**序論**

近年，人工知能（AI）の技術は多くの分野で急速に発展しており，合成音声の利用もその一環として広がっている。AIは膨大なデータから学習し，自然なイントネーションで音声を生成する技術を進化させてきた。この技術は報道機関でのニュース読み上げ，動画プラットフォームでの解説や実況に広く利用され，容易に利用可能なものとなっている（NHK，2021）。また，教育や福祉の分野においても，合成音声を活用した技術の研究が進展している(藪，2018; 山岸，2015)。こうした技術の発展により，今後も合成音声は幅広い分野で使用されることが期待される。

音声は人間にとって重要なコミュニケーション手段であると同時に，心身に直接的な影響を与える刺激でもある。音声の種類や特徴によって，人間の感情や生理的反応が変化することが多くの研究で示されている。例えば，宮本・日根（2017）の研究では，合成音声は肉声に比べて温かみや親しみやすさが低い印象を与えることが示されている。また，Engelbregt et al.（2022）の研究では，ASMR動画の視聴により心拍数が低下し，リラックス効果が得られることが示されている。しかし，合成音声が身体に与える影響についての研究は依然として十分ではなく，今後，多くの場面で合成音声を使用することが予想される中，その影響についての知見を深める必要がある。

心理学領域における音声の研究において，ポリヴェーガル理論は合成音声の心身への影響を理解する上で参考となる。この理論では，音声の特徴が自律神経系に及ぼす影響が議論されており，特に腹側迷走神経が社会的な交流や感情調整に重要な役割を果たしているとされる。この理論に基づき，本研究では，合成音声の特性が心拍変動（HRV）に及ぼす影響を検討し，その応用可能性について新たな知見を提供することを目的とする。

**予備実験**

予備実験では，実験で使用する合成音声の刺激を選定するため，一般的に好感度の高い音声と低い音声の特徴を明らかにすることを目的に実施した。17名の大学生および大学院生(男性9名，女性6名，その他1名; 平均年齢23歳(*SD*=8.56))を対象に，橋本・古屋（2019）の音声状態印象に関する尺度(Table1) を用いて，発話速度や声の高さが音声の印象にどのような影響を与えるかを評価した。この尺度は力動性，悠長さ，好感度の3因子構造から成り，音声合成ソフトウェア「VOICEVOX」と「COEIROINK」を用いて作成された8種類の音声を対象とした(Table2)。





参加者はそれぞれの音声に対してその印象を評価し，好感度因子の得点を基に使用する音声キャラクターを4種類に絞り込んだ。最終的に，サンプル1を「女声高好感度刺激」，サンプル4を「女声低好感度刺激」，サンプル5を「男声高好感度刺激」，サンプル7を「男声低好感度刺激」とした。

**本実験**

本実験では，予備実験の結果に基づき，音声の性別および好感度を操作した合成音声を用いて，自律神経活動に与える影響を検討した。対象者は20名の大学生および大学院生で，平均年齢は20.14歳(*SD*＝1.53)であった。実験ではポラール社の心拍センサー「Polar H10」とアプリケーション「Elite HRV」を用いて心拍数および心拍変動（HRV）を測定し，心拍変動解析にはKubios HRVを用いた。

刺激音声は，予備調査で得られた高好感度および低好感度の4種類の音声を使用し，再生時間を4分前後に調整した。また，実験中の生理的反応を分析するために，心拍数（HR），および心拍変動の指標であるRMSSD，高周波(HF)，LF/HF比などの生理指標を計測した。

また主観的な心理指標を測定するため，小川ら（2000）の一般感情尺度を用いて，音声再生中の感情状態を評価した。



**結果**

生理指標において，すべての刺激で心拍数（HR）は期間が進むにつれて有意に減少し，RR間隔，SDNN，RMSSD，pNN50，HFは増加していた。また，女声高好感度の刺激は，後安静期において最も強く副交感神経活動有意な変化を示した。また，HFにおいては有意な交互作用が見られ，HFは期間が進むにつれて増加し，特に女性の声において期間の影響が強く，後安静期で顕著な増加が確認されたことが明らかになった。　LFとLF/HFの結果では，低好感度の音声条件でLFが高く，LF/HF比も増加しており，交感神経活動が優位であることが示唆された。



Figure1.各刺激におけるHRの変化



Figure2.各刺激におけるRMSSDの変化



Figure3.各刺激におけるHFの変化

心理指標においては，ポジティブ感情（PA）と安静状態感情（CA）は高好感度の音声で高い得点を示した。ネガティブ感情（NA）は低好感度の音声で高得点を示し，不快感を与える傾向が見られた。また，力動性や悠長さの評価では，音声の性別や好感度が得点に影響を与え，特に低音の刺激において力強さが評価された。

