

「素材質感」の視聴覚統合の脳内機構とその発達過程

中央大学研究開発機構* 氏 家 悠 太

The brain function of audiovisual material perception and its developmental process.

Research and Development Initiative, Chuo University, Ujiie, Yuta

要 約

質感を持つ「物体」の認識は、「木」や「金属」を叩いている時の視覚映像と音が一致することにより成立する (e.g., Fujisaki et al., 2014)。本研究では、近赤外分光法 (NIRS) を用いて、素材質感の視聴覚統合に関わる脳内機構の発達過程について、言語獲得直前の乳児 (実験 1) と成人 (実験 2) を対象に検討した。実験では、Fujisaki et al. (2014) で使用された木と金属の質感の CG 動画を用いて、木や金属を叩く音とその表面材質 (視覚刺激) が一致した刺激と不一致の刺激を提示し、左右側頭領域の脳活動を測定した。実験の結果、乳児 (実験 1) と成人 (実験 2) で共通して、素材の音と視覚が一致した刺激を観察した際、右側頭領域における脳血流の有意な増加が見られた。これらの結果から、質感を持つ素材の視覚と聴覚の統合の発達には、上側頭溝を含む右側頭領域の活動が関与することが示された。

【キー・ワード】 視聴覚統合, 質感知覚, 近赤外分光法, 上側頭溝

Abstract

Adults can judge an object's material by perceiving the visual appearance of one material is combined with the impact sound of the same material (Fujisaki et al., 2014). However, its development has been little understood. This study examined the brain activity in response to audiovisual material perception ("Metal" and "Wood") in preverbal 4- to 8-month-old infants (experiment 1) and adults (experiment 2), by using functional near infrared spectroscopy (fNIRS). We measured cerebral hemoglobin concentrations in bilateral temporal brain regions, included the superior temporal sulcus (STS) area which is known to be critical for supra-additive processing of auditory and visual information (e.g., Calvert et al., 2000). Our results showed that a cerebral activation in the right temporal region increased when infants, as well as adults, observed match stimuli which is the auditory material property with the congruent visual material (i.e., "Metal"

* 現所属：中央大学心理学研究科，日本学術振興会特別研究員 (PD)

sound with “Metal” movie). This indicates that the right temporal region comes well reflect to the processing of audiovisual material matching.

【Key words】 Audiovisual integration, material perception, fNIRS, superior temporal sulcus

問題と目的

私たちは、見ること、聞くことを通して、周囲の環境にあるモノの素材や状態を認識している。例えば、ナイフやフォークなどの食器が、金属の素材で作られているのか、あるいはプラスチックにメッキ加工が施されたものであるのかを見分ける際には、素材の視覚的な光沢感だけでなく、モノを叩いた際に発する音も重要な手がかりとなる。近年では、成人を対象とした研究から、「木」や「金属」といった物体の表面材質（視覚情報）と、その物体を叩いた際に発する音の印象を統合し、物体が持つ質感のカテゴリを判断することが示されている（e.g., Fujisaki, Goda, Motoyoshi, Komatsu, & Nishida, 2014）。

質感知覚の脳内機構について、素材質感の音と視覚の統合には上側頭溝（superior temporal sulcus : STS）の働きが関与する可能性が考えられる（e.g., Calvert, Campbell, & Brammer, 2000; Kanero, Imai, Okuda, Okada, & Matsuda, 2014）。成人を対象とした fMRI 研究では、上側頭溝が視聴覚統合処理の責任領野であることが報告されており（e.g., Calvert, et al., 2000; Driver & Noesselt, 2008）、特に右側の上側頭溝は音と視覚イメージの統合（e.g., ブーバキキ効果, Kanero et al., 2014）に関与することが示されている。視覚的な素材質感（e.g., Goda et al., 2014）や、聴覚的な素材質感（e.g., Arnott et al., 2008）の知覚処理では、どちらも比較的低次な処理経路（腹側系）が関与することが示されているが、これらの処理は低次な処理経路から、上側頭溝を含む、より高次の領野へと処理が進むことが示唆されている（Goda et al., 2014）。

