まず、実験参加者には課題についての説明を行い、その後、眼球運動を計測するためのキャリブレーションを行なった。その後、実験刺激とそれに対応する選択肢の提示を行

なった (Figure 2)。刺激は、「書く・投げる・乗る・読む・飲む・開ける・食べる・登る」の順で提示し、そのうち「書く・投げる」の2動作は練習試行とした。

練習試行の「書く・投げる」ではV条件・G条件・VG条件の順で刺激を提示した(計6試行)。その後、参加者が課題を理解していることを確かめてから本試行を実施した。本試行は、6つの動作を3条件で提示した(計18試行)。3つの提示条件(VG条件、V条件、G条件)と選択肢の写真の配置はそれぞれカウンターバランスをとった。

刺激から選択肢までの提示は、ソフトウェアの制御によって自動的に行われた。ただし、選択肢から次の刺激に進む操作は実験者が手動で行った。

分析方法

発話と身振りの統合能力の分析では、統合能力の算出方法としてMulti Modai Gain Score

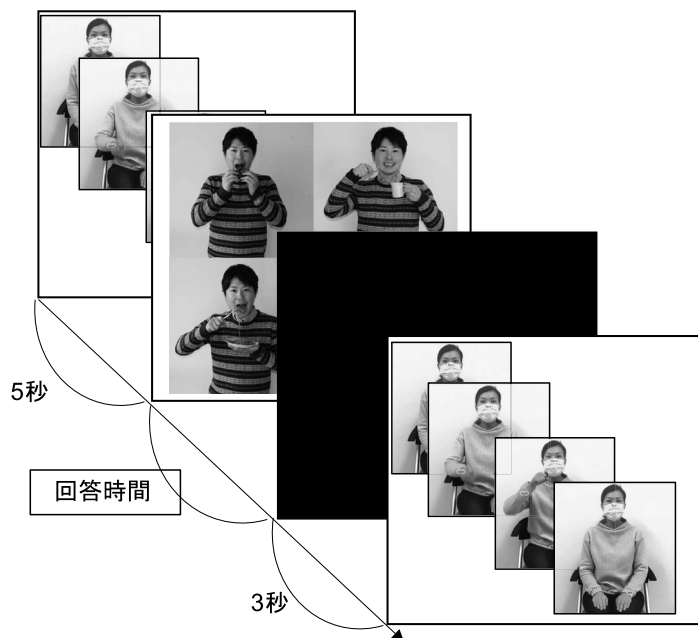


Figure 2 キャリブレーション後に実施した実験刺激の提示方法

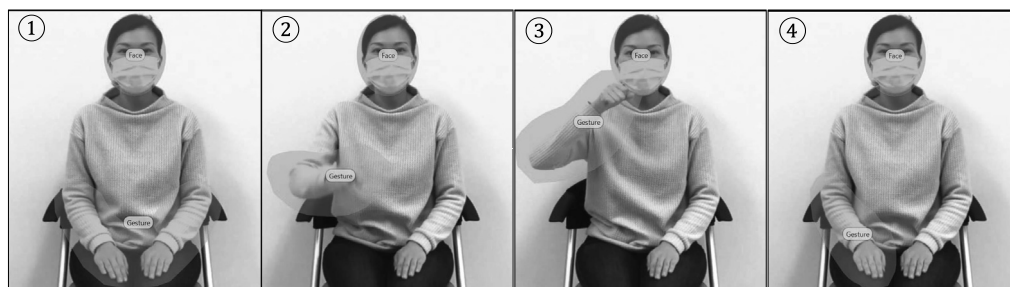


Figure 3 顔と身振りに対する Area of Interest (AOI) の例（食べる動作）

(MMG 値) を使用した (Sekine et al., 2015)。統合選択肢は、発話あるいは身振りのどちらか一方を理解している場合でもチャンスレベルを超えて選択される可能性がある。ただし、聞き手が発話と身振りを統合的に理解している場合、統合選択肢の選択割合は、VG 条件が他の条件と比べて最も高くなると考えられる。このような考え方に基づいて、MMG 値は以下のようにして算出する。

