

# 卒業論文

## 大規模言語モデルに基づく マルチエージェントによる英語学習支援

指導教官 村上 陽平 教授

立命館大学 情報理工学部  
先端社会デザインコース 4回生  
2600210040-0

稲田 匠真

2024年度（秋学期）卒業研究3（CH）  
令和7年1月31日

# 大規模言語モデルに基づくマルチエージェントによる英語学習支援

稲田 匠真

## 内容梗概

近年, ChatGPT の普及により大規模言語モデルを用いた英語学習が行われている. 例えば, 学習者のレベルに合わせた問題の自動作成や, 学習者から送信された英文をあらかじめ指定された方法に則って採点やフィードバックを用いた指導が可能である. 大規模言語モデルでは出力言語を自由に指定することができ, かつ, 人間の英語講師の時間的制約が大きく軽減されるため, 学習環境を整えることが容易である. また, 英語母語話者の講師と学習するよりも, 大規模言語モデルの方が学習者にとって心理的負荷が低くなるという効果がある. しかしながら, 現状の大規模言語モデルによる学習方法は, 学習者の回答に対するフィードバックが主であり, 学習者が他者と協力しながら学習を行うピア・ラーニングは実現できていない. これは, 現状の大規模言語モデルのユーザインタフェースがチャット形式で, ユーザとの文脈が一つに限定されるためである.

そこで本研究では, 複数の文脈を管理できるように, 複数の独立したプロンプトを組み込んだマルチエージェントシステムを用いてピア・ラーニングの実現を行う. 具体的には, 課題を作成しフィードバックを返す教師役のエージェントと, 学習者と協働しながら学習者の回答が何を意味しているのかを推測するピア役のエージェントを用いて, 大規模言語モデルと学習者のピア・ラーニングの確立を行う. 本手法の実現にあたり, 取り組むべき課題は以下の2点である.

## 大規模言語モデルを用いたマルチエージェントアーキテクチャの構築

本実験で使用する web アプリケーションでは実際の教育現場で行われているピア・ラーニングと同様の環境を構築するために教師エージェントとピアエージェントの文脈それぞれを独立して管理しながら学習者の学習支援を行う必要がある.

## ピア・ラーニングのためのプロンプト作成

web アプリ上で実際の教育現場と同様のピア・ラーニングを行うために, 学習者の支援を行う教師エージェントと学習者と協力して課題解決を行うピアエージェントのプロンプトをそれぞれ構築する必要がある.

1つ目の課題に対しては, 学習者が個人でピア・ラーニングを行えるようにするため, 教師エージェントとピアエージェントの文脈をそれぞれ独立させたマルチエージェントアーキテクチャを構築した. 具体的には, 教師エージェントと

ピアエージェントに渡す情報をサーバで制限することで文脈を独立させる。まずサーバ上のリストから教師エージェントが要素を取得し、その情報をもとに問題を作成する。作成した問題を学習者に渡すことで学習者は問題の回答を行う。次に学習者の回答を教師エージェントとピアエージェントに送信する。さらに各エージェントのフィードバックを学習者のみに提供する。このようにすることでピアエージェントには問題の内容や教師エージェントのフィードバックが共有されず文脈を独立させながらピア・ラーニングを行うことが出来る。

2つ目の課題に対しては、英語学習支援のピア・ラーニングを行う上で必要となる教師エージェントとピアエージェントのプロンプトを作成した。まず本実験の学習スタイルはタブーゲームを採用した。タブーゲームとは学習者がタブー単語を使用せずにターゲット単語について英文で説明を行う。その英文をターゲット単語やタブー単語について認識していないピア役の推論者が読み、どの英単語のことについて説明されているか推論する。このとき推論者が示した英単語とターゲット単語が一致すればゲームクリアとなる。本学習スタイルを用いたピア・ラーニングを実現するために、教師エージェントには以下の役割を持たせるプロンプトを作成した。まず、タブー単語の作成を行い、学習者が作成した英文にタブー単語が含まれていないかを確認する。さらに学習者の書いた英文に対してアドバイスを提供することで教師のように振舞う。一方でピアエージェントには学習者が作成した英文を推論することで学習者と協力して学習を進めるように行動するプロンプトを作成した。

提案手法の有用性を示すために、提案手法を組み込んだシステムを実装し、動作確認を行った。本研究の貢献は以下の通りである。

### **大規模言語モデルを用いたマルチエージェントアーキテクチャの構築**

複数の大規模言語モデルを組み込んだマルチエージェントアーキテクチャを構築することによりそれぞれのエージェントが文脈を独立しながら学習者とピア・ラーニングを行うことが出来た。

### **ピア・ラーニングのためのプロンプト作成**

被験者である日本人学生4人に対して本 web アプリケーションを用いて産出語彙の増加数の検証を行ったところ、対人でタブーゲームを行ったときと同様の産出語彙数がみられた。また人と学習を進める方が学習者は積極的に学習を進めていたが、エージェントとの学習では人と比べて心理的負担が少なく取り組むことが出来た。

## Construction of a Multi-Agent Architecture with Large-Scale Language Models

Takuma Inada

### **Abstract**

With the widespread adoption of ChatGPT, English language learning using large-scale language models has become increasingly common. These models can generate questions based on learners' proficiency levels and provide feedback, making learning more flexible and reducing the reliance on human teachers. Additionally, they ease the psychological burden on learners compared to learning with native speakers.

However, current methods focus mainly on feedback rather than peer learning, where learners collaborate. This limitation arises because chat-based interfaces restrict interactions to a single context.

To address this, we propose a multi-agent system that enables peer learning by managing multiple contexts with independent prompts. Specifically, a teacher agent generates tasks and provides feedback, while a peer agent collaborates with learners to interpret their responses, facilitating peer learning with large-scale language models. This approach presents two key challenges.

### **Establishment of a multi-agent architecture using large-scale language models**

The web application used in this experiment needed to support learners while independently managing the contexts of the teacher agent and the peer agent to create an environment similar to peer learning in real educational settings.

### **Creating prompts for peer learning**

To conduct peer learning on a web application similar to real educational settings, it was necessary to create prompts for a teacher agent to support learners and a peer agent to collaborate in problem-solving.

To address the first issue, we developed a multi-agent architecture that ensured context independence between the teacher agent and the peer agent by restricting shared information at the server level. The teacher agent retrieved elements from a server list to create questions, which were then given to the learner. The learner responded, and their answers were sent to both the teacher and peer agents.

However, each agent's feedback was provided only to the learner, preventing the peer agent from accessing the questions or teacher feedback, thus maintaining context independence in peer learning.

For the second issue, we designed prompts for the teacher and peer agents to facilitate peer learning in English language support. This study employs a taboo game, where the learner describes a target word in English without using specific taboo words. A peer predictor, unaware of both the target and taboo words, reads the description and infers the target word. The game is cleared when the guess matches the target word.

To implement peer learning within this framework, the teacher agent was tasked with generating taboo words, checking for their presence in learner responses, and providing feedback. Meanwhile, the peer agent collaborated with the learner by making inferences based on their English sentences.

To demonstrate the usefulness of the proposed method, we implemented a system incorporating the proposed method and verified its operation. The contributions of this study are as follows.

### **Construction of a Multi-Agent Architecture with Large-scale Language Models**

By building a multi-agent architecture incorporating multiple large-scale language models, each agent was able to peer-learn with the learner in a context-independent manner.

### **Creation of prompts for peer learning**

We tested the number of vocabulary words produced by this web application on four Japanese students, and found that the number of vocabulary words produced was similar to the number produced when playing tabu games with other students. The students were able to learn with less psychological burden than when learning with a person, but they were more active when learning with a person.

# 目次

<b>第 1 章 はじめに</b>	<b>8</b>
<b>第 2 章 関連研究</b>	<b>10</b>
2.1 ChatGPT による学術英語語彙の自律的学習	10
2.2 ピア・ラーニング	10
2.3 大規模言語モデルを組み込んだマルチエージェントアーキテクチャ	11
<b>第 3 章 大規模言語モデルに基づくマルチエージェントアーキテクチャ</b>	<b>13</b>
3.1 マルチエージェントアーキテクチャの構築	13
<b>第 4 章 ピア・ラーニングを行うための大規模言語モデルのプロンプト</b>	<b>15</b>
4.1 タブーゲーム	15
4.2 教師エージェント	15
4.2.1 問題作成タスク用プロンプト	15
4.2.2 チェックタスク用プロンプト	16
4.2.3 アドバイスタスク用プロンプト	16
4.2.4 アンサータスク用プロンプト	17
4.3 ピアエージェント	17
4.3.1 推論タスク用プロンプト	18
<b>第 5 章 実験</b>	<b>19</b>
5.1 被験者	19
5.2 実験手順	19
5.2.1 人・タブーゲーム	20
5.2.2 エージェント・タブーゲーム	20
5.2.3 リテリングテスト	20
5.3 評価方法	21
5.3.1 タブーゲームの好意性	21
5.3.2 リテリングテストの好意性	22
5.3.3 産出語彙	23
<b>第 6 章 実験結果</b>	<b>24</b>
6.1 タブーゲームの好意性の評価	24
6.1.1 タブーゲームの好意性のアンケート	24
6.1.2 タブーゲームのインタビュー	27
6.2 リテリングテストの好意性の評価	27
6.3 学習者の産出語彙の評価	30

