

顔と声の連合記憶の忘却特性¹⁾

遠藤信貴*

Forgetting properties of associative memory between faces and voices

Nobutaka ENDO

Abstract

The purpose of this study was to clarify the forgetting properties of associative memory between face and voice. In the learning phase, participants repeatedly learned the association between face images and voice samples obtained from unknown people until their matching accuracy rate exceeded 80%. In the recognition phase, participants were tested immediately after the learning phase, then three days later, and finally seven days later. Participants were required to judge whether they remembered both the face and voice, only the face, only the voice, or neither one of them. The results showed that the recognition performance for faces was higher than that for voices, and the retention interval from the learning phase to the recognition test did not affect the performance in recognizing either the face or voice. However, the superiority of face recognition performance was confirmed only when the association between face and voice worsened, indicating that voice recognition is not always inferior to face recognition. For these results, the forgetting properties of both face and voice were discussed in terms of the discrimination sensitivities for face and voice.

Keywords : ① face and voice memory ② forgetting property ③ discrimination sensitivity

問 題

我々は日常の行動場面において様々な情報を記録し、それらを記憶に保持している。記憶に保持される情報には、自覚的で言語化できる個人的な経験や一般的な知識だけではなく、音色や質感といった常に適切な言語化ができるとは限らない情報をも含むため、保持されている情報量は膨大である。日常生活において、他者とのコミュニケーションは不可欠であり、保持される情報には人物に関する情報も含まれる。我々は、主に言語のやり取りや表情による感情の表出などを介して対人関係を構築していくが、その過程で収集される人物に関する情報から特定の人物の記憶を形成していく。

顔は日常の対人コミュニケーション場面において接触頻度の高い人物情報である。言葉を交わしたことはなくても顔だけは知っているということはよくある。顔は表情や視線によって変化し、また、年齢に伴って変化していくが、顔の記憶にはこれらの変化に関わらず、長期的かつ安定的な人物想起が可能な頑健性がある。顔の記憶の特殊性の1つとして典型性効果が挙げられる。典型性効果とは、ある集団の中で平均的な特徴をもつ顔（プロトタイプ）は非典型的あるいは示差的な顔に比べて記憶成績が低くなるというものである。典型性効果は顔の符号化や表象の検索過程を明らかにする手掛かりとして重要な現象であるが、顔の符号化に関し

受付：令和3年5月28日 受理：令和3年6月29日

*近畿大学総合社会学部 准教授（認知心理学）

1) 本論文は、著者の指導のもとで秋山沙耶氏が近畿大学総合社会学部に提出した2018年度卒業論文で用いたデータを再分析し、再解釈のもとで改稿したものである。

て Valentine & Bruce (1986) は、平均的な顔を基準とし、その基準からのずれに基づいて個人の顔が符号化されているとするモデルを提唱している。また、近年、大脳皮質の紡錘状回には顔刺激に対して特に選択的に活動する領域として紡錘状回顔領域の存在も明らかにされており (Kanwisher, McDermott, & Chun, 1997)、この神経科学的知見は顔の記憶の他の視覚的記憶に対する特殊性を示すものと言える。

一方、声もまたコミュニケーション場面で重要な意味を持ち、顔と同様に多様性の高い人物情報である。Legge, Grosman, & Pieper (1984) は、実験参加者に未知の人物の声刺激に対する偶発学習をさせた後、学習した声刺激と未学習の声刺激の二肢強制選択による再認成績を検討した。その結果、声の学習は顔の学習よりも劣るものの、声刺激の呈示時間が長くなるに従って、学習から15分後と10日後の再認成績は向上することが明らかにされている。また、Mullennix et al. (2011) は典型性が高いと評価された音声と低いと評価された音声の偶発学習課題を行い、1週間後の再認成績を検討した。その結果、全体的に典型的な声ほど再認成績は低く、声の記憶に関しても顔と同様に典型性効果が生じることを明らかにした。また、声情報には話者の話し方の特徴として、ピッチやテンポ、イントネーション、声質などが含まれる。Mullennix et al. (2010) は、発話時の特徴として声のピッチとテンポに着目し、それらの特徴の違いが声の再認記憶に及ぼす影響を検討した。その結果、学習した声刺激（ターゲット）と未学習のディストラクタに対して、ピッチに関してはターゲットのピッチが高いときはより高いピッチのディストラクタを、ターゲットのピッチが低いときはより低いピッチのディストラクタをターゲットとして誤って選択する傾向が示され、声のピッチの高低についての記憶はあるものの、その正確なピッチを弁別することは難しいことを明らかにした。一方、発話テンポに関してはテンポの違いによる影響は確認されなかった。Mullennix et al. (2010) や Mullennix et al. (2011) の研究は、声の特性の

