

# 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

藤井 薫<sup>†</sup> 重信 智宏<sup>††</sup> 吉野 孝<sup>††,†††</sup>

異文化間コミュニケーションにおいて、言語や文化の違いは大きな障壁である。多くの人にとって、新たな言語を習得することは大きな負担であり、お互いの母国語で円滑なコミュニケーションが成立することが望ましい。本研究では、機械翻訳機能と語句への意味づけ（アノテーション）機能を持たせたチャットシステム AnnoChat を開発し、日本人、中国人、韓国人の間のコミュニケーションに適用し、異文化間コミュニケーションにおけるアノテーションの評価を行った。実験の結果、被験者によって付与されたアノテーションの約 7 割は、チャットのコンテキストに依存しない知識情報であり、蓄積し共有することで異文化間コミュニケーションにおける有用な知識情報になると考えられる。また、約 2 割はチャット中に説明できなかった情報を補足するために利用されており、リアルタイムコミュニケーションにおいてチャット中のアノテーション付与が有効な例であると考えられる。

## Evaluation of Annotation in Intercultural Communication Chat Tool Using Machine Translation

KUNIKAZU FUJII,<sup>†</sup> TOMOHIRO SHIGENOBU<sup>††</sup>  
and TAKASHI YOSHINO<sup>††,†††</sup>

In intercultural communication, there are large barriers because the languages and cultures are different. We think that it is preferable for people to communicate smoothly with their own mother tongue. Therefore, we have developed a chat system named AnnoChat. AnnoChat has an annotation function for smooth intercultural communications. AnnoChat was applied to experiments carried out among the Japanese, Chinese, and Koreans. From the results of the experiments, about 70% of the added annotations are reusable for intercultural knowledge information. About 20% of the added annotations are used to supplement information that could not be described while chatting. It is thought that this is an effective example of applying the annotation in real-time communications.

### 1. はじめに

インターネットの世界的な普及により、コンピュータを使った異文化間の交流がさかんになりつつある。しかし、異文化間コミュニケーションを行うためには様々なスキルが必要になることから、いまだ多くの人々にとって「敷居が高い」存在であるといえる。

異文化間コミュニケーションにおいて、異なる文化を持つ相手と円滑なコミュニケーションを実現するためには、主に 2 つの障壁があると考えられる。第 1

の障壁は、言語の壁である。従来から行われてきた会話ベースのコミュニケーションでは、英語や、コミュニティ内で合意されたローカルな共通言語を使ってコミュニケーションに参加することが必要であった。そのため、共通言語と母国語が異なる初心者ユーザにとっては、外国語学習という困難な道のりから始めなければならない。第 2 の障壁は、文化の壁である。円滑なコミュニケーションをとるためには、お互いの文化についての理解があることが望ましい。このためには、相手の文化について事前に学習をしたり、相手から直接教えてもらったりするなど、積極的に相手の文化を知ろうという姿勢が必要である。

本研究では、異文化間コミュニケーションに関する特別なスキルがないユーザでも、言語や文化の違いを超えた円滑なコミュニケーションに参加できることを目標とした異文化間コミュニケーション支援システ

<sup>†</sup> 和歌山大学大学院システム工学研究科  
Graduate School of Systems Engineering, Wakayama University

<sup>††</sup> 情報通信研究機構  
National Institute of Information and Communications Technology

<sup>†††</sup> 和歌山大学システム工学部  
Faculty of Systems Engineering, Wakayama University

△ AnnoChat を開発した。AnnoChat では、チャットに参加する人同士が共有する時間帯にメッセージやアノテーションを送受信している形態をリアルタイムコミュニケーションと位置づける。本論文では、異文化間コミュニケーションにおける語句への意味情報の付与が、どのような役割を果たすかについて、実際の異文化間コミュニケーションに適用した結果を報告する。

## 2. 従来研究

2.1 機械翻訳を用いたコミュニケーションツール  
情報技術を用いて異文化間コミュニケーションを支援する試みとして、これまでに機械翻訳を介したメールやチャット、電子掲示板などが開発されている。

機械翻訳を用いたチャットシステムの例として、AmiChat がある<sup>1)</sup>。このシステムでは、入力されたチャットメッセージを翻訳し、複数の言語で表示することが可能となっている。また、AmiChat の特徴的な機能として、翻訳されたメッセージが理解できない場合には、“Huh?” と書かれたボタンを押すことにより、メッセージ投稿者に対して「翻訳したメッセージの意味が分からないので、いい換えてほしい」というメッセージを表示する、翻訳結果のフィードバック機能がある。

