



パブリックスピーキング訓練のためのVRリアルタイムフィードバックシステムの開発と継続的使用効果の検証

高濱 悠作^{*1}藤本 雄一郎^{*1}澤邊 太志^{*1}神原 誠之^{*1}加藤 博一^{*1}

Abstract — プレゼンテーションはあらゆる場面での主要なコミュニケーション手段とされている。特に、非言語コミュニケーションの習熟度はプレゼンテーションの効果に大きく影響を与える。これまで多くのプレゼンテーション訓練システムが提案されてきたが、リアルタイムにフィードバックを行うシステムにおいてフィードバックのタイミングに焦点を当て訓練効果を検証した研究は未だない。本研究では、先行研究で提案されたタイミング推定モデルをバーチャルリアリティ（VR）環境でのプレゼンテーション訓練システムに実装し、被験者実験を通じてその有効性を評価した。結果として、提案モデルによってシステムが提供するフィードバックの印象を向上させることが示唆された。

Keywords パブリックスピーキング, リアルタイムフィードバック, タイミング推定, バーチャルリアリティ, 訓練

1 はじめに

プレゼンテーションは、ビジネス、教育、及びその他の多様な社会的文脈において、重要なコミュニケーション手段として広く認識されている。効果的なプレゼンテーションを行うには、プレゼンテーション資料の内容だけでなく、アイコンタクト、ジェスチャー、表情といった非言語的要素も重要である [1]。一般的に、反復練習がスキルの向上に寄与するとされているが、練習環境が本番に近いほど訓練効果は高くなるとされている。しかし、そのような本番に近い環境での練習は準備にかかるコストが高く容易ではない。

そこで、バーチャルリアリティ（VR）技術を活用したプレゼンテーション訓練システムが多数開発されている。これらのシステムは、3Dモデリングを通じてVR環境を構築し、低コストで現実に近い練習環境を提供することができ、さらに、発表者の行動を自動的に評価しプレゼンテーション中、もしくはプレゼンテーション後にフィードバックを提供する機能を有している [2]。特に非言語的要素に関しては、問題が発見されてからフィードバックするまでの時間が短いほど訓練効果が高いとされている [3]。しかし、プレゼンテーション中のリアルタイムフィードバックにおいて、不適切なタイミングでのフィードバックはプレゼンテーションの邪魔になり、結果として訓練効果が逆に低下するのではないかと我々は考えた。

そこで本研究では、VR環境でのプレゼンテーション訓練において、システムが自動で検出した問題に対する適切なタイミングでのフィードバックが訓練効果に与える影響について被験者実験を行い検証した。本

稿では、先行研究において提案されたフィードバックのタイミングの推定の方法について概説した後、実際の被験者実験を通じて得られた結果について示す。

2 関連研究

プレゼンテーション中のリアルタイムフィードバックによる訓練システムとして、栗原ら [4] は、音声情報処理と画像情報処理の両方を統合したプレゼンテーショントレーニングシステム「プレゼン先生」を開発した。このシステムは、マイクロフォンとWebカメラを用いて発表者の音声と身振りをリアルタイムで分析し、話速度や声の抑揚、聴衆とのアイコンタクトの度合いといった指標に基づいてフィードバックを行う。

一方、Damianら [5] は、Microsoft KinectとGoogle Glassを活用した「Logue」というシステムを開発した。このシステムは、スマートグラスを装着した発表者に対し、体の動きや話速度などのパラメータをシンプルなアイコンベースでリアルタイムにフィードバックを行う。

これらのシステムは、発表者が自身の今まで気づかなかった問題に気づき、それによってパフォーマンスが向上するという効果が報告されている。しかし、これらのシステムを継続的に利用することが訓練効果にどの程度寄与するのかについては調査されていない。

フィードバックのタイミングや頻度が訓練効果に与える影響について、Akizukiら [6] は、心肺蘇生（CPR）のトレーニングにおいて、フィードバックデバイスの使用のタイミングと頻度が訓練効果にどのように影響するかを調査した。この研究では、実際にCPRトレーニングを行なっている途中でリアルタイムでフィード

