

幼児指導者が示す顔の表情が子どもの言語発達に果たす役割： 初期発達時の認知・適応能力をより理解するための支援研究

東京大学 国際高等研究所 ニューロインテリジェンス国際研究機構 バーベア モニカ
東京大学 医学部 鴻野 芽依
理化学研究所 脳神経科学研究センター・ミドルベリー大学 脳科学専攻 辰巳 由夏
オックスフォード大学 実験心理学部 レシュト サミュエル
東京大学 医学部 藤村 優
トゥウェンテ大学 行動・経営・社会科学部 シヴァクマー キショール
理化学研究所 脳神経科学研究センター 馬塚 れい子
東京大学 国際高等研究所 ニューロインテリジェンス国際研究機構 辻 晶

The role of facial cues in children's language development: Understanding cognitive mechanisms and adaptive abilities

International Center for Neurointelligence, University of Tokyo, BARBIR, Monica
Faculty of Medicine, University of Tokyo, KONO, Mei
Center for Brain Science, RIKEN; Program in Neuroscience, Middlebury College,
TATSUMI, Yuka
Department of Experimental Psychology, University of Oxford, RECHT, Samuel
Faculty of Medicine, University of Tokyo, FUJIMURA, Yu
Faculty of Behavioural, Management, and Social Sciences, University of Twente,
SIVAKUMAR, Kishore
Center for Brain Science, RIKEN, MAZUKA, Reiko
International Center for Neurointelligence, University of Tokyo, TSUJI, Sho

要 約

言語を学ぶことは乗り越えられない困難な仕事のように思える。子どもは環境から受ける少ないインプットから、言語の多くの側面を学ばなければならないが、言語発達に不可欠なインプットについてまだ殆ど解明されていない。顔の表情等の要因が重要であると考えられるが、これまでの研究では子どもがめったに経験しない不自然な設定で比較実験が行われてきた。従って、ここでは日常の現実的な社会環境において、乳児期から幼児期までの語彙習得における顔の表情が果たす役割を調査した。サングラスで視線方向を遮り、マスクで唇の動きを遮ったときの子どもの単語学習を評価した。顔の表情が隠されていても乳幼児期の単語認識に差がないことがわかった。これらの結果は、どの要因が言語発達に不可欠であるか理解を深め、幼児教育の証拠に基づくガイドラインを築く基礎として役立つ

つ。

【キー・ワード】 言語発達, 単語, マスク, オンライン実験

Abstract

Learning a language can seem to be an almost insurmountable task. Children have to learn the many aspects of language, from the little input they get. Yet, we still know little about which, if any, kinds of input are essential for language development. Some factors, such as facial cues, have been deemed important, but have only been studied in comparison to unnatural settings children rarely experience. Here, thus, we investigate the role facial cues may play in vocabulary acquisition from infancy to toddlerhood, in everyday social settings. Our study assesses word learning when gaze direction is occluded with sunglasses and the lip movements with a surgical mask. We find that there are no differences in word recognition at 20 months of age when facial cues are occluded. These results advance our understanding of which factors are essential for language development, and can serve as a foundation for evidence-based guidelines for early childhood education.

