アバタの見た目および音声の変化によるマルチ文脈効果の検証

宮本 健太　【 教育情報工学研究室 】

A Validation of Multi-Context Effects Induced by Changes in Avatar Appearance and Voice

MIYAMOTO, Kenta　【 Educational Information Systems Lab. 】

1　はじめに

近年，オンライン学習やデジタル教材の普及に伴い，遠隔講義が広く実施されている．遠隔講義特有の手法の活用により，対面講義を越える教育効果が期待される．

記憶研究の分野では，複数の異なる環境的文脈（部屋や空間等）で学習を行うことで検索手がかりが増加し，想起成績が向上するマルチ文脈効果が報告されている．また，教育分野では，話者に関する情報が社会的文脈として符号化され，学習内容の記憶に影響を及ぼす可能性が指摘されている．実際に，教員の見た目をエイリアンやロボット等のアバタに変化させることでマルチ文脈効果が生起し，講義におけるテスト成績が向上したと報告されている [1]．しかし，すべてを人アバタ（外見がリアルな人）とした場合には，外見的差異が小さく，学習者にとって異なる文脈として十分に知覚されず，マルチ文脈効果が生起しなくなる可能性も指摘されている [1]．一方で，すべてが人アバタであっても，教員の外見の違いに基づく弁別は容易であり，マルチ文脈効果が生起する可能性がある．また，音声合成技術の発展により，同一内容を複数の話者音声で呈示でき，音声の変化によって教員の弁別が強化され，マルチ文脈効果がより強く生起する可能性がある．しかし，アバタの外見と音声という複数の社会的文脈を同時に操作した際の相乗的なマルチ文脈効果については明らかにされていない．

そこで本研究では，アバタおよび音声という社会的文脈に着目し，マルチ文脈効果が生起する条件やその限界の検討を目的とする．これは，デジタル教材設計および記憶研究の双方に示唆を与えることを意味する．

2　実験

本研究では 2 つの仮説を設定した．

仮説 1：人アバタのみを使用した場合でも，マルチ文脈効果は生起する．

仮説 2：記憶量は，アバタの外見と音声が変化しない場合 < 一方のみ変化する場合 < 両方とも変化する場合の順になる．

実験参加者は，高知工科大学情報学群に所属する学部生および修士生 102 名であった．そのうち，10 名（男性 5 名，女性 5 名）は本実験とは独立して事前アンケートに参加し，92 名（平均年齢 21.70 ± 1.40 歳，男性 72 名，女性 20 名）が単語暗記の本実験を行った．

人アバタ画像および音声を日本人男性に統制した複数パターンを生成し，リップシンク技術を用いて動画刺激を作成した．作成した刺激は，PC 画面上に表示するとともに，音声をイヤホンまたはヘッドホンで呈示した．

事前アンケートは，アバタおよび音声に対する主観的印象の差異が想起数に及ぼす影響の最小化を目的とした．作成したアバタ画像および音声について，2 段階の印象評価（7 件法）を実施し，評価が極端に偏らない 4 パターンの組み合わせを本実験刺激として選定した．

本実験は自由再生課題における想起数を従属変数とし，有意水準は 5%で 3 要因混合計画による比較を行った．1 つ目はアバタの外見であり，1-Avatar 条件（4 セッションすべて同一アバタ）と 4-Avatar 条件（セッションごとに異なるアバタ）の 2 水準を設定した．2 つ目は音声であり，1-Voice 条件（4 セッションすべて同一音声）と 4-Voice 条件（セッションごとに異なる音声）の 2 水準を設定した．これら 2 要因の組み合わせにより，非変化条件（1-Avatar 条件と 1-Voice 条件），音声変化条件（1-Avatar 条件と 4-Voice 条件），アバタ変化条件（4-Avatar 条件と 1-Voice 条件），複合変化条件（4-Avatar 条件と 4-Voice 条件）の 4 条件を構成し，参加者間で比較した．3 つ目はセッションであり，4 つのセッションで参加者内で比較した．

参加者は 4 条件のいずれかに無作為に割り当てた（各 23 名，男性 18 名，女性 5 名）．想起数が 3 点以下の参加者を外れ値として，2 条件において各 1 名（いずれも男性），計 2 名を分析対象から除外した．実験は 1–8 名で同時実施したが，各参加者は個別の機器を使用したため，参加者間の相互影響は生じなかった．

単語暗記課題は 4 つのセッションから構成され，各セッションでは約 2 分間の動画を用いて 25 語の単語を呈示した．今回使用した単語は，藤田ら [2] により作成された 5 段階評定 (1. 低–5. 高) における熟知価が高いものから，清音ひらがな 5 文字の名詞 100 語を選出した．単語はアバタ映像とともに音声および字幕で呈示した．セッション間には 5 分間の休憩を設け，4 セッション終了後は再度 5 分間の休憩の後，8 分間の自由再生課題を

実施した．その後，各セッションで使用されたアバタおよび音声を 8 種類から選択する社会的文脈テストと，事前アンケートと同様の印象評価アンケートを実施した．これは，単語内容の記憶がアバタや音声と結び付いて保持されているかを検討するための探索的指標であった．