機械翻訳では、つねに正しい結果が得られるとは限らないため、翻訳結果のフィードバック機能があることが望ましい。電子掲示板システム TransBBS では、メッセージを入力し、確認用のボタンをクリックすることで、投稿前に翻訳結果を確認することが可能である<sup>2)</sup>。これにより、メッセージ投稿前に翻訳結果が意図したおりの意味を示しているかどうかを確認することが可能であり、意図どおりでない場合、ユーザ自身が投稿メッセージを書き換えることで、翻訳ミスの少ない投稿メッセージを作成できる。このようなユーザの操作を「翻訳リペア」と呼ぶ<sup>3)</sup>。しかし、翻訳結果をそのまま表示するだけでは、翻訳先の言語を理解できないユーザは翻訳リペアを行えない。よって、目的の言語に翻訳した文章を再度ユーザの言語に再翻訳する「折り返し翻訳」機能を提供することで、機械翻訳を介したメッセージが相手言語でどのように伝わっているかを、おおむね確認することができる<sup>4)~6)</sup>。

### 2.2 アノテーションによる情報共有

アノテーションは、様々なメタデータを表現するための方法として、近年研究が活発に行われている。中でも、ヒューマンコミュニケーションの支援を目的とした適用事例として、議論支援システムにおける論拠の共有がある<sup>7)</sup>。

議論支援システムでは、議論の論拠に該当する補足情報をアノテーションとして外形化することで、参加者同士で、お互いの知識や意見などの差異を補完することを目的としている。

異文化コラボレーションを対象とした知識情報のフレームワークとして、BBS やチャットなどコミュニケーションに使うツールの情報をプロジェクト単位で横断的に取得し、情報共有を行うフレームワークが提案されている<sup>8)</sup>。これは、先に行われた異文化間におけるソフトウェア共同開発実験 (ICE2002) の結果をもとに提案されており、BBS やチャットなどコミュニケーションに使うツールの情報を横断的に取得し、意味情報を蓄積・共有することで、継続的なコラボレーションの支援になると提案されている。

このように、アノテーションをヒューマンコミュニケーションに適用することで、意味を補足できる可能性が高いことから、異文化間で誤解を生じる可能性がある語句に適切なアノテーションを付与することで、異文化の壁を低くすることができると考えられる。しかし、現時点では異文化間コミュニケーションにおけるアノテーションの適用例は報告されていないため、本実験では、どのようなアノテーションデータが付与されるか、また、異文化間のリアルタイムコミュニケーションにおいて、どのような影響を及ぼすかについての調査を行う。

## 3. 異文化間コミュニケーションツール AnnoChat

### 3.1 設計方針

本研究では、異文化間コミュニケーションの初心者でも、言語や文化の違いを超えた、円滑なコミュニケーションを可能とするため、異文化間コミュニケーション支援システム AnnoChat を開発した。本システムの設計方針として次の2点をあげる。

- (1) 機械翻訳を用いた言語バリアの克服支援
- (2) 語句へのアノテーション付与機能による文化的情報の共有支援

理由

- (1) 言語バリアを軽減するためには、参加者が各々の母国語でコミュニケーションに参加できることが望ましい。そこで、機械翻訳を用いることで、相手の入力言語にかかわらず、つねにユーザの母国語でログを表示することが可能になる。しかし、チャットにおける入力メッセージは、しばしば誤訳につながる口語調や主語などの省略が生じる。入力者が意図した内容を相手に伝え

るためには、折り返し翻訳などの正しく翻訳できる文章を書くための支援が必要である。本システムでは、折り返し翻訳にかかる時間を短くするため、ユーザによる明示的な操作を不要とした「リアルタイム折り返し翻訳機能」により、短時間で機械翻訳に適した送信メッセージを作成できるように支援する。

- (2) 異文化間コミュニケーションでは、参加者間でそれぞれ異なった文化的背景を持っているため、正しく翻訳されたメッセージであっても、意味が分からなかったり、送信者と受信者が違った解釈をするおそれがあると考えられる。本システムでは、チャット中の任意の語句に対するアノテーション付与機能により、意味情報を共有し、異文化間に起因する理解のずれを減らすことを支援する。

本システムでは、チャットログの任意の語句に対して、複数のアノテーションを付与できるようにする。1つの語句には、複数の意味が含まれていることが多く、必要に応じて複数のアノテーションを付与できることが望ましいからである。

### 3.2 システム構成

図1に異文化間コミュニケーションツール AnnoChat のシステム構成を示す。システムは、サーバ・クライアントモデルで構成しており、サーバは、AnnoChat サーバおよび、翻訳サーバからなる。

図1の実線は、クライアント・サーバ間のメッセージデータの流れを示している。AnnoChat クライアントから投稿されたメッセージやアノテーションは、AnnoChat サーバが受信し、翻訳サーバを用いて他の言語への翻訳を行い、チャットセッションの参加者へ配信する。また、図1の破線は、折り返し翻訳のメッセージデータの流れを示している。リアルタイム折り返し翻訳を実現するためには、翻訳結果が早く表示されることが望ましい。そこで、折り返し翻訳のクエリは AnnoChat サーバを介さず、翻訳サーバと直接通信をする形態をとった。

