

顔と声からの多感覚感情認知におけるフェイスマスク着用の影響： COVID-19 パンデミック前後の比較

Impact of Wearing a Face Mask in Multisensory Emotion Perception: Comparison Between the Pre and Post COVID-19 Pandemic

河原 美彩子^{1, 2, 3}, 田中 章浩³
Misako Kawahara, Akihiro Tanaka

¹東京大学, ²日本学術振興会, ³東京女子大学

University of Tokyo, Japan Society for the Promotion of Science, Tokyo Woman's Christian University
misako-kawahara@g.ecc.u-tokyo.ac.jp

概要

The COVID-19 pandemic has led to widespread use of face masks, which can reduce some of the facial cues used in emotion perception. To investigate whether the partial loss of visual cues impacts multisensory emotion perception from face and voice and whether frequent exposure to masked faces affects it, we conducted an emotion perception task using audio-visual stimuli that partially occluded speakers' faces. Participants were shown videos of faces and voices expressing congruent or incongruent emotions, and were asked to judge whether the person was happy or angry. Our results showed that participants relied more on vocal cues for emotion perception when facial cues were partially occluded. Furthermore, when the mouth was covered, participants relied less on vocal cues after the pandemic than before. These findings suggest that partial face masking and prolonged exposure to masked faces can affect multisensory emotion perception.

キーワード：Multisensory, emotion perception, face mask

1. はじめに

円滑なコミュニケーションを実現するためには、言語情報を手がかりとして相手の発話内容を理解するだけでなく、顔の表情や声の調子といった非言語情報を手がかりとして他者の感情状態を正しく読みとることが必要不可欠である。他者の感情を知る手がかりとして顔や声は重要な情報源であり、我々はこれらの情報を単独で受けとるのではなく、むしろ同時に受けとり、多感覚的に認知しているといえる。実際、顔と声の表す感情が矛盾している場合に比べて、感情が一致している場合に感情認知が促進され [e.g., 1, 2], 視覚情報と聴覚情報は相互に影響を及ぼすことが示されている [3, 4, 5]。しかし、日常生活におけるコミュニケーション場面において、顔や声の手がかりが常に明瞭であるとは限らない。たとえば、感情表出者がサングラスやマスクを着用している場合、顔の情報が部分的に欠損してしまう。特に COVID-19 のパンデミック下では、感染のリスクを最小限に抑えるため、口と鼻を覆うフ

ェイスマスク（以下、マスク）を着用することが広く推奨されてきた。口は顔からの感情認知において重要な手がかりであり [e.g., 6, 7], 最近の研究では、マスクによって顔が部分的に覆われていると、覆われていない場合と比較して、顔からの感情認知精度が低下することが示されている [e.g., 8, 9]。これらの知見は、顔の部分的な遮蔽が顔からの単感覚的な感情認知に影響することを示すものであるが、顔と声からの多感覚な感情認知にも影響を及ぼす可能性がある。実際、視覚情報である顔の情報の劣化が、顔と声からの多感覚感情認知における視聴覚統合に影響することが報告されている [10]。この Kokinou らの研究 [10]では、ガウシアンフィルタによって劣化させた視覚情報（つまり顔表情）を刺激として用いているが、多感覚的な感情認知について部分的に遮蔽された顔刺激を用いて検討した研究は少ない。

そこで本研究では、多感覚感情認知における顔と声に対する情報の重み付けが、顔の部分的な遮蔽によってどのように影響を受けるのか検討した。実験では、顔と声によって感情を表現した動画を刺激として用い、動画話者の感情（怒り、喜び）を回答させる感情認知課題を実施した。顔と声の感情が一致する刺激に加え、感情が矛盾する刺激も使用することで、顔と声のどちらの情報を重視して感情を判断したかを測定した。顔の遮蔽の影響を検討するため、動画話者の目または口を遮蔽した刺激も使用し、顔全体が見える場合と比較して、顔と声の情報に対する重みづけがどのように変化したか調べた。動画話者の目または口が遮蔽された場合、視覚的な手がかりが減少することによって、顔全体が見えている場合と比較して、話者の声の情報への重みづけが大きくなると仮説を立てた（仮説1）。

