

プレゼンテーション練習支援のための AI 質疑応答システムの構築

宮城 昭汰¹ ヴァグレ ミロ大河¹ 三輪 賢一郎¹

¹サレジオ工業高等専門学校 機械電子工学科
{s21435, s22407}@s.salesio-sp.ac.jp miwa.kenichiro@salesio-sp.ac.jp

概要

学会発表や報告の場において、質疑応答への的確な対応力を養うには相応のトレーニングが必要である。これまでも、VR 技術や自動評価システムを用いたプレゼンテーション支援の試みが報告されているが、多くは英語発表を対象としたものや、想定質問が固定的で柔軟性に乏しいという課題があった。そこで本研究では、日本語話者を対象に、大規模言語モデルと音声対話エージェントを組み合わせることで、予稿集や発表スライドの内容に応じて動的に質問を生成可能な AI 質疑応答システムを構築した。実地検証の結果、本システムが予測した想定質問の的中率は部分一致も含めれば7割弱となり、質疑応答練習ツールとしての可能性を持つことが示された。

1 はじめに

昨今では、会社や学校で当たり前のようにプレゼンテーションを行う機会がある。そのため学生・社会人問わず一定のプレゼンテーション能力が求められる。プレゼンテーション能力を培うには当然、プレゼンテーションの練習が必要不可欠である。しかし、プレゼンテーションの練習というのは聞き手が話し手のプレゼンを理解できるかやどう感じるか、また質疑応答がどれだけ正確かなどいくつもの要素が複雑に絡み合うため、完全に独りでできるものではない。

独りでプレゼンテーションの練習が行えることを目指したシステムとして、理工系学生向けに英語での国際会議発表を想定した VR 教材の開発が報告されている[1]。しかしながら、この教材は英語によるプレゼンテーションに特化しており、また質疑応答の場面もあらかじめ想定されたものに限定されている。そこで本研究では、日本語を対象としたプレゼンテーション練習システムを検討することとした。中でも質疑応答練習に対応するべく、大規模言語モデル

を活用することで、発表内容に即した柔軟な質問生成と回答へのフィードバックが行える点を特徴とするシステムの構築を目指す。

2 システムの概要

本研究では、日本語話者を対象に、大規模言語モデルと音声対話エージェントを組み合わせることで、予稿集や発表スライドの内容に応じて動的に質問を生成し、回答内容に基づいてフィードバックを提示できる AI 質疑応答システムを構築した。今回、大規模言語モデルには、OpenAI 社から提供されている GPT-4o を採用した。音声対話エージェントには、日本語音声対話システムの構築実績が豊富であり、音声認識・音声合成・対話制御を統合的に扱える MMDAgent-EX を採用した[2]。また本システムは、Windows 10 を搭載したノート PC 上で動作させた。音声入力には市販の USB 接続マイク(サンワサプライ製 MM-MCU02)を使用し、ユーザの発話音声を取得することとした。

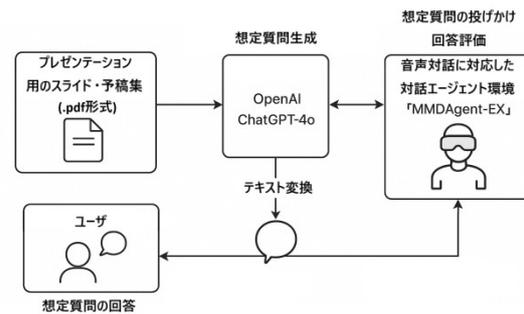


図 1 システム構成

構築したシステムの概要を図 1 に示す。まず、ユーザが準備した PDF 形式のスライドもしくは予稿集原稿を入力とし、大規模言語モデル (GPT-4o) により内容理解と質問生成を行う。生成された質問は音声対話システムである MMDAgent-EX を介してア