一方で、素材質感の音と視覚の統合がいつから、どのように獲得されるのかは明らかでない。Lickliter, Bahrick, & Vaillant-Mekras（2017）によれば、生後から 2 ヶ月の間に、刺激強度や呈示時間、時間的同期など感覚間に共通した物理特性の統合（アモーダル統合）が発達し、その後、単一の感覚処理の発達に伴い、各感覚モダリティに固有の情報を統合するクロスモーダル処理が発達するとしている。先行研究では、生後 4 ヶ月の乳児において顔と音声の統合が見られることが報告されており（e.g., Patterson & Werker, 1999; Werker & Tees, 1983）、同様の時期に視聴覚統合の神経基盤が発達していることが予測できる。

本研究では、近赤外分光法（NIRS）を用いて、素材質感の視聴覚統合に関する神経基盤の発達過程について、言語獲得直前の生後 4–8 ヶ月児（実験 1）と成人（実験 2）を対象に検討を行った。

実験 1

方法

実験参加者

生後 4–8 ヶ月 32 名（女児 19 名）が本実験に参加した。生後 4–5 ヶ月児 16 名（女児 11 名，平均日齢 137 日）を低月齢群，生後 6–8 ヶ月児 16 名（女児 8 名，平均日齢 200 日）を高月齢群とした。実験参加者のリクルートならびに本実験は，中央大学倫理審査委員会の承認の上，実施した。

刺激と手続き

実験刺激は，Fujisaki et al. (2014) で使用された，木や金属の材質を叩く CG 動画 (800 × 600 pixels) と音を用いた。材質から発せられる音とその表面材質（映像）が一致した刺激と不一致の刺激を，それぞれ 2 種類作成し，音刺激を基準として，Wood-sound 条件（木を叩く音と木を叩く映像を一致刺激，木を叩く音と金属を叩く映像を不一致刺激），Metal-sound 条件の 2 つを設定した。

低月齢群，高月齢群をそれぞれ 8 名ずつ，Wood-sound 条件，Metal-sound 条件にランダムに割り振り，実験を実施した。観察距離を約 40cm として 21 インチの CRT モニタ（解像度 1,024 × 768 pixels）から視覚刺激を呈示し，聴覚刺激はモニタの左右に配置したスピーカーから約 60dB で呈示した。刺激呈示はブロックデザインを使用し，ベースライン試行を挟んで，一致刺激あるいは不一致刺激のターゲット試行を交互に呈示した。ターゲット試行では，一致刺激（あるいは不一致刺激）が 8 回呈示し（12.4 秒），ベースライン試行では，ランダムドット画像 (800 × 600 pixels) とホワイトノイズを繰り返し呈示した（12.8 秒）。

NIRS 測定

乳児の脳血流の測定には，ETG-4000 (Hitachi Medical, Japan) を使用した。2 種類の光波長（695 and 830 nm）により，酸化ヘモグロビン (oxy-Hb)，脱酸化ヘモグロビン (deoxy-Hb)，総ヘモグロビン (total-Hb) の相対的な変化量を測定した。測定チャンネル数は左右各 12 チャンネルであり，シリコン製の乳児用プローブ（3 × 3 のファイバー配列，配置間隔 2cm）を用いて，T3，T4（国際 10–20 法）を基準とした両側頭領域を計測部位とした。測定データにバンドパスフィルタ（0.02–1.0 Hz）を適用した後，刺激呈示前 3 秒のデータの平均値と標準偏差を用いて Z-score を算出した。有効データの選定基準として，先行研究と同様に（e.g., Kobayashi, Macchi Cassia, Kanazawa, Yamaguchi, & Kakigi, 2018），ターゲット試行中のディスプレイの注視時間が 6 秒以下，あるいは，測定データに体動ノイズの波形が含まれる場合は分析対象から除外した。