中でも発話時に伴う特徴による再認成績への影響に着目したものであるが、声記憶の忘却特性については触れられていない。実際、Mullennix et al. (2010) も Mullennix et al. (2011) も声記憶の忘却特性を検討するには保持期間の設定が短かったことを問題として指摘しており、Mullennix et al. (2010) は学習フェーズ後の一定の保持期間を設けることで、直後再認の成績に比べて再認成績は低くなる可能性に言及している。これらを踏まえると声記憶の特性を明らかにするうえで、記銘プロセスや保持期間などの要因が声の再認に及ぼす影響という観点からの検討が必要であろう。

我々はコミュニケーションを通じて初対面の人物を熟知していく過程において、通常、顔あるいは声だけというように、その人物の特徴を個別に覚えていくということはない。コミュニケーションを重ねることにより、特定の人物の顔と声の特徴の結びつきは強化され、それにより特定の人物の記憶は形成されていくものである。人物の連合記憶に関して、視覚的情報である顔と聴覚的情報である声が記憶の形成過程においてどのように影響するのかについて、Stevenage, Howland, & Tippelt (2011) の研究が挙げられる。Stevenage et al. (2011) は、顔画像あるいは音声のみを記銘刺激として呈示する単独条件と、顔画像と音声の両方を記銘刺激として呈示する連合条件を設け、学習フェーズでは呈示される刺激の魅力度を評価させる偶発学習を行った。その後、テストフェーズでは学習フェーズで呈示されたターゲット刺激と未呈示のディストラクタ刺激に対して、学習フェーズで呈示されたものであるか否かの弁別課題を行った。その結果、連合条件における音声の弁別成績は単独条件に比べて低かった。しかし、顔画像の弁別成績は単独条件と連合条件の間に差は見られなかった。このことから、顔と声の連合記憶の形成過程において、顔情報は声情報の符号化に干渉する可能性が示唆されている。Stevenage et al. (2011) に先立ち、Cook & Wilding (1997) も音声のみを呈示する条件と顔と音声を同時に呈示する条件において、記銘

後1週間の遅延期間において音声に対する再認成績を検討している。その結果、記銘時に音声のみが呈示された条件に比べ、顔と音声と同時に呈示される条件の方が再認成績は低かった。これらの結果は顔と声の同時呈示により、顔情報が声情報の符号化に干渉し、声の正再認を阻害したと解釈することができ、顔情報に比べて声情報の保持機能は脆弱であることが推察される。

声記憶は顔記憶に比べて脆弱であることは示唆されているが (Legge et al., 1984), Kerstholt, Jansen, van Amelsvoort, & Broeders (2004, 2006) は、声記憶における保持時間の経過に伴う再認成績の変化について検討している。2つの研究において設定された保持期間は異なっており、保持期間以外に発話の呈示時間や音の環境 (屋内, 屋外), 発話アクセント, 音声収録方法 (直接, 電話越し) など, 操作された要因は研究間で異なるが, いずれも声の再認成績は全体的に低いものの, 保持期間が長くてもその影響は受けにくいという一貫した結果が得られている。しかし, これらの研究ではターゲットとして用いられた声刺激の発話内容は統一されていなかった。ターゲット刺激によって発話内容が異なることで, 声そのものではなく発話内容に注意が向けられた可能性が考えられる。それにより, 再認時において声質そのものに聞き覚えがあるかという判断ではなく, 発話内容に覚えがあるかどうかによって再認判断がなされた可能性は否定できない。また Kerstholt et al. (2004, 2006) の研究は耳撃証言に関するものであり, 未知の人物の声への限られた接触状況に焦点を当てたものである。そのため, 通常的人物記憶の形成が顔と声の連合学習に基づくことを踏まえると, 両者の連合学習を経たうえでの時間経過に伴う忘却特性については検討の余地がある。