**考察**

**【AI合成音声が主観的感情に与える影響】**

本実験の結果，好感度が高い音声はポジティブ感情（PA）や安静感（CA）の得点を高め，リラックス効果を強く引き出すことが示された。特に女声高好感度刺激は他の条件に比べ，ポジティブ感情やリラックス感の指標で有意な効果を示した。また，男声刺激においては女声刺激に比べて効果がやや弱く，声の高さが感情反応に影響を与える要因である可能性が示唆された。この結果は，特定の音響刺激が腹側迷走神経を活性化し，自律神経のバランスを整えるというListening Project Protocol (LPP) やSafe and Sound Protocol (SSP) の研究成果に一致した（Porges et al., 2013; Porges, 2022）。

**【測定中の大まかな自律神経活動の変化】**

実験進行に伴いHRは下降し，RR間隔やSDNN，RMSSD，pNN50，HFは上昇した。これらの変化は副交感神経活動が活性化したことを示唆した。一方，LF/HF比に有意な変化はなく，刺激は交感神経系に影響を与えないことが示唆された。

**【好感度の効果】**

音声の好感度が生理指標に与える影響を検討した結果，課題期において低好感度音声の方が高好感度音声よりもHRが低下していた。この現象は，低好感度音声が注意を引きやすいことに起因する可能性が考えられた。Lacey＆Lacey(1974)によれば，注意の集中により心拍が一時的に減少し，副交感神経活動が促進されることが示唆されている。また，Berntson et al. (1996)の研究では，注意が迷走神経の活性化を引き起こし，これが刺激への注意と関連することが示されている。今回の低好感度音声は聞き慣れないため，注意が集中し，HRが低下したと考えられた。この結果は，低好感度音声がNAを上昇させる性質や，悠長さ得点の低さとも一致していた。

**【性別の効果】**

性別の効果に関しては，女声高好感度刺激が副交感神経活動を大きく活性化させたことが示唆された。具体的には，HRは後安静期に最も低く，RR間隔，RMSSD，pNN50，HFは後安静期で最も高い値を示した。特にHFは女声高好感度刺激においてのみ有意な上昇が認められ，男性参加者に対するリラックス効果が示された。

この結果の背景には社会的交互作用の影響があると考えられる。Hildebrandt et al.(2016)は，仮想空間での脅威体験中，HRVが高い参加者ほど自律神経が安定し，冷静さを保てることを示した。また，Lehrer(2018)やThayerら(2018)は，HRVが感情制御や幸福感と関係し，適応や回復の指標となることを報告している。

ポリヴェーガル理論によれば，音声を含む社会的交互作用を司る迷走神経（腹側迷走神経）は心臓の制御にも関与する。本研究では性別の効果がHFでのみ確認されたが，これは用いた音声刺激が抑揚の調整を行わなかったためと推察される。リラクゼーション効果を最大化するためには，合成音声に適切な抑揚を加えることが有効である可能性が示唆される。

**【総合考察】**

本研究の生理指標と心理指標を総合的に考察すると，合成音声が生理的リラクゼーションに一定の影響を与える可能性が示唆された。特に課題期から後安静期にかけて，女声高好感度刺激が副交感神経活動を促進し，リラックス効果を高める傾向が見られた。しかし，生理指標においては課題期で音声の性別や好感度による有意な効果が確認されず，心理指標との一致性が低かった。これは以下の要因によるものと考えられた。

まず，音声再生時間とテンポが統一されていたことが，性別や印象の効果を抑制した可能性がある。福本ら(2004)の研究では，音楽のテンポが心拍と同期する現象が確認されている。しかし，本実験では全ての音声が一定テンポで再生されたため，テンポの変化による影響が排除され，性別や好感度の効果が明確に現れなかったと考えられた。

次に，合成音声の情動喚起が低かったことが要因として挙げられる。Gong C(2023)の研究では，AI合成音声は人間の声と比較して脳の認知活動を引き起こす力が弱いことが示されている。特にニュース放送では，人間の声がβ波の活動を強化し，聴覚情報の処理や記憶保持を促進する一方で，合成音声はその効果が限定的であることが確認された。このような合成音声の特性が，性別や好感度の効果を弱めた可能性が考えられた。

さらに，合成音声は一般的に広く利用されることを目的として作成されているため，比較的良質な声が使用される傾向にある。本実験では，この範囲内で好感度の高低を設定したため，実験参加者に対して強い不快感や否定的な印象を与える音声が含まれにくかった。その結果，有意差が得られなかった可能性が考えられた。