MMG 値 = (VG 条件における統合選択肢の選択割合) - (V 条件と G 条件で統合選択肢の選択割合の高い方の値)

聞き手が発話と身振りを統合的に理解している場合、MMG 値は 1 ~ 0 の値をとり、両者を統合していない場合は 0 ~ -1 の値をとる。本研究では、MMG 値が 0 を越えた参加者は、発話と身振りを統合的に理解したとみなし統合有群とした。MMG 値が 0 以下だった参加者は、発話と身振りを統合的に理解しなかったとみなし統合無群とした。

視線の分析では、提示刺激の顔と身振りにそれぞれ Area of Interest (AOI) を設定した (Figure 3)。顔と身振りの AOI 内で注視した頻度と、注視時間の合計時間をそれぞれ算出した。注視の定義は、Tobii Pro Lab 内の I-VT フィルタ (Olsen, 2012) に基づき、60ms 以上の停滞を 1 回として計測した。

結果

課題の正答率と MMG 値

各条件の正答率と MMG 値を参加者ごとに算出した (Table 1)。それぞれの条件における正答の選択肢は、V 条件が発話選択肢と統合選択肢、G 条件が身振り選択肢と統合選択肢、VG 条件が統合選択肢であった。そのため、V・G 条件のチャンスレベルは 50%、VG 条件のそれは 25% となる。

課題の正答率に対してチャンスレベルの検定を行なった。その結果、すべての条件において、どちらの参加者もチャンスレベルを 1% 水準で有意に上回った。次に、MMG 値が 0 を越えているかを判断するためのチャンスレベルの検定を行なった。その結果、成人はチャンスレベルを 1% 水準で有意に上回ったが、6 歳児は有意傾向であった ($t(20) = 1.98, p = .061$)。

VG 条件の正答率に対して、参加者 (6 歳児, 成人) の t 検定を行なったところ、群間に有意な差がみられ、課題の正答率は成人の

Table 1 条件ごとの課題の正答率と MMG 値の平均値 (SD)

	6歳児	成人
V	0.79 (0.20)	1.00 (0)
G	0.87 (0.14)	0.94 (0.10)
VG	0.65 (0.22)	0.94 (0.11)
MMG	0.10 (0.24)	0.19 (0.17)

方が6歳児よりも高かった ($t(40) = 5.41, p < .01$)。一方, MMG値に対して同様の t 検定を行なったところ, 群間に有意差はみられなかった。

話者のメッセージを見ている最中の聞き手の視線

メッセージを見ている最中の視線について, 各条件をAOIごとに分け, それぞれで注視の頻度と注視時間を算出した (Table 2)。

VG条件における注視時間と注視頻度の分析

注視時間に対して, 参加者 (6歳児, 成人)

× 統合 (無, 有) × AOI (顔, 身振り) の三要因分散分析を実施した (Figure 4)。その結果, 参加者 × 統合 × AOIの交互作用と, AOIの主効果が有意であった (順に, $F(1, 38) = 4.90, p < .05, partial \eta^2 = .11$; $F(1, 38) = 9.65, p < .01, partial \eta^2 = .20$)。参加者 × 統合 × AOIの交互作用について, Holm法の多重比較を行ったところ, 6歳児は統合無群で顔の注視時間が身振りよりも長かったのに対して, 成人では, 統合有群で顔の注視時間が身振りよりも長かった。顔の注視時間について, 6歳児では統合の有無による差はみられなかったが, 成人では統合有群の方が無群より

Table 2 条件ごとのAOIに注視した時間 (秒) と頻度 (個) の平均値 (SD)

条件	注視	AOI	6歳児	成人
VG	時間	顔	1.62 (0.78)	1.67 (1.09)
		身振り	0.98 (0.47)	1.06 (0.60)
	頻度	顔	2.94 (1.05)	3.06 (1.98)
		身振り	2.68 (1.18)	3.10 (1.57)
V	時間	顔	1.92 (0.95)	1.96 (1.31)
		身振り	0.22 (0.19)	0.24 (0.22)
	頻度	顔	3.23 (1.14)	3.17 (2.47)
		身振り	0.80 (0.71)	0.88 (0.76)
G	時間	顔	1.78 (0.83)	1.64 (1.08)
		身振り	1.02 (0.57)	1.16 (0.60)
	頻度	顔	2.98 (1.10)	3.09 (2.04)
		身振り	2.85 (1.23)	3.22 (1.48)