<b>第7章 考察</b>	<b>32</b>
7.1 英単語の産出語彙増加の有用性.....	32
7.2 人とエージェントでタブーゲームを行った場合の好意性の比較.....	32
7.3 人とエージェントでタブーゲームを行った場合のリテリングテストへの好意性の比較 .....	33
<b>第8章 おわりに</b>	<b>34</b>
<b>謝辞</b>	<b>35</b>
<b>参考文献</b>	<b>36</b>
<b>付録</b>	<b>37</b>
プロンプトの実例 .....	37
問題作成タスク用プロンプト .....	37
チェックタスク用プロンプト .....	38
アドバイスタスク用プロンプト.....	40
アンサータスク用プロンプト .....	42
推論タスク用プロンプト .....	43

## 第1章 はじめに

近年 ChatGPT の普及により大規模言語モデルを用いた英語学習が行われている。英語学習における大規模言語モデルの主な利用例としては翻訳や英文の自動生成などがある。大規模言語モデルでは出力言語や採点・フィードバック方法を自由に指定することができ、かつ、人間の英語講師の時間的制約が大きく軽減されるため、学習環境を整えることが容易である。また、英語講師と学習するよりも、大規模言語モデルの方が学習者にとって心理的負荷が低くなるという効果がある。このような点から大規模言語モデルを用いた英語学習の支援が実際の教育現場でも活用されている。関連研究ではプロンプトに役割や英語学習における特性を説明し、学習者のレベルに合わせた問題を自動で作成させ、学習者の英文を指定された方法に則って採点・フィードバックさせることを与えることにより、生徒が大規模言語モデルと対話しながら自律的学習を行う指導がされている。

しかし、1つの大規模言語モデルを用いて学習者の回答に対してフィードバックを作成する学習方法だと、文脈が1つしかないため他者と協力しながら学習するピア・ラーニングを行うことができないという課題が生まれてしまう。

そこで本研究では、複数の大規模言語モデルを用いてピア・ラーニングの実現を行う。具体的には、大規模言語モデルにプロンプトを組み込んで人間のように自立して動くエージェントを複数もちいて学習者の英語学習支援を行う。課題を作成しフィードバックを返す教師役のエージェントと、学習者と協働しながら学習者の回答が何を意味しているのかを推測するピア役のエージェントを用いて、エージェントと学習者のピア・ラーニングの確立を行う。本手法の実現にあたり、取り組むべき課題は以下の2点である。

### 大規模言語モデルを用いたマルチエージェントアーキテクチャの構築

本実験では複数のエージェントと会話をしながら学習をするために1つの web アプリケーションに3つのエージェントを組み込む。学習者が自身の回答のフィードバックを受けながら他のエージェントとピア・ラーニングできるようにするためには教師役のエージェントとピア役のエージェントが連携するアーキテクチャの構築が必要となる。

### ピア・ラーニングのためのプロンプト作成

本実験で使用する3つの大規模言語モデルエージェントのプロンプトを作成する。本実験で使用する学習方法はタブーゲームである。ターゲットの英単語と使用してはいけないタブーとなる英単語が複数提示され、それらを使用せずに学習者は英単語を説明する英文を作成する学習スタイルである。このタブーゲームにおいて教師役・ピア役のエージェントそれぞれのプロンプトを作成し、ピ

ア・ラーニングの確立を行う必要がある。

以下本論文では、第2章では大規模言語モデルを用いた英語学習支援の関連研究とピア・ラーニングの説明をしている。第3章ではアーキテクチャの構築について説明する。続いて、第4章ではそれぞれのエージェントのプロンプト作成について述べる。第5章では提案手法の評価を行い、第6章では提案手法に対する考察を行う。第7章では本稿をまとめ、今後の展望やさらなる課題について結論とする。

## 第2章 関連研究

本章では、大規模言語モデルがどのようにして英語学習支援を行われているかを関連研究の紹介を交えながら説明していく。また、本研究で検証するピア・ラーニングについても説明していく。

### 2.1 ChatGPT による学術英語語彙の自律的学習

柳瀬[1]は、ChatGPT を活用した語彙学習の可能性を探求し、大規模言語モデルを用いた学術英語語彙の自律的学習を促進する指導実践を提案した。この研究では、著者の言語観を反映させたプロンプトを設計した。学習者が選択した単語に関する定義・連語・例文・類義語を ChatGPT が提供し、それを基に例文を作成・修正する反復的な学習方法を採用した。京都大学の1年生クラス30人を4クラス対象に実施された実験では、実験参加者の92%がこの学習により英語力を高めたと回答している。さらに学習者が自らのペースで学習を進める中で語彙の深い理解やニュアンスの違いに気づき、学習意欲が向上するなど、肯定的な効果が確認された。一方で、負荷の高い学習方法が一部の学生には適応しづらいという課題も示唆されており、学習者の多様なニーズに応じた柔軟な指導方法の必要性が指摘されている。この研究は著者の言語観と学習方法をプロンプトに組み込んだ大規模言語モデルを用いて学習者の学術英語語彙の自律的学習方法を提案しており、参加者の92%が学習を通じて英語力を高めることが出来たと回答している。この研究から大規模言語モデルを用いた英語学習は学習者の自律的学習を支援し、さらに学習意欲向上を促進することが可能であることが分かった。

### 2.2 ピア・ラーニング

ピア・ラーニングとは、同じような立場の仲間とともに支え合いながら、ともに関わり合いを持ちながら、知識・スキルを身に付けていくことである。そこで教師やコーチなどの上下関係に基づく指導方法の特質とは異なり相互の関係所いう側面においてともに学ぶピアが持ちうる特質としては「互恵性」「対等性」「自発性」の3つ挙げられる[2]。

まず「互恵性」とは指導者が隔週者に対して一方的に恵を与えるのではなく、それぞれのピアがお互いに学び手にも教え手にもなり、恵みをもたらしながら学習が深まっていくという特質を持っている。

次に「対等性」とはピアが同じような立場の仲間だということである。同じ

立場にあることで意見を表明しやすく、相互理解が深まることを指す。また多様な能力や個性を持つピア同士が関わることで、互いの力を引き出し合う関係が生まれる。

さらに「自発性」とは自ら考えを出し合い、協力して問題を解決する学びを指す。ピアそれぞれが主体となり、関わり合いながら独自の方法で課題を達成しようとするのである。

以上のことは教師には持ちにくい特性として考えられ、教師と生徒という立関係において実現しにくい側面であるといえる。「類似性」や「対等性」をもとにピア同士が親密な信頼関係を築き、「自発的」で「互惠的」な形で学び合いを深めていく活動は、自律的に学び成長を遂げていく学習者を育てていく上で大きな力を持つものといえる。

## 2.3 大規模言語モデルを組み込んだマルチエージェントアーキテク

### チャ

加藤ら[3]は、仮想空間上のエージェントに LLM を組み込むことで人間らしい行動・計画を実行させることで、現実世界に近い状況でシミュレーションを行った。具体的には LLM との連携により意思決定が可能な形式で歩行者エージェントを組み込んだ MAS 環境を構築し、その上で歩行者エージェントが実際の人間らしい意思決定を表現できるか検証を行った。この研究では、エージェントに基づいた汎用のモデリング・シミュレーション基盤である GAMA を用いてシミュレーション環境 (CDL Street) を構築し、GAMA と LLM を連携させるために Flask を用いた Web アプリケーションを開発した。GAMA はネットワーク機能を介して外部の Python プログラムと連携し、サーバを通じてプロンプト作成に必要なデータを送信する。Python 側では、GAMA から受け取ったデータを基にプロンプトを作成し、OpenAI の API を用いて LLM により意思決定のためのテキストを生成する。生成された思考は GAMA に返され、エージェントの行動計画に反映される。

この研究の検証により、エージェントが自身の性格や経験などの記憶に基づき、環境情報を考慮した意思決定を行う様子が確認でき、人間らしい思考が表現できることが示された。また、エージェントの属性情報の変更が集団的な行動特性に及ぼす影響を分析した結果、行列を避ける性格のエージェントが多い場合には、待ち行列の人数が減少することが確認された。これは、従来のエージェントのパラメータでは扱いが困難であった人間の個性という抽象的な情報を、自然言語で表現し、その情報に基づいた意思決定を LLM により実現することで、

シミュレーション結果に反映できることを示している。

さらに、シミュレーション終了後のエージェントへのアンケート調査を実施し、エージェントが自身の経験に基づいて環境改善の提案を行うことが可能であることを確認した。これにより、従来の MAS では得られなかった新たな指標を提供し、より良い施策や空間デザインの検討に活用できる可能性を示した。以上の結果から、加藤らの研究で構築した MAS 環境は、エージェントの個性や経験を考慮した意思決定を可能にし、シミュレーションの説明性を向上させるとともに、環境改善の示唆を得るための有用な手法となり得ることが示された。