以上を踏まえ, 本研究では, 未知の人物の顔と声の連合学習を経たうえでの保持期間に伴う顔と声に対する再認成績の変化に注目し, それぞれの忘却特性を検討することを目的とする。実験は2つのセッションで構成され, 学習セッ

ションでは顔と声の連合学習を行った。再認セッションでは忘却特性を明らかにするために保持期間が異なる3回の再認テストを行った。また, Kerstholt et al. (2004, 2006) の研究の問題点として挙げた学習時の発話内容の違いが再認判断に及ぼした可能性を考慮し, 学習セッションも再認セッションも呈示する声刺激の発話内容はすべて同一としたが, 3回の再認テストで呈示する声の発話内容は変化させた。一般的な記憶の忘却特性に従えば, 顔と声の連合記憶が形成されてから再認時までの保持期間が長いほど, 顔刺激と声刺激に対する再認成績はいずれも低下することが予測される。また, 人物記憶における声情報の保持の脆弱性 (Cook & Wilding, 1997; Stevenage et al., 2011) を踏まえると, 保持期間に関わらず顔刺激よりも声刺激の再認成績は低くなることが予測される。しかし, 声刺激の再認成績が保持期間の長さの影響を受けにくいという知見 (Kerstholt et al., 2004, 2006) に基づけば, 声刺激の再認成績は保持期間の長さに関係なく同様であり, 顔刺激の再認成績を含めると, 顔と声の再認成績と保持期間には交互作用が生じると予測される。さらに, 本研究では顔と声の再認判断時における弁別感度 (d') に基づいたそれぞれの忘却特性についても検討する。

方法

実験参加者

大学生20名が実験に参加した。全員が裸眼もしくは矯正により健全な視力を有していた。また, 実験者の口頭による実験説明および教示を通じて, 聴力も問題はないことを確認した。本実験計画は近畿大学総合社会学部の研究倫理審査委員会が定める倫理基準を満たしたものであり, すべての実験参加者に対して実験への参加は任意であること, また実験参加の同意は実験中いつでも撤回できることを説明したうえで同意書への署名を求めた。

実験器具

実験はワークステーション (DELL 社製

Precision T3600) を用い、記銘刺激の呈示および測定データの収集は MATLAB ソフトウェア (The Mathworks, Inc) と心理物理実験用の関数ライブラリである Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007) で記述されたプログラムによって実行された。記銘刺激である顔刺激は 22 インチの CRT モニタ (MITSUBISHI 社製 Diamondtron Flat RDF221S) に呈示し、声刺激はヘッドフォン (SONY 社製 MDR-ZX770BN) を通して呈示した。CRT モニタの解像度は横×縦が 1024 × 768 ピクセルであり、実験参加者からの観察距離はあご台を用いて約 57 cm になるように設定した。課題遂行時の反応取得にはマウスとキーボードを用いた。

実験刺激

本実験で用いる顔画像と音声データを収集するために、本実験参加者とは異なる大学生の男女 144 名 (男性 92 名, 女性 52 名) に顔撮影と音声データの録音を依頼した。顔画像と音声データは実験刺激として呈示すること、本実験以外の目的で使用したり、外部に公表したりすることはないことを説明し、全員から撮影と録音の同意を得た。

顔画像の表情は真顔とした。呈示する顔画像の顔領域はほぼ同一になるようにサイズを調整し、また、顔部分が画像の中心に配置されるように画像の上下左右の両端をトリミングすることで、画像全体のサイズは横×縦が 900 × 675 ピクセルに統一した。画像はグレースケールで呈示された。なお、実験者と他 1 名の判断により、示差性の高いと思われる画像は顔刺激からは除外した。音声データの収集では、できるだけ普段の声かけと同じように発話し、可能な限り標準語で発音するように求めた。発話内容は「おはよう」「こんにちは」「ひさしぶり」「ありがとう」の 4 種類であった。音声データについても示差性の高い声質や発話のものは声刺激からは除外した。以上により、男性 49 名, 女性 49 名の計 98 種類の顔刺激と声刺激の組み合わせを実験刺激とした。

