

機械翻訳サービスを用いた専門知識伝達サービスの分析

An analysis of expert knowledge transmission service
using machine translation services鈴木宏[†]

Hiroshi Suzuki

菱山玲子[†]

Reiko Hishiyama

1. はじめに

グローバル化が進み、ビジネスや学問などの様々な領域で各国固有の専門知識、技能やノウハウ、文化の情報を他国の人々に伝達する、ないし、各国の人々が協同して専門知識の補完的共有を行う機会が増大している。同時に、情報化社会の進展と共にICT技術が技能・知識伝承や共有といった知的活動で用いられる場面も増えている。今後はグローバル化する社会経済を背景に、こうしたICT技術の多言語間での利用も一般的なものになると考えられる。

グローバル環境での知識伝達や知識共有では、機械翻訳サービスの利用場面も増えることが予想されるが、一方で、現状では必ずしも常に十分な翻訳精度が保証されない機械翻訳サービスを用いた知識伝達コミュニケーションでは、知識が誤まって伝達されたり、伝達すべき知識が欠落したりするといった知識伝達の失敗リスクも増大する。そこで本研究では、こうした知識伝達における失敗リスク低減のための、利用者からの適応的アプローチを検討する。具体的には、機械翻訳サービスを用いた多言語チャットと専門用語コーパスを併用した多言語知識伝達システムにより専門知識の知識伝達実験を行い、実験から得られたデータ、並びに知識を学び取る側の理解度評価を行った結果を分析し、機械翻訳サービスを利用した知識伝達においてどのような伝達方法が確実かつ効果的かを明らかにする。

本稿の構成は、次のとおりである。続く2章では知識伝達に関する研究を概観すると共に、多言語環境で機械翻訳サービスを用いて情報伝達を行っている一連の研究を説明し、本研究の位置づけを明らかにする。第3章では、本研究の提案の概要と実験内容について述べ、第4章では実験結果を報告する。第5章では4章の実験結果を考察し、機械翻訳サービスを用いた専門知識の効果的な伝達方法を明らかにする。最後に、結論として本研究のまとめと今後の課題を述べる。

2. 関連研究

グローバルな環境で実際に機械翻訳サービスを用いて知識や情報の伝達・共有を行っている事例として、YMC-Viet プロジェクト [1] が挙げられる。本プロジェクトの目的は日本からベトナムへの稲作知識支援であり、言語グリッド [2] の機械翻訳サービスを利用したシステムを用い、日本人専門家による農業知識の共有や稲作の技法・ノウハウ提供を行っている。Kitaら [3] はこの活動に対し、機械翻訳や人による翻訳作業、原文の書き換えなど、多言語コミュニケーションにおける情報伝達のための様々な手法を組み合わせて実践するこ

とによるパタンごとの伝達効果の評価・分析を行っている。更に、大谷ら [4] は、この活動におけるコミュニケーションログ、及び、知識の提供者側と享受者側双方に対するインタビュー結果を分析している。専門家は農業に対する専門知識を所持し、ベトナム人は現地の気候状態をはじめとした情報を有することで、専門家による知識伝達と同時にベトナム人からの情報の共有も行っているが、このように所持する情報の相違から生じる知識伝達における状況認識差の存在を明らかにすると共に、その認識差を埋める方法としてコミュニケーションプロトコルの改善や知識享受者側のコミュニティにおける専門家の支援を提案している。

これらの研究はいずれも、多言語環境での知識伝達サービスにおける知識伝達の有益性と共に、その困難さを指摘している。特に後者はコミュニケーションのための組織の在り方や機械翻訳の技術的改良に及んでおり、その両面からの改善が図られつつあるが、組織や翻訳サービス構成の違いに起因する比較効果については十分な知見がなく課題が残されている。この点で、伝達のための組織の在り方と機械翻訳サービス利用の両面から基礎的な改善効果を明らかにする必要がある。

一方、知識伝達と相互理解に焦点を当てて、伝達のための機械翻訳サービスの利用効果を実験室実験から調べたものとして、山下ら [5][6] の研究や、稲葉ら [7] の研究が存在する。山下ら [5] の研究では、もともと予め用意した原文に対し、折り返し翻訳を用いた原文書き換え作業を、Arendseら [8] により提供されている機械翻訳システムに関するガイドラインを元に定義した良い翻訳結果を得るための書き換えルールと共に行うことで、書き換えユーザの機械翻訳に対する適応とルールの使用効果の分析を行った。原文の書き換え実験により、定義したルールを書き換えを行うユーザに知識伝達することで、ユーザが母国語に関する知識が豊富でない場合でも機械翻訳への適応を容易に行えたことを報告している。こうした原文の書き換えレベルの研究は、機械翻訳サービスの利用において利用者の適応的アプローチの重要性を指摘するものであり、本研究と狙いを同じくするものである。

更に、より高次の機械翻訳を介した相互理解に重要な参照表現としての対象指示コミュニケーションに焦点を当てた文献 [6] や文献 [7] がある。文献 [7] の研究では、コミュニケーション行動を、母語の異なる被験者同士による図形のマッチングという、同じ図形群を異なる番号順で並べたものの順序を合わせていくタスクに表現し、機械翻訳を用いた多言語チャットを介途中で生じる相互理解の状況を分析している。前者の文献 [6] では異言語ユーザ2者間の研究として行われ、後者