## 第3章 大規模言語モデルに基づくマルチエージェント トアーキテクチャ

本章では、2つのエージェントを1つのwebアプリケーションケーションに組み込むマルチエージェントアーキテクチャの構築について説明する。

### 3.1 マルチエージェントアーキテクチャの構築

2つのエージェントをwebアプリケーションに組み込み、それぞれのエージェントの文脈を独立させながらピア・ラーニングを実装するアーキテクチャの構築について説明する(図1)。

まず、あらかじめサーバ内に用意された単語リストの中から1つ単語を選択して教師エージェントに渡す。その単語をもとに教師エージェントが問題の作成を行う。そして選択された単語と教師エージェントが作成した問題を学習者に提示する。

次に、学習者が提示された単語と問題に対して回答を作成する。そして問題を作成した教師エージェントとピアエージェントのそれぞれに学習者の作成した回答を渡す。それに対して教師エージェントとピアエージェントはそれぞれのフィードバックを学習者に送信することで学習を進める。このようにアーキテクチャを構築することで、教師エージェントは問題の作成と学習者の回答をもとにフィードバックの作成を行い、またピアエージェントには学習者の作成した回答のみを提示することで、教師エージェントとサーバの単語リストの情報を一切開示せずに学習者のみとインタラクションを進行させる。このようにして教師エージェントとピアエージェントのインタラクションを独立させるマルチエージェントアーキテクチャを構築することで、実際の教育現場で行うピア・ラーニングと同様に学習者の支援を行う教師と学習者に協力する協働者の交流を確立する。

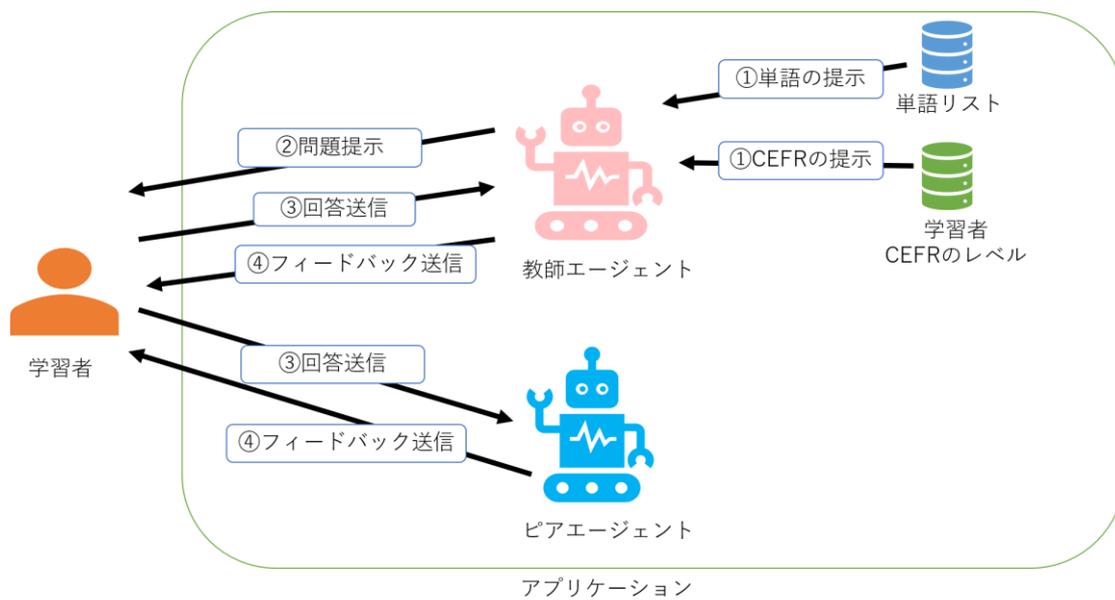


図 1. マルチエージェントアーキテクチャ図

## 第4章 ピア・ラーニングを行うための

### 大規模言語モデルのプロンプト

本章では学習者が本 web アプリケーションを使用してピア・ラーニングを行うためのプロンプト作成について説明する。また本実験で使用した ChatGPT のモデルはすべて GPT-4o-mini である。

本学習ではタブーゲームにより英語学習支援を行う。また全てのプロンプトには本学習スタイルであるタブーゲームのルールを組み込んでいる。実際に組み込んだプロンプトは付録を参照されたい。

#### 4.1 タブーゲーム

本実験では学習者の産出語彙の変化を測定するためにタブーゲームを採用した。タブーゲームとは、学習者がタブー単語とターゲット単語を用いずに、ターゲット単語を英文で説明する学習スタイルである。ターゲット単語とタブー単語を知らない第三者が英文を読んで推測し、推測結果とターゲット単語が一致すればクリアとなる。学習者が使用できる単語に制限をかけることで、ターゲット単語を説明するための単語を想起させることで産出語彙の増加を狙う。また問題の提供と回答の支援を行う教師エージェントと学習者の書いた英文を推論するピアエージェントを作成することでピア・ラーニングの確立を行う。

#### 4.2 教師エージェント

教師エージェントは学習者にタブー単語を提供する。また学習者の書いた英文に対してアドバイスを提供することにより実際の教育指導で行うピア・ラーニングにおける教師のような行動を行う。

##### 4.2.1 問題作成タスク用プロンプト

問題作成タスク用プロンプトは図 2 のように学習者の CEFR のレベルに対応したターゲット単語に関連する英単語を 3 つタブー単語として生成する役割を組み込んでいる。このときタブーゲームで使用するターゲット単語はサーバ上に用意されている英単語を使用している。また、出力単語をより明確にするためにターゲット単語を 1 つ例として取り上げ、それに関連するレベル別のタブー単語を提供する。さらに実際の CEFR レベルとターゲット単語が入力された場合と出力する場合の例を与える。

- ・ タブーゲームのルール説明
- ・ 役割 | CEFR のレベルに応じてタブー単語を 3 つ出力する
- ・ CEFR のレベル別に出力するタブー単語の例
- ・ 実際の入力内容と出力内容の例

図 2. 問題作成タスク用プロンプトの構造

#### 4.2.2 チェックタスク用プロンプト

チェックタスク用プロンプト図 3 のように学習者の英文にタブー単語が含まれていないかを確認する役割となっている。このとき、タブー単語をチェックするようにプロンプトを組み込むと複数形・活用形や品詞転換および群前置詞をタブーとして認識しない恐れがある。そのため、それらをエージェントにタブーとして認識させるために確認単語の定義を行う。

- ・ タブーゲームのルール説明
- ・ 役割 | 学習者が書いた英文に確認単語が含まれていないか確認
- ・ 確認単語の定義 | タブー単語及びそれらの活用形・群前置詞・品詞転換
- ・ 出力例 | ・ 英文にタブー単語が含まれていた場合の出力例  
・ 英文にタブー単語が含まれていない場合の出力例

図 3. チェックタスク用プロンプトの構造

#### 4.2.3 アドバイスタスク用プロンプト

図 4 のようにアドバイスタスク用プロンプトは学習者の英文にアドバイスを提供する役割を組み込んでいる。アドバイスは主に 3 つである。

1 つ目は、文法エラーの修正である。英文に文法エラーが含まれている場合は該当箇所の指摘と文法の修正をした英文の提供を行う。

2 つ目は、語彙の指導である。語彙のスペルミスの修正や英文に曖昧な表現があった場合は代案となる単語の提供を行う。

3 つ目は、タブーゲームのアシストである。まず、学習者の英文を読んでどの英単語について述べられているのか推測を行う。次に、その出力内容とターゲット単語を見比べてターゲット単語を説明するための英文に必要な単語や語彙を提供する。最後に、その単語を用いたフレーズや短い節を学習者に提供することで、英文の修正の支援を行う。またタブー単語やタブー単語の活用形などをアドバイスで使用しないために avoid 単語の定義も行う。

- ・ タブーゲームのルール説明
- ・ 役割 | 学習者の書いた英文がターゲット単語について説明できるように  
アドバイスを提供する
- ・ avoid 単語定義 | タブー単語の活用形・群前置詞・品詞転換を  
アドバイスで避けるべき単語を定義
- ・ 文法指導 | 英文に文法ミスがあった場合の指導
- ・ 語彙指導 | 英文に誤った語彙がないか確認及び指導
- ・ アシスト | ①学習者の書いた英文がどの英単語を示すか推論  
②①で説明に不足している単語を提示  
③②の単語を使って短いフレーズをヒントとして提示
- ・ 例の提示 | 問題例と間違えた英文に対するアドバイスの出力例

図 4. アドバイスタスク用プロンプトの構造

#### 4.2.4 アンサータスク用プロンプト

図 5 のようにアンサータスク用プロンプトには学習者の回答が 4 回目の場合、アンサーエージェントがタブー単語を使わずにターゲット単語を説明する英文を正解例として提供する役割を組み込んでいる。このとき (4.2.2) と同様にタブー単語の活用形などを文章に含まないために avoid 単語の定義を行う。

- ・ タブーゲームのルール説明
- ・ 役割 | タブー単語を使わずにターゲット単語について説明する英文作成
- ・ avoid 単語定義 | タブー単語及びそれらの活用形・群前置詞・品詞転換
- ・ 例の提示 | 問題例に対する正解例となる出力例

