

こえ射的：声で遊ぶARトレーニングの検討

辻野 圭亮^{1,a)} 丸本 啓太^{1,b)} 井上 創太^{1,c)} 足立 大生^{1,d)} 小笠原 一磨¹ 馬場 保仁¹
石崎 仁湖¹ 大井 翔^{1,e)}

概要：本研究では、音声インタラクション技術とAR（拡張現実）を組み合わせた新しい発声トレーニングシステム「こえ射的」を提案する。本システムは、ユーザーの声を仮想弾として活用し、AR空間内で景品を狙う射的ゲームの形式をとる。特に、高齢者の健康維持や吃音・社交不安症を抱える人々の発話支援を目的とし、発声行為を「義務」や「治療」ではなく「遊び」として促進することを目指している。実験では、システムの楽しさやリピート性が高く評価された一方で、リアリティの不足や音声認識のタイムラグが課題として挙げられた。

1. はじめに

音声インタラクション技術は、エンターテインメントだけでなく、リハビリや教育現場でも注目されている。特に、発声トレーニングへの応用は、多様な対象者に有益であると考えられる。

高齢者が声を出す機会が減少することは、健康面における課題と深く結びついている [1]。具体的には、口腔機能の低下、認知症の進行、社会的孤立感の増大などが挙げられる。音声インタラクションを取り入れたシステムは、高齢者が声を出す機会を提供する。これにより、健康維持や認知機能の向上に寄与する可能性がある。特に、楽しいゲームや対話的な活動を通じて声を出すことで、自然に積極的な発声習慣を促すことができる。

吃音や社交不安症を抱える患者にとって、発語に対する心理的なハードルは大きい。これらの患者は、他者とのコミュニケーションに苦手意識を持つことが多く、発声練習やトレーニングを避ける傾向がある。音声インタラクションを活用したシステムは、楽しく声を出す練習を行う場を提供することで、発語に対する不安を軽減し、徐々に発話の自信を高めることができる。このようなシステムは、音声トレーニングの敷居を下げ、自己効力感を高める手段として有効である。

以上の対象者に共通する課題として、発声やコミュニケーションの練習が「義務」や「治療」として認識される

場合、モチベーションの維持が困難である点が挙げられる。音声インタラクションを用いたシステムは、楽しさを通じてこれらの課題を克服し、トレーニングの定着を促進する可能性を秘めている。特に、ゲームやAR技術を組み合わせることで実際の行動とその結果を結びつけた直感的なフィードバックを提供できる点は、新しい発声支援ツールとしての可能性を広げるものである。そこでAR空間を活用して発話を視覚化し、ゲーム要素を取り入れた「こえ射的」という体験型システムを提案する。このシステムは、図1に示すように音声が発射の弾として扱われ、発話の成果が視覚的・物理的反応としてフィードバックされることで、ユーザーに楽しさと達成感を提供する。

本システムは、発声やコミュニケーションに苦手意識を持つ人々に対し、発声を促す新たなアプローチを提供するものである。

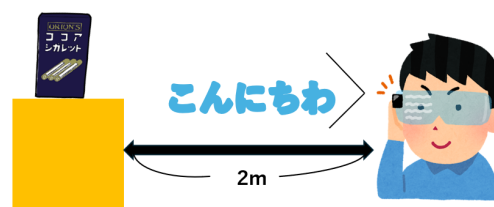


図1 こえ射的：システム概要

2. 関連研究

2.1 高齢者の声帯機能と発声課題

平野 [2] によると図2のように咽喉頭の加齢変化は50歳前後から始まり、最初は粘膜の繊毛機能低下や粘液産生低

¹ 大阪工業大学情報科学部

a) e1n20054@oit.ac.jp

b) keita.marumoto@miz-lab.net

c) e1b21010@oit.ac.jp

d) e1q21015@oit.ac.jp

e) sho.ooi@outlook.jp

下により咽喉頭の乾燥や痰がらみが起こる。声帯萎縮が起こると嗄声が発生するが、その後咽喉頭全体の筋力低下、知覚低下が起こると嚥下障害を来すようになり、高齢者の嚥下性肺炎の原因となりより状況は深刻化する。

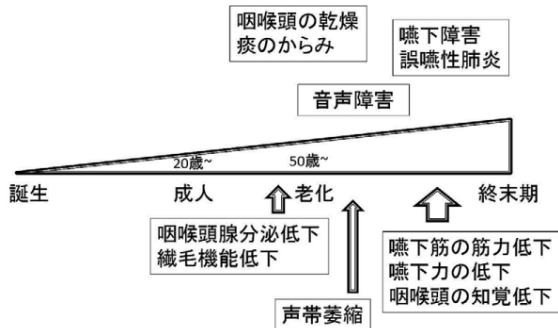


図2 年齢による声帯の変化 [2]