¹奈良先端科学技術大学院大学

バックが提供される場合とトレーニングセッションが終わった後にフィードバックが与えられる場合での訓練効果の影響を調査した。その結果、フィードバックのタイミングは、トレーニング中のパフォーマンスには強く影響するものの、スキルの維持にはそれほど影響を与えないことが明らかになった。

この研究では、訓練中と訓練後のタイミングの違いについて調査を行っているが、訓練中の様々なタイミングでのフィードバックによる訓練効果については調査されていない。

そこで我々は、これら問題に焦点を当て、VR技術を活用したプレゼンテーション訓練システムについて、フィードバックのタイミングが訓練効果に及ぼす影響、およびフィードバックへの印象に関する調査を行い、その結果にもとづいて新たな知見を提供する。

3 先行研究

藤本ら [7] は、VR環境下でプレゼンテーションを行う発表者へのリアルタイムフィードバックの適切なタイミングを推定するモデルを開発した。このモデルは、プレゼンテーションを行っているユーザの身体情報（体の動き、心拍数、発汗など）から二つの主要な指標「邪魔だと感じる度合い（DL: Disturbance Level）」と「訓練に役立つと感じる度合い（PBC: Perceived Behavior Changing）」を推定することができる。これらの指標からフィードバックに対するユーザの受容度を求め、ある閾値を超えた場合にフィードバックが行われる。モデルの有効性については、訓練データをテストデータとして用いた場合、明らかに不適切であるタイミング（例：発話が止まる、同じ内容を繰り返す）でのフィードバックを81.5%回避することができ、リコール率は74.14%であったと報告されている。しかし、これをプレゼンテーションの自動評価システムに実装し、継続的に使い続けたことによる訓練効果については未検証である。

そこで本研究では、この先行研究によって提案されたモデルを基に、システムを継続的に利用することによって生じる訓練効果の量的評価、およびフィードバックのタイミングが訓練効果に与える影響について検証を行った。また、フィードバックに対する受容度は個人差が大きいということが予備実験の結果から明らかになったため、フィードバックを提供する閾値については、実験初日に得られたデータから個人に合わせて設定し実験を実施した。

4 問題の評価指標と検出方法

プレゼンテーションの良し悪しに影響を与える指標は様々な文献 [1] で挙げられているが、本研究では発話

速度、発話音量、笑顔、ジェスチャー、アイコンタクト、姿勢の6つを主要な評価指標として定め、各種文献を参考に問題（できていない）と判断する閾値を定めた。ただし、笑顔、姿勢、ジェスチャーについては、閾値の参考になる文献が未だ少なかった為、我々の経験則と予備実験の参加者の意見を基に閾値を定めた。

4.1 発話速度の問題検出

プレゼンテーション中に普通の人は緊張や不安などから早口になる傾向があるが、適切な速度でゆっくりと話した方がよい。一般的に発話速度は単位時間あたりのモーラ数で表され、モーラとは音韻論上、一定の時間的長さをもった音の分節単位であり、現代の日本語では五十音図のそれぞれのかな（濁・半濁音、拗音を含む）を一単位としている。適切な発話速度の閾値については、松村ら [8] と志村ら [9] の文献を基に3.6[モーラ/秒]以上、9[モーラ/秒]以下を適切な発話速度と定めた。発話速度の検出は、Microsoft Speech To Text [10] を用いて発話内容と発話時間を取得し、形態素解析ライブラリである MeCab [11] を用いて仮名に変換後、モーラ数を発話時間で割って算出している。

