

卒業論文

多言語コミュニケーションのための ファシリテータエージェント

指導教官 村上 陽平 准教授

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
2600180341-6

横手 詩弓

2021年度（秋学期）卒業研究3（CH）
令和3年12月20日

多言語コミュニケーションのためのファシリテータエージェント

横手 詩弓

内容梗概

現代社会において、異なる国の文化や人々を理解する力が求められている。この力を育むためには、多様な文化に触れる機会を設け、文化差や言語差がある中でコミュニケーションを経験することが重要である。例えば、NPO パンゲアが主催する『KISSY』というサマースクールでは、様々な国の児童が集まり、機械翻訳を用いた多言語コミュニケーションチャットシステムでグループワークを行っている。しかしながら、既存の機械翻訳では、対訳データなどの言語に関するリソースが少ない言語（以降、低資源言語とよぶ）の翻訳精度が低く、低資源言語話者が多言語コミュニケーションに積極的に参加できない問題がある。

そこで、本研究では、既存の対話支援研究をもとに、低資源言語話者による多言語コミュニケーションを支援するためのファシリテータエージェントを構築する。具体的には、多言語コミュニケーション環境において、低資源言語話者のコミュニケーション促進に有効であると考えられるファシリテータエージェントの発話を定義し、多言語チャットシステムである LangridChat にファシリテータエージェントを組み込んで、その効果を検証する。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。

ファシリテータエージェントの発話の定義

多言語コミュニケーションにおける低資源言語話者の発話を促進するためのファシリテータエージェントの発話を定義する必要がある。また、その発話をファシリテータエージェントに実装するために、発話の実行条件も明確にする必要がある。さらに、被験者実験を通じた主観的評価と客観的評価によって、定義したコミュニケーションの有効性を検証する必要がある。

ファシリテータエージェントの実装

既存の多言語コミュニケーションツール LangridChat において、ファシリテータエージェントが発話ルールに基づいて、低資源言語話者やその他の話者の発話に応じて振る舞うことができるように、多言語コミュニケーションの観測と発話機能を実装する必要がある。

本研究では、前者の課題を解決するために、単言語コミュニケーションにおける発話促進のための傾聴システムを参考にして、低資源言語話者が発話しやすい場を形成するための発話として、相槌などの単純なフィードバックと同意な

どのポジティブなフィードバックを選択した。さらに、多言語コミュニケーションの特徴である、他者の発話の理解の困難さを解消するために、他者に言い換えや議論の要約を依頼する発話を新たに加え、合計で 4 つの発話を定義した。なお、言い換えは他者の発話メッセージが閾値よりも長い場合に行い、要約は低資源言語話者の発話が閾値よりも長く行われなかった場合に行う。

次に、4 つの発話のうち、ポジティブなフィードバック以外の 3 つの発話をファシリテータエージェントの振るまいとして実装した。具体的には、低資源言語話者や他者の発話メッセージに応じた相槌や言い換え依頼は、配信されるメッセージを観測するためにサーバ上で実装し、要約依頼は低資源言語話者の発話間隔を計測するためにブラウザ上で実装している。一方、ポジティブなフィードバックは、低資源言語話者の発話がアイデアや自分の考えであるかを判定しなければならないため、Wizard Of Oz 法を用いた。ファシリテータエージェントの有効性を検証するために、低資源言語話者の被験者を 5 名、その他の言語話者の被験者を 14 名集め、グループディスカッションによる実験を行った。グループディスカッションは低資源言語話者 1 名、その他の言語話者 3 名の合計 4 名で行い、ファシリテータエージェント無しを含めた 5 種の実験を各 3 回実施した。実験後にアンケート調査による主観的評価と、実験中の被験者の発話数やメッセージ量など客観的評価を行った。本研究の貢献は以下の通りである。

ファシリテータエージェントの発話の定義

低資源言語話者の理解促進を目的として、議論の要約を依頼する発話（以降、A とよぶ）と他者に言い換えを求める発話（以降、B とよぶ）を定義した。更に、低資源言語話者の発話促進を目的として、発話に相槌などの反応を返す発話（以降、C とよぶ）と低資源言語話者が意見を出した際にポジティブな反応を返す発話（以降、D とよぶ）を定義した。実験の結果、主観的評価では「ファシリテータからのコミュニケーションが発話のきっかけになったか」の項目が 5% 有意水準で有意性があったが、客観的評価では有意性のある結果が得られなかった。

ファシリテータエージェントの実装

A は、低資源言語話者がメッセージを配信した時間を観測し、一定時間ごとに実行される。B は、サーバに配信された他者の発話メッセージを観測し、閾値より長い場合に実行され、C は、サーバに配信されたメッセージを観測し、送信者が低資源言語話者であった場合に実行される。実験中、いずれの発話も 100% 正しく動作した。

Facilitator Agent for Multilingual Communication

Shiyumi Yokote

Abstract

In present-day society, there is a need for the ability to understand different cultures and people. To develop this skill, it is necessary to understand different cultures and communicate with people from various cultures and languages. For example, in NPO Pangaea's "KISSY" workshop, children from various countries do group work using a multilingual communication chat system. However, machine translation accuracy is low in languages with few language-related resources (hereafter, low-resource languages), and low-resource language speakers are unable to actively participate in multilingual communication.

Therefore, this study builds on existing research to create a facilitator agent to aid multilingual communication low-resource language speakers. To define facilitator agent utterances that are effective in facilitating communication among low-resource language speakers environment. The following are the two issues that must be addressed in order for this approach to be realized.

Definition of a facilitator agent's utterance

It is necessary to define the utterance of facilitator agents. The conditions for producing utterances should also be made clear. Subjective and quantitative evaluations obtained from experiments are required to validate the effectiveness of communication.

Implementing a Facilitator Agent

It is necessary to implement the utterance functions of the chat system so that the facilitator agent can be moved according to the user's speech in the LangridChat.

In this study, to solve the first problem, we referred to an existing work. Then, simple feedback responding, and positive feedback, such as agreement, were selected to low-resource language speakers could utterance easily. Furthermore, to solve the difficulty of understanding others' utterances, we newly proposed utterances in which we ask others to paraphrase or summarize the discussion and defined a total of four utterances. Paraphrasing is done when another person's message is longer than the threshold, and summarizing is done when a low-resource language speaker has not responded for a time longer than the threshold.

Three utterances other than positive feedback were implemented as a system of facilitator agents. The paraphrase request is implemented on the server to observe the delivered message, and the summary request is implemented on the browser to measure the utterance interval of low-resource language speakers. On the other hand, for positive feedback, we used the Wizard of Oz method because needed to know whether the content of the low-resource language speaker's utterance was an idea. To test the facilitator agent effectiveness, we gathered five low-resource language speakers and 14 other language speakers for a group discussion. The group consisted of four speakers (one low-resource language speaker and three speakers of other languages), and five different experiments were conducted three times each, including one without a facilitator agent. The subjective evaluation was carried out through a questionnaire survey of the subjects, and the quantitative evaluation was carried out through the number of utterances and volume of messages of the subjects during the experiment. The contributions of this study are as follows.