翻訳サーバは、J-Server（高電社）を用いた。J-Server は、日本語と中国語、日本語と韓国語、日本語と英語間において、双方向の翻訳が可能な翻訳機である。また、直接翻訳ができない言語対（たとえば、中国語と韓国語間）の翻訳は、日本語を介したマルチホップ翻訳により、日本語、中国語、韓国語、英語の間を相互に翻訳することが可能である。

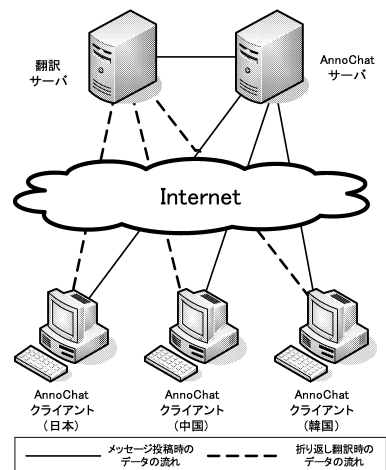


図1 AnnoChat のシステム構成

Fig. 1 System configuration of AnnoChat.

### 3.3 AnnoChat クライアント

AnnoChat クライアントは、多言語入力・表示、リアルタイム折り返し翻訳機能、アノテーション付与機能を持ったチャットプログラムである。操作方法および画面構成はインスタントメッセンジャ方式とした。図2に AnnoChat クライアントの実行画面を示す。以下に AnnoChat クライアントの機能について説明する。

- (1) 多言語入力・表示対応

表示言語切替えボタンを押すと、本システムが対応している言語一覧が表示される。1つの言語を選択すると、チャットログ、アノテーションリスト、アノテーションの内容が指定した言語で表示されるようになり、メッセージ入力ボックスが空であれば、メッセージ入力言語の指定も変更される。図2(a)と(b)は同じチャットセッションのスクリーンショットであり、(a)は日本語表示、(b)は中国語表示である。

- (2) リアルタイム折り返し翻訳機能

メッセージ入力ボックスに入力された文字は、数秒おきに折り返し翻訳を行い、結果を折り返し翻訳結果ボックスに表示する。

- (3) アノテーション付与機能

チャットログボックスにおいて、下線つき太字で表示されている部分（語句）はアノテーションが付与されていることを示す「アノテーションリンク」である。この語句の上にマウスポインタをあわせると、現在の表示言語設定に応じた言語でアノテーション内容がアノテーションバルーンとして表示される。

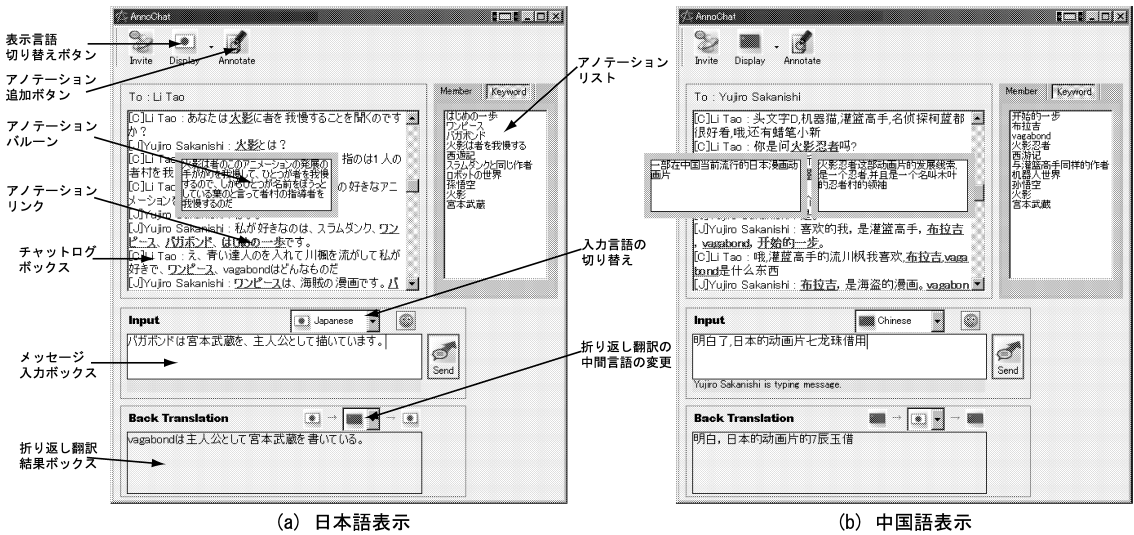


図 2 AnnoChat クライアントの画面例  
Fig. 2 An example screen of AnnoChat client.

