

大規模言語モデルとロボットを用いたアクティブラーニングにおけるリヴォイシングタイミングの研究

Active Learning Utilizing Robot Speech Generation

大場 璃々^{1*} 松岡 竜輝² 井上 南² 今井 倫太¹
Riri Oba¹ Ryuki Matsuoka² Minami Inoue² Michita Imai¹

¹ 慶應義塾大学理工学部

¹ Faculty of Science and Technology, Keio University

² 慶應義塾大学大学院理工学研究科

² Faculty of Science and Technology, Keio University

Abstract: In the field of education, active learning with robots is becoming increasingly popular as a means to enhance children's proactivity. However, there is a problem that children might not develop their critical thinking skills if they directly ask robots for answers. In this paper, we propose the use of a robot named RevoicingRobot, which vocalizes revoicing sentences generated by large-scale language models during group learning activities. "Revoicing" means listeners respond to speakers by paraphrasing or building upon their statements. The robot determines response tags from the conversation and performs revoicing accordingly. We test whether discussions become more lively when students solve problems with the robot.

1 序論

教育分野では、相手が人間でないことの気軽さから子供たちの積極的な反応を引き出しやすいことや、教員が不得意としている科目の補助として、ロボットを取り入れたアクティブラーニングの取り組みが活発化している。また、リヴォイシング (revoicing) という技法をアクティブラーニングで用いることで、学習者は概念的な理解を形成し、文章理解を深めることが明らかになっている。O'Connor らによれば、リヴォイシングとは、「議論の中で他の参加者によって行われる、口頭もしくは書き言葉での、ある児童の発話の、ある種の再発話」[1] と定義されている。さらに、人同士の会話では、全員が同じ頻度で発言をしている場合よりも、発言の頻度が少なかった人が発言を行う場合の方が、一つの発言に対して発言者以外の人の注意が向きやすくなるという場面がある。その結果、一つの発言から話の展開が変化することがある。

本稿では、リヴォイシングを行うロボットをアクティブラーニングに用いることにより、議論の活発化の支援を目指す。

既存研究では、教師が生徒に対してリヴォイシングを行うと、生徒の理解が向上することが示されている [2]。生徒同士でのリヴォイシングも、表現の具体化に役立つ [3]。しかしながら、人間のリヴォイシングは、リヴォイシングを行う人の理解度に応じた内容になってしまうという問題点がある。特に、単一の正解の存在

する良定義課題では、相対的に知識がある学習者に追従したリヴォイシングになってしまう。また、ロボットのリヴォイシングでは、リヴォイシングの内容を設計者が事前に準備する [4] 必要があり画一的なリヴォイシングになる。そのため、実際の授業でのグループ活動では実用的ではない。

そこで本稿では自然かつ多様性のある発話を出力できる大規模言語モデルを用いて、リヴォイシング内容を生成し、生成内容をロボットが発言することにより学習者が議論展開を支援する、RevoicingRobot を提案する。RevoicingRobot では、カテゴリ判定を行い、カテゴリに適したリヴォイシングを生成する。実験では、リヴォイシングの頻度を変更することで議論の様子が変化するかを調査する。

2 提案手法

RevoicingRobot は、会話に応じたリヴォイシング内容を作成し、発声するロボットである。図 1 は参加者の会話毎にリヴォイシングを行う RevoicingRobot のシステム構成図である。まず、RevoicingRobot はシステム開始時に「この問題どうやって考えていけばいいかな?」と発声する。そして、人間の発話を認識するとテキストに変換する。変換後、カテゴリ判定モジュールでカテゴリ判定を行い、判定カテゴリに応じて発話文生成モジュールまたは単縦発話文生成モジュールでリヴォイシングの文章を生成する。カテゴリ判定モジュールと発話文生成モジュールでは OpenAI の API である gpt-3.5-turbo [5] を使用した。

*連絡先：慶應義塾大学理工学部情報工学科
〒 223-0061 神奈川県横浜市港北区日吉 3 丁目 14 - 1
E-mail: ooba@ailab.ics.keio.ac.jp

カテゴリ判定モジュールでは、大規模言語モデルは、会話の応答として適当と考えた一柳の研究 [6] で用いられたリヴォイシングのカテゴリ (【課題の提示】、【役割指名】、【主張】、【消極的主張】、【応答説明】、【消極的応答】、【反論】、【疑問】、【問題】、【換言】、【解釈】、【確認】、【質問】、【同意】、【相槌】) を判定する。