実験計画

本実験は学習セッションと再認セッションから構成された。学習セッションでは、ランダムに選ばれた男女 6 名ずつの計 12 種類の顔刺激と声刺激の組み合わせ (OLD ペア) に対して意図的な記銘課題を行った。再認セッションでは、学習セッションで記銘した OLD ペアと、学習セッションで記銘していない男女 6 名ずつの計 12 種類の顔刺激と声刺激の組み合わせ (NEW ペア) に対する再認テストを行った。再認セッションの再認テストは、学習セッション終了直後、学習セッションから 3 日後、7 日後の計 3 回行った。以上により、再認セッションは刺激の種類 (OLD, NEW) と再認テスト時期 (直後, 3 日後, 7 日後) の 2 要因実験参加者内計画で行われた。

手続き

実験参加者は CRT モニタの正面に着席し、あご台によって観察距離を一定に保った状態でモニタを観察した。学習セッションは実験参加者がマウスの左右どちらかのボタンをクリックすることで開始された。はじめに灰色背景の画面中央部に白色の注視点 (+) が 500 ms 呈示された。それに続いて顔刺激と同時にヘッドフォンからその人物の声刺激が呈示された。このとき、声刺激はマウスを左クリックすることで何度でも聴き直すことができ、その間は顔刺激も呈示されたままであった。実験参加者にはそれぞれの人物の顔と声の組み合わせをしっかりと覚え込むように教示した。学習セッションで用いた声刺激の発話内容はすべて「おはよう」であった。12 種類の顔と声の組み合わせを記銘した後、記銘内容の正確さを測るために確認テストを行った。確認テストでは、記銘した 12 種類の顔と声の組み合わせのうち、半数 (男女 3 名ずつ) はそのままの組み合わせとし、残り半数は記銘時には呈示しなかった同性の人物の声刺激に変化させた計 12 種類の刺激に対して、その組み合わせの正誤の判断課題を行った。注視点を 500 ms 呈示した後、12 種類の刺激をランダムな順で呈示した。声刺激は 2 回繰

り返し呈示され、顔刺激はその間呈示されたままであった。確認テストでは正答率が80%未満であった場合は記銘が不十分であるとみなし、組み合わせの正誤判断を誤った顔刺激と声刺激のオリジナルの組み合わせを再学習させ、その後確認テストを行った。この手続きは正答率が80%以上になるまで繰り返され、学習セッションは終了した。

再認セッションでは学習セッション終了直後と3日後、7日後に3回の再認テストを行った。学習セッションで記銘した12種類の顔画像と音声の組み合わせ(OLDペア)に加え、未学習の12種類の組み合わせ(NEWペア)の計24種類の組み合わせを呈示した。3回の再認課題で呈示される12種類のNEWペアは未呈示のペアから選択され、毎回異なるものであった。音声刺激のメッセージは、学習セッション直後の1回目の再認テストでは「こんにちは」、3日後の2回目では「ひさしぶり」、7日後の3回目では「ありがとう」とした。実験参加者には呈示される顔刺激と声刺激の組み合わせが学習セッションで記銘したものであるかどうかを正確に判断するように教示した。再認テストは、注視点(+)が500ms呈示された後に顔刺激と声刺激が同時に呈示された。音声は3回繰り返し呈示され、顔刺激はその間呈示されたままであった。刺激呈示後、呈示された人物の顔と声に覚えがあるかどうかの判断を求めた。判断は、(1)顔と声の組み合わせに覚えがある、(2)顔に見覚えはあるが声に聞き覚えはない、(3)声に聞き覚えはあるが顔に見覚えはない、(4)顔にも声にも覚えはない、の4つの選択肢から強制選択させた。実験に要した時間は、学習セッションが約30min、再認セッションの3回の再認テストはそれぞれ約15minであった。

結果

20名の実験参加者のうち、再認セッションで実施された3回の再認テストすべてに参加できなかった3名のデータは分析対象から除外したため、17名分のデータを分析対象とした。学習セッション終了に要した平均学習回数

は2.41回($SD = 1.42$)、最終の平均正答率は87.3%($SD = 5.80$)であった。再認データは、OLDペアに対する正再認(Hit)とNEWペアに対するフォールスアラーム(FA)の割合を再認テスト時期ごとに算出し、その変化を分析した。また、OLDペアとNEWペアに対する弁別感度(d')を算出し、その変化についても分析した。

