

スピーカーの外見は 新奇音によるタスク遂行への妨害効果を変えるか

Do novel sounds from a speaker always distract a visual task?: A study on the influence of speaker appearance

土橋 一斗¹ 福田 聡子¹ 中島 亮一² 大澤 正彦^{1*}
Kazuto Dobashi¹, Satoko Fukuda¹, Ryoichi Nakashima², Masahiko Osawa¹

¹ 日本大学
¹ Nihon University
² 京都大学
² Kyoto University

Abstract: ヘッドホンより定常的に音出力される中での単純なタスクにおいて、新奇音によりタスク遂行は妨害される。また、音の印象は、それが発せられた出力元の外見によって異なる。これらの知見を踏まえ、本研究では、出力元（スピーカー）の外見によって、新奇音によるタスク遂行への妨害効果が異なるかを調査する。具体的には、音とともに画面に表示される数字の偶奇判断タスクにおいて、通常のスピーカーとロボット型のスピーカーを用いた場合の成績を比較した。結果、スピーカーの外見がロボットの場合に新奇音による妨害効果が引き起こされなかったことから、スピーカーの外見がロボットの際はタスクへの音による妨害効果を緩和させると考えられる。

1 はじめに

本研究では、音の出力元の外見の違いによって、聴覚刺激（音）が視覚タスクのパフォーマンスに与える影響が異なるかを検討する（図1）。

日常的に、オフィスや公共の場では、他者の会話や電話などの自分とは関係のない音が存在している状況でタスクを行うことも多い。そのような環境音は、自分が現在行っているタスクのパフォーマンスに影響を与える。例えば、聴覚刺激と視覚タスクの関係について、視覚刺激とほぼ同時に提示される聴覚刺激が視覚タスクのパフォーマンスを促進させることが報告されている [1]。一方、視覚タスクとは無関係な聴覚刺激が視覚タスク中に発生する場合、視覚タスク遂行者の注意がその聴覚刺激に向くことで、一時的に視覚タスクのパフォーマンスが低下することを報告している研究も存在する [2]。本研究では、特に聴覚刺激による視覚タスクパフォーマンスへの妨害効果に注目する。妨害効果の例として、SanMiguel らは、定常的に音出力されている環境で、モニターに提示される視覚刺激が人間の顔か、その顔をスクランブルにしたものかを判別するタスクを行う際、稀に新奇音出力されること



図 1: 本研究の概要

で、タスクへの妨害効果が引き起こされ、正答時の反応時間が長くなると報告している [3]。

新奇音によってタスクへの妨害効果が引き起こされたと報告している SanMiguel らの研究 [3] では、聴覚刺激をヘッドホンから出力させていた。しかし、日常場面を想定する場合、電話の着信音は電話から聞こえてくる等、環境音はヘッドホン以外から聞こえることも多い。では、音が出力される場所がヘッドホンではなく、被験者が視認できる場所に配置されたスピーカーであっても、新奇音による妨害効果が引き起こされるのだろうか。

HAI 領域の研究では、エージェントの外見により、出

*連絡先：日本大学文理学部
〒156-8550 東京都世田谷区桜上水 3-25-40
E-mail: osawa.masahiko@nihon-u.ac.jp

力された音の印象が変わることを報告した研究がある。小松らは、事前実験にてポジティブと評価された人工音 4 種類およびネガティブと評価された人工音 4 種類を、外見の異なる出力元から出力させた際に、音の印象は変化するのかを検討した。その結果、ロボットの外見条件では、前実験でポジティブと設定された人工音においてもネガティブと推定されてやすくなることが確認された [4]。では、異なる外見のスピーカーから出力した際、外見によってタスクに対する音の妨害効果に差異は起きるのだろうか。