3　結果

各条件全体の想起数は，非変化条件が $M = 19.05$（$SD = 11.26$），音声変化条件が $M = 19.87$（$SD = 8.10$），アバタ変化条件が $M = 20.77$（$SD = 9.28$），複合変化条件が $M = 19.96$（$SD = 11.52$）であった．各条件および各セッションにおける単語想起数の分布（0–25 語）を図 1 に示す．×印は平均値を表す．

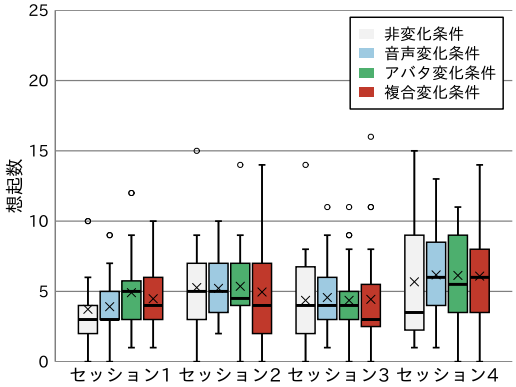

図 1　各セッションにおける条件別単語想起数

各条件ごとに正規性の検定を行った結果，一部の条件において正規性の仮定が棄却されたため（Shapiro-Wilk test，$p < .05$）整列ランク変換を施した後，アバタ×音声×セッションの 3 要因混合計画分散分析を実施した．分析の結果，セッション要因の主効果が有意であった（$F(3, 258) = 9.67, p < .001, \eta_p^2 = .10$）．一方で，アバタ要因の主効果（$F(1, 86) = 0.24, p = .628, \eta_p^2 = .003$），音声要因の主効果（$F(1, 86) = 0.01, p = .917, \eta_p^2 < .001$），および 3 要因間すべての交互作用はいずれも有意ではなかった（いずれも $ps > .10$）．

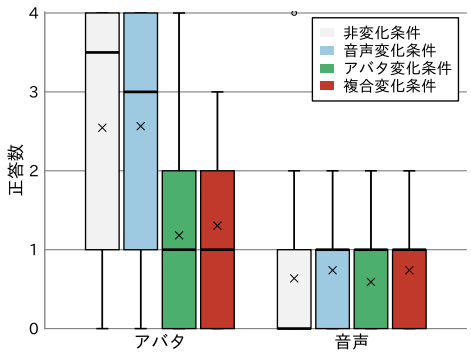

図 2　アバタ，音声における条件別社会的文脈テスト正答数

次に，条件ごとの社会的文脈テストの正答数（0–4）を図 2 に示す．×印は平均値を表す．アバタに関するテストでは，複数条件において正答数が低くなる傾向が見られた．一方で，音声に関するテストでは，条件間で大きな差は見られず，全体として低い正答数であった．

また，社会的文脈テストの正答数と，自由再生課題における合計想起数との関連について，スピアマンの順位相関分析を行った．その結果，アバタおよび音声の同定成績はいずれも合計想起数との間に有意な相関は認められず，相関係数も小さかった（いずれも $|\rho| < .20$, $ps > .10$）．

4　考察

アバタの外見および音声の変化が想起数に与える影響を検討した結果，いずれの要因の主効果および交互作用も有意ではなく，マルチ文脈効果は確認されなかった．

仮説 1 については条件間で想起数に有意差は認められず，この仮説は支持されなかった．本研究で使用したアバタはすべて日本人男性であり，年齢層や服装，背景といった要素も共通していたため，アバタ間の視覚的差異が限定的となり，参加者にとって異なる人物として知覚されにくかった．また，社会的文脈テストにおいて，アバタの同定成績が複数条件で低かった点とも整合的であり，人物アバタの外見が有効な文脈手がかりとして十分に機能しなかったことが，マルチ文脈効果が生起しなかった一因であると考えられる．

仮説 2 については条件間に明確な順序関係は認められず，アバタ要因と音声要因の交互作用も有意ではなかったことから，この仮説は支持されなかった．社会的文脈テストにおいて音声の同定成績が全体的に低かったことから，音声情報は視覚情報と比較して識別が困難であり，音声の変化は有効な文脈手がかりとして機能せず，アバタ要因と音声要因の組み合わせによる相乗的なマルチ文脈効果は生じなかったと考えられる．

社会的文脈テストの正答数と合計想起数との間に有意な相関が認められなかったことから，本研究の条件下では，単語内容の記憶とアバタや音声といった文脈情報は相互に独立して保持されていた可能性が示唆される．

5　まとめ

本研究では，アバタの外見および音声を変化させ，マルチ文脈効果が単語記憶に与える影響を検討した．その結果，本研究の条件設定においては，アバタおよび音声の変化によるマルチ文脈効果は，自由再生課題における単語想起数に明確な影響を及ぼさないと示唆された．

今後は，アバタや音声の差異をより明確に操作し，マルチ文脈効果が生起する条件を検討する必要がある．

参考文献

[1] 瑞穂 嵩人, 雨宮 智浩, 鳴海 拓志, 葛岡 英明, “Virtual Omnibus Lecture: 多様な講師アバタを使った遠隔講義が学生の記憶に与える効果,” 日本バーチャルリアリティ学会論文誌, Vol. 29, No. 3, pp. 109–118, 2024.

[2] 藤田 哲也, 齊藤 智, 高橋 雅延, “ひらがな清音 5 文字名詞の熟知価について,” 京都橘女子大学研究紀要, No. 18, pp. 79–93, 1991.