図 5. アンサータスク用プロンプトの構造

### 4.3 ピアエージェント

ピアエージェントは学習者の書いた英文に対して推論を行う。推論した英単語を学習者に提供する。また教師エージェントとは異なりターゲット単語・タブー単語を提供しないことで実際の教育指導で行うピア・ラーニングにおける生徒のように行動する。

### 4.3.1 推論タスク用プロンプト

推論タスク用プロンプトには，学習者の英文がどの英単語について述べられているかを推測する．そして推測した英単語を学習者に提供する役割を組み込んでいる．ピアエージェントが日本語を用いて推論することを防ぐために本プロンプトはすべて英文で記す．

- ・ 役割 | 送られてきた英文がどの英単語を述べているか推測

図 6. 推論タスク用プロンプトの構造

## 第5章 実験

本章ではピア・ラーニングにより学習者の産出語彙増加検出の評価を行う。

### 5.1 被験者

実験に協力していただいたのは、日本人母語話者の学生が 4 名である。また全員の CEFR は B1 となっている。CEFR とは学習者、教授する者及び評価者が、外国語の熟達度を同一の基準で判断しながら、学び・教え・評価できるように開発されたものである。C2 が最も高いレベルに位置付けられている。今回は図 7 のように学習者の TOEIC の点数を参考にして CEFR のレベルを決定した<sup>1</sup>。

	CEFRレベル*		TOEIC L&R	
			Listening	Reading
難しい (上級) ↑ やさしい (初級)	Proficient User	C1	490～	455～
	Independent User	B2	400～	385～
		B1	275～	275～
	Basic User	A2	110～	115～
		A1	60～	60～

図 7. TOEIC スコアと CEFR のレベル相関関係

### 5.2 実験手順

本実験の手順は以下のとおりに進めていく。

- ① リテリングテスト 1 回目
- ② 人・エージェント
- ③ リテリングテスト 2 回目
- ④ エージェント・タブーゲーム
- ⑤ リテリングテスト 3 回目

<sup>1</sup> TOEIC 公式 Web サイト「TOEIC Program 各テストスコアと CEFR との対照表」より

評価結果が人タブーゲームとエージェントタブーゲームの順序に依存しないようにするために上記の手順で行う実験グループを A として、②と④の順序が入れ替わって実験を進めるグループを B とする。

またリテリングテストの好意性とタブーゲームの好意性を調べるために毎テスト・毎ゲーム後にアンケートを行う。

### 5.2.1 人・タブーゲーム

学習者・推論者・アドバイザーが対面でタブーゲームを行う。ゲーム時間は 10 分間であり、学習者は 1 つのターゲット単語につき 4 回の回答権がある。また、ターゲット単語とタブー単語はあらかじめ学習者の CEFR のレベルを参考にして用意されたものを出题する。手順は以下の流れに沿って進めていく。

まず学習者がターゲット単語を説明する英文を PC にタイピングして作成する。

次に推論者が英文を読んで英単語の推測を行う。ここで推論結果とターゲット単語が一致するか 4 回目の回答を終えると次の問題に進むことが出来る。推論者の結果と一致しなければ推論者が PC をアドバイザーに渡す。アドバイザーは英文を読んでそれに対するアドバイスを別ウィンドウに記入する。

そして学習者はアドバイスを見ながら英文の作成に取り組む。これを繰り返すことでタブーゲームを進めていく。このときルールとして、学習者が推論者に見せる画面は自身の書いた英文のみである。また推論者に学習者の英文以外の情報を与えないためにアドバイザーと学習者は口頭でのやり取りを禁止する。

### 5.2.2 エージェント・タブーゲーム

学習者がエージェントを組み込んだ web アプリケーションを用いてタブーゲームを行う。ここでもターゲット単語は人・タブーゲームと同じ難易度の単語を用意している。またタブー単語は学習者がログイン時に自身の CEFR レベルを入力することでそれに応じた単語を問題作成エージェントが出力する。こちらもゲーム時間は 10 分間とする。

### 5.2.3 リテリングテスト

本実験ではそれぞれのタブーゲームを行うことで学習者の産出語彙がどの程度増加したかを測定するためにリテリングテストを行う。またタブーゲームを行うことでリテリングテストの好意性が変化するかをアンケートで測る。

リテリングテストは、「英語熟達度によるリテリングパフォーマンスの違いの複雑さ、正確さ、流暢さとの関係と発話内容の分析」を参考にして行う[4]。先行研究で行われたリテリングテストの手順は以下のようになっている。

まず、被験者に 2 分間英文の黙読を行う。次に文章の内容を確認するための問題を 3 問出題し、文章を参考にしながら解き進める。さいごに被験者が英文を

伏せて 2 分 30 秒で文章の内容の説明と新しい自分の意見をスピーキングする。また被験者は大学生 25 名が集められており、TOEIC の平均スコアは 450 点となっている。被験者の英語熟練度はリテリングで使用した英文は英検 2 級対策問題集に載っている 60 語程度の文章を使用している。結果として、英語熟練度が高い生徒程発話回数が多かったが被験者全員が問題に正解しており、さらに全員が時間内にリテリングを完了していた。

そこで本実験で行うリテリングテストはこの方法を参考にして進めていく。まず被験者には、英検 2 級の対策問題集に載っている 90 語程度の英文を 2 分間黙読してもらい、次に文章の内容を確認するための問題を 3 問出題し、文章を参考にしながら解き進める。さいごに文章を伏せて、文章に対する自分の意見をタイピングによりライティングしてもらい、出題する文章の単語数およびリテリング内容が先行研究と異なる理由は、先行研究において「直前に読んだ英文を英語で説明する」という課題が与えられていた点にある。この課題では、読んだ英文を基に話す内容を組み立てることが可能であり、また記憶力に頼って覚えている文やフレーズをそのまま使用することもできると考察されている。本研究では、被験者に英文の説明を求める課題を課さないことで、英文の記憶に依存しないように配慮した。これにより、被験者が提示された英文に基づいて記憶力ではなく創造力を働かせ、新しい英文を作成するプロセスを促進することを目的としている。さらに、この方法を通じて、タブーゲームを活用した産出語彙の増減をより正確に測定できると考えた。

## 5.3 評価方法

### 5.3.1 タブーゲームの好意性

学習者のタブーゲームの取り組みやすさを測るためアンケートとそれぞれの会話記録を用いて行う。また本結果を用いて人とエージェントどちらの方が取り組みやすいと感じたのか比較検証を行う。

まずアンケート項目は、「ピア・ラーニングに対する学習者の認識と学びのプロセス」[5]を参考にして作成した。先行研究では、日本語授業に参加している留学生 20 名を対象に、ピア・ラーニングを取り入れた授業を実施した。留学生はテーマに対する意見を日本語で作文し、各作文に対して 3 人グループで交互にピアレビューを行った。この授業スタイルは、同じテーマで 5 週間にわたり継続された。全ての授業が終わったあとに全体の振り返りアンケートを行ったところ全員がピア・ラーニングに対して効果があると評価をしており、その理由として異なる視点から見てもらえるので、推敲に役立つこととグループワークに対する自分自身やピアの協力の方法について考えるようになったと回答した。

このことから学習者が楽しく積極的にタブーゲームに取り組めたか、相手に伝わりやすい文章を作成しようと心がけたかをアンケートで尋ねたいと考えた。以下が実際のアンケート項目である。

- ① タブーゲームを楽しみながら行ったか
- ② 積極的に取り組もうと思ったか
- ③ ターゲット単語を説明するときに間違いを恐れずに文章を書こうと思ったか
- ④ ターゲット単語を説明するときに相手にわかりやすい文章を書こうと意識したか

また、アンケート項目数は中立的な意見を避け、意見のニュアンスを細かく反映させるために、リッカート尺度の6件法を用いて作成した。

次にタブーゲームの記録について説明する。人と行うタブーゲームは学習者に文章をタイピングしてもらい、推論者の推論結果も紙面に記録として残す。またエージェントと行うタブーゲームは、問題内容・学習者の英文また教師エージェント・ピアエージェントの出力内容を逐次データベース上に保管していく。

アンケート結果とタブーゲームのやり取りを比較することで、人とエージェントで学習の取り組みやすさに違いが出るかを検証する。

### 5.3.2 リテリングテストの好意性

タブーゲームを経ることでリテリングテストの取り組みやすさをアンケートで評価する。また本結果を用いて人とエージェントどちらのタブーゲームを経ることで取り組みやすさに違いが生じるか比較検証を行う。

何も行わずに始めたリテリングテストと、人もしくはエージェントとタブーゲームを行った後に取り組むリテリングテストにおいて、学習者の取り組む姿勢を尋ねるアンケートを実施する。以下にアンケート項目を示す。こちらもアンケート項目数は、中立的な意見を避け、意見のニュアンスを細かく反映させるために、リッカート尺度の6件法を用いて作成した。

- ① 今回のリテリングテストは楽しかったですか？
- ② リテリングテストに自分から積極的に取り組みましたか？
- ③ 今回のリテリングテストは取り組みやすかったですか？
- ④ 内容に自信がなくても文章を書き進めようと思いましたか？

このアンケートでは、タブーゲームを通じてリテリングテストへの姿勢の変化を測定する。さらに、人とエージェントそれぞれと行うタブーゲームを経た際に、

これらの変化に差が生じたかを比較する。

### 5.3.3 産出語彙

タブーゲームを実施することで産出語彙の評価をリテリングテストによって行う。リテリングテストを用いた産出語彙の評価方法は、先行研究[4]において被験者のリテリングの語彙数を用いて複雑さと正確さの評価が行われていたことを参考にして、学習者のリテリングテストの内容から産出語彙の評価を表1に則って行う。また、学習者の主観で産出語彙が増えたと感じるのかを調査するために5.3.2のアンケート項目に加えて、リングテスト後に「今回のリテリングテスト中に、文章を組み立てるときに使う単語やフレーズが増えたと感じましたか？」というアンケートをリッカート尺度の6件法で尋ねる。

表1. リテリングの評価観点と指標

観点	ラベル	指標
複雑さ	C1	Token の数 (全発話語数)
	C2	type の数 (発話語のうち, 異なる語の総数)
	C3	$C2 \div T\text{-unit}$ の数 ( T-unit あたりの語彙数)
正確さ	*A1	全誤り数
	*A2	$\text{全誤り数} \div T\text{-unit}$ (T-unit あたりの誤り数)

注. 左上の\*印は, スコアが大きいほど評価が低い指標.