これらのことを踏まえ、本研究では次のような問いを立てた。音出力される場所がヘッドホンではなく、被験者が視認できる場所に配置されたスピーカーであっても、新奇音による妨害効果が引き起こされるのか。またスピーカーの外見によって出力された音の印象が変化するならば、新奇音によるタスクの妨害効果もスピーカーの外見によって違うではないか。本研究では、タスクとは無関係な音が鳴っている状況（具体的には、定常的に音が鳴っていて、たまに異なる新奇音が鳴る状況）での視覚タスクにおいて、新奇音による妨害効果が、スピーカーの外見によって異なるのかを調査する。定常的な音（標準音）と新奇音が鳴った際の、タスクのパフォーマンスを比較することで、妨害効果を評価する。具体的には、新奇音時のパフォーマンスが標準音時よりも低ければ妨害効果と想定される。

2 背景

2.1 新奇音によるタスク遂行への妨害効果

SanMiguel ら [3] は、視覚タスクのパフォーマンスに対するタスク非関連の新奇音の影響を調査した。実験では、モニターに画像が連続して提示されるので、実験参加者はその画像が人間の顔かスクランブルした顔画像かを判別した。各試行は 1100ms で構成されており、各視覚刺激の提示後 800ms まで判別の応答が受け付けられた。実験中、視覚刺激の提示の 300ms 前にそれぞれ 80%、20% の割合で異なる音がヘッドホンから出力され、200ms 間続いた。80% および 20% の割合で出力される音をそれぞれ標準音、新奇音とした。その結果、標準音出力時のタスクの正答率と新奇音出力時のタスクの正答率に有意な差は見られなかったが、新奇音において正答時の回答の反応時間が長くなる傾向が確認された。

2.2 スピーカーの外見による音の印象

小松ら [4] は、音声の出力元の外見がユーザーの情報解釈に及ぼす影響について焦点をあてた研究を行った。

44 種類の三角波をノート PC から出力し、被験者はそれぞれの音声の印象を回答する実験が行われた。この実験結果をもとに、ポジティブと評価された人工音 4 種類およびネガティブと評価された人工音 4 種類を、外見の異なる出力元から出力させた際に音声の印象が変わるのかが検証した。その結果、ノート PC の外見条件では前実験結果と差異がほぼなかったが、ロボットの外見条件では、前実験でポジティブと設定された人工音においてもネガティブと推定されてしまいやすくなってしまうことが確認された。

3 実験方法

3.1 実験タスク

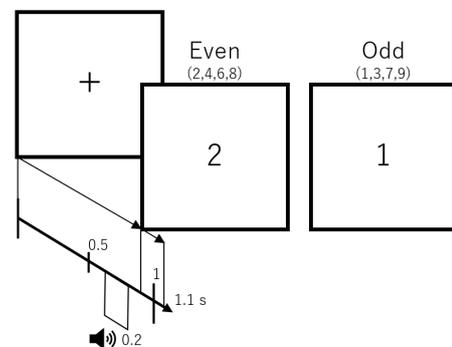


図 2: 刺激提示順序

本実験では SanMiguel らの実験 [3], Escera らの実験 [5] を参考にして、タスクを設計した。モニターに視覚刺激として 1 桁の数字 1,2,3,4,6,7,8,9 のいずれかを提示し、被験者はその数字が偶数か奇数かを判断した。1 試行の刺激提示順序を図 2 に示す。各試行は 1100ms で構成されており、視覚刺激後 800ms までの判別の応答を受け付けた。応答はテンキーの 4 キーを人差し指、6 キーを中指で押すことで受け付けた。半数の被験者は奇数の際に 4 キーを、残り半数の被験者は奇数の際に 6 キーを押すように教示した。実験中、視覚刺激の出力 300ms 前にビープ音がスピーカーから 200ms 間出力された。出力されるビープ音について、80% の割合で 600Hz、20% の割合で 700Hz、それぞれ約 75db のビープ音出力された。それぞれ標準音、新奇音とよぶ。実験は 3 つのブロックあり、各ブロック 80 試行を行った。各ブロック初めの 4 試行では標準音を出力した。ブロック間には最低でも 1 分の休憩を挟み、その後任意のタイミングで次のブロックを開始することができるようにした。本実験の前にはビープ音の出力がない練習の試行が 10 回行われた。被験者は偶数か奇数かの判断の正答率が 70% を超えるまで、練習を繰り返した。