さらに本研究では、上述の感情認知課題を COVID-19 パンデミック前後に実施し、結果を比較することで、多感覚感情認知における顔の部分遮蔽の影響が

COVID-19 パンデミックの前後で異なるかどうかについても検討した。COVID-19 パンデミック下における対面コミュニケーションでは、マスク着用によって顔の下部が遮蔽された顔（以下、マスク顔）に暴露する機会が劇的に増加したといえる。アメリカ人を対象とした先行研究では、マスク顔への接触が相手の顔の目領域を手がかりとすることの依存度を高め、表情認知パターンを変容させる可能性を示唆している [11]。マスクの着用は日本で特に広く普及していることから [12]、日本人はマスク顔への接触頻度が欧米諸国の人々よりもさらに高い可能性があり、COVID-19 パンデミックの前後で多感覚感情認知における顔と声の手がかりの重み付けにも変化が生じている可能性がある。マスク顔への接触の影響として、二通りの仮説を立てた。第一に、パンデミック前よりも後の方が、多感覚感情認知における声に対する重みづけが全般的に増加すると予測した（仮説 2a）。マスク顔に接触する中で、減少した視覚的手がかりを補い聴覚情報に注意を向けて感情を読みとるよう変化した可能性がある。一方で、パンデミック前よりも後の方が、多感覚感情認知における声に対する重みづけが、特に口元が遮蔽された他者に対して減少すると予測も可能である。マスク顔への接触に伴い感情表出者の目領域への依存が促進され [11]、顔上部からの感情認知精度が向上し、その結果声の手がかりへの重みづけが少なくなる可能性がある。したがって、パンデミック後の方が前と比較して、口元が遮蔽された他者から感情を読みとる際に声を重視する程度が小さいと仮説を立てた（仮説 2b）。

2. 方法

参加者 COVID-19 パンデミック前（2019年6月）に28名 ($Mage = 20.4; SD = 0.6$)、パンデミック後（2022年11月～2023年1月）に29名 ($Mage = 20.2; SD = 1.1$)に実験を実施した。参加者は全員日本語を母国語とする東京女子大学の女子学生で、正常な聴力と視力を持っていた。なお、本研究は東京女子大学の研究倫理委員会の承認を受けている（承認番号：A2020-5）。

刺激 日本語話者2名が、顔と声によって感情を表現した動画を刺激として使用した。発話内容は意味的に中立な短いセリフ4種（「これなに?」「そうなんですか」「さようなら」「はいもしもし」）で、怒りまたは喜びの感情を込めて発話された。顔と声の感情が一致する刺激（以下、感情一致刺激：怒り顔+怒り声、喜

び顔+喜び声）に加え、感情一致刺激の映像と音声を入れ替えて感情が矛盾する刺激（以下、感情不一致刺激：怒り顔+喜び声、喜び顔+怒り声）も作成した。

顔の部分的な遮蔽が多感覚的な感情認知にどのように影響するかを検討するため、動画話者の鼻背の中心より上部（目遮蔽条件）または下部（口遮蔽条件）に Premiere Elements 2018 (Adobe Inc.)で白い正方形を追加し、顔の一部を遮蔽した刺激も作成した。加工前の元の映像が、遮蔽なし条件の刺激として使用した。したがって、計96個の多感覚刺激が作成された（話者2名×セリフ4種×感情2種×感情一致性（一致、不一致）×顔遮蔽3条件）。また、顔または声からの単感覚的な感情認知精度も検討するため、感情一致刺激の映像または音声から作成した単感覚刺激（顔のみ刺激計48個、声のみ刺激16個）も作成した。

手続き 実験室にて個別に実験を実施した。実験では、顔と声による動画刺激を呈示する多感覚セッションに続いて、顔のみ刺激または声のみ刺激を呈示する単感覚セッションが実施された。各試行では、まず注視点（500ms）と合図音（440 Hz の純音、持続時間 100ms）が同時に呈示され、500ms の空白を挟んだ後、刺激が呈示された。刺激の映像はPCモニタ（Latitude3540, Dell）から、音声はヘッドホン（HDA300, SENNHEISER）からそれぞれ呈示された。参加者は、各試行において呈示された刺激の動画話者の感情が怒りと喜びのどちらであるか判断するよう教示された。なお、顔と声のどちらの感情に注目するかについて指示はなかった。各刺激が1回ずつ呈示されたため、実験は計160試行から構成された。

