

卒業論文

ニューラル機械翻訳を用いた 3 言語間対話に おける参照方法の分析

指導教官 村上 陽平 准教授

立命館大学 情報理工学部
先端社会デザインコース 4 回生
2600170400-0

松原 駿樹

2021 年度（秋学期）卒業研究 3（CH）
令和 3 年 2 月 1 日

ニューラル機械翻訳を用いた 3 言語間対話における参照方法の分析

松原 駿樹

内容梗概

近年、機械翻訳の精度向上に伴い、機械翻訳を用いた母国語による多言語コミュニケーションの需要が増大している。例えば、昨今のコロナウイルスの感染拡大によって、テレビ会議が一般的になり、地理的な壁が取り除かれたことで、母語の異なる人同士のコミュニケーション機会が増え、母国語による機械翻訳を用いたコミュニケーションが必要とされている。

既存の機械翻訳の研究では、ある言語から別の言語への一方向の翻訳が評価されるため、対話の文脈などを考慮してこなかった。そのため、多言語コミュニケーションでは対話の文脈に起因した問題が多く見られた。また 3 言語間での多言語コミュニケーションの研究では、2 言語とは異なり、機械翻訳により他者の間での対話が阻害され、相互理解の基盤の形成が困難になっている。このように従来のフレーズベース統計機械翻訳を用いた多言語コミュニケーションは分析されているが、翻訳精度の向上したニューラル機械翻訳でもコミュニケーションに同様な影響を与えるのかが明らかではない。特にニューラル機械翻訳は従来の機械翻訳とは異なり流ちょうさが格段に向上しており、相互理解の構築方法に異なる影響を与えていることが予想される。

そこで本研究では、日本語、英語、中国語を母国語とする話者を対象に、従来手法のフレーズベース統計機械翻訳 (PBMT) とニューラル機械翻訳

(NMT) とで相互理解の構築方法にどのような違いが表れるかを比較し分析を行った。具体的には、機械翻訳を用いた三者対話による図形マッチング実験を、説明者を交替しながら繰り返し、PBMT と NMT とでそれぞれ話者がどのように参照語を形成し、図形を同定しているのかを分析する。相互理解の構築方法の違いを明らかにするために、本研究で以下の二つの分析に取り組んだ。

同定プロセスの効率性

PBMT と NMT の翻訳精度の違いが、相互理解の円滑さに与える影響を調べるために、図形の同定に要するコストを比較する必要がある。これにより、NMT が多言語コミュニケーションをどれだけ効率化したかを理解でき、それぞれの同定プロセスにおける利点と欠点を明らかにできる。

参照語の形成プロセス

PBMT と NMT の翻訳精度の違いが、相互理解の構築に重要な参照語の形成に

どのような影響を与えるのか調べるために、実験回数に応じた参照語の変化を調査し、比較する必要がある。これにより、NMTが多言語コミュニケーションの相互理解にどれだけ影響を与えているかを理解できる。

これらの分析を行うために、本研究では、機械翻訳を用いたチャットツール Langrid Chat を用いて日英中の 3 言語間対話による図形のマッチング実験を行った。実験では、各言語話者に順不同の図形群を配布し、一人の説明者が母国語で図形の説明を行って、他の二人がその図形を同定する。全ての図形の同定が終わると説明者を交替し、全被験者が 2 回説明を担当するまで繰り返す。実験を通して、3 言語間でどのようにして、それぞれの図形の同定が行われ、参照語が形成されていったのかを PBMT および NMT で比較する。

1 つ目の同定プロセスの効率性という課題に関しては、一つの図形をマッチングするために要した発話数や発言の長さで評価する。次に 2 つ目の参照語の形成プロセスという課題に関しては、図形のマッチングを繰り返す中で生じる参照語の変化を「同じ」、「短い (simplify)」、「短縮 (narrow)」、「長い」、「違う表現」の 5 種類に分類し、PBMT と NMT で各種類の出現頻度で比較した。実験の結果、各課題に対して得られた知見は以下の通りである。

同定プロセスの効率性

PBMT を用いたグループより、NMT を用いたグループの方が、最短の呈示受理プロセスによって図形を同定するケースが多かった。また、両方の機械翻訳でそれぞれ発話カテゴリを分割して調べたところ、参照語による表現が NMT の方が多かったことから、NMT の方が、効率性が高いといえる。

参照語の形成プロセス

NMT を用いたグループは PBMT を用いたグループと比べて、参照語として違う表現を用いるケースが少なく、同じ表現を使い続ける傾向にあることが分かった。また、表現を変える場合は「短縮 (narrow)」が多く、「短縮 (simplify)」が行われるのが PBMT を用いたグループよりも遅いという実験結果を得た。これは、NMT が PBMT よりも流ちょうさが優れているため、全体的な意味を残す短縮 (narrow) が多く用いられたと考えられる。

Analysis of Reference Method in Trilingual Dialogue Using Neural Machine Translation

Takaki MATSUBARA

Abstract

In recent years, with the improvement of the accuracy of machine translation, the demand for multilingual communication that communicates in the native language using machine translation has increased. For example, by coronavirus, video conferencing has become commonplace and geographical barriers have been removed. This will increase opportunities for communication between people with different mother tongues, and communication using machine translation in the mother tongue is required.

In research using existing machine translation, there are many studies between two languages, we have not considered the context of the dialogue as one-way translations from one language to another are valued. Therefore, many problems were found due to the context of conversation. Also, in studies using machine translation between three languages, unlike bilingual, machine translation may hinder dialogue between others. Therefore, it is difficult to form a foundation for mutual understanding. In this way, communication has been analyzed using conventional phrase-based statistical machine translation, but it is not clear whether neural machine translation with improved translation accuracy has the same effect on communication. In particular, neural machine translation is different from conventional machine translation in that its fluency is significantly improved. It is thought that this has a different influence on the method of constructing mutual understanding.

In this study, we compared and analyzed the differences in the construction of mutual understanding between the conventional method of phrase-based statistical machine translation (PBMT) and neural machine translation (NMT) for native speakers of Japanese, English, and Chinese. Specifically, in an experiment of figure matching by a three-way dialogue using machine translation, the explanation was reported while changing the explainers. Then, we analyze how the speakers of PBMT and NMT form reference words and identify figures. In order to clarify the difference in the method of constructing mutual understanding, we worked on the following two analyzes in this study.