結果と考察

各チャンネルの Oxy-Hb 値について，刺激呈示前 3 秒を基準として Z-score を算出し，刺激呈示後

9~13 秒を分析区間とした左右各 12ch の平均値を図 1 に示した。

条件ごと、ターゲット試行（一致刺激，不一致刺激）における左右の脳血流量の平均値と，ベースライン（Z score = 0）との t 検定を行った結果，Wood-sound 条件では両月齢群で，一致刺激を観察した際の右側頭領域の脳血流量の有意な増加が示され（低月齢群： $t(7) = 2.37, p = .049, d = .71$ ；高月齢群： $t(7) = 2.91, p = .02, d = 1.06$ ），一方で，不一致刺激に対する脳血流量の有意な増加は示されなかった（低月齢群： $t(7) = 1.06, n.s.$ ；高月齢群： $t(7) = .78, n.s.$ ）。左側頭領域では，いずれの刺激に対しても脳血流量の有意な増加が示されなかった（一致刺激：低月齢群： $t(7) = 1.94, n.s.$ ；高月齢群： $t(7) = .65, n.s.$ ，不一致刺激：低月齢群： $t(7) = .58, n.s.$ ；高月齢群： $t(7) = .78, n.s.$ ）。Metal-sound 条件では，高月齢群においてのみ，一致刺激を観察した際の右側頭領域の脳血流量の有意な増加が示されたが（低月齢群： $t(7) = .14, n.s.$ ；高月齢群： $t(7) = 2.66, p = .03, d = 1.44$ ），不一致刺激では脳血流量の有意な増加は示されなかった（低月齢群： $t(7) = .11, n.s.$ ；高月齢群： $t(7) = .35, n.s.$ ）。左側頭領域では，Wood-sound 条件の結果と同様に，いずれの刺激に対しても脳血流量の有意な増加が示されなかった（一致刺激：低月齢群： $t(7) = .49, n.s.$ ；高月齢群： $t(7) = 1.37, n.s.$ ，不一致刺激：低月齢群： $t(7) = .40, n.s.$ ；高月齢群： $t(7) = .29, n.s.$ ）。

実験 1 の結果から，素材質感の音と視覚の統合には，音と視覚イメージの統合に関わる右側の上側頭溝の活動（Kanero et al., 2014）が関与することが示された。さらに，このような素材質感の脳内機構は，言語獲得前の生後 4 ヶ月児においても見られることが明らかとなった。

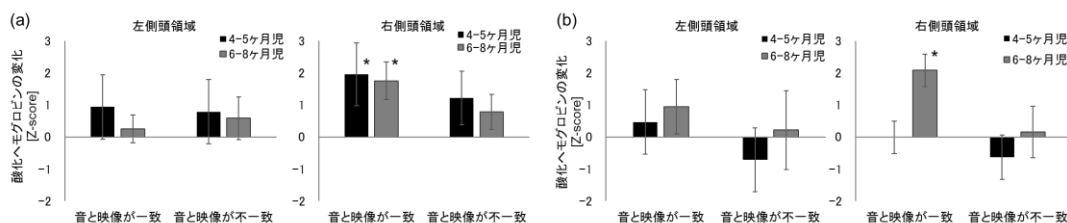


図 1 ターゲット刺激呈示後 9~13 秒間の Oxy-Hb の平均変化量 (Z-scores)

(a) Wood-sound 条件, (b) Metal-sound 条件。エラーバーは標準誤差，アスタリスクはベースライン (Z-score=0) との有意差 (有意確率 5%)。

実験 2

方法

実験参加者

日本人の成人 10 名（女性 7 名，男性 3 名，平均年齢 21.9 歳）が実験に参加した。すべての参加者は正常な視力（矯正含む），および正常な聴力であった。

刺激と手続き

観察距離を約 60cm として 21 インチの CRT モニタ（解像度 1,024 × 768 pixels）から視覚刺激を呈示し、聴覚刺激はモニタの左右に配置したスピーカーから約 60dB で呈示した。実験刺激ならびに刺激呈示の手続きは、実験 1 と同様とした。本実験では、各被験者に対して、Wood-sound 条件と Metal-sound 条件の 2 つのブロックに分けて NIRS 測定を実施した。

NIRS 測定

成人の脳血流の測定には、実験 1 と同様に、ETG-4000 (Hitachi Medical, Japan) を使用した。乳児の測定方法と同様に、測定チャンネル数を左右各 12 チャンネルであり、T3, T4（国際 10-20 法）を基準とした両側頭領域を計測部位とした。測定プローブは成人用プローブ（3×3 のファイバー配列、配置間隔 2cm）を用いた。測定データにバンドパスフィルタ（0.02-1.0 Hz）を適用した後、刺激呈示前 3 秒のデータの平均値と標準偏差を用いて Z-score を算出した。測定データに体動ノイズの波形が含まれる場合は分析対象から除外した。

