

# 博士論文

## 機械翻訳を介したコミュニケーションにおける ファシリテーションによる 低資源言語話者の参与支援 (Enhancing Low-Resourced Language Speakers' Involvement through Facilitation in MT-Mediated Communication)

2025 年 3 月

立命館大学大学院 情報理工学研究科  
情報理工学専攻 博士課程後期課程

元澤 海月



立命館大学審査博士論文

機械翻訳を介したコミュニケーションにおける  
ファシリテーションによる  
低資源言語話者の参与支援  
(Enhancing Low-Resourced Language Speakers'  
Involvement through Facilitation in  
MT-Mediated Communication)

2025 年 3 月

March 2025

立命館大学大学院 情報理工学研究科  
情報理工学専攻 博士課程後期課程  
Doctoral Program in Advanced  
Information Science and Engineering  
Graduate School of Information Science and Engineering  
Ritsumeikan University

元澤 海月  
MOTOZAWA Mizuki

研究指導教員：村上 陽平教授  
Supervisor : Professor MURAKAMI Yohei



# 内容梗概

持続可能な開発目標 (SDGs: Sustainable Development Goals) の1つに、「質の高い教育をみんなに」がある。ここでは、グローバル化が進展する現代社会において、文化の多様性を尊重する態度の重要性が主張されている。このような態度を涵養するため、機械翻訳を用いた異文化コラボレーションが行われている。異文化コラボレーションでは、文化や言語の異なる相手と相互理解を試みる体験が可能となる。このようなコラボレーションでは、共通言語を持たない場合もあり、機械翻訳をもちいたコミュニケーションも盛んに行われている。また、近年のニューラル機械翻訳による翻訳精度の大幅な向上によって、機械翻訳を用いた多言語でのコミュニケーションは円滑化している。実際に行われている異文化コラボレーションでは、機械翻訳を用いた多言語の話者同士による議論が行われている。しかしながら、そこに言語資源が少なく翻訳精度が低い言語 (低資源言語) 話者が参加する場合、この話者にとっては、議論内容を理解することが困難なため、発言が少ないことが報告されている。翻訳精度の向上には、言語資源を拡大する必要があるが、膨大なコストや労力がかかるため、困難である。

そこで、本研究では、議論のファシリテーションに着目し、低資源言語話者による議論の理解を支援することで、多言語コミュニケーションへの参与を促進することを目的とする。そのために、まずはフィールド調査を行い、機械翻訳を介したファシリテーションが低資源言語話者の議論への参与に与える影響を分析した。次に、低資源言語話者の議論の理解を支援するために、議論の内容を要約するファシリテーターエージェントを構築した。本論文で取り組んだ課題は以下の二つである。

## 1. 異文化コラボレーションにおけるファシリテーターの対話行為分析

機械翻訳を用いた多言語での話者達による議論の場では、ファシリテータの対話行為が聞き手の応答に作用する。そこで、低資源言語話者の発話を促進させるために、実際の異文化コラボレーションの場におけるファシリテータの対話行為を分析した。まず、発話を定量的に分析するために、既存の対話行為に基づく発話アノテーション手法を用いて、ファシリテータの発話を分類した。既存の汎用的な対話行為の分類では、ファシリテーションに固有の参加者に投げかける質問や、グループワークを進行するための発話の区別ができないため、各分類の発話例の追加や区別が難しい分類の区別方法を示すことでタグを詳細化し、アノテーションスキーマを定義した。その後、ファシリテータの各対話行為への応答数を比較することで、対話行為による応答への効果を分析した。その結果、日本語話者に対しては、「依頼」の対話行為が応答を促進した。一方、低資源言語であるクメール語話者に対しては、促進する対話行為が明らかにならなかったものの、日本語話者の応答を促進する「依頼」が機械翻訳によって翻訳は正しいが、他の対話行為に変化することを統計的に明らかにし、これが低資源言語話者の応答を阻害する要因となりうることを発見した。

## 2. ファシリテータエージェントによる異文化コラボレーション支援

ファシリテータが低資源言語話者の応答を促進させるために有効的な対話行為を行ったとしても、対話内容を理解できていなければ発話することは困難である。このことから、低資源言語話者の対話理解を促進させるため、要約に着目した2種類のファシリテータエージェントを実装した。具体的には、高資源言語話者に要約を要求するエージェントと、要約を提供するエージェントである。これらのエージェントを用いたコミュニケーション実験を通じて、低資源言語話者の対話の理解度、参加度、満足度を向上させるかの有効性を検証した。また、要約要求エージェントを用いた場合は、高資源言語話者が要約を行うため、負担が大きくなる可能性がある。そのため、このエージェントを用いた場合の作業負荷や低資源言語話者の配慮度にどう影響するかも検証した。

その結果、要約要求エージェントは、高資源言語話者からの要約に含まれる情報の少なさや、エージェントからの要約要求に応じない場合があったため、低資源言語話者の理解度の向上は認められなかった。さらに、議論の参加度や満足度の向上も認められなかった。また、高資源言語話者への効果については、不満を増加させるものの、高資源言語話者による配慮発話を増加させることで、低資源言語話者のグループ議論への関与を高める可能性があることが示された。次に、要約提供エージェントは、低資源言語話者の議論の参加度や満足度の向上は見られなかったものの、理解した議論項目が増加する効果が明らかになった。



# 謝辞

担当教授である村上陽平教授には、本研究を進めるにあたって、学部からの7年間に渡って多大なるご指導とご助言を頂きまして、心より感謝申し上げます。また、当研究室の Mondheera Pituxcoosuvann 助教には、研究活動におけるご助言や、英語論文執筆や発表資料作成の際は、英文の確認や改善についてご助言を頂きまして、心より感謝申し上げます。

特定非営利活動法人 パンゲアからは研究に必要なデータの提供していただき、理事長の森由美子様、副理事長の高崎俊之様には常に親切なご助言や激励を頂きまして、厚く感謝申し上げます。

京都情報大学院大学の Amit Pariyar 准教授には、コミュニケーション実験のネパール語話者の被験者募集にご助力を頂きまして、厚く感謝申し上げます。

社会知能研究室の荒巻光里さんには、コミュニケーション実験の際、実験準備や運営の協力など、多大なる助力を頂きました。また、研究室の西村一球さんにも、論文執筆における技術的な助言や、研究活動における助言など、多大なる助力を頂きました。

最愛の両親には、博士課程を修了するために、今まで温かい支援と励ましを頂いたことに、心より感謝いたします。

本研究は、日本学術振興会科学研究費基盤研究 (B)(21H03561, 令和3年度～令和5年度, 23K21730, 令和6年度, 21H03556, 令和3年度～令和5年度) 及び日本学術振興会科学研究費若手研究 (21K17794, 令和3年度～令和6年度) の支援を受けた。



# 目次

|            |                                                    |           |
|------------|----------------------------------------------------|-----------|
| <b>第1章</b> | <b>序論</b>                                          | <b>1</b>  |
| 1.1        | 研究背景 . . . . .                                     | 1         |
| 1.2        | 異文化コラボレーションの課題 . . . . .                           | 1         |
| 1.3        | 研究目的 . . . . .                                     | 2         |
| 1.4        | 研究の位置付け . . . . .                                  | 2         |
| 1.5        | 論文構成 . . . . .                                     | 3         |
| <b>第2章</b> | <b>多言語コミュニケーション支援</b>                              | <b>5</b>  |
| 2.1        | 非母国語話者のコミュニケーション支援 . . . . .                       | 5         |
| 2.2        | 機械翻訳を用いたコミュニケーション支援 . . . . .                      | 6         |
| 2.3        | 異文化コラボレーション支援 . . . . .                            | 8         |
| 2.4        | 異文化コラボレーションのフィールド . . . . .                        | 9         |
| 2.4.1      | Kyoto International Summer School for Youth(KISSY) | 9         |
| 2.4.2      | 多言語チャットツール . . . . .                               | 10        |
| 2.4.3      | ファシリテーション . . . . .                                | 11        |
| <b>第3章</b> | <b>ファシリテーション分析のための発話アノテーション手法</b>                  | <b>13</b> |
| 3.1        | 対話行為の分析の課題 . . . . .                               | 13        |
| 3.2        | 対話アノテーション . . . . .                                | 13        |
| 3.3        | 予備実験 . . . . .                                     | 15        |
| 3.4        | ファシリテータの対話行為タグ . . . . .                           | 16        |
| 3.4.1      | 必要条件 . . . . .                                     | 17        |
| 3.4.2      | ISO 24617-2 (SemAF) . . . . .                      | 17        |
| 3.5        | ファシリテータの発話へのアノテーション . . . . .                      | 19        |
| 3.5.1      | アノテーションの曖昧さ . . . . .                              | 19        |
| 3.5.2      | 曖昧なアノテーションタグの区別の明確化 . . . . .                      | 20        |

|            |                                      |           |
|------------|--------------------------------------|-----------|
| 3.5.3      | アノテーションフォームデザイン . . . . .            | 21        |
| 3.6        | アノテーション実験 . . . . .                  | 22        |
| 3.6.1      | 実験環境 . . . . .                       | 22        |
| 3.6.2      | 評価 . . . . .                         | 23        |
| 3.7        | 考察 . . . . .                         | 30        |
| 3.8        | まとめ . . . . .                        | 32        |
| <b>第4章</b> | <b>ファシリテータの対話行為分析</b>                | <b>33</b> |
| 4.1        | 対話行為に着目したファシリテーション . . . . .         | 33        |
| 4.2        | 機械翻訳を活用した多言語間のコミュニケーション . . . . .    | 34        |
| 4.2.1      | 対話行為に着目したコミュニケーション支援 . . . . .       | 34        |
| 4.2.2      | 機械翻訳の現状 . . . . .                    | 34        |
| 4.2.3      | リサーチクエスション . . . . .                 | 35        |
| 4.3        | 多言語コミュニケーション対話ログ . . . . .           | 36        |
| 4.3.1      | コミュニケーション環境 . . . . .                | 36        |
| 4.3.2      | 対話ログ . . . . .                       | 36        |
| 4.4        | 発話アノテーション . . . . .                  | 37        |
| 4.4.1      | 対話行為に基づいたタグセット . . . . .             | 37        |
| 4.4.2      | アノテーション方法 . . . . .                  | 38        |
| 4.4.3      | 隣接応答ペアの抽出 . . . . .                  | 39        |
| 4.5        | RQ1: 各対話行為への児童の応答数の比較 . . . . .      | 40        |
| 4.5.1      | 対話行為への児童の応答数の集計 . . . . .            | 42        |
| 4.5.2      | 統計的検定 . . . . .                      | 42        |
| 4.6        | RQ2: 児童の各使用言語間での応答数比較 . . . . .      | 45        |
| 4.6.1      | 統計的検定 . . . . .                      | 46        |
| 4.7        | RQ3: 翻訳前後の対話行為比較 . . . . .           | 46        |
| 4.7.1      | 使用言語ごとの受け取る対話行為数の比較 . . . . .        | 47        |
| 4.7.2      | 機械翻訳が介入する言語ペアの特定 . . . . .           | 49        |
| 4.7.3      | 各言語ペアにおける対話行為の組み合わせパターンの集計 . . . . . | 49        |
| 4.7.4      | マクネマー検定 . . . . .                    | 50        |
| 4.8        | 考察 . . . . .                         | 52        |
| 4.9        | まとめ . . . . .                        | 56        |

|            |                                          |            |
|------------|------------------------------------------|------------|
| <b>第5章</b> | <b>ファシリテータエージェントを用いた異文化コラボレーション支援</b>    | <b>57</b>  |
| 5.1        | 要約エージェントによる対話理解促進 . . . . .              | 57         |
| 5.2        | ファシリテータエージェントを用いたコミュニケーション支援研究 . . . . . | 58         |
| 5.2.1      | 仮説 . . . . .                             | 58         |
| 5.3        | ファシリテータエージェント . . . . .                  | 60         |
| 5.3.1      | システム構成 . . . . .                         | 60         |
| 5.3.2      | 要約要求エージェント . . . . .                     | 62         |
| 5.3.3      | 要約提供エージェント . . . . .                     | 64         |
| 5.4        | 実験環境 . . . . .                           | 65         |
| 5.4.1      | 概要 . . . . .                             | 65         |
| 5.4.2      | 被験者 . . . . .                            | 65         |
| 5.4.3      | タスク . . . . .                            | 66         |
| 5.5        | 分析 . . . . .                             | 69         |
| 5.5.1      | 議論の理解度の評価 . . . . .                      | 69         |
| 5.5.2      | 議論への参加度の評価 . . . . .                     | 74         |
| 5.5.3      | 議論の満足度の評価 . . . . .                      | 80         |
| 5.5.4      | 作業負荷の評価 . . . . .                        | 82         |
| 5.5.5      | 低資源言語話者への配慮度の評価 . . . . .                | 84         |
| 5.6        | 考察 . . . . .                             | 87         |
| 5.7        | まとめ . . . . .                            | 93         |
| <b>第6章</b> | <b>結論</b>                                | <b>95</b>  |
| 6.1        | 本研究のまとめ . . . . .                        | 95         |
| 6.2        | 今後の展望 . . . . .                          | 96         |
|            | <b>参考文献</b>                              | <b>99</b>  |
|            | <b>研究業績</b>                              | <b>111</b> |



# 目次

|     |                                                         |    |
|-----|---------------------------------------------------------|----|
| 2.1 | グループワークの様子 (写真提供：NPO 法人 パンゲア)                           | 10 |
| 2.2 | 多言語チャットツール (写真提供：NPO 法人 パンゲア)                           | 11 |
| 3.1 | アノテーションのための Excel フォーム                                  | 22 |
| 3.2 | 各アノテータによる基本的な発話分類のタグ付け回数の分布                             | 27 |
| 3.3 | 各アノテータによる指示の細分類のタグ付け回数の分布                               | 29 |
| 3.4 | 各アノテータによる修辞関係のタグ付け回数の分布                                 | 29 |
| 3.5 | 各アノテータによる宛先のタグ付け回数の分布                                   | 30 |
| 3.6 | 各分類の確信度と Kappa 係数の相関図                                   | 30 |
| 4.1 | 日本語話者の各対話行為への応答数と無応答数の集計表                               | 43 |
| 4.2 | 英語話者の各対話行為への応答数と無応答数の集計表                                | 43 |
| 4.3 | クメール話者の各対話行為への応答数と無応答数の集計表                              | 43 |
| 4.4 | 「情報探索」への各言語話者の応答数と無応答数の集計表                              | 47 |
| 4.5 | 「依頼」への各言語話者の応答数と無応答数の集計表                                | 47 |
| 4.6 | 「命令」への各言語話者の応答数と無応答数の集計表                                | 48 |
| 4.7 | 「提案」への各言語話者の応答数と無応答数の集計表                                | 48 |
| 4.8 | 各対話行為の出現数                                               | 50 |
| 5.1 | Langrid Chat の UI 画面                                    | 60 |
| 5.2 | Langrid Chat のシステム構成図                                   | 61 |
| 5.3 | ファシリテータエージェントの処理                                        | 63 |
| 5.4 | ネパール語話者の議論の理解できたアイテム数の集計表<br>(要約要求エージェント条件とエージェントなし条件間) | 71 |

|      |                                                         |    |
|------|---------------------------------------------------------|----|
| 5.5  | ネパール語話者の議論の理解できたアイテム数の集計表<br>(要約提供エージェント条件とエージェントなし条件間) | 71 |
| 5.6  | 言語話者ごとの議論の理解できたアイテム数の集計表(要<br>約要求エージェント条件) . . . . .    | 72 |
| 5.7  | 言語話者ごとの議論の理解できたアイテム数の集計表(要<br>約提供エージェント条件) . . . . .    | 73 |
| 5.8  | 言語話者ごとの議論の理解できたアイテム数の集計表<br>(エージェントなし条件) . . . . .      | 73 |
| 5.9  | ネパール語話者の条件間における議論理解度 Q2 の箱ひ<br>げ図 . . . . .             | 73 |
| 5.10 | ネパール語話者の条件間における議論理解度 Q3 の箱ひ<br>げ図 . . . . .             | 74 |
| 5.11 | ネパール語話者の条件間における議論理解度 Q4 の箱ひ<br>げ図 . . . . .             | 74 |
| 5.12 | ネパール語話者の議論に参加できたアイテム数の集計表<br>(要約要求エージェント条件とエージェントなし条件間) | 76 |
| 5.13 | ネパール語話者の議論に参加できたアイテム数の集計表<br>(要約提供エージェント条件とエージェントなし条件間) | 76 |
| 5.14 | ネパール語話者の条件間における参加度 Q2 の箱ひげ図                             | 77 |
| 5.15 | ネパール語話者の条件間における参加度 Q3 の箱ひげ図                             | 77 |
| 5.16 | ネパール語話者の条件間における 1 タスクの文字数偏差<br>の箱ひげ図 . . . . .          | 79 |
| 5.17 | ネパール語話者の条件間における 1 タスクの発話数の箱<br>ひげ図 . . . . .            | 79 |
| 5.18 | ネパール語話者の条件間における満足度 Q1 の箱ひげ図                             | 81 |
| 5.19 | ネパール語話者の条件間における満足度 Q2 の箱ひげ図                             | 81 |
| 5.20 | ネパール語話者の条件間における客観的満足度の箱ひげ図                              | 82 |
| 5.21 | 日本語話者の条件間における作業負荷の箱ひげ図 . . .                            | 83 |
| 5.22 | 日本語話者の条件間における 6 つの負荷の箱ひげ図 . .                           | 84 |
| 5.23 | 日本語話者の条件間における低資源言語話者への主観的<br>配慮度の箱ひげ図 . . . . .         | 85 |
| 5.24 | 日本語話者の条件間における低資源言語話者への客観的<br>配慮度の箱ひげ図 . . . . .         | 87 |

# 表 目 次

|      |                                                   |    |
|------|---------------------------------------------------|----|
| 3.1  | Searle の 5 つの発話分類 . . . . .                       | 15 |
| 3.2  | Vanderveken の遂行動詞 (行為指示型の細分類) . . . . .           | 16 |
| 3.3  | 修飾関係のタグ分類 . . . . .                               | 18 |
| 3.4  | ISO24617-2 に対応づけた対話行為のタグセット . . . . .             | 19 |
| 3.5  | ISO24617-2 に対応づけた指示の細分類 . . . . .                 | 20 |
| 3.6  | 対話行為分類のタグ付け数 . . . . .                            | 23 |
| 3.7  | 指示の細分類のタグ付け数 . . . . .                            | 24 |
| 3.8  | 修飾関係のタグ付け数 . . . . .                              | 24 |
| 3.9  | 宛先のタグ付け数 . . . . .                                | 24 |
| 3.10 | 基本的な 5 つの対話分類の Fleiss' kappa 係数 . . . . .         | 25 |
| 3.11 | 指示の細分類の Fleiss' kappa 係数 . . . . .                | 25 |
| 3.12 | 修飾関係の Fleiss' kappa 係数 . . . . .                  | 25 |
| 3.13 | 宛先の Fleiss' kappa 係数 . . . . .                    | 26 |
| 3.14 | 修正後の修飾関係分類タグ . . . . .                            | 32 |
| 4.1  | ログデータの会話事例 . . . . .                              | 35 |
| 4.2  | グループの構成, 使用言語内訳 . . . . .                         | 37 |
| 4.3  | ログデータの会話事例 . . . . .                              | 38 |
| 4.4  | 基本的な 5 つの対話行為分類の定義 . . . . .                      | 40 |
| 4.5  | 「指示」の 4 つの細分類の定義 . . . . .                        | 41 |
| 4.6  | ファシリテータの発話で統合された発話事例 . . . . .                    | 41 |
| 4.7  | ジョージア国籍とカンボジア国籍の英語話者児童の各対<br>話行為への応答数集計 . . . . . | 45 |
| 4.8  | 日本語から英語での出現数の集計表 . . . . .                        | 50 |
| 4.9  | 英語から日本語での出現数の集計表 . . . . .                        | 51 |
| 4.10 | 英語からクメール語での出現数の集計表 . . . . .                      | 53 |

|                                               |    |
|-----------------------------------------------|----|
| 4.11 集計表の一例 . . . . .                         | 54 |
| 4.12 マクネマー検定結果 . . . . .                      | 55 |
| 5.1 エージェントの要約要求メッセージ . . . . .                | 62 |
| 5.2 議論の理解度に関する質問項目 . . . . .                  | 70 |
| 5.3 議論の参加度に関する質問項目 . . . . .                  | 75 |
| 5.4 各シナリオの1タスクあたりのネパール語話者の平均文<br>字数 . . . . . | 78 |
| 5.5 各シナリオの1タスクあたりのネパール語話者の平均発<br>話数 . . . . . | 79 |
| 5.6 議論の満足度に関する質問項目 . . . . .                  | 80 |
| 5.7 議論の配慮に関する質問項目 . . . . .                   | 85 |
| 5.8 日本語話者によるネパール語話者への配慮する発話の種類                | 86 |
| 5.9 対話事例1 (要約提供エージェント/北極サバイバル) . .            | 88 |
| 5.10 対話事例2 (エージェントなし/海洋サバイバル) . . . .         | 89 |
| 5.11 要約例1 (要約提供エージェント/無人島サバイバル) . .           | 91 |
| 5.12 要約例2 (要約要求エージェント/北極サバイバル) . . .          | 91 |

# 第1章 序論

## 1.1 研究背景

グローバル化が発展する現代社会において、文化の多様性を尊重する態度が重要である。2015年に実施された国連サミットで採択された「持続可能な開発目標 (SDGs)」の一つである「質の高い教育をみんなに」にも、文化の多様性を理解し、受け入れる態度を養うことの必要性が言及されている。また、そのような態度を涵養するため、言語や文化の異なる相手との相互理解を深めようと試みる経験が重要である。そのような相手とのコミュニケーションでは、共通言語を持たない場合や、共通言語を用いた場合でも母国語よりもコミュニケーションが難しくなる。そのため、機械翻訳を用いることが有用である。また、近年ではニューラル機械翻訳の登場によって、翻訳精度も飛躍的に向上し、異なる言語間のコミュニケーションはさらに円滑化している。

## 1.2 異文化コラボレーションの課題

実際の異文化コラボレーションの場では、機械翻訳を用いた多言語コミュニケーションが行われている。ここでは、多言語の話者たちがグループで議論を行なっている。このような場において、低資源言語と呼ばれる言語資源の乏しい言語話者の発話が少ないことが報告されている。低資源言語は、言語資源の少なさから翻訳精度が低く、対話に介入することが難しい。しかし、言語資源を拡張することは、コストや労力が膨大にかかるため、困難である。また、機械翻訳による翻訳精度を向上させたとしても、一般的な会話や文章でほとんど使われないような希少語の翻訳は正しく翻訳できない場合や、入力された文章

とは完全に無関係な内容を出力する場合があります，異文化コラボレーションにおける相互理解の課題は残されている [30].

### 1.3 研究目的

実際の異文化コラボレーションの場の一つでは，異なる言語の話者達が集まり，機械翻訳を組み込んだチャットツールを用いてグループで議論を行う．グループ内で議論を活性化させるため，アイデアや意見を引き出すことを行うファシリテータは重要な役割を担う．そのため，本研究ではファシリテーションに着目し，機械翻訳を用いた異文化コラボレーションの場を支援するためのファシリテーションの特定を目的とする．具体的には，実際の機械翻訳を用いて行われた多言語のチャットコミュニケーションにおけるファシリテータの発話を分析し，発話を促進するために有効なファシリテーションを明らかにする．さらに，有効的なファシリテーションを行ったとしても，対話の理解が乏しいと発言が難しい．そのため，対話理解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．

### 1.4 研究の位置付け

低資源言語話者の異文化コラボレーション支援の研究では，低資源言語話者の発言数が少なさが報告されている [45, 67] が，具体的なアプローチ方法については研究されていない．また，ファシリテータエージェントの研究では，同一言語間や非同期コミュニケーションに着目した支援や，ファシリテータを支援するためのエージェントなどがある．一方で，本研究では機械翻訳を用いた多言語間の同期コミュニケーションに着目した，低資源言語話者を支援するためのエージェントの研究である．

## 1.5 論文構成

本論文は次のように構成されている。第2章では、多言語コミュニケーション支援について紹介する。第3章で、本研究で提案するファシリテータの対話行為を分析するためのアノテーションスキーマについて説明する。第4章で、有効なファシリテータの対話行為の分析について説明する。第5章で、議論の要約に着目したファシリテータエージェントの実装と検証について説明する。最後に、第6章で結論を述べる。



## 第2章 多言語コミュニケーション支援

### 2.1 非母国語話者のコミュニケーション支援

グローバル化やデジタル化が発展する現代社会において、母国語が異なる話者同士のコラボレーションの機会が増加している。このような場合、共通言語を用いてコミュニケーションすることが多い [25]。特に、英語を共通言語として使用することが多い [13]。このようなコミュニケーションのうち、母国語話者 (NS) と非母国語話者 (NNS) 間のコミュニケーションの場合、NNS は NS ほど流暢には話すことができないため、NS 同士のコミュニケーションでは見られない問題が生じる [32]。そのため、NNS を支援するための研究が多く行われている。

例えば、NNS の対話を理解を支援する研究である。NS と NNS の多人数での音声会議での会話において、NS の方が多い場合は議論が白熱するほど会話が早くなる傾向がある [6, 10]。しかし、NS の発話速度が速くなるほど、NNS は会話を理解しにくくなり [1, 51]、会話についていけないことが多い [68]。そこで、Duan ら [11] は、自動音声認識ツールを用いてリアルタイムの発話速度を計測し、可視化するツールを提案した。これにより、NS は自分の発話速度を認識し、遅くするように努める。また、Gao ら [16] は、自動音声認識ツールを用いた会話の書き起こしを公開するツールを提案した。これにより、NS は明瞭に話すようになり、コミュニケーションの質に関する評価を向上させた。しかしながら、テキスト化された NNS の発話を会話をしながら読むには負担が大きいため、同時に会話を聞くことや話すことが困難になる。そこで、Echenique ら [12] は、重要部分のみに着目し、NS が重要部分をテキスト化することの効果を検証した。その結果、NNS の理解が促進し

たことを確認した。この重要部分の書き起こしを自動で行うのは、重要部分の検出の精度が低い [57] ため、難しい。このことから、Pan ら [43] は、書き起こされたトランスクリプトの重要部分に NS がハイライトすることを提案し、相互理解を高めることを確認した。

また、NNS が発話することを支援するための研究もある。例えば、NNS は会話の流れの中で、自身の考えを述べる機会を逃すことが多い [52]。そのため、Guo ら [17] は NNS に発話の順番を与えるために会話に参加するエージェントを提案した。これによって、NNS は発話ターンの頻度を向上させることができる。また、Yamashita ら [65, 90] は、発話と発話の間に無音を挿入することの効果を検証した。その結果、NNS は会話を処理するための時間ができ、理解の正確性が向上した。また、笹島ら [77, 78] は、発話を入力すると、それに対する相手の回答を予測するためのツールを提案した。さらに、福島ら [87] は多言語辞書を用いて議論内容を図解化することで NNS の議論理解を支援する、多言語の討論支援システム PaneLive を開発した。