Definition of a facilitator agent's utterance

This study defined several types of facilitator agent's utterances. The first type of utterance is the utterance that requests a summary of the conversation (strategy A). The second one is the utterance that asks someone else to paraphrase (strategy B). To promote the utterance of low-resource language speakers, this study also defined utterances responding to a previous utterance made by a low-resource language speaker (strategy C) and utterances which is a positive response to an opinion by a low-resource language speaker (strategy D). The experiment's results revealed that Subjective valuations item "Did the communication with the facilitator trigger your comments?" was significance. There was no significance in the objective evaluations.

Implementing a Facilitator Agent

Strategy A monitors the time it takes for a low-resource language speaker to deliver a message and executes it at regular intervals. According to the results of the experiment, all strategies worked correctly 100 percent of the time.

多言語コミュニケーションのための ファシリテータエージェント

目次

第 1 章 はじめに	1
第 2 章 関連研究	3
第 3 章 ファシリテータエージェントの発話	5
3.1 発話の定義	5
3.2 発話ルール	6
第 4 章 ファシリテータエージェントの実装	8
4.1 全体のシステム構成	8
4.2 サーバ側の実装	12
4.3 クライアント側の実装	14
第 5 章 実験	16
5.1 実験設定	16
5.2 相互作用の効果	19
5.3 発話の評価	21
5.4 主観的評価	22
第 6 章 考察	24
第 7 章 おわりに	25
謝辞	26
参考文献	27

第1章 はじめに

年々グローバル化が進んでいる現代社会では、異なる国の文化や人々を理解し、コミュニケーションをとる力が求められている。昨今は、機械翻訳の技術が進歩しており、グローバル化に伴って、異なる国の人々が翻訳機を通してコミュニケーションをとる場も増えている。例えば、NPO パンゲアが主催する『KISSY』というサマースクールでは、様々な国の児童が集まり、機械翻訳を用いた多言語コミュニケーションチャットシステムを利用したグループワークが行われている。こうした場で、自分とは違う国、文化、言語の人々とコミュニケーションを行うことで、グローバル化する社会の中で多様な価値観を理解し、受容する力を身に付けることができる[1]。

対話を支援するという研究はこれまでも数多く行われてきたが、そのほとんどが単言語コミュニケーション環境を想定したものである。上記の例のような多言語コミュニケーション環境との違いとしては、意思疎通が難しい、文化や価値観が違う、国によって独特の言い回しがあるなどが挙げられる。こうした違いを理解するためには、より活発なコミュニケーションを行うことで、相手の情報や考えを知ったり、自らも情報を発信し相手に伝えたりすることで、意思疎通を図ることが重要である。

しかし、既存の機械翻訳技術では、対訳データなどの言語資源が少ない低資源言語への機械翻訳精度が高くないため、低資源言語話者は相手の発言の内容が理解できない、また、伝えたい内容を正しく伝えられないなどの問題が起りやすい。その結果、低資源言語話者は積極的に会話に入ることができず、発言数が他の言語話者に比べて少ないという傾向がある。

本研究では、言語や文化が異なる多言語コミュニケーション環境において、低資源言語話者の発話促進に有効なコミュニケーションを明らかにする。低資源言語話者の発話促進を目的としたファシリテータエージェントのふるまいを定義し、母国語の異なる人同士でのグループディスカッションを通して実験することで、その効果を検証する。また、本研究において取り組む課題については、低資源言語話者の発話促進に有効であると考えられるファシリテータエージェントの発話の定義、定義したファシリテータエージェントの実装の2点である。

以下本論文では、2章で関連研究についての説明を行う。次に、3章でファシリテータエージェントの発話の定義と発話ルールについて、そして4章では、

ファシリテータエージェントの全体システムの構成とサーバ側の実装と，クライアント側の実装について記述する．5章では，実験の設定について，発話の評価と被験者による主観的評価の結果を示す．最後に，6章では全体を通しての分析結果についての考察を示す．

第2章 関連研究

本章では、コミュニケーション支援に関する既存の研究を示す。

既存のコミュニケーション支援として、傾聴対話システムというシステムが提案されている[2][3]。このシステムは、雑談対話の一種であり、ユーザの、対話したいという欲求を満たすことや高齢者の認知機能の維持などの効果が期待できる傾聴対話をシステムで実現することを目標としている。そして、よりユーザにとって発話しやすい傾聴対話を実現するために、システムのふるまい(以降、聞き手応答とよぶ)を設計している。同じ分野の従来の研究では、相槌に関して取り扱われている[4][5]が、これに加えて、ユーザの発話に応じて、慣用的表現の発話を返す語彙的応答、発話の一部を繰り返す応答、内容の詳細を問う掘り下げ質問の選択と生成に取り組んでおり[6]、[7]ではこれらの聞き手応答を、6種類に分類している。これらの過去の研究を踏まえ、石田らの研究では、より自然で発話しやすい傾聴対話を行うために、掘り下げ質問、繰り返し応答、語彙的応答、評価応答という聞き手応答に加えて、ユーザの発話の内容に応じてシステムが自分自身の考えや情報を提示する自己開示という応答の生成に取り組んでいる。そして、上記の各応答を生成する手法を提案したのち、ユーザ発話の音声認識、焦点解析の結果や、生成された応答などの情報を特徴量として、それぞれの応答が傾聴的に妥当であるかの判定を行い、適切な応答の種類を選択する手法も提案している。

さらに、被験者実験のデータを利用して、上記の各聞き手応答を統合したシステムの印象評定実験を行っている。被験者実験は、被験者1人対アンドロイド1体の Wizard Of Oz 法の対話形式で行われており、1セッション10分程度の対話データを13セッション使用している。比較対象として、提案システムとは別に、ベースラインシステムを用意している。なお、印象評定実験に用いられた音声データは、コーパス内の対話の一部の1分から3分程度を切り取り、アンドロイドの操作者が発話した部分を、提案システムとベースラインシステムそれぞれで生成した応答文をアンドロイド用の Text-To-Speech で音声合成したものに置き換えたものである。これらの音声対話データのペアを8組用意し、8人の被験者それぞれに対して、以下の手順で2種類の評価をさせている[2]。

評価 1:下記の9項目に関する、各音声対話データの全体の印象評定

1. データ A の音声を聴き，印象評定を行う
2. データ B の音声を聴き，印象評定を行う

評価 2:音声対話データ内での 2 つのシステム応答の対比較

1. 評価 1 と同じ対話部分について，語り手の音声のみを聴く
2. それぞれの応答箇所での応答としてシステム A と B のどちらによるものの方がより適切かを選択する

以上の流れを音声対話データ 1 から音声対話データ 8 まで行っている．また，評価 1 の印象評価では，表 1: 評価項目 [2] についてそれぞれ-3 から 3 の 7 段階で評価させ，その結果をベースラインシステムと比較することで提案システムの有意性を示している．