Efficiency of the identification process

In order to investigate the effect of the difference in translation accuracy between PBMT and NMT on the smoothness of mutual understanding, it is necessary to compare the cost of figure identification. This makes it possible to understand how much the development of NMT has improved the efficiency of communication using machine translation, and to clarify the advantages and disadvantages of each identification process.

The process of forming referents

In order to investigate how the difference in translation accuracy between PBMT and NMT affects the formation of referents, which is important for the construction of mutual understanding, it is necessary to investigate and compare the change in referents according to the number of repetitions of the experiment. This makes it possible to understand how the development of NMT affects mutual understanding of communication using machine translation.

In order to conduct these analyses, we conducted a figure matching experiment using a chat tool (Langrid Chat) with machine translation in a trilingual dialogue between Japanese, English and Chinese. In the experiment, using a set of unordered figures distributed to speakers of each language, one explainer described the figures in their native language, and the other two identified the figures. After all the figures have been identified, the explainer is changed, and the process is repeated until all the subjects have been explained twice. Through the experiment, we will compare how the identification of each figure and the formation of the referent between the three languages were carried out by PBMT and NMT.

Specifically, with regard to the first issue of the efficiency of the identification process, we measure the number of utterances and the length of utterances required to match a single figure between the two machine translations in each experiment and investigate how the efficiency of the reference word identification process changes between PBMT and NMT by producing an average value. As for the second issue, the formation process of reference words, we divided the process of changing reference words into the patterns of "same", "shortened(simplify)", "shortened(narrow)", "long", and "different expression", and investigated the frequency of occurrence of each pattern in PBMT and NMT.

The results of the experiments and the findings for each task are as follows.

Efficiency of the identification process

The group using NMT identified more figures by the shortest presentation acceptance process than the group using PBMT. In addition, both machine translations were examined by dividing the utterance categories respectively, and it was found that NMT was also more efficient in expressing them with reference words.

The process of forming referents

The NMT group used fewer different expressions as referents than the PBMT group and tended to keep using the same expressions. In addition, the experimental results show that when changing expressions, abstraction is more frequent and shortening takes place later than in the group using PBMT. This is because NMT is more fluent than PBMT, and there are more expressions that have contractions that preserve the overall meaning.

卒業論文タイトル

目次

第1章 はじめに	1
第2章 機械翻訳を用いたコミュニケーション	3
2. 1 先行研究	3
2. 1. 1 参照	3
2. 1. 2 同定	4
2. 2 同定プロセスでの障害	4
2. 3 参照語の形成	4
2. 4 本研究の意義	5
第3章 比較実験	6
3. 1 参加者	6
3. 2 Langrid Chat	6
3. 3 図形マッチング実験	7
3. 4 実験手順	8
第4章 分析手法	10
4. 1 効率性の定量化	10
4. 2 参照語の形成カテゴリ	10
第5章 分析結果	12
5. 1 同定プロセスの効率化	12
5. 1. 1 翻訳の非対称性	12
5. 1. 2 同定プロセスの効率性	14
5. 2 参照語の形成プロセス	17
第6章 考察	19
第7章 おわりに	21
謝辞	23
参考文献	24

第1章 はじめに

近年、コロナウイルスの影響から対面でのコミュニケーションが困難になり、テレビ会議の需要が増加している。特に他国の人と話す場合には、機械翻訳を用いたオンラインでのコミュニケーションが必要になっている。このときに母語の異なる人同士のコミュニケーション機会が増え、多言語間同士での母国語を使用する機械翻訳を用いたコミュニケーションを必要としている。既存の機械翻訳を用いた研究では、2言語間の研究が多く、ある言語から別の言語への一方向の翻訳が評価されるため、対話の文脈などを考慮してこなかった。そのため、対話の文脈に起因した問題が多く見られた。また3言語間での機械翻訳を用いた研究では、2言語とは異なり、機械翻訳により他者の間での対話が阻害される可能性があるため、相互理解の基盤の形成が困難になっていることが報告されている。このように従来のフレーズベース統計機械翻訳を用いたコミュニケーションの分析はされているが、翻訳精度が向上したニューラル機械翻訳でもコミュニケーションに同様な影響を与えるのかが明らかではない。特にニューラル機械翻訳は従来の機械翻訳と異なり流ちょうさが格段に向上しており、相互理解の構築方法に異なる影響を与えていることが予想される。

そこで本研究では、日本語、英語、中国語を母国語とする話者を対象に、従来手法のフレーズベース統計機械翻訳 (PBMT) とニューラル機械翻訳

(NMT) とで相互理解の構築方法にどのような違いが表れるかを比較し分析を行った。具体的には、機械翻訳を用いた三者対話による図形マッチング実験では、説明者を交替しながら繰り返し行い、PBMT と NMT とでそれぞれ話者がどのように参照語を形成し、図形を同定しているのかを分析する。相互理解の構築方法の違いを明らかにするために、本研究で以下の二つの分析に取り組んだ。

同定プロセスの効率性

PBMT と NMT の翻訳精度の違いが、相互理解の円滑さに与える影響を調べるために、図形の同定に要するコストを比較する必要がある。これを比較することは、NMT の発達が機械翻訳を用いたコミュニケーションにどれだけの効率性を向上させたのかが理解でき、それぞれの同定プロセスにおける利点と欠点を明らかにできる。

参照語の形成プロセス

PBMT と **NMT** の翻訳精度の違いが、相互理解の構築に重要な参照語の形成にどのような影響を与えるのか調べるために、実験の繰り返し回数に応じた参照語の変化を調査し、比較する必要がある。これを比較することは、**NMT** の発達が機械翻訳を用いたコミュニケーションの相互理解にどれだけの影響を与えているかを理解できる。

以下、本研究では 2 章において機械翻訳を用いたコミュニケーションにおける参照に関する関連研究を紹介し、これに基づいて本研究を行う意義を述べ、第 3 章では、本研究において行った実験の概要について詳細を記述する。第 4 章では、参照語の形成プロセスが **PBMT** 及び **NMT** でどう行われているかを調査し参照内容の同定プロセスがどう効率化されているかを記述する、第 5 章では、本論文でのアプローチを述べ、提案手法を記述し、第 6 章で本稿をまとめる。