## 2.2 機械翻訳を用いたコミュニケーション支援

異文化コラボレーションで大きな障壁となるのは、言語の壁である。英語などの共通言語を用いたコミュニケーションは、英語が母国語でない言語話者にとっては、新たな言語を学ばなければならないため、負担が大きい。そこで、有用となるのが機械翻訳である。機械翻訳を用いることで、共通言語を持たない話者同士のコミュニケーションが可能になる。しかしながら、機械翻訳の翻訳精度は未だ完璧ではない [92]。例えば、単語の多義性によって起こるコミュニケーション齟齬についての研究である。具体的には、1 対 1 のコミュニケーションの場合、話し手による発話と聞き手による応答が同じ多義語でも異なった語義で翻訳されるという非対称性が起こる。また、会話の中で同じ多義語の翻訳が一貫せずに異なってしまうという非一貫性が起こる [67]。さらに、3 者でのコミュニケーションでは、話し手による発話に多義語が含まれる場合、それが聞き手間で異なる語義に翻訳されてしまい、情報の共有

ができない。また、聞き手同士ではそれを認識することが難しい [66]. そのため、機械翻訳を用いたコミュニケーション支援のための研究が多く行われている。

例えば、コミュニケーションを支援するためのツールについての研究もある。AmiChat[14] は、入力されたメッセージを多言語に翻訳し、複数の言語で表示することができるツールである。このシステムの特徴として、理解できない翻訳後のメッセージがあった場合に、送信者に言い換えを依頼するメッセージを送信するためのボタンが配置されている。他にも、電子掲示板システム TransBBS[15] は、投稿するメッセージを入力する際、それを投稿する前に翻訳リペアを行うためのボタンがある。翻訳リペアとは、翻訳後のメッセージを確認し、それが自分の意図していない内容の場合にメッセージを書き換える作業のことである。この作業により、翻訳精度が改善することが分かっている [38, 84]. この翻訳リペアについて、有効的である一方で、何度も修正を重ねるなどの大きな負担となる場合がある。そこで、山下ら [91] は、原文作成時の有効なルール [5] の教示効果を検証する実験を行った。その結果、原言語の知識レベルによって効果が異なることが明らかになった。さらに、翻訳リペアは、ユーザの判定により修正されるため、修正が不十分なまま送信されてしまうという問題も発生する。そのため、宮部ら [83] は、修正が不十分なメッセージが送信された場合に、送信者に修正を促すシステムを提案した。また、林田ら [19] は、翻訳リペア支援機能として、折り返し翻訳機能と、強調表示機能を定義し、有効性を確認した。折り返し翻訳とは、ユーザが翻訳先の言語を理解できず、翻訳リペアができない場合に、再度ユーザの言語に再翻訳することでおおむねの翻訳後の意図を確認することができる [70] 機能である。AnnoChat[56, 72] は、リアルタイム折り返し翻訳の機能が実装された多言語チャットツールである。さらに、このツールには、チャット中の任意の語句にアノテーションを付与する機能がある。これは、参加者間で異なる文化的背景を持っている場合、正しく翻訳されたメッセージでも異なる解釈をするおそれがあるため、必要に応じてユーザが意味情報を付与し、それを共有することができる。これらの支援ツールはメッセージの原文を修正することや、テキストでの情報補完に着目し、相互理解の向上を目指している。他にも、Langrid Blackboard[21] は、多言語ユーザ間での議論に

において、議論内容を整理するための多言語共有支援ツールである。これは、共有ウィンドウ内にテキストを入力し、それらをグループ化することで議論内容を整理し、共有する。訳してネット [41, 53, 59] は、ドメイン志向翻訳を実現した機械翻訳サイトである。これは、ユーザが独自に専門用語の対話辞書を作成し、既存の機械翻訳を連携させることで翻訳精度の向上を可能にした。船越ら [96] は、自動翻訳機能を備えた電子掲示板を介した異文化間ソフトウェア開発プロジェクトである Intercultural Collaboration Experiment2002(ICE2002)[42, 69, 86] を観察し、多言語での協調作業の支援ツールに必要な要件を定義した。また、吉野ら [71] は、コミュニケーション支援のための Wikipedia のデータを利用した語句の文化差判定手法を提案している。さらに、文化差検出の手法は他にも画像の比較技術を用いた研究 [47, 48, 49] や、言語モデルを用いた研究 [50] がある。

さらに、言語能力が未熟な児童に焦点を当てた研究もある。Matsuda ら [37] は、日本人児童向けの新しい翻訳システムを開発した。

## 2.3 異文化コラボレーション支援

言語の壁を超えて、正しい意図でメッセージを伝えることができたとしても、文化の差によって異なる意味で解釈され、コミュニケーション齟齬が生じる可能性がある。そのような課題を解決するための研究が行われている。例えば、岡本ら [73] は、異文化間のインフォーマルなコミュニケーションにおいて、それぞれが持つ文化的背景が異なることによる相互理解の困難さを支援するためのシステムを開発した。具体的には、音声認識エンジン Julius[88] と形態素解析 MeCab[31] を用いて名詞の抽出を行い、その名詞の画像と説明を Web から取得し、提示することで視覚情報を共有する。

さらに、表情などの視覚的な表現ができるアバタを活用したコミュニケーションはユーザ相互のインタラクションを円滑にする [33, 44, 58]。そこで、神田らは [29, 89]、異文化間コミュニケーションの場において、このアバタの表情や顔つきの解釈に齟齬があるのかの比較検証を行っ

た。また、Cho ら [9, 20] は、日米の児童達の同一の絵文字に対する解釈の違いを調査した。さらに、高崎ら [61] は、多言語対応し、文化的背景を反映させた、様々な国の児童でも利用可能な PictNet というオンラインの絵文字コミュニケーションシステムを提案している。このシステムは実際の児童の異文化間コミュニケーションのツールとして使用されている [94]。宗森ら [80, 81] は、絵文字のみでチャットコミュニケーションを行うシステムを開発し、コミュニケーション実験を重ねている。

他にも、視覚的にコミュニケーションを円滑化する研究があり、宮部 [82] らは、リアルタイムのコミュニケーションを円滑にするために、相手のメッセージを入力する状況を提示することで、相手の応答を待つ時間が長くなることを確認した。

## 2.4 異文化コラボレーションのフィールド

### 2.4.1 Kyoto International Summer School for Youth(KISSY)

NPO 法人 パンゲアは KISSY というサマースクールを毎年開催している。KISSY では、さまざまな国籍の児童たちによる国際問題に対する解決策の議論がグループで行われる [2]。図 2.1 は実際の KISSY の様子である。図に示されている国旗は、各参加者の国籍を示しており、この図の通り、各グループは多国籍である。グループ内には大人のファシリテータも所属し、児童たちによる議論を支援する。



図 2.1: グループワークの様子 (写真提供：NPO 法人 パンゲア)

## 2.4.2 多言語チャットツール

KISSYでは、言語の異なる参加者達がコミュニケーションを行うため、多言語のチャットツールが用いられる。このツールには、言語グリッド [22, 23, 85] が用いられている。言語グリッドとは、機械翻訳とドメイン固有の辞書を組み合わせた翻訳を提供するためのプラットフォームである。言語グリッド上で機械翻訳に構築したパンゲア独自の辞書を連携させることで、パンゲアで活動するときの議論でのメッセージを翻訳する精度が向上する [40]。図 2.2 は実際に KISSY で使われる多言語のチャットツール画面である。画面の上部はファシリテータの名前と顔写真、現在の議論テーマ、ファシリテータからのメッセージが表示される。また、画面の下部分には、各児童の名前、顔写真のアイコン、

国籍，発話したメッセージが表示されている．図の画面は英語ユーザのための画面であり，図に示されるテキストは全て英語で表示される．つまり，英語以外の言語話者によるメッセージは全て英語に翻訳され，表示される．また，このチャットツールでは，ピクトンという絵文字を使用できる．これは，言語の障壁なく伝え合うことを目的に開発された，パンゲア独自のピクトグラムである [39, 62]．さらに，各児童のアイコンの下にカラーバーが表示されている．これは，発言回数の偏りを示している．これによりファシリテータがあまり発話していない児童がいれば気づき，彼らの発話を促すためのファシリテーションを行うことができる．また，児童自らが自分の発言の多少度合いに気づき，振る舞いを正すことができるようユーザインタフェースが設計されている [79]．

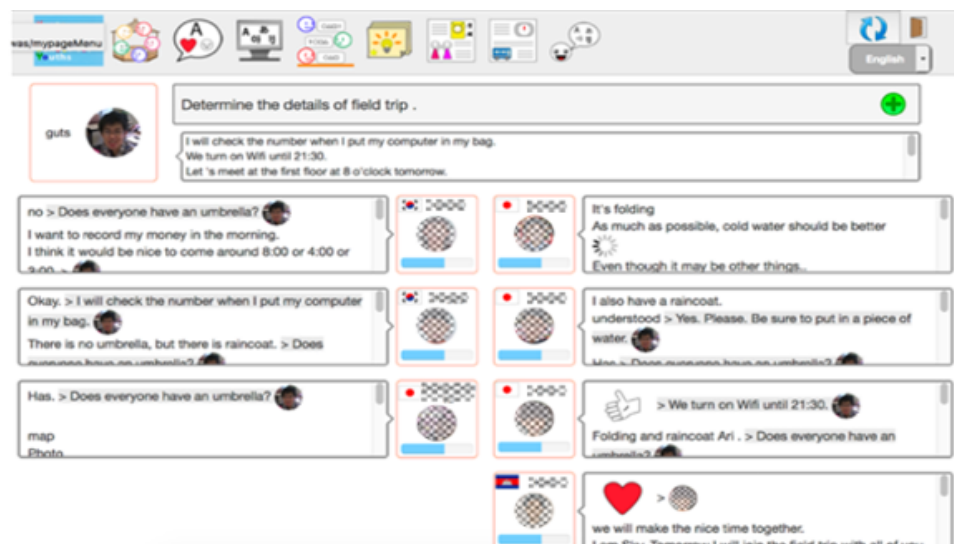


図 2.2: 多言語チャットツール (写真提供：NPO 法人 パンゲア)

### 2.4.3 ファシリテーション

KISSYでは，それぞれのグループの議論を円滑にするため，各グループに1名ずつファシリテータが所属する．ファシリテータはグループ

に向けて意見や指示や質問の発話を行い，それに対して参加者が応答する形式で対話し，議論を進める．各ファシリテータは事前に講習を受け，同じメソッドの基にファシリテーションを行うよう指示されている．ファシリテータの他にもタイピングやツールの使用方法などをサポートするためのスタッフも待機し，児童たちのコミュニケーションを支援する [46]．また，グループ構成は，グループによって言語話者の偏りがないように考慮されている．

## 第3章 ファシリテーション分析 のための発話アノテー ション手法

### 3.1 対話行為の分析の課題

本研究では、ファシリテータの発話の種類によって児童の応答数が異なることに着目し、低資源言語話者の発話を促進や抑制する対話行為を明らかにすることを目的とする。対話行為とは、発話によって伝達される意図や行為のことである [8, 55, 64]。発話を促進させたための対話行為を定量的に分析するためには、アノテーションを行うことが必要である。しかしながら、対話行為に基づく既存のアノテーション手法は、汎用的に用いられることを目的に作成されているため、ファシリテータの発話分析に適用するには、いくつかの課題点があった。そこで、本章では新たなファシリテータの発話を分析するためのアノテーション手法を提案し、それを評価した。

### 3.2 対話アノテーション

Austin は、発話によって遂行される行為を、発話行為 (*locutionary act*)、発話内行為 (*illocutionary act*)、発話媒介行為 (*perlocutionary act*) の3つの分類を定義した [26]。発話行為とは、コミュニケーションを目的に音声、単語、文、その他の言語表現を用いて、特定の意味を持つ文を発する行為のことを示す。発話内行為とは、話し手が聞き手に対して指示や要求などの別の行為を行わせるための行為のことを示

す。発話媒介行為とは、聞き手の感情や考え、行動に影響を与えるために発話する行為のことを示す。また、発話内行為による効果は発話内効力 (*illocutionary forces*) と呼ばれる。これらの分類は、排他的に定義されているわけではなく、関連し合っている。Searle は、発話内行為の 5 つの分類である断定型 (*Assertives*)、行為拘束型 (*Commissives*)、行為指示型 (*Directives*)、感情表現型 (*Expressives*)、宣言型 (*Declarations*) を定義した [55]。この 5 分類には、それぞれ基本となる遂行動詞がある。Vanderveken は、その遂行動詞に基づき、特定の条件を割り当てることで他の遂行動詞を導出し、定義した [64]。つまり、Vanderveken によって、Searle の発話行為の分類が階層的に細分化され、体系化された。

このような言語学的な理論に基づき、対話分析のためのいくつかのアノテーション手法が提案されている。例えば、DIT スキームは、特定の分野に依存しない情報探索の対話のために Bunt によって開発された [7]。また、DAMSL スキームは、アプリケーションに依存しない多次元アノテーションスキームである [3]。さらに、DIT++ スキーム、DIT スキームに DAMSL やその他のスキームが統合された、汎用的なアノテーションスキームである [7]。これは、フィードバックやターン割り当て、発話管理などのコミュニケーションに関するサブカテゴリが導入されている。

ISO 24617-2 には、言語資源管理のための標準スキームが定義されている [8]。ここには、DIT++ のコミュニケーション機能に基づき、対話行為のアノテーションするためのいくつかのアノテーション項目が追加されている。追加されたのは、以下の項目である。次元 (*Dimensions*) は、複数のコミュニケーション機能が同時に発生する会話の分類を示す。修飾子 (*Qualifiers*) は、発話行為に条件、偽り、または特定の感情を伴って行われているかどうかを示す。コミュニケーション機能 (*Communicative functions*) は、発話行為がコミュニケーションの一形態であるかどうかを示す。機能およびフィードバック依存性 (*Functional and feedback dependences*) は、発話が単独では意味をなさず、他の発話に依存して意味を持つことを示す。修辭的關係 (*Rhetorical relations*) は、ある発話が他の発話につながる関係を示す。

### 3.3 予備実験

予備実験では，Searle の分類と，それを細分化した Vanderveken の遂行動詞に基づいて定義したタグセットを用いて，2019 年度の KISSY の対話ログ内のファシリテータの 207 発話にアノテーションを行った．本研究は，ファシリテータの発話分析に用いるための，アノテーションスキーマの定義を目的としている．そのため，Searle の 5 つの分類の一つである行為指示型 (*Directive*) のみ細分化した Vanderveken の遂行動詞を用いた．この分類は，聞き手の応答を促すために用いられるからである．分析に使用したタグセットは，表 3.1，表 3.2 に示している．

これらのタグセットを用いてアノテーションを行った後，それに対する各言語話者からの応答数を集計した．その後，どのファシリテータ発話分類が低資源言語話者の応答を促進したかを確認するため，各分類間で低資源言語話者の応答数の差を比較分析した．その結果，聞き手の同意を条件とした指示である「*Request*」の発話が，低資源言語話者の応答を促進していた．一方，聞き手の同意に関係なく行う指示である「*Tell*」の発話は，低資源言語話者の応答が，他の言語話者の応答よりも抑制される傾向が見られた．

この予備実験での定性的な分析によって，2 つの考慮すべき追加点が明らかになった．一点目は発話の宛先である．ファシリテータの発話は，必ずしもグループ内全体に向けられているわけではなく，特定の聞き手に向けたものがあった．例えば，特定の名前を挙げた発話など

表 3.1: Searle の 5 つの発話分類

| 分類                  | タグ定義                           |
|---------------------|--------------------------------|
| 断定型 (Assertives)    | あることが真実であると話し手に表明させるための発話      |
| 行為拘束型 (Commissives) | 話し手の将来の行為を約束するための発話            |
| 行為指示型 (Directives)  | 聞き手に何かをさせようとするための発話            |
| 感情表現型 (Expressives) | 心理的状況を表現するための発話                |
| 宣言型 (Declarations)  | ある状態を宣言することによって成立させることを目的とした発話 |

表 3.2: Vanderveken の遂行動詞 (行為指示型の細分類)

| 分類      | タグ定義                                               |
|---------|----------------------------------------------------|
| Request | 聞き手に指示を与えることを目的とした発話。聞き手はその指示に応じることにも拒否することにも許容する。 |
| Tell    | 聞き手に指示を与えることを目的とした発話。聞き手がその指示に拒否したり反応しないことを許容しない。  |
| Insist  | 持続的にしつこく指示することを目的とした発話。                            |
| Suggest | 聞き手にある行為を行うことを検討させることを目的とした発話。                     |

である。このような発話は、特定の聞き手からの応答を期待している。聞き手の応答を促進するための発話を明らかにするには、各発話に対する聞き手の応答数と無応答数を集計するため、どの聞き手からの応答を期待していたかを集計する必要がある。しかしながら、予備実験でのアノテーションでは、誰に向けられた発話であるかはタグ付けされなかった。その結果、適切な応答率を算出することができず、適切に評価できていないと考えられる。

二点目は、連続発話であるかどうかである。ファシリテータが聞き手の応答を期待した発話を行い、それに対する応答がなかった場合、ファシリテータは表現を言い換えた発話や、詳細に説明した発話を行っていた。このような発話は、同じ分類の対話行為であっても、既出發話に後続する発話であるかどうかで聞き手への発話の効果は異なる可能性がある。また、我々の分析では、ファシリテータの発話とそれに対する児童の応答を対応づけた、一対一の発話を対象としている。そのため、連続発話であるかどうかを分類する必要がある。

### 3.4 ファシリテータの対話行為タグ

本章では、予備実験で得られた知見から、ファシリテータの対話行為を分析するために必要なアノテーション項目と、それを実現するために用いられた ISO タグセットを説明する。

### 3.4.1 必要条件

ファシリテータの発話に対して、応答された発話数を「応答数」、応答されなかった発話数を「無応答数」とし、これらを正確に集計するためには、ファシリテータはどの聞き手からの応答を期待したかを考慮する必要がある。例えば、「〇〇さん、意見を教えてください。」のような、名指しをした発話の場合、名指しされた相手以外は応答する必要はないため、それらが無応答数として数えるのは不適切である。したがって、各ファシリテータの発話がグループ内全員に向けられた発話なのか、特定の聞き手に向けられた発話かを分類する必要がある。また、特定の聞き手に向けられた発話の場合は、誰であるかを明確にする必要がある。

さらに、連続発話の場合、一度目とそれ以降の発話では、応答への効果が異なる可能性がある。また、一度目の発話とどのように関係する発話かによっても、効果が異なる可能性がある。そのため、既出發話と関係がある発話であるか、また、既出發話とそのような関係がある発話かを区別必要がある。さらに、どの既出發話と関係するかも指定する必要がある。

### 3.4.2 ISO 24617-2 (SemAF)

ISO24617-2は、対話行為を分析するために標準化されたアノテーションスキーマである。対話行為には、対話行為を生み出す話し手と、それを受け取る聞き手の立場が存在する。例えば、2者の対話では、1名の話し手と1名の聞き手がいる。また、3者以上の対話では、話し手と聞き手に加えて、会話に参加しない参加者や、少しだけ会話に参加する参加者など、様々な種類の参加者がいる。つまり、複数人での会話の場合、発話者は特定の聞き手に話しかける場合がある。ISO24617-2では、特定の聞き手を指定するために「宛先 (*Addressee*)」属性が定義されている。この属性のタグ付けは、全ての発話を対象にしている。

また、対話行為の分析において、説明、原因、譲歩など、発話同士の関係性については広く研究されており、これらは修辭関係や談話関係と呼ばれている。ISO24617-2では、発話間の特定の修辭関係を分類

表 3.3: 修飾関係のタグ分類

| 関係                    | タグ定義                                                                             |
|-----------------------|----------------------------------------------------------------------------------|
| 詳細化<br>(Elaboration)  | 既出の発話について説明する関係.<br>例) 「あなたの役割は脚本係です。」の後の「脚本係はみんなで<br>作成するクレイアニメーションのストーリーを考えます」 |
| 言い換え<br>(Restatement) | 既出の発話を言い換える関係.<br>例) 「どの映画を観ますか？」の後の「どちらの映画を観ますか<br>？」                           |
| その他<br>(Other)        | どの既出の発話との関係がない, または, 上記以外の関係.                                                    |

するための「修辞関係 (*Rhetorical Relations*)」属性が定義されており, 19 種類の修辞関係のタグ分類がある. 本研究では, 予備実験で多く観測されたファシリテータの「詳細化 (*Elaboration*)」および「言い換え (*Restatement*)」のみを用いることとし, それ以外の関係性の場合を分類するための「*Other*(その他)」を追加した3つのタグを定義した. 全ての発話が既出發話と関係する訳ではないため, 関係性がない発話も「その他 (*Other*)」に分類することとした. また, どの既出發話と関係するかも指定する必要があるため, 各発話には ID を付与し, その ID を明記するようにした. 定義した修辞関係のタグ定義は, 表 3.3 に示す. この属性のタグ付けは, 全ての発話を対象にしている. なお, 既出發話の指定は関係性がある場合のみタグづけ対象とする.

ISO24617-2 には, 対話行為のタグ分類が「コミュニケーション機能 (*Communicative Functions*)」として包括的に定義されている. アノテーションに使用するタグセットを ISO アノテーションスキーマと統一するため, 予備実験で使ったタグセットを ISO24617-2 に定義されているタグセットに対応づけた. 新たなタグセットは表 3.4 および表 3.5 に示す.

表 3.4: ISO24617-2 に対応づけた対話行為のタグセット

| 分類                              | タグ定義                                                                                 |
|---------------------------------|--------------------------------------------------------------------------------------|
| 情報提供<br>(Information-providing) | 話し手が特定の情報を相手に知らせるための発話.<br>例) 「私は発表の準備をしなければなりません。」                                  |
| 約束 (Commissive)                 | 話し手が自身のこれからの行為を約束する発話.<br>例) 「これからグループワークで行うことを言います。」                                |
| 情報探索<br>(Information-seeking)   | 聞き手の持つ情報を聞き出そうと質問する発話.<br>例) 「あなたの役割は何ですか?」「あなたのアイデアは何ですか?」                          |
| 指示 (Directive)                  | 聞き手に何かを行わせようと指示することを目的とした発話.<br>例) 「あなたの意見は何ですか?」「何かアイデアはありますか?」                     |
| 社会的機能<br>(Dimension-specific)   | 挨拶, お礼, 謝罪, 感情表現, 間投詞, 既出の発話に対する確認などの社会的機能を果たすための発話.<br>例) 「おはようございます.」「ありがとうございます.」 |

## 3.5 ファシリテータの発話へのアノテーション

### 3.5.1 アノテーションの曖昧さ

ISO 24617-2 はドメインに依存せず、汎用的に使用されることを目的に作成されている。そのため、ファシリテータの発話を分類する際に、曖昧な点が2つあった。一点目は、ファシリテータが参加者の意見を引き出すための「何かアイデアはありますか?」や「あなたの考えは何ですか?」の質問の分類である。これらの発話は、文面に着目すると、前者は相手にアイデアがあるかどうかを尋ねる質問であり、後者は相手の考えを尋ねる質問である。そのため、どちらも話し手が質問によって聞き手から知りたい情報を得ようとする発話である情報探索に分類されるように思える。しかしながら、ファシリテータの発話であることを考慮すると、ファシリテータ自身が知りたい情報を引き出すために質問している訳ではなく、他のメンバーとアイデアを共有するために意見を促すことを目的としている。そのため、これらは意見を言うように指示するための発話である指示に分類されるという解釈もできる。

表 3.5: ISO24617-2 に対応づけた指示の細分類

| 分類                       | タグ定義                                                                                                                      |
|--------------------------|---------------------------------------------------------------------------------------------------------------------------|
| 依頼<br>(Request)          | 相手の同意を条件として、ある行為を遂行するよう相手に指示する発話。聞き手が指示内容を拒否したり応答しなかったりすることを許容しつつ指示する発話。間接的な指示もこれに分類される。<br>例) 「アイデアを教えてください。」「意見はありますか？」 |
| 命令<br>(Instruct)         | 相手の同意によらず、ある行為を遂行するよう相手に指示をする発話。直前の発話からの文脈を踏まえて催促する発話もこれに分類される。<br>例) 「アイデアを言って。」「意見を出しなさい。」                              |
| 提案<br>(Suggestion)       | 相手にある行為を遂行するかを検討させるために伝達する発話。<br>例) 「みんなで意見を共有しませんか?」「みんなで意見を出し合いましょう」                                                    |
| 提案の返答<br>(Address Offer) | 相手の申し出た行為を行ってほしい、もしくは望んでいないことを相手に対して伝達する発話。<br>児童の「〇〇をやりましょうか?」という発話に対する、「はい。お願いします。」                                     |

二点目は、ファシリテータがグループワークを進行するための「役割分担を決めましょう」の発話である。これは、次のアジェンダを提供するための「情報提供」に分類する場合や、この発話をきっかけに話し合いをするように指示するための「指示」に分類する場合、ファシリテータ自身も含めた全員で話し合うことを約束するための「約束」に分類する場合がある。このように、議論の促進を意図したファシリテータの発話は、分類が難しいものがある。

### 3.5.2 曖昧なアノテーションタグの区別の明確化

前節で述べたようなファシリテータの発話のアノテーションタグの曖昧さに対処するため、以下のようにアノテーション基準を定義した。

一点目のファシリテータが参加者に投げかける質問について、グループでのタスクを達成するための発話であるか、あるいはファシリテータ自身が知りたい情報を得るための発話であるかによって分類を区別する。例えば、グループでの議論において、参加者同士の意見を共有

するための質問の場合は、ファシリテータ自身がそれを知りたい訳でなく、議論を促進しグループでのタスクの進行を促すための発話であるため、「指示」に分類するべきである。一方で、「今、何について話合っていますか?」、「今朝は何を食べましたか?」などの単にファシリテータ自身が知りたい情報を聞き出す発話の場合は「情報探索」に分類するべきである。

二点目のファシリテータがグループワークを進行するための発話について、このような発話を分類するためには、文脈の流れを考慮する必要がある。例えば、議論の流れから、言及したタスクを始めるように指示しているという意図があると考えた場合は「指示」に分類し、単にグループ作業の内容を知らせるだけの発話であると考えた場合は「情報提供」に分類するべきである。一方で、ファシリテータはグループワークのタスクの進行をサポートする立場にあるため、グループの一員として共にタスクを行うことを意図する「約束」に分類するのは適切ではない。

### 3.5.3 アノテーションフォームデザイン

各対話行為タグの定義は、できるだけ簡潔で明確な表現を使用する必要がある。各定義の説明が複雑な場合はアノテータが混乱してしまう可能性がある。また、説明が曖昧な場合は、幅広い解釈が可能になり、アノテータ間で解釈が異なる可能性がある。そのため、アノテーションタグセットには、アノテーションタグセットと、先述で示した曖昧で区別が難しい発話の分類基準、発話例を追記した。