[†]早稲田大学大学院創造理工学研究所経営システム工学専攻

の稲葉ら [7] の研究では異言語ユーザ 3 者間の研究として拡張的に行われている。様々な伝達要素の中の参照表現に注目した以上の研究では、被験者である異言語ユーザに予め図形の形や並びという一定の情報を与えることで指示参照表現に注目した研究を行うためのタスクとして設計している一方、本研究では専門知識伝達に関わる様々な情報伝達行動の出現を想定している。つまり、より一般的な多言語環境での知識伝達の方法を探るタスクとして設定しており、文献 [6] や [7] とは異なる情報環境を前提とし、知識の提供者側と享受者側とで知識やスキルの差があり、事前の共有情報がないような非対称性のある情報環境を想定している。

なお、Szulanski ら [9] は、知識伝達は行動やモデルとしてではなく、伝達プロセスとして考えることの重要性を指摘したうえで、伝達する知識自体に加え、知識伝達の管理としての介在方法、知識伝達を支援するための組織的なメカニズムデザインを考えることの必要性に言及している。本研究は特に、企業などでは一般的と考えられる専門用語や専門知識を含む情報環境を想定したタスクを想定している。そのうえで、実験から機械翻訳サービスの利用に適した多言語環境での組織メカニズムを考察し、機械翻訳サービスの介在によって知識伝達の効果を高める方法の提案を目指す点で、既に用意した書き換えルールを実践しその効果を検証している文献 [5] や参照表現に注目した文献 [6] や文献 [7] とは、位置付けが異なっている。

3. 提案

3.1. 概要

本研究では、知識を提供する側を知識提供者とし、知識を享受する側を知識享受者とし、母語が異なる知識提供者から知識享受者への専門知識を含む知識伝達の効果を扱う。前章で述べたように、知識伝達では、知識伝達の管理としての介在方法や知識伝達を支援するための組織的なメカニズムデザインが円滑な知識伝達に重要であることが指摘されている。一方、企業や社会的な組織において機械翻訳サービスを介したコミュニケーションを通じて知識が伝達される過程では、同言語ないし多言語のコミュニケーションが混在し、組織の言語環境の違いにより知識伝達の効果が異なることが予測される。また、多言語サービスによる支援や介在の在り方もより複雑になり、組織の言語環境に最も適した介在を行う必要が生じる。更に、組織の中に伝達される知識は単一で断片的かつ独立に扱える知識ばかりではなく、知識が相互に関連づけられているケースや専門性の高い知識を含むケースを想定しておかなければならない。例えば、想定される具体的な知識伝達の場面とは、企業活動における業務フローや工程・作業手順の説明などが該当する。現時点での機械翻訳サービスの翻訳精度については、文献 [7] が指摘しているように、翻訳精度の向上を待つだけではなく、それ以外の支援方法を考案することも重要である。従って、サービス利用者の側での適応として、知識伝達の方法を正確かつ効率的に考案することも必要となっている。本研究は、こうしたより実用的な場면을想定した知

識伝達の効果を考えることを目的としている。すなわち、本研究で扱う課題は以下にまとめられる。

1. 知識提供者と享受者からなる組織が多言語環境であることを前提に、実用的な場면을想定しつつ機械翻訳サービスを利用したどのような知識伝達が効果的かを、利用者側の適応から明らかにする。
2. 知識享受者側における組織学習の枠組みを組合せて、機械翻訳サービスによる知識伝達の正確性を最大化する方法を明らかにする。
3. 相互に関連づけられた複雑な知識や専門性の高い知識を共有するために効果的な方法を、具体的な機械翻訳サービスの利用方法として考案する。

これらの課題に取り組むため、本研究では多言語による専門知識伝達システムを用い、知識提供者から知識享受者への知識提供実験を行い、その伝達効果の測定を行うこととした。実験の被験者は、専門知識を教える知識提供者と専門知識を学ぶ知識享受者である。知識提供者は日本人 1 名とし、知識享受者は日本人ないし中国人からなる 3 名で集団パターンを構成する。3 名の集団パターンの詳細は次章で説明する。

3.2. 実験システム

本研究で用いた専門知識伝達サービスのシステム構成を、図 1 に示す。本システムは、言語グリッド [2] 基盤上の機械翻訳サービスを用いて実現した多言語チャットシステムに、専門用語コーパスやメモ記録機能といった、知識伝達と学習を行うために効果的だと考えられる機能を追加したアプリケーションである。言語グリッドの接続には多言語工房 [10] を利用し、高電社の J-Server (機械翻訳 Web サービス) を用いて日中翻訳サービスを提供する。なお、本システムの多言語チャット部には翻訳リペア [11] 機能が付加されているが、翻訳リペア機能の利用は被験者の任意である。文献 [11] によれば、日本語入力から英語・中国語・韓国語への文の翻訳リペアを複数回行った結果、全体の 65% の文の平均翻訳精度を意味伝達可能なレベルまで改善することができたことが報告されており、被験者が有効に利用すれば翻訳精度の向上が得られる可能性がある。

システムのインタフェースを図 2 に示す。システムのインタフェースは 4 パートに分かれており、画面左が専門用語コーパス部、中央が多言語チャット部、画面右上が知識提供者専用でビデオにより専門知識を学ぶ部分、右下が各自でメモを記録し表示させる部分である。実験で扱う専門知識は「日本酒の造り方」とし、専門用語コーパスはそれに対応して予め意味の検索が予想される専門用語 20 個を抽出し、日本語と中国語に対応したものとして作成した。