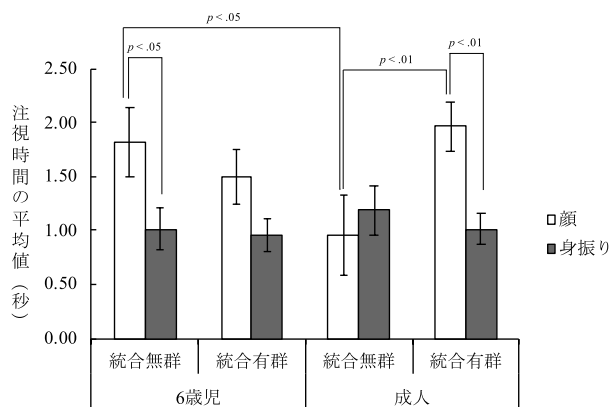


Figure 4 VG条件における統合の有無とAOI別に分けた注視時間の平均値

も長かった。また、6歳児と成人の間では、統合無群の6歳児が統合無群の成人よりも顔の注視時間が長かった。身振りの注視時間では、いずれも差がみられなかった。AOIの主効果については、顔の注視時間が身振りよりも長かった。

次に、注視頻度に対して、参加者（6歳児、成人）×統合（無、有）×AOI（顔、身振り）の三要因分散分析を実施した（Figure 5）。その結果、統合×AOIの交互作用のみが有意であった（ $F(1, 38) = 6.38, p < .05, \text{partial } \eta^2 = .14$ ）。Holm法による多重比較を行なったところ、統合有群の参加者は、顔に対する注

視頻度が無群よりも多かった。

V条件における注視時間と注視頻度の分析

注視時間に対して、参加者×統合×AOIの三要因分散分析を実施した（Figure 6）。その結果、統合×AOIの交互作用、AOIと統合の主効果が有意であった（順に、 $F(1, 38) = 5.86, p < .05, \text{partial } \eta^2 = .13$ ； $F(1, 38) = 43.41, p < .01, \text{partial } \eta^2 = .66$ ； $F(1, 38) = 2.99, p < .05, \text{partial } \eta^2 = .12$ ）。