3.2 実験環境

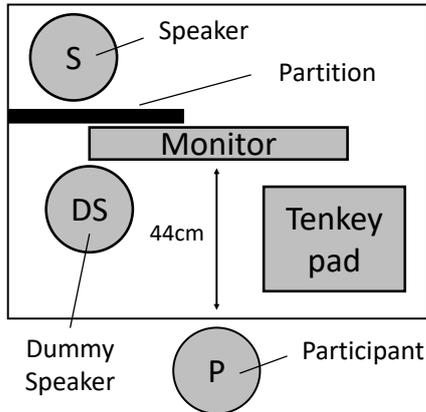


図 3: 机上レイアウト

実験中に出力されるピープ音以外の外部の環境音の実験結果に与える影響を小さくするために、第一著者の所属する大学施設内の比較的静かな部屋を使用して実験を行った。モニターやスピーカーの位置は固定され、被験者による相違がないように配慮した(図3)。スピーカーは被験者が手を伸ばして触れることができる程度の距離に配置した。タスク試行中は部屋の電気を消し、スピーカーの形状が視認できる程度の明るさにした。また、スピーカーの種類による音質の違いが実験結果に影響を与えないように、スピーカーの外見条件によって異なるダミーのスピーカーを被験者が視認できる場所に配置し、実際に音出力されるスピーカーをダミースピーカーの後方に配置した。なお、実際に音出力されるスピーカーはパーティションによって被験者が視認できないように配置した。

3.3 ダミースピーカーの外見

ダミースピーカーとして、一般的な丸型スピーカー(Normal 条件)とロボット型スピーカー(Robot 条件)を使用した(図4)。Normal 条件にて使用するスピーカーは、Creative 製の Pebble V3 を使用した。Robot 条件にて使用するスピーカーの外見は、川崎らが開発したロボットを使用した[6]。このロボットは全体的に丸みを帯びている形状となっており、2 頭身の外見によって親しみやすいデザインとなっている。

被験者への実験の説明時に、ピープ音が各条件のスピーカーから出力されると教示した。これにより被験者に音の出力元を認識させた。



(a) Normal 条件

(b) Robot 条件

図 4: ダミースピーカーの外見

3.4 被験者

本実験には正常な視力または矯正視力とスピーカーからの音を聞くのに十分な聴力を持ち、かつ利き手が右である 18 ~ 29 歳までの男性 19 名と女性 19 名、計 38 名が参加した。被験者はスピーカーの外見の各条件にランダムで割り当てられた。本試行のタスクの正答率が 70% 以下だった男性 1 名、また実験環境の不備で女性 1 名のデータが無効となった。その結果、Normal 条件に 18 名(うち女性 10 名)、Robot 条件に 18 名(うち女性 8 名)のデータを分析対象とした。

3.5 分析方法

本研究では、タスクのパフォーマンスを測るため、Inverse Efficiency Score (IES) の分析を行う。IES は、タスクの正答率と正答時の反応時間を組み合わせたパフォーマンスの指標である[7]。IES はタスクの反応時間を正答率で割ることで数値が算出される。これによって、条件間で速さと正確さのトレードオフ(回答速度の速さを意識すると雑にタスクをこなすようになり、逆に丁寧にタスクを行うことを意識すると、回答速度は遅くなる場合がある)が生じていたとしても、それを考慮して条件間のパフォーマンスを比較できる。

$$IES = \frac{\text{正答時の反応時間}}{\text{正答率}}$$

スピーカーの外見について Normal 条件と Robot 条件を被験者間要因、タスクとは無関係に出力されるピープ音の種類(標準音、新奇音)を被験者内要因とする 2×2 の 2 要因混合計画の分散分析を行う。また、正答率、正答時の反応時間についても同時に分析する。なお、先行研究[3]にならい、新奇音が出力された試行の直後の試行は標準音を出力する試行とし、分析から除外した。