第2章 機械翻訳を用いたコミュニケーション

本章では、まず多言語間でのコミュニケーション分析に必要なチャットログを収集するために行った実験の概要と、後に記述する、参照と同定がどう行われているのかを示す。さらに、先行研究とは異なる翻訳機を用いて行った、本研究の意義を明記する。

2.1 先行研究

先行研究では、母国語の異なる 2 人のユーザが翻訳付きのチャットシステムを介して参照内容を同定する実験を行った。課題の内容は、「図形のマッチング」とした。この課題では、2 人のユーザに異なる順番に並べられた同一の図形セットを渡し、これをマッチングする。ここで、2 人に配布された図形はタングラムと呼ばれ、抽象的な図形であるため、2 人のユーザが初めてマッチングする際には、互いに図形の見方に対する合意形成が行われる。また、再びマッチングする際には、参照語を用いて、同定されることが多い。以下では、先行研究の結論をもとに、本研究の課題を記述する。

2.1.1 参照

	Game 1 PBMT	Game 2 PBMT	Game 3 PBMT 以降
	中： <u>2つ目は少し不格好なロボット</u> 。右向きに歩く。 <u>頭と体が正方形</u> 英： <u>ロボット</u> やMinecraftのキャラクターのような？	英：9番目は長方形の頭と本体。 <u>ロボットのように見える。</u>	日： <u>ロボット</u> のようです

図 1：参照表現の利用例

本研究において、参照はユーザが相互理解を構築するために重要な役割を果たしている。ユーザは「図形のマッチング」において、あるタングラムを参照内容としたときにそれを同定するプロセスを共有することで、見方や考え方を共有し、その後のゲームでもその共通の認識として使うことが多い。

2.1.2 同定

本研究における同定とは、「図形のマッチング」において、参照語を用いて他ユーザと相互理解を図ることである。本実験では一度、参照語が使われて同定できた図形は、以後、同じ表現や短縮した参照語が度々見られた。

以下では、2.2節から2.3節に先行研究での同様の実験を通して得た知見を紹介し、これをもとに、2.5節で本研究の試みの意義を記述する。

2.2 同定プロセスでの障害

先行研究の中では、共通言語を用いて同様の実験を行う場合と、機械翻訳を用いて行う場合を比較している。その中で同定プロセスの効率性について論じられている障害は、以下の点である

- ・ 翻訳の非対称性

翻訳の非対称性とは、言語 A から言語 B に翻訳した言葉を再び言語 A に翻訳しなおすと、もとの意味とは違った意味で返ってくるという現象である。例えば、英語の話者が「arm」という単語を使った際に、日本語の話者には「腕」と正しく伝わっていたが、中国語の話者には「武器」という意味で伝わっていた。しかし、中国語の「武器」を英語や日本語に翻訳した場合には、「arm」や「腕」と翻訳される。このため、中国語話者のみが図形を同定出来ず、別の参照表現が必要になる。このような、非対称性のために機械翻訳を用いたコミュニケーションでは、同一の参照表現への収束が難しくなる。

2.3 参照語の形成

本研究の実験において、一度形成された参照語は次のゲームで、同様または似た表現及び、より短くなった表現が使われる場面が多く見られた。参照語が使われるパターンとして、「一度使われた参照語と同じ」、「参照内容の全体的な意味を残して詳細を省略する方法 (simplification)」、「参照内容の特徴的な情報を残

す方法(narrowing)」が存在する。例えば、「手を 180°広げて走っている子供」が、「手を広げている子供」に短縮された場合は、全体的な意味を残す短縮(simplify)に該当し、「子供」に短縮された場合は、特徴的な情報を残す短縮(narrow)に該当する。先行研究によると、共通言語を用いた場合では、全体的な意味を残す短縮方法が大半を占めていた。しかし、機械翻訳を介した対話では短縮された参照表現が正しく翻訳されるとは限らず、短縮された参照表現や参照語を用いて図形を同定出来ない可能性がある。これから、誤解を避けるために、一度使われた表現を用いてゲームを繰り返す。すなわち、特徴的な情報をそのまま残す、短縮(narrow)がとられることが多い。

2.4 本研究の意義

これまでの先行研究では、共通言語を用いた場合と、機械翻訳を用いて母国語を使用した図形のマッチング実験が行われてきた。現在ではニューラル機械翻訳(NMT)は従来手法であったフレーズベース統計機械翻訳(PBMT)とは異なり、流ちょうさが格段に向上している。先行研究で論じられていた結果は PBMT を用いた結果である。しかし、翻訳精度が向上した NMT でもコミュニケーションに同様な影響を与えるかは明らかではない。NMT を使用することによって、PBMT で生じていた問題が NMT に変更することで解決され、逆に NMT を使用したことによって新たな障害が生じた部分もあるのではないかと考える。例えば、NMT では訳抜けしやすいという特性もあり、それが対話に影響している場面があるとも仮定できる。すなわち、本研究において、先行研究と同様の実験を NMT と PBMT 両方で調査することによって、相互理解の構築に異なる影響を与えることが明確化できると考える。

第3章 比較実験

3.1 参加者

本実験の参加者は日本語、英語、中国語の話者がそれぞれ一人ずつの計 3 名であった。参加者は翻訳機の利用経験については経験ある者とない者が混在していたが、本実験に差支えはないものだった。

3.2 Langrid Chat

本実験で用いた翻訳付きチャットシステムでは、ユーザが入力する言語と出力する言語を選択でき、それぞれが担当する言語に統一した。さらに、ユーザが入力している場合には他のユーザにそれが表示されて認識できるものとなっていた。この Langrid Chat を用いて、フレーズベース統計機械翻訳 (PBMT) とニューラル機械翻訳 (NMT) の両方の実験を行った。図 2 は Langrid Chat の実際の画面である。対話は画面左側で行われ、自身の発話はそのうちの右側、他 2 人の発話は左側に翻訳されて自身の言語で表示される。また、画面右側には参加しているユーザが表示される。