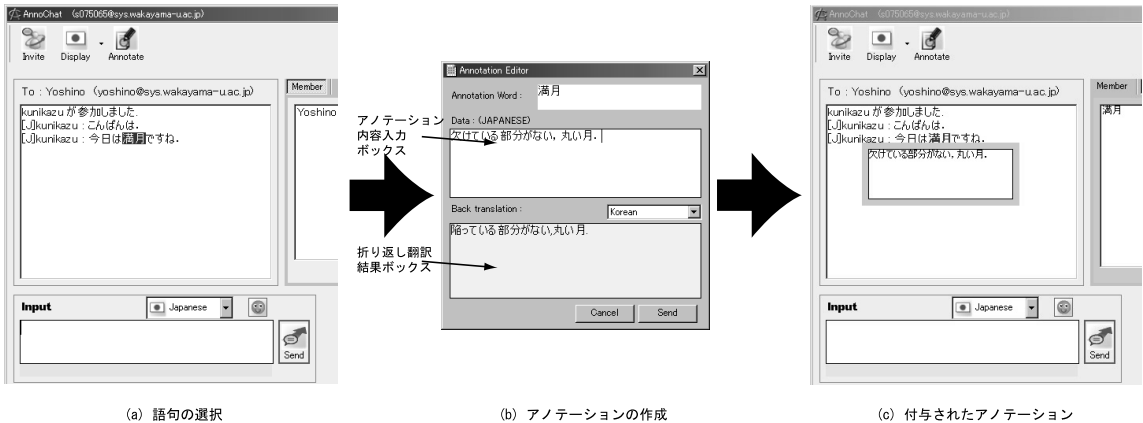


図 3 アノテーションの付与手順  
Fig. 3 Making procedure of an annotation.

新しくアノテーションを付与する方法を図 3 に示す。

- (a) チャットログボックスで付与したい語句を選択し、アノテーション追加ボタンをクリックする。
- (b) アノテーション編集ウィンドウに語句の説明を母国語で入力する。
- (c) Send (送信) ボタンを押すと、アノテーション内容が翻訳され、チャットセッションに属するすべてのユーザにアノテーションが配信される。

また、アノテーションバルーンをクリックすることで、既存アノテーションを編集することができる。

AnnoChat クライアントは Microsoft .Net Framework 上で動作する Windows アプリケーションであり、C#言語を用いてプログラミングを行った。

#### 4. 適用実験と実験結果

##### 4.1 適用実験

3章で開発したシステムを異文化間コミュニケーションに適用し、アノテーション付与において得られたデータについて調査を行った。

実験フィールドは、アジア 5 カ国 (中国、韓国、タイ、マレーシア、日本) の研究機関が共同で実施した異文化コラボレーション実験 (ICE2005)<sup>9)</sup> において、中国語、韓国語、日本語を母国語とする各国の大学生および大学院生の方々が被験者とした。実験は、被験

者同士に面識のない組合せで1対1のチャットを、韓国人と日本人で3組、中国人と日本人で6組、中国人と韓国人で10組の計19組実施した。組合せごとに実施数が異なるのは、被験者数や実験時間の割振りによって制約が生じたためである。

本実験では、アノテーションの付与を確実に進めようとするため、タスクを次のように3段階に分けた。

- (1) 文化の違いが表れるようなチャットログを得るため、20分間のチャットを行う。
- (2) チャットログに対してアノテーションを5つ付与する。
- (3) アノテーションが付与されたチャットログを参照し、アンケートを行う。

本システムでは、チャット中にアノテーションを付与することが可能である。しかし、チャットの合間を見計らってアノテーションを付与する操作は、本システムの操作に十分慣れていない被験者にとって負担が大きいことから、操作に手間どることが考えられる。そこで、本システムの習熟の程度、キー入力速度、チャット経験の多少にかかわらず、確実にアノテーションを付与してもらう実験方法として、チャット開始から一定時間(20分間)経過後に会話を中断してもらい、両者が同時にアノテーションを付与するための間を作り出した。20分間と定義した理由は、アノテーションを付与することができる程度のチャットログが蓄積されたところという想定である。

上記の実験手順では、アノテーションの付与を含めたコミュニケーションに制約を加えた形となっており、厳密なリアルタイムコミュニケーションであるとはいえない。しかし、一般的なチャットにおけるリアルタイム性とは、チャット参加者が同じ時間帯に集まり、ある程度の間隔をもってメッセージをやりとりする形態を指す場合が多い。とくに、一般的なチャットでは、送信ボタンを押すまではメッセージが送信されないため、face to faceの会話に比べると、発話が完了してから相手の返答が到着するまでの時間が長くなる。そのため、会話分析などの観点からは非リアルタイムで一方的であるとの議論もある<sup>10)</sup>。しかし、チャットを経験したことのあるユーザは、ある発話から応答までにはある程度時間がかかることを承知しており、遅延時間が許容範囲内であれば、リアルタイムなコミュニケーション手段として見なされている<sup>11)</sup>。本論文では、チャット参加者が同じ時間帯において、チャットメッセージを送受信している形態をリアルタイムチャットと呼ぶ。