表 1: 評価項目

	評価項目
1	全体を通して聞き手の応答は自然でしたか
2	全体を通して聞き手は真面目に聞いていると感じましたか
3	全体を通して聞き手は集中して聞いていると感じましたか
4	全体を通して聞き手は積極的に聞いていると感じましたか
5	全体を通して聞き手は親身に聞いていると感じましたか
6	全体を通して聞き手は理解していると感じましたか
7	全体を通して聞き手は関心を持っていると感じましたか
8	全体を通して聞き手は話し手に共感していると感じましたか
9	全体を通して聞き手は語りを促していると感じましたか

第3章 ファシリテータエージェントの発話

本章では、まず、低資源言語話者のコミュニケーションを支援するために定義したファシリテータエージェントのふるまいについて説明する。その後、定義した発話がどのような条件でどのような発話を行うかについて説明する。

3.1 発話の定義

関連研究について調査を行ったのち、低資源言語話者のコミュニケーション支援に有効であると考えられる発話を定義した。

まず、低資源言語話者が議論の内容や他者の発話の意味を理解することを促進する目的の発話として、議論のまとめを要求する発話を定義した。この発話は、その時点での議論の内容を、低資源言語話者以外の被験者に簡単にまとめてもらうことで、低資源言語話者にとってより議論の状況を理解しやすくするという狙いがある。この発話によって、低資源言語話者が議論の内容や他者の意見を理解しやすくなる、意見を発話しやすくなるなどの効果が期待できる。

同じく、低資源言語話者の議論に対する理解の促進を目的として、低資源言語話者以外の被験者に対して言い換えを求める発話を定義した。この発話は、文字数が長く、上手く機械翻訳されない可能性がある発話について、低資源言語話者以外の被験者に簡潔に言い換えてもらうものである。これによって、低資源言語話者にとって理解しづらいと考えられる発話についても、理解を深めることができる。この発話についても、低資源言語話者が議論の内容や他者の意見を理解しやすくなる、意見を発話しやすくなるなどの効果が期待できる。

次に、低資源言語話者がより発話しやすくなることを目的とした発話として、低資源言語話者の発話に反応を返す発話を定義した。これは、単純な相槌を発話するものであり、低資源言語話者が発話しやすい空気をつくる、議論を活発化させるなどの狙いがある。これによって、低資源言語話者が発話しやすくなる、より活発に発話を行いやすくなるなどの効果が期待できる。なお、これは、第2章で説明した関連研究でも有効性が認められている[2][1]ふるまいである。

また、上記と同じ発話促進を目的として、低資源言語話者が意見を出した際に肯定的な反応を返す発話を定義した。この発話は、低資源言語話者の発話に対して同意などの肯定的な発話を行うもので、低資源言語話者が発話しやすい空気をつくる、より活発に発話させるなどの狙いがある。これによって期待できる効果

としては、低資源言語話者が発話しやすくなる、意見を発話しやすくなるなどが挙げられる。また、この発話についても、先行研究で有効性が認められている[1]。

3.2 発話ルール

3.1 で定義した発話の実行条件とその内容を定義した。

まず、議論のまとめを要求する発話は、低資源言語話者の発話が一定時間ないことを実行条件とした。実行内容は、被験者に対して、その時点での議論の状況をまとめるように要求するものである。具体的には、「今までの議論をおさらいしましょう」「いったん話をまとめましょう」といった内容である。また、試験実験の結果、3分ごとに実行すると効果的であると判断したため、低資源言語話者の最後の発話から3分間新たな発話がなかった場合に経過する実行するように設定している。なお、実行はbotによる自動返信で行う。

次に、低資源言語話者以外の被験者に対して言い換えを求める発話は、低資源言語話者以外の被験者が一定文字数以上を発話した場合に実行される。実行内容は、被験者に対して、一定文字数以上の発話に関して簡潔に言い換えるように要求するものである。具体的には、「簡単な言葉でまとめましょう」「簡潔に言い換えてください」といった内容が発話する。試験実験の結果、90文字以上を発話した場合、わかりにくい内容になっている可能性が高いと判断したため、低資源言語話者以外の被験者が90文字以上を発話した場合に実行するように設定している。実行は、botによる自動返信で行う。

低資源言語話者の発話に反応を返す発話は、低資源言語話者の発話があった場合に実行する。単純な相槌を返すものであるため、実行は発話の内容にかかわらず行われる。具体的には、「そうなんですね」「うんうん」「なるほど」といった内容のインドネシア語であり、botによる自動返信で実行する。この発話については、低資源言語話者に対するものであるため、被験者実験で低資源言語話者として参加してもらったインドネシア人に伝わりやすいように、インドネシア語で発話するように設定している。

低資源言語話者が意見を出した際に肯定的な反応を返す発話は、低資源言語話者が自分の考えやアイデアを発話した際に実行するものである。具体的には、挨拶、自己紹介、相槌、を除いた全ての発話に対して実行する。発話内容としては、「いいですね」「素敵な考えですね」「とてもいいと思います」といったものである。この発話については、低資源言語話者の発話内容を理解・判断して実行

する必要があるため、人間が **SYSTEM MESSAGE** のふりをして実行する **Wizard Of Oz** 法で実行した。

本章で説明した 4 種類の発話の内容とルールを、表 2：定義した発話の内容とルール に示す。

表 2：定義した発話の内容とルール

目的	内容	実行条件	実行方法	発話
低資源言語話者の理解促進	議論のまとめを要求する	低資源言語話者の発話が一定時間ない	bot	「今までの議論をおさらいしましょう」 「いったん話をまとめましょう」
	言い換えを求める	低資源言語話者以外の被験者が一定文字数以上を発話する		「簡単な言葉でまとめましょう」 「簡潔に言い換えてください」
低資源言語話者の発話促進	発話に反応を返す	低資源言語話者が発話する	bot	「そうなんですね」 「うんうん」 「なるほど」
	肯定的な反応を返す	低資源言語話者が自分の意見を発話する		WOZ 法

第4章 ファシリテータエージェントの実装

本章では、定義した発話の、多言語チャットシステム LangridChat への実装について説明する。

4.1 全体のシステム構成

LangridChat とは、Django と React をベースに開発された Web アプリケーションである。初期画面は図 1 : LangridChat の初期画面のようになっており、「Input your nick name」で名前を入力し、「Select Language」で母国語を選択することができる。現在は英語、日本語、タイ語、ベトナム語、インドネシア語、ネパール語、韓国語、簡体字中国語、繁体字中国語を選択することができる。多言語チャットシステムであるため、送信側のユーザが母国語で入力した内容はサーバで翻訳され、受信側のユーザそれぞれの母国語となって送信・表示される。また、図 2 : チャットルームを新規作成する場合の画面のように、「you also can create new room」をクリックすると、「JOIN ROOM」というボタンが「CREATE ROOM」というボタンに変化する。このボタンをクリックすることによって、新しいチャットルームを作成することができる。なお、この手順であらかじめ作成しておいたルームの ID を「Room」に入力することで、別のブラウザからもそのルームに入室することができる。これによって、複数人で同じチャットルームを使用しチャットを行うことができる。