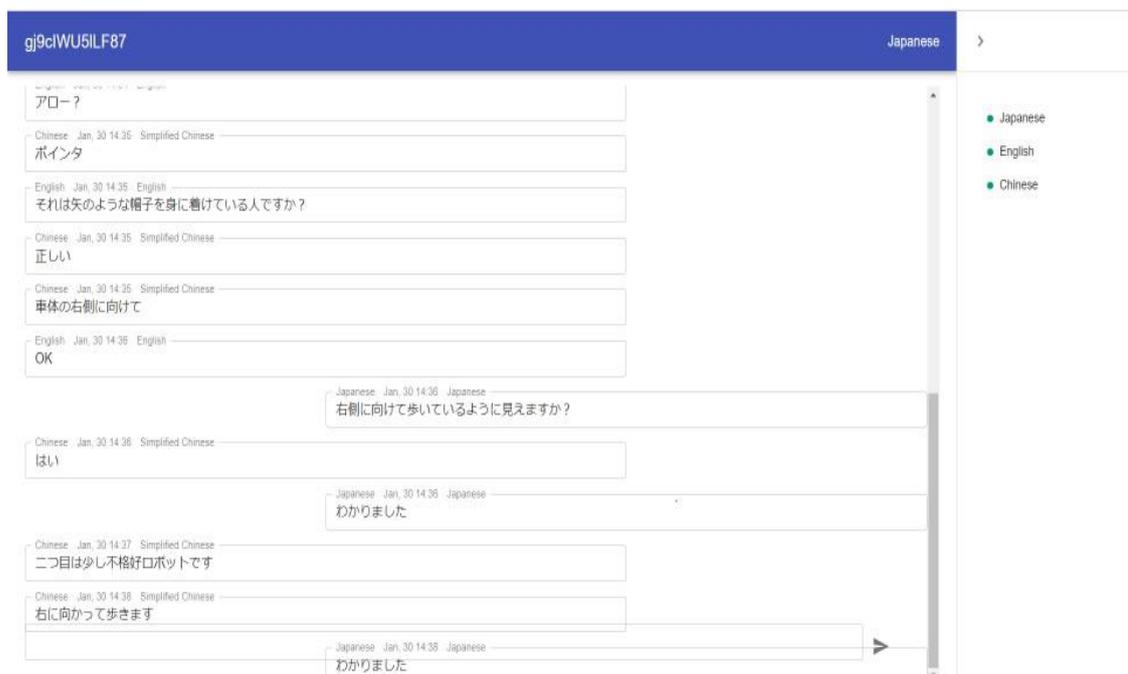


図 2 : Langrid Chat

3.3 図形マッチング実験

本実験では、前節で記した **Langrid Chat** を用いて図形のマッチングを行った。図 3 は本実験で用いた図形セットの一例である。参加者はそれぞれ異なった順番に並べられた図形のセットを持ち、1 名が正解の順番を指示する指示者となる。そして、他の 2 名が指示者の説明から自身の図形セットを並び替える照合者となる。指示者は、照合者に対してどのような図形の順番で並べられているかを 1 枚ずつ説明し、照合者はその説明を聞き、手元にある異なる並び順の図形を指示者の説明通りに並び替える。指示者が 1 枚ずつ説明し、照合者が合意を行い、10 枚すべてのタングラムを並び換え終わるまでを 1 ゲームとし、それが終わると、指示者を入れ替えて再び並び換える。この実験は照合者をそれぞれの話者が 2 セット行う、計 6 ゲーム行う。さらに本実験では、これを **PBMT** 及び **NMT** の両方の機械翻訳で行う。**NMT** では先行実験で使われた **PBMT** と同様に、初めて図形のマッチングを行うときには図形の見方に関する合意形成が行われた。さらに、合意形成を繰り返すうちに、やがて同じ参照表現に収束し、それをもとに同定されることが多くなる。