3.3. 実験内容

実験の手順は以下のとおりである。

1. まず知識提供者 1 名が単独で、「日本酒の造り方」のビデオをビデオライブラリから再生することで知識享受者に教える知識を学ぶ。

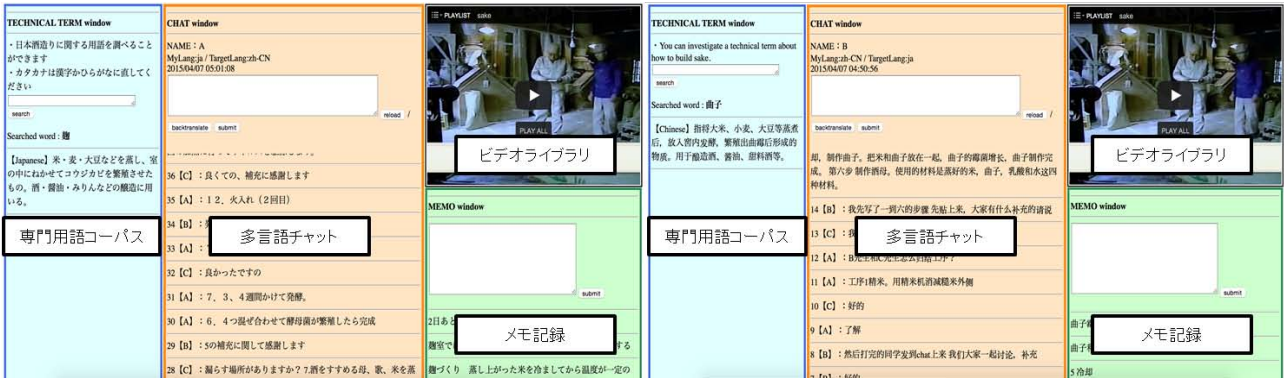


図 2: システム・インターフェース (左: 日本語表示, 右: 簡体字表示)

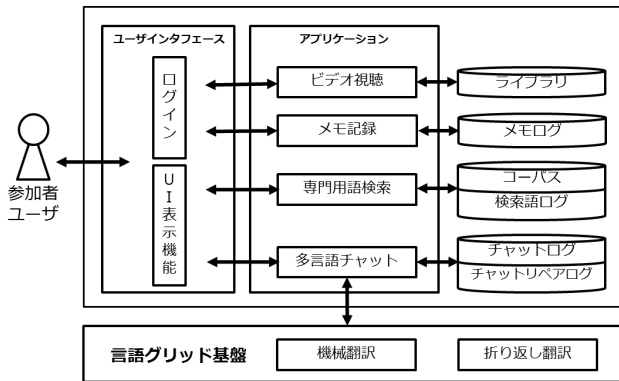


図 1: システム構成図

- 知識提供者の学習後, 知識享受者 3 名を参加させ, システムを用い知識提供者側から知識享受者側への知識提供を多言語チャットをはじめとした実験システムの諸機能を用いて行う (本過程が知識提供過程とする). 知識提供は 40 分間とし, この間, 知識享受者は知識提供者への質問や専門用語コーパスの使用が可能である.
- 知識提供者から知識享受者への知識提供後, 理解度確認のため, 知識享受者は知識提供によって学んだ「日本酒の造り方」に関する基礎的な知識の理解度調査として工程の並び替え問題に解答する. 工程の並び替え問題とは, 日本酒造りの一つ一つの工程の名称と写真をプリントしたカード (図 3) を, 順番が正しくなるよう並び替えるものである.



図 3: 並び替え問題で使用するカード例

- その後, 知識享受者のみでグループワーク (チャット) に取り組む. グループワークの課題は「日本酒の造り方を, それを知らない第三者になるべく詳細に教える」としたときの説明文をグループで一

つ作成せよ。」というものであり, 説明文はチャット画面上で投稿する. このアウトプットとなる説明文も理解度評価, また知識伝達による伝達効果の評価対象とする. このグループワークは情報共有の場にもなっており, 途中で先に解いた並び替え問題の正解情報が開示されることはなく, 情報は全てグループの各メンバーの享受した知識を合わせたものの範囲内で共有される (本過程を知識共有確認過程とする).

- グループワーク後, 知識享受者は再び並び替え問題に解答する. 最後に内容に関する質問とアンケートに回答後, 実験終了となる.

知識享受者 3 名の集団パタン構成は, 中国人 3 名による実験, 日本人 3 名による実験, 中国人 2 名に日本人 1 名を混在させた集団で行う実験の 3 パタンである. 各構成による実験を 4 回ずつ行い, 計 12 回実施した. 知識提供者は各回異なる日本人 1 名である. 各実験パタンについて, 次章で提案する理解度評価方法による結果比較を行うと共に, 多言語チャットログを用い知識提供者の知識伝達方法を分析することで, 機械翻訳を用いる際の正確かつ効果的な伝達方法を考察する.

4. 実験結果

4.1. 誤誤を感じた割合

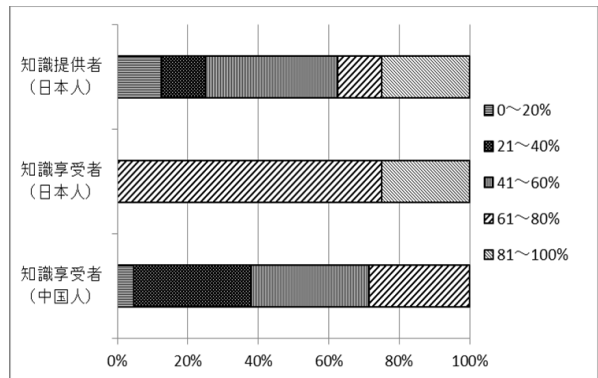


図 4: 誤誤を感じた割合 (国別)

まず初めに, 直感的な翻訳精度の印象として, 被験者が誤誤を感じた割合[‡]を図 4 に示す. 図より, 中国

[‡] アンケート調査により得られたデータである.