統合×AOIの交互作用について、Holm法による多重比較を行なったところ、顔の注視時間は、統合有群の方が無群よりも長かった。

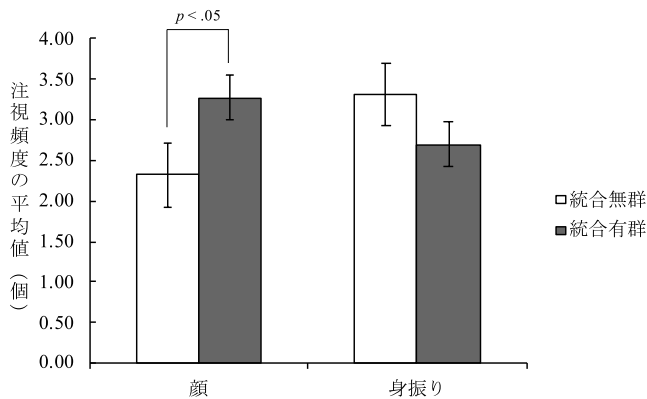


Figure 5 VG条件における統合の有無とAOI別に分けた注視時間の平均値

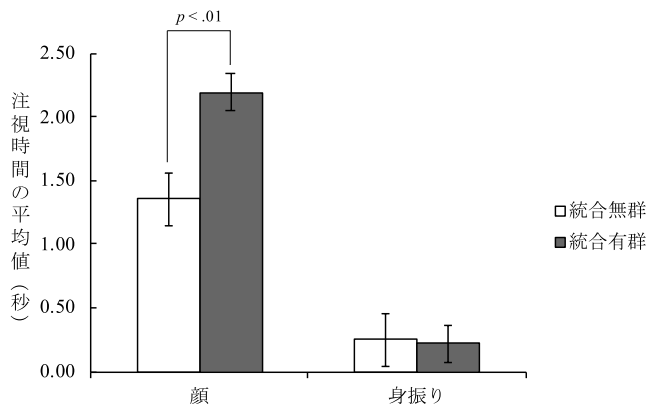


Figure 6 V条件における統合の有無とAOI別に分けた注視時間の平均値

AOIの主効果については、顔の注視時間が身振りよりも長かった。統合の主効果については、統合有群の注視時間が無群より長かった。

次に、注視頻度について参加者×統合×AOIの三要因分散分析を実施した (Figure 7)。その結果、参加者×統合×AOIの交互作用、統合×AOIの交互作用、AOIと統合の主効果がそれぞれ有意であった (順に、 $F(1, 38) = 9.76, p < .05, partial \eta^2 = .12$; $F(1, 38) = 9.57, p < .05, partial \eta^2 = .12$; $F(1, 38) = 77.34, p < .01, partial \eta^2 = .52$; $F(1, 38) = 4.31, p < .05, partial \eta^2 = .10$)。

参加者×統合×AOIの交互作用について、Holm法による多重比較を行なったところ、AOI別の注視頻度について、統合無群の成人では、顔と身振りの注視頻度に差が見られず、そのほかの群では顔の注視頻度が身振りよりも多かった。顔の注視頻度について、成人では、統合有群の方が無群よりも多かった。また、統合無群に限り6歳児は成人よりも顔の注視頻度が多かった。

次に、統合×AOIの交互作用についてHolm法による多重比較を行なったところ、AOIの注視頻度について、統合の有無にかかわらず顔の方が身振りよりも多かった。また、顔の注視頻度では、統合有群が無群よりも多

かった。AOIの主効果については、顔の注視頻度が身振りよりも多かった。統合の主効果については、統合有群の方が無群よりも注視頻度が多かった。

G条件における注視時間と注視頻度の分析

注視時間に対して、参加者×統合×AOIの三要因分散分析を実施した。その結果、AOIの主効果のみが有意であった ($F(1, 38) = 8.74, p < .01, partial \eta^2 = .19$)。具体的には、顔の注視時間が身振りよりも長かった。注視頻度について参加者×統合×AOIの三要因分散分析を行なったところ、有意差はみられなかった。

聞き手の視線とMMG値の相関分析

6歳児と成人に分けて、VG条件の注視時間・注視頻度とMMG値の間で相関分析を行なった (Table 3, 4)。その結果、6歳児と成人の両方で、顔に対する注視頻度と顔に対する注視時間、身振りに対する注視頻度と身振りに対する注視時間に正の相関がみられた。MMG値と視線の関連について、6歳児は有意な相関はみられず、成人は顔に対する注視頻度とMMG値が有意傾向で正の相関を示した ($p = .054$)。

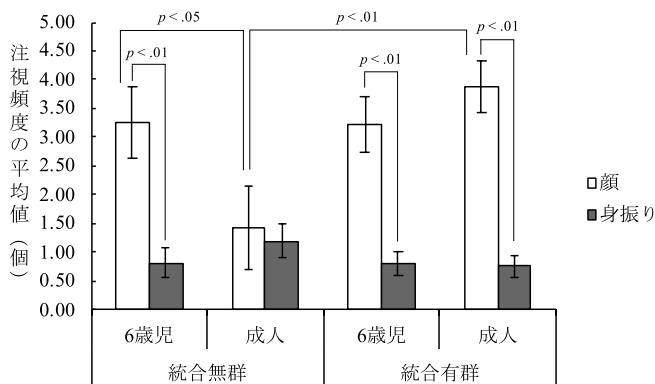


Figure 7 V条件における統合の有無とAOI別に分けた注視頻度の平均値

考 察

本研究は、指示詞による統合的理解の促進メカニズムを明らかにするための予備調査として、発話と身振りに対する聞き手の視覚的注意の様相を明らかにし、幼児と成人で比較することが目的であった。具体的には、発話と身振りを組み合わせた実験刺激を提示し、刺激に対する実験参加者の視線を計測して分析を行なった。