また、本システムにおけるアノテーションを付与す

る流れは、(1)チャットの実施、(2)アノテーションの付与、(3)アノテーションの確認、で1サイクルと考えると、実際のチャットは、上記サイクルが不規則に何度も行われる形態であるといえる。本実験では、アンケート調査時にアノテーションの確認を行い、それ以降のコミュニケーションは継続しなかったことから、上記コミュニケーションの1サイクルを行ったといえる。1サイクルの実験では、(1)から(2)への影響(チャットログからアノテーションへの影響)を調査できるが、(3)から(1)への影響(付与されたアノテーションから以降のチャットへの影響)については知ることができない。そこで、(3)の後に、アノテーションが必要な(欲しかった)語句を聞き出し、マッチングをとることで、付与されたアノテーションが以降のチャットサイクルに影響を及ぼすものかどうかについての判断を行うことにした。

本実験で話し合われたチャットのテーマは、文化の違いが表れるようなチャットコンテキストが必要であることから、特定分野に関する専門知識を必要としない範囲で、有名な観光地や食べ物紹介、流行しているゲームなど6つのテーマの中から被験者双方が話し合いのうえ、1つ選択してもらい、チャットを実施した。

アンケート調査用紙は英語および日本語で書かれたものを準備した。アンケートの自由記述欄に関しては、回答者の母国語もしくは英語で記入してもらった。

上記の異文化間コミュニケーション実験と比較をするため、日本人同士によるチャット実験も行った。条件が同じになるように、お互いに面識がないユーザ同士をペアとして、実験中はお互いの様子が直接分からないように、2つの部屋に分けて実験を行った。

実験では、図1に示した翻訳サーバを1台、Anno-Chatサーバを1台、ともに和歌山大学内に設置した。

## 4.2 実験結果

### 4.2.1 アンケート結果

アンケートにおける5段階評価の結果を表1に示す。表は、(a)チャットペアの組合せ別(日・中、日・韓、中・韓、日本人同士)に集計したもの、および、(b)被験者の国別(日本、韓国、中国、ただし日本人同士の実験で得られた回答は含めない)に集計したものを示す。

アンケート評価の点数は、数値が大きいほど質問項目に同意していることを示せるように、次のように段階分けした。

- (1) 強く同意しない (Strongly disagree)
- (2) 同意しない (Disagree)

表 1 アンケートの結果

Table 1 Result of the questionnaires.

質問項目	(a) 組合せ別				(b) 国別		
	日・韓	日・中	中・韓	日・日	日本	韓国	中国
(1) 相手と円滑にコミュニケーションがとれた。	4.0	3.3	2.5	4.3	3.7	2.7	2.8
(2) 機械翻訳を介した相手のメッセージを母国語のように理解できた。	3.8	2.8	2.0	—	3.2	2.0	2.6
(3) アノテーションはお互いのメッセージの理解に役立つ。	3.8	4.5	3.7	3.9	4.3	3.0	4.3
(4) アノテーションの追加は難しい。	2.0	2.8	2.3	3.1	2.8	2.8	1.8

注) 日本人同士による比較実験のアンケートは (b) 表には含めていない。また、日・日欄の—は、質問項目の内容が異文化間での実験を対象としているため、質問していない。

- 最後にまとめてつけるのではなく、会話の最中に書きたい。(日本)
- チャットしながら、アノテーション機能を使用するのは難しいと思う(日本)
- 相手の国の固有名詞に付けてもらおうと知らないことが分かりやすかった(日本)
- 相手が見つけたアノテーションを見て、読んで理解できたものがあり、とても面白かった(日本)
- アノテーションは役立つと思う。利用者の交流にとって必要です(中国)
- 使いやすいと感じる。アノテーションを使うと、さらに理解できる(中国)
- 会話は同じ時間に行うので、アノテーションを使って説明する必要はあるのか?(韓国)
- 翻訳精度が悪く、役に立たなかった(韓国)

(a) 異文化間実験での感想

(a) Comments of the intercultural experiments.

- 今回は分かる単語ばかりだったので、あまり使用しなかったが、難しい単語などを使用する場合は必要だと感じた。
- 同じ日本人同士だと、だいたい理解できるため、アノテーションを付けにくかった。
- 分からない言葉があれば(チャット中に)文章で聞かれないと思う。

(b) 参考実験(日本人同士)での感想

(b) Comments of the reference experiments (Japanese).

図 4 アノテーション機能に対する感想(アンケート記述, 抜粋)  
Fig. 4 Comments of the annotation function.

- (3) どちらともいえない(Neutral)
- (4) 同意する(Agree)
- (5) 強く同意する(Strongly agree)

図 4 にアノテーション機能に対する自由記述アンケートの結果を示す。自由記述アンケートは、各国の被験者に母国語もしくは英語で記述するように指示したため、記述された項目は日本語に翻訳したうえで集計を行った。

#### 4.2.2 アノテーションに関する集計

本実験で得られたチャット・アノテーションログと、アンケートデータを照合比較してアノテーションに関する傾向を調査した。調査項目は、以下の2点である。

- (1) 付与アノテーションと要求アノテーションの一致

表 2 要求アノテーションと付与アノテーションの一致数

Table 2 Number of equivalent to the requested annotations and the attached annotations.