【Key words】 Language acquisition, vocabulary, surgical mask, online study

問題と目的

音, 単語, 文法に至る言語のすべての側面を習得するには, 乳幼児が保護者から受け取る言語情報だけでは不十分に思えるかもしれない。これまでの研究では口の動きや視線方向といった話し手の音声以外の補足的な手がかりが言語習得に不可欠であり, とても重要な役割を果たす可能性があることを示唆してきた (e.g., McGurk & MacDonald, 1976; Lobmaier, Fischer, & Schwaninger, 2006; Hollich, Newman, & Jusczyk, 2005; Westlund et al, 2017)。一方このような手がかりを使うことのできない状況, たとえばパンデミックによりマスクをすることが普通になるといった状況も十分に存在し, 子どもに接する周りの大人や教育を行う者はこのことが乳幼児の言語発達に影響を及ぼすかもしれないと懸念している。しかし, 顔の手がかりが乳幼児の言語発達に与える直接の影響を調べるのは難しい。顔の手がかりは至るところにあるため, それを除くと不自然, もしくは奇妙になることさえある。たとえば, 言語習得の過程での視線の重要性を測定するために, 話し手がある対象物について話しながらそれを見る条件 (自然) と, 話している間に全く対象の方を見ない条件 (不自然) を比較する, といった研究がよく行われる (e.g., Westlund et al, 2017)。ゆえに我々の研究では, 特に顔の手がかりの役割に焦点を合わせ, 乳幼児の言語学習環境の自然さを完全に保つ。そうすることではじめて, 日常の言語学習の中での顔表情の役割を測定することができるのである。

我々の1つ目の研究では, 話し手の顔の手がかりをマスクやサングラスなどの一般的な服飾品で隠された状態にして子どもに提示する。それから 12, 24, 28 ヶ月の乳幼児が既知の単語をどの程度認

識するのか(理解効果), 未知の単語をどの程度学習するか(学習効果)を評価する。次の研究では大人がマスクやサングラスなどの一般的な服飾品を身に着け, その状態で小さい学習者に向かって単語を教える。話し手が顔の手がかりを隠された状態になると話し方を調整するのか, 調整するならどの程度なのかを明らかにする予定である。もし話し手が話し方を調節するならば, それにより顔の手がかりが隠された影響が相殺されるかもしれない。これらの研究は両方ともオンラインで実施する予定だったが, 技術的な問題から1つ目の研究は研究室で行うことになった。1つ目の研究のための調査はほぼ終了した。更にこの2つの研究をオンラインで実現するために全ての研究者が使用できる乳幼児のオンライン調査実施方法のデータベース(親への説明ビデオや正確に動画を記録できるコードなど)を作成した。

言語発達と顔を見ることに相互関係が報告されていることから(Pons, Bosch, & Lewkowicz, 2019; Tsang, Atagi, & Johnson, 2018; Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan, 2015), この効果の背景にある原因をより詳しく理解することは重要である。本研究の結果によって, 子どもに接する大人や教育者に対して, どのように子どもの言語の発達を促せばよいのか, 特に多くの人がマスクを着用しているパンデミック期間にどうすればよいのかについての情報を与えることができるだろう。

方 法

実験1：顔表情の手がかりが言語習得に与える影響

3つの条件, すなわち視線を隠す(サングラスをかけた話し手), 口元を隠す(フェイスマスクを着けた話し手), 顔の手掛かりを隠さない(帽子をかぶる話し手, 図1)において単語学習課題を実施しその結果(言葉習得の可否とその結果の正答率)を比較する。



図1 学習の条件

参加者

参加者はラボデータベースからリクルートし, 合計126人だった。COVID-19流行により予定の人数の参加者を集めることが困難だったため, 12ヶ月と28ヶ月の参加者はマスク条件のみの調査を行った。全員が一言語を話す日本語を用いる参加者であり, 他の言語を聞いた割合は30%に満たない。試行回数が4回に満たない参加者と視線のデータが50%以下の試行は除外した。残りの参加者は, 12ヶ月: 平均月齢: 12.29(12.02-13.29); 女児率: 16/30; 20ヶ月: 平均月齢: 20.28(20.01-22.00); 女児の人数: 30/68; 28ヶ月: 平均月齢: 28.30(28.04-29.29); 女児の人数: 15/28である。全ての群を完了するためさらに11人の調査を12月までに行う予定である。調査前に参加者の親に調査について説