## 第6章 実験結果

本章では人とエージェントでタブーゲームを行い、その後リテリングテストをすることで産出語彙が増加とタブーゲームの好意性・リテリングテストの好意性を評価する。

### 6.1 タブーゲームの好意性の評価

タブーゲームにおいて、学習者の好意的な反応が人間とエージェントのピア・ラーニングにおいて見られるかを評価していく。本評価はタブーゲーム後に実施したアンケートの結果とタブーゲームへの回答回数を記録したデータを基に分析を行う。またアンケートの回答理由について被験者にインタビューを行った。

#### 6.1.1 タブーゲームの好意性のアンケート

表 2 に記載されているグループ A のアンケート結果と表 3 のタブーゲームの回答回数の評価を行う。

まず学習者のゲームに対する満足度を評価する。アンケート項目の「タブーゲームは楽しかったか?」「積極的に取り組もうと思ったか?」に対しては、人タブーゲームの方が肯定的な意見が 2 つ入っている一方で、エージェントで行ったタブーゲームでは否定的な意見と肯定的な意見が 1 つずつ入った。次に「間違いを恐れずに英文を作成できたか?」という項目では、人タブーゲームとエージェントタブーゲームでどちらも肯定的な意見が見られたが、実際の回答回数を比較するとどちらの被験者もエージェントタブーゲームの回答回数の方が人タブーゲームよりも 2 倍程度多かった。このことから人よりもエージェントの方がグループ A の学習者はタブーゲームに取り組みやすかったといえる。最後に学習者が受け手を考慮した文章作成の意識を評価する。「英文を作成するときに相手にわかりやすい文章を書こうと意識したか」という項目では、推論者が人の場合は両者ともに「意識した」と回答していたが、エージェントの場合だと「やや意識した」「意識していない」と受け手への考慮が人の場合に比べて弱まっていた。このことから学習者は受け手が人間である場合の方が、相手を意識した文章作成を心がける傾向があるとわかった。

表 2. グループ A のタブーゲーム後アンケート

質問項目	選択肢	人(人)	Agent(人)
今回のタブーゲームは 楽しかったですか？	とても楽しかった	1	0
	楽しかった	1	1
	やや楽しかった	0	0
	あまり楽しくなかった	0	1
	楽しくなかった	0	0
	全く楽しくなかった	0	0
タブーゲームに積極的に 取り組もうと 思いましたか？	非常にそう思う	1	1
	そう思う	0	0
	ややそう思う	1	0
	あまりそう思わない	0	1
	そう思わない	0	0
間違いを恐れずに 英文を作成 できましたか？	非常にそう思う	1	0
	そう思う	0	2
	ややそう思う	1	0
	あまりそう思わない	0	0
	そう思わない	0	0
英文を作成するとき 相手にわかりやすい 文章を書こうと 意識しましたか？	強く意識した	0	0
	意識した	2	0
	やや意識した	0	1
	あまり意識していない	0	0
	意識していない	0	1
	全く意識していない	0	0

表 3. グループ A タブーゲームの回答回数

ゲーム方式	被験者①(回)	被験者②(回)
人	6	5
エージェント	9	10

表 4 と表 5 に記載されているグループ B のアンケート結果とタブーゲームの回答回数の評価を行う。まず学習者のゲームに対する満足度を評価する。アンケート項目の「タブーゲームは楽しかったか?」「積極的に取り組もうと思ったか?」に対しては、人およびエージェントどちらの項目にも「楽しかった」や「非常にそう思う」「そう思う」といった肯定的な意見が入っている。次に「間違いを恐れずに英文を作成できたか?」という項目では、人とエージェントの場合で

も「非常にそう思う」「そう思う」といった肯定的な意見がどちらにも見られた。しかし実際の回答回数を比較すると、被験者④はややエージェントタブーゲームの方がやや多かったのに対して、被験者③はエージェントタブーゲームの方が人タブーゲームと比べて 9 回多く回答していた。このことから学習者の主観では取り組みやすさに違いは見られなかったが、実際にエージェントの方が人よりも気楽に回答できたことがわかる。最後に学習者が受け手を考慮した文章作成の意識を評価する。「英文を作成するときに相手にわかりやすい文章を書こうと意識したか」という項目では、推論者が人の場合は両者ともに「強く意識した」と回答していたが、エージェントの場合だと「強く意識した」「やや意識した」と受け手への考慮が人の場合に比べてやや弱まる被験者がいた。このことからグループ B の学習者はエージェントの場合でも相手を意識した文章を作成するように心掛けるが、受け手が人間である場合の方がより相手を意識してゲームに取り組んでいるとわかった。

表 4. グループ B のタブーゲーム後アンケート

質問項目	選択肢	Agent(人)	人(人)
今回のタブーゲームは楽しかったですか？	とても楽しかった	0	0
	楽しかった	2	2
	やや楽しかった	0	0
	あまり楽しくなかった	0	0
	楽しくなかった	0	0
	全く楽しくなかった	0	0
タブーゲームに積極的に取り組もうと思いましたか？	非常にそう思う	1	0
	そう思う	0	1
	ややそう思う	1	1
	あまりそう思わない	0	0
	全くそう思わない	0	0
間違いを恐れずに英文を作成できましたか？	非常にそう思う	1	1
	そう思う	1	1
	ややそう思う	0	0
	あまりそう思わない	0	0
	全くそう思わない	0	0
英文を作成するとき相手にわかりやすい文章を書こうと意識しましたか？	強く意識した	1	2
	意識した	0	0
	やや意識した	1	0
	あまり意識していない	0	0
	全く意識していない	0	0

表 5. グループ B タブーゲームの回答回数

ゲーム方式	被験者③(回)	被験者④(回)
人	1	4
エージェント	10	6

### 6.1.2 タブーゲームのインタビュー

人タブーゲームとエージェントタブーゲームのアンケート回答への理由に対するインタビューを行った。

まず、グループ A・B とともに人タブーゲームがエージェントタブーゲームに比べて楽しく・積極的に取り組めたという意見が見られた理由についてインタビューを行った。インタビューでは、人タブーゲームでは「推論者と一緒にタブーゲームに取り組んでいる」や「相手の表情を見ながら取り組めて親近感がわいた」といった意見が得られた。一方でエージェントタブーゲームでは「web アプリケーションだと推論者が無機質であったため淡々とタスクを進めていた」「可もなく不可もなくタブーゲームを進めた」といった意見が見られた。

次に、グループ A・B とともに人タブーゲームとエージェントタブーゲームで回答回数に違いが出た理由については、人タブーゲームだと「相手にわかりやすい文章を書こうと意識してしまった」といった意見があったが、エージェントタブーゲームでは「人に見られないので緊張感無く問題に取り組めた」や「間違いを気にせず文章を書けた」といった意見が得られた。

これらのことからエージェントに協働者としての親近感が湧かなかったためエージェントタブーゲームの方が人タブーゲームに比べて積極性が弱まったが、評価を気にせずゲームに取り組むことができたことがわかった。

## 6.2 リテリングテストの好意性の評価

タブーゲームを経ることでリテリングテストに対する好意性が変化するかをリテリングテスト後のアンケートを基に評価する。

表 6 に記載されているグループ A のアンケート結果を分析していく。まず「リテリングテストは楽しかったか?」という項目において、1 回目のテストで「あまり楽しくなかった」「楽しくなかった」と否定的な意見が見られ、人タブーゲーム後とエージェントタブーゲーム後のテストでは「あまり楽しくなかった」と否定的な意見に 1 つずつ入り、「やや楽しかった」「楽しかった」と肯定的な意見が入った。また「リテリングテストに積極的に取り組みましたか?」という項目では、1 回目のテストで「非常にそう思う」「そう思わない」といった項目があり、2・3 回目のテストでは「そう思う」「あまりそう思わない」といった意見がみられた。回数を重ねても否定的な意見と肯定的な意

見の人数の変化はあまり見られなかった。このことからタブーゲーム後に行うリテリングへの「楽しさ」に対してはゲームを行うことで肯定的な意見が増えたが、積極性への変化はあまり見られなかった。