4.2 音量の問題検出

発話音量の閾値は、長井今ら [12] の文献を基に、少し音が小さいと感じる音圧レベルを33dB、うるさいと感じる音圧レベルを72dBとして定め、過去10秒間のうち最も検出頻度の高い音圧レベルをその時の声の大きさと定めた。声の大きさの取得は、本研究での被験者実験で用いた HTC VIVE PRO Eye に内蔵されているマイクロフォンから取得を行い、本システムの構築に用いたゲーム開発エンジンである Unity 内で処理される音圧レベルを用いて問題検出を行った。

4.3 アイコンタクトの問題検出

プレゼンテーション中の適切なアイコンタクトの時間については、Fullwoodら [13] の文献を基に現時点から過去20秒間のうち聴衆の方を見ている時間が30秒以上（6秒以上）と定めた。聴衆の方を見ているかの判断は、Unity内での頭部オブジェクトから目線方向にRayを放射し、聴衆であるエージェントの顔付近と衝突判定が起きている場合をアイコンタクトができていると判定した。

4.4 表情の問題検出

プレゼンテーションでの笑顔は発表者の印象に繋がる重要な要素である。笑顔の判定は、VIVE フェイシャルトラッカーという笑顔や悲しい表情など38種類の顔の動きをリアルタイムに取得可能なデバイスを用いて笑顔検出を行い、現時点から過去10秒間のうち1秒以上笑顔を行った場合を笑顔ができている状態として定めた。

4.5 姿勢とジェスチャーの問題検出

姿勢とジェスチャーの判定は、Microsoft Azure Kinectを用いて体の動きをトラッキングすることで行った。具体的に、姿勢については背筋が真っ直ぐ伸びているかどうかを判定するために、骨盤座標を中心とし頭部方向に伸びた円柱内に頭部座標が収まっている割合が現時点から過去10秒間のうち7割以上の場合を姿勢が良い状態として定めた。ジェスチャーについては、現時点から過去30秒間のうち左右どちらかの手がへそより上にあり、その手が一定距離動いている場合を適切なジェスチャーとして定めた。

5 システム概要

本研究では、前章で定義された6つのプレゼンテーションの評価指標である発話速度、発話音量、笑顔、ジェスチャー、アイコンタクト、姿勢をリアルタイムで評価しフィードバックするシステムをVR環境下に開発した。

5.1 フィードバックインターフェース

VR空間内には各評価指標に対応するアイコンが配置される。これらのアイコンはユーザの頭の動きに連動し、図1に示すように常に視界内に存在している。これにより、プレゼンテーション中に即座にフィードバックを知覚することができる。

5.2 リアルタイム評価とフィードバック

評価指標については前章で設定した閾値に基づいてリアルタイムに評価される。問題が検出された場合、図1に示すように対応するアイコンが赤く点滅することでフィードバックが提供される。複数の問題が同時に検出されフィードバック可能な場合は、継続時間が最長の問題が優先してフィードバックされる。

5.3 フィードバックの初期化と待ち時間

一度フィードバックが提供された場合、そのフィードバックに該当する問題の評価内容は初期状態にリセットされる。また、システムがフィードバックを行った場合、その後20秒間は次のフィードバックが提供されないように待ち時間が設定されている。これは、発表者がフィードバックによる問題を修正するための待ち時間であり、次のフィードバックに対応するための準備時間を確保するためである。

5.4 タイミング推定に基づくフィードバック

本システムは、先行研究で提案されたタイミング推定モデルを採用しており、第3章で詳述した推定方法に基づき発表者の受容度を計算する。計算された受容度がある閾値以上（つまり、受容度が高いと判断された場合）にのみ、フィードバックが発表者に提供される。ただし、受容度が一貫して低い場合、フィードバックが一切行われないう問題が生じる可能性がある。



図1 VR環境下における発表者視点

これを防ぐために、各問題（評価指標）に「最長継続時間（20秒間）」を設定した。この時間を超えた場合、受容度とは無関係に強制的にフィードバックが提供される。

6 被験者実験

6.1 仮説

本実験の目的は、VR環境下でのプレゼンテーション訓練におけるリアルタイムフィードバックにおいて、先行研究で提案されたタイミング推定モデルによるフィードバックが訓練効果に与える影響について調査することである。仮説として以下の2つを立てた。