さらに、本研究では、ファシリテータの発話にアノテーションを行うための Excel フォームを作成した(図 3.1)。このフォームでは、ファシリテータの各発話をハイライトしている。これは、予備実験において、タグ付け漏れが起きやすかったことから、アノテータがタグ付けすべき発話を認識しやすいようにするためである。また、各発話には発話 ID を付与し、修辭関係のタグ付けの際にどの既出發話と関係があるかを指定できるように設定した。さらに、タグの選択順序も考慮したデザインとなっている。具体的には、最初に対話行為の 5 つの分類のタ

| 発話ID | 発話時間             | 発話メッセージ                                                      | ID: 送信者        | 発話言語     | 発話分類 | 指示の細分類                       | 関係する発話ID | 発話間関係性 | 確信度 | 備考 |
|------|------------------|--------------------------------------------------------------|----------------|----------|------|------------------------------|----------|--------|-----|----|
| 1675 | 2019/8/5 8:04:11 | 自分が選んだ投票したミッションを解決するアイデアはありますか？                              | 1: Facilitator | Japanese |      |                              |          |        |     |    |
| 1688 | 2019/8/5 8:05:23 | ジェスチャーで表現して友達になる。                                            | 3: Child A     | Japanese |      | 情報提供 (Information-providing) |          |        |     |    |
| 1697 | 2019/8/5 8:06:03 | 私はそれが素晴らしいと思います                                              | 9: Child G     | Khmer    |      | 情報探索 (Information-seeking)   |          |        |     |    |
| 1699 | 2019/8/5 8:06:16 | 象のための保護区域を増やす                                                | 7: Child E     | Japanese |      | 約束 (Commissive)              |          |        |     |    |
| 1701 | 2019/8/5 8:06:35 | 自分の家の畑じゃないところに象の餌をまいて誘導させる                                   | 4: Child B     | Japanese |      | 指示 (Directive)               |          |        |     |    |
| 1702 | 2019/8/5 8:06:42 | 高い建物の上に畑を作る。その建物は4本の柱で支えられている。建物の下は象が通れるようになっている。            | 5: Child C     | Japanese |      | 社会的機能 (Dimension-specific)   |          |        |     |    |
| 1703 | 2019/8/5 8:06:45 | お互いに言葉を教えあってしゃべるようになる方法。                                     | 3: Child A     | Japanese |      |                              |          |        |     |    |
| 1704 | 2019/8/5 8:06:53 | 人々は象と仲良くするための方法を考えました。おもいついたのは音楽のすばらしさを象たちに伝えお互いの気持ちを音楽で伝え合う | 6: Child D     | Japanese |      |                              |          |        |     |    |
| 1705 | 2019/8/5 8:07:03 | 象と農家はやるのがいい                                                  | 9: Child G     | Khmer    |      |                              |          |        |     |    |
| 1706 | 2019/8/5 8:07:04 | チームメンバーが書いた解決策に賛                                             |                |          |      |                              |          |        |     |    |

図 3.1: アノテーションのための Excel フォーム

グを付け、それが「指示」の場合は細分類もタグを付ける。次に、既出発話との修辭関係をタグ付け、関係がある場合にどの既出发話であるかを指定する。次に、発話の宛先タイプをタグ付け、特定の聞き手宛てである場合は特定できるように欄に示す。最後に、それらのタグ付けの確信度を付与する。このように、アノテータがアノテーションする順序に沿って左から右に順にフォームを埋めるようにデザインした。

## 3.6 アノテーション実験

### 3.6.1 実験環境

我々は、定義したアノテーションスキーマを評価するため、アノテーション実験を行った。この実験には KISSY(2.4.1) のチャットツールから収集した対話データを用いた。この対話データにある、ファシリテーター

タによる 15 の発話をアノテーション対象とした。アノテーションには、日本語の対話ログを使用した。つまり、日本語の原文と他の言語の発話の日本語訳のメッセージにアノテーションしている。

### 3.6.2 評価

6 名のアノテータによってファシリテータの 15 発話にアノテーションが行われた。つまり、合計で 90 のタグ付けが収集された。それぞれの項目のタグ付け数を表 3.6～3.9 に示す。

#### 信頼性

アノテータ間のタグ付けの一致率を確認するため、Kappa 係数を算出した。本分析では、評価者が 3 名以上であり、名義尺度のカテゴリも 3 つ以上あるため、Fleiss の Kappa 係数を使用した。各タグセット内全体と、タグセット内の各分類のアノテータ間の Kappa 値は表 3.10～3.13 の通りである。

表 3.6: 対話行為分類のタグ付け数

| 分類    | タグ付け数 |
|-------|-------|
| 情報提供  | 12    |
| 約束    | 12    |
| 情報探索  | 23    |
| 指示    | 35    |
| 社会的機能 | 8     |
| 合計    | 90    |

表 3.7: 指示の細分類のタグ付け数

| 分類    | タグ付け数 |
|-------|-------|
| 依頼    | 11    |
| 命令    | 14    |
| 提案    | 10    |
| 提案の応答 | 0     |
| 合計    | 35    |

表 3.8: 修辞関係のタグ付け数

| 分類   | タグ付け数 |
|------|-------|
| 詳細化  | 22    |
| 言い換え | 9     |
| その他  | 59    |
| 合計   | 90    |

表 3.9: 宛先のタグ付け数

| 分類    | タグ付け数 |
|-------|-------|
| 参加者全員 | 79    |
| 特定の個人 | 11    |
| 合計    | 90    |

表 3.10: 基本的な 5 つの対話分類の Fleiss' kappa 係数

| 分類    | Kappa 値 |
|-------|---------|
| 全体    | 0.538   |
| 情報提供  | 0.192   |
| 約束    | 0.308   |
| 情報探索  | 0.825   |
| 指示    | 0.579   |
| 社会的機能 | 0.561   |

表 3.11: 指示の細分類の Fleiss' kappa 係数

| 分類 | Kappa 値 |
|----|---------|
| 全体 | 0.344   |
| 依頼 | 0.109   |
| 命令 | 0.086   |
| 提案 | 0.055   |

表 3.12: 修辞関係の Fleiss' kappa 係数

| 分類   | Kappa 値 |
|------|---------|
| 全体   | 0.201   |
| 詳細化  | 0.230   |
| 言い換え | -0.012  |
| その他  | 0.262   |

表 3.13: 宛先の Fleiss' kappa 係数

| 分類    | Kappa 値 |
|-------|---------|
| 全体    | 0.648   |
| 参加者全員 | 0.648   |
| 特定の個人 | 0.648   |

この結果を、Landis の目安 [35] に基づいて評価すると、この結果より、4つのタグセット内の全体のアノテーションの一致性があると評価された。その中でも、基本的な対話行為の分類は中程度の一致、宛先は高い一致であった。つまり、これらのタグセットの一貫性が保証され则认为られる。一方で、指示の細分類と修辭関係のタグセットは、一定の一致度は認められたが、いくつかのタグ分類において不明瞭さが残っており、改善の余地がある则认为られる。特に、指示の細分類は、アノテーションが一致するには、基本の対話行為の分類が指示で一致し、かつ、指示の細分類でも一致する必要があるため、特に一致しにくくなっている。次に、実際の各発話のアノテーションを確認すると、指示の細分類のタグ付けにおいて、「依頼」と「提案」は、アノテータ同士でどちらかのタグで分かれ、最も混同しやすいことが明らかになった。このことから、この2つのタグを明確に区別できるよう、より明確な説明が必要である则认为られる。また、修辭関係においては、「関係なし(およびその他の修辭関係)」のタグ付けはアノテータ間で一致していたものの、「言い換え」や「詳細化」のタグ付けは一致しにくかった。このことについて、アノテータからのフィードバックによると、「その他の修辭関係」のタグ付けの判断は難しく、「言い換え」や「詳細化」以外の修辭関係がどのようなものかが不明確であることから、提示されている2つの修辭関係とどのように区別してタグ付けすれば良いか理解できなかった。そのため、「言い換え」、「指示」とそれ以外の修辭関係の区別について考える余地を持たさないように、この2つの修辭関係に該当する場合のみにタグ付けするように改善する必要がある则认为られる。

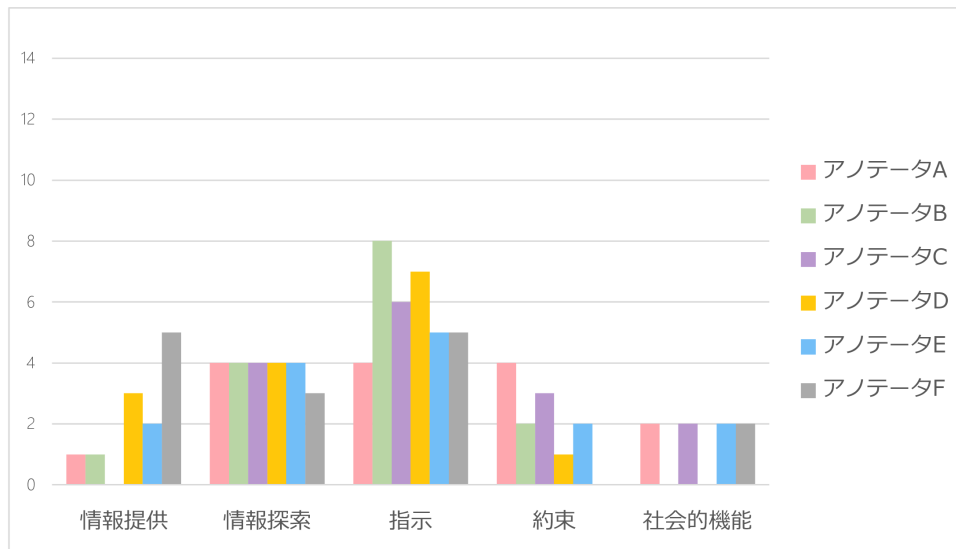


図 3.2: 各アノテータによる基本的な発話分類のタグ付け回数の分布

### タグ付けの分布

各アノテータによるタグ付けの違いを確認するため、タグ付けられた回数を各分類で集計した。図 3.2 から図 3.5 の集計結果のグラフを示す。この図の横軸が各分類タグ、縦軸がその使用回数を示している。これらの図から、「修辞関係」のタグ付けでは、ある傾向が確認できる。「言い換え」のタグ付けを用いなかったアノテータ 2 名は、「関係なし」または「その他の関係」のタグ付け回数が他のアノテータと比較して多かった。そのため、アノテータ達は「連続発話」以外の関係性を混同して解釈する可能性がある。

### アノテーションの困難さ

アノテータ達はタグ付けを行う際、各発話に対するタグ付けに対する確信度を付与するように指示された。確信度は 1(全く自信がない) から 5(非常に自信がある) の 5 段階尺度で付けた。

アノテータ間のタグ付けの一致度が低いタグは、アノテーションが

難しい可能性がある．そのため，各タグ分類のカッパ係数と確信度の相関関係を確認した．各タグセットの Kappa 値と確信度の平均値から計算された相関係数  $r$  を以下に示す．

#### 各タグセットの相関係数

対話行為分類:  $r = 0.0900$

指示の細分類:  $r = 0.8817$

修辞関係:  $r = 0.8767$

宛先:  $r = -$

また，各タグセットの分類別の Kappa 値と相関係数をプロットした図を 3.6 に示す．各タグセットの相関係数の絶対値が 0.7 以上であれば，強い相関があると示される．つまり，「指示の細分類」と「修辞関係」のタグセットには正の相関が認められた．それらのタグセットのそれぞれの分類に着目すると，Kappa 値が高いタグ分類は確信度が高い傾向があった．さらに，対話行為分類には，相関関係が認められず，一致度が低いにもかかわらず確信度が高い分類が含まれていた．このことから，アノテータによって異なる解釈がされ，他の分類と混同してしまった可能性がある．このタグ分類については，定義を明確化する必要があると考えられる．

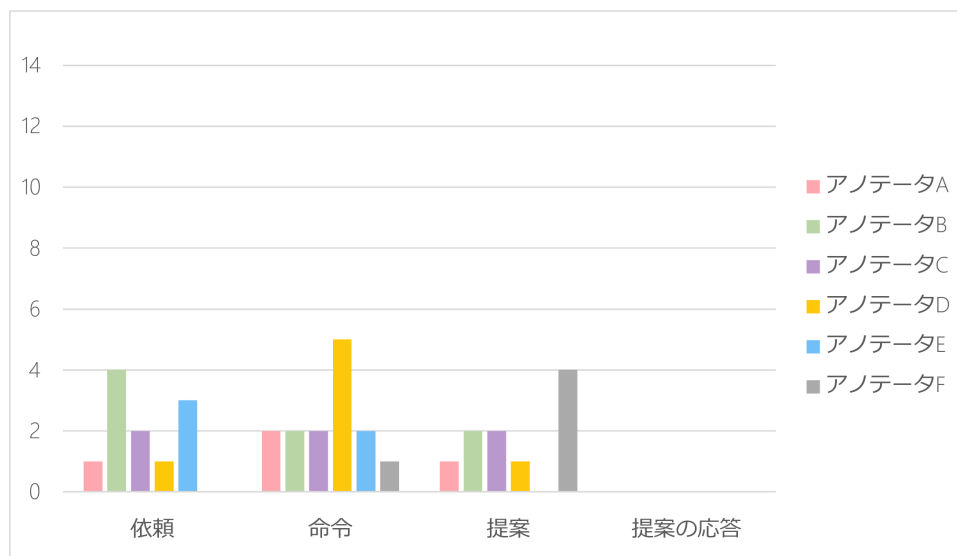


図 3.3: 各アノテータによる指示の細分類のタグ付け回数の分布

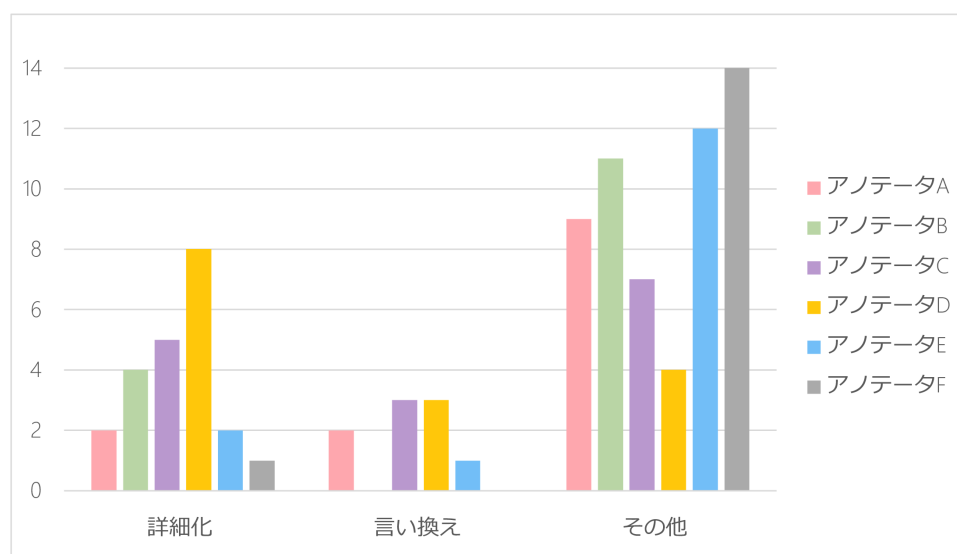


図 3.4: 各アノテータによる修辞関係のタグ付け回数の分布

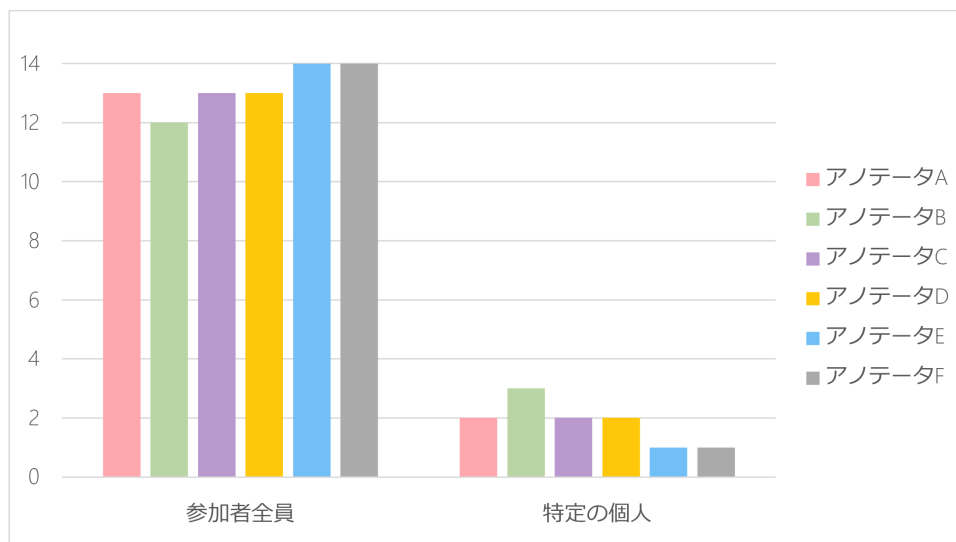


図 3.5: 各アノテータによる宛先のタグ付け回数の分布

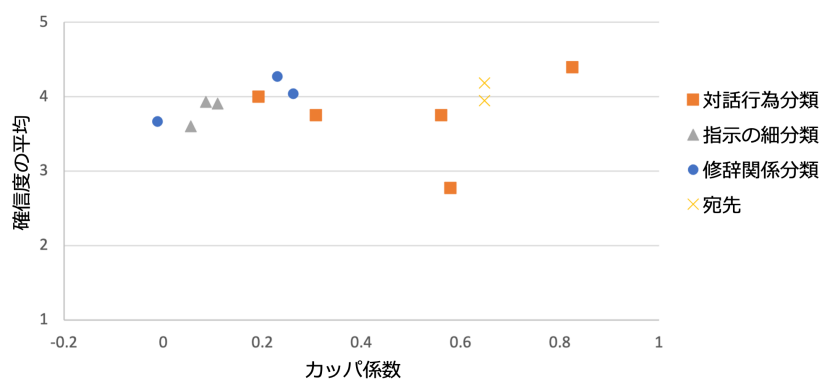


図 3.6: 各分類の確信度と Kappa 係数の相関図

### 3.7 考察

対話行為分類において、最も Kappa 値の低いタグ分類は「情報提供」と「約束」であった。これらの分類はアノテータによって異なり、混同

してしまうことが多くあった。例えば、「ストーリーを話し合います」のような未来の行動についての発話である。このような発話は、ファシリテータが話し合うように指示する発話、ファシリテータがこれから話し合うことを伝達する発話、ファシリテータがこれから話し合うことを約束する発話の3つの解釈ができる。まず、ファシリテータはあくまで議論を支援するための立場であり、一緒に議論するために参加していないため、これから話し合うことを約束する発話である「約束」に分類するべきではない。次に、その発話の前後の対話を考慮し、「情報提供」か「指示」のどちらかを区別するべきである。

指示の細分類においては、「依頼」と「提案」の一致率が低い傾向にあった。これらの分類の区別については、具体的な発話事例を示すことで、解釈の相違を解消する必要がある。さらに、「提案」のタグ付けは、特に一貫性がなく、指示の細分類の中で確信度の平均が最も低かった。各指示の細分類を付けた発話への確信度の平均は、依頼が4.00、命令が3.93、提案が3.78である。そのため、定義をより詳細に記し、明確化する必要がある。

修辞関係の分類については、「言い換え」や「詳細化」のタグ付けのKappa値が低く、「その他の分類(または関係なし)」の分類と混同されることが多くあった。それは、2つの関係性以外の分類についての知識がなく、消去法などを適用できず、区別するのが難しいと感じていたためであった。そのため、既出發話の中で修辞関係があるかどうかを判断し、ある場合に分類を選択するように、誘導するような指示を行う必要がある。また、ファシリテータの発話として、対話ログから2つの修辞関係分類の他に、既出の発話の事例を補足する発話があった。そのため、「例示」という分類を追加することを検討する必要がある。改善された修辞関係のタグセットを表3.14に示す。

宛先の分類においては、全体的な一致度は高かったものの、いくつかの発話で合意がされていない発話があった。これは、特定の児童の既出發話に関する質問する場合に多くあった。このような発話を行う場合、名前を含むなど、誰に対しての質問かを明示することが多いが、まれに名前を明示しないこともあった。このような発話を正確にタグ付けするためには、対話の流れを考慮する必要がある。そのため、ファシリテータには対話の流れの考慮するように指示を追加する必要がある。

表 3.14: 修正後の修辭関係分類タグ

| 分類                      | タグ定義                                                                         |
|-------------------------|------------------------------------------------------------------------------|
| 詳細化<br>(Elaboration)    | 既出の発話について説明する関係.<br>例) 「あなたの役割は脚本係です。」の後の「脚本係はみんなで作成するクレイアニメーションのストーリーを考えます」 |
| 言い換え<br>(Restatement)   | 既出の発話を言い換える関係.<br>例) 「どの映画を観ますか?」の後の「どちらの映画を観ますか?」                           |
| 例示<br>(Exemplification) | 既存の発話に関連する事例を示す関係.<br>例) 「しりとりでは、語尾に「ん」がつくと負けです。」の後の「例えば、みかんです。」             |

### 3.8 まとめ

異文化コラボレーションの場におけるファシリテータの対話行為を分析するため、既存のアノテーションスキームを適用することは困難であった。例えば、対話行為を分類するタグ定義が明確でないため、アノテータがタグ付けをする際に複数の分類の解釈を混同する場合があった。また、対話行為を分析するために、応答率を求める必要があった。そのため、応答数を正確に集計するための発話の宛先を付与する必要であった。そこで、既存の対話行為分類を用いたアノテーション実験を通して、タグセットの一貫性の確認や、改善が必要なタグを抽出し、改善することで、異文化コラボレーションの場におけるファシリテータの対話行為を分析のためのアノテーションスキーマを定義した。

## 第4章 ファシリテータの対話行為分析

### 4.1 対話行為に着目したファシリテーション

コミュニケーションにおいて、聞き手がどのように応答するかは、話し手から受け取る対話行為に作用される。そこで、本章では、ファシリテータの対話行為に着目し、低資源言語話者が発言のしやすい場を構築するためのファシリテーションや、機械翻訳を介すことで対話行為にどのような影響があるかを明らかにする。そのために、実フィールドから収集した対話ログデータの分析を行った。具体的には、実際の各言語で対話行為に基づいたアノテーションを行う。その後、対話行為によって応答数が変わるのかの比較や、使用言語によって応答数が変わるのかを比較する。さらに、機械翻訳の介入した前後で対話行為が変わるのかを比較し、その傾向を評価する。

以下の節では、4.2 節で関連する研究とそこから導出した仮説を説明し、4.3 節で分析した対話ログのコミュニケーション環境、4.4 節ではファシリテータの発話へのアノテーション、4.5 節から 4.7 節ではそれぞれの分析手法と結果について説明し、4.8 節に考察を行う。

## 4.2 機械翻訳を活用した多言語間のコミュニケーション

### 4.2.1 対話行為に着目したコミュニケーション支援

対話行為に着目したコミュニケーション支援研究では、話し手によって話しやすくなる対話を作るために、聞き手側に着目した対話行為のパターンが分析されている。目黒らは、このパターンを対話制御システムに組み込み、評価した。その結果、聞き手役の友好的な対話ルールの提案を行っている [93]。また、話し手による発話と聞き手による応答の対を隣接ペアと言う [54, 76]。これは、特定の発話の種類によってどのように応答されるかが類型化されている [75]。これらは、同一言語でのコミュニケーションにおける対話行為の効果の研究であり、多言語でのコミュニケーションにおける対話行為の効果は明らかでない。また、これらの着目するのは、話し手が意図する対話行為である。しかし、本研究で対象とする多言語間のコミュニケーションは、機械翻訳が介入するため、翻訳前後で対話行為が変化する可能性もある。そこで、本研究では、聞き手側が受け取る話し手の対話行為に着目し、機械翻訳が作用することによる対話行為への影響を分析する。

### 4.2.2 機械翻訳の現状

機械翻訳を活用することで、多言語間でのコミュニケーションが可能になる。特に、共通言語を持たない話者同士のコミュニケーションの場合は有効性が高い。近年では、ニューラル機械翻訳 [4, 60] が出現したことによって、翻訳精度は著しく向上した [63]。このことから、異なる言語話者同士のコミュニケーションはスムーズになっている。一方で、多言語間のコミュニケーションには未だに課題が残されている。表 4.1 には KISSY(2.4.1 を参照) で観測された対話の一部を示す。ファシリテータは「移民の人が学校に来た時に問題になるのは何だと思いますか?」と参加者に意見を促している。その発話への応答として、日本語の参加者はアイデア自体を発話している。その一方、クメール語

表 4.1: ログデータの会話事例

| No. | 発話者 (使用言語)   | メッセージ          |
|-----|--------------|----------------|
| 1   | ファシリテータ (ja) | 何かアイデアはありますか？  |
| 2   | 児童 (ja)      | おもてなしをするおかしをだす |
| 3   | 児童 (km)      | はい、持っています      |
| 4   | 児童 (ja)      | 遊んで言葉のをひくくする   |

話者はアイデアを持っているかどうかを発話している。これらの応答は、どちらも文面からすると適切なものである。しかしながら、どの言語話者かによって、話し手の発話の意図をどのように認識したかが異なっていることがわかる。

### 4.2.3 リサーチクエスチョン

本分析では、聞き手が認識する対話行為に着目し、以下のリサーチクエスチョンについて調査する。

- RQ1** 聞き手が受け取る対話行為によって、応答のしやすさに影響を与えるか。また、影響するのであれば、どの対話行為への応答がしやすいか。
- RQ2** 聞き手の話者言語によって、各対話行為への応答のしやすさに影響を与えるか。また、影響するのであれば、どの言語話者にとってどの対話行為に応答がしやすいか。
- RQ3** 機械翻訳による翻訳前後で認識される対話行為は変わるのか。また、変わるのであれば、どの言語間の翻訳でどのように対話行為が変わるか。

## 4.3 多言語コミュニケーション対話ログ

### 4.3.1 コミュニケーション環境

ファシリテータの対話行為を分析するため、実際の多言語コミュニケーションである KISSY での多言語チャットツールの対話ログを収集した。この多言語チャットツールでのメッセージの翻訳には、Google Translation が使用されている。この時の KISSY では、合計 32 名の参加者があり、その内、4 名がファシリテータ、28 名が小学 3 年生から中学 3 年生までの児童であった。各参加者の国籍は日本、ケニア、ジョージア、カンボジアである。カンボジアで話されるクメール語が低資源言語である。参加者は 8 名ずつの 4 チームに分けられる。各チームは児童 7 名とファシリテータ 1 名が所属する。チーム分けは国籍や言語が偏らないように考慮される。各チームに所属する参加者の国籍と母国語、チャットで使用した言語の内訳は表 4.2 に示す。カンボジア人の児童は各チームに 1 名ずつ所属していた。その内、A と C のチームに所属するカンボジア人の児童のみがクメール語を使用した。よって、参加した児童のうち、日本語を使用したのが 20 名、英語を使用したのが 6 名、クメール語を使用したのが 2 名であった。