Hit率とFA率

OLDペアに対して、顔と声の組み合わせに覚えがあると反応した場合はBoth-Hit、顔に見覚えはあるが声に聞き覚えはないと反応した場合はFace-Hit、声に聞き覚えはあるが顔に見覚えはないと反応した場合はVoice-Hitとして、それぞれの反応数をOLDペア数12で割った値をHit率の分析対象とした。また、NEWペアに対して、顔と声の組み合わせに覚えがあると反応した場合はBoth-FA、顔に見覚えはあるが声に聞き覚えはないと反応した場合はFace-FA、声に聞き覚えはあるが顔に見覚えはないと反応した場合はVoice-FAとして、それぞれの反応数をNEWペア数12で割った値をFA率の分析対象とした。

3回の再認テストにおける反応タイプ(Both-Hit, Face-Hit, Voice-Hit)ごとのHit率はFigure 1に示すとおりである。反応のタイプと再認テスト時期(直後、3日後、7日後)の2要因分散分析の結果、反応のタイプの主効果が有意であった($F(2, 32) = 131.49, p < .001, \eta_p^2 = .892$)。多重比較の結果、Both-Hit(.632)はFace-Hit(.322)とVoice-Hit(.016)に比べて有意に高く($ts(16) > 6.18, ps < .001, ds > 2.17$)、Face-HitはVoice-Hitに比べて有意に高かった($t(16) = 12.12, p < .001, d = 3.94$)。再認テスト時期の主効果は有意ではなく($F(2, 32) = 0.04, p = .952$)、交互作用も有意ではなかった($F(4, 64) = 2.13, p = .122$)。

各再認テスト時期における反応のタイプ(Both-FA, Face-FA, Voice-FA)ごとのFA率はFigure 2に示すとおりである。2要因分散分析の結果、反応のタイプの主効果が有意であっ

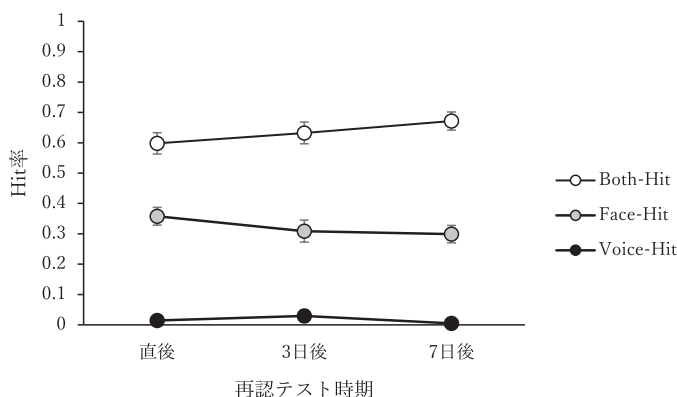


Figure 1 各再認テスト時期における Hit 率の変化

た ($F(2, 32) = 140.62, p < .001, \eta_p^2 = .898$). 多重比較の結果, Voice-FA (.314) は Both-FA (.007) と Face-FA (.023) に比べて有意に高く ($t(16) > 11.43, p < .001, ds > 3.61$), Face-FA は Both-FA に比べて有意に高かった ($t(16) = 3.41, p = .004, d = 0.75$). 再認テスト時期の主効果は有意ではなく ($F(2, 32) = 2.45, p = .107$), 交互作用も有意ではなかった ($F(4, 64) = 1.83, p = .162$).

顔刺激と声刺激に対する弁別感度

Hit 率と FA 率より, 顔刺激と声刺激に対する各再認テスト時期での弁別感度 (d') を求めた. d' とは信号検出理論における信号 (OLD ペア) を含まないノイズ分布と信号を含む信号