人よりも日本人の方が相対的に誤訳を感じる割合が高いことがわかるが、中国人の割合を見ても半数近くの人々が40%以上の誤訳を感じていることから、専門知識の提供を行う本実験では被験者の国籍に関係なく誤訳を感じた割合が大きいといえる。この点で、多言語環境での翻訳精度に起因するユーザ適応の問題を検討することの意義を見出すことができる。

4.2. 基礎的な理解度

知識享受者の基礎的な知識の理解度調査に用いた工程並び替え問題の実験パターンごとの平均正答率(%)を図5に示す。1回目は知識提供者による提供後、2回目は情報共有としてのグループワーク後に、各知識享受者から解答を得たものである。

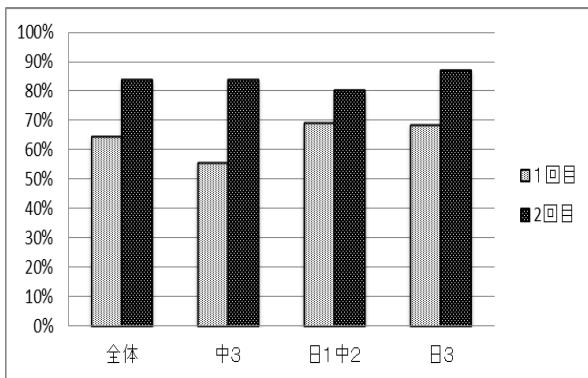


図5: 並び替え問題正答率 (パターン別)

図5より、情報共有となるグループワーク前に比べ、グループワーク後は主に順序が誤っていた部分が修正されるなどしていずれのパターンにおいても正答率が80%以上まで向上した。また、1回目と2回目の各パターン毎の正答率にばらつきは見られず(分散分析: $p > 0.1$)、情報共有のプロセスとしてはどの実験においても初めの工程から順番に皆で詳細な情報も含め確認しながら共有を行っていた。

次に、全知識享受者を日本人と中国人の国別に分けた時の正答率(%)を以下の図6に示す。

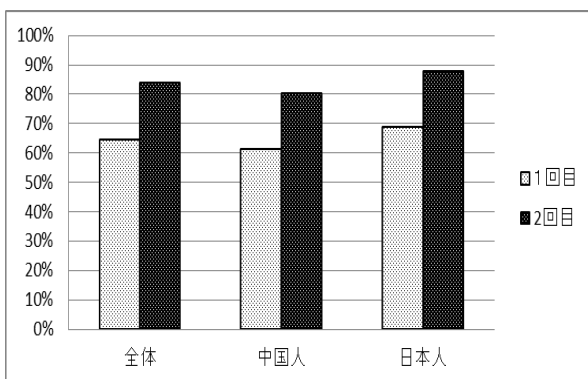


図6: 並び替え問題正答率 (国別)

図6より、中国人でも1回目で60%の正解率を残し、グループワークによる情報共有後には80%まで伸ばすことができている。1回目と2回目共に日本人と比べても正答率にばらつきは見られなかった(t 検定: $p >$

0.05)。以上より、基礎的な知識に関しては中国人の場合でも知識提供者からの知識提供過程において内容をある程度享受できており、彼らで情報共有を行うことで80%程度まで正答率を上げることができるといえる。

ここで、さらに中国人について、グループワークが同言語環境である中国人3名のパターンと多言語環境である中国人2名に日本人1名を加えたパターンでの正答率を比較するため、以下に図7として示す。

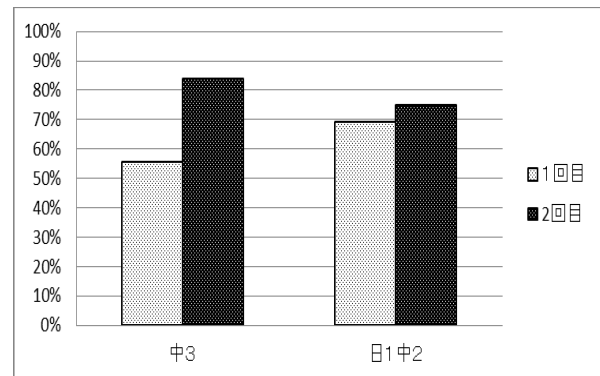


図7: 中国人の並び替え問題正答率 (実験パターン別)

図7より、同言語環境である中国人3名で情報共有を行う方がグループワークによる正答率の向上が高い傾向にあり、両者の向上の割合に差が確認できた(t 検定: $p < 0.05$)。このようになった理由として、日本人が入らないことにより多言語環境にならず翻訳精度に関する配慮を行う必要がないため、その分情報共有を円滑に行えたことが考えられる。

4.3. 知識伝達全体に対する理解度

次に、詳細な情報も含めた知識伝達過程全体に対する理解度を表すために、知識提供者が知識提供過程で提供した知識範囲に対し、知識享受者がグループでどれほど学習できているかを「学習度」として数値化したものを示す。

ここで「学習度」の値は、グループワークのアウトプットである知識享受者による「日本酒の造り方」の説明文を独自に考案した採点基準に基づいて採点したものと、知識提供者が知識伝達過程で知識享受者に教えている時のチャットログを同基準で採点したものの商(割合)で定義され、各知識教示者それぞれの理解度のもと、それらに対しどれ程の内容を各学習者グループに伝達できたかを算出する。この採点において中国語で書かれた説明文は日中バイリンガルに翻訳を依頼し、日本語へ翻訳したものをを用いて採点した。ここで以下に考案した採点基準を示す。