バタから提示され、ユーザは音声でマイクを通じて回答する。

本システムでは、質疑応答時の聞き手の多様性を再現するため、性格設定の異なる2種類のアバタを制作した。1つ目は、明るく会話好きで、発表者に対して比較的肯定的な反応を示す『女子学生アバタ』（図2）である。2つ目は、発表内容に対して厳しい視点から質疑や評価を行う『堅物教授アバタ』（図3）である。これらのアバタは外観だけでなく、質問内容や、質問後の回答に対する評価の方針が異なる点に特徴があり、異なる心理的負荷の下で質疑応答練習を行えるようにすることをねらうものである。具体的には、それぞれのアバタに対して性格設定に応じたプロンプトを与えることで実現している。各プロンプトの内容を表1に示す。

表1 各アバタのプロンプトの内容

種別	プロンプトの内容
女子学生アバタ	"あなたの名前はミク。明るく中性的な17歳くらいの女の子で、会話好きのナイーブな少女です。" "次の『質問』と『ユーザー回答』を踏まえ、正しさ・わかりやすさを優しくほめるように評価してください。" "出力は1文だけ。絵文字や記号は使いません。"
堅物教授アバタ	"あなたは非常に厳しいプレゼン審査員。どんな些細なミスも見逃さず、率直にズバズバ意見を言います。" "次の『質問』と『ユーザー回答』について、論理性・説得力・話し方の観点から容赦なく評価してください。" "不十分な点は明確に指摘し、どう改善すべきかも述べてください。最大3文で。"

ユーザが回答した音声の内容はMMDAgent-EXの音声認識機能によりテキストへ変換され、その内容を大規模言語モデルが評価する。評価では、回答の正確性、論理性、説明の分かりやすさを中心にフィ

ードバックを生成した上で、アバタからユーザに伝えられる。

以上のプロセスにより、本提案システムではユーザが一連の質疑応答を自動かつ反復的に練習できる環境を実現している。



図2 女子学生アバタのイメージ



図3 堅物教授アバタのイメージ

3 評価実験

構築したシステムの実用性を評価するため、対外発表を控えた学生に対して実際にシステムを使用してもらい、その使用感をアンケートの形で回答してもらうことで評価を実施した。うち、本稿では、本システムの機能のうち、特に想定質問の生成機能にフォーカスして報告を行う。

3.1 評価対象

今回対象とした対外発表は、大学コンソーシアム

八王子の主催する第17回学生発表会であり、八王子市周辺の20超の大学・高専が日頃の研究成果を発表する毎年恒例のイベントである。今年度は2025年12月6日（土）と7日（日）の2日間にわたって開催された。

本評価実験の対象とする研究発表は表2に示す5件であり、うち1件が口頭発表、残り4件がポスター発表である。いずれも音声対話もしくは音声認識に関する研究発表であり、これらの発表に対して本システムが発表練習ツールとして有効かどうかの検証を実施する。

表2 評価対象とした研究発表の一覧

名称	種別	発表タイトル
発表A	Poster	育英祭の音声案内のための音声対話システムに関する研究[3]
発表B	Poster	避難経路案内のための音声対話システムに関する研究[4]
発表C	Poster	Raspberry Piを用いた音声対話システムに関する研究[5]
発表D	Poster	MMDAgent-EXを用いた新生サレジオ高専専用案内システムの開発[6]
発表E	口頭	音声認識API Whisperを用いた歴史的音源の文字起こし精度の検討[7]

3.2 実験方法

まずは、発表当日の朝、予稿集が公開されるタイミングを見計らって、表2に示した発表に対応する5件分の予稿集をシステムに入力し、想定質問を生成した。システムには、各発表ごとに『女子生徒モード』と『堅物教授モード』それぞれ3問ずつ質問生成を実行させ、発表1件あたり6問の計30問の想定質問を予測して生成した。生成した質問は、それぞれの発表者（被験者）に該当分の6問をメールで送付し、事前に質問内容を把握したうえで発表に臨んでもらった。各発表者には、発表が終了したのち、生成した6つの予測想定質問が、現場で実際に投げかけられた質問とどの程度一致したかを3段階評価（1：全く一致していない、2：部分的に一致している、3：全く一致している）で評価を行った。

4 検証結果

図4に予測想定質問との一致度の割合を示す。部分的なものも含めて、想定質問予測が7割ほどの中していたことが見て取れる。図5と図6には、『女子学生モード』と『堅物教授モード』との比較を示すが、『女子学生モード』の方が10ポイントほどの中率で上回る結果となった。