### 4.3.2 対話ログ

本分析に用いた対話ログには、155 件のファシリテータによる発話と、777 件の児童による応答が含まれていた。表 4.3 には対話ログを一部抜粋を示す。これは、移民の子供を学校に受け入れるためにどうするか？というテーマの議論をしている。基本的に、ファシリテータが質問や指示の発話で児童の意見やアイデアを促し、児童がそれに対して応答する相互作用を繰り返した議論の進め方である。この事例は日本語で記載されているが、実際には、全ての発話は多言語に翻訳されるため、日本語以外で発話されたメッセージは日本語に翻訳されたものが示されている。

表 4.2: グループの構成, 使用言語内訳

|        | 日本<br>(日本語) | ケニア<br>(英語) | ジョージア<br>(グルジア語) | カンボジア<br>(クメール語) |
|--------|-------------|-------------|------------------|------------------|
| Team A | 6*<br>(日本語) | 1<br>(英語)   |                  | 1<br>(クメール語)     |
| Team B | 6*<br>(日本語) | 1<br>(英語)   |                  | 1<br>(英語)        |
| Team C | 5<br>(日本語)  | 1*<br>(英語)  | 1<br>(英語)        | 1<br>(クメール語)     |
| Team D | 5<br>(日本語)  |             | 1<br>(英語)        | 2*<br>(英語)       |

\*は大人ファシリテータが含まれた人数であることを示す。  
( )内には使用言語を示す。

## 4.4 発話アノテーション

定量的にファシリテータの発話を分析するため, 対話ログデータへのアノテーションが必要となる。そこで, 本節では, アノテーションに用いたタグセットとアノテーションの手順の説明を示す。

### 4.4.1 対話行為に基づいたタグセット

アノテーションに用いたタグセットは, 第3章で定義したアノテーション手法を用いている。これは, 対話アノテーションの国際標準規格の ISO/DIS24617-2[8] に基づいて定義されている。具体的には5つの基本的な対話行為分類と, その中の一つの「指示」を細分化した4つの分類である。この「指示」のみ細分化を用いたのは, ファシリテータがアイデアや意見を促す時に使用され, 重要であると考えられるためである。そのため, 2種類のタグセットを用いてアノテーションを行った。以下で, 実際に分析に用いたタグセットを紹介する。

表 4.3: ログデータの会話事例

| 発話者 (使用言語*)     | メッセージ                                    |
|-----------------|------------------------------------------|
| ファシリテータ<br>(ja) | 「移民の人が学校に来た時に問題になるのは何だと思いますか？」           |
| 児童 A(ja)        | 「話しかける」                                  |
| 児童 B(ja)        | 「彼らはそこの言語を理解している教師、またはグーグル翻訳さえ与えられるべきです」 |
| 児童 A(ja)        | 「一緒に遊ぶ」                                  |
| 児童 B(ja)        | 「児童 A 一緒に遊ぶということはどういう意味ですか」              |
| 児童 C(ja)        | 「国ごとのクラスをつくる」                            |
| 児童 B(ja)        | 「私たちはインターナショナルスクールのシステムを採用しています」         |

基本的な 5 つの対話行為分類は、「情報提供」「情報探索」「指示」「約束」「社会的機能」であり、発話意図によって区別される。また、そのうちの「指示」の 4 つの細分類は、「依頼」「命令」「提案」「提案の返答」であり、指示の方法や意図によって区別される。これらの分類の定義を表 4.4、表 4.5 に示す。各定義には発話例も示している。

#### 4.4.2 アノテーション方法

前節で示したタグセットを使用し、対話ログのファシリテータの発話にアノテーションを行った。本研究では、聞き手が認識する対話行為に着目するため、児童達が使用する言語である、日本語、英語、クメール語の 3 言語でアノテーションを行った。4 名の日本語話者、4 名の英語話者が 4 名、3 名のクメール語話者によって、各言語の対話ログにアノテーションされた。アノテーションを一貫して行うため、ISO で

規定されたタグセットに基づいて作成されたマニュアルとフォームが使用された(第3章). それぞれのアノテータはこのマニュアルに則り, ファシリテータの各発話に対し, 該当する対話行為分類と, その確信度を5段階評価を注釈した. その後, 各言語内でアノテータ間のアノテーション一致率を算出することで, アノテーション結果の信頼性を評価した. 一致率には Kappa 係数が用いられた. 各言語の Kappa 係数は, 日本語で 0.518, 英語で 0.660, クメール語で 0.727 であった. これら进行评估するため, Landis[35] の目安を適用したところ, 適度に一致していると認められた. 複数人のアノテータによるアノテーションを一意に決定するため, アノテーションが一致しなかった発話については, 各アノテータの確信度を確認し, その総和が最も高い分類に決定した. 例えば, 3 名のアノテータによる 5 つの基本的な対話行為分類と確信度のアノテーションが(提案, 4), (情報探索, 1), (情報探索, 2), の時, 分類ごとの確信度の総和は「情報提供」が 3, 「提案」が 4 となる. そのため, その発話は「提案」の分類とされる.

#### 4.4.3 隣接応答ペアの抽出

本研究では, それぞれの対話行為に対してどれだけの児童が応答したかを比較し, 分析する. そのためには, ファシリテータによる各発話と, それに対する児童たちの応答を対応付ける必要がある. 一般的には, 連続する発話を抽出し, 隣接応答対とする. しかし, 児童たちのタイピング能力や自らの意見を言語化する能力には個人差がある. そのため, 議論の流れや各発話の内容を確認しながら対応付けを行った. この時, メッセージ内容に基づき, 複数回に分けて送信された発話や, タイプミスなどによって言い直したと考えられる発話については, 一発話として統合された. 統合された発話事例は表 4.6 に示す. さらに, 児童のメッセージ内容から, どのファシリテータの発話への応答とも判断されない場合は分析対象外とした.

表 4.4: 基本的な 5 つの対話行為分類の定義

| 対話行為  | 分類の定義                                                         |
|-------|---------------------------------------------------------------|
| 情報提供  | 話し手が持つ特定の情報を聞き手に知らせることを目的とした発話<br>例: 「議論のテーマは〇〇です。」           |
| 情報探索  | 話し手が聞き手の持つ情報を聞き出すために質問することを目的とした発話<br>例: 「あなたの意見は何ですか？」       |
| 指示    | 話し手が聞き手に何らかの行為を行わせようと指示することを目的とした発話<br>例: 「あなたのアイデアを教えてください。」 |
| 約束    | 話し手が聞き手に何らかの行為を行うと約束することを目的とした発話<br>例: 「これからのスケジュールを発表します。」   |
| 社会的機能 | 話し手が挨拶やお礼, 感情表現などの社会的機能を果たすことを目的とした発話<br>例: 「おはよう」「楽しいです。」    |

## 4.5 RQ1: 各対話行為への児童の応答数の比較

ファシリテータの対話行為は、児童の応答のしやすさに影響するか。また、影響するのであれば、どのファシリテータによる対話行為に応答しやすいかを明らかにするため、各対話行為に対する児童の応答を言語別で集計し、比較した。対話行為の分類の中で、聞き手の応答を促すために使用されるのは「情報探索」と「指示」の細分類(「依頼」,「命令」,「提案」,「提案の応答」)の5つである。そのため、これらの発話を分析対象とした。この分析では、それぞれの言語における

表 4.5: 「指示」の4つの細分類の定義

| 細分類   | 分類の定義                                                                    |
|-------|--------------------------------------------------------------------------|
| 依頼    | 聞き手が同意することを条件とした、特定の行為の遂行を聞き手に指示する発話。<br>例: 「あなたのアイデアを教えてください。」          |
| 命令    | 聞き手が同意するかによらず、特定の行為の遂行を聞き手に指示する発話。<br>例: 「意見を教えなさい。」                     |
| 提案    | 聞き手に特定の行為の遂行を提案する発話。<br>例: 「あなたの意見をみんなに共有しましょう。」                         |
| 提案の返答 | 聞き手による特定の行為の遂行の申し出に対し、同意や反意を伝える発話。<br>例: (児童からの「〇〇しましょうか?」に対する)「お願いします。」 |

表 4.6: ファシリテータの発話で統合された発話事例

| 発話の順序 | メッセージ             |
|-------|-------------------|
| 1     | 残 15 分の時間の使い方です   |
| 2     | 1つはフィールドトリップの話をする |
| 3     | 2つ目はストーリーを考える     |
| 4     | 3つ目は寝る            |
| 5     | 4つ目は部屋に戻る         |

対話行為と応答数の関連性と、それぞれの対話行為がどのように関連するかを明らかにすることを目的としている。以下では、どのように児童の応答数を集計したのかと、分析に使用した統計的検定と結果を説明する。

### 4.5.1 対話行為への児童の応答数の集計

まず、ファシリテータの各発話に対して期待する児童の応答見込み数を「期待応答数」、その数値のうち、応答がどれだけあったかで「応答数」と「無応答数」として集計した。この集計は、言語ごとに行われた。例えば、日本語話者の児童が5名所属するグループにおいて、ファシリテータが全員に向けて行った発話の場合、全員からの返答を期待するため、期待応答数は5である。その内、2名からの応答があった場合は応答数が2、無応答数が3である。各グループのファシリテータ達は、事前にパンゲアによる講習を受けており、彼らは同様のメソッドに基づいたファシリテーションを行っていたと考えられるため、ファシリテータの国籍や使用した言語、グループで議論されたトピックは異なっていたものの、各グループにおけるファシリテーションに大きな差はなかったと考えられる。そのため、ファシリテータによって区別せず、グループを超えてデータを集計した。図4.1～4.3に各言語話者の集計結果を示す。また、英語話者の集計には、英語が母国語でないジョージア国籍とカンボジア国籍のデータも含まれている。彼らの個々のデータは表4.7に示す。彼らは母国語でない英語を使用していた。そのことから、彼ら自身の中で母国語と英語間の翻訳を行い、それが対話行為を解釈をする際に影響した可能性も考えられる。そのため、母国語が英語であるケニア人のデータと、英語を使用していたジョージア人とカンボジア人のデータを比較分析した。具体的には、ケニア人と英語を使用したジョージア人、ケニア人と英語を使用したカンボジア人の各国籍間のデータをフィッシャーの正確性検定を用いた有意差検出を行った。その結果、どの対話行為においても、有意差は認められず、彼ら自身の中で翻訳は行っていた可能性はあるものの、対話行為の解釈に影響はないと考えられる。よって、彼らの集計データも英語話者のデータとして分析対象とした。

### 4.5.2 統計的検定

図4.1～4.3に示されている通り、「提案の返答」の発話は観測されなかった。そのため、この分類以外の対話行為4つの集計表を用いて独立

| 使用言語<br>対話行為 | 日本語 |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|-----|------|-----------------------------|
|              | 応答数 | 無応答数 |                             |
| 情報探索         | 16  | 54   |                             |
| 依頼           | 50  | 130  |                             |
| 命令           | 1   | 24   |                             |
| 提案           | 3   | 52   |                             |

図 4.1: 日本語話者の各対話行為への応答数と無応答数の集計表

| 使用言語<br>対話行為 | 英語  |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|-----|------|-----------------------------|
|              | 応答数 | 無応答数 |                             |
| 情報探索         | 13  | 38   |                             |
| 依頼           | 2   | 9    |                             |
| 命令           | 4   | 16   |                             |
| 提案           | 2   | 12   |                             |

図 4.2: 英語話者の各対話行為への応答数と無応答数の集計表

| 使用言語<br>対話行為 | クメール語 |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|-------|------|-----------------------------|
|              | 応答数   | 無応答数 |                             |
| 情報探索         | 8     | 12   |                             |
| 依頼           | 0     | 1    |                             |
| 命令           | 0     | 1    |                             |
| 提案           | 0     | 7    |                             |

図 4.3: クメール話者の各対話行為への応答数と無応答数の集計表

性検定を行った。さらに、クメール語の「依頼」、「命令」の発話について、どちらの分類も1発話ずつしか観測されておらず、データ数が非常

に小さい。そのため、これらも分析対象外とした。そのため、クメール語話者は「情報探索」、「提案」の2つの分類のみの独立性検定である。それぞれの表の各セルは、実測値から算出される残差値に基づいてハイライトされている。残差値が0に近いほど白に、残差値が大きいほどピンクに、残差値が小さいほど紫に色付けられている。独立性検定については、期待値が小さいセルがあることから、フィッシャーの正確性検定 ( $\alpha = 0.05$ ) を用いた。その結果、日本語話者のみに有意差が認められ ( $p < .001$ )、英語話者 ( $p = .905$ ) とクメール語話者 ( $p = .068$ ) は有意差が認められなかった。次に、有意な差が認められたデータは、どの対話行為が特に有意に差が認められるのかを確認するため、調整済み残差を用いた残差分析 ( $\alpha = 0.05$ ) を行った。残差分析において算出された残差値が 1.96 以上である場合は有意差ありとなる。この時、応答数に有意差ありと認められると、児童の応答を促進させたとと言える。一方、無応答数に有意差ありと認められると、児童の応答を抑制したと言える。分析の結果、有意差が認められた表 4.1 内のセルは、下線で示されている。この結果から、日本語話者話者の、応答を促進する傾向があるのは「依頼」の発話であり、応答を抑制する傾向があるのは「命令」、「提案」の発話であった。また、クメール語話者のデータにおいては、 $p$  値が有意水準 5% 下では、有意な差は認められなかった。しかし、有意水準 10% であれば有意差が認められる程度の値であることと、クメール語話者の各対話行為への応答数を考慮すると、「情報探索」に対する応答が「提案」よりも応答していたため、「情報探索」の発話は応答を促す傾向があると考えられる。これらの結果より、聞き手が受け取る対話行為によって、応答のしやすさに影響を与えられられる。特に、日本語話者においては「依頼」、クメール語話者においては「情報探索」が応答を促進する。つまり、言語によって応答数に影響する対話行為が異なっていた。このことは、使用する言語によって、応答しやすい対話行為が異なる可能性を示唆している。

この結果について、児童たちの参加に対するモチベーションや言語能力、タイピングの速さなどの対話行為以外の要因が応答数に影響した可能性も考えられるかもしれないが、その影響は少ない。なぜなら、KISSY は学校の授業のように参加を強制されている訳ではなく、児童自らの意思で参加している。また、言語化やタイピングの能力におい

表 4.7: ジョージア国籍とカンボジア国籍の英語話者児童の各対話行為への応答数集計

| 国籍    | 対話行為  | 応答数 | 無応答数 |
|-------|-------|-----|------|
| ジョージア | 情報探索  | 4   | 8    |
|       | 依頼    | 0   | 1    |
|       | 命令    | 1   | 4    |
|       | 提案    | 0   | 6    |
|       | 提案の返答 | 0   | 0    |
| カンボジア | 情報探索  | 3   | 11   |
|       | 依頼    | 1   | 3    |
|       | 命令    | 2   | 4    |
|       | 提案    | 0   | 4    |
|       | 提案の返答 | 0   | 0    |

ても、各グループに所属するファシリテータの他に、児童たちを支援するための大人のスタッフが待機しており、児童たちの発話するスピードに差はあったかもしれないが、応答数への影響は少ないと言える。さらに、KISSY の実施は対面で行われており、グループメンバーとは同じテーブルを囲んで座り、各自に提供される PC を用いて機械翻訳を組み込んだ多言語チャットツール上での議論が行われる。他のグループメンバーは同一のテーブルにいるため、チャットツールを介さない対面での会話は可能であり、児童の応答に影響した可能性も考えられる。しかし、実際には、同一言語話者の児童たちが会話をする様子も確認されたが、それは議論されている内容について確認するなどの最低限のものであったため、影響は少ないと考えられる。

## 4.6 RQ2: 児童の各使用言語間での応答数比較

前節の 4.5.2 の結果から、児童が使用する言語によって、各ファシリテータによる対話行為への応答のしやすさに影響するのか、また、影

響するのであれば、どの言語にとってどの対話行為に応答しやすいかを明らかにするため、RQ1での集計をもとに、使用言語間でそれぞれの対話行為への応答数を集計した表を作成し、独立性検定を行った。以下では、その統計的検定と結果を説明する。

#### 4.6.1 統計的検定

各集計において、期待値が小さいデータにはフィッシャーの正確性検定 ( $\alpha = 0.05$ )、それ以外のデータには  $\chi^2$  検定 ( $\alpha = 0.05$ ) を行った。なお、クメール語話者のデータにおいては、「依頼」、「命令」がどちらも1発話ずつしか観測されておらず、データが非常に小さいため、これらの対話行為については、日本語話者と英語話者のみのデータを分析対象として検定を行った。各表の集計結果を図4.4～4.7に示す。ここで、「提案の返答」は、観測されず0値だったため、分析対象外としている。それぞれの表の各セルは、実測値から算出される残差値に基づいてハイライトされている。残差値が0に近いほど白に、残差値が大きいほどピンクに、残差値が小さいほど紫に色付けられている。検定の結果、「情報探索」は  $p = .303$ 、「依頼」は  $p = .730$ 、「命令」は  $p = .155$ 、「提案」は  $p = .423$  であり、どの対話行為においても有意差が認められなかった。つまり、聞き手の話者言語によって各対話行為への応答のしやすさには影響を与えない。このことから、使用言語によって応答数に有意差が認められるほどの差はなく、日本語話者の児童は、他の言語より「依頼」の発話への応答が多かったり、「命令」や「提案」の発話への応答が少ない訳ではないと言える。

### 4.7 RQ3: 翻訳前後の対話行為比較

RQ1, RQ2の分析では、機械翻訳の介入は考慮しなかった。これらの分析結果より、使用する言語によってそれぞれの対話行為への応答数に有意な差は見られなかった。しかし、各言語でのそれぞれの対話行為に対する応答数には、日本語話者のみの応答数に有意差が認められた。このことについて、機械翻訳が介入することが対話行為に影響し、そ

| 対話行為<br>使用言語 | 情報探索 |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|------|------|-----------------------------|
|              | 応答数  | 無応答数 |                             |
| 日本語          | 16   | 54   |                             |
| 英語           | 13   | 38   |                             |
| クメール語        | 8    | 12   |                             |

図 4.4: 「情報探索」への各言語話者の応答数と無応答数の集計表

| 対話行為<br>使用言語 | 依頼  |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|-----|------|-----------------------------|
|              | 応答数 | 無応答数 |                             |
| 日本語          | 50  | 130  |                             |
| 英語           | 2   | 9    |                             |
| クメール語        | 0   | 1    |                             |

図 4.5: 「依頼」への各言語話者の応答数と無応答数の集計表

れぞれの言語で受け取った対話行為に差が生じたためだと考えられる。そのため、RQ3 では、機械翻訳の対話行為への影響について分析する。具体的には、機械翻訳によって対話行為が変化するのか、変化するとすれば、どの対話行為がどう変化するのかを明らかにした。

#### 4.7.1 使用言語ごとの受け取る対話行為数の比較

まず、使用した言語によって受け取った対話行為が異なるかを明らかにするため、全てのファシリテータの発話が各言語ではどの対話行為で受け取られるかを抽出した。その後、言語ごとにそれぞれの対話行為が何発話出現したかを集計した。集計結果を図 4.8 に示す。なお、「提案の返答」の発話はどの言語においても観測されなかったため省略してい

| 対話行為<br>使用言語 | 命令  |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|-----|------|-----------------------------|
|              | 応答数 | 無応答数 |                             |
| 日本語          | 1   | 24   |                             |
| 英語           | 4   | 16   |                             |
| クメール語        | 0   | 1    |                             |

図 4.6: 「命令」への各言語話者の応答数と無応答数の集計表

| 対話行為<br>使用言語 | 提案  |      | 残差値<br>≥1.96<br>0<br>≤-1.96 |
|--------------|-----|------|-----------------------------|
|              | 応答数 | 無応答数 |                             |
| 日本語          | 3   | 52   |                             |
| 英語           | 2   | 12   |                             |
| クメール語        | 0   | 7    |                             |

図 4.7: 「提案」への各言語話者の応答数と無応答数の集計表

る。その後、言語間で比較するために独立性検定 ( $\alpha = 0.05$ ) を行った。このとき、期待値が 5 以下のデータが含まれていたため、フィッシャーの正確性検定を用いている。その結果、 $p < .001$  であったため、有意差が認められた。このことから、使用した言語で受け取ったそれぞれ対話行為数に差があったと言える。その後、特にどの言語のどの対話行為に有意差が認められるかを確認するため、調査済み残差を用いた残差分析 ( $\alpha = 0.05$ ) を行った。残差分析の結果、有意差が認められたセルを図 4.8 内の下線で示している。また、表の各セルは、実測値から算出される残差値に基づいてハイライトされている。残差値が 0 に近いほど白に、残差値が大きいほどピンクに、残差値が小さいほど紫に色付けられている。残差値が 1.96 以上であり、正に有意である場合、その対話行為が他の言語よりも有意に多く受け取られたと言える。一方、残差値が -1.96 以下であり、負に有意である場合、その対話行為が他の

言語よりも有意に少なく受け取られていたと言える。これらのことから、日本語話者の児童は他の言語話者の児童よりも多くの「依頼」を受け取っていたことがわかる。つまり、日本語話者の児童は「依頼」の発話に対し、応答する機会が多くあった。このことがRQ1の結果で、日本語話者の児童の応答数に有意差が認められた可能性がある。それぞれの児童は言語は異なっていたが同じファシリテータからの発話を受け取っていた、しかし、各言語で受け取った対話行為の数が異なっていた。これは、機械翻訳によって対話行為が変化することを示唆している。

#### 4.7.2 機械翻訳が介入する言語ペアの特定

前節の分析結果から、機械翻訳によって対話行為が変化することが示唆されたため、機械翻訳の介入する前後の言語において、対話行為の認識が変化するか。また、そうであるとすれば、どの言語ペアにおいて、どの対話行為が他のどの対話行為に変化するかを分析する。そのために、まず機械翻訳の介入する前後の言語ペアを抽出した。具体的には、KISSYでのファシリテータと児童の使用言語から、機械翻訳が介入する言語ペアを特定した。ファシリテータが使用した言語は日本語と英語であった。また、児童の使用した言語は、日本語、英語、クメール語であった。このことから、抽出された言語ペアは、日本語から英語、英語から日本語、英語からクメール語である。ここで、日本語からクメール語の言語ペアについては、英語を介した翻訳がされていたため、言語ペアとして扱わなかった。この確認には、Google 翻訳で日本語話者のファシリテータの発話を、クメール語に翻訳したものと、英語を介してクメール語に翻訳したもので比較して行った。

#### 4.7.3 各言語ペアにおける対話行為の組み合わせパターンの集計

各言語ペアごとに、それぞれの翻訳前後の対話行為の組み合わせパターンの出現数を集計した。表 4.8, 表 4.9, 表 4.10 を示す。

| 対話行為<br>使用言語 | 情報<br>提供 | 情報<br>探索 | 依頼 | 命令 | 提案 | 約束 | 社会的<br>機能 | 残差値          |
|--------------|----------|----------|----|----|----|----|-----------|--------------|
| 日本語          | 32       | 14       | 36 | 5  | 11 | 2  | 19        | $\geq 1.96$  |
| 英語           | 35       | 37       | 7  | 14 | 10 | 12 | 4         | 0            |
| クメール語        | 46       | 34       | 4  | 2  | 9  | 3  | 15        | $\leq -1.96$ |

図 4.8: 各対話行為の出現数

表 4.8: 日本語から英語での出現数の集計表

| 翻訳後   | 翻訳前      |          |    |    |    |    |           | 合計 |
|-------|----------|----------|----|----|----|----|-----------|----|
|       | 情報<br>提供 | 情報<br>探索 | 依頼 | 命令 | 提案 | 約束 | 社会的<br>機能 |    |
| 情報提供  | 16       | 0        | 0  | 0  | 0  | 0  | 1         | 17 |
| 情報探索  | 1        | 10       | 15 | 0  | 0  | 0  | 1         | 27 |
| 依頼    | 0        | 1        | 5  | 0  | 0  | 0  | 1         | 27 |
| 命令    | 0        | 0        | 4  | 3  | 3  | 0  | 0         | 10 |
| 提案    | 1        | 1        | 0  | 0  | 3  | 0  | 0         | 5  |
| 約束    | 0        | 0        | 0  | 0  | 1  | 1  | 0         | 2  |
| 社会的機能 | 0        | 0        | 0  | 0  | 0  | 0  | 10        | 10 |
| 合計    | 18       | 12       | 24 | 3  | 7  | 1  | 13        | 78 |

#### 4.7.4 マクネマー検定

表 4.7.3 をもとに、それぞれの言語ペアの対話行為別に「翻訳の前後で対話行為が同じだった発話数」「翻訳後に他の対話行為へ変わった発話数」、「翻訳後に他の対話行為から変わった発話数」、「その他」を集計し、集計表を作成した。表 4.11 は作成した集計表の一例である。これを対話行為ごと、かつ言語ペアごとに作成した。その後、各集計表を使用し、マクネマー検定を行った。有意水準は 5% である。表 4.12 はその結果を示す。なお、「提案の返答」については、どの集計表においても値が 0 であったため、結果を省いている。また、英語から日本語の言

表 4.9: 英語から日本語での出現数の集計表

| 翻訳後   | 翻訳前  |      |    |    |    |    |       | 合計 |
|-------|------|------|----|----|----|----|-------|----|
|       | 情報提供 | 情報探索 | 依頼 | 命令 | 提案 | 約束 | 社会的機能 |    |
| 情報提供  | 27   | 0    | 0  | 1  | 0  | 0  | 0     | 28 |
| 情報探索  | 0    | 9    | 0  | 0  | 0  | 0  | 0     | 9  |
| 依頼    | 0    | 12   | 1  | 6  | 0  | 0  | 0     | 19 |
| 命令    | 1    | 1    | 0  | 1  | 1  | 0  | 0     | 4  |
| 提案    | 0    | 0    | 0  | 1  | 6  | 0  | 0     | 7  |
| 約束    | 0    | 0    | 0  | 0  | 0  | 2  | 0     | 2  |
| 社会的機能 | 3    | 0    | 0  | 0  | 0  | 0  | 4     | 7  |
| 合計    | 31   | 22   | 1  | 9  | 7  | 2  | 4     | 76 |

語ペアにおける「約束」は、表内に0の値が多く、算出ができなかったため、表記していない。この検定結果より、いくつかの表で有意差が認められた。このことは、特定の言語ペアにおいて、特定の対話行為が変化しやすいことを示唆している。つまり、機械翻訳の翻訳前後で認識される対話行為は変わる。各言語ペアのそれぞれの結果を以下で示す。