1. タイミング推定ありの方がタイミング推定なしよりもフィードバックに対して良い印象を持つ
2. タイミング推定ありの方がタイミング推定なしよりも訓練効果が高い

ここでの「訓練効果」は、訓練前と訓練後でのプレゼンテーションの評価結果の差、および訓練過程での問題点の改善速度を指す。

6.2 実験条件

本実験には、22歳から26歳までの同様の学生20名（男性17名、女性3名）が参加した。事前に参加者に対してプレゼンテーションの巧拙に対する自己評価を1（苦手）～7（得意）の尺度で行ってもらい、その結果を用いて、2つの実験条件のグループ間でプレゼンテーションのスキルレベルが均等になるように参加者を配分した。

条件1：タイミング推定なし

この条件下では、タイミング推定モデルは用いず、第5.2節で詳述したフィードバック方法に従い、検出された問題がフィードバック可能な状態であれば、即座にフィードバックが提供される。

条件2：タイミング推定あり

この条件下では、タイミング推定モデルを用いて



図2 VRシステムを用いない環境

フィードバックのタイミングを調整する。フィードバックの方法については、第5.4節で示したタイミング推定によるフィードバックに従う。

6.3 評価指標

本実験では、実験参加者によるシステム及び自身のプレゼンテーションに対する主観的評価と本実験に参加していない2名の評価者による客観的評価の2つの側面から行う。以下にそれぞれの具体的な評価内容を示す。

主観的評価では、参加者は各実験セッション後に、プレゼンテーション中に受けた全てのフィードバックに対して以下の項目で5段階評価を実施した。

集中力：1（ほとんど集中していなかった）～5（かなり集中していた）

フィードバックを邪魔だと感じたか：1（全く気にならなかった）～5（気になった）

フィードバックによって問題を修正できたか：1（何も変わらなかった）～5（すぐに行動を変え、発表の間意識していた）

加えて、システム全体に対する印象やその他の感想もアンケートで収集した。

客観的評価では、本実験に参加していない2名の評価者に実験参加者の訓練前と訓練後に撮影したVRシステムを用いず普段通りプレゼンテーションを行う動画を視聴してもらい評価を実施した。評価にはPeetersら[14]によって提案されたルーブリックから、非言語コミュニケーションの評価項目のみを使用し採点を行った。採点后、各評価者の訓練後のスコアと訓練前のスコアの差を平均化することで、訓練効果を評価した。

6.4 実験手順

5日間の実験は以下の流れで行われた。まず参加者は、事前にこちらで用意したプレゼンテーション資料の内容の理解を行ってもらい初日の実験に参加した。初日は、図2に示す環境で5分程度の一般的なプレ

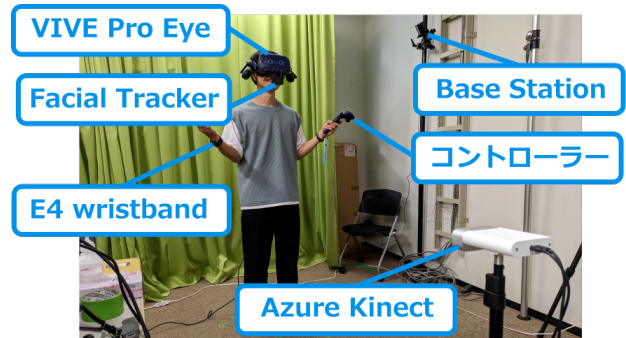


図3 VRシステムを使った訓練環境

ゼンテーションを1度行った後に、図3に示すようなヘッドマウントディスプレイ（HMD）を装着してもらいVR環境に慣れてもらった。その後、VRシステムを用いたプレゼンテーションを一度行う。2日目～4日目は、各日3回VRシステムを使ったプレゼンテーションを行い、最終日の5日目は最初にVRシステムを用いたプレゼンテーションを一度行ったあと、初日と同様図2に示す環境で一般的なプレゼンテーションを行い、5日間の実験が終了となった。