その後表 1 のように、【課題の提示】、【主張】、【応答説明】、【反論】、【疑問】、【問題】、【換言】、【解釈】、【確認】、【質問】と判定された場合、発話文生成モジュールで処理が行われ、【役割指名】、【消極的主張】、【消極的応答】、【同意】、【相槌】と判定された場合、単純発話文選択モジュールで処理が行われる。発話文生成モジュールでは判定したカテゴリに応じたリヴォイシングを生成する。また、単純発話文選択モジュールでは、事前実験で生成の種類が多くなかったため、処理時間削減の観点から、表 2 のようにカテゴリ毎に限定した応答文章を生成する。

最後に、生成文は自動的に読み上げられ、音声として出力される。

音声認識モジュールが会話を認識できなかった場合は、【消極的主張】の「うーん」を発話し、再度会話を認識する。また、大規模言語モデルが生成するプロンプト用意のため、カテゴリ判定の際はカテゴリ名とカテゴリ毎の定義を、リヴォイシング文章生成の際はカテゴリの定義と例を大規模言語モデルに与えた。

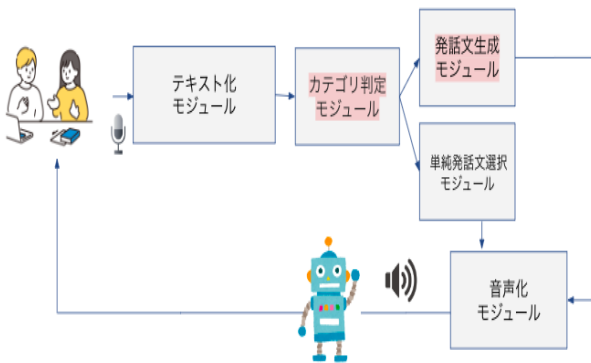


図 1: システム概要

表 1: 一柳の研究で用いられたリヴォイシングカテゴリ [6] とモジュールの対応

モジュール	カテゴリ	
発話文生成モジュール	【課題の提示】	【主張】
	【応答説明】	【反論】
	【疑問】	【問題】
	【換言】	【解釈】
	【確認】	【質問】
	【役割指名】	【相槌】
単純発話文選択モジュール	【消極的主張】	【同意】
	【消極的応答】	

表 2: 単純発話文選択モジュールの応答文章

カテゴリ	生成内容
役割指名	「これ/それ/あれ 書いて」
消極的主張	「うーん」 / 「全然わからない」
消極的応答	「わからない」
同意	「確かに」 / 「そうだね」
相槌	「なるほど」

3 実験

RevoicingRobot が議論活発化の支援ができていないかを調査するために、リヴォイシング頻度を変化させた。

3.1 実験環境

実験には卓上型ロボット Sota[7] を用いた。授業での議論の場を模すために、ロボットは 2 人の参加者の間に参加者側を向いて設置し、1 人 1 枚白紙の紙を用意した。また、問題を表示するためにスクリーンを使用した。まず、実験は毎セット、2 人の参加者とロボットで行い、参加者は議論がしやすいように横並びで着席した。実験環境の様子を図 2 に示す。

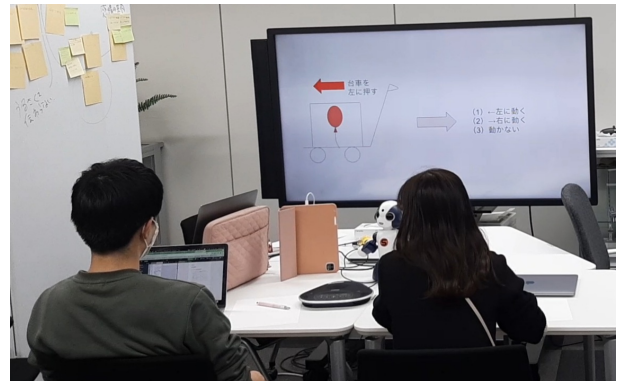


図 2: 実験の様子

3.2 実験条件

問題の議論中に、RevoicingRobot は参加者の発話の頻度に応じてリヴォイシングを行った。「フリーリヴォイシング」条件では、音声認識が発話を認識する度にリヴォイシングを行った。「インターバルリヴォイシング」条件では、音声認識が 3 回発話を認識する毎にリヴォイシングを行った。

3.3 実験参加者

参加者は慶應義塾大学理工学部在籍する、大学 4 年生 8 名が参加し、各試行において 2 人の参加者が参

加した. 実験参加者のうち男性は5名, 女性は3名である.