#### 日本語から英語の言語ペアについて

翻訳後に他の対話行為から「情報探索」や「命令」に変化しやすく、翻訳前の「依頼」が翻訳によって他の対話行為に変化しやすいことが明らかになった。表 4.8 をもとに、これらの変化について確認すると、翻訳前の「依頼」の発話は、「情報探索」へ変化している発話数が一番多く、それに次いで「命令」が多くあった。さらに、翻訳後の「情報探索」の発話は、翻訳前は「依頼」である場合が一番多くあった。これらのことから、日本語から英語への翻訳では、「依頼」が「情報探索」または「命令」に変化しやすい傾向があると言える。

## 英語から日本語の言語ペアについて

翻訳後に他の対話行為から「依頼」に変化しやすく、翻訳前の「情報探索」が翻訳によって他の対話行為に変化しやすことが明らかになった。表 4.9 をもとに、これらの変化について確認すると、翻訳前の「情報探索」の発話は、「依頼」へ変化している発話数が一番多くあった。さらに、翻訳後の「依頼」の発話は、翻訳前は「情報探索」である場合が一番多くあった。これらのことから、英語から日本語への翻訳では、「情報探索」が「依頼」に変化しやすい傾向があると言える。

## 英語からクメール語の言語ペアについて

翻訳後に他の対話行為から「情報提供」に変化しやすく、翻訳前の「命令」が翻訳によって他の対話行為に変化しやすことが明らかになった。表 4.10 をもとに、これらの変化について確認すると、翻訳前の「命令」の発話は、「情報提供」へ変化している発話数が一番多くあった。さらに、翻訳後に「情報提供」の発話は、翻訳前は「命令」である場合が一番多くあった。これらのことから、英語からクメール語への翻訳では、「命令」が「情報提供」に変化しやすい傾向があると言える。

## 4.8 考察

日本語と英語間の言語ペアにおいては、どちらの言語へ翻訳する場合においても、「情報探索、命令」(英語)-「依頼」(日本語)のパターンで対話行為が変化した。例えば、日本語話者による「移民を学校に受け入れるための解決策についてアイデアはありますか？」の発話の場合、日本語話者はこれを「依頼」として受け取るが、それが英語に翻訳された「Do you have an idea for a solution to accept immigrants in school?」の発話を、英語話者は「情報探索」として受け取った。これらのメッセージについて、それぞれの言語話者のアノテーションにインタビューし、この発話をどのように認識したかを尋ねた。日本語話者は、この発話への応答として自分のアイデアの内容を期待されていると認識し、英語話者は、この発話への応答として自分はアイデアを持っているかどうかの有無

表 4.10: 英語からクメール語での出現数の集計表

| 翻訳後   | 翻訳前  |      |    |    |    |    |       | 合計  |
|-------|------|------|----|----|----|----|-------|-----|
|       | 情報提供 | 情報探索 | 依頼 | 命令 | 提案 | 約束 | 社会的機能 |     |
| 情報提供  | 44   | 4    | 5  | 9  | 7  | 1  | 0     | 70  |
| 情報探索  | 0    | 42   | 0  | 1  | 0  | 0  | 0     | 43  |
| 依頼    | 1    | 1    | 2  | 4  | 0  | 0  | 0     | 8   |
| 命令    | 0    | 2    | 0  | 2  | 0  | 0  | 0     | 4   |
| 提案    | 0    | 0    | 1  | 3  | 5  | 0  | 0     | 9   |
| 約束    | 1    | 0    | 0  | 0  | 0  | 3  | 0     | 4   |
| 社会的機能 | 2    | 0    | 0  | 0  | 0  | 0  | 14    | 16  |
| 合計    | 48   | 49   | 8  | 19 | 12 | 4  | 14    | 154 |

を期待されていると認識していた。つまり、日本語話者はこのメッセージを間接的な指示だと認識し、英語話者は直接的な指示であると認識した。さらに、「Tell me your position」という発話を英語話者は「命令」として受け取り、それが日本語に翻訳された「あなたの立場を教えてください」を「依頼」として受け取った。これらのメッセージについて、英語では命令形の文型で表現されているが、日本語では「～ください」と丁寧な文末で表現されている。このことは、翻訳によって丁寧度が変化する傾向にあることを示唆している。

日本語と英語間の言語ペアにおける対話行為の変化は、聞き手からの応答を期待する対話行為同士での変化であった。一方で、英語からクメール語の言語ペアにおいては、「命令」の発話が翻訳によって「情報提供」に変化した。この「情報提供」は、聞き手に自分の持っている情報を提供することを目的としており、聞き手からの応答を期待しない。このような場合、クメール語者は応答することを期待されていると感じないため、応答しないと考えられる。このような変化が要因となり、クメール語話者の児童の発言数が少なくなっていた可能性も考えられる。この対話行為の変化の例として、英語話者のファシリテータによる

表 4.11: 集計表の一例

|     |        | 翻訳後  |            | Total |
|-----|--------|------|------------|-------|
|     |        | 情報提供 | 他の<br>対話行為 |       |
| 翻訳前 | 情報提供   | 16   | 2          | 18    |
|     | 他の対話行為 | 1    | 59         | 60    |
|     | Total  | 17   | 61         | 78    |

「命令」の発話である「About how the story ends. I want the original blue team idea!」が、クメール語に翻訳され、「情報提供」として受け取られた。この発話をクメール語に翻訳されたものを折り返し翻訳すると、「I want the blue team's original idea how to end the story!(折り返し翻訳文)」となる。この発話の原文を確認すると、間接的に要求するような表現となっている。これがクメール語に翻訳されることで、アイデアが欲しいという事実を知らせているだけの発話に変化し、受け取られている。

さらに、それぞれの言語ペアにおける対話行為の変化の発話事例を確認すると、メッセージに間接的な要求が含まれている場合が多くあった。このことから、間接的な要求は翻訳によって対話行為が変化しやすいと言える。そのため、ファシリテータが間接的な要求を含む発話を行った場合、それが翻訳によって他の意図に認識されやすいと言える。そのため、機械翻訳を介したコミュニケーションの場では間接的な表現を含まないような発話を行うことが必要である。

表 4.12: マクネマー検定結果

| 翻訳ペア* | 対話行為  | <i>p</i> 値 |
|-------|-------|------------|
| 日→英   | 情報提供  | 1.000      |
|       | 情報探索  | .001 (-)   |
|       | 依頼    | <.001 (+)  |
|       | 命令    | .023 (-)   |
|       | 提案    | .683       |
|       | 約束    | 1.000      |
|       | 社会的機能 | .248       |
| 英→日   | 情報提供  | .371       |
|       | 情報探索  | <.001 (+)  |
|       | 依頼    | <.001 (-)  |
|       | 命令    | .227       |
|       | 提案    | .479       |
|       | 約束    | -(算出不可)    |
|       | 社会的機能 | .248       |
| 英→ク   | 情報提供  | <.001 (-)  |
|       | 情報探索  | .077       |
|       | 依頼    | .772       |
|       | 命令    | .001 (+)   |
|       | 提案    | .546       |
|       | 約束    | .479       |
|       | 社会的機能 | .479       |

\*日→英: 日本語から英語, 英→日: 英語から日本語, 英→ク: 英語からクメール語

†(+) 他の対話行為へ変化, (-) 他の対話行為から変化

## 4.9 まとめ

機械翻訳を介した児童達の多言語コミュニケーションにおいて、低資源言語を話す児童の発話数の少なさが課題である。そのため、本研究では、ファシリテータの発話意図がどのように伝わるかに着目し、低資源言語を話す児童の応答を促すためのファシリテータによる対話行為や、機械翻訳を介したことによる対話行為への影響を分析した。具体的には、対話行為への児童の応答数の比較や、使用言語による児童の応答数の比較、翻訳前後の対話行為の比較を行った。その結果、翻訳を介したことによって、対話行為が変化することが明らかになった。日本語と英語間の翻訳の場合、どちらの言語への翻訳においても、「依頼」(日本語)-「情報探索, 命令」(英語)のパターンで変化する傾向が見られた。また、英語からクメール語の翻訳の場合は、「命令」が「情報提供」へ変化する傾向が見られた。この変化について、「情報提供」は聞き手の応答を必要としない対話行為である。そのため、このような変化が起きた場合、クメール語話者から応答されないと考えられる。このことが、クメール語話者の発話数の少なさを引き起こしている可能性がある。また、本分析での日本語話者の児童の応答を促進する傾向があった「依頼」の発話は、機械翻訳を介したことによって変化しやすかった。そのため、日本語話者以外の言語話者の児童は、この対話行為の発話を受け取ることが少なかったと考えられる。そのため、クメール語話者の発話を促進することが困難であったことを示唆している。また、対話行為が変化する対話には、原文に間接的な要求が含まれている場合が多く、このような表現が対話行為の変化に影響した可能性がある。そのため、機械翻訳を介した多言語のコミュニケーションの場では、このような表現を使用しないことが重要である。

最後に、今後の課題としては、分析に用いたデータ数に限りがあった。特に、低資源言語であるクメール語の児童の応答数の分析に用いたデータは29件とデータ数が少なかった。そのため、さらに多くのデータを収集し、分析することで、有効なファシリテーションにつながるための知見が得られると考える。また、分析に用いたデータは実際のフィールドから得られたものであるため、統制実験を行い、明らかになった有効的なファシリテーションの検証を行う必要があると考える。

## 第5章 ファシリテータエージェントを用いた異文化コラボレーション支援

### 5.1 要約エージェントによる対話理解促進

前章までは、異文化コラボレーションの場での低資源言語話者の発話を促進するために有効的なファシリテータの対話行為を明らかにした。しかしながら、ファシリテータが有効的な対話行為を行なってファシリテーションを行ったとしても、低資源言語話者が発言するためには、低資源言語話者が対話を理解することが不可欠である。そのため、本章では、対話の理解を促進するための要約に着目し、要約するための要約要求エージェントと要約提供エージェントの2種類のエージェントを定義した。要約要求エージェントは、高資源言語話者に対話の要約を要求する。一方で、要約提供エージェントは、エージェント自身が対話を要約し、提供する。これによって、低資源言語話者の対話の理解を促進し、より積極的に対話に参加し、会話が活性化する。これらのエージェントを用いたコミュニケーション実験を行い、有効性の評価を行った。

## 5.2 ファシリテータエージェントを用いたコミュニケーション支援研究

Li ら [36] は、非母国語話者と母国語話者の共通言語を用いたコミュニケーションにおける発話ターンに着目し、非母国語話者に発話を促すように設計された自動と手動のエージェントを設計した。一方、本研究では、議論の要約に着目しており、機械翻訳を用いた多言語コミュニケーションの支援を目的としている。

また、伊藤ら [24, 74] は、ファシリテータ支援機構に基づく、大規模な意見集約システム COLLAGREE を開発した。これは、電子掲示板型の Web 議論システムである。また、伊美ら [95] は、このシステムに実装されたファシリテーションの支援システムを用いた実証実験により、有効性を検証した。このシステムは、多数の参加者の意見を取り込み、合意形成を促進するものである。これらの研究は、同一言語間のファシリテータ支援システムとして設計されている。また、Kikihira ら [28] は、議論された話題と各発話者の発話量を可視化するツールを実装した。一方、本研究では、リアルタイム多言語コミュニケーションにおいて、低資源言語話者を支援するエージェントを実装する。

Kim ら [27] は、オンラインのグループチャットシステムにおいて、意見を言わない参加者や、発話の少ない参加者に意見を促し、その応答に対しての反応するファシリテータエージェントの機能を実装し、検証した。このエージェントは、参加者の発話数の平等性を促進することに着目している。一方で、本研究では、対話の理解促進を目的としている。

### 5.2.1 仮説

本研究では、低資源言語話者に要約を提供する機能を持つ 2 種類のファシリテータエージェントの有効性を検証するため、各エージェントを用いた条件タスクと、エージェントを用いない条件タスクでのコミュニケーションタスクを行った。本研究では、対話を要約するためのファシリテータエージェントとして、2 種類のエージェントを定義した。要

約要求エージェントと要約提供エージェントである。前者は、グループ全体に対話を要約するように要求するメッセージを送信し、高資源言語話者がそれに応答することを期待する。また、後者は、エージェント自身が議論を要約し、グループ全体に提供する。これらのエージェントは、どちらも低資源言語話者が議論についていけず、一定回数発話がなかった場合に機能する。具体的には、高資源言語話者の発話が4回連続した場合に機能する。低資源言語話者は、議論の要約を受け取ることで、議論の理解度を向上させることができると考えられる。また、議論の内容を理解することで、積極的に発話するようになり、議論が活性化する。さらに、議論が活性化することにより、議論への貢献度が向上し、議論によって決まった結論に対する満足度も上がると考えられる。これらのことから、低資源言語話者がこれらのエージェントを用いることによる効果に関する仮説を以下のように定義した。

**H1.1** エージェントを用いた条件では、低資源言語話者の議論の理解度が向上する。

**H1.2** エージェントを用いた条件では、低資源言語話者の議論への参加度が向上する。

**H1.3** エージェントを用いた条件では、低資源言語話者の議論への満足度が向上する。

要約要求エージェントは、高資源言語話者が議論を要約し、提供することを期待する。そのため、このエージェントを用いる条件では、高資源言語話者の作業負担が増加すると考えられる。このことにより、高資源言語話者は、低資源言語話者に対してより配慮していると感じる。これらのことから、高資源言語話者がこれらのエージェントを用いることによる効果に関する仮説を以下のように定義した。

**H2.1** 要約要求エージェント条件では、他の条件と比較して高資源言語話者の作業負担を増加させる。

**H2.2** 要約要求エージェント条件では、他の条件と比較して高資源言語話者は低資源言語話者への配慮を感じる。

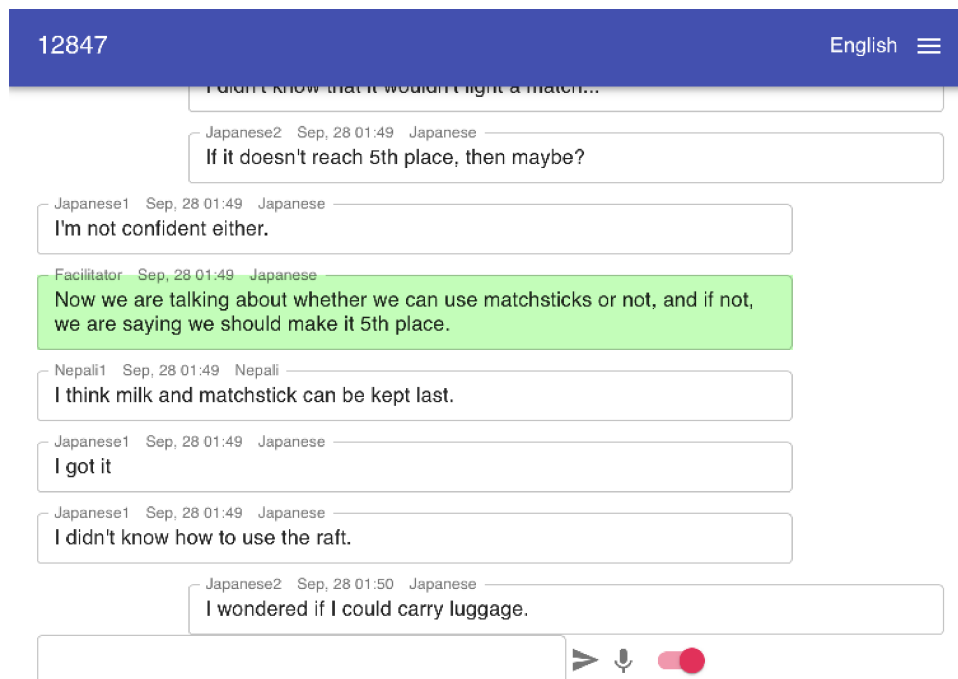


図 5.1: Langrid Chat の UI 画面

## 5.3 ファシリテータエージェント

本研究で定義した要約要求エージェントと、要約提供エージェントは、研究室内で運営する多言語チャットツール Langrid Chat を拡張して実装した。本節では、これらのエージェントの実装について説明する。

### 5.3.1 システム構成

Langrid Chat は機械翻訳が組み込まれた多言語チャットツールである。各ユーザから送信されたメッセージは、他の言語に翻訳され、各クライアントに配信される。つまり、各ユーザはそれぞれの母国語で全てのメッセージを見ることができる。図 5.1 は Langrid Chat のインターフェース画面である。英語話者向けの画面のため、メッセージは全て英語で表示される。それぞれのユーザ自身が送信したメッセージが右

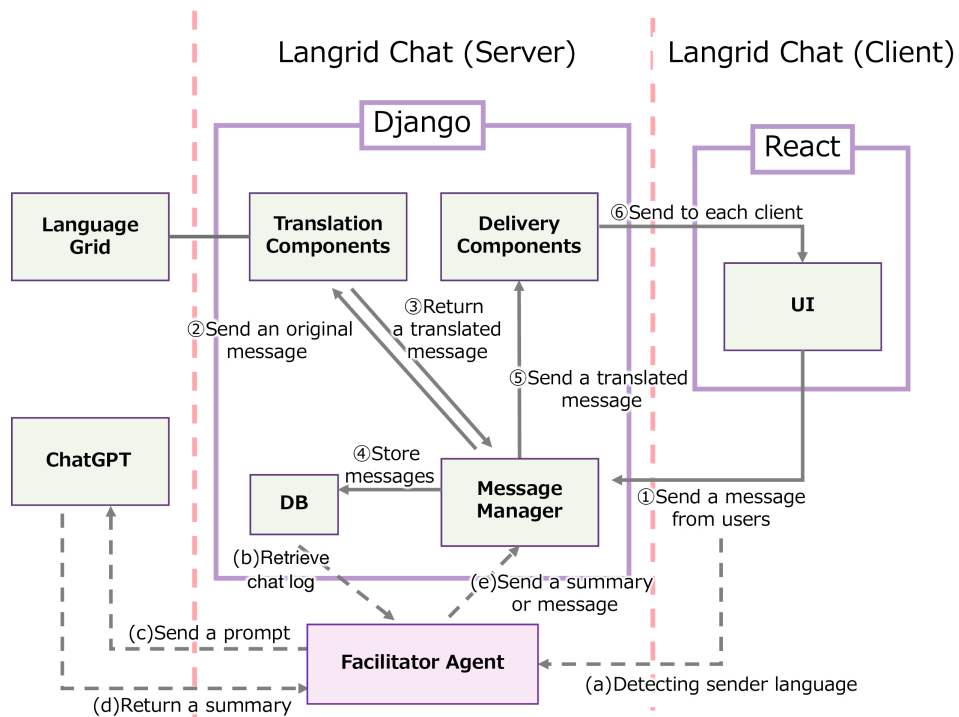


図 5.2: Langrid Chat のシステム構成図

側に、他のユーザから受信したメッセージは左側に表示される。このシステムの対応言語は、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、簡体中国語、繁体中国語である。図 5.2 は Langrid Chat のシステム構成図である。このツールのシステムの動作は、サーバ側とクライアント側に分かれている。サーバ側は Django をベースに構築されており、主にメッセージの翻訳やメッセージのデータベースへの保存、各クライアントへの翻訳後のメッセージ送信などを行う。クライアント側 (UI) は React をベースに構築されており、主に受信メッセージの表示を行う。

### 5.3.2 要約要求エージェント

このエージェントは、高資源言語話者からの発話が4回連続した場合に要約を要求する発話を送信する。図5.2にはLangrid Chat内のこのエージェントの実装を示し、図5.3にはエージェント内の一連の処理のフローチャートを示す。ユーザがメッセージを送信すると、エージェントが送信者の言語を検出する。この検出された言語が高資源言語の場合、連続発話のカウンタがインクリメントされる。また、検出された言語が低資源言語の場合、カウンタはリセットされる。このカウンタが4以上になった場合、エージェントは要約を要求する内容のメッセージを送信し、カウンタをリセットする。エージェントから送信されるメッセージは表5.1に示す。このメッセージは送信者が「Facilitator」、送信言語が「日本語」に設定される。また、このエージェントの発話は、前章で日本語話者の発話を促す効果があると認められた、「依頼」に分類される発話であり、異文化コラボレーションの場として適切な直接的な表現の発話である。また、表5.1の下段のメッセージは、日本語からネパール語に人手で翻訳されたエージェントの発話であり、低資源言語であるネパール語話者が受け取るメッセージである。つまり、エージェントからのメッセージは機械翻訳を介さず、人手で翻訳されたものが受け取られる。

表 5.1: エージェントの要約要求メッセージ

| 言語    | メッセージ                                                                                                          |
|-------|----------------------------------------------------------------------------------------------------------------|
| 日本語   | 議論についていけなくなったので、今どんな議論をしているのか要約してください。                                                                         |
| ネパール語 | म छलफलसँगसँगै अगाडि बढ्न सकिरहेको छैन, त्यसैले कृपया हामीले हालै छलफल गरिरहेको कुरालाई संक्षेपमा बताइदिनुहोस्। |

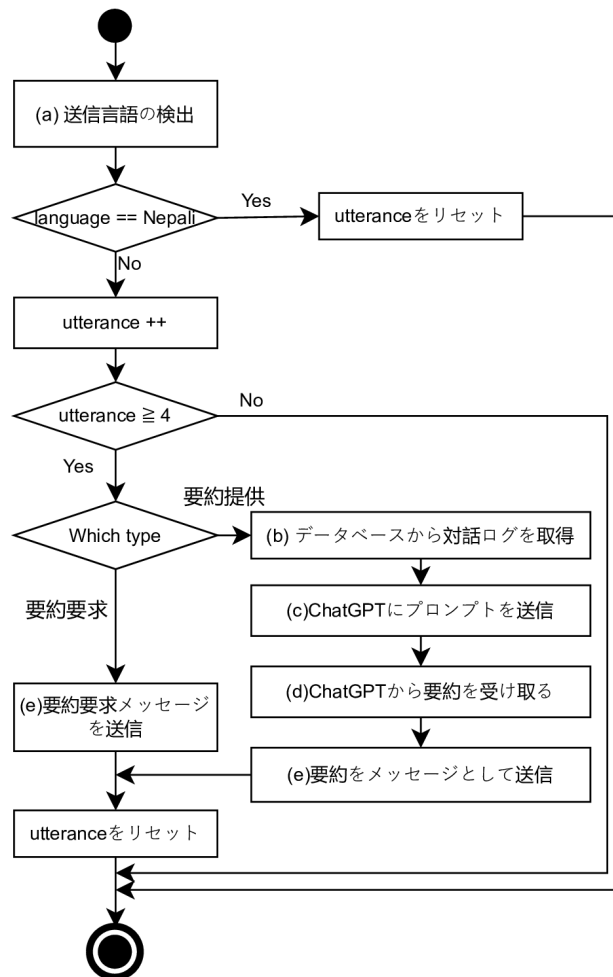


図 5.3: ファシリテータエージェントの処理

### 5.3.3 要約提供エージェント

このエージェントは、高資源言語話者からの発話が4回連続した場合、議論を要約した内容の発話を送信する。図 5.2 には Langrid Chat 内のこのエージェントの実装を示し、図 5.3 にはエージェント内の一連の処理のフローチャートを示す。このエージェントの実装には ChatGPT(モデル:GPT-4o)を用いている。ユーザがメッセージを送信すると、エージェントは送信者の言語を検出する。この検出された言語が高資源言語の場合、連続発話のカウンタがインクリメントされた。また、検出された言語が低資源言語の場合、カウンタはリセットされる。このカウンタが4以上になった場合、エージェントはデータベースから対話ログを取り出し、それをプロンプトに挿入して ChatGPT に送信する。以下は ChatGPT に送信されるプロンプトである。

**プロンプト：**あなたは、機械翻訳を用いた多言語コミュニケーションの対話ログを要約するエージェントです。参加者たちがサバイバルゲームを行なっている途中ですが、今どのようなことが議論されているかを要約してください。以下の指示に従って、現在の議論内容を要約したエージェントによる発話を作成してください。

**条件：**

- 1) (ネパール語話者のユーザ名)の最後の発話以降の内容を要約してください。
- 2) 要約は 60 字以内に行ってください。
- 3) 翻訳しやすいような簡単な言葉を使用してください。

**要約例：**今は～～について議論しています。このアイテムは～～という理由で重要であるという意見があります。

**対話ログ：**(ここに対話ログを挿入)

その後、ChatGPT から送られてきた要約結果をメッセージとして送信する。この時、送信者は「Facilitator」、送信言語は「日本語」に設定される。このエージェントのメッセージは、他のメッセージと同様に翻訳されて低資源言語話者に送信される。その後、高資源言語話者の発

話数のカウンタはリセットされる。

## 5.4 実験環境

本節では、実験の概要や、被験者について、実験の際に用いたタスクについて説明する。

### 5.4.1 概要

本研究で実装したファシリテータエージェントの効果を検証するため、機械翻訳を用いた多言語での共同タスクを行うコミュニケーション実験を行った。この実験において検証する項目は、低資源言語話者の「議論の理解度」と「議論への参加度」と「議論の満足度」、高資源言語話者の「作業負荷」と「低資源言語話者への配慮」である。被験者達は、高資源言語話者2名と低資源言語話者1名のグループで Langrid Chat を用いた共同タスクを行った。各グループは、要約要求エージェント条件、要約提供エージェント条件、エージェントなし条件の2つの条件下でそれぞれのタスクを行った。つまり、各グループは3タスクずつ行った。それぞれのグループで行う各条件の順序は、偏りが出ないように、ラテン方格法に基づいてランダムイズし、設定した。各評価項目に対する評価を収集するため、被験者は各タスク後にアンケートに回答した。また、各被験者の発話数と発話長を測定するために、Langrid Chat の対話ログを取得した。

### 5.4.2 被験者

本実験では、高資源言語話者として日本語話者12名、低資源言語話者としてネパール語話者6名を募集した。彼らは日本語話者2名とネパール語話者1名のグループに分けられた。日本語話者は日本語を母国語とする日本在住の大学生から選出した。また、ネパール語話者は日本在住でネパール語話者を母国語としている留学生から選出した。被

験者達は、同一の部屋で実験に参加したが、タスク中はオフラインでの被験者間の干渉を避けるため、他の被験者との口頭での会話は禁止された。そのため、被験者達は Langrid Chat 上でのみコミュニケーションが許された。