6.5 実験環境

VRシステムを使わない実験環境を図2に示す。参加者は、事前に共有したプレゼンテーション資料をプロジェクターで壁に投影する形でプレゼンテーションを行った。スライドの操作は参加者の前に配置されたラップトップまたは手持ちのレーザーポインタを用いて行われた。聴衆は、VR環境内のエージェントの数に合わせマネキン2体と実験担当者1人の計3人が配置された。動画の撮影にはiPad Air（第4世代）が用いられており、参加者からカメラまでの距離は約3.4m離れている。

VRシステムを使った訓練環境を図3に示す。参加者はHTC VIVE PRO EyeにVIVE フェイシャルトラッカーを装着したHMDを頭部に装着し、左手にはVR環境内でスライド操作やレーザーポインタに用いるVIVEコントローラを装着する。右手にはフィードバックのタイミング推定に用いる指標である心拍数や発汗量を計測するための装置であるE4wristbandを装着する。HMDの位置トラッキングに用いるSteamVR Base Stationは3台使用しており、それぞれ参加者の初期位置から半径約1.8mのところの間隔を空けて設置された。ボディトラッキングにはMicrosoft Azure Kinectを使用し、参加者の初期位置から1.8m離れた場所に、参加者と対面する形で設置された。

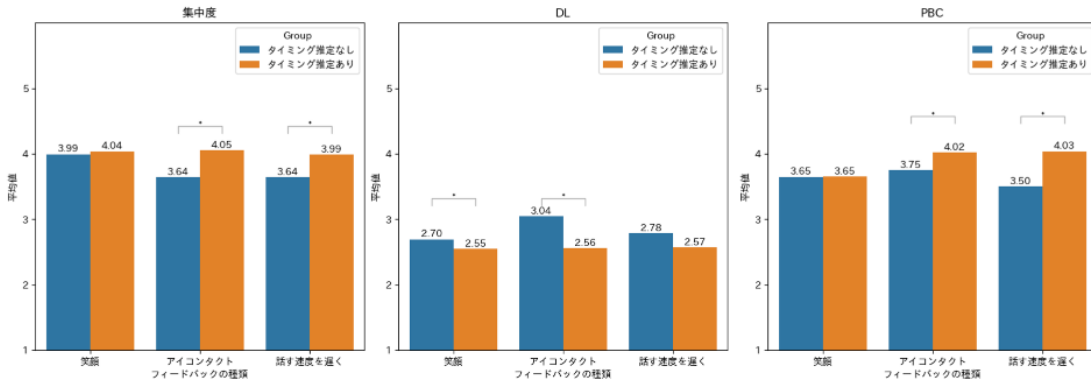


図4 システムからフィードバックに対する主観的評価

7 結果

7.1 主観評価の結果

本実験における主観評価の結果を図4に示す。本実験では全部で6つの問題検出が可能だったが参加者が受けた全てのフィードバックのうち、笑顔のフィードバックが全体の54.3%，次にアイコンタクトが30.2%，発話速度が速いというフィードバックが10.3%という結果となり、その他のフィードバックは2%以下という結果となった。この結果から参加者の訓練効果に寄与する可能性の高い笑顔、アイコンタクト、発話速度の3つのフィードバックについて図4では示す。

結果から、ほぼ全ての評価項目においてタイミング推定ありの方がスコアが高いという結果となった。また、各評価項目の笑顔、アイコンタクト、発話速度の3つについて、タイミング推定の有無で有意な差があるかマンホイットニーU検定を用いて分析を行った。その結果、 $p < 0.05$ でアイコンタクトについては全ての評価項目で有意な差が確認された。また、発話速度については集中度とPBC、笑顔についてはDLにおいて $p < 0.05$ で有意な差が確認された。これらのことから仮説1が支持された。

