

# 翻訳機を介して相互理解を構築するときの難点について

山下 直美<sup>\*1</sup>  
Naomi Yamashita

石田 亨<sup>\*2</sup>  
Toru Ishida

<sup>\*1</sup> NTTコミュニケーション科学基礎研究所  
NTT Communication Science Labs.

<sup>\*2</sup> 京都大学 情報学研究科  
Department of Social Informatics, Kyoto University

Reaching mutual understanding over lean media using low quality machine translation is a difficult task. To support such communication, it is important to understand how machine translation affects communication. In this study, we focused on referring behavior, which is closely related to mutual understanding. We compared referring behavior of pairs from different language regions in their shared language, English, and in their native languages using a machine translation embedded chat system. Analysis results show that in machine translation-mediated communication, participants have trouble efficiently identifying referents through basic exchanges. Participants identify a figure using the same referential expressions in the prior trial, when iteratively identifying the same referent.

## 1. はじめに

国際的な協調作業が年々増加する中、相互理解を実現する機械翻訳システム(翻訳機)が切望されている。翻訳機に関する従来研究の大半は翻訳精度の向上に焦点を当ててきた。最近では、翻訳機を用いた協調作業に関する研究も活発に行われている[石田 2006]。しかし、未だ翻訳機がユーザ間の相互理解の構築にどのような影響を及ぼすかは明らかになっていない。

今後、円滑な相互理解を実現できる翻訳システムを開発する上で、翻訳機を介した対話における相互理解の構築方法を理解することは意義深いと考える。

そこで本研究では、相互理解の構築に重要な役割を果たす参照に的を絞って、異言語ユーザがどのようなプロセスを経て参照内容を同定し相互理解を構築するかを調べる。具体的には、異言語ユーザが翻訳付きチャットシステムを介して抽象図形のマッチングを行う実験を実施し、彼らが母国語ではない共通言語を用いてマッチングする場合と翻訳機を用いて母国語でマッチングする場合における抽象図形の参照、同定プロセスを分析した。

## 2. 参照と同定

参照は、メンバが相互理解を構築する上で重要な役割を果たしている[Clark 1986]。メンバは、参照内容を同定するプロセスを共有することによって、参照内容に対して共通の見解を得ることができる。

### 2.1 仮説

- 同定プロセスの効率性に関する仮説:

参照内容を同定する最短プロセスは、1. 一方のメンバによる参照語の"呈示"と 2. 他方のメンバによる参照語の"受理"である[Clark 1986]。特に同じ物事の同定を繰り返す場合、2 度目以降は、この呈示受理プロセスによって同定することが多い[Clark 1986]。しかし、異言語ユーザが翻訳機を介して図形のマッチングを行うとき、彼らは同一の参照表現やキーワードを共有することができず、また翻訳に一貫性がないため、翻訳機を介した対話では呈示受理プロセスで参照内容を同定することは困

難であろう。

- 同定プロセスにおける発話内容に関する仮説:

通常の対話において、相手の発話内容に関して十分に理解できないと感じたとき、聞き手は以下のいずれかの行動をとる。

(1) 質問をして説明を求める。(2) 発話者が説明を追加してくれるまで待つ。(3) 発話者の言いたかったことを自分の言葉で確認する。過去の研究において、インストラクタが電話を介して初心者リアタイムな指示を出すとき、初心者は自分の理解が正しいか否かを何度も質問したり確認したりすることが報告されている。翻訳機を介したコミュニケーションでも、対話に誤訳が混入するため、メッセージの受け手は自分の理解に不安を感じるだろう。このため、翻訳機を介した対話では、互いの発言内容を質問したり確認したりする発言が頻出するだろう。

- 参照表現の短縮に関する仮説:

通常の対話では、ある事柄について深く議論されるとき、その事柄を指す参照表現が短縮され最終的には短い参照語(noun phrase)になる。参照表現の短縮方法には、1. 参照内容の全体的な意味を残して詳細を省略する方法と 2. 参照内容の特徴的(部分的)な情報だけを残す方法が存在する[Clark 1986]。通常の対話では、全体的な意味を残す短縮方法が大半を占める。しかし、翻訳機を介した対話では短縮後の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できない可能性がある。翻訳付き掲示板を用いたコミュニケーションでは、ユーザは誤訳による誤解を避けるために同じ言い回しを繰り返し用いていた。本実験でも、ユーザは 1 度目にマッチングができた表現となるべく同じ表現を用いて 2 度目のマッチングをしようと考えられる。参照表現を短縮する場合には、1 度目にマッチングした表現と部分的に同じフレーズを使う方法、すなわち特徴的な情報をそのまま残す方法がとられるだろう。

## 3. 比較実験

### 3.1 参加者

本実験の参加者は 16 名(日本人 8 名、中国人 5 名、韓国人 3 名)であった。実験では日本人と中国人、日本人と韓国人がペアになって図形のマッチングを行ったが、どのペアも実験前に面識はなかった。すべての参加者は翻訳機の利用経験が殆ど

なく、また日常的にメールやチャットを使っていた。参加者は全員、日常生活に必要な最低限な英語を使いこなすことができた。

### 3.2 図形のマッチング

本実験では、参加者のペアが翻訳付きチャットシステムを介して抽象図形のマッチングを行う。参加者のペアは、共通言語(英語)と翻訳を介して母国語で図形のマッチングを2度ずつ行う。ただし、1度目と2度目の図形セット(10個の抽象図形)は同じ図形が異なる順序で並んでいる。また、翻訳機を用いてマッチングする図形セットと、共通言語を用いてマッチングする図形セットは異なる。