Room

Input your nick name

Select language
Japanese

JOIN ROOM

you also can create new room :

図 1 : LangridChat の初期画面



図 2 : チャットルームを新規作成する場合の画面

チャットルームに入室後の画面は図 3 : チャットルーム入室後の画面のようになっている。左上に表示されている文字列がルーム ID であり，右上には選択した言語が表示されている。画面右側には，現在そのチャットルームに誰が入室しているかを表している。画面下部の入力欄からメッセージを入力でき，その右側の紙飛行機のアイコンをクリックすることで送信が可能である。また，メッセージを入力中のユーザは，入力欄の下に「Inputting user」として表示される。実際に LangridChat で会話をを行うと，以下の図のようになる。日本語話者とインドネシア語話者の会話を例としており，図 4 : 日本語話者のチャット画面が日本語話者のチャット画面，図 5 : インドネシア語話者のチャット画面がインドネシア語話者のチャット画面である。日本語話者が日本語で送信したメッセージが，インドネシア語話者の画面ではインドネシア語に翻訳され，表示されている。

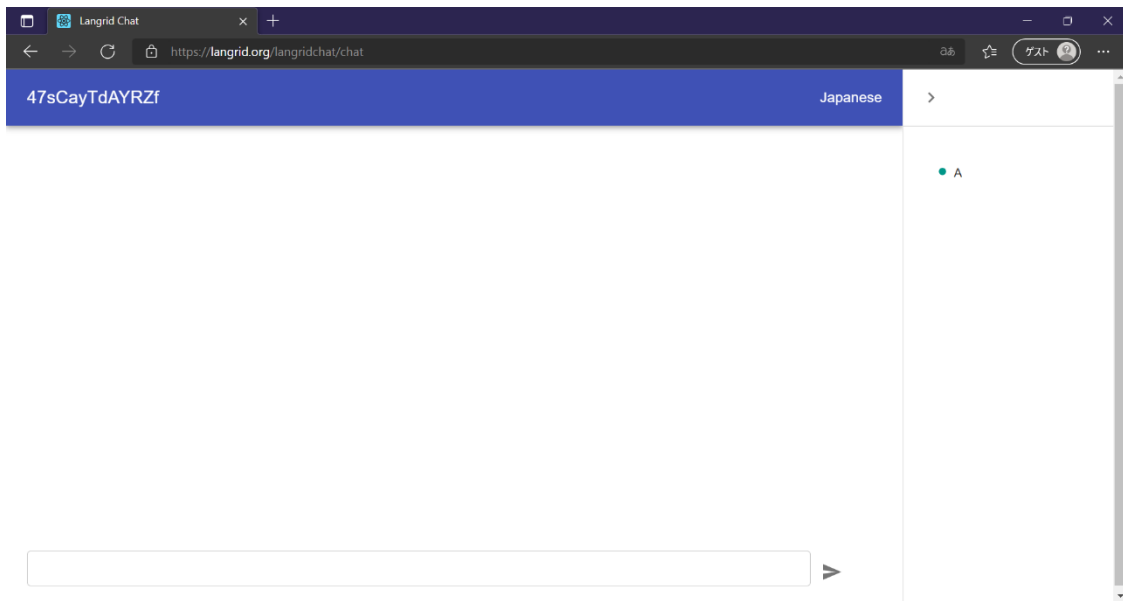


図 3：チャットルーム入室後の画面

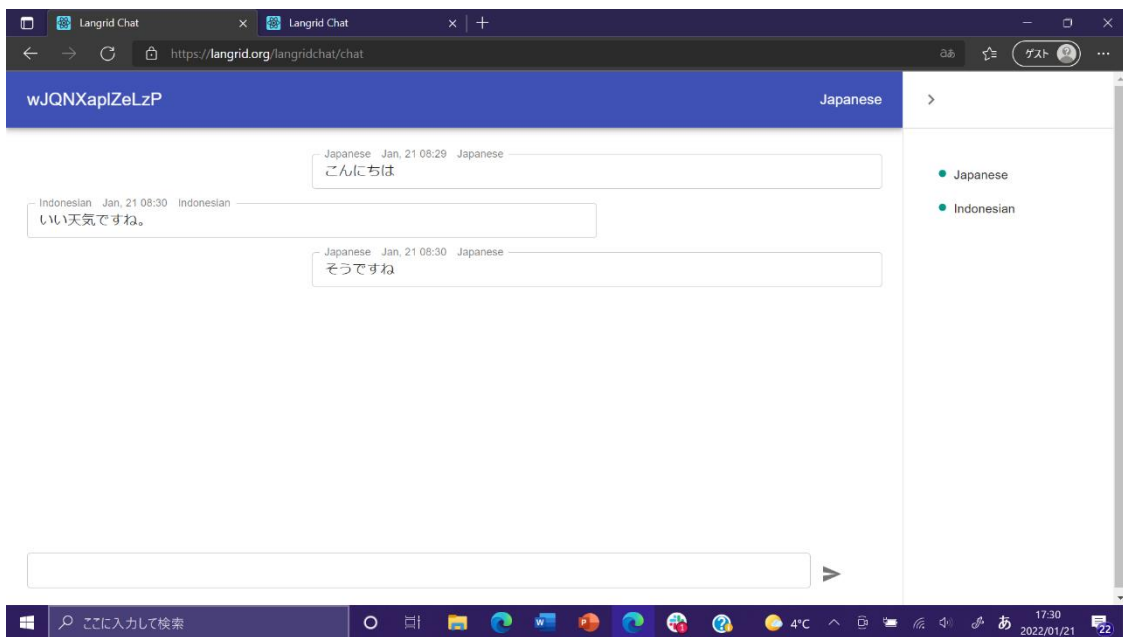


図 4：日本語話者のチャット画面

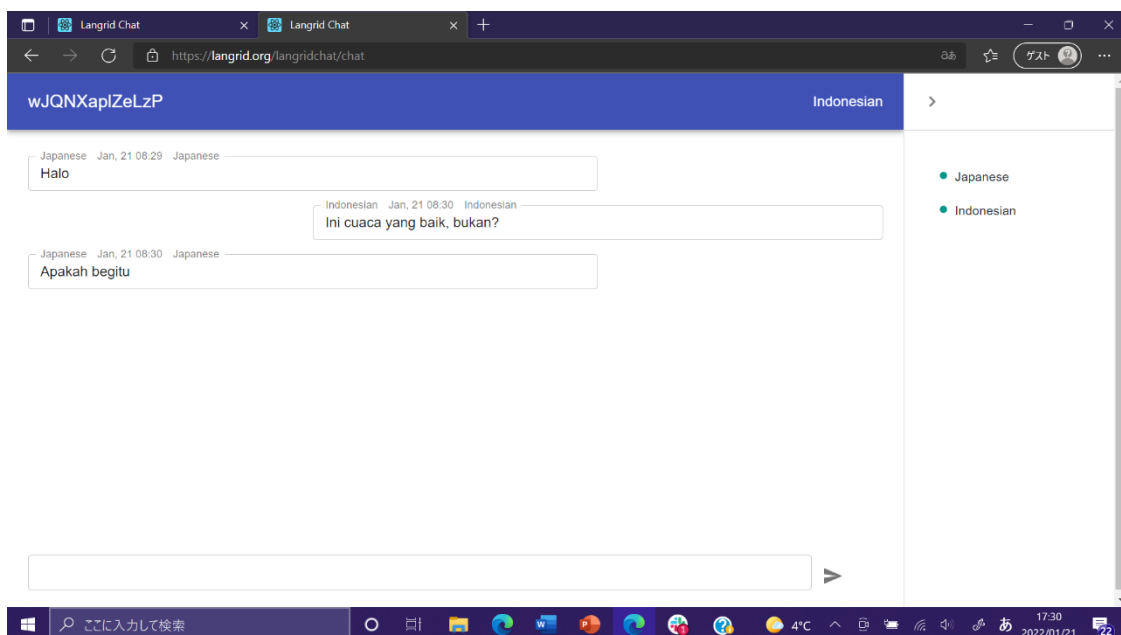


図 5：インドネシア話話者のチャット画面

このシステムは、サーバ側で動作する部分とクライアント側で動作する部分に分かれている。図 6 に示すように、前者は **Django** をベースに、メッセージ配信 **API** と翻訳コンポーネント、配信コンポーネントがある。そのため、チャットルームの作成、メッセージの翻訳と送信、チャットログの記録、ユーザ情報の取得などはサーバ側が行っている。後者は **React** ベースの **UI** があり、ユーザ情報の記録、送信されたメッセージなどのディスプレイ表示などを行っている。そのため、サーバ上で動作する部分を観測して実行する発話はサーバ上に、クライアント側で動作する部分を観測して実行する発話はクライアント上に実装する必要がある。