### 5.4.3 タスク

各被験者グループは、3つの本番タスクと1つのトレーニングタスクを行った。各タスクは、既存のグループでの協調議論を通して課題を解決する形式のタスク [34] を修正したものである。このタスクは、困難な環境下で、手元にある複数のアイテムを生存するために重要なランキングを決定するためにグループで議論する。オリジナルのタスクシナリオは砂漠であり、アイテムは15個である。しかし、本実験では本番タスクとトレーニングタスクで4種類のタスクが必要であったため、オリジナルのシナリオをもとに、海洋、北極、無人島のシナリオを追加した。オリジナルの砂漠のシナリオはトレーニングタスクで使用され、その他のシナリオは本番タスクで使用された。また、ランク付けするアイテム数は本番タスクでは5つ、トレーニングタスクでは3つに修正された。さらに、予備実験において、ランキングのみを合意するタスクの場合、メンバー同士のランキングが一致しているアイテムは議論されないことがあったことから、ランキングだけでなく、その順位の理由も合意するように指示した。各被験者グループは本番タスクの前にトレーニングタスクを行った。これは、タスクを経験をしたかどうかの差によるノイズを最小限にするためである。また、実施するシナリオの順序や条件との組み合わせの偏りが最小限になるように考慮して設定した。

それぞれの本番タスクは、3ステップで構成されている。1つ目のステップは個人タスクであり、各被験者は最初の5分間(トレーニングタスクは3分間)で個人のランキングとその理由を決定した。その後、次のステップで、Langrid Chat 上でグループメンバーと20分間(トレーニングタスクでは5分間)議論し、グループでのランキングとその理由を決定した。この時、被験者達はグループの順位が最初の個人ランキン

グに近くなるよう、他の被験者を説得するように指示された。最後のステップでは、10分間でアンケートに回答した。このアンケートは、被験者からの各項目に対する評価を収集するための質問が含まれている。具体的には、議論内容の理解度、議論への参加度、議論の満足度、作業負荷、高資源言語話者から低資源言語話者への配慮に関する質問であり、アンケートには先述にある項目順で記載された。作業負荷に関する質問は、主観的な作業負荷を測定するための既存のツールである、NASA-TLX (NASA Task Load Index) [18] を使用した。トレーニングタスクにおける各ステップでは、時間内で終わることができるように被験者達に見える場所にカウントダウン画面を表示した。各本番タスク、およびトレーニングタスクの状況、シナリオ、アイテムを以下に示す。

#### **トレーニングタスク / 砂漠サバイバル**

**シナリオ：** あなた達は小型飛行機でアリゾナ砂漠のどこかに不時着しました。現在は8月中旬の午前9時頃です。今日の地表温度は54℃（130°F）に達するでしょう。あなたと他の2名が唯一の生存者です。飛行機はいつ爆発してもおかしくないため、すぐに避難しなければなりません。避難前に、墜落現場から3つのアイテムを回収することができました。すべてのアイテムは良好な状態です。今、その3つのアイテムを生存に重要な順にランク付けする必要があります。

**アイテム：** 45口径のピストル(弾入り)、赤と白のパラシュート、メイク用の鏡

### 本番タスク1 / 海洋サバイバル

**シナリオ：** あなたと他の2名のメンバーは、大西洋のプライベートヨットで漂流しています。原因不明の火事によって、ヨットと船内の物の多くが破壊されました。ヨットは現在、ゆっくりと沈んでいっています。重要な航法機器も破壊されたので、位置情報は不明ですが、最寄りの陸地からは数百マイル離れているようです。以下の5つのアイテムが火事の後は無傷に回収できました。さらに、あなた達とアイテムを運ぶのに十分な大きさのゴム製のいかだもあります。あなたと他の2名のメンバーは、これらの5つのアイテムを生存に重要な順にランク付けする必要があります。

**アイテム：** 大西洋の地図、15 フィート (約4.5m) のロープ、サメ避けスプレー、剃刀用の鏡、20 平方フィート (約1.86m<sup>2</sup>) の不透明なプラスチック製シート

### 本番タスク2 / 北極サバイバル

**シナリオ：** あなた達は、小型飛行機の墜落から生還しました。パイロットと副操縦士は墜落で亡くなりました。現在は1月中旬で、北カナダにいます。日中の気温はマイナス25℃、夜間の気温はマイナス40℃です。地面には雪が積もっており、周囲は木々に覆われ、いくつかの小川が交差しています。最寄りの町まで20マイルの距離があります。あなた達は皆、ビジネス会議用の服装をしています。以下の5つのアイテムは墜落の後に無傷で回収できたものです。あなた達はこれらの5つのアイテムを生存に重要な順にランク付けしてください。

**アイテム：** 小さな斧、板チョコレート3つ、予備のシャツとパンツ3セット、新聞3部、ガスライター (燃料切れ)

### 本番タスク3 / 無人島サバイバル

**シナリオ：** あなた達は熱帯地域の無人島に流れ着きました。乗ってきた船は沈没してしまっています。災難なことに、嵐が近づいているようです。あなた達は、救助隊に見つけてもらうまで、島でしばらく生き延びなければなりません。以下の5つのアイテムは沈没しそうな船から回収しました。あなた達はこれらの5つのアイテムを生存に重要な順にランク付けしてください。

**アイテム：** 金属製の器、寝袋2つ、ペニシリン入りの応急処置キット、7日分のフリーズドライ食品、シャベル

## 5.5 分析

各エージェントの効果を明らかにするため、コミュニケーション実験を通じて得られたアンケートによる被験者の評価データや対話ログデータを用いて定量的な分析と定性的な分析を行った。本節では、それぞれの分析手法と結果を説明する。

### 5.5.1 議論の理解度の評価

ネパール語話者の議論の理解度に対する主観的評価を分析した。具体的には、3つの条件間で議論の理解度を比較するため、アンケートの議論内容の理解度についての質問項目に対するネパール語話者の回答を条件ごとに収集した。質問項目は表 5.2 の通りである。

Q1 の回答を用いた分析については、それぞれのタスクの対話ログを確認したところ、被験者達は議論開始時に個々のランキングを提示し、そこで3者間で順位が一致したアイテムについて議論されない場合など、各タスクにおいて議論されたアイテム数が異なっていた。そのため、各タスク内で全てのアイテムについて議論されたかどうかを考慮

表 5.2: 議論の理解度に関する質問項目

| 番号 | 質問文                                                                                                     |
|----|---------------------------------------------------------------------------------------------------------|
| Q1 | どのアイテムの議論(順位とその理由, 使い道など)を理解できましたか?(アイテムで複数選択)                                                          |
| Q2 | 他の参加者のメッセージをどの程度理解できたか?<br>(11 段階評価 / 0%~100%の 10%刻み)                                                   |
| Q3 | 他の参加者はあなたのメッセージをどの程度理解できていたと感じましたか?(11 段階評価 / 0%~100%の 10%刻み)                                           |
| Q4 | 参加者同士の母語が異なることが議論の理解度に影響したと思う。(5 段階評価 / 1: 全くそう思わない, 2: あまりそう思わない, 3: どちらとも言えない, 4: ややそう思う, 5: とてもそう思う) |

する必要がある。そこで、各タスクの対話ログから、議論されたアイテムを抽出した。このとき、単に被験者が自分のランキングやその理由を提示するだけの発話は、そのアイテムについて議論したとは言えないため、そのようなケースは対象外としている。その後、議論されたアイテムのうち、ネパール語話者の Q1 の回答から、理解できたアイテム数と理解できなかったアイテム数を算出した。次に、各エージェントを用いることで、理解できたアイテム数が向上するのを確認するため、各エージェント条件とエージェントなし条件の 2 条件間ごとに理解できたアイテム数と理解できなかったアイテム数を集計し、 $\chi^2$  検定 ( $\alpha = 0.05$ ) を行った。図 5.4, 図 5.5 は、検定に使用したクロス集計表を示している。それぞれの表の各セルは、実測値から算出される残差値に基づいてハイライトされている。残差値が 0 に近いほど白に、残差値が大きいほどピンクに、残差値が小さいほど紫に色付けられている。 $\chi^2$  検定の結果、要約要求エージェント条件とエージェントなし条件間の関連はなかった ( $\chi^2(1) = 3.300, p = .069$ ) が要約提供エージェント条件とエージェントなし条件間では関連性が認められた ( $\chi^2(1) = 5.576, p = .018$ )。次に、 $\chi^2$  検定で関連性が認められたクロス集計表について、どのセルに有意な差があるかを確認するため、調整済み残差を用いた

| 条件         | 理解できた<br>アイテム数 | 理解できなかった<br>アイテム数 | 残差値          |
|------------|----------------|-------------------|--------------|
| 要約要求エージェント | 13             | 9                 | $\geq 1.96$  |
| エージェントなし   | 7              | 15                | $\leq -1.96$ |

図 5.4: ネパール語話者の議論の理解できたアイテム数の集計表 (要約要求エージェント条件とエージェントなし条件間)

| 条件         | 理解できた<br>アイテム数 | 理解できなかった<br>アイテム数 |
|------------|----------------|-------------------|
| 要約提供エージェント | <u>16</u>      | <u>8</u>          |
| エージェントなし   | <u>7</u>       | <u>15</u>         |

図 5.5: ネパール語話者の議論の理解できたアイテム数の集計表 (要約提供エージェント条件とエージェントなし条件間)

残差分析を行った。残差分析で有意差が認められたセルを図 5.5 内の下線で示している。この結果から、要約提供エージェント条件では、エージェントなし条件よりも理解できたアイテム数が多くなることが示された。

さらに、話者言語間での理解できたアイテム数との関連も確認するため、条件別に言語話者ごとの理解できたアイテム数と理解できなかったアイテム数を集計し、 $\chi^2$  検定を行った。図 5.6, 図 5.7, 図 5.8 は、検定に使用したクロス集計表を示している。それぞれの表の各セルは、実測値から算出される残差値に基づいてハイライトされている。残差値が 0 に近いほど白に、残差値が大きいほどピンクに、残差値が小さいほど紫に色付けられている。その結果、要約要求エージェント条件 ( $\chi^2(1) = 0.852, p = .356$ ) と要約提供エージェント条件 ( $\chi^2(1) = 0.061, p = .804$ ) においては関連はなかったが、エージェントなし条件では関連性が認められた ( $\chi^2(1) = 11.005, p = .001$ )。次に、 $\chi^2$  検定で関連性が認められたクロス集計表について、どのセルに有意な差があるかを

| 言語話者    | 理解できた<br>アイテム数 | 理解できなかった<br>アイテム数 |
|---------|----------------|-------------------|
| 日本語話者   | 31             | 13                |
| ネパール語話者 | 13             | 9                 |

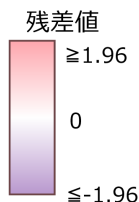


図 5.6: 言語話者ごとの議論の理解できたアイテム数の集計表 (要約要求エージェント条件)

確認するため、調整済み残差を用いた残差分析を行った。残差分析で有意差が認められたセルを図 5.8 内の下線で示している。この結果から、エージェントなし条件においては、ネパール語話者よりも日本語話者の方が議論を理解できたアイテム数が多いことが明らかになった。このことから、エージェントを用いた条件は、エージェントなし条件よりもネパール語話者の理解できたアイテム数が増加したと考えられる。

次に、Q2～Q4 の回答を用いた分析では、各エージェントを用いることで、理解度に関する評価が向上するのを確認するため、各エージェント条件とエージェントなし条件間でウィルコクソンの符号順位検定 ( $\alpha = 0.05$ ) を行った。図 5.9, 図 5.10, 図 5.11 に条件間の評価を比較する箱ひげ図を示している。また、それぞれの質問項目の分析結果は以下の通りである。

#### 要約要求エージェント条件とエージェントなし条件間

Q2  $W = 6.0, p = .850$

Q3  $W = 10.5, p = .272$

Q4  $W = 5.5, p = .341$

#### 要約提供エージェント条件とエージェントなし条件間

Q2  $W = 3.0, p = .498$

Q3  $W = 4.0, p = .850$

Q4  $W = 3.0, p = .571$

これらの結果から、どの質問項目においても条件間に有意な差は見られなかった。

| 言語話者    | 理解できた<br>アイテム数 | 理解できなかった<br>アイテム数 |
|---------|----------------|-------------------|
| 日本語話者   | 32             | 14                |
| ネパール語話者 | 16             | 8                 |

残差値

≥1.96

0

≤-1.96

図 5.7: 言語話者ごとの議論の理解できたアイテム数の集計表 (要約提供エージェント条件)

| 言語話者    | 理解できた<br>アイテム数 | 理解できなかった<br>アイテム数 |
|---------|----------------|-------------------|
| 日本語話者   | <u>32</u>      | <u>11</u>         |
| ネパール語話者 | <u>7</u>       | <u>15</u>         |

残差値

≥1.96

0

≤-1.96

図 5.8: 言語話者ごとの議論の理解できたアイテム数の集計表 (エージェントなし条件)

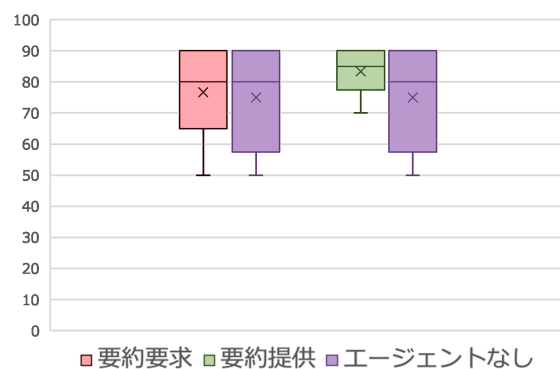


図 5.9: ネパール語話者の条件間における議論理解度 Q2 の箱ひげ図

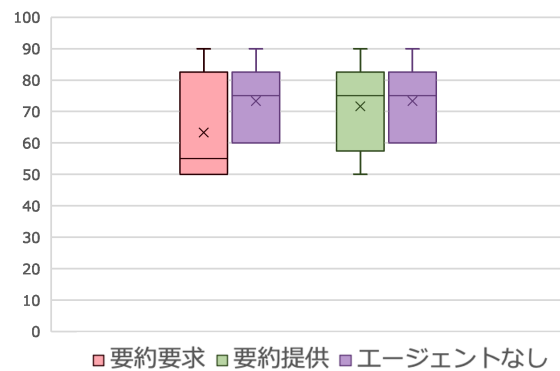


図 5.10: ネパール語話者の条件間における議論理解度 Q3 の箱ひげ図

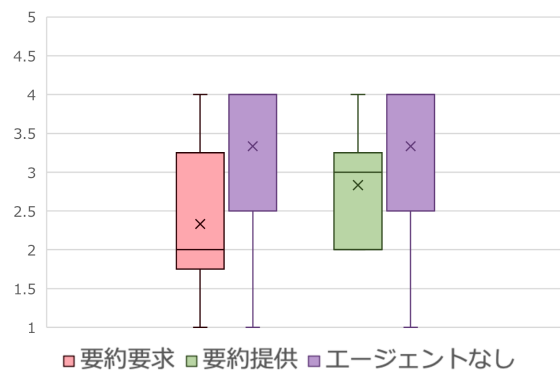


図 5.11: ネパール語話者の条件間における議論理解度 Q4 の箱ひげ図

## 5.5.2 議論への参加度の評価

ネパール語話者の議論への参加度に対する主観的評価と客観的評価を分析した。主観的評価には、アンケートの議論内容の参加度についての質問項目に対するネパール語話者の回答を使用し、客観的評価には、議論中のネパール語話者の1タスクあたりの文字数と発話数を使用した。

まず、主観的評価の分析については、アンケートの議論内容の参加度についての質問項目に対する回答を条件ごとに収集し、分析データ

表 5.3: 議論の参加度に関する質問項目

| 番号 | 質問文                                                                                                     |
|----|---------------------------------------------------------------------------------------------------------|
| Q1 | どのアイテムについての議論に参加できたか？<br>(アイテムで複数選択)                                                                    |
| Q2 | 自分の意見をどの程度発言することができたか？<br>(11 段階評価 / 0%～100%の 10%刻み)                                                    |
| Q3 | 参加者同士の母語が異なることが発話のし易さに影響したと思う。(5 段階評価 / 1: 全くそう思わない, 2: あまりそう思わない, 3: どちらとも言えない, 4: ややそう思う, 5: とてもそう思う) |

として使用した。質問項目は表 5.3 の通りである。

Q1 についての分析は、この質問項目への回答と、前項 5.5.1 の Q1 の分析の際に抽出した、各タスクにおいて議論されたアイテムを基に、ネパール語話者が議論に参加できたアイテム数とそれ以外のアイテム数を算出した。次に、各エージェントを用いることで議論に参加できたアイテム数が増加するかを確認するため、各エージェント条件とエージェントなし条件の 2 条件間ごとに議論に参加できたアイテム数と参加できなかったアイテム数を集計し、各クロス集計表を用いた  $\chi^2$  検定 ( $\alpha = 0.05$ ) を行った。図 5.12, 図 5.13 は集計結果のクロス集計表である。それぞれの表の各セルは、実測値から算出される残差値に基づいてハイライトされている。残差値が 0 に近いほど白に、残差値が大きいほどピンクに、残差値が小さいほど紫に色付けられている。 $\chi^2$  検定の結果、要約要求エージェント条件とエージェントなし条件間 ( $\chi^2(3) = 1.17, p = .761$ )、要約提供エージェント条件とエージェントなし条件間 ( $\chi^2(3) = 1.56, p = .668$ ) の、どちらの条件間でも関連性は認められなかった。つまり、エージェントを用いたことによる議論に参加できたアイテム数の増加は認められなかった。

Q2 と Q3 の回答を用いた分析は、各エージェントを用いることで、参加度に関する評価が向上するのを確認するため、各エージェント条件とエージェントなし条件間でウィルコクソンの符号順位検定 ( $\alpha = 0.05$ )

| 条件         | 参加アイテム数 | 非参加アイテム数 | 残差値          |
|------------|---------|----------|--------------|
| 要約要求エージェント | 13      | 9        | $\geq 1.96$  |
| エージェントなし   | 12      | 10       | $\leq -1.96$ |

図 5.12: ネパール語話者の議論に参加できたアイテム数の集計表 (要約要求エージェント条件とエージェントなし条件間)

| 条件         | 参加アイテム数 | 非参加アイテム数 | 残差値          |
|------------|---------|----------|--------------|
| 要約提供エージェント | 14      | 10       | $\geq 1.96$  |
| エージェントなし   | 12      | 10       | $\leq -1.96$ |

図 5.13: ネパール語話者の議論に参加できたアイテム数の集計表 (要約提供エージェント条件とエージェントなし条件間)

を行った。図 5.14 と図 5.15 は条件間の評価を比較する箱ひげ図を示している。また、それぞれの質問項目の分析結果は以下の通りである。

#### 要約要求エージェント条件とエージェントなし条件間

Q2  $W = 7.0, p = 1.000$

Q3  $W = 10.0, p = .572$

#### 要約提供エージェント条件とエージェントなし条件間

Q2  $W = 6.0, p = .784$

Q3  $W = 6.0, p = .773$

これらの結果から、どの条件間においても、有意差は認められなかった。つまり、エージェントを用いることによって参加度は向上しなかったと言える。

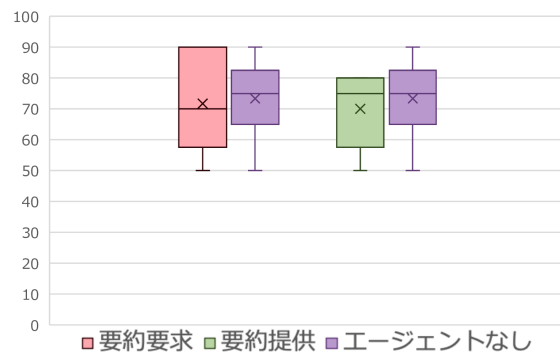


図 5.14: ネパール語話者の条件間における参加度 Q2 の箱ひげ図

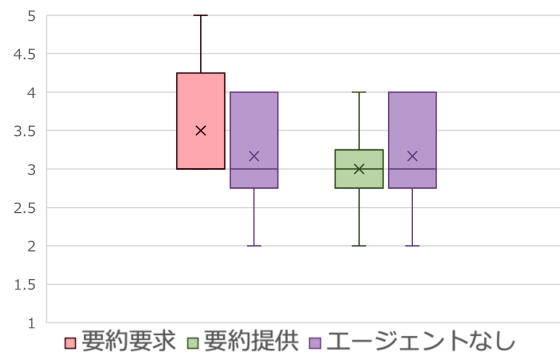


図 5.15: ネパール語話者の条件間における参加度 Q3 の箱ひげ図

次に、客観的評価については、対話ログから、1 タスクあたりのネパール語話者の発話数と文字数を収集し、評価データとして使用した。このとき、議論に用いた発話のみを対象とするため、議論前後の挨拶や、誤タイプなどによる言い直しの発話は除外した。

1 タスクあたりの文字数を用いた分析では、各サバイバルタスクのシナリオによって、議論に必要な文字数が異なることが考えられる。そのため、各シナリオにおける 1 タスクあたりのネパール語話者の平均文字数を算出した。この平均値を元に、各タスクにおける 1 タスクの文字数の偏差を算出し、分析データとして条件別に収集した。このとき、

ネパール語の対話ログを用いて文字数を収集した。表 5.4 は各シナリオの 1 タスクあたりのネパール語話者の平均文字数、図 5.16 は条件ごとに収集したデータを箱ひげ図で示している。これらのデータを用いて、各エージェントを用いることで 1 タスクあたりのネパール語話者の文字数は増加するのを確認するため、各エージェント条件とエージェントなし条件間に対応あり t 検定 ( $\alpha = 0.05$ ) を行った。その結果、要約要求エージェント条件とエージェントなし条件間 ( $t(10) = 0.035, p = .973$ )、要約提供エージェント条件とエージェントなし条件間 ( $t(10) = 0.637, p = .552$ ) の、どちらの条件間でも関連性は認められなかった。

1 タスクあたりの発話数を用いた分析では、文字数での分析と同様に、各サバイバルタスクのシナリオによって、議論に必要な発話数が異なることが考えられる。そのため、各シナリオにおける 1 タスクあたりのネパール語話者の平均発話数を算出した。この平均値を元に、各タスクにおける 1 タスクの発話数の偏差を算出し、分析データとして条件別に収集した。表 5.5 は各シナリオの 1 タスクあたりのネパール語話者の平均文字数、図 5.17 は条件ごとに収集したデータを箱ひげ図で示している。これらのデータを用いて、各エージェントを用いることで 1 タスクあたりのネパール語話者の発話数は増加するのを確認するため、各エージェント条件とエージェントなし条件間に対応あり t 検定 ( $\alpha = 0.05$ ) を行った。その結果、要約要求エージェント条件とエージェントなし条件間 ( $t(10) = 0.324, p = .759$ )、要約提供エージェント条件とエージェントなし条件間 ( $t(10) = 0.000, p = 1.000$ ) の、どちらの条件間でも関連性は認められなかった。

表 5.4: 各シナリオの 1 タスクあたりのネパール語話者の平均文字数

| シナリオ | 平均文字数  |
|------|--------|
| 海洋   | 232.00 |
| 北極   | 298.33 |
| 無人島  | 251.67 |

表 5.5: 各シナリオの 1 タスクあたりのネパール語話者の平均発話数

| シナリオ | 平均発話数  |
|------|--------|
| 海洋   | 10.167 |
| 北極   | 13.167 |
| 無人島  | 10.167 |

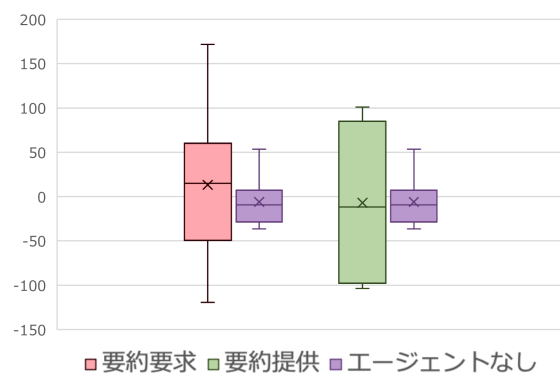


図 5.16: ネパール語話者の条件間における 1 タスクの文字数偏差の箱ひげ図

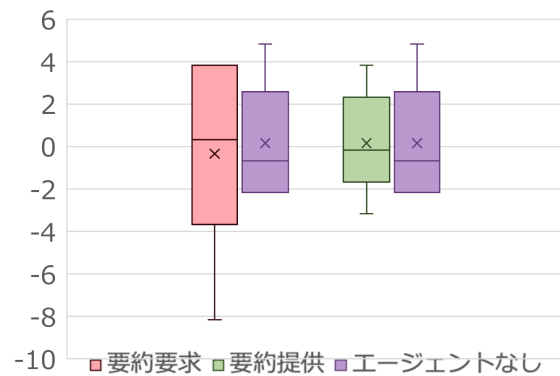


図 5.17: ネパール語話者の条件間における 1 タスクの発話数の箱ひげ図

表 5.6: 議論の満足度に関する質問項目

| 番号 | 質問文                                                         |
|----|-------------------------------------------------------------|
| Q1 | どのアイテムのランキング付けに貢献できたと思いますか？<br>(アイテムで複数選択)                  |
| Q2 | メンバーとの議論を通じて決定したランキングや理由の中で、<br>どのアイテムに満足していますか？(アイテムで複数選択) |

### 5.5.3 議論の満足度の評価

ネパール語話者の議論の満足度に対する主観的評価を分析した。具体的には、アンケートの議論によって決定した結論の満足度についての質問項目に対するネパール語話者の回答を使用した。

まず、主観的評価の分析については、アンケートの議論によって決定した結論の満足度についての質問項目に対するネパール語話者の回答で選択されたアイテム数を条件ごとに収集し、分析データとして使用した。質問項目は表 5.6 の通りである。

これらの質問への回答を用いて、各エージェントを用いることで、ネパール語話者の議論の満足度に関する評価が増加するのかを確認するため、各エージェント条件とエージェントなし条件間で対応あり t 検定 ( $\alpha = 0.05$ ) を行った。図 5.18 と図 5.19 は条件間の評価を比較する箱ひげ図を示している。また、それぞれの質問項目の分析結果は以下の通りである。