+ノイズ分布の頂点間の距離であり, d' の値が大きいほど OLD ペアと NEW ペアの弁別が良好であることを意味する. また, d' は実験参加者の判断基準に影響されない純粋な信号検出力の指標とされる. d' は各実験参加者の顔刺激と声刺激に対する Hit 率と FA 率をプロビット変換し, 標準正規分布上での Hit 率に対応する z 得点と, FA 率に対応する z 得点の差分から求められるが, Hit 率や FA 率が 0% や 100% の場合, d' を定義することができない. Hit 率と FA 率の分析において, Voice-Hit, Face-FA, Both-FA の平均は非常に低く, 0 に近い値であった. そこで d' の分析は, 各再認テスト時期における Both-Hit と Face-Hit の合計を顔刺激への Hit 率, Both-Hit と Voice-Hit の合計を声刺激への

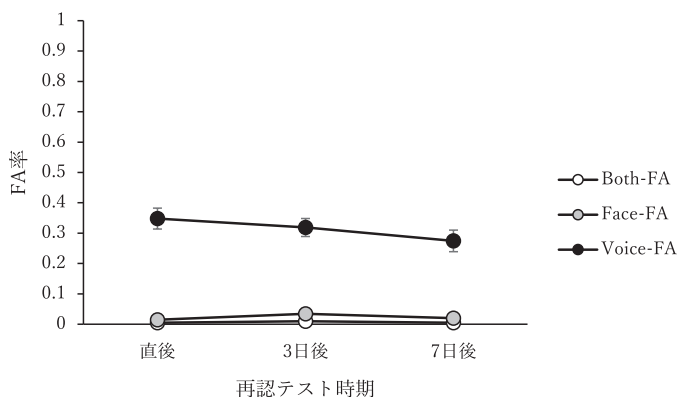


Figure 2 各再認テスト時期における FA 率の変化

Hit 率とした。また、Both-FA と Face-FA の合計を顔刺激への FA 率、Both-FA と Voice-FA の合計を声刺激への FA 率とした。しかし、この手順により各実験参加者の顔刺激と声刺激に対する Hit 率と FA 率を算出したものの、一部の実験参加者において Hit 率が 100%、FA 率が 0% となったため、Macmillan & Creelman (1991) の方法によってデータの補正を行った。本実験では、OLD ペアと NEW ペアはそれぞれ 12 種類であったことから、Hit 率が 100% のときの補正值は 95.83% (1/24 を 1 から引いたもの)、FA 率が 0% のときの補正值は 4.17% (1/24) として分析した。

各再認テスト時期における顔刺激と声刺激に対する d' は Figure 3 に示すとおりである。刺激の種類（顔、声）と再認テスト時期の 2 要因分散分析の結果、刺激の種類の主効果が有意であり ($F(1, 16) = 361.26, p < .001, \eta_p^2 = .958$)、顔刺激に対する d' (3.15) は声刺激 (0.93) よりも有意に高かった。また、交互作用に有意傾向が見られた ($F(2, 32) = 2.64, p = .087, \eta_p^2 = .142$)。下位検定の結果、声刺激における再認テスト時期の単純主効果が有意であった ($F(2, 64) = 4.61, p = .022, \eta_p^2 = .233$)。多重比較の結果、直後の再認テスト時 (0.71) に比べて 3 日後 (0.93) と 7 日後 (1.16) の d' は有意に高かった ($t_s(16) > 2.22, ps < .05, ds > 0.40$)。3 日後と 7 日後の d'

に差は見られなかった。

考察

本研究の目的は、未知の人物の顔と声を連合させる反復学習を経たうえでの顔と声の記憶の忘却特性を明らかにすることであった。再認テストでは、OLD ペアと NEW ペアのそれぞれに対して、顔と声の組み合わせに覚えがある、顔のみに見覚えがある（声には聞き覚えがない）、声のみに聞き覚えがある（顔には見覚えがない）、そして顔と声の組み合わせに覚えはない、のいずれかの判断を求めた。さらに、再認判断時における実験参加者の判断基準に影響されない純粋な OLD ペアに対する弁別感度の観点からも忘却特性を検討した。

実験の結果、再認テストの時期に関係なく一貫して、顔と声の OLD ペアに対する Hit 率 (Both-Hit) は高く、FA 率 (Both-FA) は極めて低かった。また、顔のみの Hit 率 (Face-Hit) は声のみの Hit 率 (Voice-Hit) に比べて一様に高く、声のみの FA 率 (Voice-FA) は顔のみの FA 率 (Face-FA) に比べて一様に高かったが、Face-FA は極めて低かった。Face-Hit が Voice-Hit よりも高く、Face-FA がほとんど観察されなかったことは顔記憶が声記憶に比べてその正確さが安定的に高いことを示唆するものであり、声記憶の顔記憶に対する脆弱さを示した Legge et al. (1984) の研究結果に整合するもの