明し、同意手続きを行った。

実験方法

最初に、サングラス、マスク、帽子のいずれかを着用した女性の話し手が新しい言葉（例、ドッチョ）を教えるトレーニングビデオを視聴させる（図 2A）。次に、テストフェーズでは、画面に 2 つの画像を表示しトレーニングビデオで教えた画像の方を見るように指示する（例、「見て見て、ドッチョだよ。」図 2B）。乳幼児の視線の焦点（画面上で見ている場所）と視線の分布（視線のパターン）は、それぞれ乳幼児のパフォーマンス（言葉の習得）とコンフィデンス（自信）の指標となる。

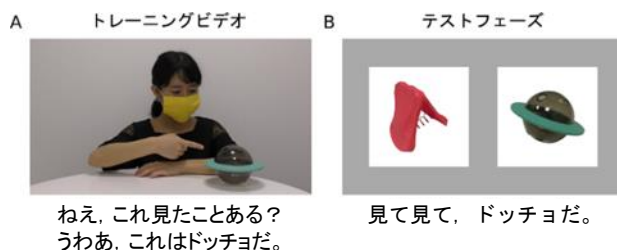


図 2 実験 1 の手法

結果

人数が十分でない群があるため、新規の単語についての分析は行われていない。そこで、既知の単語（車・本など）の認識についての分析を行った結果を提示する。

線形混合モデルで分析したところ、マスク条件において偶然と比較したときの「ターゲット（正解の画像）を見た率 - ディストラクタ（不正解の画像）を見た率」、偶然と比較すると有意に正解を選ぶ確率が高かった ($M = 0.571$, $SD = 0.255$, $\beta = -0.066$, $SE = 0.032$, $t = -2.046$, $p = 0.041$)。さらに月齢が上がるにつれて「ターゲットを見た率 - ディストラクターを見た率」が大きくなった ($B = 0.007$, $SE = 0.001$, $t = 4.49$, $p < 0.001$, 図 3)。この結果をより理解するため各年齢ごとに分析を行った。全ての分析は多重比較のためボンフェローニ法で補正済み ($\alpha = 0.05 / 2$)。12 ヶ月では正解と不正解を見る割合に有意差は見られなかった ($M = 0.512$, $SD = 0.282$, $\beta = 0.012$, $SE = 0.019$, $t = 0.63$, $p = 0.53$)。20 ヶ月 ($M = 0.585$, $SD = 0.245$, $\beta = 0.082$, $SE = 0.011$, $t = 7.16$, $p < 0.001$) と 28 ヶ月 ($\beta = 0.12$, $SE = 0.015$, $t = 7.93$, $p < 0.001$, $M = 0.623$, $SD = 0.217$) で有意に差があった。20 ヶ月以上では単語を認識できることがわかる。

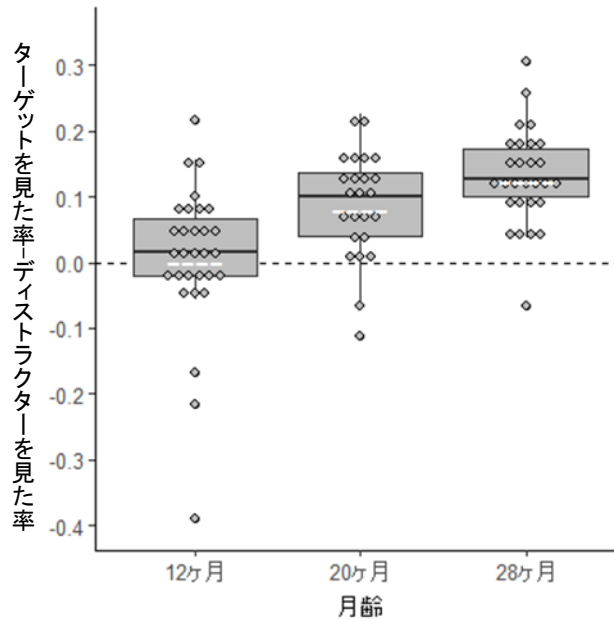


図3 単語認識（年齢別）

年齢とターゲットを見た率 - デストラクターを見た率。点は参加者、白破線は平均、箱の上部と下部は第一四分位数・第三四分位数を表す（25-75パーセントイル）。上のひげは第三四分位数から1.5四分位間の中での最大値、下のひげは1.5四分位間の中での最小値から第一四分位数までを表す。黒の破線は参加者が写真を見たときに正答と誤答を同じ割合で見たことを指す。