要約要求エージェント条件とエージェントなし条件間

Q1  $t(10) = 1.464, p = .203$

Q2  $t(10) = 1.195, p = .286$

要約提供エージェント条件とエージェントなし条件間

Q1  $t(10) = 1.112, p = .317$

Q2  $t(10) = 0.237, p = .822$

これらの結果から、どちらの質問項目においても条件間に有意な差は見られなかった。

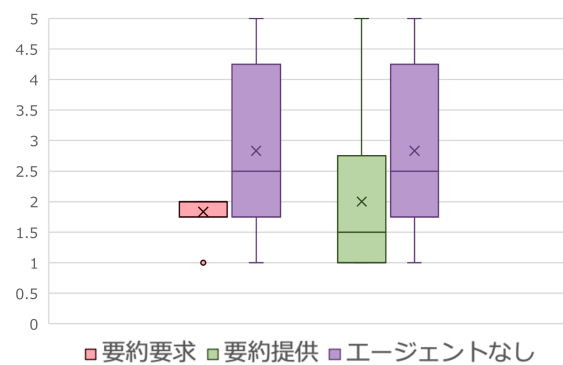


図 5.18: ネパール語話者の条件間における満足度 Q1 の箱ひげ図

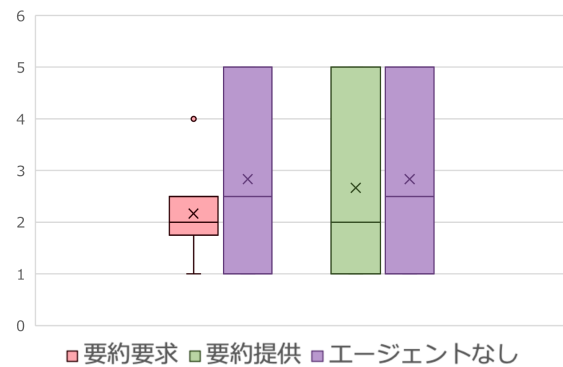


図 5.19: ネパール語話者の条件間における満足度 Q2 の箱ひげ図

次に、客観的評価の分析については、グループの議論によって決まったランキングと、議論を踏まえて決定したネパール語話者の最終ランキングの相関係数を算出し、収集した。このとき、ランキングが一致したアイテムについては議論されない傾向があったことから、ランキングが一致したアイテムを除くアイテムを対象とした。算出した相関係数を各条件ごとに集計し、分析データとして使用した。図 5.20 は条件間の評価を比較する箱ひげ図を示している。その後、各エージェントを用いることで、ネパール語話者の議論の満足度に関する評価が増加するのを確認するため、各エージェント条件とエージェントなし条件間で対応あり t 検定 ( $\alpha = 0.05$ ) を行った。その結果、要約要求エージェント条件とエージェントなし条件間 ( $t(10) = 1.697, p = .150$ )、要約提供エージェント条件とエージェントなし条件間 ( $t(10) = 1.746, p = .141$ ) の、どちらの条件間でも関連性は認められなかった。

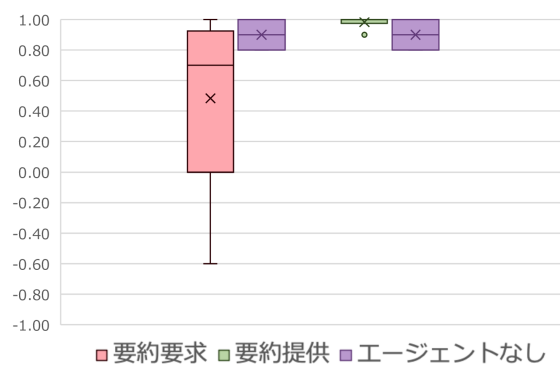


図 5.20: ネパール語話者の条件間における客観的満足度の箱ひげ図

#### 5.5.4 作業負荷の評価

日本語話者の作業負荷に対する評価を分析した。作業負荷の評価には、NASA-TLX[18]を使用した。これは、作業負荷を測定するための使用される一般的な手法である。具体的には、精神的負荷、身体的負荷、時間的圧迫感、パフォーマンス、努力、不満の6つの負荷の評価値を基に総合的な作業負荷を算出する方法である。日本語話者のアンケート

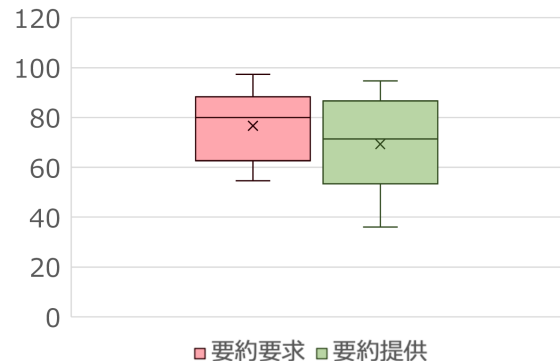


図 5.21: 日本語話者の条件間における作業負荷の箱ひげ図

での評価を基に算出した作業負荷を各条件で集計した。このとき、被験者実験を通して、要約要求エージェントによる要約要求メッセージは35回送信され、日本語話者はその内19回しか対応しておらず、要約要求メッセージ全てに対応したわけではなかった。しかし、全てのグループにおいて最低1回以上は要約要求エージェントの要約要求に対して要約を返したため、全ての日本語話者のデータを使用した。これらのデータを用いて、要約要求エージェントを用いることで日本語話者の作業負荷が増加するのを確認するため、要約要求エージェント条件と要約提供エージェント条件間で対応ありt検定( $\alpha = 0.05$ )を行った。図 5.21 は条件間の評価を比較する箱ひげ図を示している。分析の結果、関連性は認められ( $t(22) = 2.300, p = .043$ )、特に、要約要求エージェント条件の方が、作業負荷が高いことが明らかになった。

先述の分析によって、要約要求エージェントを用いることで、日本語話者の総合的な作業負荷が向上することは明らかになったが、具体的にどのような負荷が向上したのかは明らかにできない。そのため、NASA-TLX を用いて総合的な作業負荷を算出ために用いる、精神的負荷、身体的負荷、時間的圧迫感、パフォーマンス、努力、不満の6つの負荷の評価値を使用し、これらの負荷のうち、日本語話者のどの負荷を向上させたのかを確認した。具体的には、6つの負荷に対する各被験者の評価を収集し、各負荷で要約要求エージェント条件と要約提供エージェン

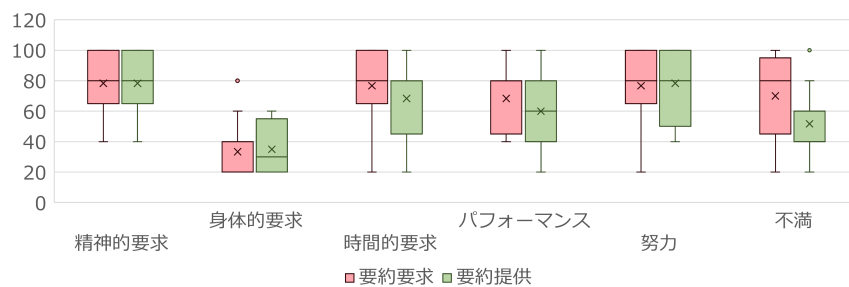


図 5.22: 日本語話者の条件間における 6 つの負荷の箱ひげ図

ト条件間で対応あり t 検定 ( $\alpha = 0.05$ ) を行った。図 5.22 に条件間の評価を比較する箱ひげ図を示している。また、それぞれの負荷についての分析結果は以下の通りである。

|         |                            |
|---------|----------------------------|
| 精神的負荷   | $t(10) = 0.000, p = 1.000$ |
| 身体的負荷   | $t(10) = 0.364, p = .723$  |
| 時間的圧迫感  | $t(10) = 1.000, p = .338$  |
| パフォーマンス | $t(10) = 0.890, p = .392$  |
| 努力      | $t(10) = 0.432, p = .674$  |
| 不満      | $t(10) = 2.421, p = .034$  |

これらの結果から、「不満」の負荷のみに関連性が認められ、要約要求エージェントの方が高く評価されていた。

### 5.5.5 低資源言語話者への配慮度の評価

日本語話者の低資源言語話者への配慮度に対する主観的評価と客観的評価を分析した。主観的評価には、アンケートのネパール語話者への配慮度についての質問項目に対する日本語話者の回答を使用し、客観的評価には、議論中の日本語話者によるネパール語話者への配慮発言数を使用した。

まず、主観的評価の分析については、アンケートのネパール語話者への配慮度についての質問項目に対する日本語話者の回答を条件ごとに

表 5.7: 議論の配慮に関する質問項目

| 番号 | 質問文                                                                                                        |
|----|------------------------------------------------------------------------------------------------------------|
| Q1 | ネパール語話者が議論内容を理解できるようにどのくらい配慮しましたか？(5段階評価 / 1: 全く配慮しなかった, 2: ほとんど配慮しなかった, 3: 多少配慮した, 4: よく配慮した, 5: 十分に配慮した) |

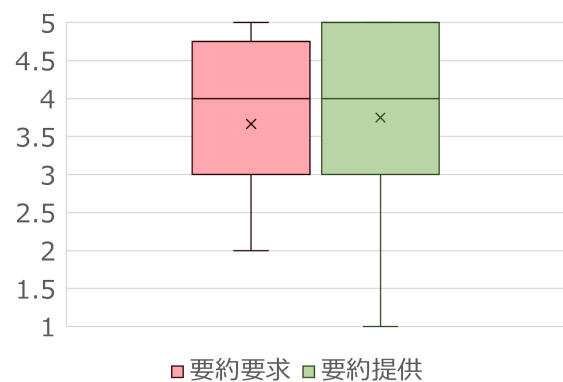


図 5.23: 日本語話者の条件間における低資源言語話者への主観的配慮度の箱ひげ図

収集し、分析データとして使用した。質問項目は表 5.7 の通りである。要約要求エージェントを用いることで、日本語話者のネパール語話者への配慮する発話が増加するのを確認するため、要約要求エージェント条件と要約提供エージェント条件間でウィルコクソンの符号順位検定 ( $\alpha = 0.05$ ) を行った。図 5.23 に条件間の評価を比較する箱ひげ図を示している。分析結果、要約要求エージェント条件と要約提供条件間 ( $W = 7.5$ ,  $p = 1.000$ ) の、どちらの条件間でも関連性は認められなかった。

次に、客観的評価については、対話ログから、日本語話者によるネパール語話者への配慮する発話数をグループ別に集計し、評価データ

表 5.8: 日本語話者によるネパール語話者への配慮する発話の種類

| 種類 | 説明                                                                                                    |
|----|-------------------------------------------------------------------------------------------------------|
| 質問 | ネパール語話者の意見を引き出そうと質問する発話.<br>例) 「〇〇さんはどう思いますか?」「(ネパール語話者の質問に対して) なぜそのように思いますか?」<br>など                  |
| 確認 | ネパール語話者の発話を理解しようと確認する発話.<br>例) 「(ネパール語話者の発話に対して, ) どういう意味ですか?」 など                                     |
| 反応 | ネパール語話者の発話に対して理解していると示すための反応の発話.<br>例) 「～～ということですね.」 など                                               |
| 要約 | 自主的に議論を要約した発話. ただし, 要約要求エージェントからの要約要求に対応するための要約は含まない.<br>例) 「(議論で決まったことについてまとめると) ランキングは～～でいいですか?」 など |

として使用した. このとき, 対象とした発話は表 5.8 に示した 4 種類である. 要約要求エージェント条件と要約提供エージェント条件間で日本語話者によるネパール語話者への配慮する発話数と関連性があるかどうかを確認するため, 対応あり  $t$  検定 ( $\alpha = 0.05$ ) を行った. 図 5.24 に条件間の評価を比較する箱ひげ図を示している. その結果, 要約要求エージェント条件と要約提供エージェント条件間で関連性が認められ ( $t(10) = 2.803, p = .037$ ), 特に, 要約要求エージェント条件の方が配慮する発話数が多くあった.

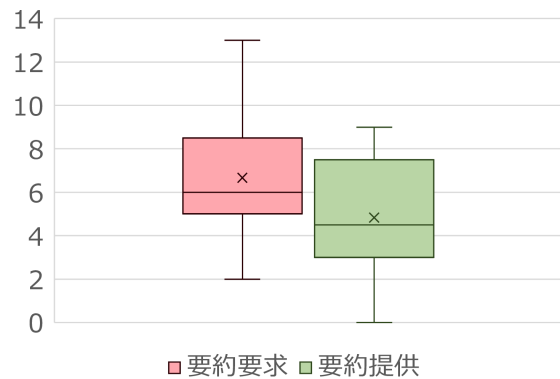


図 5.24: 日本語話者の条件間における低資源言語話者への客観的配慮度の箱ひげ図

## 5.6 考察

議論の理解度については、前節での分析から、要約提供エージェント条件でエージェントなし条件よりも、ネパール語話者の理解できるアイテム数が増加したことが明らかになった。つまり、要約提供エージェントを用いることで理解できた項目数が増加したと言える。一方で、理解度の Q2 の他の参加者のメッセージの理解度の評価においては、要約提供エージェント条件とエージェントなし条件間で有意差は認められなかった。また、図 5.9 から、エージェントなし条件でも平均 80% ほど理解できており、メッセージは概ね理解できたことを示している。これらのことから、ネパール語話者は他の参加者のメッセージ自体は理解できるものの、そのメッセージ同士が複雑に関係し合う議論を把握することは困難であり、要約提供エージェントが提供する要約によって、議論を把握でき、理解できた議論の項目数が増加したと考えられる。さらに、表 5.9 は要約提供エージェントを用いた条件でのタスクの対話ログの一部である。発話 1 でネパール語話者はアイテムの 1 つである「予備のシャツとパンツ 3 セット」について言及している。その後、発話 6 でファシリテータが要約を提供する発話を行っている。要約では、「予備のシャツとパンツ 3 セット」と「板チョコレート 3 つ」のアイテム

表 5.9: 対話事例 1 (要約提供エージェント/北極サバイバル)

| No. | 発話者     | メッセージ                                                          |
|-----|---------|----------------------------------------------------------------|
| 1   | ネパール語話者 | とても寒かったので先に <u>服</u> を洗濯しました                                   |
| 2   | 日本語話者 1 | 確かに寒さを防ぐのは大事なので、1 番は <u>シャツ</u> にしますか                          |
| 3   | 日本語話者 2 | <u>新聞</u> はくしゃくしゃにして包まれば、暖かいかなと思いました                           |
| 4   | 日本語話者 1 | <u>チョコレート</u> は生き延びるための食料として最適ですね                              |
| 5   | 日本語話者 2 | 一番 <u>シャツ</u> にしましょ!!                                          |
| 6   | ファシリテータ | 今は寒さ対策として <u>シャツ</u> を一番にすること、 <u>チョコ</u> は食料として大切であると議論しています。 |
| 7   | ネパール語話者 | <u>チョコレートボール</u> の前にベルト食べた分のミルク                                |
| 8   | 日本語話者 2 | 二番目は <u>チョコレート</u> にしましょ!                                      |
| 9   | 日本語話者 2 | <u>食べ物</u> がなかったら死んじゃいます                                       |
| 10  | 日本語話者 1 | 3 番目はどうしましょうか、 <u>斧</u> があれば便利だと思うのですがどうでしょうか?                 |
| 11  | ネパール語話者 | 私も同意します、ハニー                                                    |
| 12  | 日本語話者 2 | そうしましょ～～!                                                      |
| 13  | 日本語話者 1 | いいね                                                            |

について議論されていると示されている。その後、ネパール語話者は、要約に示されていたアイテムの内、言及していなかった「板チョコレート 3 つ」について言及した。一方で、表 5.10 は、同グループのエージェントなし条件でのタスクの対話ログの一部である。日本語話者達の議論に着目すると、はじめは「20 平方フィート (約  $1.86m^2$  の不透明なプラスチック製シート)」について議論し、その次に「剃刀用の鏡」につい

表 5.10: 対話事例 2 (エージェントなし/海洋サバイバル)

| No. | 発話者     | メッセージ                                                                          |
|-----|---------|--------------------------------------------------------------------------------|
| 1   | 日本語話者 1 | <u>シート</u> は雨を避け、風を受けてボートを進めることができます。                                          |
| 2   | ネパール語話者 | <u>スプラリー</u> に黄色い蜂蜜を混ぜてみました                                                    |
| 3   | 日本語話者 2 | あー確かにそうですね… <u>シート</u> は何に使うのかと思いました                                           |
| 4   | 日本語話者 2 | <u>鏡</u> は太陽を反射させて助けを呼ぶことができます<br>と思いました                                       |
| 5   | 日本語話者 1 | (ネパール語話者) さん、どの順番が一番いい<br>と思うか教えて欲しいです                                         |
| 6   | 日本語話者 1 | <u>鏡</u> の使い方がわからなかったのですが、使え<br>そうですね                                          |
| 7   | ネパール語話者 | ボートが何かをしても止まらなくなったら、<br><u>プラスチック</u> 全体が <u>プラスチック</u> になってしま<br>います。         |
| 8   | 日本語話者 2 | <u>シート</u> があれば船を漕ぐことができますね！                                                   |
| 9   | 日本語話者 1 | <u>プラスチックのシート</u> は何番目がいいですか<br>？                                              |
| 10  | ネパール語話者 | <u>スプリング</u> < <u>プラスチック</u> < <u>ロープ</u> < <u>ノッ<br/>ク</u> < <u>ミラー</u>       |
| 11  | 日本語話者 1 | <u>サメのスプレー</u> < <u>プラスチック</u> < <u>ロープ</u> <<br><u>大西洋マップ</u> < <u>鏡</u> ですか？ |
| 12  | 日本語話者 2 | じゃあ被っているもので、5 位は <u>鏡</u> にしまし<br>よ！                                           |
| 13  | ネパール語話者 | はい、先生                                                                          |
| 14  | 日本語話者 2 | <u>サメのスプレー</u> がどんなものか分からないの<br>ですが、 <u>プラスチック</u> の方が何かと使えそう<br>だなと思いました      |
| 15  | 日本語話者 1 | わたしも賛成です。                                                                      |

での議論を行っている。しかし、ネパール語話者は、「20 平方フィート (約  $1.86m^2$ ) の不透明なプラスチック製シート」についてしか言及していない。このことから、要約提供エージェントを用いることによって、議論されているアイテムの議論に包括的に参加できたと考えられる。さらに、対話ログから、要約提供エージェント条件では、日本語話者達が「板チョコレート 3 つ」について議論している最中に、遅れることなくネパール語話者もこのアイテムについて意見している。一方で、エージェントなし条件では、日本語話者達が「20 平方フィート (約  $1.86m^2$ ) の不透明なプラスチック製シート」について議論した後に、次のアイテムの議論に移ってから意見している。このことから、要約提供エージェントを用いることによって、議論をすぐに理解でき、議論についていくことができたと考えられる。

一方で、要約要求エージェントは要約提供エージェントと同様に、ネパール語話者に要約を提供する目的のエージェントである。しかし、要約要求エージェントは議論を理解できたアイテム数との関連性は認められなかった。このことについて、各エージェント条件での要約の仕方が異なる可能性が考えられるため、エージェントによる要約と日本語話者による要約を比較した。表 5.11 は、要約提供エージェントの要約例である。要約には、議論されているアイテムと、それについての意見が提示されている。一方で、表 5.12 は、要約要求エージェントの要約例である。要約には、1 位から決めているということしか提示されていない。つまり、どのアイテムについてどんな意見があるのかを理解するには、日本語話者の発話を確認するほかない。このことから、ネパール語話者にとって、日本語話者からの要約は、議論を理解するには不十分であったと考えられる。また、このタスク中には、他にもエージェントからの要約要求により、日本語話者は要約を送信したが、「次は 3 位から話合います」や、「3 位を決めています」などの表 5.12 に示した要約例と同じような要約であった。さらに、日本語話者が要約を送信することによって、単純にチャット画面内のメッセージが増える。そのため、ネパール語話者が議論を理解しようと読む、チャットメッセージ数は多くなる。そのため、要約要求エージェントを用いたタスクでは、理解するのに時間がかかり、議論に追いついていけなくなることが考えられる。これらのことから、仮説の H1.1 は要約要約エージェントに

表 5.11: 要約例 1 (要約提供エージェント/無人島サバイバル)

| No. | 発話者     | メッセージ                                                            |
|-----|---------|------------------------------------------------------------------|
| 1   | 日本語話者 1 | 夜寒そうですね                                                          |
| 2   | 日本語話者 2 | 雨に濡れると体温下がりますよね                                                  |
| 3   | 日本語話者 1 | 寝袋も防水ではないでしょう                                                    |
| 4   | 日本語話者 2 | それでもある程度は凌げます                                                    |
| 5   | ファシリテータ | 今は寝袋の重要性について議論しています。<br><u>寝袋は寒さや体温低下対策として重要だとの<br/>意見があります。</u> |

表 5.12: 要約例 2 (要約要求エージェント/北極サバイバル)

| No. | 発話者     | メッセージ                                      |
|-----|---------|--------------------------------------------|
| 1   | 日本語話者 1 | 予備のシャツパンツって防寒具になりますか                       |
| 2   | 日本語話者 2 | 1 位はシャツですか？                                |
| 3   | 日本語話者 1 | 一応防寒具にはなりそうですけど                            |
| 4   | ネパール語話者 | 寒いのでまずは暖かくしてください                           |
| 5   | 日本語話者 2 | 効果は薄そうです                                   |
| 6   | 日本語話者 1 | 薄いですがもんね生地が                                |
| 7   | 日本語話者 2 | そうですね                                      |
| 8   | 日本語話者 1 | やはりこれも食料ではないですか                            |
| 9   | ネパール語話者 | 2 番目の斧                                     |
| 10  | ファシリテータ | 議論についていけなくなったので、今どんな<br>議論をしているのか要約してください。 |
| 11  | 日本語話者 1 | <u>現在、上位から何がランクインしているか話<br/>合っています</u>     |
| 12  | 日本語話者 2 | <u>1 位を決めています</u>                          |

関しては支持されず、要約提供エージェントに感じては支持される。

議論の参加度について、要約要求エージェントと要約提供エージェントのどちらにおいても、向上させる効果が認められなかった。このことから、仮説の H1.2 は支持されなかった。このことについて、本実験においては3名での少数グループの議論であり、発話の順番の意識しやすく、ネパール語話者は発言しなければいけないと感じ、ファシリテータエージェントを用いた場合と用いない場合で、統計的に有意差が認められるほどの発話数や文字数の差が現れなかった可能性がある。

議論の満足度について、要約要求エージェントと要約提供エージェントのどちらにおいても、向上させる効果が認められなかった。このことから、仮説の H1.3 は支持されなかった。このことについて、満足度の評価は、自分の意見がグループの意見に反映された、自分意見であるかによらず納得できたか、他の参加者と合意できたかなど、多角的に評価できる。そのため、本実験で用いた議論の満足度に関する質問項目では、正確に満足度を測ることができなかった可能性がある。

高資源言語話者の作業負荷について、要約要求エージェントを用いる方が、要約提供エージェントよりも総合的な作業負荷が有意に高かった。このことから、仮説の H2.1 は支持される。特に、作業負荷のうち、不満が増加し、時間的圧迫感も増加傾向が見られた。このことについて、日本語話者に、ファシリテータの印象についてインタビューを行ったところ、要約要求エージェントに対し、「特に困っていないタイミングだったので要約する手間ができただけに感じた」、「ネパールの方の発言量が多ければ役に立ったのかもしれないけど、ほとんど自分の発言の要約だったので必要に感じなかった。」、「少し話を遮ることがあった」、「無い方が議論の時間は伸びていたが、無ければ議論を整理することなかったため、どちらが良いか分からないから」、「議論があまりうまく進めていなかったから、構っていられなかった」などと回答し、時間的圧迫や不満についてが多く言及された。これらのことから、要約を作成するための思考などの精神的な負担の増加ではなく、限られた時間でタスクを行う中でタスクが増えてしまうという時間的圧迫感や、自分の役に立たない余分なタスクを行わないといけないことや議論を遮られることへの不満が増加した。

低資源言語話者への配慮について、日本語話者は、要約要求エージェ

ントを用いることで、主観的には配慮していることを感じていないものの、客観的に見ると配慮する発言が増加していた。このことから、仮説の H2.2 は支持される。このことについて、要約要求エージェントによるメッセージには、議論についていけなくなったと言及している。そのため、日本語話者は、このメッセージによってグループ内に議論についていけない参加者がいることを認識し、配慮する発言をしたと考えられる。

最後に、今後の研究の展望として、要約要求エージェントは日本語話者にとっては、エージェントからの要約要約メッセージには要約にどのような情報が必要かが示されていなかったため、どのように要約すべきか理解していなかった可能性があり、要約例などの要約方針を明示することで、日本語話者から提供される要約が改善され、理解度を向上する効果する可能性がある。さらに、この要約要求エージェントについて、今回の被験者実験を通して 35 回の要約要求メッセージを送信したが、その内、19 回しか日本人は要約を行わなかった。つまり、要約要求メッセージに対し、半数ほどしか応答しなかった。たとえ、エージェントからの要約要求メッセージに適切な要約文を示し、日本語話者が適切な要約ができるようになったとしても、そもそも要約されなければ無意味である。そのため、メッセージ改善と共に、日本語話者による要約を促進するための機能の追加を検討する必要があると考えられる。要約要求エージェント条件の対話ログから、はじめの何度かは要約要求メッセージには応じたが、それ以降は応じなくなったケースが多かった。このことから、日本語話者が要約した際に、お礼などの何らかの反応を示し、継続して要約してもらいやすくすることが有効かもしれない。

## 5.7 まとめ

本研究では、低資源言語話者の議論理解を促進させるために、要約に着目した、2 種類のファシリテータエージェントを実装した。具体的には、高資源言語話者に要約するように要求するエージェントと、エージェントが要約を提供するエージェントである。これらのエージェント

を用いることによる低資源言語話者の議論の理解度と参加度、満足度の向上の効果と、要約要求エージェントを用いることによる高資源言語話者の作業負荷、低資源言語話者への配慮度の向上の効果の有効性を検証するため、コミュニケーション実験を行った。その結果、要約提供エージェントは、低資源言語話者の理解できた議論項目数を増加させる効果が示された。しかし、議論の参加度や満足度の向上は示されなかった。また、要約要求エージェントは、高資源言語話者による要約に含まれる情報量の少なさから、低資源言語話者の議論の理解度、参加度、満足度を向上させる効果は認められなかった。また、このエージェントを用いることで高資源言語話者から低資源言語話者への配慮する発言は増加する一方で、高資源言語話者の作業負荷が増加し、特に不満を強く感じる事が示された。

また、今後の課題としては、ファシリテータエージェントを検証するためのコミュニケーション実験は6グループ行った。今回の実験での独立変数は3である。今回の実験試行回数は、実験計画法に基づく検証するために必要な最低限のデータであった。そのため、さらに多くのデータを収集して分析の方が好ましい。また、今回は、低資源言語話者が要約を受け取ることのみに着目しており、要約の質までは考慮していない。そのため、議論を理解するためにはどのような要約文が適切かに焦点をあて、要約要求のメッセージや、ChatGPTのプロンプト文を改善することで、さらに有効性の高いファシリテータエージェントが実現できると考えられる。

最後に、今後の展望としては、本研究で構築した要約に着目したファシリテータエージェントは、異文化コラボレーションの場での低資源言語話者のみに限らず、同一言語話者間や高資源言語話者間でのコミュニケーションの場においての議論の理解が追いつけないユーザに対しても効果的であると考えられる。そのため、異文化コラボレーションに限らず、他のコミュニケーションの場においても適用し、その有効性を検証することで、幅広いコミュニケーションを支援するためのエージェントとして貢献できる可能性がある。

## 第6章 結論

### 6.1 本研究のまとめ

本論文では、異文化コラボレーションの場での低資源言語話者を支援するためのファシリテータの対話行為の分析や、議論理解支援のためのファシリテータエージェントの構築に向けて行ってきた研究を述べた。

第3章では、ファシリテータの対話行為を分析するためのアノテーションスキーマの提案を行った。ファシリテータの対話行為を定量的に分析するためには、アノテーションが必要になる。しかしながら、既存の発話アノテーションのスキーマは、汎用的に使用できるように設計されている。予備実験として、実際の異文化コラボレーションの場でのファシリテータの発話アノテーションに適用してみた結果、定義に曖昧性や不明瞭な点があり、アノテータが困難であったり、アノテータによって解釈が異なることがあり、適用が難しかった。そのため、本研究では、既存のアノテーションスキーマを基に修正や追記などを行い、ファシリテータの対話行為の分析のための発話アノテーションに適用可能なスキーマの提案を行った。

第4章では、先行研究で提案した発話アノテーションスキーマを用いて、実際の異文化コラボレーションの場での対話ログに適用し、ファシリテータの発話を定量的に分析した。その結果、既存の異文化コラボレーションの場でのファシリテーションの課題点と、低資源言語話者の発話を促進するために重要なファシリテータの発話を明らかにした。

第5章では、低資源言語話者の対話内容の理解を支援するためのファシリテータエージェントの構築を行った。先行研究で分析で述べた、低資源言語話者の発話促進のためのファシリテータによる対話行為は、対

話内容を理解できていることを前提としている。つまり、低資源言語話者は対話内容を理解していなければ、意見を言うことができない。そのため、本研究では、要約要求エージェントと要約提供エージェントの2種類のエージェントを実装した。また、これらを用いたコミュニケーション実験を行い、有効性を確認した。

本研究は、ファシリテータの対話行為分析に適用可能なアノテーションスキーマの提案、異文化コラボレーションの場での既存のファシリテーションへのフィードバック、議論理解を支援するファシリテータエージェントを構築したことにより、異文化コラボレーション研究において、低資源言語話者の対話を理解し発話を促進させるための支援に貢献したと考えられる。

## 6.2 今後の展望

フィールド分析では、翻訳前後で対話行為が変わってしまい、送信者の意図を伝わらないことがあることが明らかになったことから、正確な意味の翻訳精度の向上だけでなく、話し手の意図が正確に聞き手に伝わるかを考慮した翻訳を目指す。

また、ファシリテータエージェントを用いた異文化コラボレーション支援では、低資源言語話者の発話促進や対話理解を支援することを目的として、低資源言語話者が翻訳品質が低い中でも、他の言語話者とのコミュニケーションに参加できることを目標としている。しかし、多言語の話者同士のコミュニケーションを活性化させるためには、単に参加するだけでなく、今後は低資源言語が高資源言語話者に向けてどれだけ自分の意見を正確に伝えられるための支援のための研究も行っていくべきである。

さらに、本研究では、要約を提供するエージェントを実装し、低資源言語話者の対話の理解度に寄与する効果が示された。一方で、高資源言語話者にとっては、話を遮ったり、不必要な要約を提供されることから、不満を感じながらの対話になりかねない。そのため、多様な言語の言語話者による円滑なコミュニケーションを実現するには、多くのコストや労力をかけたとしても、低資源言語の言語資源を拡大し、翻

訳精度の向上することが最も理想的であると考える.