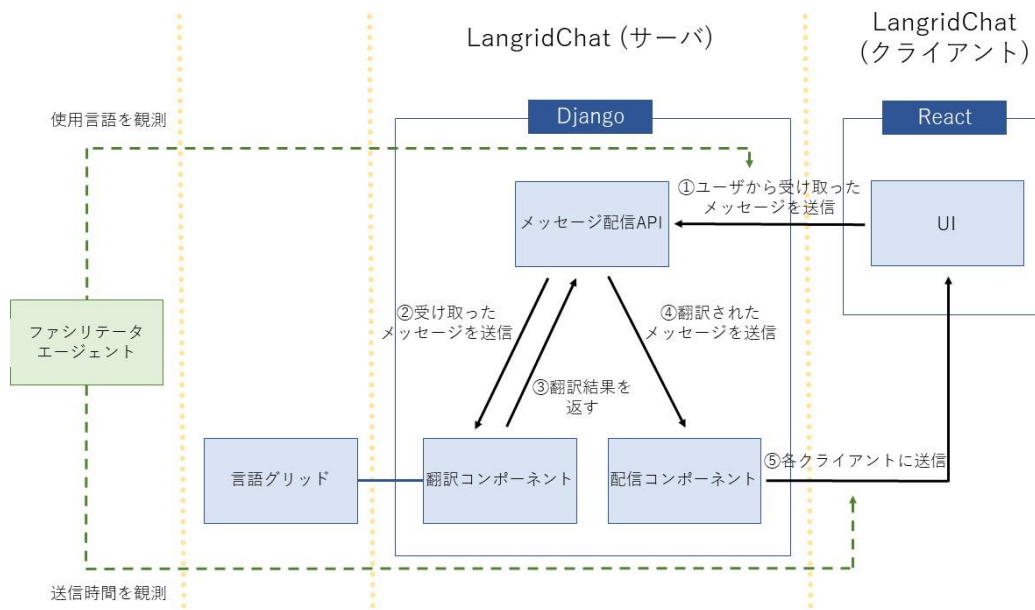


図 6 : LangridChat のシステム構成図

4.2 サーバ側の実装

サーバ上で実装した発話は、言い換えを求める発話と、低資源言語話者の発話に反応を返す発話である。図 7: 言い換えを求める発話 は、言い換えを求める発話のフローチャートで、図 8: 低資源言語話者の発話に反応を返す発話のは低資源言語話者の発話に反応を返す発話のフローチャート、図 9: 議論のまとめを要求する発話のは議論のまとめを要求する発話のフローチャート である。

言い換えを求める発話は、低資源言語話者以外の被験者が 90 文字以上を発話した場合に実行するため、メッセージの送信者の使用言語とメッセージの文字数を観測する必要がある。これらの情報は、サーバ側でやりとりされるため、サーバ側に実装した。メッセージの送信者の使用言語を確認し、低資源言語、すなわちインドネシア語もしくはタイ語以外であった場合に、さらにそのメッセージの文字数を確認する。そして、その文字数が 90 文字以上であった場合は、「簡単な言葉でまとめましょう」という発話か「簡潔に言い換えてください」というメッセージを **SYSTEM MESSAGE** というユーザ名でランダムに送信するように設定している。

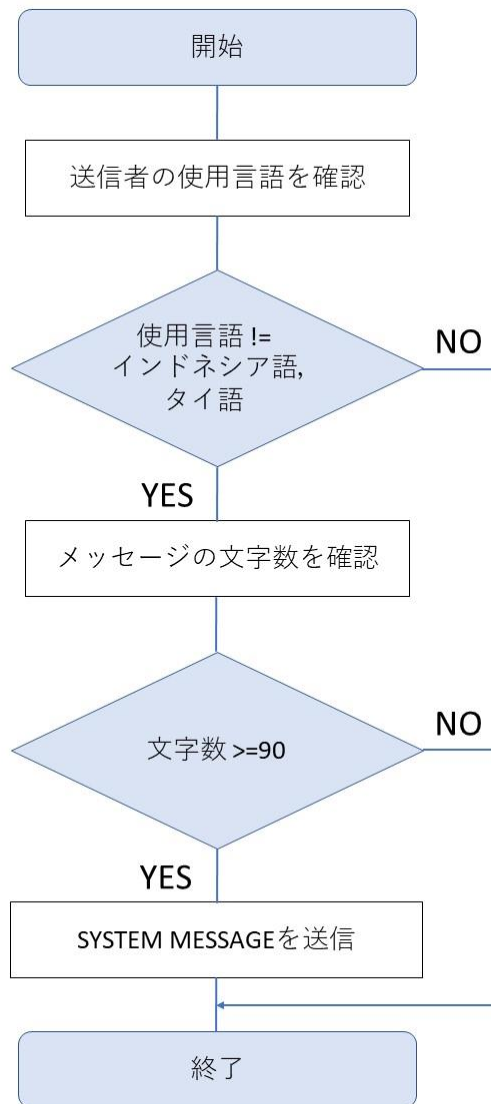


図 7: 言い換えを求める発話のフローチャート

低資源言語話者の発話に反応を返す発話は、低資源言語話者の何かしらの発話が条件になっている。つまり、エージェントはメッセージの送信者の使用言語を観測する必要があり、先述の通りその情報はサーバ側が司っているため、サーバ上に実装する必要がある。メッセージの送信者の使用言語を確認し、インドネシア語もしくはタイ語であった場合、「**seperti itu ya** (そうなんですね)」か「**oh iya juga** (なるほど)」か「**Iya iya** (うんうん)」というメッセージを **SYSTEM MESSAGE** というユーザ名でランダムに送信する。

