

仮想プレゼン空間の聴衆者のリアクションに基づく声の表現力 向上支援システムの提案

吉田圭吾¹ 佐野睦夫² 砂川顕子³

概要: 社会人になると学生の頃と比べて、上司とのコミュニケーション、プレゼンや会議での発言など人前で話す機会が圧倒的に増えるため、プレゼンテーション力が大きく求められる。また、人前で話すのが苦手という人の割合が多いことから、本研究では、発話音声、発話内容、視線、に着目し、プレゼンテーションの評価フィードバックを行い、トレーニングを通してプレゼンテーション力向上を図る。結果として視線、音声パワー、については4人中4人とも評価が向上し発話速度、ドッグワードについては4人中3人が向上する結果となった。また、オクターブについてはあまり優位な結果が得られないものとなった。今後の展望として発表中の間の取り方や間隔を分析し評価していく必要がある。

1. はじめに

近年コロナウイルスの影響やオンライン化が促進したことによりプレゼンをする機会が増え、プレゼンテーション力[1]が大きく求められる。そこで、本番の環境に慣れるために経験や練習が必要だが、練習と本番では話すスピードが速くなってしまい、伝わりにくい発表になってしまうなど、1人で練習することは難しい。というのも、1人練習をしても実際に自分が良いプレゼンテーションができていないかの判断をすることは難しい。また、1人で練習する際はpcの画面を常に見ながら練習する人が多いが本番では、常にスクリーンを見る発表はよく無いため、本番に近い環境での発表練習が必要となる。そこで今回はこれらの問題を解決するため、VR内でプレゼンテーショントレーニングや発表を評価[2]し、ユーザにフィードバック[3]することでプレゼンテーションスキル向上させる、支援システムを開発する。

私はこれまで、対話エージェントを用いたQOL(生活の質)支援システムの開発に取り組んできた。QOL支援システムでは、ユーザがその日の出来事やコミュニケーション内容について対話を行った後、エージェントとの話し方トレーニングを実施する。トレーニングの内容は、短い文章を伝えたい部分を強調して発音することで会話の改善を図るものである。その後、エージェントが腹式呼吸の方法を説明し、ユーザはプリントに書かれた絵本の文を腹式呼吸に意識しながら読む。この過程で、腹式呼吸、表情、感情の測定を行い、次のトレーニング時に振り返りやアドバイスを提示することで、さらなる改善を促進するというものである。それらに応用し、今回はVRトレーニング内に実装し、開発を行う。トレーニングにはVRを用いてNPC一人一人に感情を持たせ、一人一人異なるリアクションをするような実際の発表に近い空間での発表トレーニングや話し方トレーニング、発表内容の改善トレーニングなどを行う。

2. 関連研究

ビジネスプレゼンテーションにおける言語・非言語的能力の自動推定

八木ら[4]は、プレゼンテーションデータから得られたマルチモーダル情報を用いて「プレゼンテーション能力」を推定するモデルの構築と評価を行った。発話内容に含まれる品詞情報、韻律情報、動作量をマルチモーダル特徴量として抽出し、個人のプレゼンテーション能力値を機械学習により推定するモデルを構築した。評価実験の結果、「目的網羅性」「内容ロジック」「Visual&Vocal」「効果的演出要素」に関する評定値について、それぞれ回帰タスクで最大0.59, 0.51, 0.38, 0.29の決定係数を得た。

3. 提案手法

3.1 目的

提案するVRプレゼンテーションスキル向上支援システムを通して、発表時のコツ[5]やポイントを伝え、VR内で本番に近い状況で練習をし、発表時のポイント可視化させユーザに伝えることによって、プレゼンテーション力を向上させることを本研究の目的としている。

3.2 システム構成

今回提案する手法として現実空間でプレゼン発表後これらのフィードバックを行う。VRトレーニングでは複数のトレーニングを行い、その後自信で作った発表資料[6]を使用してVR内で実際の環境に近い環境で発表し、それらの評価をフィードバックすることでプレゼンテーションスキルをアップさせるというものである。具体的に発表の評価は音声パワー、発話速度、オクターブ、視線、ドッグワード、聞き手への問いの投げかけで行う。VRトレーニングでは初めに