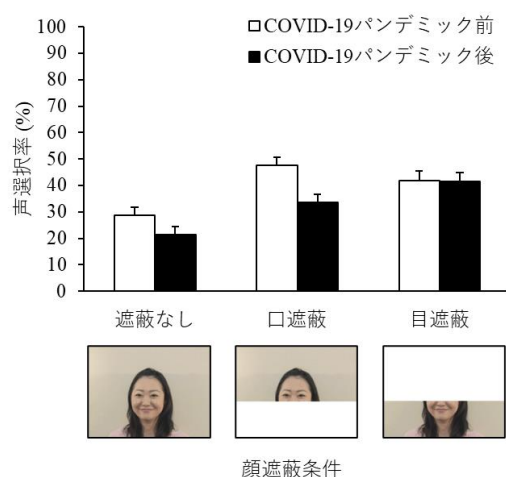
3. 結果

多感覚感情認知における顔と声に対する情報の重みづけに焦点を当てる本研究では、感情不一致刺激に対する回答に着目し、声の感情を選択した割合（以下、声選択率）を算出した。たとえば、怒り顔と喜び声の組み合わせに対して「喜び」と回答した場合、声の感情を選択した回答とみなした。顔の遮蔽と実験実施時期の各条件における声選択率を Figure 1 に示す。

顔の部分的な遮蔽およびマスク顔への接触が、多感覚感情認知にどのように影響を及ぼしているか検討するため、声選択率に対して時期（パンデミック前後）×顔遮蔽条件（遮蔽なし、口遮蔽、目遮蔽）のANOVAを実施した。その結果、顔遮蔽条件の主効果 ($F(1.80,$

Figure 1

顔遮蔽と時期の各条件における声選択率



注) エラーバーは標準誤差を示す。

99.12)=44.25, $p<.001$, $\eta_p^2=.45$) が有意だった。Bonferroni法による下位検定を行ったところ、遮蔽なしの場合と比較して、口または目が遮蔽された場合には声選択率が高いことが示され ($ps<.001$), 仮説 1 が支持された。時期の主効果は有意傾向であり ($F(1, 55)=3.22$, $p=.08$, $\eta_p^2=.06$), 声選択率はパンデミック前に比べて後の方が低い傾向にあった。

重要な結果として、顔遮蔽条件と時期の交互作用が有意だった ($F(1.80, 99.12)=6.00$, $p=.005$, $\eta_p^2=.10$)。単純主効果検定を実施した結果、パンデミック前後での声選択率の差異は、口遮蔽条件でのみみられ ($F(1, 55)=10.47$, $p=.002$, $\eta_p^2=.16$), 目遮蔽条件 ($F(1, 55)=0.001$, $p=.98$, $\eta_p^2<.001$) と遮蔽なし条件 ($F(1, 55)=2.72$, $p=.11$, $\eta_p^2=.05$) ではみられなかった。交互作用効果の検出力を調べるために G*Power を使用して事後の検定力検定を行ったところ、 $\alpha=.05$ の場合検定力 ($1-\beta$)=0.99 となり、分析に際して十分なサンプルサイズが確保されていると判断された。この結果は、仮説 2b を支持するものである。パンデミック下におけるマスク顔への接触が感情表出者の目領域の情報への依存を促進し、顔上部からの感情認知精度が高くなった結果、聴覚の手がかりである声に対する情報の重みづけが小

Table 1 多感覚感情一致および単感覚刺激への正答率

	多感覚セッション 感情一致		単感覚セッション 顔のみ		単感覚セッション 声のみ	
	前	後	前	後	前	後
遮蔽なし	96.2 (1.0)	96.3 (1.4)	96.7 (0.8)	96.3 (0.9)		
口遮蔽	93.1 (1.0)	95.0 (1.4)	87.3 (0.9)	90.7 (1.4)	93.5 (1.1)	94.4 (1.2)
目遮蔽	97.3 (0.7)	95.7 (1.4)	96.4 (0.7)	95.5 (1.1)		