表1: 説明文の採点基準

語句	加点数(寡少の間違いの時)
日本酒造りに重要な各工程を言えている	1語につき3点(2点)
工程ごとに詳細説明をする文言を含んでいる	1語につき2点(1点)

この採点基準をもとに算出した学習度(%)を実験パターンごとに以下の図8に示す。

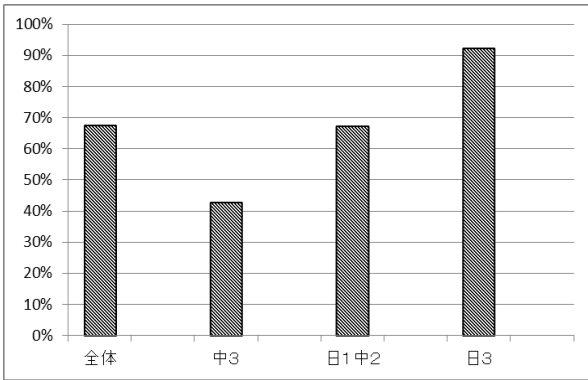


図 8: 知識伝達に対する学習度 (パターン別)

図 8 より、中国人 3 名よりも日本人 1 名と中国人 2 名のパターンの方が学習度の値が高く、そのパターンよりも日本人 3 名の方がさらに高くなるという結果になった (分散分析: $p < 0.01$)。従って、知識提供者と母国語の異なる中国人を 3 名そろえた実験では、母国語が同じ日本人 3 名の場合と比べると知識内容全体を対象とした学習度に大きな差が存在していることがわかる。

しかし、ここで中国人 3 名のパターンよりも日本人 1 名と中国人 2 名のパターンの方が学習度が高いことに注目すると、知識提供者と母国語が異なる知識享受者の集団に母国語が同じ知識享受者を 1 名でも加えることで、グループとして学習できる知識範囲を向上させることができることが知見として得られた。

機械翻訳サービスを用いた多言語コミュニケーション環境では、現状の翻訳精度から生じる誤訳による悪影響が想定されるが、この評価結果は、その悪影響よりも知識提供者と同じ母国語で知識を得た知識享受者からの知識共有の効果の方が結果として大きいことを示しているといえる。

5. 考察

5.1. 効果的な教事例

機械翻訳サービスを用いた専門知識伝達を行う際の効果的な伝達方法について考察する。ここで知識伝達全体に対する理解度として、学習度の実験別の結果を以下の図 9 に示す。

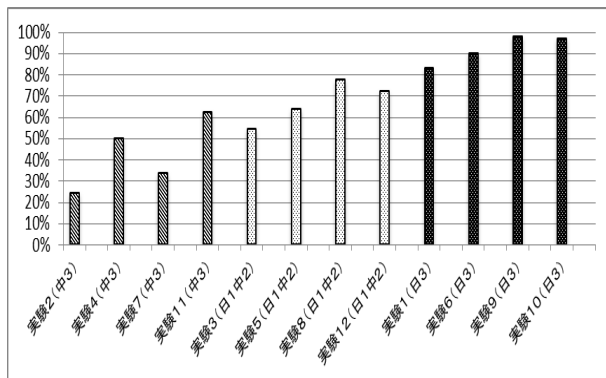


図 9: 知識伝達に対する学習度 (実験別)

実験の番号は実験を実施した順に振られ、図 9 ではそれらを実験パターン毎にまとめて並べている。図 9

より、中国人に効果的に教えている実験として実験 11, 8, 12 に注目し、実験で行われた知識提供者による知識伝達過程でのチャットログの一例を原語スレッド及び翻訳スレッド [12][§] で示す。

● 実験 11

表 2[¶] に、実験 11 の原語スレッドと翻訳スレッドを示す。実験 11 は日本人の知識提供者に対し、知識享受者を中国人 3 名とする実験パターンである。

表 2: 実験 11 の知識伝達例

原語スレッド	翻訳スレッド
「米を洗う」→「浸漬」→「米を蒸す」	「洗米を蒸す」→「浸漬」→「米」
「麴作り」→「酒母作り」→「仕込み」→「搾り・ろ過」→	「麴の制作」→「酒母制作」→「訓練」→「搾る・ろ過」
「火入れ (1 回目)」→「熟成」→「調合・加水」→「火入れ (2 回目)」→日本酒の完成です	「ファイヤボット (1 回)」→「成熟」→「調合・加水」→「ファイヤボット (2 回)」→日本酒の完成
「浸漬」=米に吸水させること	「浸漬」=コメの吸水
「米を蒸す」=1 時間米を蒸します	「米を蒸す」=一時間米
「麴作り」=米に麴菌をかけ米のでんぷんを糖化させます	「麴の制作」=コメの上に麴菌を掛けるコメのでんぷんを糖化する
「酒母作り」=麴、蒸し米、乳酸、水を混ぜ酒母を作ります	「酒母の制作」=麴、蒸し混ぜ米、乳酸、水が酒母を作る
「仕込み」=酵母、麴、蒸し米、水を 4 日間で 3 回に分けて混ぜると『もろみ』が出来ます 糖化と発酵を同時に行う=並行複発酵	「教育」=酵母、麴、蒸し米、水 4 日で 3 回に分け混ぜると、ろ過していない酒も同時に糖化と発酵が進行する=並行複式発酵
「搾り・ろ過」=『もろみ』を木綿袋で搾り、ろ過します	木綿袋を使い絞る「ろ過」=「ろ過をしていない酒」、ろ過
「火入れ (1 回目)」=搾った上澄みを加熱殺菌します	加熱殺菌「ファイヤボット (1 回)」=絞った上澄みの部分
「熟成」=半年間タンクで熟成させます	「成熟」=半年タンクで熟成にさせる
「調合・加水」=品質の異なるタンク内の酒を調合しアルコール濃度を水で調整します	調合「調合・加水」=質が異なるタンク内の酒を水を用いてアルコール濃度を調整する
「火入れ (2 回目)」=調合し加水した酒を加熱殺菌します	「ファイヤボット (2 回)」=調合加熱殺菌加水下の酒