結果と考察

分析の指標は実験 1 と同様に、酸化ヘモグロビン (oxy-Hb) とした。各チャンネルの Oxy-Hb 値について、刺激呈示前 3 秒を基準として Z-score を算出し、刺激呈示後 3 秒ごとを分析区間とした Oxy-Hb の Z-score の平均値を表 1 に示した。

刺激呈示に対応した脳血流の変化量について、実験 1 と同様に、各分析区間の Z-score とベースライン (Z score = 0) との t 検定を行った。その結果、Metal-sound 条件では、一致刺激を観察した際、右側頭領域の脳血流量が、刺激呈示後 6 秒以降に有意に増加することが示され（刺激呈示後 6~9 秒： $t(9) = 3.13, p = .01, d = 1.40$ ；刺激呈示後 9~12 秒： $t(9) = 1.84, p = .09, d = .86$ ）、左側頭領域においても、同様の脳血流量の増加が示された（刺激呈示後 6~9 秒： $t(9) = 2.85, p = .02, d = 1.27$ ；刺激呈示後 9~12 秒： $t(9) = 2.67, p = .03, d = 1.19$ ）。一方で、Wood-sound 条件では、一致刺激を観察した際、刺激呈示後 9 秒以降に、左右側頭の脳血流量が増加する傾向が示された（右側頭領域、刺激呈示後 6~9 秒： $t(9) = .51, n.s., d = .23$ ；刺激呈示後 9~12 秒： $t(9) = 1.92, p = .09, d = .82$ 、左側頭領域、刺激呈示後 6~9 秒： $t(9) = .58, n.s., d = .26$ ；刺激呈示後 9~12 秒： $t(9) = 1.92, p = .09, d = .86$ ）。

これらの結果から、成人においては、素材質感の視聴覚統合に両側頭領域の活動が関与することが示された。本実験では脳血流量の増加が刺激呈示後 6~12 秒で見られたが、成人における脳血流量の増加は乳児と比べてピークが早まることが知られており（Minagawa-Kawai et al., 2011）、先行研究の結果とも整合的であると考えられる。また、Wood-sound 条件における結果は有意傾向ではあるものの、分析により示された効果量は大きな値であり（Cohen, 1993）、本実験のサンプルサイズが統計結果に影響した可能性が考えられる。

表 1 刺激呈示後 3 秒ごとを分析区間とした 0xy-Hb の Z-score の平均値 (標準偏差)

Condition	Hemisphere	Stimulus	0-3s	3-6s	6-9s	9-12s	12-14s
Metal sound	Left	match	-.06 (.731)	.61 (1.198)	1.50 (1.661)	1.73 (2.050)	1.44 (1.409)
		mismatch	-.19 (.662)	-.41 (1.637)	.22 (2.854)	.94 (3.381)	.79 (3.694)
	Right	match	-.27 (.742)	.12 (.860)	.87 (0.877)	.98 (1.695)	1.03 (1.704)
		mismatch	-.28 (.744)	-.52 (2.555)	-.32 (3.951)	.46 (3.296)	.63 (2.734)
Wood-sound	Left	match	-.22 (1.107)	-.14 (1.793)	.31 (1.677)	1.17 (1.923)	1.30 (2.886)
		mismatch	.14 (1.042)	.32 (1.836)	.49 (2.787)	1.67 (3.361)	2.71 (4.030)
	Right	match	.05 (1.659)	.27 (1.900)	.16 (1.003)	.97 (1.594)	1.52 (2.785)
		mismatch	-.11 (0.992)	.33 (1.768)	.68 (2.014)	1.29 (3.317)	2.19 (3.765)

総合考察

本研究では、素材質感の視聴覚統合における脳内機構の発達過程を明らかにすることを目的とし、近赤外分光法 (NIRS) を用いて、実験 1 では言語獲得直前の乳児 (生後 4-8 ヶ月児) を対象に、実験 2 では成人を対象に検討を行った。