1 大阪工業大学 大学院
Osaka Institute of Technology, Graduate School

2 大阪工業大学
Osaka Institute of Technology

3 VOICE 美 studio 合同会社
VOICE Beauty Studio Limited Liability Company

プレゼンを行う上で気をつけたいといけないポイントをいくつか説明し、良いプレゼンとされている動画を見てもらう。その後、適切な発話の速度トレーニングをしてもらう。一般的に発表時の適切な発話速度は1分間に約300字程度とされている。トレーニングでは、発表のポイントの一部を300時にまとめたものを1分間で発話してもらい適切な発話速度を練習してもらうというものである。また、このトレーニングと並行して、音声パワーのトレーニングも行う。プレゼン時の適切な声の大きさは65dB~75dBとなっており、図1のように声の大きさが表示される空間でトレーニングを行ってもらうことで自身の声を65dB~75dBに調節してもらうものである。

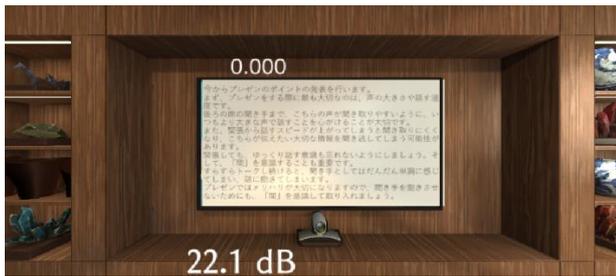


図1 発話速度トレーニング

その後発話トレーニングとして、私は夏に母とアメリカに1ヶ月間、初めての海外旅行に行きました。という文の誰と？どこに？どれくらいの期間？何を？などの質問に応じて返答になっている部分の強弱を意識して答えてもらうトレーニングを行う。具体的には誰と夏を過ごしていましたか？という質問には母との部分を強めて発話してもらうものである。そしてプレゼンのつかみの部分として発表で自信が最も伝えたい部分の中から1つをプレゼンの初めに問いとして投げかけてみる練習を行う。最後にVR空間で複数のNPCがいる前で発表をしてもらう。尚、その発表中にNPCがユーザの発話に応じて行動が変化する。具体的には55dBに満たない大きさに発話している場合、聞こえにくく、内容が伝わっていないものとし、図2のようなダルそうな伸びをするポーズや発表中に寝ているポーズ、残念がるポーズをし、逆に65dB以上の声の大きさに発表していると聞き取りやすい発声と評価し、NPCは図3のような発表中にはうなずきが多くなり、発表後は拍手をしてくれるといったものである。

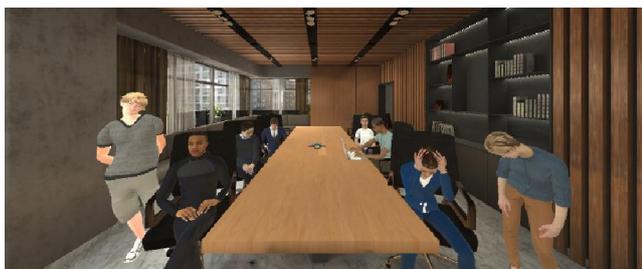


図2 評価の低いNPCの反応



図3 評価の高いNPCの反応

3.3 視線計測

視線計測[7]では、現実空間とVR空間で計測方法が異なり、現実空間の際はOpenCV[8]を用いて発表中の聴講者を見ている割合を計測した(図4)。VR空間内ではNPCの居る方向とスクリーンの方向にそれぞれ見えないボックスを用意し、それらを見ている割合を計測した。