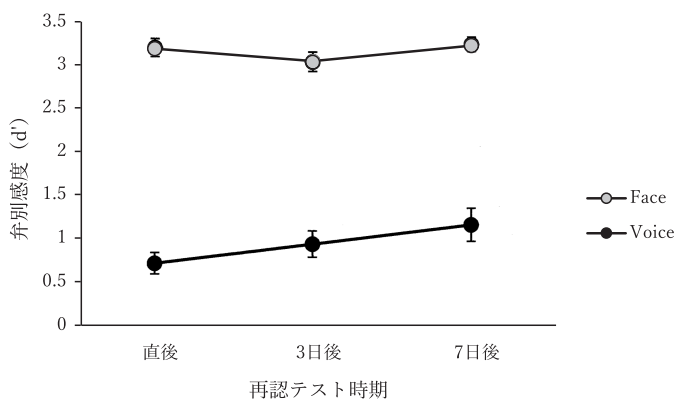


Figure 3 各再認テスト時期における顔刺激と声刺激に対する弁別感度

である。Cook & Wilding (1997) や Stevenage et al. (2011) は、顔と声の連合記憶の形成過程において、顔情報による声情報への干渉が生じることで、声の再認は阻害されることを示唆しているが、本研究の結果からも Voice-Hit はほとんど観察されなかったのに対して Face-Hit が一貫して高いことが示され、この声の再認成績の低さは符号化時における顔刺激の干渉によって説明できるかもしれない。しかし、本研究では Both-Hit は他に比べて安定的に高く、Both-FA もほとんど観察されていないことを踏まえると、声刺激の符号化時における顔刺激の干渉は常に生じるわけではないと考えられる。つまり、顔と声の連合が維持されている限りは声の再認も安定的になされる可能性が示唆される。耳撃証言研究においても音声ラインナップからターゲットを再認させる際、音声と同時に話者の顔を呈示すると再認成績が低下するという顔隠蔽効果 (Heath & Moore, 2011) が報告されているが、本研究の結果を踏まえると、顔隠蔽効果は単に再認時に顔を呈示することによる特有の現象ではなく、声刺激の符号化、つまり耳撃時の状況にかなり依存している可能性が考えられる。

本研究では、顔と声の記憶の忘却特性を明らかにするために、再認テストを3回に渡って行った。その結果、顔と声のペアに対する再認成績について再認テスト時期による影響は確認されなかった。また、Face-Hit が示す顔のみの再認成績も同様の傾向であった。顔は日常的な人物同定の際に重要な手掛かりであり、一般的にもその頑健性は確認されている。本研究では保持期間を最長で7日に設定したが、一般的な記憶の忘却特性からすれば一定の忘却は生じる長さである。少なくとも7日程度の保持期間では学習した顔を忘却するまでには至らなかったと考えられ、顔記憶は保持期間の影響を受けにくいという特性をもつと言えるかもしれない。一方、声の再認成績については顔と声の連合が維持されている限りは安定的に高い再認成績を示しており、保持期間による影響も見られなかったが、Voice-Hit が示す声のみの再認成績

が極めて低かったことは3回の再認テストにおいて一貫していた。3回の再認テストにおける Voice-Hit の平均 Hit 率は 1.63% であったが、補足的に行った分析により、この Hit 率は 0 より高い傾向であった ($t(16) = 1.61, p = .063$)。これを踏まえると、本実験における声のみの再認成績は極めて低いものの、保持期間の影響は受けないとする Kerstholt et al. (2004, 2006) の研究と整合する結果と言えるだろう。