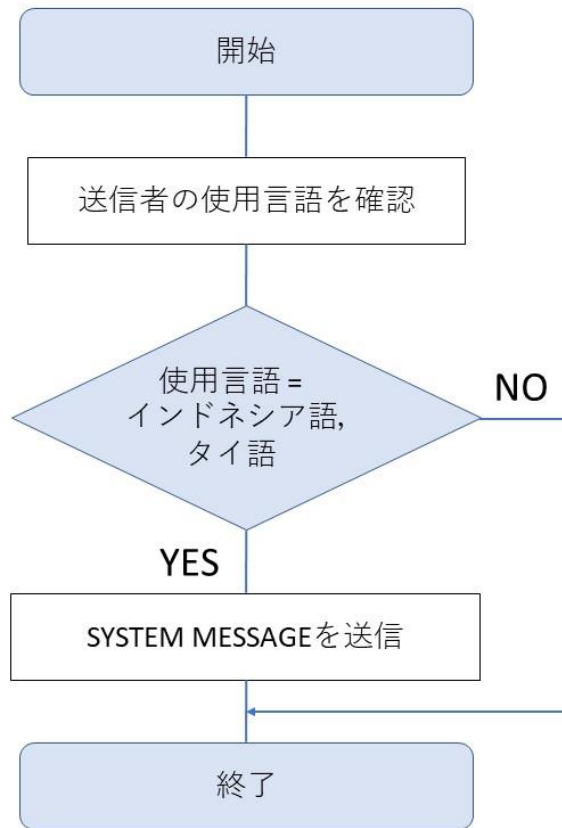


図 8 : 低資源言語話者の発話に反応を返す発話のフローチャート

4.3 クライアント側の実装

クライアント上で実装した発話は、議論のまとめを要求する発話である。この発話は、低資源言語話者の発話が一定時間ないことを条件に実行する。そして、ユーザがメッセージを送信した時間は、クライアント側で管理する部分であるため、クライアント側に実装した。ユーザによるメッセージの送信があった際に、その使用言語がタイ語、もしくはインドネシア語であることをサーバ側で確認する。そして、その二つのいずれかに当てはまる場合は、メッセージが送信された時間から 3 分間計測する。もし、3 分間を計測中に、低資源言語話者から新たな発言があった場合は、その時点で計測をリセットし、新たに 3 分間計測する。そして、低資源言語話者の発話が無いまま 3 分経った場合に、「今までの議論をおさらいしましょう」もしくは「いったん話をまとめましょう」というメッセージを **SYSTEM MESSAGE** というユーザ名でランダムに送信する。

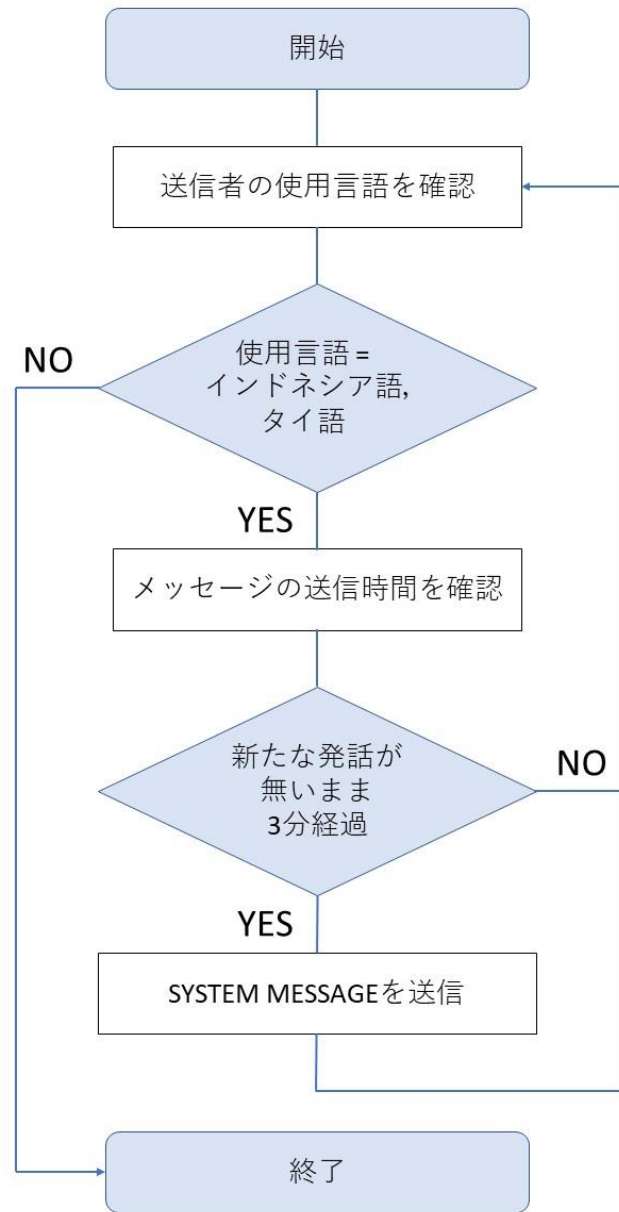


図 9 : 議論のまとめを要求する発話のフローチャート

第5章 実験

本章では、定義した各ファシリテータエージェントの発話の効果を検証するために実施した被験者実験について、被験者の人数、実験タスクなどの実験設定と評価の結果について記述する。

5.1 実験設定

被験者実験は、低資源言語話者としてインドネシア語話者が5人、その他の言語話者として中国語話者が5人、日本語話者が14人の合計19人の被験者を集めて行った。この被験者を、以下の表4：から表8：に示すように各4人の5グループに分け、グループディスカッションを通してファシリテータエージェントの発話の効果を検証した。なお、1グループにつき低資源言語話者は1人、中国語話者は1人、日本語話者は2人という内訳になっている。実験は、2日間に分けて行い、一日目は表4：低資源言語話者1の実験設定と

表5：低資源言語話者2の実験設定、二日目は、表6：低資源言語話者3の実験設定と表7：低資源言語話者4の実験設定、表8：低資源言語話者5の実験設定のグループを集めた。グループディスカッションは、各発話条件につき3回ずつ、全15回行った。そのため、1グループあたり3回実験を行っている。また、1回あたり15分の制限時間を設け、残り10分と5分の時点でアナウンスを行った。

実験タスク（ディスカッションテーマ）は以下の全5種類を用意し、タスクの難易度による差が出ないように、ディスカッションが1回終わるごとにシャッフルした。なお、タスクそれぞれに複数の選択肢を用意しており、その中から結論を1つ選び、その理由も考えることをタスクの最終目標とした。