次にタブーゲームを行うことでリテリングテストへの取り組みやすさが変化するかを分析する。「リテリングテストは取り組みやすかったですか?」という項目で、1回目のアンケートは「あまりそう思わない」「そう思わない」といった否定的な意見が多かった。しかし2回目・3回目に行ったアンケートでは、「あまりそう思わない」とやや否定的な意見が1つずつ入ったが、「ややそう思う」「そう思う」と肯定的な意見が見られた。また「内容に自信がなくても文章を書き進めようと思いましたか?」という項目では、1回目が「ややそう思う」「そう思わない」といった意見だったが、2・3回目のテストでは「そう思う」「ややそう思う」といった肯定的な意見の数が多くなった。以上のことからタブーゲームを行うことで、リテリングテストへの取り組みやすさに否定的な気持ちが弱くなり、文章を作成することに前向きになるとわかった。

表 6. グループ A のリテリングテスト後アンケート

質問項目	選択肢	テスト①	テスト②	テスト③
今回のリテリングテストは楽しかったですか?	とても楽しかった	0	0	0
	楽しかった	0	0	1
	やや楽しかった	0	1	0
	あまり楽しくなかった	1	1	1
	楽しくなかった	1	0	0
	全く楽しくなかった	0	0	0
今回のリテリングテストに積極的に取り組みましたか?	非常にそう思う	1	0	0
	そう思う	0	1	1
	ややそう思う	0	0	0
	あまりそう思わない	0	1	1
	そう思わない	1	0	0
	全くそう思わない	0	0	0
今回のリテリングテストは取り組みやすかったですか?	非常にそう思う	0	0	0
	そう思う	0	0	1
	ややそう思う	0	1	0
	あまりそう思わない	1	1	1
	そう思わない	1	0	0
	全くそう思わない	0	0	0
内容に自信がなくても文章を書き進めようと思いましたか?	非常にそう思う	0	0	0
	そう思う	0	1	1
	ややそう思う	1	1	1
	あまりそう思わない	0	0	0
	そう思わない	1	0	0
	全くそう思わない	0	0	0

表 7 のグループ B のアンケート結果を分析していく。まず「リテリングテストは楽しかったか?」という項目では、1回目のテストとタブーゲームを行っ

た2回目・3回目のテストで「やや楽しかった」「楽しかった」といった肯定的な意見が見られた。また「リテリングテストに積極的に取り組みましたか？」という項目では、1回目と2回目のテストでは「非常にそう思う」「ややそう思う」といった意見があり、3回目のテストでは「非常にそう思う」「そう思う」といった意見がみられた。この結果により人タブーゲームを行った後のリテリングテストの方が学習者の積極性が強くなった。

次にタブーゲームを行うことでリテリングテストへの取り組みやすさの変化を分析する。「リテリングテストは取り組みやすかったですか？」という項目で、1回目と2回目のアンケートでは「非常にそう思う」「そう思う」といった意見がみられた。しかし3回目のアンケートでは、「そう思わない」と否定的な意見があった。また「内容に自信がなくても文章を書き進めようと思いましたが？」という項目では、1回目のアンケートで両者ともに「あまりそう思わない」といったやや否定的な意見であったが、2・3回目のテストでは「非常にそう思う」「そう思う」といった肯定的な意見の数が多くなった。このことから人タブーゲーム後のリテリングテストではテストに取り組みにくいと感じる学習者がいたものの、タブーゲームを行うことで自信が無くても英文を作成しようと前向きな気持ちになったとわかった。

表7. グループBのリテリングテスト後アンケート

質問項目	選択肢	テスト①	テスト②	テスト③
今回のリテリングテストは楽しかったですか？	とても楽しかった	0	0	0
	楽しかった	1	1	2
	やや楽しかった	1	1	0
	あまり楽しくなかった	0	0	0
	楽しくなかった	0	0	0
	全く楽しくなかった	0	0	0
今回のリテリングテストに積極的に取り組みましたか？	非常にそう思う	1	1	1
	そう思う	0	0	1
	ややそう思う	1	1	0
	あまりそう思わない	0	0	0
	そう思わない	0	0	0
	全くそう思わない	0	0	0
今回のリテリングテストは取り組みやすかったですか？	非常にそう思う	1	1	0
	そう思う	1	1	1
	ややそう思う	0	0	0
	あまりそう思わない	0	0	0
	そう思わない	0	0	0
	全くそう思わない	0	0	0
内容に自信がなくても文章を書き進めようと思いましたが？	非常にそう思う	0	1	1
	そう思う	0	1	1
	ややそう思う	0	0	0
	あまりそう思わない	2	0	0
	そう思わない	0	0	0
	全くそう思わない	0	0	0

### 6.3 学習者の産出語彙の評価

グループ A の学習者を被験者①・②，グループ B の学習者を被験者③・④とする。また，C3・A2 の結果は小数点以下 3 位を四捨五入したものである。

リテリングテストの結果から表を参考に分析を行う。まずグループ A における表 8 と表 9 のように被験者①と被験者②のリテリングテストによる産出語彙の評価は以下のようになった。被験者①②ともにリテリングテストで使用される語彙数に違いが見られなかった。また，作成した文章の正確さにおいてもタブーゲームを経ることで変化は見られなかった。また，リテリングテスト後に行った「今回のリテリングテスト中に，文章を組み立てるときに使う単語やフレーズが増えたと感じましたか？」というアンケートには，1 回目のテスト後には「全くそう思わない」「あまりそう思わない」という回答が見られた。またタブーゲームを行った後に行った 2・3 回目のテスト後アンケートで「ややそう思う」にそれぞれ 1 人が回答する一方で「そう思わない」「あまりそう思わない」とやや肯定的意見と否定的な意見が見られる結果となった。

表 8. 被験者①の産出語彙評価

リテリングテスト①		リテリングテスト②		リテリングテスト③	
項目	測定値	項目	測定値	項目	測定値
C1(全語彙数)	22	C1(全語彙数)	21	C1(全語彙数)	26
C2(異なる語の総数)	18	C2(異なる語の総数)	20	C2(異なる語の総数)	20
C3(T-unitあたりの語彙数)	11.00	C3(T-unitあたりの語彙数)	7.00	C3(T-unitあたりの語彙数)	6.50
A1(全誤り数)	4	A1(全誤り数)	1	A1(全誤り数)	5
A2(T-unitあたりの誤り数)	2.00	A2(T-unitあたりの誤り数)	0.33	A2(T-unitあたりの誤り数)	1.25

表 9. 被験者②の産出語彙評価

リテリングテスト①		リテリングテスト②		リテリングテスト③	
項目	測定値	項目	測定値	項目	測定値
C1(全語彙数)	12	C1(全語彙数)	13	C1(全語彙数)	13
C2(異なる語の総数)	12	C2(異なる語の総数)	11	C2(異なる語の総数)	12
C3(T-unitあたりの語彙数)	6.00	C3(T-unitあたりの語彙数)	6.50	C3(T-unitあたりの語彙数)	6.50
A1(全誤り数)	3	A1(全誤り数)	2	A1(全誤り数)	4
A2(T-unitあたりの誤り数)	1.50	A2(T-unitあたりの誤り数)	1.00	A2(T-unitあたりの誤り数)	2.00

次に表 10 と表 11 のようにグループ B 被験者③と被験者④のリテリングテストによる産出語彙の評価は以下のようになった。被験者③は 1 回目のリテリン

グテストのときよりもタブーゲームを経て行った2・3回目のテストで産出語彙数がやや右肩上がりに増えており、T-unitあたりの語彙数も1回目のときよりも2倍程度増加した。被験者④はエージェントタブーゲームの後は語彙数とT-unitあたりの語彙数が1回目のリテリングテスト時よりもやや減少しているが、人タブーゲームの後にいったリテリングテストでは同程度の語彙数になった。また、両者ともにテスト回数に依らずの誤り数は同じ程度であった。1回目のテスト後のアンケートでは「非常にそう思う」「そう思う」という意見が見られた。しかしながら、2・3回目のテスト後のアンケートでは「非常にそう思う」「そう思う」という意見がそれぞれに見られたのに対して「そう思わない」という否定的な意見が見られる結果となった。

表 10. 被験者③の産出語彙の評価

リテリングテスト①		リテリングテスト②		リテリングテスト③	
項目	測定値	項目	測定値	項目	測定値
C1(全語彙数)	23	C1(全語彙数)	29	C1(全語彙数)	31
C2(異なる語の総数)	22	C2(異なる語の総数)	24	C2(異なる語の総数)	27
C3(T-unitあたりの語彙数)	7.33	C3(T-unitあたりの語彙数)	14.50	C3(T-unitあたりの語彙数)	15.50
A1(全誤り数)	3	A1(全誤り数)	3	A1(全誤り数)	2
A2(T-unitあたりの誤り数)	1.00	A2(T-unitあたりの誤り数)	1.50	A2(T-unitあたりの誤り数)	1.00