本研究では、Hit 率と FA 率に基づいた再認判断時における弁別感度についても検討した。その結果、顔刺激の弁別感度は再認テスト時期に関わらず一貫して声刺激の弁別感度よりも高かった。しかし、顔と声の弁別感度の保持期間による変化には異なる傾向が示された。顔の弁別感度は再認成績と同様に保持期間に関係なく一様であったのに対して、声刺激の弁別感度は直後の再認テスト時に比べて3日後と7日後の再認テスト時では上昇した。このことは、声の記憶と顔の記憶の異なる特性の1つとして解釈できるかもしれない。本研究では3回の再認テストにおいて OLD ペアは繰り返し呈示された。顔記憶の優位性の観点から学習時に形成された顔刺激の表象が非常に安定的であったとしても、再認テストで繰り返し OLD ペアに接することでその表象はより強化されることになる。声刺激の表象も同様に OLD ペアに繰り返し接することが表象の強化をもたらしたとすれば、弁別感度の上昇は顔と声の両方において確認されるはずである。しかし声に対する弁別感度のみには上昇が見られたことは、声刺激の表象が精緻化され再認判断に寄与するには顔刺激よりも長い時間を要することを示唆しているのかもしれない。本研究では、Voice-Hit, Face-FA, Both-FA の割合が 0 に近かったことから、Both-Hit や Both-FA を含めた分析を行っているため、今後詳細な検討が必要ではあるものの、声刺激の弁別感度の保持期間に伴う上昇傾向は、声記憶の長期持続性を明らかにするうえで重要な意味をもつと考えられる。

顔の記憶における典型性効果は記銘対象となる顔のもつ特徴が記憶成績に強く影響すること

を意味するが、これは声の記憶においても同様である (Mullennix et al., 2011)。本研究では顔刺激と声刺激に対する典型性や声質についての厳密な統制は行わなかった。このことが顔や声の OLD ペアと NEW ペアの弁別に影響した可能性は否定できないため、十分な統制手続きの下での再検証は必要である。また、本研究は顔と声の連合記憶を促す目的で意図的な反復学習を行い、顕在的な再認識課題によって忘却特性を検討した。日常における人物記憶の形成や人物同定では、意識的な記銘や想起を伴わない自動的な処理に基づくことも多い。記憶研究において記憶の顕在成分と潜在成分の区分 (Jacoby, 1991) については長く議論が続いており、想起時における顕在意識による記憶成績への影響も指摘されていることから、声記憶の長期持続性、忘却特性のさらなる理解に向けて潜在記憶の枠組みでの検討も必要である。

引用文献

- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433-436.
- Cook, S., & Wilding, J. (1997). Earwitness testimony 2: Voices, faces, and context. *Applied Cognitive Psychology*, **11**, 527-541.
- Heath, A. J., & Moore, K. (2011). Earwitness memory: Effects of facial concealment on the face overshadowing effect. *International Journal of Advanced Science and Technology*, **33**, 131-140.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, **30**, 513-541.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, **17**, 4302-4311.
- Kerstholt, J. H., Jansen, N. J. M., van Amelsvoort, A. G., & Broeders, A. P. A. (2004). Earwitness: Effects of speech duration, retention interval and acoustic environment. *Applied Cognitive Psychology*, **18**, 327-336.
- Kerstholt, J. H., Jansen, N. J. M., van Amelsvoort, A. G., & Broeders, A. P. A. (2006). Earwitness: Effects of accent, retention and telephone. *Applied Cognitive Psychology*, **20**, 187-197.
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). "What's new in psyctoolbox-3?". *Perception (ECVP Abstract Supplement)*, **36**, 14.
- Legge, G. E., Grossmann, C., & Pieper, C. M. (1984). Learning unfamiliar voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 297-303.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- Mullennix, J. W., Ross, A., Smith, C., Kuykendall, K., Conard, J., & Barb, S. (2011). Typically effect on memory for voice: Implication for earwitness testimony. *Applied Cognitive Psychology*, **25**, 29-34.
- Mullennix, J. W., Stern, S. E., Grounds, B., Kalas, R., Flaherty, M., Kowalok, S., May, E., & Tessmer, B. (2010). Earwitness memory: Distortion for voice pitch and speaking rate. *Applied Cognitive Psychology*, **24**, 513-526.
- Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.
- Stevenage, S. V., Howland, A., & Tippelt, A. (2011). Interference in eyewitness and earwitness recognition. *Applied Cognitive Psychology*, **25**, 112-118.
- Valentine, T., & Bruce, V. (1986). The effect of distinctiveness in recognizing and classifying faces. *Perception*, **15**, 525-535.