3.4 実験手順

まず, 参加者が, 答えのみを聞き出し議論を行わないことを防ぐために, ロボットと一緒に学習する仲間と事前に伝えた.

次に画面に議論する図3の「糸巻き問題」もしくは図4の「台車と風船問題」を表示し, 最初3分間, 人間のみで議論を行い, その後7分間でロボットを含めて議論を行った. ロボットを含めた議論を始めると同時にシステムの音声入力を開始した. 人間のみでの議論後とロボットを含めた議論後に, 参加者は解と解の選択理由を記述した. 最後に3人での議論後は, 理由の記述後, 問題の解説動画を視聴した. リヴォイシングは, フリーリヴォイシングとインターバルリヴォイシングの2種類の条件を用い, フリーリヴォイシングで実験を行ったのちに, インターバルリヴォイシングで, 問題を変更した上で同様の実験を行った. 最後にリヴォイシング内容に関するアンケートと谷島らの「学習目標志向測度」[8]のアンケートを行った. 実験の様子はビデオ, システムのログで記録した.

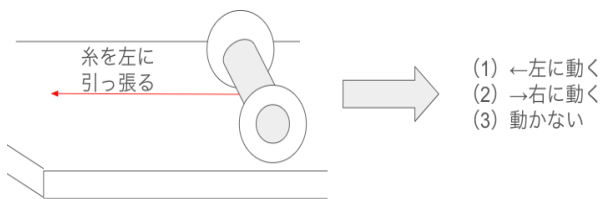


図 3: 糸巻き問題

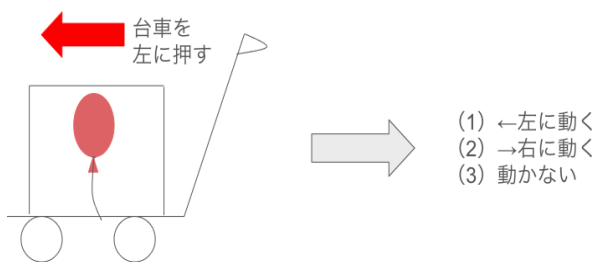


図 4: 台車と風船問題

3.5 評価方法

参加者はアンケート内で議論に集中できたと感じたリヴォイシング条件を選択し, RevoicingRobot の意見を参照した場面とその理由を自由記述形式で回答した. 加えて, 表3の谷島らのアンケートの一部である生徒の学習意欲や学業成績を類推できる4段階のアンケー

ト [8] を実施し, 勉強意欲に応じてリヴォイシングに影響されやすいかを判定した. 1は「全く当てはまらない」, 2は「あまり当てはまらない」, 3は「少し当てはまる」, 4は「とても当てはまる」となっている.

表 3: 谷島らの学習目標志向測度 [8] の一部

項目	質問内容
Q1	どんなに興味のある問題でも, 自分で解くよりも, 他の人に聞いたり, 答えを見たりしたくなる.
Q2	自分でできたと思った問題が間違っていた時は, 恥ずかしくなる.

4 結果

4.1 議論に集中できたと感じたリヴォイシング条件

議論に集中できたと感じたリヴォイシング条件を表4に示す. また, χ^2 検定より $\chi^2(1) = 0.5 < 3.84$ (有意水準5%) で有意差があるとはいえないことが分かった. 一方で, 表4から, フリーリヴォイシングの方が議論に集中できたと回答する人が多かった.

表 4: 議論に集中できたと感じたリヴォイシング条件

条件	人数
フリーリヴォイシング	5人
インターバルリヴォイシング	3人

4.2 RevoicingRobot の意見の参考理由

表5によると, 知識の確認や2人で出した解が正しいかどうかの確認を行う際に RevoicingRobot の意見を参照していることがわかった.