7.2 客観的評価の結果

評価者による客観的評価結果を図5に示す。各評価項目においてタイミング推定の有無で訓練効果に有意な差があるかを確かめるためにマン・ホイットニーU検定を用いて分析を行ったが有意な差は確認されなかった。このことから仮説2は支持されなかった。

7.3 相関分析の結果

タイミング推定に基づいた訓練効果が個々の特性とどの程度相関するかを調査するために、参加者の特性と訓練効果の相関分析を実施した。その結果、いくつかの有意な相関が確認された。タイミング推定なしの条件では、発話速度の訓練効果と特性不安との間に負の相関(相関係数: -0.68 , $p=0.02$)が見られた。一

方で、タイミング推定ありの条件では、アイコンタクトの訓練効果と特性不安との間(相関係数: -0.74 , $p=0.02$)、ならびに訓練前の状態不安との間(相関係数: -0.69 , $p=0.04$)に負の相関が確認された。

8 考察

8.1 実験参加者によるフィードバックの主観的評価結果の考察

集中度とDLを用いた分析に基づく結果から、タイミング推定ありのグループは、タイミング推定なしのグループよりも、フィードバックを受けた際のプレゼンテーションに対する集中度が高く、それでいてフィードバックが邪魔だと感じなかった。これは、システムが適切なタイミングでフィードバックを提供した可能性を示唆している。また、アイコンタクトについて全評価項目において有意な差があったのに対し、発話速度と笑顔は部分的にしか有意差が確認されなかった。これは、タイミング推定が特にアイコンタクトに対して影響を与える可能性が高いことを示唆している。システムに対する主観的な印象というものは重要であり、特に訓練システムのように長期的な利用を考えた場合、ユーザの継続率を高めることはその効果性に大いに寄与すると考えられる。

8.2 ルーブリックを用いた客観的評価結果の考察

客観的評価において、有意な差が確認されなかった一因として、サンプルサイズの小ささが挙げられる。具体的には、タイミング推定の有無の条件でそれぞれ10人の被験者を対象としており、このサンプルサイズが統計的な検証が不十分である可能性がある。さらに、本研究で用いたルーブリックの評価尺度が1~4点という狭い範囲であり、この制約がより細かい評価を妨げている可能性も考慮する必要がある。このような狭い評価尺度では、1点の差が結果に大きな影響を与えるため、微細な訓練効果が反映されにくいという問題が存在する。

8.3 相関分析結果の考察

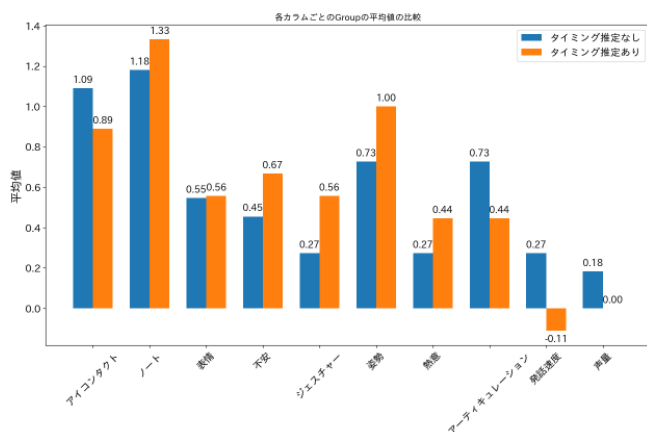


図5 プレゼンテーションループリックによる客観的評価

相関分析の結果において、特性不安と訓練効果には相関があり、また、タイミング推定の有無によって訓練効果があったスキルに違いが見られた。このことから、参加者のもつ不安と関連して、タイミング推定の有無が異なるスキルの改善に寄与する可能性があると考えられる。