注) 括弧内は標準誤差を示す。

さくなった可能性がある。もしそうであるならば、口が遮蔽された顔からの単感覚的な感情認知精度がパンデミック前よりも後に向上しているはずである。そこでこの可能性を検証するため、単感覚セッションにおける顔のみ刺激に対する正答率 (Table 1) について、時期×顔遮蔽条件の ANOVA を実施した。その結果、顔遮蔽条件の主効果 ($F(1.65, 90.88)=47.60$, $p<.001$, $\eta_p^2=.46$) と、時期×顔遮蔽条件の交互作用 ($F(1.65, 90.88)=3.88$, $p=.03$, $\eta_p^2=.001$) が有意だった。単純主効果検定を実施したところ、パンデミック前後での顔からの感情認知精度の差異は、口遮蔽条件でのみみられ ($F(1, 55)=4.26$, $p=.04$, $\eta_p^2=.07$), 目遮蔽条件 ($F(1, 55)=0.08$, $p=.78$, $\eta_p^2=.001$) と遮蔽なし条件 ($F(1, 55)=0.50$, $p=.48$, $\eta_p^2<.001$) ではみられなかった。口が遮蔽された顔に対する感情認知正答率のみパンデミック後に向上していたことから、パンデミック下において、口が遮蔽されたマスク顔に接触したことにより、顔上部からの感情認知精度が向上したと解釈することができる。したがって、パンデミック後に多感覚感情認知における声に対する情報の重みづけが小さくなったのは、顔上部からの感情認知精度が向上したことが影響している可能性が示唆された。

4. 考察

本研究では、多感覚感情認知における顔と声に対する情報の重み付けが、顔の部分的な遮蔽によってどのように影響を受けるのか、またその影響の受け方は COVID-19 パンデミックの前後で変化がみられるのかを検討することを目的とした。実験の結果、まず感情表出者の顔の一部が遮蔽されていない場合と比較して、口または目の遮蔽によって視覚的手がかりが減少した場合は、多感覚感情認知において声の感情を重視する程度が大きくなることが明らかとなった。これらの結果は、視覚情報の劣化が多感覚感情認知に影響することを報告した先行研究 [10] の知見を拡張し、顔の部分的な欠損も多感覚感情認知に影響することを示唆するものである。

さらに本研究では、多感覚感情認知における顔と声の情報に対する重みづけが COVID-19 パンデミック前後で変化したことが示され、パンデミック前に比べて後の方が、口が遮蔽された感情表出者に対して声の感情を重視する程度が小さいことが明らかとなった。マスク顔に接触することによって、感情の手がかりとし

て相手の目領域への依存度が高まり、表情認知パターンが変化することが先行研究によって報告されており [11]、本研究の結果もこれに合致している。顔からの単感覚的な感情認知精度は、口が遮蔽された顔に対してのみ、パンデミック後に向上していたことから、パンデミック下におけるマスク顔への接触が、目領域を手がかりとした感情認知とその精度を促進し、その結果として多感覚感情認知における声の情報への重みづけが小さくなった可能性がある。マスク顔への接触によって、表情認知だけでなく、顔と声からの多感覚的な感情認知様式も変容することが本研究によって明らかとなった。パンデミック前後で多感覚感情認知における情報の重みづけが異なるということは、感情表出者が普段のマスク非着用状態と同じように感情を表出しても、マスク着用時は口元が遮蔽されていることによって、声の感情に注意が向けられにくく、顔が遮蔽されていない場合とは異なった感情の受けとられ方がなされる可能性がある。普段の感情表現においては、視聴覚のモダリティ間で感情が一致しない場合も少なくないことが報告されている [13]。また、顔の表情は実際とは別の感情を示すことができ、特に微笑みは感情を隠蔽する際に最も頻繁に使われる一方、音声は表情よりも感情を隠すことが難しいと指摘されている [14]。したがって、たとえば笑顔によって隠蔽されたネガティブな感情が声に表れていた場合、マスク着用時にはそのネガティブ感情に気づくことができず、コミュニケーションの齟齬につながる恐れもある。ただし、実際にマスクを着用した対面コミュニケーションにおいては、顔の可視領域が減少するだけでなく、声の明瞭度も低下する [15]。聴覚的手がかりの明瞭さの操作を行っていないことが本研究の限界として指摘でき、今後の研究では、視聴覚情報のいずれの明瞭さも低下した場合に、多感覚感情認知における顔と声の情報に対する重みづけがどのように変化するか検討する必要がある。さらに、マスク顔への接触頻度の増加に伴い、顔表情の見方とは独立に聴覚情報（つまり音声）の処理そのものに変化が生じ、それが結果に影響した可能性や、本研究の結果が他の文化圏で再現されるかについては未検討であり、今後の検討課題である。