表 5: RevoicingRobot の意見を参考にした場面

理由	人数
知識を確認したい時	2人
ロボットの意見が参加者の意見と同じか異なるかの確認時	4人
ロボットが理解可能かつ論理的な言葉で説明をしていた時	1人
参加者の意見を取り入れて, 考えを改めた時	1人

4.3 学習アンケート

表3の4段階評価のアンケート結果を示す。図5から「全く当てはまらない」以外に回答をした人は8人中5人いたことがわかる。また図6から「少し当てはまる」、「とても当てはまる」と回答をした人は8人中6人いたことがわかる。

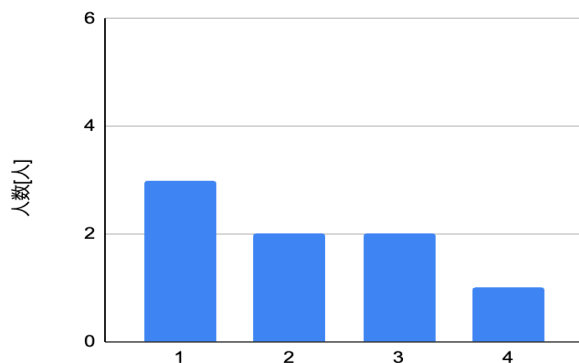


図5: Q1 結果の結果

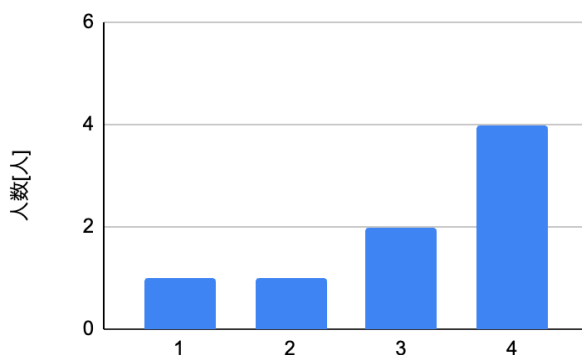


図6: Q2 結果の結果

5 考察

実験結果を元に、RevoicingRobotが生成した発話を元に議論がどのように変化したか考察する。具体的には、ロボットの呟きによって学習者同士の議論が起きたり、解答を変更したりした箇所を抽出して考察した。

5.1 議論に集中できたと感じたリヴォイシング条件

参加者は実験の順番により、フリーリヴォイシング条件でRevoicingRobotは毎回会話に反応するロボットと認識したと考えられる。そのため、インターバルリヴォイシング条件では、例えば、インターバルリヴォイシング条件で話す回数が減少したため、参加者同士がRevoicingRobotが反応していないもしくは機能が

停止したのではないかと考え、会話が止まったりロボットに呼びかけたりした様子が観察された。そして、インターバルリヴォイシング条件では、ロボットが話す間隔を掴めず、RevoicingRobotの発話と参加者の発話が被ってしまうことがあった。それにより、参加者がRevoicingRobotの発言を聞き取ることが出来ず、もう一度発話するように発言していたが、ロボットの発話回数の間隔の関係からRevoicingRobotは回答しなかった。結果として、話さなくなったことに戸惑いを感じた人がいたと考えられ、戸惑いが生じ難かったフリーリヴォイシング条件の方が議論に集中できたと回答した人が多かったと考えられる。また、有意差があるとはいえないという結果になったことについて、条件に応じて差が起きる場合は、実験において条件の順番を固定することと考えられる。しかし、本実験では、回答者8名のうち、フリーリヴォイシング条件、インターバルリヴォイシング条件でより議論に集中できたと感じた条件をそれぞれ5名、3名が回答した。どちらの条件が選ばれるかは1/2の確率なので、回答結果の数字を見ると極端に差が生じていない。そのためもう少し実験者数を増やすべきだと考えられる。

5.2 RevoicingRobotの意見が学習者に与える影響

実験結果から、参加者はRevoicingRobotの課題に関する発言を重視していた。RevoicingRobotを含めた議論が始まる前の3分間の議論で、参加者は感覚的に解答を決めていたり、解答が曖昧な状態だった。そのため、実験がアクティブラーニングという複数人数で議論をし、意見を述べ合っ解決を決める状況ということと、表5、図5、図6から、RevoicingRobotが問題に関する発言をした時は、その意見も解答の根拠の材料にしたということが考えられる。

次に実験参加者は、知識を確認する際にもロボットの意見を取り入れていた。これは、議論の主題ではないところの疑問を解決できる手段としてロボットを活用したと考えられる。また、参加者はRevoicingRobotを知識を有しているものとしてみなしていたためか、図6から参加者が分からないことがあった時に、ペアの相手に直接質問することが恥ずかしいためにロボットを介して解決しようとしたことが考えられる。そのため、ロボットは人間同士の会話の媒介者として利用されていると考えられる。

6 今後の課題

6.1 不良定義課題での検討

本稿では、グループ学習で解が一つに定まる不良定義課題を解いた際にRevoicingRobotがどのように活用されるか検証した。一柳の課題の種類に応じて主に使用されるリヴォイシングのカテゴリが異なるということが

明らかとなっている [6]. 不良定義課題は正解が一義に決まらないため、良定義課題に比べて、参加者は大規模言語モデルの意見に反応しないのではないかと考えられる. そのため今後、不良定義課題で RevoicingRobot の効果を検証する際に、良定義課題では、一つの会話に対する大規模言語モデルの返答内容を複数にならないように設定したが、不良定義課題では正解が複数考えられるため、複数種類のリヴォイシング文を生成することを検討する.