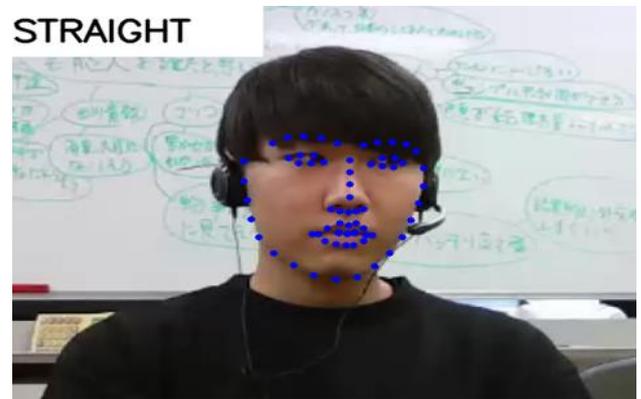


図4 OpenCVでの視線計測

3.4 音声パワー、オクターブ

音声パワーは録音された音声データの平均dB(音声パワーの平均値)で計算し、オクターブは、音楽や音響学の用語で、特定の基準となる周波数に対して、周波数が2倍または1/2倍になる音の関係を表す。オクターブの値を計算するには、基準となる周波数と比較して、指定された周波数が何オクターブ離れているかを以下の対数計算[9]で求めることができる。

$$\text{オクターブ値} = \log_2 \left(\frac{\text{周波数}}{\text{基準周波数}} \right)$$

3.5 発話速度

発表時の適切な発話速度は、1分間に約300文字が目安とされている。この範囲は、聞き手が内容を理解しやすく、集中力を保ちやすい速度である。今回は発表時の音声を録音し、それらを文字お越ししその文字数を発表時間で割ったものを発話速度として扱っている。

3.6 ドッグワード

ドッグワードとは『えー』や『んー』などの意味を持たない言葉を表し、これらが多すぎると話が長く感じたり、聞きづらいことがある。今回は発話速度同様、発表時の音声を

文字お越ししドットワードの数を検出した。

3.7 アンケート

アンケートではプレゼンテーションを通して自分自身と他の参加者を聞いてどうだったかの2種類のアンケートに回答してもらう(20問程度)。具体的には声の大きさやドットワードの多さ、緊張していたかななどを1~5の5段階で評価してもらうものである。

4. 実験

4.1 実験対象者

参加者は20代の大阪工業大学に在籍している学生4名である。参加者には事前に実験の手順が書かれているプリントを渡し、大まかな説明している。

実施期間は2024年12月、実験場所は執筆者の所属するゼミ室である。実験の実施に当たり、全員に研究内容を説明し、実験参加への同意を得た。ライフサイエンス実験承認番号2023-10を得た上で実験を行った。

4.2 実験手順

今回の実験ではVRトレーニングシステムの効果を検証するべく、現実空間、VR空間、現実空間の順に発表を3回行う。

参加者にはまず、コンセンサスゲームというゲームを行なってもらう。コンセンサスゲームとは、与えられた課題について、チームで協力しながら答えを見つけていくものである。具体的には砂漠に遭難したメンバーが、不時着した飛行機の中から見つけた12個のアイテムの中で必要なものから重要度の高い順に1番から12番までの順位をつけ、それらが自分の考えているものと他の参加者が考えているものを比較し、チームで話し合い一つの答えを導き出すというものである。その後、コンセンサスゲームを自分がどのように考え答えを出したかという発表資料とコンセンサスゲームとは別に大学での授業についての発表資料の2つを作成してもらう。その後、現実空間で発表を行い、参加者には発表者の評価についてのアンケートを回答してもらう。また、アンケート回答時間にそれぞれの発表時の視線、発話音声、発話内容についてのフィードバックも同時に行った。その後、VRトレーニングを行ってもらい、最後に発表を再度行い最終アンケートに回答後実験は終了という流れである。VRトレーニングの実験中の様子については図5に示す。



図5 実験の様子

5. 結果

5.1 発表時間

表1 発表時間

発表時間(秒)	1回目	VR内	2回目
参加者1	184	200	145
参加者2	238	183	161
参加者3	120	109	105
参加者4	169	157	148