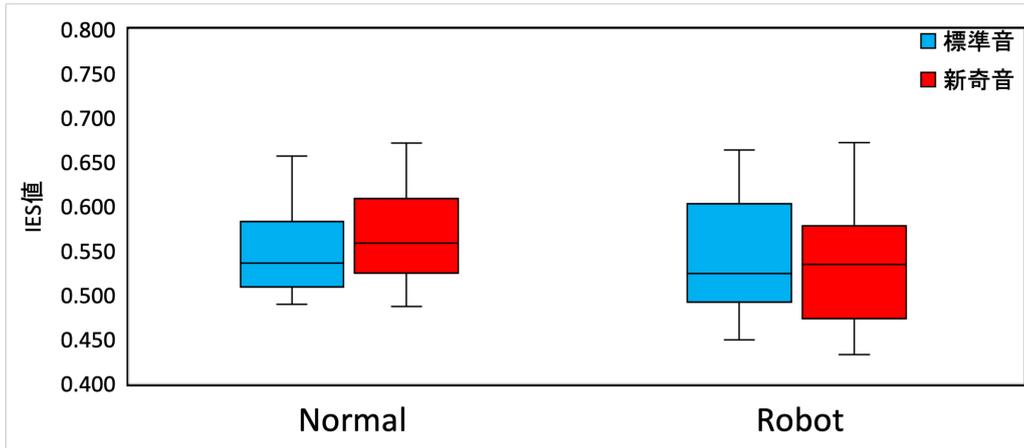


図 5: 標準音時と新奇音時の IES の比較. ひげの上端から下端まで 4 つの区間に区切られており, それぞれの区間が全体の 1/4 のデータを収容している. 箱中の線は中央値.

表 1: 各条件によるタスクの正答率, 正答時の反応時間

	Normal		Robot	
	標準音	新奇音	標準音	新奇音
正答率	90.4%	88.8%	90.7%	92.2%
正答時の反応時間	495ms	500ms	491ms	490ms

4 実験結果

各条件によるタスクの IES を図 5 に示す. 分散分析の結果, スピーカーの外見とビープ音の種類 of 交互作用が認められた ($F(1, 34) = 5.393, p = 0.026$). 単純主効果検定の結果, Normal 条件で, 新奇音時の IES が標準音時の IES よりも増加した (ただし, 有意傾向であった. ($F(1, 34) = 3.814, p = 0.067$)). また, Robot 条件では, 音の種類における単純主効果が有意ではなかった ($F(1, 34) = 1.777, p = 0.200$). よって, ビープ音出力されるスピーカーの外見が Robot 条件の場合, 新奇音によるタスクへの, エラー率で重み付けをした反応時間の増加させる妨害効果が, Normal 条件に比べて緩和されることが示唆される. また, スピーカーの外見の条件間での標準音時の IES は差は見られなかった ($F(1, 34) = 0.070, p = 0.793$). よって, タスクパフォーマンス自体が, スピーカーの外見の条件間で大きく違っているわけではなかった.

本研究のメインの目的ではないが, IES の違いがどのように生じたかについても検討する. 各条件によるタスクの正答率, 正答時の反応時間を表 1 に示す. 正答率について, 分散分析の結果スピーカーの外見とビープ音の種類で交互作用が認められた ($F(1, 34) = 5.258, p = 0.028$). 単純主効果検定の結果, Normal 条件で, 新奇音時の正答率が標準音時の正答率よりも低下した (ただし, 有意傾向であった. ($F(1, 34) = 3.137, p = 0.094$)). また, Robot 条件では, 音の種類における単純主効果が

有意ではなかった ($F(1, 34) = 2.234, p = 0.153$). よって, 正答率を見ても, ビープ音出力されるスピーカーの外見が Robot 条件の場合, Normal 条件に比べて新奇音による妨害効果が緩和されることが示唆される. また, スピーカーの外見の条件間での標準音の正答率は差は見られなかった ($F(1, 34) = 0.041, p = 0.841$). よって, タスクパフォーマンス自体が, スピーカーの外見の条件間 (つまり 2 つのグループ間) で大きく違っているわけではなかった.