次に、20ヶ月の参加者で条件ごとに差があるのかを調べた。画像を見る回数は条件間で同程度だった（帽子:M=0.586, SD=0.246; マスク:0.585, SD=0.246; サングラス:M=0.559, SD=0.257, 図4）。単語認識に条件間での有意な差は見られなかった（マスク: $\beta=0.001$, SE=0.025, $t=0.034$, $p=0.97$, コーエンの $d=0.03$; サングラス: $\beta=0.026$, SE=0.031, $t=0.859$, $p=0.39$, コーエンの $d=0.10$ ）。有意差が見られなかったためベイズ因子を求めたが、値はマスク 0.045, サングラス 0.859であり、それぞれの条件間での差がないと考えられる。

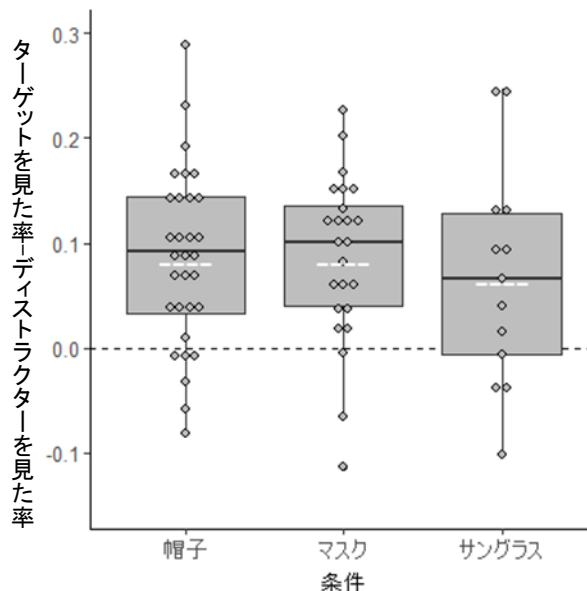


図 4 単語認識 (条件別)

条件とターゲットを見た率 - デストラクターを見た率。点は参加者、白破線は平均、箱の上部と下部は第一四分位数・第三四分位数を表す (25-75 パーセントイル)。上のひげは第三四分位数から 1.5 四分位間の中での最大値, 下のひげは 1.5 四分位間の中での最小値から第一四分位数までを表す。黒の破線は参加者が写真を見たときに正答と誤答を同じ割合で見たことを指す。

また視線分布の分析も実施した。子どもがターゲットとデストラクタの間で視線を切り替える回数 : 帽子 : $M = 9.197$, $SD = 6.893$; マスク : $M = 8.475$, $SD = 4.465$; サングラス : $M = 8.227$, $SD = 4.302$ 。分析の結果, 条件間で有意差はなかった (マスク : $\beta = 0.86$, $SE = 1.367$, $t = 0.629$, $p = 0.52$, コーエンの $d = -0.079$; サングラス : $\beta = 1.438$, $SE = 1.671$, $t = 0.861$, $p = 0.381$, コーエンの $d = 0.139$)。更にベイズ因子分析により条件間に差がない可能性が高いことが確認された (値はマスク 0.055, サングラス 0.067)。視線の焦点と視線の分布の両者とも単語認識において条件間の違いはなかった。