知識提供者は知識伝達を行う流れとして、はじめに「日本酒の造り方」について最初の工程から最後の工程までの一連の流れを簡単に紹介し、その後各工程に関して詳細に説明を加えている。この時、一連の流れを紹介するときの知識伝達方法

[§] 原語スレッドとは、多言語チャットでのコミュニケーションで実際に発話された日本語文であり、知識提供者が発した文章を原文のまま扱う。つまり、母国語を日本語としてシステムを利用する時、チャット画面に表示されるものと同じものである。翻訳スレッドとは、母国語を中国語とした時にシステムに表示されるチャット内容を、日本語に翻訳したものである。つまり、機械翻訳を介して中国語に翻訳された知識提供者による発話を、中日バイリンガルにより日本語に翻訳し直したものである。

[¶] 表中の「仕込み (日本語)」が機械翻訳により「教育 (簡体字)」に翻訳されていた。教育 (簡体字) には「訓練」や「仕立てる」といった「仕込み」にニュアンスが似た意味があるため、このように翻訳されたと考えられる。

において、工程名を「」の記号で囲み、順序を表すときには「→」の記号を用い工程を並べている。この部分の機械翻訳文を見てみると、これらの記号は機械翻訳を経て形を変えず、「」は引き続き中国語に翻訳された工程名を囲み、それらを「→」が原文と同様に繋いでおり、それが翻訳スレッドにも反映されている。また、紹介した工程を詳細に説明する段階において、「～は」や「～とは」といった表現をするために「＝」の記号を用い工程名と説明文を繋いでいる。この「＝」も、一部の翻訳結果では位置が変わっているものの、変わっていない発話では原語スレッドと同じ意味がある程度伝わっていることがわかる。

● 実験 8

実験 11 と同様に、以下に実験 8 の知識伝達でのチャットログの一例を示す。実験 8 は日本人の知識提供者に対し、知識享受者を中国人 2 名に日本人 1 名を加えたパターンである。

表 3: 実験 8 の知識伝達例

原語スレッド	翻訳スレッド
全部で 12 の工程を説明します。	全てで 12 工程。
1 精米	1 精米
日本酒の原料は玄米です。	日本酒の原料は玄米です。
機械を使って、玄米の一番外側の皮を取り除きます。	機械を使い、玄米の一号外側の皮を除去します。
なぜなら、玄米の外側の皮には脂肪とタンパク質が含まれており、それらは日本酒の味を損ねてしまうからです。	原因は、玄米の外側の皮は脂肪やたんぱく質が含まれており、日本酒の味を害しているから。
2 米を洗う	洗う 2 米
精米が済んだお米を水で洗います。	水で精米の終わった米を洗う。
3 浸漬	3 浸漬
洗ったお米を水に浸し、吸水させます。	洗った米を水に浸し、水を吸収させます。
4 米を蒸す	蒸 4 米
水に浸したお米を「こしき」の中に入れ、「1 時間」蒸します。	水に浸した米を「越過來」に入れ。「一時間」蒸す。
蒸したお米は、後の工程で使います。	蒸した米は、後の工程で使います。
5 麴（こうじ）作り	5 麴の制作
蒸したお米を、「麴室（こうじむしろ）」に運びます。	「麴室」へ行き、蒸した米を運送する。
6 酒母作り	6 酒母の制作
麴、蒸し米、乳酸、水をつ一つの鍋に入れて配合します	麴、蒸した米、乳酸、水をつ一つの鍋に入れて配合します
ここでは、「酵母」を増殖させて、「酒母」をつくります。	ここで、「酵母」を増殖させて、「酒母」を作ります。

初めに、「全部で 12 の工程を説明します。」というように、工程が全部でいくつあるのか数字で紹介している。知識享受者には事前に共有されている情報が一切ないため、これから学習する知識の情報量を事前に推測することは難しいと考えられる。その点で、初めに日本酒造りの工程がいくつあるのか知ることによって、学習を進めながら現在どの程度まで工程の学習が全体に対して進んでいるのか推測しながら学習できるという点で、初

めに大まかな情報量を伝達することは有効であると考えられる。

知識伝達の具体的な流れとして、実験 11 では初めに全工程を一通り紹介し、その後各工程を詳細に説明しているのに対し、ここでは初めの工程から 1 工程ずつ順に詳細な説明をしている。

チャットの投稿方法として、各工程の説明の際にまず工程の名称を英数字を付与して投稿し、次の投稿で工程に関する詳細な説明を行っていた。この際、工程の情報量が多くなる際には複数の投稿に分けることにより、1 回の投稿をなるべくシンプルな 1 文で投稿しようとしていることがわかる。

ここで翻訳スレッドを見てみると、一部に対しては誤訳が発生しているが、工程ごとの説明をする投稿はシンプルな文章が多く、それらは多くの投稿で意味がある程度伝わっていることがわかる。数字を使うという表現に対し中国人に答えてもらったアンケートでは、「数字を使うことで理解しやすかった」といった有効性を示すコメントが得られた。