表3：実験タスク

	実験タスク	選択肢
1	無人島に一つだけ持っていくなら何にするか	ライター、ナイフ、水、寝袋、釣り竿
2	オリンピックに新しい競技を採用する	ボウリング、綱引き、フリスビ

	ならどれか	一、ドッジボール、スキューバダイビング
3	一番重要な教科はどれか	国語、算数、理科、社会、英語
4	生まれ変わるなら何になりたいか	鳥、犬、猫、ライオン、イルカ
5	デートで注意すべきポイントは何か	場所、時間、服装、天気

また、被験者グループについても、同じメンバーで実験を繰り返すことで話しやすさが出てしまうことを避けるために、多数派である日本語話者の被験者を以下の表のとおりシャッフルした。そして、ディスカッションが終わるごとに、主観的評価のためのアンケートを記入させた。さらに、客観的評価も行うために、チャットのログデータも取得した。

表 4：低資源言語話者 1 の実験設定（グループ①）

	条件	実行方法	実験タスク	被験者
1回目	議論のまとめを要求する発話	bot	1	低資源言語話者 1
				中国語話者 1
				日本語話者 3
				日本語話者 4
2回目	言い換えを求める発話		2	低資源言語話者 1
				中国語話者 1
				日本語話者 2
				日本語話者 4
3回目	低資源言語話者の発話に反応を返す発話		3	低資源言語話者 1
				中国語話者 1
				日本語話者 1
				日本語話者 4

表 5：低資源言語話者 2 の実験設定（グループ②）

	条件	実行方法	実験タスク	被験者
1回目	議論のまとめを要求する発話	bot	3	低資源言語話者 2
				中国語話者 2
				日本語話者 1
				日本語話者 2
2回目	言い換えを求める発話		4	低資源言語話者 2
				中国語話者 2
				日本語話者 1

3回目	低資源言語話者の発話に反応を返す発話	5	日本語話者 3
			低資源言語話者 2
			中国語話者 2
			日本語話者 3
			日本語話者 2

表 6：低資源言語話者 3 の実験設定（グループ③）

	条件	実行方法	実験タスク	被験者
1回目	肯定的な反応を返す発話	WOZ 法	1	低資源言語話者 3
				中国語話者 3
				日本語話者 7
				日本語話者 8
2回目	議論のまとめを要求する発話	bot	2	低資源言語話者 3
				中国語話者 3
				日本語話者 7
				日本語話者 5
3回目	コミュニケーションなし	-	5	低資源言語話者 3
				中国語話者 3
				日本語話者 7
				日本語話者 3

表 7：低資源言語話者 4 の実験設定（グループ④）

	条件	実行方法	実験タスク	被験者
1回目	コミュニケーションなし	-	3	低資源言語話者 4
				中国語話者 4
				日本語話者 6
				日本語話者 3
2回目	肯定的な反応を返す発話	WOZ 法	4	低資源言語話者 4
				中国語話者 4
				日本語話者 6
				日本語話者 8

3回目	言い換えを求め る発話	bot	2	低資源言語話者 4
				中国語話者 4
				日本語話者 6
				日本語話者 9

表 8：低資源言語話者 5 の実験設定（グループ⑤）

	条件	実行方法	実験タスク	被験者
1回目	低資源言語話者 の発話に反応を 返す発話	bot	1	低資源言語話者 5
				中国語話者 5
				日本語話者 9
				日本語話者 5
2回目	コミュニケーションなし	-	2	低資源言語話者 5
				中国語話者 5
				日本語話者 9
				日本語話者 3
3回目	肯定的な反応を 返す発話	WOZ 法	3	低資源言語話者 5
				中国語話者 5
				日本語話者 5
				日本語話者 8

5.2 相互作用の効果

実験中、ファシリテータエージェントがどの程度発話したか、そして期待した反応があったかについて評価した。

議論のまとめを要求する発話については、以下の表 9 の結果が得られた。動作の正確性はいずれも 100%であるが、作用の効果はグループ①が低い値になっている。これは、プログラム通りファシリテータエージェントがシステムメッセージを送信したにも関わらず、被験者がそれに応じず無視したためである。これを受けて、実験 2 日目では、システムメッセージを「今までの議論をおさらいしましょう」「いったん話をまとめましょう」から「〇〇さん、今何について話しているか教えてください」「□□さん、今までの議論をおさらいしてください」「△△さん、今何について話していますか？」に変更した。その他の言語話者を名指しすることで、システムメッセージに応じやすくなり、実際に実験 2 日目の

グループ③では名指しされた被験者がシステムメッセージに応じ、作用の効果を
得ることができた。

なお、グループ②については、不具合により有意な結果を得ることができな
かったため、無効としている。

表 9：議論のまとめを要求する発話の評価

	グループ	実行回数	動作の正確性 (%)	作用の効果 (%)
議論のまとめを 要求する発話	①	6	100	17
	②	-	-	-
	③	1	100	100
平均		3.5	100	58.5

言い換えを求める発話は、以下の表 10 の結果となった。こちらも議論のまと
めを要求する発話と同じく、ファシリテータエージェントは正確に動作した
ものの、グループ②の被験者がシステムメッセージに応じなかったため、作
用の効果を得ることができなかった。同様に実験 2 日目のグループ④では、
システムメッセージを「簡単な言葉でまとめましょう」「簡潔に言い換えてく
ださい」から「〇〇さん、今の発言を簡単な言葉で言い換えてもらえますか？」
「□□さん、今の発言を簡潔に言い換えてもらえますか？」「△△さん、今
の発言を簡単に言い換えてもらえますか？」に変更した。しかし、グループ④で
は実験中に実行条件が満たされず、検証することができなかった。

表 10：言い換えを要求する発話の評価

	グループ	実行回数	動作の正確性 (%)	作用の効果 (%)
言い換えを要求 する発話	①	0	100	0
	②	1	100	0
	④	0	100	0
平均		0.33	100	0

低資源言語話者の発話に反応を返す発話は以下の表 11 の結果となった。これ
は、他者に直接発話を求めるものではないため、作用の効果は算出していな
い。

表 11：発話に反応を返す発話の評価

	グループ	実行回数	動作の正確性(%)
発話に反応を返す発話	①	3	100
	②	8	100
	⑤	3	100
平均		4.67	100

表 12：肯定的な反応を返す発話の評価

	グループ	実行回数
肯定的な反応を返す発話	③	6
	④	6
	⑤	3
平均		5

低資源言語話者が意見を出した際に肯定的な反応を返す発話は以上の表 12 の結果となった。これは、WOZ 法によって実行するものであり、他者に直接発話を求めるものではないため、動作の正確性や作用の効果は算出していない。

5.3 発話の評価

客観的評価に基づいて、各条件によって低資源言語話者の発話数にどのような変化があったかを分析する。

以下の表は、その結果の平均値を条件ごとにまとめたものである。

表 13：客観的評価の平均

	低資源言語話者が発話した回数	低資源言語話者が発話するまでにかかった時間(分)	低資源言語話者が意見を発話した回数	低資源言語話者が発話した文字数	その他の言語話者が低資源言語話者に対して声掛けをした回数
議論のまとめを要求する発話	4	3.14	4	333	0.67
言い換えを求める発話	4.67	2.72	3	332	1.11
反応を返す発話	4.33	3.32	2	274	0.56
肯定的な反応を返す発話	5.33	2.81	4	356	0.78
コミュニケーションなし	4.33	2.98	3.33	415	0.78