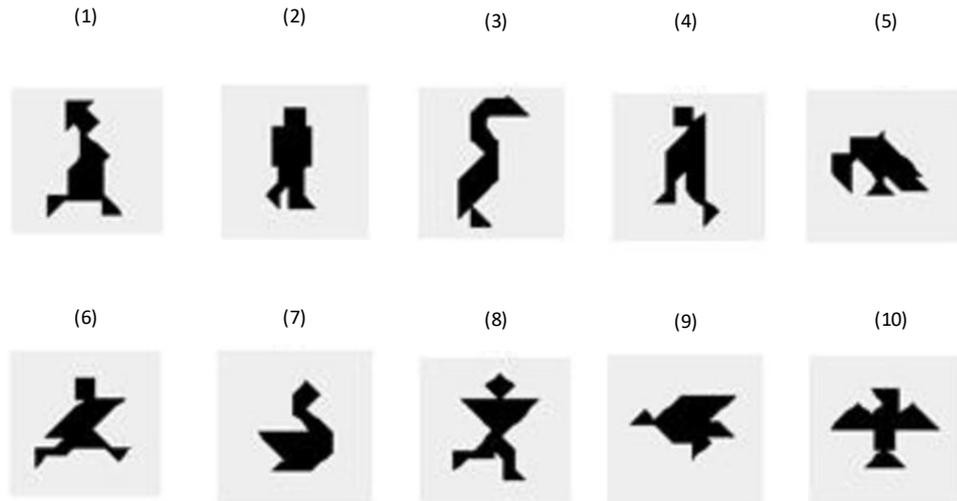


図 3：図形マッチングに用いたタングラム図形のセット

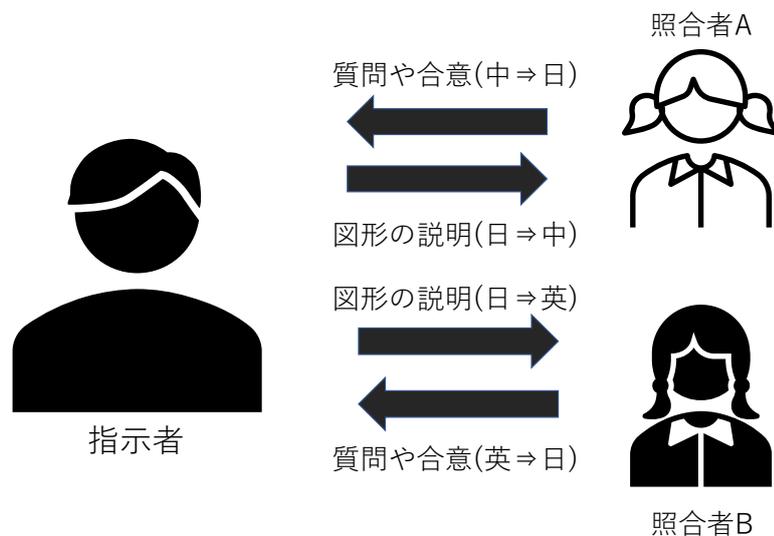


図 4：実験時のインタラクション

3.4 実験手順

- (1) 実験の説明：参加者には事前に実験手順の確認を行い、ルールを確認する。被験者には、基本的に支持者が図形の説明を行い、照合者はその指示をもとに図形の同定を行い、質問や確認といった必要な発話を行うように指示した。また、図形の同定が完了したら、合意の発話を行うよう指示した。指示者は照合者が合意出来たことを確認したら次のタングラムの説明に移るように指示した。
- (2) 図形セット：図形マッチングの実験を行う前に、実験者にタングラムのセットを配布する。これらは 3 者とも同じ図形だが順番がバラバラに配られている。
- (3) 図形のマッチング：ある一人が正解の順番に並べられた図形の順番であり、その人が指示者になる。他の 2 名は指示者が図形についての特徴を言うていくのをチャットで見て、タングラムを並び替えていく、照合者となる。照合者は指示者に対して、図形に関する質問や確認を取り、最終的には合意をする。
- (4) 指示者の交代：10 枚のタングラムを並び替え終わったら、指示者を変えてゲームを繰り返す。10 枚のタングラムを並び替えるまでを 1 ゲームとして、

これを6ゲーム繰り返す.

第4章 分析手法

4.1 効率性の定量化

参照と同定のプロセスの効率性を調べるために、本実験で得た対話ログを分析する。具体的には、図形のマッチングに要した時間や発言の長さの平均を用いて定量的評価を行う。また、マッチングに要した時間や発言の差が、どのような発言内容によって生じているのかを明らかにするために、発話を「情景描写」、「参照表現」、「質問・確認」、「合意」、「理解不能」、「その他」の6種類に分類する。これによって、実験を繰り返す毎に、発話の長さや発話の内容がどのように変化していくのかをPBMTとNMTとで比較する。

- 情景描写：図形の内容を描写する発話
- 参照表現：図形の内容を参照語で表現する発話
- 質問・確認：図形について質問や確認をする発話
- 合意：発話内容を正しく理解できたことを伝える発話
- 理解不能：発話内容が理解できないことを伝える発話
- その他：上記以外の発話

4.2 参照語の形成カテゴリ

本実験に用いたPBMTおよびNMTでは、一度形成された参照語は次のゲームで、同様または似た表現及び、より短くなった表現が使われる場面が多く見られた。特に表現が短くなる、「参照内容の全体的な意味を残して詳細を省略する短縮(simplify)」、「参照内容の特徴的な情報を残す短縮(narrow)」についてPBMTとNMTでどのような違いが行われるかを考える。すると、PBMTとNMTでは参照語が短くなる過程に違いが見られた。そこでこの違いを明らかにするために、図形をマッチングする際に用いられた参照語が一度同定されてから、次のゲームでどう変化するかをPBMTおよびNMTで調査した。本実験では、この参照語の変化をPBMTとNMTにおいて参照語の形成がゲームを繰り返すごとにどう変化しているか調査するために、以下の5点に分類して考える。

- ① 同じ（前回のゲームと同様の表現）

- ② 短縮(**simplify**) (参照内容の全体的な意味を残して詳細を省略する方法)
- ③ 短縮(**narrow**) (参照内容の特徴的な情報を残す方法)
- ④ 長い (前回のゲームより参照語が長くなる)
- ⑤ 違う表現 (新たな参照語が作られる)

②と③は同じく参照語を前ゲームより縮小するものではあるが性質が違うものである。②の方では、例えば「7番はカラスが下を向いて餌をついばんでいる」という表現が前ゲームでなされていた場合、次のゲームで「7番は鳥が餌を食べている」という風に、図形の形やどういった行動をしているかといった状態についての、短縮(**simplify**)が行われる。③では同じく「7番はカラスが下を向いて餌をついばんでいる」とあったときに、次ゲームでは「7番はカラス」といった風に、参照語の特徴的な部分である、「カラス」という一部を残した、短縮(**narrow**)が起こる。

第5章 分析結果

5.1 同定プロセスの効率化

PBMT と NMT の翻訳精度の違いが、相互理解の円滑さに与える影響を調べるために、図形の同定に要するコストを比較する必要がある。これを比較することは、NMT の発達が機械翻訳を用いたコミュニケーションにどれだけの効率性を向上させたのかが理解でき、それぞれの同定プロセスにおける利点と欠点を明らかにできる。

5.1.1 翻訳の非対称性

翻訳の非対称性とは、言語 A から言語 B に翻訳した言葉を再び言語 A に翻訳しなおすと、もとの意味とは違った意味で返ってくるという現象である。例えば、先に述べた通り、PBMT で行った実験では、英語の話者が「arm」という単語を使った際に、日本語の話者には「腕」と正しく伝わっていたが、中国語の話者には「武器」という意味で伝わっていた。しかし、中国語の「武器」を英語や日本語に翻訳した場合には、「arm」や「腕」と翻訳される。このため、中国語話者のみが図形を同定出来ず、別の参照表現が必要になる。このような、非対称性のために機械翻訳を用いたコミュニケーションでは、同一の参照表現への収束が難しくなる。

日本語話者	英語話者	中国語話者
3番は腕を広げて右に実行されている男です。	<u>No.3 is the man running to the right with arms spreading out.</u>	三是运行在正确的武器铺开的人。
彼は翼を持っているよう。	<u>Like he has wings.</u>	就像他有翅膀。
彼の腕は180度です。	<u>his arms are 180 degree.</u>	他的武器是180度。
腕？	arms?	<u>武器？</u>
彼の手はい。	<u>Yes his hands.</u>	是的他的手。
彼の手は、それは180度だということです	His hand is that it's 180 degrees	<u>他的手是180度的那个吗</u>
はい。	<u>Yes.</u>	是。