### 3.3 実験手順

- (1) 実験の説明と参加者ペアの構成: 参加者にメールで実験手順を送付する。参加者は指示された相手とペアになり、チャットシステム上で簡単な挨拶を交わす。
- (2) 図形セットの獲得: 参加者は個別に指示された URL にアクセスし、図形のセットを入手する。
- (3) 共通言語を用いた図形のマッチング(1回目): 参加者ペアは、共通言語(英語)を用いて図形のマッチングを行う。
- (4) 共通言語を用いた図形のマッチング(2回目): 図形の配置は(3)のものとは異なる。
- (5) 母国語を用いた図形のマッチング(1回目): 図形のセットは共通言語でマッチングしたものとは異なる。
- (6) 母国語を用いた図形のマッチング(2回目): 図形の配置は(5)のものとは異なる。

## 4. 分析結果

### 4.1 同定プロセスの効率性

仮説1を検証するために、共通言語を用いた場合と翻訳機を用いた場合に呈示受理プロセスで図形を同定できた割合を調べた。

その結果、1度目のマッチングでは共通言語でマッチングした場合も翻訳機を介して母国語でマッチングした場合も呈示受理プロセスで同定できた割合に差が見られなかった(共に2割)。しかし2度目のマッチングでは、共通言語を用いてマッチングする方が呈示受理プロセスで同定できる割合が多く、その割合に大きな差が生じた(共通言語を用いた場合は7割、翻訳機を用いた場合は4割)。

この結果は、翻訳機を用いた場合に参照語による同定が困難なことを示している。

### 4.2 同定プロセスにおける発話内容

仮説2を検証するために、各発話を以下の6カテゴリーに分類した:「情景描写:図形の内容を描写する発話」、「参照語で表現:図形の内容を参照語で表現する発話」、「質問や確認:質問や確認をする発話」、「合意:正しく理解したことを伝える発話」、「理解不能:理解できないことを伝える発話」、「その他:上記のどのカテゴリーにも属さない発話」。

表1: 発話のカテゴリー分類結果

カテゴリー	English 1 <sup>st</sup> trial	English 1 <sup>st</sup> trial	MT 1 <sup>st</sup> trial	MT 2 <sup>nd</sup> trial
情景描写	0.35	0.15	0.44	0.36
参照語で表現	0.06	0.31	0.03	0.14
質問や確認	0.15	0.09	0.13	0.10

合意	0.39	0.44	0.36	0.40
理解不能	0.02	0	0.03	0.01
その他	0.03	0	0.03	0

表1は発話の分類結果を示している。予想に反し、翻訳機を介した対話において「質問や確認」に関する発話が多い傾向は見られなかった。一方、「情景描写」と「参照語で表現」に属する発話の出現頻度は、共通言語を用いた場合と翻訳機を用いた場合で有意な差が検出された;共通言語を用いた場合も翻訳機を介した場合も2度目のマッチング時に「参照語で表現」に属する発話が占める割合は増えるが、共通言語を用いる方が「参照語で表現」に属する発話の全体に占める割合が高い。また、共通言語を用いた場合は「情景描写」に属する発話が2度目のマッチング時に大幅に減少するのに対し( $\chi^2=8.230$ ,  $p=.006$ ), 翻訳機を介した場合は2度目のマッチング時にも「情景描写」に属する発話あまり減少しない( $\chi^2=1.284$ ,  $p=.257$ )。

この結果は、翻訳機を介した対話では参照表現を短縮し参照語を形成しづらい状況を示唆している。

### 4.3 参照表現の短縮

仮説3を検証するために、2度目のマッチング時にユーザがどのように図形を参照していたかを調べた。

その結果、翻訳機を用いた2度目のマッチングでは、1度目に合意形成に至ったものと同じ表現を用いたり特徴的な情報だけを残した表現を用いる割合が多いことがわかった(共通言語を用いた場合:39%, 翻訳機を用いた場合:80%)。逆に、全体的な意味を保持した参照表現の短縮は少なかった(共通言語を用いた場合:50%, 翻訳機を用いた場合:14%)。

この結果は、4.2節の結果と一貫して、翻訳機を介した対話において参照表現を短縮することが困難な様子を表している。

## 5. まとめ

本研究では、相互理解に重要な役割を果たす参照に注目し、異言語ユーザが翻訳機を介して参照するときに生じる対話スタイルを調べる実験を実施した。

実験中の対話を分析した結果、次のことがわかった。1.一方が呈示した参照語を他方が受理するという最短のプロセスで同定できるケースが少なかった。2.抽象図形の様々な特徴(情報)を次々に呈示することで同定しようとしていた。3.相互理解が構築された表現を再び伝える場合は、短縮せずにそのまま伝えるか、もしくは特徴的な情報のみ伝えていた。

本実験により、翻訳機を介して参照内容を同定することが困難な様子があった。本実験で見られた対話スタイルが翻訳機を介したあらゆる対話に有効とは考えにくいので、翻訳機を介して円滑に参照や同定ができるような支援が必要である。

### 謝辞

翻訳付きチャットシステムを提供くださった和歌山大学の吉野孝先生に深く感謝します。

### 参考文献

- [Clark 1986] Clark, H. H. and Wilkes-Gibbs, D.: Referring as a collaborative process, *Cognition*, Vol.22, pp.1-39, 1986.
- [石田 2006] 石田亨, 内元清貴, 山下直美, 吉野孝: 機械翻訳を用いた異文化コラボレーション, *情報処理*, Vol.47, No.3, pp.269-275, 2006.