課題の正答率と統合能力

t 検定において6歳児と成人でMMG値に差はみられなかったものの、成人では発話と身振りからメッセージを統合的に理解していたが、6歳児では両者を統合したメッセージ理解を十分に行えるようになっていたとは言いがたい。実際に、課題の正答率においても差がみられた。実験刺激を実演ではなくビデオとして提示した場合、幼児のMMG値が低下する傾向にあることは、Sekine et al. (2015) や Miyake & Sugimura (2018) の結果とも一

致している。このことは、幼児は発達にともなって発話や身振りの提示される媒体の影響が小さくなり、やがて提示媒体に関わらずに両者を容易に統合することができるようになることを示唆している。

発話と身振りを見る聞き手の視線

聞き手は、話者のメッセージを理解する際に、身振りよりも顔の方を長く見ていることが明らかになった。参加者ごとにみると、発話と身振りを統合しなかった6歳児は顔の方を身振りよりも長く見ていた。一方、成人は両者を統合した参加者の方が、顔の方を身振りよりも長く見ていた。次に、AOIごとの注視頻度に注目すると、全体的な注視頻度では参加者間や統合の有無による違いはみられなかったが、顔の注視頻度は、発話と身振りを統合した参加者の方が、統合しなかった参加者よりも多かった。

顔に対する視覚的注意について、顔は、他の対象よりも優位に検出されやすく (Langton, Law, Burton, & Schweinberger, 2008), 行為者

Table 3 6歳児におけるAOI別の注視時間・注視頻度とMMG値の相関

		注視時間		注視頻度	
		顔	身振り	顔	身振り
注視時間	顔				
	身振り	.205			
注視頻度	顔	.480 *	.138		
	身振り	.001	.762 **	.210	
MMG値		-.138	-.165	-.198	-.365

** $p < .01$, * $p < .05$, + $p < .10$

Table 4 成人におけるAOI別の注視時間・注視頻度とMMG値の相関

		注視時間		注視頻度	
		顔	身振り	顔	身振り
注視時間	顔				
	身振り	.004			
注視頻度	顔	.694 **	-.006		
	身振り	-.205	.847 **	.002	
MMG値		.340	-.187	.427 +	-.177

** $p < .01$, * $p < .05$, + $p < .10$

の視線はメッセージ理解において有効な手がかりになる (Nummenmaa, Hyönä, & Hietanen., 2009) ことが報告されている。そのため、顔の方が身振りよりも手がかりとして優位に検出されることが、注視時間の増加につながったと考えられる。

また、親松・平山・勝二 (2014) は、手話の読み取りにおいて、聾者や手話上級者になるほど、手指動作範囲への注視が減少し、顔領域に対する注視頻度が高まることを明らかにしている。そして、手話の熟達者ほど、手指動作で伝達される情報を周辺視野で捉えていることを示唆した。Buckley, Codina, Bhardwaj, & Pascalis. (2010) は、手話経験を積むほど、視野の下方と側方領域での識別閾が高くなることを報告している。このことから、発話と身振りの統合的理解においても手話の理解と同様に、両者を安定して統合することのできる参加者ほど、顔に注視することを維持しながら身振りを視野でとらえていると考えられる。

なぜ、発話と身振りの統合的理解において顔に視覚的注意を向ける必要があるのだろうか。この背景には、統合処理における情報の優位性が関係していると考えられる。Sekine et al. (2015) は、統合に失敗した場合、メッセージの理解として発話情報を優位に処理することを示している。これはメッセージ理解において発話情報の処理の優先度が高いことを示唆し、統合処理の過程においても発話が身振りよりも優先的に処理されている可能性を示している。顔に対する視覚的注意は、この統合処理における発話情報の処理を反映していると考えられる。

ただし、6歳児では、発話と身振りを統合しなかった参加者の方が、統合した参加者よりも顔の注視時間が長かった。先述したように、身振り情報の多くは、顔に注視している

状態を維持しながら周辺視野でとらえられていると考えられる。統合できなかった6歳児は成人と比べて顔に注意を向けている時に、身振りを視野でとらえることが難しいのかもしれない。そのため、身振りに注意を向けられない分、顔を長く見ていたのかもしれない。