表 1：PBMT における非対称性（下線部が話者）

本研究では、こういった翻訳の非対称性が NMT に置いて散見できるかを調べたところ、NMT において本実験では非対称性ではないが、似た現象が見受けられた。つまり本実験にて得られた、NMT で起こった現象は PBNT の非対称性と少し違いが見られた。例えば NMT の実験において、中国語の話者が図形のマッチングを行う上で図形の特徴を説明しているときに、日本語の話者がその説明に対して疑問を投げかけている場面があった。その際に日本語の話者が投げた質問が、「走っているような人？」という疑問文だった。しかし、それが図形と異なっていたため中国語の話者は否定を返す、「没有」と返答した。つまり、「没有」とは日本語に翻訳される際に、「いいえ」という意味で返ってくるのが翻訳としては正しいものとなる。しかし、日本語に翻訳された際に、日本語の話者には「番号」という意味として誤訳されて伝わってしまった。そのため、日本語の話者はこの質問で求める返答を受け取ることが出来なかったため、別の質問をして求める答えを得るといった場面があった。しかし、日本語の「番号」を中国語に戻すと「没有」とはならない。このことから、この場面において中国語と日本語の二つだけでは翻訳の非対称性は見受けられなかった。つまり、翻訳機が誤訳してしまったとも考えられる。しかし、英語の翻訳を見てみると、中国語の話者が発した「没有」は英語では「No」と正しく伝わっていた。このことから、NMT を用いた場合にも、

日本語話者	英語話者	中国語話者
2番目の写真も右に歩いている人です	The second picture is also a person walking to the right	<u>第二张图片也是向右走的人</u>
その頭は正方形です	Its head is square	<u>它的头是正方形</u>
本体も長方形で構成されています	The body is also composed of rectangles	<u>身体也是由长方形组成</u>
<u>走っているような人？</u>	Someone who seems to be running?	有人在跑步吗？
番号	No	<u>没有</u>
歩いている	Is walking	<u>是在步行</u>
とった！一方の脚がもう一方の脚よりもはるかに大きく見える	<u>Got it! one leg looks way bigger than the other one.</u>	得到它了！一只脚看起来比另一只大
<u>歩幅が大きい人</u>	People who do not have a large stride	步伐不大的人
はい、それは彼です	Yes, it's him	<u>没错，就是他</u>
わかった	all right	行

表 2 : NMT での翻訳問題

ある言語 A から別の言語 B, C へ翻訳した際に, A から B へ翻訳した際に正しく伝わっても, A から C に翻訳した際には異なる翻訳がされる場面は見られた. しかし, この PBMT と NMT で起こった現象は本実験において, 文の違いによって見られる現象だと考えられる. それゆえに, これは機械翻訳特有の問題であると考えられるため, 機械翻訳に対しての支援が必要になると考える.

5.1.2 同定プロセスの効率性

5.1.1 で論じた PBMT と NMT それぞれにおこる問題が, 同定プロセスの効率性に対して, どのような違いが発生するかについて定量的に評価する
図形マッチングに要した時間

下の図 5 では, PBMT および NMT でそれぞれゲームを繰り返す中で, 図形 1 つを同定するのにかかった時間の推移を表している. この表から, NMT では 1 回目のゲームでは PBMT よりも図形をマッチングする時間が長かったが, 2 回目以降のゲームでは, NMT の方が時間は短くなっていった. つまり, NMT では, 最初に参照語の同定を行う場合には, PBMT と同様に参照表現を多く呈示していたために時間がかかっていたが, 2 回目以降には NMT の方が参照内容の特徴的な情報を残していたため時間も短くなった. しかし, 図 5 から見てわかるように, PBMT ではゲーム 3, NMT ではゲーム 4 に時間が増加する現象が見られた. これらはいずれも, それまでの参照語から別の表現で行われたため, 参加者同士が互いに相



図 5 : 1つの図形をマッチングするのに要した平均時間(秒)

互理解がなかなか図れずに時間が延びていた。

発言のカテゴリ、参照語の多さ（情報量）の定量的評価

PBMT と NMT で発言の違いから生じる、同定プロセスの違いを調べる上で、まず呈示・受理という最短プロセスで同定できた図形に注目し、その割合を算出した。

図 6 のグラフから 1 回目では、PBMT および NMT で最短の呈示・受理プロセスで図形を同定した割合は低かった。これはまだ参照語を形成できていない段階であるため、最短でのプロセスになるには至らなかった場合が多かったからであろう。しかし、2 回目から差が大きく見られた。PBMT の方では、1 回目から 2 回目へ進んだ際に、最短の呈示・受理プロセスに大きな差が見られなかったのに対して、NMT では 2 回目からおよそ 6 割が最短の呈示・受理プロセスで図形を同定させていた。そして、6 回目までいくと PBMT ではおよそ 4 割であったのに対して、NMT では 9 割の図形が最短の呈示・受理プロセスで同定されていた。これは一つの図形に対して用いられる参照語が、PBMT の方が NMT より多いため、最短の呈示・受理プロセスでは同定されづらいものとなる。同様に、発言の分類によって効率性を求めることが出来る。