**【今後の展望】**

本研究の結果から，女性の合成音声は男性音声よりも高いリラックス効果を示し，日常的に使用される音声には適していると考えられた。しかし，性別や好感度の効果は一部の分析手法でのみ認められたことから，感情的な働きかけを行う場合には，抑揚を高めるように音声を調整する必要があるかもしれない。一方で，抑揚が少なく感情が伝わりにくい音声が適する場面もある。例えば，公共の注意喚起では情動を刺激しないことで冷静な対応を促す効果が期待でき，抑揚の少ない合成音声の使用が適切である可能性が示唆された。

今後の研究では，実験参加者の置かれた状況を多様に変化させ，合成音声の効果をさらに検討する必要がある。また，本研究では参加者の性別や人格特性の影響は検討されていない。本実験の参加者は男性に限定されており，好感度得点の標準偏差に注目すると，女声高好感度刺激では回答のばらつきが小さい一方で，他の3刺激ではばらつきが大きい可能性が示された。これは音声に対する個人の好みが多様であることを示している。今後の研究では，参加者の性別や人格特性を含む多様な要因を考慮し，状況や相手に応じた適切な合成音声の使用法を明らかにすることが求められる。

**主な引用文献**

Berntson, G. G., Cacioppo, J. T., & Fieldstone, A. (1996). Illusions, arithmetic, and the bidirectional modulation of vagal control of the heart. *Biological Psychology*, 44(1), 1-17.

Engelbregt, H. J., Brinkman, K., van Geest, C. C. E., Irrmischer, M., & Deijen, J. B. (2022). The effects of autonomous sensory meridian response (ASMR) on mood, attention, heart rate, skin conductance, and EEG in healthy young adults. *Experimental Brain Research*, 240(7), 1727-1742.

福本 誠・楠芳 之・長島 知正. (2004). 音楽のテンポと心拍の同期現象 Synchrogram による同期状態の検出とリラクセーション効果への影響. 感性工学研究論文集, 4(2), 17-24.

Gong, C. (2023). AI voices reduce cognitive activity? A psychophysiological study of the media effect of AI and human newscasts in Chinese journalism. *Frontiers in Psychology*, 14, 1243078.

橋本 和奈実・古屋 健(2019)．発話速度と声の高さが特性推論に及ぼす影響．応用心理学研究，45(1), 15-25．

Hildebrandt, L. K., McCall, C., Engen, H. G., & Singer, T. (2016). Cognitive flexibility, heart rate variability, and resilience predict fine‐grained regulation of arousal during prolonged threat. *Psychophysiology*, 53(6), 880-890.

Lacey. B. C., & Lacey. J. 1 (1974). Studies of heart rate and other bodily processes in sensorimotor behavior. In P. A. Obrist, A. H. Black, J. Brenner & L. V. DiCara (Eds), *Cardiovascular Psychophysiology*. Chicago: Aldine pp.538-564.

Lehrer, P. M. (2018). Heart rate variability biofeedback and other psychophysiological procedures as important elements in psychotherapy. *International Journal of Psychophysiology*, 131, 89-95.

Mather, M., & Thayer, J. F. (2018). How heart rate variability affects emotion regulation brain networks. Current Opinion in Behavioral Sciences, 19, 98-104.

宮本 敬子・日根 恭子(2017)．合成音声と人声による発話の長さが音声の印象評価へ与える影響．*日本認知心理学会発表論文集*, P2-11.

NHK(2021). 視覚障害者向けテレビの研究. https://www.nhk.or.jp/strl/open2021/tenji/9-1/index.html. (2024年1月17日)

小川 時洋・門地 里絵・菊谷 麻美・鈴木 直人(2000)．一般感情尺度の作成．*心理学研究*, 71(3), 241-246.

Porges, S. W., Macellaio, M., Stanfill, S. D., McCue, K., Lewis, G. F., Harden, E. R., & Handelman, M. (2013). Respiratory sinus arrhythmia and auditory processing in autism: Modifiable deficits of an integrated social engagement system? International Journal of Psychophysiology, 88(3), 261-270.

Porges, S. W. (2022). Polyvagal theory: A science of safety. *Frontiers in Integrative Neuroscience*, 16, 871227.

藪 謙一郎(2018)．障害者が必要としている音声技術．*日本音響学会誌*, 74, 136-143.

山岸 順一(2015)．音声の障がい者のための最先端音声合成技術．*情報管理*, 57(12), 882-889.