次に、発話のみを提示したV条件の注視時間に着目すると、参加者間や統合の有無にかかわらず、顔の注視時間が身振りよりも長かったが、発話と身振りを統合した参加者の方が、統合しなかった参加者よりも顔の注視時間が長かった。そのため、統合に失敗した幼児のように、情報を補うために顔に注意を高める場合があるものの、概ね発話と身振りを統合することのできる聞き手は、統合できない聞き手よりも顔に対する注意が高いと考えられる。

次に、V条件の注視頻度に着目すると、発話と身振りを統合しなかった聞き手では、6歳児の方が成人よりも顔の注視頻度が多かった。この結果は、発話と身振りの統合が難しい幼児は、発話と身振りの統合場面に限らず、顔に対する注意が高いことを示唆している。そのため、幼児の顔に対する注意は、成人にかけてU字曲線を描きながら発達している可能性がある。これは、幼児が発話と身振りを統合することができるようになってくると、徐々に顔以外にも注意を向けられるようになることを示唆し、やがて、両者を安定的に統合できるようになるにつれて、顔に対する注意を維持しながら身振り情報をとらえているのかもしれない。

また、G条件は、発話が提示されず身振りの動作のみを提示する刺激であったが、参加者は身振りよりも顔の方に注意を向けていた。そのため、顔に対する視覚的注意は、たとえ静止画であったとしても、身振りの動作よりも注意が向きやすいことが明らかになった。

最後に、身振りに対する視覚的注意は、6歳児や成人、統合の有無に関わらず、有意差がみられなかった。そのため、身振りに対する注視は、発話と身振りの統合処理とは関係していない可能性が示唆された。

統合能力と視線の関係

6歳児と成人の両方で、顔や身振りの注視頻度と注視時間にそれぞれ正の相関がみられた。注視時間は、AOIを注視する時間の合計であるため、これらの間に関連がみられたと考えられる。

MMG値は、成人では顔の注視頻度と有意傾向で正の相関を示し、6歳児では有意な関連がみられなかった。この結果は、発話と身振りを統合した聞き手が、統合しなかった聞き手よりも、顔に注意を向けていたことと関連していると考えられる。そのため、有意傾向ではあるものの、顔に対する視覚的注意は、発話と身振りの統合能力の発達と関連することを示唆していると考えられる。

本研究の限界点と今後の課題

本研究の成果は、発話と身振りの統合的理解における、聞き手の視覚的注意の様相を明らかにしたことである。具体的に、発話と身振りを統合した聞き手が、統合しなかった聞き手よりも、顔に対する注意が高いことが明らかになった。そのため、顔に注意を向けることが発話と身振りの統合において重要な要因であることが示唆される。次に、発話と身振りを統合しなかった6歳児と成人では、顔に対する視覚的注意に違いがみられた。そのため、6歳児と成人では発話や身振りに対する視覚的注意の総量だけでなく、身振りに注意を向けるタイミング、発話と身振りへの注視の仕方についても更なる検討が必要であると考えられる。

本研究においては次の限界点が挙げられる。まず1点目として、本研究で設定したAOIは、顔や身振りの周囲をある程度含めた設定であった。顔や身振りのAOIは、それらの周囲の範囲を含めたものであったため、参加者は顔や身振りのそれぞれの方向に注視をしていたが、これらの注視がすべて顔や身振りを見ていたとは言い切れない。そのため、今後は、顔や身振りの周囲を含まずにAOIを設定する必要があるだろう。

2点目として、聞き手が発話と身振りを見たあとに行う選択行動では視線を測定していないことであった。本研究では、話者が表す発話と身振り（メッセージ）を見ている最中の聞き手の視線に限定したものであったが、本課題ではその後、メッセージに最も一致する写真を選択させている。このような選択行動は、両者の統合的理解に基づいて行われているため、聞き手の視線にも両者の統合処理の影響がみられると考えられる。発話と身振りを見ている最中の視線は、聞き手に入力される情報の処理過程に関わるものであると考えられるが、選択行動における視線は、聞き手の統合的理解に基づく出力の処理過程に関わっているものと考えられる。そのため、今後は、聞き手が写真を選択している最中の視線も検討することで、統合処理における視線の様相を聞き手の情報の出力という点から明らかにすることができると考えられる。