5.2 視線

表2 聴講者を見ている視線割合

視線	1回目	VR内	2回目
参加者1	62.30%	42.30%	78.00%
参加者2	71.90%	61.30%	72.70%
参加者3	57.70%	56.80%	66.50%
参加者4	45.00%	53.20%	70.90%

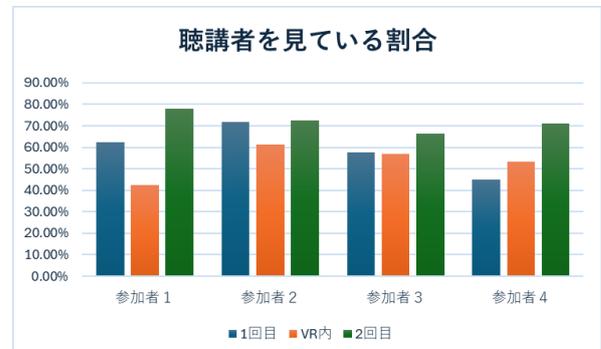


図6 聴講者を見ている視線割合比較

5.3 音声パワー

表3 声の大きさ

平均dB	1回目	2回目
参加者1	64.18	67.82
参加者2	62.48	67.69
参加者3	63.13	65.64
参加者4	57.11	60.68

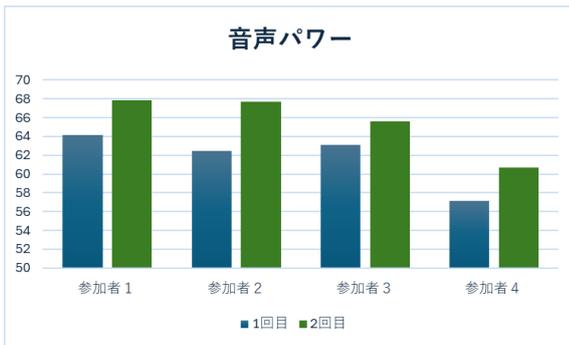


図7 声の大きさ比較

5.4 オクターブ

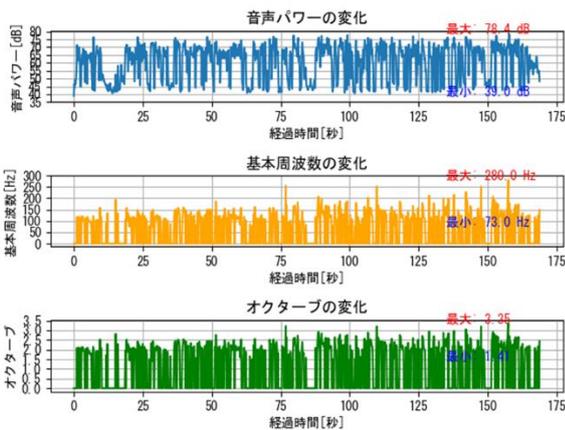


図8 音声波形 (大きさ,周波数,オクターブ)