本研究の結果、素材質感の視聴覚統合では、視聴覚統合処理の責任領野である上側頭溝の働きが関与することが示された (e.g., Calvert et al., 2000; Driver & Noesselt, 2008)。実験 1 の結果、言語獲得前の乳児では、素材を叩く音と映像が一致した刺激を観察した際、不一致の刺激と比べ、右側頭領域の脳血流の有意な増加が示された。乳児を対象とした NIRS の先行研究から、T3, T4 を中心とした左右側頭の測定領域には上側頭溝の活動が含まれることが示されており (e.g., Blasi et al., 2014; Lloyd-Fox et al., 2014)、質感を持つ素材の視聴覚統合の発達には右側の上側頭溝の活動が関与することが示された。また、成人を対象とした実験 (実験 2) では、素材を叩く音と映像が一致した刺激を観察した際、両側頭の脳血流の増加が示された。先行研究によれば、右側の側頭領域は環境音の処理、左の側頭領域は言語音の処理に対応することが示唆されており (Thierry, Giraud, & Price, 2003)、乳児と成人における活動領野の違いは、言語獲得に伴う脳活動の変化を反映している可能性が考えられる。

さらに、乳児の脳活動の結果から、素材質感の視聴覚統合の発達は、素材によって獲得時期が異なることが示された。素材質感の視聴覚統合に対応する側頭領域の脳活動は、木の素材に対しては生後 4 ヶ月から存在するが、金属の素材では生後 6 ヶ月以降に見られることが示された。質感知覚の発達では、生後 7 ヶ月から物体表面の光沢感を知覚することが示されており (e.g., Yang et al., 2011)、金属素材の視聴覚統合の獲得時期に視覚的な光沢知覚の発達が影響している可能性が考えられる。

引用文献

Arnott, S. R., Cant, J. S., Dutton, G. N. & Goodale, M. A. (2008). Crinkling and crumpling: An

- auditory fMRI study of material properties. *NeuroImage* 43, 368–378.
- Blasi, A., Lloyd-Fox, S., Johnson, M.H., & Elwell, C.E. (2014). Test–retest reliability of functional near infrared spectroscopy in infants. *Neurophotonics* 1(2), 025005.
- Calvert, G. A., Campbell, R. & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159
- Driver, J. & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron* 57, 11–23.
- Fujisaki, W., Goda, N., Motoyoshi, I., Komatsu, H. & Nishida, S. (2014). Audiovisual integration in the human perception of materials. *Journal of Vision*. 14, 1–20.
- Goda, N., Tachibana, A., Okazawa, G. & Komatsu, H. (2014). Representation of the material properties of objects in the visual cortex of non-human primates. *J. Neurosci.* 34, 2660–2673.
- Kanero, J., Imai, M., Okuda, J., Okada, H. & Matsuda, T. (2014). How sound symbolism is processed in the brain: A study on Japanese mimetic words. *PLoS ONE* 9, e97905.
- Kobayashi, M., Macchi Cassia, V., Kanazawa, S., Yamaguchi, M. K., & Kakigi, R. (2018). Perceptual narrowing towards adult faces is a cross-cultural phenomenon in infancy: A behavioral and near-infrared spectroscopy study with Japanese infants. *Developmental Science*, 21(1), e12498.
- Lickliter, R., Bahrick, L.E., Vaillant-Mekras, J. (2017). The intersensory redundancy hypothesis: Extending the principle of unimodal facilitation to prenatal development. *Developmental Psychobiology*. 59, 910–915.
- Lloyd-Fox, S., Richards, J.E., Blasi, A., Murphy, D.G.M., Elwell, C.E., & Johnson, M.H. (2014). Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics* 1(2), 025006.
- Minagawa-Kawai, Y., Lely, H.V., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cerebral cortex*, 21 2, 254–61.
- Patterson, M., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, 22, 237–247.
- Thierry, G., Giraud, AL., & Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron*. 38, 499–506.
- Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, 37, 278–286.
- Yang, J., Otsuka, Y., Kanazawa, S., Yamaguchi, M. K. & Motoyoshi, I. (2011). Perception of surface glossiness by infants aged 5 to 8 months. *Perception* 40, 1491–1502.

謝 辞

本実験に参加して下さった多くの赤ちゃんとご家族、成人被験者の方々にお礼を申し上げます。また、データの取得にご協力いただきました中央大学山口研究室の皆様へ感謝申し上げます。なお本研究の一部は、新学術領域研究「トランスカルチャー状況下における顔身体学の構築—多文化をつなぐ顔と身体表現—」（研究代表者：山口真美，課題番号：17H06343）、「多元質感知」（研究代表者：金沢創，課題番号：18H05014），ならびに科学研究費補助金（課題番号：19J00722，19K20650），そして公益財団法人発達科学研究教育センターのご支援を頂きました。