最後に、今後の展望として、発話と身振りの統合処理における指示詞の影響を、聞き手の視覚的注意から検討することが挙げられる。発話と身振りの統合と身振りに対する視覚的注意が見られなかったものの、「これくらい」や「こうやって」といった指示詞は、語用論的に指示対象である身振りに注意を促すものであるため、この作用が視覚的注意に現れる可能性は未だ残されている。もし、指

示詞が発話と身振りの統合を促すメカニズムとして、視覚的注意を促進しているのであれば、指示詞が語用論的に身振りを指示するものであるため、身振りに対する視覚的注意が高まると予測される（あるいは、視覚的注意と語用論的な認知処理の両方で身振りを強調していると考えられる）。一方、指示詞が身振りに対する視覚的注意に影響を与えないのであれば、指示詞の統合的理解を促進するメカニズムが視覚的注意の促進ではなく、聞き手の語用論的な処理において身振りを強調することにあると考えられる。これらの仮説を検証するために、今後は指示詞の有無を設定して、発話と身振りの統合と視覚的注意の関連性を検討することが求められる。

文 献

- Buckley, D., Codina, C., Bhardwaj, P., & Pascalis, O. (2010). Action video game players and deaf observers have larger Goldman visual fields. *Vision Research*, **50**, 548-556.
- Cocks, N., Morgan, G., & Kita, S. (2011). Iconic gesture and speech integration in younger and older adults. *Gesture*, **11**, 24-39.
- Cocks, N., Sautin, L., Kita, S., Morgan, G., & Zlotowitz, S. (2009). Gesture and speech integration: An exploratory study of a man with aphasia. *International Journal of Language and Communication Disorders*, **44**, 795-804.
- Hamasaki, N. (2002). The timing shift of two-year-olds' responses to caretakers' yes/no questions. In: Shirai, Y., Kobayashi, H., Miyata, S., Nakamura, K., Ogura, T. & Sirai, H. (Eds.). *Studies in Language Sciences (2)- Papers from the Second Annual Conference of the Japanese Society for Language Sciences*, 193-206.
- Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, **21**, 260-267.
- Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, **107**, 330-342.
- McNeill, D. (1992). *Hand and mind*. Chicago, IL: University of Chicago Press.
- 三宅英典. (2019). 発話と映像的身振りの統合的理解における指示後発話の効果：発話と身振りを統合する能力と課題の成績に着目して. *心理学科学*, **40**, 1-12.
- Miyake, H., & Sugimura, S. (2018). The effect of directive words on integrated comprehension of speech and iconic gestures for actions in young children. *Infant and Child Development*, **27**, doi.10.1002/icd.2096
- Nummenmaa, L., Hyönä, J., & Hietanen, J. K. (2009). I'll walk this way: Eye reveal the direction of locomotion and make passerby look and go the other way. *Psychological Science*, **20**, 1454-1458.
- Olsen, A. (2012). *The tobii I-VT fixation filter: Algorithm description*. White Paper on Tobii.
- 親松紗知・平山太市・勝二博亮. (2014). 視線パターンからみた手話読み取り能力—手話熟達度の違いによる検討—. *茨城大学教育学部紀要 (教育科学)*, **63**, 183-194.
- Sekine, K., Sowden, H., & Kita, S. (2015). The development of the ability to semantically integrate information in speech and iconic gesture in comprehension. *Cognitive Science*, **39**, 1855-1880.
- Stanfield, C., Williamson, R., & Ozçalışkan, S. (2014). How early do children understand gesture-speech combinations with iconic gesture? *Journal of Child Language*, **41**, 462-471.
- Watamaki, T., & Ogura, T. (2004). *Technical manual of the Japanese MacArthur communicative development inventory: Words and grammar*. Kyoto: Kyoto International Social Welfare Exchange Center.