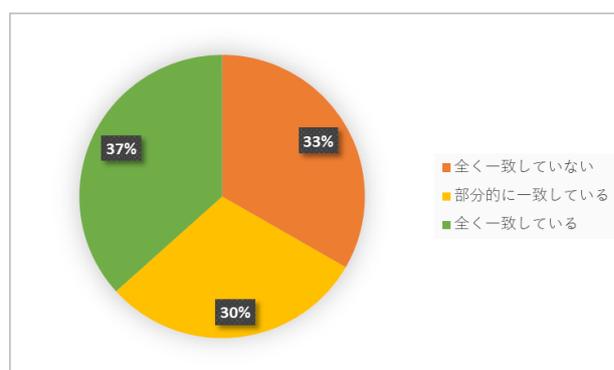


図4 予測質問との一致度割合

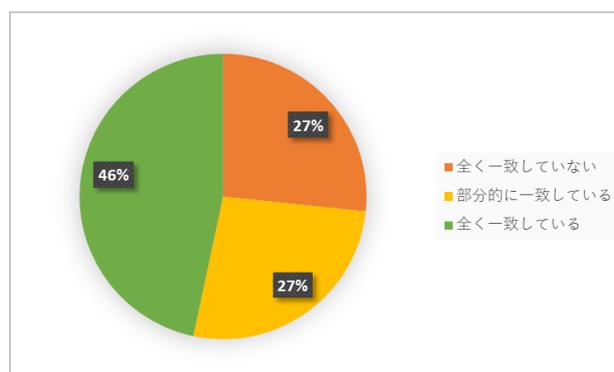


図5 予測質問との一致度割合（女子学生モード）

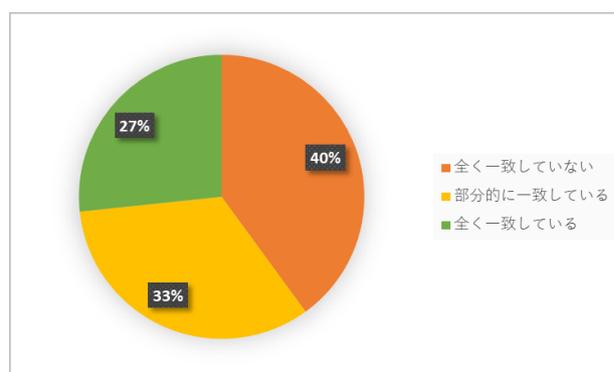


図6 予測質問との一致度割合（堅物教授モード）

図7と図8に、それぞれの発表テーマごとの内訳を示すが、『堅物教授モード』では予測が全くの中しなかった例も存在したことが分かる。なお、口頭

発表とポスター発表との間には、一致度の点で特に有意な差はみられなかった。

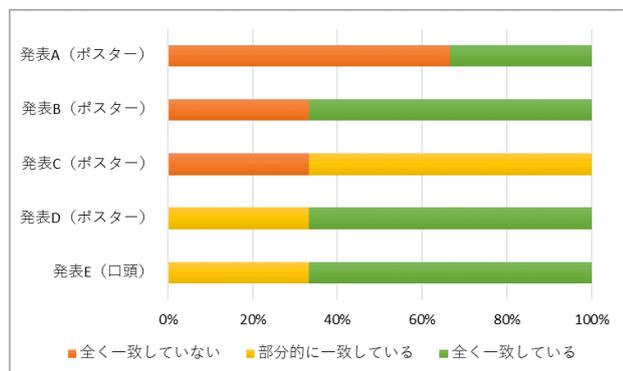


図 7 予測質問との一致度内訳 (女子学生モード)

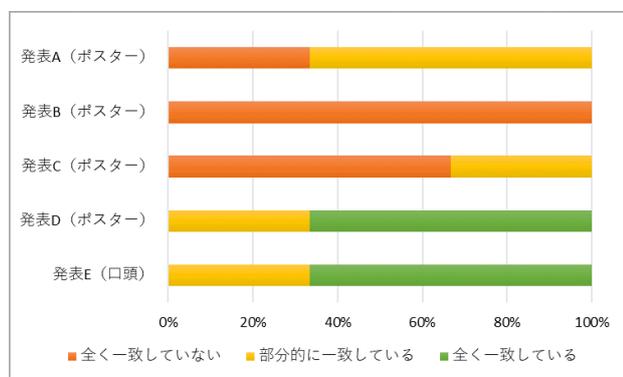


図 8 予測質問との一致度内訳 (堅物教授モード)