正答時の反応時間では, スピーカーの外見とビープ音の種類で交互作用は認められなかった ($F(1, 34) = 0.697, p = 0.410$). また, ロボットの外見条件の主効果 ($F(1, 34) = 0.161, p = 0.691$), ビープ音の種類的主効果 ($F(1, 34) = 0.268, p = 0.608$) ともに有意ではなかった.

5 考察

今回の実験ではタスクの正答率と正答時の反応時間を組み合わせたパフォーマンスの指標である IES でも分析したが, スピーカーの外見とビープ音の種類 of 要因間の交互作用が認められ, スピーカーの外見が Normal 条件において, ビープ音の種類における単純主効果が有意傾向となり, スピーカーの外見が Robot 条件において, ビープ音の種類における単純主効果は有意ではなかった. このことから, 一般的なスピーカーから新

奇音が出力された際は、先行研究でのヘッドホンから出力される際 [3] と同じく、新奇音によるタスクへの妨害効果が引き起こされるが、スピーカーの外見がロボットであることにより、新奇音によるタスクへの妨害効果が緩和されるのではないかと考える。

5.1 先行研究との比較

先行研究である SanMiguel らの実験 [3] では、タスクの反応時間の増加させる妨害効果が報告されていた。しかし、反応時間に有意差は見られなかった。では、なぜ今回は反応時間において有意差が見られなかったのだろうか。

先行研究での反応時間と今回の実験での反応時間を比較する。先行研究では、標準音出力時の反応時間は 405～420ms 程、新奇音出力時の反応時間は 415～430ms 程に対して、今回の実験では、標準音出力時の反応時間は、スピーカーの外見が Normal 条件で平均 495ms、Robot 条件で平均 491ms、新奇音出力時の反応時間は、スピーカーの外見が Normal 条件で平均 500ms、Robot 条件で平均 490ms であった。よって、スピーカーの外見条件にかかわらず、反応時間が先行研究よりも全体的に増加していることがわかった。このことから、全体的な回答時間が先行研究よりも遅くなったことで、新奇音時の反応時間がタスクの回答時間の制約である 800ms に収まりきらなくなったのかもしれない。

5.2 スピーカーの外見によるタスクへの妨害効果の緩和

今回の実験結果では、スピーカーの外見が Normal 条件の際は、IES、正答率において、新奇音によるタスクパフォーマンスの妨害効果が見られた。一方、スピーカーの外見が Robot 条件の際は、IES、正答率、反応時間に違いが見られなかった。つまり、Robot 条件では新奇音による妨害効果が見られなかった。

Robot 条件にて使用したロボットの外見について、福田らは、ロボットの外見の印象評価のアンケート調査を行っている [8]。そのなかで、「電子音を出せそう」といった質問において、今回使用したロボットの外見にて高い評定値となっていた。また、「頭を撫でたい」、「抱きしめたい」といったアトラクティブ要因について問う質問においても、高い評定値となっていた。このことから、本研究のスピーカーの外見条件にて、Robot 条件の際に使用したロボットの外見は、電子音が出力されても被験者に違和感をもたせることなく、かつ外見においてポジティブな印象を感じると考えられる。一方、小松らはスピーカーの外見がロボットであることにより、出力される音がネガティブな印象を持たれや

すくなると報告している [4] が、その実験で使用されていたロボットは、犬型の外見、機械的な印象を与える外見のロボットであった。これについて小松らは、ビープ音が、実験で使用されたロボットから表出されることが被験者から予期されず、被験者から不自然だと認識されたため、ネガティブな印象を持たれてしまいやすくなると考察している。これらのことから、「電子音を出せそう」という印象を与えるロボットの外見であれば、被験者は出力音に対するネガティブな印象を持たなかったのかもしれない。その結果、新奇音による妨害効果が緩和されたのではないかと考えられる。