6.2 実験順序の変更

本稿では同じ参加者にリヴォイシングの頻度の条件を変更して実験を行い、フリーリヴォイシング条件、インターバルリヴォイシング条件の順番で実験を行った. そのため、参加者はフリーリヴォイシング条件のロボットの反応を基準にしてインターバルリヴォイシング条件での実験を行ったと考えられる. そのため、初めにインターバルリヴォイシング条件を行ってからフリーリヴォイシング条件を行った場合、異なる結果が得られるのではないかと考えられる. 従って、今後、条件の順序を変化して実験を行うことを検討する.

6.3 RevoicingRobot の動作

本稿では、ロボットが生成文章を発話するのみの設計となっている. しかし、実験を通して、参加者から、ロボットの動きがないことから、ロボットに話掛けるタイミングが掴みにくかったという意見が挙がった. 実際、インターバルリヴォイシング条件の実験中、ロボットが話し始めるタイミングと参加者が話し始めるタイミングが被ってしまう様子が見られた. 従って、今後はロボットが話し始めるタイミングで動作をつけることを検討する. また、本稿ではリヴォイシングという聴覚に関する技法を用いているが、今後はロボットの動作の頻度を変化させるという視覚に関する技法を用いて議論がどのように変化するかを検討する.

7 結論

本稿の目的は、リヴォイシングをするロボットをアクティブラーニングで用いて議論を活発化することである. そこで、大規模言語モデルを用いてリヴォイシング内容を生成するロボット、RevoicingRobot を提案した. RevoicingRobot は既存研究のリヴォイシングのカテゴリを用いて、会話の応答に適切なカテゴリを判定し、カテゴリ内容に応じたリヴォイシングを生成する. それにより、画一的でないリヴォイシングを生成することが可能となる. 実験では、RevoicingRobot を用いると議論の展開があるかを調査するためにリヴォイシング頻度を変更した. 実験の結果から、参加者の会話を認識する毎にリヴォイシングした方が議論に集中できたと感じた人が多いと分かった. また、知識を確認

したい時や3人目の意見を聞きたい時にロボットの意見を参考にしていた. さらに、リヴォイシングの内容を参考にして議論を始めたり、ロボットにもう一度理由を尋ねる様子が見られた.

謝辞

本研究は、JST, CREST, JPMJCR19A1 の支援を受けたものである.

参考文献

- [1] M.C. O' Connor., S. Michaels.: Shifting participant frameworks: orchestrating thinking practices in group discussion, in *Discourse, Learning, and Schooling*, D. Hicks, Ed. Cambridge: Cambridge University Press, pp.63–103(1996)
- [2] 一柳智紀: 教師のリヴォイシングの相違が児童の聴くという行為と学習に与える影響, *教育心理学研究*, Vol.57, No.3, pp.373–384, (2009)
- [3] 菅井篤, 有元典文: 教え合い活動におけるリヴォイシングの効果 —アクティブ・ラーニングの発話分析から—, *横浜国立大学教育学会研究論集*, No.3, pp.13–22, (2016)
- [4] 白水始, 中原淳: 人の主体的な問題解決を促すロボットの役割, *日本ロボット学会誌*, Vol.29, No.10, p.898–901, (2011)
- [5] OpenAI: <https://platform.openai.com/docs/models/gpt-3-5-turbo>, (2024.2.19 閲覧)
- [6] 一柳智紀: 小グループでの問題解決過程における学習者によるリヴォイシングの機能: 課題構造による相違に着目して, *新潟大学教育学部研究紀要. 人文・社会科学編*, Vol.7, pp.37–48, (2014)
- [7] ヴァイストーン株式会社: <https://sota.vstone.co.jp/home/> (2024.2.19 閲覧)
- [8] 谷島弘仁, 新井邦二郎: 学習の目標志向の発達の検討および学業達成との関連, *筑波大学心理学研究*, Vol.16, pp.163–173, (1994)