	日・韓	日・中	中・韓	合計
付与アノテーション数	30	59	100	189
一致数[箇所]	12	21	30	63
割合[%]	40.0	35.6	30.0	33.3

一致数=要求アノテーションと付与アノテーションが一致した数。

#### (2) アノテーションの分類

各項目の調査結果を以下に示す。

- (1) 付与アノテーションと要求アノテーションの一致  
本実験では、アノテーションを付与する語句の数を5つ、語句の選定基準は「相手に理解してもらえないと感じる語句」と指示して実験を行った。その結果、アノテーションを付与する側(送り手)と、アノテーションを求める側(受け手)では、付与対象である語句にどれくらいの差が生じているかについて調査した。

調査方法は、チャット終了後のアンケートにおいて、相手(送り手)が送信したメッセージの中から「異文化間の理解のためにはアノテーションが必要」と思われる語句(要求アノテーション)を5つあげてもらった。

表 2 に、要求アノテーションと、付与アノテーション(送り手が実際にアノテーションを付与した語句)が一致した件数を示す。

#### (2) アノテーションの分類

アノテーションには、いくつかの用途が存在すると考えられる。付与されたアノテーションを、語句の意味やアノテーション内容および、アノテーションに関するアンケートをもとにして、次の3種類に分類した。

##### (a) 辞書タイプ

付与されたアノテーションを、他のチャットと同じ語句に付与してもアノテーションとして成立する(意味が通じる)と考

(1)	辞書タイプ ・「南怡島」: 韓国の代表的ドラマである冬のソナタの撮影地になった場所 ・「ワンピース」: 海賊が主人公の話。
(2)	チャット補足タイプ ・「携帯電話でテレビを見るのだ」: 新しい移动通信業務の一種、携帯電話でテレビ番組を見ることができる ・「マン街」: マンガ本 (タイプミスの補足と思われる)
(3)	問合せタイプ ・「世界版」: 意味が確かではない。 ・「聞く」: listening の意味だが、文脈上合わない ・「」内はアノテーションの語句、右側はアノテーションの内容を示す。 ・語句および内容は、各国語で書かれていたものを日本語に翻訳して提示している。

図 5 アノテーションの分類例

Fig. 5 Classification examples of annotations.

表 3 アノテーションの分類結果  
Table 3 Classification of annotations.

分類項目	(a) 組合せ別			(b) 国別			合計
	日韓	日中	中韓	日本	韓国	中国	
辞書タイプ	25	49	65	37	44	58	139 (74%)
チャット補足タイプ	5	6	20	8	12	11	31 (16%)
問合せタイプ	0	4	15	0	6	13	19 (10%)
合計	30	59	100	45	62	82	189

えられ、異文化間の辞書的な説明をする役割を果たしていると考えられるもの。

#### (b) チャット補足タイプ

付与されたアノテーションを、他のチャットの同じ語句に付与した場合、文脈と意味が合致しない可能性が高いと考えられるもの。

#### (c) 問合せタイプ

付与されたアノテーションに説明的な内容はなく、アノテーションを付与した理由を尋ねたアンケートにおいて「(語句の)意味が分からない」という記載があるアノテーション。

図 5 に実際に付与されたアノテーションを分類した例について、表 3 に 3 項目の集計結果をそれぞれ示す。

なお、要求アノテーションの分類に関して、今回の実験では、アノテーションが必要とされる語句を調査したため、内容に関する分類は行っていない。

## 5. 考 察

### 5.1 コミュニケーションに関して

相手と円滑にコミュニケーションが行われたかどうかについてのアンケート結果(表 1(1))および、機械翻訳を介した相手のメッセージの理解度(表 1(2))について、両方ともに日本人と韓国人の間(日・韓間)が最も高い評価であり、つづいて日本人と中国人の間(日・中間)、中国人と韓国人の間(中・韓間)という傾向が見られた。

この傾向について、機械翻訳機の精度を調べた ICE2002 の文献<sup>12)</sup>によると、ICE2002 で用いた機械翻訳の精度は、日・韓間が最も良く、つづいて日・中間、最も精度が悪かったのが中・韓間であると報告している。本実験で使用した機械翻訳機は、ICE2002 で用いたものと同一社製であることから、コミュニケーションの円滑性や理解度は機械翻訳の精度に依存したと考えられる。

### 5.2 アノテーションの機能

本実験で付与されたアノテーションについて、使われ方や機能の観点から 3 種類に分類した結果(表 3)、最も多かったのが辞書タイプ(74%)であった。辞書タイプは、チャットコンテキストに依存しない語句の説明であり、このタイプのアノテーションは蓄積・共有をすることで、異文化間の意味情報を持つ資源となり、他のチャットで再利用することで、アノテーションを新規作成する負荷を軽減できると考えられる。