今後はスピーカーの外見条件が Robot 条件時の、被験者個人の成績を細かく調査していくことで、特徴が無いかを探索していく。また、ロボットの外見の種類についても検討を行う必要がある。今回使用した電子音を出すことに対して違和感をもたせず、親しみやすくポジティブな印象を与えるようなロボットの外見ではなく、電子音を出力させることで、音の印象がネガティブと判断されやすくなってしまいう犬型ロボットの外見や機械的な印象を与えるロボットの外見 [4] など、様々なロボットの外見で新奇音によるタスク遂行への妨害効果がどうなるのかを検証する。さらに、本研究で使用したロボットは腕があり、ユーザーと手をつなぐなどの「接触インタラクション」も容易に行えるという特徴がある。そのため、ロボットとの接触により、ロボットの外見に対して感じた印象が変化 [9] することで、妨害効果がより緩和される可能性も考えられる。この可能性についても今後検証を進める。

6 おわりに

本研究は、音とともに画面に表示される数字の偶奇判断タスクにおいて、通常のスピーカーとロボット型のスピーカーを用いた場合の成績を比較することで、スピーカーの外見によって、稀に出力される新奇音によるタスク遂行への妨害効果が異なるかを探索した。結果として、通常のスピーカーでは、新奇音によるタスク遂行への妨害効果が引き起こされること確認できた。また、ロボット型のスピーカーを用いた場合では、新奇音によるタスク遂行への妨害効果が緩和され、標準音出力時と同等のタスクパフォーマンス結果となった。よって、電子音を出すのが不自然だとは感じさせず、親しみやすくポジティブな印象を与えるような外見のロボットは、新奇音によるタスクへの妨害をしづらくなると考えられる。

参考文献

- [1] Steven A. Hackley, Fernando Valle-Inclán, “Accessory stimulus effects on response selection: Does arousal speed decision making?”, *Journal of Cognitive Neuroscience*, 11, 3 pp.321-329, 1999.
- [2] Juduth Dominguez-Borrás, Manuel Garcia-Garcia, Carles Escera, “Negative emotional context enhances auditory novelty processing”, *Neuroreport*, 19, 4, pp.503-507, 2008.
- [3] Iria SanMiguel, David Linden, Carles Escera, “Attention capture by novel sounds: Distraction versus facilitation”, *European Journal of Cognitive Psychology*, 22, pp.481-515, 2010.
- [4] 小松 孝徳, 山田 誠二, “エージェントの外見の違いがユーザの態度解釈に与える影響－外見の異なるエージェントからの同一人工音の提示実験”, *知能と情報 (日本知能情報ファジー学会誌)*, 20, 4, pp.500-512, 2008.
- [5] Carles Escera, Maria-Jose Corral, Elena Yago, “An electrophysiological and behavioral investigation of involuntary attention towards auditory frequency, duration and intensity changes”, *Cognitive Brain Research*, 14, pp.325-332, 2002.
- [6] 川崎邦将, 大澤正彦, 今井倫太, 長田茂美, “認知的制約付き擬人化キャラクターに着目したコミュニケーションロボットの設計と開発”, 第6回汎用人工知能研究会, 2017.
- [7] Raymond Bruyer, Marc Brysbaert, “Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?”, *Psychologica Belgica*, 51, pp.5-13, 2011.
- [8] 福田聡子, 澤田志織, 川崎邦将, 奥岡耕平, 大澤正彦, 長田茂美, 今井倫太, “適応ギャップ理論を拡張したインタラクションデザインの提案”, *HAI シンポジウム*, 2019.
- [9] 野々村佳祐, 寺田和憲, 山田誠二, 伊藤昭, “ロボットとの感情的インタラクションは外見の悪影響を緩和する”, *情報処理学会研究報告ヒューマンコンピュータインタラクション (HCI) 2015-HCI-161*, 6, pp.1-7, 2015.