9 結論

本稿では、VR環境下でプレゼンテーションを行う発表者に対するリアルタイムフィードバックに先行研究で提案されたタイミング推定モデルを用いることによって、タイミング推定なしよりもフィードバックに対して良い印象を持つ可能性が主観的評価の結果より示唆された。一方で、客観的評価ではタイミング推定の有無の群間で有意な差は確認されなかった。これは、サンプルサイズの小ささや評価尺度の狭さに起因する可能性が高い。相関分析の結果では、タイミング推定の有無が特性不安や状態不安と訓練効果との関係性に影響を与える可能性が示唆された。

今後は、サンプルサイズの拡大と評価尺度の改善によって、本研究で明らかになった仮説と相関関係などを更に詳細に解明していく予定である。

謝辞

本研究は、JST, CREST, JPMJCR19Aの支援を受けたものである。なお、人を対象とする研究に関する先端科学技術研究科倫理審査委員会(承認番号2022-I-54)をうけて、実施されたものである。

参考文献

[1] Robert Dolan: Effective presentation skills; FEMS Microbiology Letters, 364(24) (2017.Nov)

[2] VirtualSpeech Ltd. Virtual Speech. Available online: <https://virtualspeech.com/> (accessed on 27 April 2021).

[3] Paul E. King, Melissa J. Young & Ralph R. Behnke: Public speaking performance improvement as a function of information processing in immediate and delayed feedback interventions; Communication Education, 49:4, 365-374 (2000.Mar)

[4] Kazutaka Kurihara, Masataka Goto, Jun Ogata, Yosuke Matsusaka, and Takeo Igarashi: Presentation sense: a presentation training system using speech and image processing; Proceedings of the 9th international conference on Multimodal interfaces (ICMI '07), 358-365 (2007.Nov)

[5] Ionut Damian, Chiew Seng (Sean) Tan, Tobias Baur, Johannes Schöning, Kris Luyten, and Elisabeth André: Augmenting Social Interactions: Realtime Behavioural Feedback using Social Signal Processing Techniques.; Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), 565-574 (2015.Apr)

[6] Kazunori Akizuki, Ryohei Yamamoto, Kazuto Yamaguchi, Jun Yabuki, and Yukari Ohashi: The effects of feedback timing and frequency on the acquisition of cardiopulmonary resuscitation skills of health sciences undergraduate students: A 2 x 2 factorial quasi randomized study; PLoS one 14(7), e0220004, (2019.Jul)

[7] Yuichiro Fujimoto, Hangyu Zhou, Taishi Sawabe, Masayuki Kanbara, and Hirokazu Kato: Stop Bad Real-time Feedback!: Estimation of the Timing of Feedback that Negatively Impacts Presenters for Presentation Training in Virtual Reality; Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR2023), Poster, (2023.Oct)

[8] 村松賢一: II プレゼンテーション能力測定尺度の作成(ストレートトーク番組におけるプレゼンテーション技法の研究); 研究報告 104, 13-105 (1997)

[9] 志村栄二, 筑一彦: Dysarthria 例の発話速度調節訓練に影響を与える要因の一考察(第1報); 音声言語医学, 53.4, 302-311 (2012)

[10] <https://azure.microsoft.com/ja-jp/products/cognitive-services/speech-to-text>

[11] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto: Applying Conditional Random Fields to Japanese Morphological Analysis; Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), 230-237 (2004.Jul)

[12] 長井今日子, 小寺一興, 芦野聡子, & 前田知佳子: 会話をたいする快適レベルと不快レベルの測定; Audiology Japan, 34.4, 259-264 (1991)

[13] Chris Fullwood, Gwyneth Doherty-Sneddon: Effect of gazing at the camera during a video link on recall; Applied Ergonomics, 37(2), 167-175 (2006.Mar)

[14] Michael J. Peeters, Eric G. Sahlo, and Gregory E. Stone: A standardized rubric to evaluate student presentations; American journal of pharmaceutical education, 74(9), 171 (2010.Nov).

(c) 2023 by the Virtual Reality Society of Japan (VRSJ)