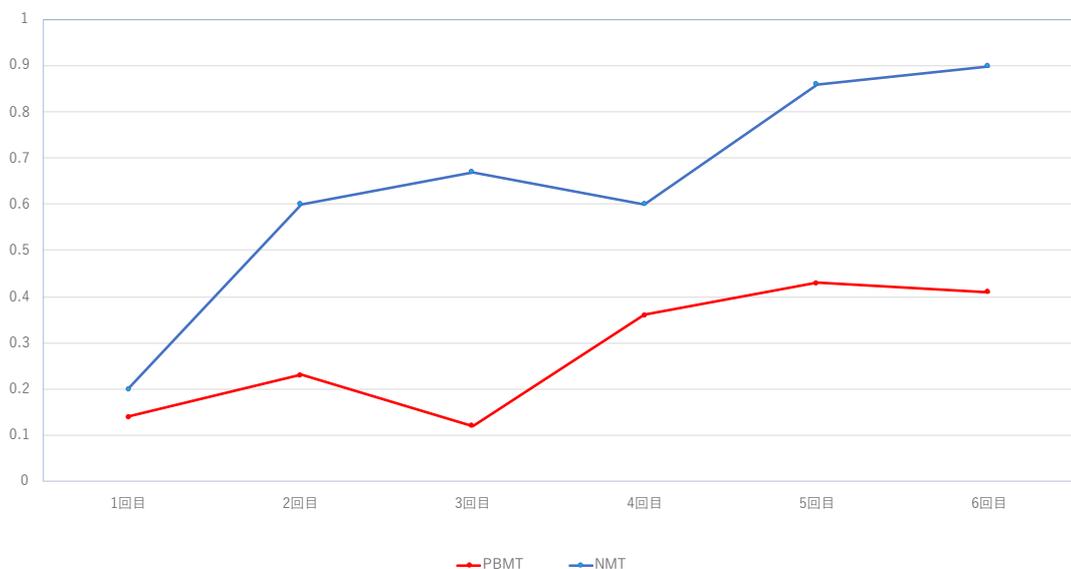


図 6 : 参照語の呈示・受理プロセスで図形を同定した割合

PBMT	1回目	2回目	3回目	4回目	5回目	6回目
情景描写	0.26	0.30	0.25	0.22	0.20	0.17
参照語	0.06	0.10	0.21	0.20	0.18	0.23
質問・確認	0.33	0.23	0.08	0.05	0.05	0.03
合意	0.35	0.34	0.45	0.48	0.51	0.56
理解不能	0.00	0.00	0.01	0.00	0.02	0.00
その他	0.00	0.03	0.00	0.05	0.04	0.01

表 3 : PBMT における発話のカテゴリ分類

表 3, 表 4 より PBMT では情景描写が多く使われていて, NMT では参照語を用いて図形のマッチングが行われていたことがわかる. このことから, 同定の効率性は NMT の方が高いと言える.

NMT	1回目	2回目	3回目	4回目	5回目	6回目
情景描写	0.28	0.16	0.11	0.08	0.07	0.07
参照語	0.04	0.13	0.21	0.25	0.30	0.28
質問・確認	0.29	0.18	0.13	0.08	0.09	0.07
合意	0.35	0.47	0.53	0.54	0.49	0.56
理解不能	0.00	0.02	0.01	0.00	0.02	0.00
その他	0.04	0.04	0.01	0.05	0.03	0.02

表 4 : NMT における発話のカテゴリ分類

5.2 参照語の形成プロセス

前章の 3 節で述べたように、参照語の変化について 5 つに分類した。さらにこれらがゲーム 1 で参照語が同定されてから、ゲーム 2 以降で 5 つの分類方法のどれに偏っているのかを調査した。以上がその結果である。表 5 では PBMT, 表 6 では NMT についての参照語の変化で分類している。表 1 および表 2 では、ゲーム 2 から 6 まででそれぞれ参照語の変化が「同じ」、「短縮(simplify)」、「短縮(narrow)」、「長い」、「違う表現」に分けたとき、各図形のマッチングの同定にどれだけの回数が用いられたかの回数を左に記述し、右には各ゲームでそれらが出現した割合を記している。

表 5 および表 6 を比較したときに得られる分析結果は以下のとおりである。

- ① PBMT では参照語の変化が「短縮(simplify)」に多く見られるのに対して、NMT では「短縮(narrow)」が多く見受けられる。
- ② 「長い」や「違う表現」を用いることが PBMT では多く見られたのに対して、NMT ではあまり見られなかった。

PBMT	同じ	短縮 (simplify)	短縮 (narrow)	長い	違う表現
Game1	/	/	/	/	/
Game2	4 0.210	3 0.157	2 0.105	2 0.105	8 0.421
Game3	7 0.350	6 0.300	1 0.050	2 0.100	6 0.300
Game4	5 0.357	5 0.357	0 0.000	1 0.071	3 0.214
Game5	10 0.714	2 0.142	1 0.071	0 0.000	1 0.071
Game6	11 0.916	1 0.083	0 0.000	0 0.000	0 0.000

表 5：PBMT における参照語の変化の分類

NMT	同じ	短縮 (simplify)	短縮 (narrow)	長い	違う表現
Game1	/	/	/	/	/
Game2	4 0.266	2 0.133	3 0.200	1 0.066	5 0.333
Game3	5 0.357	1 0.071	4 0.285	1 0.071	3 0.214
Game4	6 0.600	1 0.100	2 0.200	0 0.000	1 0.100
Game5	8 0.727	2 0.181	0 0.000	0 0.000	1 0.090
Game6	10 0.833	2 0.166	0 0.000	0 0.000	0 0.000

表 6 : NMT における参照語の変化の分類

第6章 考察

本実験において、3言語間における対話コミュニケーションではPBMTを用いた場合とNMTを用いた場合とで、似た性質もあれば、大きな違いが見受けられたケースがあった。NMTを用いた場合には、PBMTと比較した際に具体的に以下のような違いが生じるという結論を得た。

1. 翻訳の非対称性から、PBMTと同様に、3者間で共通の相互理解が得られない場面がある。
2. 抽象図形を参照語で同定した際には、それ以降にはより抽象図形の特徴的な情報を残そうとする。

現象1に関しては、例えば日本語の話者が他の2人のユーザに図形の特徴を伝えた際に、日本語の話者は自身が伝えたことが伝わっていると感じているものの、実際には英語の話者には伝わっているが、中国語の話者には伝わっていないということがあるため、対話が破綻する場面がPBMTと同様に、NMTでも見受けられた。これは、翻訳機を通した際に、日本語の話者が正しく伝わっているかを把握することが出来ないため起こる問題だと考えられる。よって、何らかの方法で日本語の話者が発信したことを、日本語の話者が正しく伝わっていることを把握する手段を用意すべきであると考えられる。つまり、この例では、先に述べてきた翻訳の非対称性から、日本語→中国語→日本語の順だけで自身が把握しようとしても不十分である。よって、「日本語→中国語→英語→日本語」と「日本語→英語→中国語→日本語」といったように自身が発信したことを2回以上日本語以外の言語に変えてから日本語に返せば、翻訳の非対称性を考慮せずとも対話がスムーズになされるのではないかと考える。つまり、日本語の話者にとっては中国語←→英語のやり取りを視覚化することで、各言語間のやり取りを共有でき、翻訳結果に齟齬が生まれたときに別の表現で言い換えることができると考えた。