● 実験 12

以下に実験 12 のチャットログの一例を示す。実験 12 は日本人の知識提供者に対し、知識享受者を中国人 2 名に日本人 1 名を加えたパターンである。

表 4: 実験 12 の知識伝達例

原語スレッド	翻訳スレッド
工程 1 精米 玄米の外側を精米機で削ります	精米機で除去する工程 1 精米 玄米外側
酒の味を損ねる微生物や、脂肪、たんぱく質を除去します	酒の味を害する微生物、脂肪、タンパク質を除去する
工程 2 米を洗う	洗う 工程 2 米
工程 3 浸漬 米を水に浸して吸水させます	工程 3 の米を水の中に浸し吸水させます
工程 4 米を蒸す 「こしき」という機械で 1 時間ほど蒸します。	工程 4 米を蒸す「腰米」の機械を用いて一時間熱する。
蒸しあがった米は、このあといくつかの工程に分けて使用されます。	蒸した米は、この後の工程で使われます。
工程 5 麴作り 蒸しあがった米を冷ましてから温度が一定の麴室に入れます	工程 5 麴政策蒸した米を冷却した後、温度が一定の麴室に入れます
麴室では、蒸した米と粉状の麴菌を混ぜて 2 日間放置します。	麴室は粉の麴菌と蒸した米を一緒に混ぜて 2 日間放置します。
2 日後に、麴カビが成長したら、麴の完成です。	2 日後、もしも麴カビが増えていたら、麴の完成です。
工程 6 酒母作り 蒸した米からもう 1 つ作られるのが、酒母です。これは酵母菌を大量に増殖させたものです。	工程 6 酒母制作蒸した米が更に作るのとは一つ、酒母。これは酵母菌を増殖させる。
酒母の材料は、蒸した米、麴、乳酸、水の 4 つです。	酒母の材料は蒸した米、麴、乳酸、水の 4 つ。
4 つを混ぜ合わせて、酵母菌が増殖したら完成です。	もしも 4 つを混ぜたら、酵母菌は増殖完成です。
乳酸を入れるのは、雑菌の繁殖を防ぐためです。	入れるのは乳酸、雑菌の繁殖を抑えるために。
工程 7 任込み 酒母、麴、蒸し米、水をタンクに入れます。4 日間で 3 回に分けて入れます。	工程 7 訓練 酒母、麴、蒸した米、水をタンクに入れます。4 日 3 回に分けて入れます。
このとき麴が、米のデンプンを糖に変えます。	この時麴が、米のデンプンが糖に変わります。

知識伝達の順序としては、実験 8 と同様に初めから 1 工程ずつ順に詳細な説明を行っている。知識伝達文の投稿方法として、こちらは順番を表すために英数字を付けた工程の紹介と詳細な説明をスペースを間に入れて同時に行っている。しかし、翻訳結果を見てみると、スペースは認識されず工程名と説明が詰まった状態になる発話がいくつか見られた。翻訳スレッドを見ると、ある程度原文と同じように意味が通っているものもあるが、中には誤訳を引き起こしている発話も数個見られた。従って、工程の名称と説明を一度に行う場合は、“=”などの記号を扱うほうが良いと考えられる。

● グループワークでの会話方法

実験 8 と実験 12 は知識享受者のグループを中国人 2 名と日本人 1 名とした実験である。前述したように、知識提供者と母国語が異なる中国人 2 名に同じ母国語である日本人を加えると、日本人が中国人に足りない情報を共有することで中国人のみの場合よりも学習度の値（グループでの学習範囲）が大きくなるという知見が得られている。そこで、同パターンで行われた各実験のグループワークのチャットログを比較してみると、実験 8 と実験 12 のいずれにおいても日本人と中国人が双方で英数字を用いながら工程に番号を付け会話をしていた。日本人にとっては中国人に知識が不足していると思える工程の追加説明を行う際に、中国人にとってはあまり学び取れなかった工程について日本人に質問する際に英数字を用いることで、互いが相手とどの工程について会話をしているのか知ることができると考えられ、それが知識伝達の確実性のさらなる向上に有効だったと考えられる。

このような動きは実験 5 のグループワークでも見られ、このパターンの中で最も学習度の低い実験 3 のグループワークでは、3 番目の工程の議論まで数字を用いていたが、その後は用いられなくなり、中盤の工程で一度だけ用いられただけであった。

5.2. 各実験事例のまとめ

以上をまとめると、知識伝達の流れとしては 2 種類の伝達順序が学習度の値の高い実験にあることがわかり、共通することとして、どの発話もシンプルな文章であり、説明が長くなる際は投稿を複数回に分けていた。また、順接を表すために“次に”や“そして”といった接続語ではなく、英数字や“→”といった記号を代わりに用い、“～は”や“～とは”といった表現をする際には、“=”の記号を用いたり、説明内容を次の発言に持ち越して投稿することなどを行っていた。

山下らの研究 [5] で用いられている、良い翻訳結果を得るための書き換えルールでは、文章に基づくルールの 1 つとして「長い文を避け、短い文にする」ことを挙げており、同時に「接続助詞を使用した長い文は接続助詞の前後で分割する」というルールを挙げています。本研究ではルールの教示は一切行っていないが、以上を自然に実践したとみられる実験として、実験 8 や実験 12 の他に実験 4 も挙げられ、今回専門知識を伝達知