表4 オクターブの変化量

オクターブ	1回目	2回目
参加者1	2.61	2.71
参加者2	2.88	2.7
参加者3	1.78	2.34
参加者4	1.91	1.85

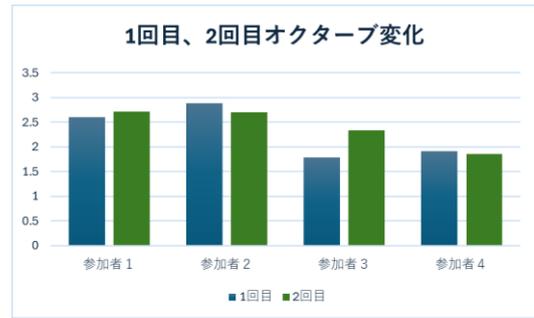


図9 オクターブの変化量比較

5.5 発話速度

表5 発表時の発話文字数

発話文字数	1回目	2回目
参加者1	1152	825
参加者2	1280	823
参加者3	659	542
参加者4	897	708

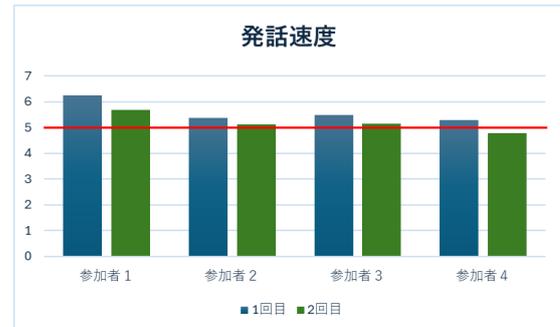


図10 発話速度比較

5.6 ドッグワード

表6 発表時のドッグワード数

ドッグワード	1回目	2回目
参加者1	15	11
参加者2	41	28
参加者3	32	24
参加者4	27	21

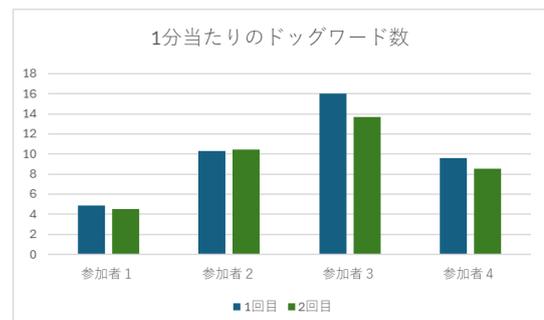


図11 1分当たりのドッグワード数比較

5.7 アンケート

表7については発表者以外の3人の実験参加者の5段階評価の合計15点満点で得点を表している。

表7 発表時の聞き取りやすさ

発表の聞き取りやすさ			発表時の話す速度		
	1回目	2回目		1回目	2回目
参加者1	13	14	参加者1	13	13
参加者2	12	14	参加者2	12	14
参加者3	11	11	参加者3	12	12
参加者4	11	12	参加者4	11	12

表8 現実空間とVR空間比較

現実空間とVR空間、どちらが緊張？		発表時のNPCの変化に気づきましたか？	
参加者1	現実空間	参加者1	気づいた
参加者2	現実空間	参加者2	気づいた
参加者3	現実空間	参加者3	気づかなかった
参加者4	VR空間	参加者4	気づいた

6. 考察

今回の結果から音声パワーに関しては表3, 図6より4人中4人とも1回目の現実空間での発表よりトレーニング後の2回目の現実空間での発表の方が平均dBが大きく、アンケート(表7)でも聞き取りやすさの項目で4人とも2回目の発表の方が高い評価となった。これは、発話トレーニングや発話速度トレーニングの際に自分の声の大きさを調節してもらったことにより、普段の話し方より少し大きな声の方が適切だと気づきその後の2回目で調節したためだと考える。同様に、発話速度についても表1, 表5, 図10より4人中3人が適切な速度に近づいており、一人はあまり変わらなかったとい結果だが、これらに関してもトレーニングの効果が大きいと考える。

視線については1回目と2回目を比較した時に3人が聴講者を見ている割合が減っているが、これは自然計測の方法が現実空間とVR空間では異なり、また、VR空間では自分の立ち位置を変えることができるため、スクリーンと聴講者の間に立っている場合180度首を動かさないといけなくなるケースがあった。このように、VR空間内で視線の割合が低いのは発表者の立ち位置によって動作が変わってしまったこともあると考える。しかし、1回目と2回目の現実空間で結果を見ると4人とも聴講者を見ている割合が増えているため、視線についてのフィードバックやアドバイスの効果が大きいと考える。オクターブについては最低オクターブと最高オクターブの差が大きい場合や変化量が適度に多い場合メリハリのある聞きやすい発話になり、今回の場合1回目と2回目を比較した時に優位な差は見られなかった。そして2回目では4人とも発表のつかみの部分で聞き手に疑問の投げかけを行うことができていた。