2.2 吃音と社交不安症における発話の障壁

吃音や社交不安症を抱える人々にとって、発話は心理的ハードルが高く、コミュニケーションへの苦手意識が強い。そのため発声練習を避ける傾向があり、症状の改善が進みにくい状況が続く。吃音症は成長とともに表面上の症状が軽減されることがあるものの、発話場面を回避することが積み重なり、思春期や青年期には社交不安障害（SAD）を合併することが少なくない [3]。吃音者は発話を回避しがちであり、この事がSADのリスクを高める要因となっている。このため、自然に声を出せる環境を提供することが吃音とSADの両方に対する効果的な支援となると考えられる。

3. システムについて

本システムは、ヘッドマウントディスプレイ（HMD）を用いた拡張現実（AR）環境において、音声インタラクションを通じて現実空間との連携を実現するシステムを構築した。提案するシステムは、ユーザーの発話内容を解析し、AR空間内に対応するオブジェクトを生成する。ユーザーは生成されたオブジェクトを発射し、現実世界の景品を狙って発射することで、物理的な景品を倒し獲得することが可能である。

3.1 発話内容のオブジェクト生成

ユーザーの声をオブジェクトに変換する手法としてまず図3に示すように音声を録音し、音声データを音声認識モデル Whisper に送信してテキスト変換を行う。Whisper が生成したテキストを基に、図4に示すようにAR空間内で発話内容のオブジェクトを生成する仕組みとなっている。



図3 発話のテキスト変換プロセス

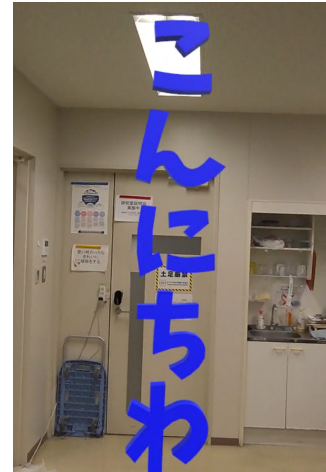


図4 音声認識によるオブジェクト生成結果

3.2 ARと実空間を連動させた景品獲得システム

景品の位置に仮想的な当たり判定領域を設置。ユーザーが発射するAR内のオブジェクトがこの領域に衝突することで、衝突イベントを発生させる。衝突イベントが発生すると、Quest デバイスから Arduino に通信が送信される。Arduino はこの信号を受け取り、図5に示すような仕掛けを作動させる制御を行う。Arduino に接続されたモーターが作動し、景品に設置された仕掛けが動くことで、お菓子が倒れる設計とした。これにより、仮想空間内のアクションが現実空間の物理的な結果に結びつく。



図5 景品動作を制御する Arduino の接続例

4. システムの評価実験

提案したシステムのユーザー体験評価を目的として、ア

ンケートを大学生3人に実施した。図6に示すように実験参加者にシステムを体験してもらい、その後アンケートを通じて主観の評価や改善点に関する意見を収集した。収集するアンケート内容は以下のとおりである。

- 緑日で射的をしたことがあるか
- 実際の射的にしているか 1: 全く似ていない~5: とても似ている
- リピータ性 1: 遊びたくない~5: もっと遊びたい
- 景品の倒れ方の自然さ 1: 自然でない~5: 自然である
- 景品の倒れ方のユニークさ 1: ユニークでない~5: ユニークである
- 音声の認識精度について 1: とても悪い~5: とてもよい
- 音声の認識レスポンス 1: とても悪い~5: とてもよい
- 楽しさについて 1: とても悪い~5: とてもよい
- どの景品の倒れ方が一番良かったか
- 上記の理由
- よかった点
- 悪かった点



図 6 実験の様子

5. 結果と考察

参加者の評価結果を表1に示す。表1から、本システムの射的としてのリアリティに対する評価は低く、実際の射的体験に比べて没入感が不足していることが示唆された。この要因として、景品の倒れ方が現実感に欠けていた可能性が考えられる。

リピータ性に関しては非常に高い評価を得ており、本システムが参加者に対して強い興味を引き出したことが確認された。特に、「声を弾として使用する」という新しい体験が参加者に楽しさを提供したと考えられる。この結果は、発声行為が「義務」や「治療」としてではなく、「遊び」として認識されることにより、モチベーションを喚起できる可能性を示している。ただし、繰り返し遊びたいと感じる要素をさらに高めるためには、景品の種類や倒れ方のバリエーションを増やし、より多様なインタラクションを提供する必要がある。