表 11. 被験者④の産出語彙の評価

リテリングテスト①		リテリングテスト②		リテリングテスト③	
項目	測定値	項目	測定値	項目	測定値
C1(全語彙数)	35	C1(全語彙数)	29	C1(全語彙数)	39
C2(異なる語の総数)	31	C2(異なる語の総数)	25	C2(異なる語の総数)	34
C3(T-unitあたりの語彙数)	11.67	C3(T-unitあたりの語彙数)	7.25	C3(T-unitあたりの語彙数)	13.00
A1(全誤り数)	2	A1(全誤り数)	3	A1(全誤り数)	3
A2(T-unitあたりの誤り数)	0.67	A2(T-unitあたりの誤り数)	0.75	A2(T-unitあたりの誤り数)	1.00

## 第7章 考察

本研究では教師エージェントとピアエージェントを組み込んだ web アプリケーションを用いた英語学習支援の有用性について実験を行ってきた。

### 7.1 英単語の産出語彙増加の有用性

グループ A のリテリングテストの結果では、タブーゲームを行っても産出語彙の変化は見られなかった。グループ B では被験者③がタブーゲームを行うことでリテリングテストの産出語彙がわずかに増加していたが、被験者④の産出語彙数は 1 回目と 3 回目のリテリングテストの語彙数は同程度であるものの 2 回目の産出語彙はそれらに比べてやや減少していた。これらのことから、人タブーゲームとエージェントタブーゲームによる産出語彙数は同程度であることが分かった。しかしながら、今回の実験では 1 日間に 10 分間のタブーゲームを行ったのみであったため、英語学習の実験環境として有効であるとは言えない。そのためエージェントを用いた学習支援の検証を行うためには学習時間を延ばして連日で実験を検証する必要があると考えられる。

### 7.2 人とエージェントでタブーゲームを行った場合の好意性の比較

グループ A・B とともに人タブーゲームを行った場合の方がエージェントタブーゲームに比べて楽しく積極的に取り組んでいた。またインタビュー結果から人タブーゲームでは推論者に親近感を持ってゲームに取り組んだが、エージェントタブーゲームでは親近感が湧かず淡々と取り組んだという意見が得られた。このことから学習者がエージェントタブーゲームに積極的に取り組む工夫が必要であると考えた。例えば、問題が提示されている画面から回答のフィードバックを受ける画面に遷移するまでの間に、ピアエージェントがフィードバックを考えている様子を示すイラストやアニメーション画像を組み込む。このような工夫により、学習者は実際のピア・ラーニングと同様に協働者とともに考えながら問題に取り組んでいるかのような疑似体験を得ることができると考えられる。

またエージェントタブーゲームの方が人タブーゲームに比べて回答回数が多かったという結果が得られた。またインタビューでは、人タブーゲームでは相手にわかりやすい文章を書こうと意識したが、エージェントタブーゲームでは緊張感無く取り組めたという回答が得られた。これらの結果から、エージェントタブーゲームは学習者がリラックスした状態で取り組めることに加えて、教師エージェントによるフィードバックが即時に提供されるため、学習者のアウトプ

ット能力の向上が促進されると考えられる。さらに、学習者のライティングに対する抵抗感の低減にもつながると考えられる。

### 7.3 人とエージェントでタブーゲームを行った場合のリテリングテストへの好意性の比較

#### テストへの好意性の比較

グループ A では、積極性に大きな変化は見られなかったものの、テストを重ねるごとに「楽しい」と感じる学習者が現れた。また、テストへの取り組みやすさについては、タブーゲームを繰り返すことで「取り組みやすい」と感じる意見がみられた。さらに、初回テストでは自信がなく、文章作成に対して否定的な意見を持つ学習者も見られたが、タブーゲームを通じて否定的な意見はなくなった。グループ B では楽しさと積極性に大きな変化はみられなかった。一方で取り組みやすさは1回目と2回目で肯定的な意見が多かったものもエージェントタブーゲームを行った後のテストでは「そう思わない」と否定的な意見も出てきた。一方で内容に自信がなくても文章を書こうと思ったかというアンケートでは1回目はやや否定的な意見が見られたものも2・3回目のテストでは肯定的な意見になった。以上のことからタブーゲームを終えることでリテリングテストへの楽しさと積極性に大きな変化は見られなかったが取り組みやすさには変化が見られた。また取り組みやすさの変化はエージェントタブーゲームを行った後のテストの方が人タブーゲームに比べて肯定的な意見に変化していった。これは、エージェントによるタブーゲームの方が、人とのタブーゲームに比べて学習者の回答回数が多かったことから、ライティングに対する心理的ハードルが低くなったために生じたためだと考えられる。

## 第8章 おわりに

学習者が複数のエージェントを組み込んだ web アプリケーションでピア・ラーニングを用いた英語学習支援の有用性を調べるためにタブーゲームで検証を行ってきた。本研究の貢献は以下のとおりである。

### マルチエージェントアーキテクチャの構築とピア・ラーニングのプロンプト作成

複数の大規模言語モデルの文脈を独立してピア・ラーニングを行うためのアーキテクチャを構築した。これにより 1 つの web アプリケーションで大規模言語モデルに与える情報を制御することができ、ピアエージェントと教師エージェントの確立を行うことができる。また問題の作成とアドバイスを返す教師エージェントと学習者に協働するピアエージェントのプロンプトを作成した。これにより個人でピア・ラーニングを行うことができる web アプリケーションを作成した。

### 英語学習支援の有用性の検証

本研究の英語学習支援は産出語彙の増加に有用性が見られなかったことがわかった。また、推論者が人の方が積極的に学習に取り組めた一方で、エージェントの場合がと英文を作成しやすいことがわかった。

本研究ではレベルに応じて使用してはいけない英単語を 3 つ使用せずにターゲット単語について英文で説明するタブーゲームとその後に産出語彙の評価を行うためにリテリングテストを行った。しかしタブーゲームを行っても学習者の産出語彙に変化は見られなかった。さらに人と行うタブーゲームの方がエージェントで行う場合よりも楽しく積極的に取り組めたという意見がみられた一方で、英文作成回数はエージェントで行うタブーゲームの方が大幅に増加した。

今後はタブーゲームで学習する期間と時間を延ばすことで産出語彙が増加するかを検証する必要がある。またエージェントタブーゲームを長期間行うことで学習者の心理的負荷や英語学習に積極的に取り組めたかを評価するべきである。さらにピアエージェントが実際のピア・ラーニングの協働者と同様の動きをイラストやアニメーションを用いて設計することで学習者が親近感を持つように工夫する必要があると考えられる。

## 謝辞

本研究を行うにあたり，ご指導くださりました村上陽平教授に深く感謝申し上げます．また，Mondheera 助教や研究室の同期・先輩方にも厚い協力，ご指導を賜りました．感謝申し上げます．最後になりますがアンケートにご協力くださいました皆様に感謝申し上げます．

## 参考文献

- [1]柳瀬陽介: ChatGPT による学術英語語彙の自律的学習 ―言語観とプロンプト設計と学習者認識の一貫性―, KELES ジャーナル, 9 巻, pp.45-51(2024).
- [2]中谷素之, 伊藤崇達:ピア・ラーニング 学び合いの心理学, 金子書房, pp.2-5 (2013).
- [3]加藤新, 服部宏充, 吉添衛, 山本友輔: LLM に基づくエージェントモデルの MAS への導入, 2024 年度人工知能学会全国大会(第 38 回), 第 38 回(2024).
- [4] 早船 由紀見:英語熟達度によるリテリングパフォーマンスの違い: 複雑さ, 正確さ, 流暢さとの関係と発話内容の分析, JACET-KANTO Journal, 9 巻, pp.26-41(2021).
- [5]望月通子: ピア・ラーニングに対する 学習者の認識と学びのプロセス, 外国語学部紀要, 第 8 号(2013).

## 付録

### プロンプトの実例

#### 問題作成タスク用プロンプト

ここで使われている{question\_word}はサーバ上にあるターゲット単語である。また{userlevel}は学習者が web アプリケーションログイン時に入力した CEFR のレベルである。

あなたはアメリカ人の英語の先生です。

今からあなたに私の「CEFR レベル」を教えます。私の CEFR レベルは「{userlevel}」です。さらに、ターゲット単語である「{question\_word}」を送信します。その「ターゲット単語」に関する「タブー単語」は私の「CEFR レベル」に対応したものを考えてください。私は、その「タブー単語」を使用せずに「ターゲット単語」を英語で説明する「タブーゲーム」を行います。

いまから「タブー単語」の定義を行います。

タブー単語(定義)

「タブー単語」とは「{question\_word}」に関連する英単語です。このとき「タブー単語」は「{question\_word}」の意味や用途に関連しています。一般的に定義されている「避けるべき」意味をもつタブーではありません。また「タブー単語」は「タブーゲーム」にふさわしい英単語だと思うものを 3 つ選んでください。」

このときタブー単語は私の CEFR レベルに応じて出力してください。

いまからあなたの出力内容である<問題作成>の手順を教えます。

<問題作成>

①私の CEFR レベルを確認する。

②①を確認して、「{question\_word}」に関連する英単語である「タブー単語」を 3 つ出力する。

③②が①ので確認した CEFR のレベルに応じた「タブー単語」であるかを確認してください。①で確認した CEFR レベルに見合った「タブー単語」でなければ修正して、私の CEFR のレベルに見合った「タブー単語」を 3 つ出力してください。