ドッグワードについては表6, 図11より4人中3人が1分間のドッグワードを減らすことができていた。従来では

ドッグワードのフィードバックを個人で行なっていたが、その際にドッグワードを言わないようにすることに意識が行き、発表が詰まってしまったり、逆にドッグワードが増えてしまう人がいた為、今回はトレーニングの際にアドバイスやポイントとして説明することで少し意識してもらおうようにした。これらが結果として良いものとなったと考える。また、現実空間とVR空間でどちらの方が緊張しましたかというアンケートに関して、3人が現実空間、1人はVR空間と回答しており、VR空間の方が緊張したと回答した人は少し声の大きさが小さくNPCの行動で図2のような評価の低い反応をしたことでNPCが圧に感じ、VR空間の方が緊張したと回答したのではないかと考える。

7. 今後の展望

今後の展望として今回発表時のつかみや強弱の評価を疑問の投げかけができていないか？とオクターブ変化のみで行なったが、発表中の間の取り方や間隔を分析し評価していきたいと考えている。また、NPCの反応についても現在は声のパワーでの反応のみだが、発話時の強弱や間の取り方の違いでNPCの反応の違いを追加していきたいと考えている。

8. まとめ

トレーニングとフィードバックを通して現実空間での発表にどのような変化が起こるのかということで検証を行った。結果として視線、音声パワー、については4人中4人とも評価が向上した。また発話速度、ドッグワードについては4人中3人が向上する結果となった。また、オクターブについてはあまり優位な結果が得られないものとなった。今回の結果からNPCの行動とユーザの緊張に相関があるのではないかとこの予測がたつた。今後は発表時のつかみや強弱の評価を疑問の投げかけができていないかとオクターブ変化のみで行なったが、発表中の間の取り方や間隔を分析し評価していく必要がある。

参考文献

- [1] 竹田尚彦, et al. プレゼンテーション演習における学生間相互評価の分析. 情報処理学会研究報告コンピュータと教育 (CE), 2005, 2005.36 (2005-CE-079): 55-62.
- [2] 山下祐一郎. プレゼンテーションにおけるスライド評価と発表評価の一致率. 日本教育工学会論文誌, 2021, 44.Suppl.: 5-8.
- [3] 山下祐一郎; 中島平. プレゼンテーションスキルと分かりやすさの関係分析: レスポンスアナライザによる評価とアンケート分析の比較. 日本教育工学会論文誌, 2010, 34.Suppl.: 5-8.
- [4] 八木悠太郎, et al. ビジネスプレゼンテーションにおける言語・非言語的能力の自動推定. In: 人工知能学会全国大会論文集 第33回 (2019). 一般社団法人 人工知能学会, 2019. p. 4F2OS11a02-4F2OS11a02.
- [5] GLOBIS CAREER NOTE, 仕事が上手くいく話し方のコツとは? 自分の考えや意見を正確に伝えよう!, <https://mba.globis.ac.jp/careernote/1104.html> (最終閲覧日 2024 12 月 19 日)
- [6] 栗原一貴, et al. スライド提示型プレゼンテーション方法論の拡張手法を定量的に評価する研究. 情報処理学会論文誌,

2010, 51.2: 391-403.

- [7] HAKKY ,Dlib による視線方向推定, <https://book-st-hakky.com/data-science/eye-deirection-overview/>(最終閲覧日 2024 12 月 19 日)
- [8] @fujino-fpu(秀則 藤野), Python と OpenCV を使った笑顔認識 <https://qiita.com/fujino-fpu/items/99ce52950f4554fbc17d>(最終閲覧日 2024 12 月 19 日)
- [9] ONO SOKKI 技術レポート騒音計 (サウンドレベルメータ) とは第 3 版 https://www.onosokki.co.jp/HP-WK/c_support/newreport/noise/souon_11.htm (最終閲覧日 2024 9 月 19 日)