## 参考文献

- [1] Hayati Abdolmajid. The effect of speech rate on listening comprehension of efl learners. *Creative Education*, Vol. 01, pp. 107–114, 2010.
- [2] Milam Aiken. Multilingual communication in electronic meetings. *ACM SIGGROUP Bulletin*, Vol. 23, No. 1, pp. 18–19, 2002.
- [3] James Allen and Mark Core. *Draft of DAMSL Dialog Act Markup in Several Layers*, 2007.
- [4] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.
- [5] Arendse Bernth and Claudia Gdaniec. Mtranslatability. *Machine translation*, Vol. 16, pp. 175–218, 2001.
- [6] Ann R. Bradlow and Tessa Bent. Perceptual adaptation to non-native speech. *Cognition*, Vol. 106, No. 2, pp. 707–729, 2008.
- [7] Harry Bunt. The dit++ taxonomy for functional dialogue markup. In *Proc. of AAMAS Workshop “Towards a Standard Markup Language for Embodied Dialogue Acts (EDAML 2009)”*, pp. 13–24, 2009.
- [8] Harry Bunt, Jan Alexandersson, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Volha Petukhova, Andrei Popescu-Belis, and David Traum. ISO 24617-2: A semantically-based standard for dialogue annotation. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, pp. 430–437, 2012.

- [9] Heeryon Cho, Toru Ishida, Naomi Yamashita, Rieko Inaba, Yumiko Mori, and Tomoko Koda. Culturally-situated pictogram retrieval. In *Intercultural Collaboration*, pp. 221–235. Springer Berlin Heidelberg, 2007.
- [10] Wen Duan, Naomi Yamashita, and Susan R. Fussell. Increasing native speakers’ awareness of the need to slow down in multilingual conversations using a real-time speech speedometer. Vol. 3, No. CSCW, 2019.
- [11] Wen Duan, Naomi Yamashita, and Susan R. Fussell. Increasing native speakers’ awareness of the need to slow down in multilingual conversations using a real-time speech speedometer. *Proc. ACM Hum.-Comput. Interact.*, Vol. 3, No. CSCW, 2019.
- [12] Andy Echenique, Naomi Yamashita, Hideaki Kuzuoka, and Ari Hautasaari. Effects of video and text support on grounding in multilingual multiparty audio conferencing. In *Proceedings of the 5th ACM International Conference on Collaboration across Boundaries: Culture, Distance Technology, CABS ’14*, pp. 73–81. Association for Computing Machinery, 2014.
- [13] Alan J. Feely and Anne-Wil Harzing. Language management in multinational companies. *Cross Cultural Management: An International Journal*, Vol. 10, pp. 37–52, 2003.
- [14] Raymond S. Flournoy and Chris Callison-Burch. Secondary benefits of feedback and user interaction in machine translation tools. In *Workshop on MT2010: Towards a Road Map for MT*, 2001.
- [15] Kaname Funakoshi, Akishige Yamamoto, Saeko Nomura, and Toru Ishida. Lessons learned from multilingual collaboration in global virtual teams. In *HCI (4)*, pp. 1098–1102. Citeseer, 2003.
- [16] Ge Gao, Naomi Yamashita, Ari MJ Hautasaari, Andy Echenique, and Susan R. Fussell. Effects of public vs. private automated transcripts

- on multiparty communication between native and non-native english speakers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 843–852. Association for Computing Machinery, 2014.
- [17] Zixuan Guo and Tomoo Inoue. Using a conversational agent to facilitate non-native speaker’s active participation in conversation. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pp. 1–6. Association for Computing Machinery, 2019.
  - [18] Peter A. Hancock and Najmedin Meshkati. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Human Mental Workload*, Vol. 52 of *Advances in Psychology*, pp. 139–183, 1988.
  - [19] Naoko Hayashida and Toru Ishida. Supporting urban communication with machine translation: Towards intercultural communication support agents. 2005.
  - [20] Cho Heeryon, 石田亨, 山下直美, 稲葉利江子, 高崎俊之, 神田智子. 絵文字解釈における人間の文化差判定. ヒューマンインタフェース学会論文誌, Vol. 10, No. 4, pp. 427–434, 2008.
  - [21] Rieko Inaba. Usability of multilingual communication tools. In *Usability and Internationalization. Global and Local User Interfaces*, pp. 91–97. Springer Berlin Heidelberg, 2007.
  - [22] Rieko Inaba, Yohei Murakami, Akiyo Nadamoto, and Toru Ishida. Multilingual communication support using the language grid. In *Intercultural Collaboration*, pp. 118–132. Springer Berlin Heidelberg, 2007.
  - [23] Toru Ishida, Yohei Murakami, Donghui Lin, Takao Nakaguchi, and Masayuki Otani. Language service infrastructure on the web: The language grid. *IEEE*, Vol. 51, No. 6, pp. 72–81, 2018.

- [24] Takayuki Ito, Yuma Imi, Takanori Ito, and Eizo Hideshima. Collagree: Facilitator-mediated large-scale consensus support system. In *The 2nd Collective Intelligence Conference*, 2014.
- [25] Sirkka L. Jarvenpaa and Dorothy E. Leidner. Communication and trust in global virtual teams. *Organization science*, Vol. 10, No. 6, pp. 791–815, 1999.
- [26] Austin Langshaw John. *How to do things with words*. Oxford University Press, 1975.
- [27] Soomin Kim, Jinsu Eun, Joseph Seering, and Joonhwan Lee. Moderator chatbot for deliberative discussion: Effects of discussion structure and discussant facilitation. *Proc. ACM Hum.-Comput. Interact.*, Vol. 5, No. CSCW1, p. 26, 2021.
- [28] Ryunosuke Kirikihira and Kazutaka Shimada. Discussion map with an assistant function for decision-making: A tool for supporting consensus-building. In *Collaboration Technologies and Social Computing*, pp. 3–18. Springer International Publishing, 2018.
- [29] Tomoko Koda. Interpretation of emotionally expressive characters in an intercultural communication. In *Knowledge-Based Intelligent Information and Engineering Systems*, pp. 862–868. Springer Berlin Heidelberg, 2004.
- [30] Philipp Koehn and Rebecca Knowles. Six challenges for neural machine translation. In *Proceedings of the First Workshop on Neural Machine Translation*, pp. 28–39. Association for Computational Linguistics, 2017.
- [31] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp. 230–237. Association for Computational Linguistics, 2004.

- [32] Salla Kurhila. Correction in talk between native and non-native speaker. *Journal of Pragmatics*, Vol. 33, No. 7, pp. 1083–1110, 2001.
- [33] David Kurlander, Tim Skelly, and David Salesin. Comic chat. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 225–236, 1996.
- [34] James Clayton Lafferty, Human Synergistics, and Alonzo William Pond. *The desert survival problem*. 1974.
- [35] Richard Landis and Gary Grove Koch. The measurement of observer agreement for categorical data. *Biometrics*, Vol. 33, No. 1, pp. 159–174, 1977.
- [36] Xiaoyan Li, Naomi Yamashita, Wen Duan, Yoshinari Shirai, and Susan R. Fussell. Improving non-native speakers’ participation with an automatic agent in multilingual groups. *Proc. ACM Hum.-Comput. Interact.*, Vol. 7, No. GROUP, pp. 1–28, 2022.
- [37] Masafumi Matsuda and Yasuhiko Kitamura. Development of machine translation system for japanese children. In *Proceedings of the 2009 International Workshop on Intercultural Collaboration, IWIC ’09*, pp. 269–272. Association for Computing Machinery, 2009.
- [38] Mai Miyabe, Takashi Yoshino, and Tomohiro Shigenobu. Effects of undertaking translation repair using back translation. In *Proceedings of the 2009 International Workshop on Intercultural Collaboration, IWIC ’09*, pp. 33–40. Association for Computing Machinery, 2009.
- [39] Yumiko Mori, Toshiyuki Takasaki, and Toru Ishida. Patterns in pictogram communication. *Proceedings of the 2009 International Workshop on Intercultural Collaboration*, pp. 277–280, 2009.
- [40] Yohei Murakami, Toru Ishida, and Takao Nakaguchi. Infrastructure for language service composition. In *2006 Semantics, Knowledge and Grid, Second International Conference on*, pp. 5–5, 2006.

- [41] Toshiki Murata, Mihoko Kitamura, Tsuyoshi Fukui, and Tatsuya Sukehiro. Implementation of collaborative translation environment ‘yakushite net’. In *Proceedings of Machine Translation Summit IX: System Presentations*, 2003.
- [42] Saeko Nomura, Toru Ishida, Mika Yasuoka, Naomi Yamashita, and Kaname Funakoshi. Open source software development with your mother language: Intercultural collaboration experiment 2002. In *HCI (4)*, pp. 1163–1167, 2003.
- [43] Mei-Hua Pan, Naomi Yamashita, and Hao-Chuan Wang. Task rebalancing: Improving multilingual communication with native speakers-generated highlights on automated transcripts. *CSCW ’17*, pp. 310–321. Association for Computing Machinery, 2017.
- [44] Per Persson. Exms: an animated and avatar-based messaging system for expressive peer communication. In *Proceedings of the 2003 ACM International Conference on Supporting Group Work*, GROUP ’03, pp. 31–39. Association for Computing Machinery, 2003.
- [45] Mondheera Pituxcoosuvann, Toru Ishida, Naomi Yamashita, Toshiyuki Takasaki, and Yumiko Mori. Machine translation usage in a children’s workshop. In *Proc. of 10th International Conference on Collaboration Technologies and Social Computing (CollabTech 2018)*, Vol. 11000 of *Lecture Notes in Computer Science*, pp. 59–73. Springer, 2018.
- [46] Mondheera Pituxcoosuvann, Toru Ishida, Naomi Yamashita, Toshiyuki Takasaki, and Yumiko Mori. Supporting a children’s workshop with machine translation. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. Association for Computing Machinery, 2018.
- [47] Mondheera Pituxcoosuvann, Donghui Lin, and Toru Ishida. A method for automated detection of cultural difference based on image similarity. In *Collaboration Technologies and Social Computing*, pp. 129–143. Springer International Publishing, 2019.

- [48] Mondheera Pituxcoosuvann and Yohei Murakami. *Confidence Value Calculation for Cultural-Based Misunderstanding Warning in Multilingual Chat*, pp. 173–187. 2022.
- [49] Mondheera Pituxcoosuvann, Yohei Murakami, Donghui Lin, and Toru Ishida. Effect of cultural misunderstanding warning in mt-mediated communication. In *Collaboration Technologies and Social Computing*, pp. 112–127. Springer International Publishing, 2020.
- [50] Mondheera Pituxcoosuvann, Yohei Murakami, and Kaede Miwa. Detection for cultural difference in impression using masked language model. In *Culture and Computing*, pp. 569–579. Springer Nature Switzerland, 2023.
- [51] Griffiths Roger. Speech rate and nns comprehension: A preliminary study in time - benefit analysis\*. *Language Learning*, Vol. 40, pp. 311 – 336, 2006.
- [52] Pamela Rogerson-Revell. Participation and performance in international business meetings. *English for Specific Purposes*, Vol. 27, No. 3, pp. 338–360, 2008.
- [53] Miki Sasaki and Toshiki Murata. A pattern-based machine translation system-yakushite net mt engine. In *Proceedings of the Second International Workshop on Spoken Language Translation*, 2005.
- [54] Emanuel A. Schegloff and Harvey Sacks. *Opening up closings*, Vol. 8, pp. 289–327. Semiotica, 1973.
- [55] John R. Searle. *Speech acts*. Cambridge University Press, 1969.
- [56] Tomohiro Shigenobu, Kunikazu Fujii, and Takashi Yoshino. The role of annotation in intercultural communication. In *Usability and Internationalization. HCI and Culture*, pp. 186–195. Springer Berlin Heidelberg, 2007.

- [57] Rachit Shukla. Keywords extraction and sentiment analysis using automatic speech recognition, 2020.
- [58] Marc A. Smith, Shelly D. Farnham, and Steven M. Drucker. The social life of small graphical chat spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, pp. 462–469. Association for Computing Machinery, 2000.
- [59] Tatsuya Sukehiro, Mihoko Kitamura, and Toshiki Murata. Collaborative translation environment'yakushite. net'. In *NLPRS*, pp. 769–770, 2001.
- [60] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pp. 3104–3112. MIT Press, 2014.
- [61] Toshiyuki Takasaki. Pictnet: Semantic infrastructure for pictogram communication. In *The 3rd International WordNet Conference (GWC-06)*, pp. 279–284, 2006.
- [62] Toshiyuki Takasaki and Yumiko Mori. Design and development of a pictogram communication system for children around the world. In *Intercultural Collaboration*, pp. 193–206. Springer Berlin Heidelberg, 2007.
- [63] Zhaopeng Tu, Zhengdong Lu, Yang Liu, Xiaohua Liu, and Hang Li. Modeling coverage for neural machine translation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers*. The Association for Computer Linguistics.
- [64] Daniel Vanderveken. *Meaning and Speech Acts: Volume 1, Principles of Language Use*. Cambridge University Press, 1990.

- [65] Naomi Yamashita, Andy Echenique, Toru Ishida, and Ari Hautasaari. Lost in transmittance: how transmission lag enhances and deteriorates multilingual collaboration. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work, CSCW '13*, pp. 923–934. Association for Computing Machinery, 2013.
- [66] Naomi Yamashita, Rieko Inaba, Hideaki Kuzuoka, and Toru Ishida. Difficulties in establishing common ground in multiparty groups using machine translation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 679–688. Association for Computing Machinery, 2009.
- [67] Naomi Yamashita and Toru Ishida. Effects of machine translation on collaborative work. In *Proc. of the 20th anniversary conference on Computer supported cooperative work*, pp. 515–524, 2006.
- [68] Lin Yao, Ying xin Pan, and Dan ning Jiang. Effects of automated transcription delay on non-native speakers’ comprehension in real-time computer-mediated communication. In *Human-Computer Interaction – INTERACT 2011*, pp. 207–214, 2011.
- [69] Mika Yasuoka. 機械翻訳を介したオープンソースソフトウェア開発：intercultural collaboration experiment 2002. 2003.
- [70] Shoichi Yokoyama, Hideki Kashioka, Masaki Matsudaira, Yoshiko Shirokizawa, Shuji Kodama, Terumasa Ehara, Terumasa Ehara, Shinichiro Miyazawa, and Yuzo Murata. An automatic evaluation method for machine translation using two-way mt. 2001.
- [71] Takashi Yoshino, Mai Miyabe, and Tomohiro Suwa. A proposed cultural difference detection method using data from japanese and chinese wikipedia. In *2015 International Conference on Culture and Computing (Culture Computing)*, pp. 159–166, 2015.

- [72] 藤井薫和, 重信智宏, 吉野孝. 機械翻訳を用いた異文化間チャットコミュニケーションにおけるアノテーションの評価. 情報処理学会論文誌, Vol. 48, No. 1, pp. 63–71, 2007.
- [73] 岡本健吾, 吉野孝. 会話中の名詞の関連情報を用いた対面型異文化間コミュニケーション支援システムの構築と評価. 情報処理学会論文誌, Vol. 52, No. 3, pp. 1213–1223, 2011.
- [74] 伊藤孝行, 奥村命, 伊藤孝紀, 秀島栄三. 多人数ワークショップのための意見集約支援システム collagree の試作と評価実験. 日本経営工学会論文誌, Vol. 66, No. 2, pp. 83–108, 2015.
- [75] 伝康晴. 会話・対話・談話研究のための分析単位：隣接ペア (<連載チュートリアル>多人数インタラクションの分析手法〔第4回〕). 人工知能, Vol. 23, No. 2, pp. 271–276, 2008.
- [76] 高梨克也. 基礎から分かる会話コミュニケーションの分析法. ナカニシヤ出版, 2016.
- [77] 笹島宗彦, 井本和範, 下森大志, 山中紀子, 矢島真人, 福永幸弘, 正井康之. 発話意図理解と回答誘導による異言語間会話支援ツール「グローバルコミュニケーター」. インタラクション, pp. 119–126, 2005.
- [78] 笹島宗彦, 井本和範, 下森大志, 山中紀子, 矢島真人, 福永幸弘, 正井康之ほか. 発話意図理解と回答誘導による異言語間会話支援ツールの試作. 情報処理学会論文誌, Vol. 48, No. 3, pp. 1262–1273, 2007.
- [79] 高崎俊之, 肥田昌平, 大谷雅之, 中口孝雄. 児童のための多言語議論支援システムのユーザインタフェース設計. 電子情報通信学会和文論文誌 D, Vol. J99-D, No. 10, pp. 1013–1021, 2016.
- [80] 宗森純, 大野純佳, 吉野孝ほか. 絵文字チャットによるコミュニケーションの提案と評価. 情報処理学会論文誌, Vol. 47, No. 7, pp. 2071–2080, 2006.

- [81] 宗森純, 川津美菜穂, 伊藤淳子. 協力型オンライン rpg コミュニケーションのための絵文字定型文を用いたチャットシステム emo-g コミュニケーター ii. 情報処理学会論文誌, Vol. 62, No. 2, pp. 560–571, 2021.
- [82] 宮部真衣, 吉野孝. リアルタイム遠隔テキストコミュニケーションにおける対人許容応答時間の評価. 情報処理学会論文誌, Vol. 50, No. 3, pp. 1214–1223, 2009.
- [83] 宮部真衣, 吉野孝. 機械翻訳を介したチャットコミュニケーションにおける精度判定に基づく送信拒否の適用可能性. 情報処理学会論文誌, Vol. 51, No. 3, pp. 784–795, 2010.
- [84] 宮部真衣, 吉野孝, 重信智宏. 折返し翻訳を用いた翻訳リペアの効果. 電子情報通信学会論文誌 D, Vol. 90, No. 12, pp. 3141–3150, 2007.
- [85] 石田亨, 村上陽平, 稲葉利江子, 林冬恵, 田仲正弘. 言語グリッド: サービス指向の多言語基盤. 電子情報通信学会論文誌 D, Vol. 95, No. 1, pp. 2–10, 2012.
- [86] 野村早恵子, 石田亨, 船越要, 安岡美佳, 山下直美. アジアにおける異文化コラボレーション実験 2002: 機械翻訳を介したソフトウェア開発. 情報処理, Vol. 44, No. 5, pp. 503–511, 2003.
- [87] 福島拓, 吉野孝, 喜多千草. 共通言語を用いた対面型会議における非母語話者支援システム panelive の構築. 電子情報通信学会論文誌 D, Vol. 92, No. 6, pp. 719–728, 2009.
- [88] 河原達也, 李晃伸. 連続音声認識ソフトウェア julius (<特集> 研究のツールボックス (2)). 人工知能, Vol. 20, No. 1, pp. 41–49, 2005.
- [89] 神田智子, 石田亨ほか. アバタ表情解釈の異文化間比較. 情報処理学会論文誌, Vol. 47, No. 3, pp. 731–738, 2006.
- [90] 山下直美, 葛岡英明, 石田亨ほか. 国際電話会議の負担を軽減する手法の提案と評価. 情報処理学会論文誌, Vol. 54, No. 6, pp. 1794–1806, 2013.

- [91] 山下直美, 坂本知子, 野村早恵子, 石田亨, 林良彦, 小倉健太郎, 井佐原均ほか. 機械翻訳へのユーザの適応と書き換えへの教示効果に関する分析. 情報処理学会論文誌, Vol. 47, No. 4, pp. 1276–1286, 2006.
- [92] 中澤敏明. 機械翻訳の新しいパラダイム：ニューラル機械翻訳の原理. 情報管理, Vol. 60, No. 5, pp. 299–306, 2017.
- [93] 目黒豊美, 東中竜一郎, 堂坂浩二, 南泰浩. 聞き役対話の分析および分析に基づいた対話制御部の構築. 情報処理学会論文誌, Vol. 53, No. 12, pp. 2787–2801, 2012.
- [94] 森由美子. 異文化コラボレーション:2. 子供たちの異文化間コミュニケーション. 情報処理, Vol. 47, No. 3, pp. 276–282, 2006.
- [95] 伊美裕麻, 伊藤孝行, 伊藤孝紀, 秀島栄三ほか. オンラインファシリテーション支援機構に基づく大規模意見集約システム collagree—名古屋市次期総合計画のための市民議論に向けた社会実装. 情報処理学会論文誌, Vol. 56, No. 10, pp. 1996–2010, 2015.
- [96] 船越要, 藤代祥之, 野村早恵子, 石田亨ほか. 機械翻訳を用いた協調作業支援ツールへの要求条件-機-日中韓馬異文化コラボレーション実験からの知見. 情報処理学会論文誌, Vol. 45, No. 1, pp. 112–120, 2004.

# 研究業績

## 主要論文

### 論文誌

1. 元澤 海月, 村上 陽平, Mondheera Pituxcoosuvann, “児童の多言語コミュニケーションにおけるファシリテータの対話行為分析”, ヒューマンインタフェース学会論文誌, vol. 26, no. 1, pp. 77-86, 2024

### 国際会議

1. Mizuki Motozawa, Yohei Murakami, Mondheera Pituxcoosuvann, Toshiyuki Takasaki, and Yumiko Mori, “Conversation Analysis for Facilitation in Children ’ s Intercultural Collaboration,” in *Proc. of the 20th Annual ACM Interaction Design and Children Conference (IDC ’21)*, pp. 62-68, 2021.
2. Mizuki Motozawa, Yohei Murakami, and Mondheera Pituxcoosuvann, “Annotation References for Facilitation Analysis in Intercultural Collaboration,” in *Proc. of the 24th HCI International Conference (HCII 2022)*, pp. 157-172, 2022.
3. Motozawa, Mizuki and Murakami, Yohei and Pituxcoosuvann, Mondheera, “Dialogue Act Analysis of Facilitator-Children Multilingual Communication,” in *Collaboration Technologies and Social Computing: 30th International Conference (CollabTech 2024)*, pp. 97-113, 2024.

4. Motozawa, Mizuki and Murakami, Yohei and Pituxcoosuvann, Mondheera, Amit Pariyar, Hikari Aramaki, “Building a Facilitator Agent to Promote Understanding in Multilingual Communication,” in *Proc. of the 8th International Conference on Information Technology (InCIT 2024)*, pp. 804-809, 2024.

## その他

### 国内会議

1. 元澤海月, 村上陽平, Mondheera Pituxcoosuvann, “児童の異文化コラボレーションにおけるファシリテーションの分析”, ヒューマンコミュニケーション基礎 (*HCS*), 電子情報通信学会技術研究報告, vol. 120, no. 136, pp. 23-28, 2020.
2. 元澤海月, 村上陽平, Mondheera Pituxcoosuvann, “児童の異文化コラボレーションにおけるファシリテータの対話行為分析”, ヒューマンコミュニケーション基礎 (*HCS*), 電子情報通信学会技術研究報告, vol. 121, no. 438, pp. 67-73, 2022.
3. ピタクスワン モンティージャー, 村上 陽平, 元澤 海月, 西村 一球, “ChatGPT とタブーゲーム: 英語スピーキング力向上のためのシステムの提案”, *インタラクション 2024*, 情報処理学会, pp. 1475-1478, 2024.
4. 元澤海月, 村上陽平, Mondheera Pituxcoosuvann, Amit Pariyar, 荒巻光里, “要約提示による多言語議論支援の効果分析”, ヒューマンコミュニケーション基礎 (*HCS*), 電子情報通信学会技術研究報告, vol. 124, no. 360, pp. 28-33, 2025.

### 国際会議

1. Pituxcoosuvann Mondheera, Motozawa Mizuki, Murakami Yohei, Yokote Shiyumi, “Facilitator Agent to Support Low-Resource Lan-

- guage Speakers in MT-Mediated Communication,” in *Collaboration Technologies and Social Computing (CollabTech 2022)*, pp. 231-244, 2024.
2. Mondheera Pituxcoosuvann, Mizuki Motozawa, Yohei Murakami, “Designing Facilitator Assistant Functions to Support Intercultural Children’s Group Work,” in *Proc. of the 24th International Conference on Artificial Intelligence in Education(AIED 2023)*, pp. 21-26, 2023.