つづいてチャット補足タイプ(16%)は、特定の発言を補足することなどを目的としたアノテーションで、チャットコンテキストが違くと語句の説明として成立しないものである。このタイプのアノテーションは、チャット中に伝えることができなかつた意味や説明を、チャットとは別の流れ(チャンネル)で付け足したと考えることができる。

上記 2 種類のアノテーションについて、辞書タイプの蓄積・再利用が進むと、有益な異文化間の意味情報を持つ資源に成長する可能性が高い。しかし、蓄積された意味情報だけですべてのケースを説明できるとは考えにくいことから、チャット補足タイプのアノテーションや、辞書タイプの蓄積されたアノテーションが存在しない場合には、その場でアノテーションを付与する必要が生じてくるため、リアルタイムに付与するアノテーションも必要である。

残る 10%のアノテーションは問合せタイプであった。このタイプのように、意味を問い合わせることを目的にアノテーションを付与するという使用法は、当初

予定していた実験手順に反した点がある。

問合せタイプと分類した 19 件のアノテーションについて、付与語句の前後に位置するコンテキストを確認したところ、13 件が、翻訳ミスと思われる語句に付与されていたことが分かった。本実験では、アノテーションを語句の問合せに使うことは想定していなかったため、数の大小に関しては評価できないが、機械翻訳のミスや、異文化間での意味を理解できない語句が多数発生することは容易に想定できるため、語句の意味を問い合わせるようなフィードバック要素を持つアノテーションが今後必要になると考えられる。

### 5.3 アノテーションの有効性

実験後のアンケート(表 1(3)(b))より、「アノテーションがお互いのメッセージ理解に役立つ」という質問には、日本人被験者と中国人被験者はいずれも 4.3 ポイントで、おおむね役立つという評価を得ることができた。アノテーション機能に関する感想(図 4(a))からは、「アノテーションが理解に役立った」という肯定意見が多く見られたが、一部被験者からは「疑問があれば会話内で解決できるためアノテーションは必要ない」といった意見も得られた。しかし、チャット内で疑問を解決しようとした場合、チャットの流れをさえぎることになるため、多くの質問はできない可能性が高い。意味が分からない語句や様々な疑問が発生することが予想される異文化間コミュニケーションでは、細かな疑問や意味を自己解決できるアノテーション機能や、アノテーションが付与されていない場合には、チャットの流れに影響されずに問合せを行う機能が必要と考えられる。

また、アノテーションを付与する一連の過程について、難しさを感じたかどうかについてのアンケートを実施したところ、表 1(4)より、異文化間(日・韓、日・中、中・韓)ではいずれも 3.0 ポイントを下回り、難しくないと感じる被験者が多かったといえる。しかし、日本人同士の実験では 3.1 ポイントと、異文化間よりもわずかに難しいという結果が出た。日本人同士でのアノテーション付与が、異文化間よりも難しいと感じさせた理由について、アノテーション機能に対する感想(図 4(b))より、同じ日本人同士では、相手が知らないと思う語句を探すことに苦労した、という主旨の感想が多く見られた。異文化間コミュニケーションでは、相手が知らないと思う語句が多数あることから、補足すべき内容も多いと考えられ、同じ文化間よりも有益に機能すると考えられる。

### 5.4 円滑なコミュニケーションに向けて

5.1 節の考察より、今回の実験ではコミュニケーショ

ンの円滑さは、機械翻訳の精度に依存するという結果が得られた。これは、チャットメッセージとアノテーションデータのいずれも機械翻訳を介していることから、基盤となる翻訳部分の精度が低いと、正しく伝わらなくなるためだといえる。そこで、翻訳精度の問題とは別に、アノテーションを用いた円滑なコミュニケーションの支援について検討を行う。今回の実験で、付与されたアノテーションと、ユーザが要求するアノテーションとの一致が 33%(表 2)という結果が得られた。この値についての良し悪しは、比較対象がないため判断することができない。しかし、この数値が高くなればなるほど、付与されたアノテーションが有効に活用され、コミュニケーションの円滑さを向上させる可能性があると考えられる。

## 6. おわりに

本論文では、機械翻訳を用いた異文化間コミュニケーションのためのチャットツール AnnoChat を開発し、中国人、韓国人、日本人の間でユーザ主導によるアノテーション付与の適用実験をした結果について検討を行った。

実験より、リアルタイムコミュニケーション(チャット)へのアノテーションの付与では、以下に示す 3 種類のアノテーションを得ることができた。最も多かったものが、異文化間の意味情報となる辞書タイプのアノテーションが 74%を占め、続いて、チャットのコンテキストを補足するチャット補足タイプのアノテーションが 16%であった。前者の辞書タイプは、蓄積し共有(再利用)することで、アノテーションを新規作成する際の負荷を下げ、簡単にアノテーションを付与することが可能になると考えられる。また、意味情報として蓄積されていない語句や、後者のチャット補足タイプに対しては、チャット中にアノテーションを付与できる機能が重要になると考えられる。