実験 2 : 顔の手がかりを隠した状況における言語指導者の教え方適応法

この実験では指導者 (大人の実験参加者) が, 視線を隠す (サングラス), 口元を隠す (マスク), または顔の手掛かりを隠さない (比較対照, 帽子) 条件下で学習者へ話しかける際の客観的な発話方法を評価する (図 5)。オンライン環境で正確に時間を固定してビデオと音声を記録するのが困難だったため, この調査を実行することはまだできなかった。代わりに, このようなオンライン調査を正常に行う前提になる乳幼児のオンライン調査実施方法のデータベースを作成することにした。

言葉を教える



これは猫です。

図5 実験2の手法

これは乳幼児（やさらに大きい子ども，大人）でのオンライン調査を行う研究者が無料で使うことができるものである。このオンラインのデータベースは2つの部分に分かれる。1) 説明ビデオと2) オンライン調査実装用コードである。親に調査手順と親の役割を明確に説明するため，説明ビデオは2歳未満の子ども用と2歳以上の子ども用に分けている（図6）。



図6 説明映像：実験中の子どもを抱く方法

これにより，近くに調査スタッフがいてもオンラインで調査を行うことができる。オンライン調査用のコードを使うことで子どもの視線をビデオで記録するために顔を中心に配置するテンプレートを作り（図7），正確にビデオ記録を始め，同意書を匿名化しつつ帰属を記録することができるようになる。このデータベースは完成次第 Open Science Framework で利用可能になる (<https://osf.io/kn7hg/>)。

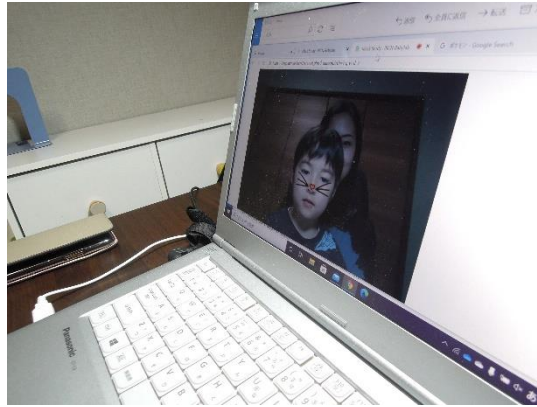


図 7 ウェブカメラにおける画面調整

大人のオンライン調査は広く普及しており、これまでに世界中から 60 万人の参加者による調査例も存在する (Hartshorne, Tenenbaum, & Pinker, 2018)。世界中の多くの参加者にたとえ研究室に来ることができなくても (居住地や仕事の都合など) アプローチすることができる。一方乳幼児のオンライン調査はそれよりずっと珍しい。今回作ったりソースにより研究者が乳幼児のオンライン調査の難しさ (たとえば乳幼児が応答するのにキーボードを使わず視線を記録しなければならないこと) を乗り越えることができるだろう。

これにより研究者は日本中、さらには世界中の乳幼児を対象とした調査が可能になり、これまで見過ごされがちだった層にもアプローチできるようになる。加えて、冬のインフルエンザシーズンや世界的な感染症の流行が起こっているときも、調査を家庭で安全に実施できるため研究を継続することが可能である。乳幼児オンライン調査は、発達科学者がこれまで何年も我慢しなくてはならなかった調査限界の一部を解決することができる。

考 察

マスクやサングラスで顔を隠した条件と、帽子のみを被り顔を隠さなかった条件では、有意差は見られなかった。自然な状況下においては、顔の一部が隠れていることの単語認識に対する影響は小さいと言える。

今回は自然な状況で行った調査であることがこれまでの調査との違いであり、条件間で単語認識に違いがなかったのは自然さが関わっているかもしれない。この研究ではトレーニングビデオ中にできるだけ多くの自然な合図を入れた。とりわけ、乳幼児向けの話し方をしたり、ビデオの中の話し手と乳幼児の間で視線を合わせたりなど、明確な合図 (ostensive cue) を入れた。これらの合図は学習において非常に重要だと考えられている (Senju, & Csibra, 2008)。よって、乳幼児がトレーニングビデオ中の指差しやジェスチャーに注意を向けるのは可能であり、顔の手がかりがない状況を補うことができたのだ。