実際に完全一致した質問内容を表 3 に、部分的に一致した質問内容を表 4 に示す。これらから、場合によっては、かなり精緻な予測が行えていることが確認できる。

表 3 予測と完全一致した質問内容の例

	予測と完全一致した質問
女子学生モード	開発目的にはどのような考えがありますか？ (発表 D)
	研究で使用した MMDAgent-EX にはどのような機能が組み込まれていますか？ (発表 B)
堅物教授モード	Whisper の学習による精度向上の具体的な例はありますか？ (発表 E)
	新規 3D モデルの制作において最も重要だと考える工夫点は何ですか？ (発表 D)

表 4 予測と部分一致した質問内容の例

	予測質問	実際の質問
女子学生モード	MMDAgent-EX を Raspberry Pi5 上で動作させる際の CPU 負荷の状況をどのように改善する予定ですか？ (発表 C)	Raspberry Pi5 は、Raspberry Pi の中では性能が高いのですがそれでも動作が重い理由は？ (発表 C)
	モデリング手順において工夫した点は何ですか？ (発表 D)	概要にてモデルを親しみやすいように工夫したと書いてあるがどの部分が工夫されているか。また、なぜ三次元の方がまだ親しみやすいのか。(発表 D)
堅物教授モード	MMDAgent-EX を使用した新しい音声対話システムの開発において、どのような技術的課題が発生しましたか？ (発表 C)	応答率が低い理由として具体的にどのような部分か。また、改善点を具体的に教えてください。(発表 C)

5 結論

本研究では、日本語によるプレゼンテーション練習を対象とし、大規模言語モデルと音声対話エージェントを組み合わせることで、一人でも質疑応答を含む発表練習が行える AI 質疑応答システムを構築した。

検証の結果、システムが予測した想定質問的中率は部分一致も含めれば7割弱となり、本システムが質疑応答練習ツールとしての可能性を持つことが示された。

今後は、より高精度な音声認識モデルの導入や、ユーザの回答意図を補完する仕組みの検討、さらには多様な発表分野・話者での検証を進めることで、より実践的で信頼性の高い質疑応答練習環境の構築を目指す。

謝辞

本研究においては、音声対話のフレームワークとして、名古屋工業大学の研究成果であるMMDAgent-EX[2]を利用しております。

参考文献

- [1] 山崎敦子, 村上嘉代子, 山中司, 長谷川浩志, "理工系英語プレゼンテーション VR 教材の開発と評価方法," 工学教育研究講演会講演論文集, 公益社団法人日本工学教育協会, 2024 年.
- [2] MMDAgent-EX, 名古屋工業大学, <https://mmdagent-ex.dev/ja/>.
- [3] 若林暉, 三輪賢一郎, "育英祭の音声案内のための音声対話システムに関する研究," 大学コンソーシアム八王子 第 17 回学生発表会, 25PC05, Dec. 2025.
- [4] 野村悠翔, 三輪賢一郎, "避難経路案内のための音声対話システムに関する研究," 大学コンソーシアム八王子 第 17 回学生発表会, 25PB06, Dec. 2025.
- [5] 山崎雄介, 三輪賢一郎, "Raspberry Pi を用いた音声対話システムに関する研究," 大学コンソーシアム八王子 第 17 回学生発表会, 25PA04, Dec. 2025.
- [6] 太刀川皆仁, 三輪賢一郎, "MMDAgent-EX を用いた新生サレジオ高専専用案内システムの開発," 大学コンソーシアム八王子 第 17 回学生発表会, 25PC03, Dec. 2025.
- [7] 清野貴一郎, 三輪賢一郎, "音声認識 API Whisper を用いた歴史的音源の文字起こし精度の検討," 大学コンソーシアム八王子 第 17 回学生発表会, 25C221, Dec. 2025.