5. 謝辞

本研究は、JSPS 科研費 No. JP 22H01106 の助成を受けたものです。

6. 引用文献

- [1] de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, 14(3), 289–311. <https://doi.org/10.1080/026999300378824>
- [2] de Gelder, B., Vroomen, J., & Bertelson, P. (1998). Upright but not inverted faces modify the perception of the emotion in the voice. *Neural Substrates of Cognitive Processes. Meeting*, 17, 1021–1031.
- [3] Pell, M. D. (2005). Prosody–face Interactions in Emotional Processing as Revealed by the Facial Affect Decision Task. *Journal of Nonverbal Behavior*, 29(4), 193–215. <https://doi.org/10.1007/s10919-005-7720-z>
- [4] Pourtois, G., De Gelder, B., Vroomen, C. J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11(6), 1329–1333. <https://doi.org/10.1097/00001756-200004270-00036>
- [5] Tanaka, A., Koizumi, A., Imai, H., Hiramatsu, S., Hiramoto, E., & De Gelder, B. (2010). I Feel Your Voice. *Psychological Science*, 21(9), 1259–1262. <https://doi.org/10.1177/0956797610380698>
- [6] Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, 37(11), 2049–2058. <https://doi.org/10.1037/0022-3514.37.11.2049>
- [7] Roberson, D., Kikutani, M., Döge, P., Whitaker, L., & Majid, A. (2012). Shades of emotion: What the addition of sunglasses or masks to faces reveals about the development of facial expression processing. *Cognition*, 125(2), 195–206. <https://doi.org/10.1016/j.cognition.2012.06.018>
- [8] Carbon, C. (2020). Wearing face masks strongly confuses counterparts in reading emotions. *Frontiers in Psychology*, 11(2020), 566886. <https://doi.org/10.31234/osf.io/sd5mj>
- [9] Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PLOS ONE*, 16(4), e0249792. <https://doi.org/10.1371/journal.pone.0249792>
- [10] Kokinous, J., Tavano, A., Kotz, S. A., & Schröger, E. (2017). Perceptual integration of faces and voices depends on the interaction of emotional content and spatial frequency. *Biological Psychology*, 123, 155–165. <https://doi.org/10.1016/j.biopsycho.2016.12.007>
- [11] Barrick, E. M., Thornton, M., & Tamir, D. I. (2021). Mask exposure during COVID-19 changes emotional face processing. *PLOS ONE*, 16(10), e0258470. <https://doi.org/10.1371/journal.pone.0258470>
- [12] Lu, J. G., Jin, P. J., & English, A. S. (2021). Collectivism predicts mask use during COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*, 118(23). <https://doi.org/10.1073/pnas.2021793118>
- [13] 高木幸子, 田部井賢一, 田中章浩. 2013 表情と音声の示す感情が一致していない刺激からの感情知覚: 異文化間バーチャル・リアリティ・コミュニケーションへの応用 (<特集> 基礎心理学の実用, 応用展開), 基礎心理学研究, 32(1), pp.29-39. <https://doi.org/10.14947/psychono.KJ00008988953>
- [14] Ekman, P. (1985). *Telling lies: Clues to Deceit in the Marketplace, Politics and Marriage*. New York: W. W. Norton & Company.
- [15] Goldin, A., Weinstein, B., & Shiman, N. (2020). How do medical masks degrade speech perception. *Hearing Review*, 27(5), 8–9. <https://hearingreview.com/hearing-loss/health-wellness/how-d-o-medical-masks-degrade-speech-reception>