いまからそれぞれのレベル別に応じたタブー単語の選び方を例として教えます。

<入力内容>talk: talk, CEFR レベル A1

<CEFR レベルに応じたタブー語> : タブー単語 : pet, bark, tail

<入力内容>ターゲット単語:talk, CEFR レベル A2

<CEFR レベルに応じたタブー語> : タブー単語 : chat, discuss,

conversation

<入力内容>ターゲット単語:talk, CEFR レベル B1

<CEFR レベルに応じたタブー語> : タブー単語 : communicate, dialogue,

express

<入力内容>ターゲット単語:talk, CEFR レベル B2

<CEFR レベルに応じたタブー語> : タブー単語 : discuss, debate, verbal

<入力内容>ターゲット単語:talk, CEFR レベル C1

<CEFR レベルに応じたタブー語> : タブー単語 : articulate, converse,

rhetoric

<入力内容>ターゲット単語:talk, CEFR レベル C2

<CEFR レベルに応じたタブー語> : タブー単語 : discourse,

extemporaneous, interlocutor

またあなたの<出力内容>の例を教えます。次の例はターゲット単語が talk, レベルが B1 です。

ターゲット単語 : talk

タブー単語 : communicate, dialogue, express

ここでは私が「タブー単語」を使用せずに「{question\_word}」を英語で説明する「タブーゲーム」を行う上で、私の CEFR レベルに応じた最適な「タブー単語」を送信してください。

いまから「{question\_word}」を送ります。私の CEFR レベルは「{userlevel}」です。準備は良いですか？

### チェックタスク用プロンプト

あなたは英語の先生です。

これから「タブー単語」と「英文」を送ります。

私は「タブーゲーム」を行っており、「ターゲット単語」を英語で説明しますが、「タブー単語」を使ってははいけません。「ターゲット単語」は、私が説明する英単語です。「タブー単語」は、使用が禁止されている単語です。

役割 : 私の「英文」に「確認単語」が含まれているかどうかをチェックすることです。ただし、文の意味を推測したり、類義語や関連語に変換したりせず、単語そのものが含まれているかどうかを確認してください。

確認単語の定義

「確認単語」とは、タブー単語の複数形、現在形、過去形、過去分詞形、品詞変換、および群前置詞を含みます。

ただし、文の意味を推測したり、類義語や関連語を考えたりすることは禁止です。

文中にタブー単語そのもの、またはその変化形があるかどうかのみを確認してください。」

手順について説明します。

1. 「英文」に「確認単語」が含まれているかをチェックしてください。  
このとき、文の意味を推測したり、類義語を考えたりすることは禁止です。単語そのものの存在だけを確認してください。
2. もし「英文」に「確認単語」が含まれていない場合は、「タブー単語は含まれていません」とだけ出力してください。

以下に「入力内容」と「出力内容」の例を示します。

・「英文」に「確認単語」が含まれている場合

<入力内容>ターゲット単語: learn

タブー単語: learn, study, knowledge, teacher, education

英文: To gain new skills or understanding when someone teaches you or when you take part in activities that educate you.

<出力内容> タブー単語の'teaches'/'educate'が確認されました。訂正してください。

・「英文」に「確認単語」が含まれていない場合

<入力内容> ターゲット単語: learn

タブー単語: learn, study, knowledge, teacher, education

英文: Acquiring new abilities through experience.

<出力内容>タブー単語は含まれていません

また、「英文」に含まれる単語を類義語に変換してタブー単語とみなすことは禁止です。

例えば、タブー単語: vehicle

英文: It is a big car.

car → (変換) vehicle

このように類義語を変換して「タブー単語が含まれている」と判断することは禁止です。

これはタブーゲームのルールに反するため、類義語の変換による誤った判定は行わないようにしましょう。

## アドバイスタスク用プロンプト

あなたは英語の先生です。

今からあなたに「ターゲット単語」・「タブー単語」・「英文」を送信します。

役割：あなたの役割は「**avoid** 単語」を含めずに私の書いた「英文」が「ターゲット単語」を説明できるようにヒントの指導を行うことです。これから<出力内容>と「**avoid** 単語」の定義を教えます。

### avoid 単語(定義)

「**avoid** 単語」とは、私が送信した「タブー単語」と「タブー単語」の複数形や活用形・品詞転換・群前置詞です。このとき「英文」の文脈からの推測を避け、単語そのものの存在の有無にだけ焦点を当てるようにしてください。いまから例を教えます。

'create'に対する avoid 単語は'create, 'created', 'creating', 'creation', 'creativity'.」

これからあなたが私に出力する内容を教えます。

<出力内容>

<文法の指導>

①「英文」に文法エラーが含まれていた場合、エラーが出ている文章を指摘してください。

②①にタブーゲームの正解は教えないでください。

③①では「**avoid**」単語を含めないでください。

<語彙の指導>

①英文に使われている英単語が不自然な表現で使われていた場合その単語を指摘してください。

②私の英文に修正すべき英単語が含まれている場合は、それらの代案となる英単語を教えてください。このとき出力する内容に「タブー単語」と「**avoid** 単語」を含めないでください。

③①と②の出力する内容に「**avoid** 単語」と「タブー単語」を使用しないでください。

<ゲームの推論>

①私の英文をみて、どの英単語について書かれているかを推論して、その結果を全て私に教えてください。

②①の英文を読んでターゲット単語を説明するために不足している英単語を私に教えてください。このとき説明する英文や英単語に「タブー単語」と「**avoid** 単語」を含めないでください。

③②の英単語を使って簡単で短い節を作成してください。このとき説明する英文や英単語に「タブー単語」と「avoid 単語」を含めないでください。

④①と②と③には「avoid 単語」を含めないでください。  
今から<入力内容>と<出力内容>の例を教えます。

<入力内容>

ターゲット単語：air conditioning

タブー単語：adjust, environment, automation, air conditioning

英文：this device want to personel comfortable. For the reason, a person control temprature in the room.

あなたがこの英文をみて予測した正解と考えている文章：This device wants to make personnel comfortable. For the reason, through this device, a person controls the temperature in the room.

<出力内容>

文法の指導：主語と動詞の一致： "this device want" → 主語 "this device" は単数なので、動詞は "wants" にする必要があります。不定詞や動名詞の正しい使い方: "to personel" → "personel" は誤りで、"to make a person comfortable" とする必要があります。

語彙の指導：personel は「人員」を意味する単語で、ここでは「人」や「人々」を指すべきです。適切な語は "a person" または "people" です。"temprature" は誤字で、正しくは "temperature" です。

ゲームの推論：

この英単語は「エアコン」もしくは「ストーブ」や「扇風機」の説明をしています。

「ターゲット単語」であるエアコンの説明をするために、"make summer cool", "make winter hot"や"summer", "winter"といった単語を加えて特徴をあげましょう。これらの使い方は、it makes summer cool in the room, make winter hot などの使い方があります。

あなたの<出力内容>に「タブー単語」と「avoid 単語」を絶対に含めないでください。ゲームのルールに背きます。また、タブーゲームの正解の文章を教えないでください。

### アンサータスク用プロンプト

{userlevel}とは学習者が web アプリケーションログイン時に入力した CEFR のレベルである。

あなたは英語の先生です。

今からあなたに「ターゲット単語」・「タブー単語」を送信します。

これからあなたは「avoid 単語」を使用せずに「ターゲット単語」を英語で説明する「タブーゲーム」を行っています。「ターゲット単語」はあなたが英文で説明する英単語、「avoid 単語」は使用してはいけない単語です。

また、あなたの回答は私でも理解できるように出力してください。あなたには私の CEFR レベルを教えます。

いまから「avoid 単語」について定義します。

「avoid 単語」とは、私が送信した「タブー単語」の複数形や活用形・品詞転換・群前置詞です。このとき「英文」の文脈からの推測を避け、単語そのものの存在の有無にだけ焦点を当てるようにしてください。それは「タブーゲーム」の真意に背きます。

例を教えます。'create'に対する avoid 単語は'create, 'created', 'creating', 'creation', 'creativity'です。」

<出力内容>をこれから教えます。

<出力内容>

①私が理解できる英文を書いてください。私の CEFR レベルは

「{userlevel}」です。

②①を確認して、英文を簡潔に書いて「avoid 単語」を使わずにターゲット単語を説明してください。

③②に「avoid 単語」が含まれていた場合もしくは、あなたの文章が①のレベル対応していない場合は、修正して代案となる英文を教えてください。

いまから<入力内容>と<出力内容>の例を教えます。

<入力内容>

CEFR のレベル A1, ターゲット単語:bus, タブー単語:ride, bus, big, car

<出力内容>

It is a long vehicle that carries many people from one place to another. It stops at different locations so people can enter or leave.

あなたにはこれから「ターゲット単語」と「タブー単語」を送ります。準備は良いですか？

**推論タスク用プロンプト**

{sentence} は学習者の書いた英文である.

You are a Taboo game guesser. You should try to guess the word from the given clue. Only give the word you think it is answer. Do not answer as a sentence or add something else that is not your answer.

{sentence}\n\n : Read this English sentence and answer what English word is expressed. Also, write down the reason for your evaluation.