残る 10%のアノテーションは、語句の意味を問い合わせるアノテーションであった。機械翻訳を介した異文化間コミュニケーションでは、語句の意味を問い合わせたり、不適切な翻訳に対応したりするためのフィードバック機能が必要とされていることが分かった。

被験者の感想から、アノテーションがコミュニケーションを円滑にする可能性が示唆されたが、今回の実験では、アノテーションが異文化間コミュニケーションを円滑にしたという評価までは行えなかった。今後は、今回得られたアノテーションの機能分類に基づいて、アノテーションの利用を向上させる支援を行う予定である。



謝辞 実験に協力していただきました，京都大学石田亨研究室，中国・Shanghai Jiao Tong University，Southwest University，韓国・Kyung Hee Universityの皆様へ深く感謝いたします。

本研究の一部は，「独立行政法人情報処理推進機構」の「平成16年度未踏ソフトウェア創造事業」の支援を受けた。

### 参 考 文 献

- 1) Flournoy, R.S. and Callison-Burch, C.: Secondary Benefits of Feedback and User Interaction in Machine Translation Tools, *Workshop paper for "MT2010: Towards a Roadmap for MT" of the MT Summit VIII* (2001).
- 2) 船越 要, 山本晃成, 藤代祥之, 野村早恵子, 石田亨: 異文化コラボレーション支援システムの設計, 情報処理学会第65回全国大会, 4A-5 (2003).
- 3) 安岡美佳, 中小路久美代, 大平雅雄, 石田亨, 野村早恵子: 異文化協調作業における共有理解構築の機会としてのコミュニケーションエラー現象の利用, 情報処理学会研究報告, 2003-HI-103, pp.47-54 (2003).
- 4) 内元清貴, 林田尚子, 石田亨, 井佐原均: 機械翻訳可能性の自動評価, 言語処理学会第11回年次大会発表論文集, pp.9-12 (2005).
- 5) Yokoyama, S., Kashioka, H., Kumano, A., Matsudaira, M., Shiokizawa, Y., Kodama, S., Ehara, T., Miyazawa, S. and Murata, Y.: An Automatic Evaluation Method for Machine Translation using Two way MT, *Proc. 8th MT Summit Conference* (2001).
- 6) Frederking, R.E., Black, A.W., Brown, R.D., Moody, J. and Steinbrecher, E.: Field Testing the Tongues Speech-to-Speech Machine Translation System, *Proc. 3rd International Conference on Language Resources and Evaluation (LREC 2002)* (2002).
- 7) 東 大介, 武井恵雄, 荒井正之: アノテーションを活用した議論支援ツールの作成, 情報処理学会研究報告, 2003-CE-72, pp.63-68 (2003).
- 8) 杉山香織, 船越 要, 神田智子, 藤代祥之, 石田亨: セマンティックアノテーションに基づく多言語コラボレーション支援, 情報処理学会研究報告, 2005-ICS-139, pp.157-162 (2005).
- 9) ICE2005.  
<http://www.ai.soc.i.kyoto-u.ac.jp/ice/2005/>
- 10) 伊藤 勇, 徳川直人 (編著): 相互行為の社会学, 北樹出版, pp.185-187 (2002).
- 11) 村田和義, 川口 修, 渋谷 雄, 倉本 到, 辻野嘉宏: チャット参加者の応答期待時間の延長を旨とした情報提示とそのタイミング, ヒューマンインタフェース学会論文誌, Vol.8, No.3, pp.423-433 (2006).
- 12) 小倉健太郎, 林 良彦, 野村早恵子, 石田亨: 目的志向の異言語間コミュニケーションにおける機械翻訳の有効性の分析—異文化コラボレーション ICE2002 実証実験から, 情報処理学会第65回全国大会, 2T6-4 (2003).

(平成18年5月30日受付)

(平成18年11月2日採録)



藤井 薫和 (学生会員)

昭和58年生。平成18年和歌山大学システム工学部卒業。同年同大学大学院システム工学研究科博士前期課程入学。現在、同研究科にて異文化間コミュニケーションに関する研究に従事。



重信 智宏 (正会員)

昭和52年生。平成12年鹿児島大学工学部電気電子工学科卒業。平成15年和歌山大学大学院システム工学研究科博士前期課程修了。平成18年同大学院システム工学研究科博士後期課程修了。博士(工学)。現在(独)情報通信研究機構専攻研究員。異文化コラボレーションに関する研究に従事。



吉野 孝 (正会員)

昭和44年生。平成4年鹿児島大学工学部電子工学科卒業。平成6年同大学大学院工学研究科電気工学専攻修士課程修了。平成16年より和歌山大学システム工学部デザイン情報学科助教授。博士(情報科学)東北大学。平成17年IPA未踏ソフトウェア創造事業スーパークリエイター認定。コラボレーション支援に興味を持つ。