一方 12 ヶ月の参加者は既知の単語を認識することができず、顔の手がかりを隠すことには懸念が

残る。しかし、親の報告によるとほとんどの12ヶ月の参加者はこれらの単語を知らなかったこと、語彙の増加は生後約20ヶ月で起こると報告されているということを考慮すべきである。我々は乳幼児が少数の単語しか知らない月齢（12ヶ月）、語彙爆発が起こっている月齢（20ヶ月）、多くの単語を知っている月齢（28ヶ月）という風に目的を持ってそれぞれの月齢を選んだ。

そして月齢が上がるにつれ正解を選んだ割合と選ばなかった割合の差が上昇した。単語を認識する能力は月齢が上がるにつれて高くなる。幼児の語彙は20ヶ月頃に増えると言われているので、予想通りと言える（Fernald, Perfors, & Marchman, 2006）。

文献では口を隠すことは知覚にネガティブな影響を与えるというエビデンスがあるが、特定の層においてサージカルマスクの着用が理解力に与える影響、という点に着目している研究ではほとんど影響が確認されていない（Radonovich, Yanke, Cheng, & Bender, 2009; Atcherson et al, 2017; Saito, 2018; Singh, Tan, & Quinn, 2021）。今後、話を聞きながら参加者が顔のどこを見ているのか分析する予定だが、それにより、マスクで口が隠れている状態の話し手を見るとときに聞き手がどのようなテクニックを使っているのか解明できるだろう。

本調査では自然な条件下では顔の表情の一部がわからなくても単語の認識に差がないということがわかった。そのため、保育園などで大人がマスクを着けることの子どもへの影響は、単語認識に対しては大きな心配がないと考えられる。

引用文献

- Atcherson, S. R., Mendel, L. L., Baltimore, W. J., Patro, C., Lee, S., Pousson, M., & Spann, M. J. (2017). The effect of conventional and transparent surgical masks on speech understanding in individuals with and without hearing loss. *Journal of the American Academy of Audiology, 28*(1), 58-67.
- Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental psychology, 42*(1), 98.
- Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition, 177*, 263-277.
- Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child development, 76*(3), 598-613.
- Lobmaier, J. S., Fischer, M. H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53*(2), 117-122.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews, 4* (1), 1.

- Pons, F., Bosch, L., & Lewkowicz, D. J. (2019). Twelve-month-old infants' attention to the eyes of a talking face is associated with communication and social skills. *Infant Behavior and Development, 54*, 80-84.
- Radonovich Jr, L. J., Yanke, R., Cheng, J., & Bender, B. (2009). Diminished speech intelligibility associated with certain types of respirators worn by healthcare workers. *Journal of Occupational and Environmental Hygiene, 7*(1), 63-70.
- Saito, O. (2018). マスク着用による聞き取りへの影響と聴力について(The effect of masks on listening comprehension and hearing ability). *Audiology Japan, 61*(5), 136.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*(9), 668-671.
- Singh, L., Tan, A., & Quinn, P. C. (2021). Infants recognize words spoken through opaque masks but not through clear masks. *Developmental Science*.
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42*(6), 1173-1190.
- Tsang, T., Atagi, N., & Johnson, S. P. (2018). Selective attention to the mouth is associated with expressive language skills in monolingual and bilingual infants. *Journal of Experimental Child Psychology, 169*, 93-109.
- Westlund, J. M. K., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction, 13*, 1-9.

本研究は東京大学の倫理委員会の承認を得ています。

謝 辞

ご参加いただくすべてのご家族に感謝致します。また、参加者の募集および調査に関してご協力いただいた幡野由里様、小野理恵様、大塚知子様、高島智子様、そして機器設置にご協力いただいた山根直人様にも感謝の意を伝えさせていただきます。