景品の倒れ方に関しては、自然さよりもユニークさが評

価された。自由記述では、図7の仕組みの景品落下音や図8のような特定の景品が勢いよく飛ぶ仕組みが満足感を高める要因として挙げられた。一方で、「重力で落ちたように見え自分で倒した感覚がない」といった課題も指摘された。景品の倒れ方には更なる改善が必要となる。



図 7 Arduino による景品制御 (1 番)

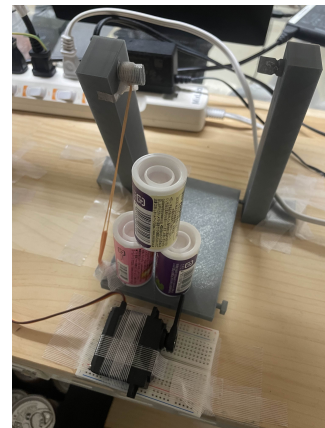


図 8 Arduino による景品制御 (4 番)

音声認識精度とレスポンスに関して、認識精度に対しては概ね良好な評価が得られたものの、一部で認識ミスやタイムラグが発生したことが報告された。また、参加者からは、コントローラーと弾の飛ぶ方向が一致しないことへの不満が指摘された。これは、音声認識結果に基づく弾の生成位置や発射方向の調整が不十分であったことが原因と考えられる。今後は、処理の高速化を図るとともに、ユーザーの発声方向と仮想弾の動きを連動させる設計が必要である。

「楽しさ」については高い評価を得た。特に、音声インタラクションを活用するという新しい試みによって、参加者は従来の射的ゲームでは得られない体験を享受したと考えられる。この結果は、音声をういた AR ゲームが発声トレーニングの枠を超え、エンターテインメント要素を強化する可能性を示している。今後は、音声インタラクション

表 1 実験後アンケートの結果

	参加者 A	参加者 B	参加者 C
緑日で射的の経験	ある	ある	ない
実際の射的ににているか	2	2	3
リピート性	4	5	5
景品の倒れ方 自然さ	3	3	2
景品の倒れ方 ユニークさ	2	3	5
音声について 認識精度	2	4	5
音声について レスpons	2	3	5
楽しさについて	4	5	4
どの景品の倒れ方が	1 番	1 番	4 番
一番良かったかまたその理由 (自由記述)	ブリッツが落ちたときの 音が大きく満足感がある	真正面向いても文字が左の方に ばかりたっていてなかなか 右に当たらなかったから	最も勢いよく飛んだ 1 番は 重量で落ちた感じがする
良かった点	自分が話した音声を射的の 玉とするのが新しく新鮮に感じた		声の認識精度が高い
悪かった点	タイムラグが生じる時があり、 意図した文字でない弾で 射撃になってしまった		1 番の倒れ方

のスムーズさを向上させるだけでなく、さらなるゲーム性の工夫を加えることで、ユーザーの興味を維持する取り組みが必要である。

症度尺度 (LSAS-J) の検討」, 耳鼻と臨床, 63 巻 2 号, p.41-46, 2017

6. まとめ

本研究では、音声インタラクションと AR 技術を用いた新しい発声トレーニングシステム「こえ射的」を提案し、ユーザー体験を評価した。本システムは、音声を仮想弾として利用し、現実の景品を倒す AR 射的ゲームとして設計されており、楽しさを通じて発声トレーニングを促進することを目的としている。

大学生 3 名を対象とした実験では、楽しさやリピート性に対して高い評価が得られた一方、射的としてのリアリティや音声認識の精度、レスポンスには課題があることが確認された。特に、景品の倒れ方のユニークさが肯定的に評価される一方で、倒れる動作のリアリティが不足しているとの指摘があった。また、音声認識のタイムラグや意図しない認識結果が発生する場面も報告された。

これらの結果から、本システムが発声行為を「義務」ではなく「遊び」として認識させる可能性が示された。これにより、発声トレーニングの心理的ハードルを下げ、楽しい体験を通じて継続的なトレーニングを促進することが期待される。一方でリアリティの向上や音声認識技術の改良がさらなるユーザー体験の向上に必要であると考えられる。

参考文献

- [1] 爲我井 壽一：高齢者の歌唱学習とその効果について，秋田大学教養基礎教育研究年報，17 巻，p.63-72，2015.
- [2] 平野 滋：専門医を悩ませる高齢疾患への対応 声帯萎縮症
- [3] 菊池 良和，梅 俊郎，澤津橋 徹，山口 優実，安達 仁美，佐藤 俊哉，中川 正法：「吃音症における社交不安障害の重