現象2に関しては、PBMTは単語ベースの翻訳を行うのに対して、NMTではより文脈に沿った翻訳がなされる。このことから、NMTの方がPBMTより流ちょうな対話が行われていた。しかし、対話において、例えば母国語同士で会話する際には短縮(simplify)が多く生じる。例として「眼鏡をかけて勉強をしている男性」のような文章があった際に、それをもう一度表現するときには「勉強している男性」といった風に、以前使われた参照語を用いて伝えようとする短縮

(simplify)がよく生じる。このことから本実験を考えると、結果として比較的 PBMTの方が NMTより抽象図形を参照語で形成する際には、短縮(simplify)を利用してゲームを進めており、より共通言語に近い現象が起こった。しかし、前章の結果から NMTの方が効率良く対話が進められた。つまり機械翻訳を用いた対話において、共通言語を用いた場合とは異なった合意形成の方法が成されたと考えられる。これは、本実験で3言語を用いたため、機械翻訳を用いた場合には3人のユーザが共通の参照語を共有しにくい。したがって NMTではそれぞれの話者が前回使われた参照語から特徴的な情報を残そうとする短縮(narrow)を用いてより短い表現で伝えようという、ユーザの意図的な工夫がより多く行われた。例えば、「左に走っている男」とあったときには、試行を繰り返す毎に、「左向き」として、特徴を捉えた短縮(narrow)で合意形成を行っていた。それに対して、PBMTでは翻訳精度の低さから3言語での合意形成が難しく、一度用いられた参照語と同じ表現や全体的な意味を残す短縮(simplify)が使われ、その上で NMTより多くの情報量を出すことで翻訳精度の低さを補ったと考えられる。つまり、NMTでは参照語の共有の難しさから短縮(narrow)が、PBMTでは翻訳精度の低さから短縮(simplify)がより多く用いられたと考えられる。そのため NMTでは PBMTと比較して短縮(narrow)が多く用いられたが、効率良く対話が進んだと予想できる。

今回の実験から図形のマッチングを行った際には NMTの方が PBMTより効率よく対話が進んでいったといえる。しかし NMTでは比較的短い発話しかなされていなかったために、NMTに起こりうる訳抜けなどの問題が特に発生せず影響しなかったが、発話が長くなればなるほど NMTでもコミュニケーションの破綻も考えられる。よって、NMTの翻訳精度をさらに高めるには、こうした特有の問題を軽減していく必要があると考える。

第7章 おわりに

機械翻訳の精度向上に伴い、機械翻訳を用いて母国語によりコミュニケーションを行う多言語コミュニケーションの需要が増大している。そこで本実験では、PBMTとNMTの比較から、翻訳機を用いたコミュニケーションにおける円滑な相互理解を実現する上でのNMTの有用性や解決しきれていないことについて調査した。その上で、翻訳機を用いた対話では相互理解を行う際にはそれぞれの翻訳機に関する支援が必要であることがわかった。したがって、本研究では、3言語間での円滑なコミュニケーションを行う際にはどのような支援が適切か、あるいはPBMTからNMTに変化させたところ、翻訳精度の向上が3言語間コミュニケーションに対してどれだけの影響を与えるか、あるいはPBMTと同様の結論が得られるのか、調べる実験を参照語に注目して実施した。

同定プロセスの効率性

参照語を用いて、図形のマッチングを行う際には、NMTとPBMTには翻訳の非対称性とそれに近い問題が見受けられた。よって、3言語間での対話を行う際には、この問題を防ぐアプローチが必要であると考えられ、それによって同定プロセスの効率化が可能である。

参照語の形成プロセス

NMTでは参照語を形成する上で、それまで用いられてきた参照語の特徴的な表現を抽出して、その情報を残す方法がPBMTに比べて多く見受けられた。これから、NMTでは対話の流ちょうさが向上したために、より対話らしさが増したと言える。このことから参照語の形成プロセスから、PBMTに比べてNMTが対話において有用であると考えられる。

総合的に見て、NMTを用いた場合ではPBMTと比べて、参照語の同定プロセスは効率がよく、参照語の形成プロセスはより特徴的な情報を取る方法を取っていたことから、より自然な対話になっている。しかし、機械翻訳というフィルタを挟んでいる以上、本実験では見受けられなかった問題も存在していることも考えられる。また本実験で発見できた問題に対しても、翻訳精度の向上ではなく、別のアプローチを考えた。したがって、機械翻訳を用いたコミュニケーションにおいて、円滑な相互理解を図るには、それにあった支援が必要になってくる。そこで、本研究では1対1の独立した翻訳を行うのではなく、3言語間で機械翻訳を用いたコミュニケーションを行う際には、1対1対1の他2言語の翻訳

を考慮した翻訳支援が必要だと考えた.

謝辞

本研究を行うにあたり，熱心なご指導，ご助言を賜りました村上陽平准教授に深謝申し上げます．また，ご協力いただいた被験者の **Mondheera Pituxcoosuvarn** 助教，西村一球さん，叶峰さん，普段からお世話になっている社会知能研究室の皆様にも心より感謝申し上げます．

参考文献

- [1] 山下直美 , 石田亨 , 平田圭二 : 機械翻訳を用いた対話における思い違いに関する分析, 情報処理学会論文誌, Vol. 41, No6, June (2000)
- [2] 山下直美 , 石田亨 : 機械翻訳を用いた対話における参照方法に関する分析, 情報処理学会論文誌, Vol. 48, No2, Feb (2007)
- [3] 稲葉利江子, 山下直美 , 石田亨 , 葛岡英明 : 機械翻訳を用いた3言語間コミュニケーションの相互理解の分析, 電子情報通信学会論文誌, 2009/6, Vol. J92-D , No6
- [4] 藤井真 , 新納浩幸 , 古宮嘉那子 : 文を持つ情報量を用いたニューラル機械翻訳の訳抜け検出, 言語処理学会 第25回年次大会 発表論文集 (2019年3月)
- [5] 高田凌平 , 秋葉友良 , 塚田元 , 豊橋技術科学大学 : ニューラル機械翻訳における文書トピック情報の利用, 言語処理学会 第25回年次大会 発表論文集 (2019年3月)