これらの値について、一元配置分散分析を行った。しかし、有意性のある項目は無かった。

5.4 主観的評価

主観的評価は、被験者に対するアンケートで行った。低資源言語話者には、以下の 6 項目について、その他の言語話者には以下の 5 項目について 5 段階で評価をしてもらった。数値は、いずれも被験者の回答の平均を示している。

表 14：低資源言語話者の主観評価の平均

	議論の内容を理解できたか	他の参加者の意見を理解できたか	全体を通して話しやすいと感じたか	自分の意見を他の参加者に伝えることができたか	ファシリテーターからのコミュニケーションが発話のきっかけになったか	議論に参加しているという実感があつたか
議論のまとめを要求する発話	4.5	4.5	4.5	4.5	4.5	5
言い換えを求める発話	4.67	4.67	4	4.67	2.67	4.33
反応を返す発話	4.67	4.33	4.67	4.33	4	4.67
肯定的な反応を返す発話	5	4.67	4.33	4.33	4	4.67
コミュニケーションなし	5	4.67	4.33	4.67	2	4.67

表 15：その他の言語話者の主観評価の平均

	インドネシア人の参加者が話を理解できていると感じたか	話しづらそうな人がいると感じたか	他の参加者が議論に参加できるようにサポートできたか	自分の意見をわかりやすく伝えることができたか	全員が議論に参加していると感じたか
議論のまとめを要求する発話	4.67	4	4	4.5	3.83
言い換えを求める発話	4.33	3.11	3.78	3.89	3.89
反応を返す発話	4.33	3.67	3.78	4.22	3.89
肯定的な反応を返す発話	4.56	4	3.22	4.33	4.33
コミュニケーションなし	4	2.89	3.11	4	3.33

これらの結果について、クラスカル・ウォリス検定を行った。その結果、低資源言語話者は、「ファシリテーターからのコミュニケーションが発話のきっかけになったか」という項目についての P 値が **0.047100756523030654** となり、5%有意水準で有意性があつた。その他の言語話者のアンケートについては、同じ有意

水準で有意性のある項目は無かったが、10%有意水準では「話しづらそうな人がいると感じたか」について有意性があった。

第6章 考察

議論のまとめを要求する発話について、実験 1 日目では被験者がシステムメッセージに応じることが少なかったが、システムメッセージを「今までの議論をおさらいしましょう」「いったん話をまとめましょう」から「〇〇さん、今何について話しているか教えてください」「□□さん、今までの議論をおさらいしてください」といったものに変更することで被験者が応じるようになった。これは、名指しで要求することによって、無視しづらくなったためだと考えられる。そのため、ファシリテータエージェントの発話内容としては、応じなければいけないという義務感がある程度生じる内容にするべきであると考えられる。

言い換えを求める発話については、実行回数が少なく、実行されてもシステムメッセージに対して被験者が応じない事例があった。そのため、その有効性を検証することができなかった。これは、実験タスク自体が長い発話を必要とするものではなく、難易度が適切でなかったことが原因であると考えられる。そのため、もう少し難しい内容でのグループディスカッションか、もしくはグループディスカッション以外での実験を通して、この条件の有効性を検証する必要があると考えられる。

第7章 おわりに

本研究では、多言語コミュニケーションにおける機械翻訳の精度によるコミュニケーション問題を解消するために、低資源言語話者のコミュニケーション活性化に効果的であると考えられる発話を定義した。そして、それらを多言語チャットシステム LangridChat に実装し、グループディスカッションによる被験者実験を通してその効果を検証・分析した。本研究の貢献は以下の通りである。

ファシリテータエージェントの発話の定義

低資源言語話者の理解促進を目的として、議論のまとめを要求する発話と言い換えを求める発話を定義した。更に、低資源言語話者の発話促進を目的として、発話に反応を返す発話と低資源言語話者が意見を出した際に肯定的な反応を返す発話を定義した。実験の結果、主観的評価では「ファシリテータからのコミュニケーションが発話のきっかけになったか」について5%有意水準で有意性があった。客観的評価では、有意性のある結果は得られなかった。

ファシリテータエージェントの発話の実装

議論のまとめを要求する発話は、低資源言語話者がメッセージを配信した時間を観測し、一定時間ごとに実行するように実装した。言い換えを求める発話は、サーバに配信された他者の発話メッセージを観測し、閾値より長い場合に実行され、発話に反応を返す発話は、サーバに配信されたメッセージを観測し、送信者が低資源言語話者であった場合に実行するように実装した。実験中、ファシリテータエージェントはいずれも100%正しく動作した。しかし、言い換えを求める発話については、実験タスクが適切ではなかったため、その有効性を検証することができなかった。

謝辞

本研究を行うにあたり，指導教官の村上陽平准教授から熱心なご指導，ご助言を賜りました．深謝申し上げます．また，LangridChat システムに関して多大なるご協力を賜りました社会知能研究室の張禹王さん，実験にご協力頂いた Mondheera Pituxcoosuvorn 助教，被験者の皆様に厚く感謝申し上げます．

参考文献

- [1] Mizuki Motozawa, Yohei Murakami, Mondheera Pituxcoosuvann, Toshiyuki Takasaki, Yumiko Mori: Conversation Analysis for Facilitation in Children's Intercultural Collaboration, in Proc. of the 20th ACM Interaction Design and Children (IDC 2021), pp. 62-68, (2021).
- [2] 石田真也, 井上昇治, 中村静, 高梨克也, 河原達也: 共感表出と発話促進のための聞き手応答を生成する傾聴対話システム, 人工知能学会資料, Vol. 82, pp. 7-12, (2018).
- [3] 石田真也, 井上昇治, 高梨克也, 河原達也: 共感・発話促進のための多様な聞き手応答を生成する傾聴対話システム, 第80回全国大会講演論文集, pp. 403-404, (2018).
- [4] 上里美樹, 吉野幸一郎, 高梨克也, 河原達也: 傾聴対話における相槌の韻律的特徴の同調傾向の分析, 言語・音声理解と対話処理研究会, Vol. 70, pp. 7-13, (2014).
- [5] 山口貴史, 井上昇治, 吉野幸一郎, 高梨克也, Nigel G. Ward, 河原達也: 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成, 人工知能学会研究会資料, Vol. 76, pp. 49-54, (2016)
- [6] 石田真也, 井上昇治, 中村静, 高梨克也, 河原達也: 傾聴対話システムのための発話を促す聞き手応答の生成, 人工知能学会研究会資料, Vol. 77, No. 01, pp. 1-6, (2016).
- [7] 伝康晴: 対話への情報付与, 日本語コーパス 3: 話し言葉コーパス-設計と構築-, pp. 101-130, (2015).