識として扱う中で、実験 8, 12 で見られる数字を扱う方法や、実験 11 の様に記号を多用する方法も見られた。

ここで、学習度の値が低い実験として実験 2 と実験 7 の知識伝達過程時のチャットログを見てみると、どちらも上記の表現をするために接続詞や助詞を用いており、英数字や記号は一部で使用する以外にほぼ用いていなかった。よって、一般的な言語の表現で知識伝達を行う方法は、前述した、実験で被験者が感じる誤訳の割合が高いという知見から考えても好ましい方法ではないと考えられる。

5.3. 伝達方法レシピ

今回、学習度の高い実験から数字や記号を扱う例が多く見られた。数字や記号が効果的であった理由として、文字自体は翻訳の影響をほぼ受けないことや、国際的にそれらの持つ意味がある程度共有されていることが考えられる。本研究では日本酒の製造工程という、しっかりと決まった作業や工程の順序が存在する知識情報を伝達する専門知識として扱っている。このような知識情報を扱う際に、数字や記号を用いて説明する方法は効果的であると考えられる。

以上のような行動は、人間の側からの機械翻訳への適応を表す行動とみなすことができる。ここで、これまでに述べた効果的だと考えられる知識伝達方法をまとめ、本研究で得られた知見に基いた機械翻訳を用いる際の効果的な伝達を実現する知識伝達レシピを考案した。以下、前節で述べた 2 種類の知識伝達順序で示す。

表 5: 効果的な知識伝達レシピ（流れを紹介後、詳細説明する場合）

手順 1	”「工程名」と”→”を用い流れを紹介。投稿は長ければ何回かに分けても良い。
(例)	「精米」→「洗米」→「浸漬」→「蒸米」→…
手順 2	”=”を用い紹介した工程を 1 つずつ説明
(例)	洗米=米を洗い、削った粃殻を除去する。
注意 1	なるべくシンプルな文章になるよう心掛け、投稿は 1 回あたり 1 文とする。
注意 2	従って、説明が長くなる際は必要に応じ投稿を複数回行う。

表 6: 効果的な知識伝達レシピ（初めから 1 工程ずつ順に説明する場合）

手順 1	”英数字”を用い工程に番号を付け、”=”を用い 1 工程ずつ順に説明。
(例)	2 洗米=米を洗い、削った粃殻を除去する。 3 浸漬=米を水に浸し、吸水させる 4 蒸米=米を「こしき」と呼ばれる蒸釜に入れ、1 時間ほど蒸す。
注意 1	なるべくシンプルな文章になるよう心掛け、投稿は 1 回あたり 1 文とする。
注意 2	従って、説明が長くなる際は必要に応じ投稿を複数回行う。

以上のレシピを実行する前段階として、事前に伝達する情報量が分かっている場合はそれを伝達すること

を提案する。また、表の注意書きにもあるように、文章を短く分割する方が翻訳精度の影響を受け難いと考えられる。従って、“=”の後ろで投稿を区切ることや、単語を適宜”|”で囲んで表示させるなど、翻訳リペアが行える際はその結果に応じて対応することがより精度の高い翻訳結果を得るために好ましいと考えられる。

今回レシピとして提案した機械翻訳サービスの利用方法は、実験全体を通じた知識伝達の有効性を表す基準として、独自に定義した指標の値が高い実験のシステム使用履歴から法則的に抽出したものである。これに対し、指標の値が低い実験からは、これに近い利用方法を法則的に抽出することはできなかった。日中翻訳における誤訳の感じ易さを考慮すると、抽出した利用方法が常に有効に作用するとは限らない。しかし、レシピの内容を法則的に実施した教示を行うことで、一部で誤訳が発生した場合でも知識伝達行動全体としては知識享受者にとってより適応し易くなると考えられる。

6. おわりに

本研究では、多言語環境を前提に実用的な場面を想定し、機械翻訳サービスを利用した知識伝達をどのように行うことが正確かつ効果的かを、利用者側の適応から明らかにすることを目的として取り組んだ。

結果の分析から、知識提供者と母国語の異なる中国人のみの集団パターンにおいて、基礎的な知識に関しては正答率を8割まで延ばす情報共有を行えるほどの知識を知識共有過程で集団的に享受できていることがわかった。また、知識提供者が教える知識範囲全体に対するグループ毎の理解度を評価・分析した結果、知識提供者と同言語の知識享受者を1名でも加えることで、機械翻訳による悪影響がある中でもグループとして全体的な学習範囲を向上させられるという知見が得られた。このように、知識享受者側における組織学習の枠組みを組合せる場合、集団パターンの特性を考慮した知識伝達・共有環境を考慮することが機械翻訳サービスによる知識伝達の正確性の最大化に役立つことを明らかにした。更に、これらの評価結果を会話履歴と共に分析し、相互に関連づけられた複雑な知識や専門性の高い知識を共有するために効果的な方法として、機械翻訳を用いる際の効果的であると考えられる伝達方法をレシピとして考案するに至った。

今後の課題としては、次の事項が挙げられる。今回は実験設定として、知識提供者に実験の場で始めて知識を取得させており、今後は提供者に事前知識があるケースとの比較が必要である。また、日本語と中国語による多言語環境で実験を行ったのみであり、扱う言語によって翻訳精度が異なることから、今回得られた専門知識の伝達に関する知見が他の言語を扱った場合にも得られるか検証することが必要である。さらに、様々な品質を備える機械翻訳サービスを試し、レシピを実践する知識伝達実験の蓄積を通じてレシピの効果を検証する。この検証において新たな知見が得られることで、より機械翻訳サービスの利用者適応として有効なレシピへと改善をしていくことが必要である。

謝辞

本研究は、JSPS 科研費(S)(24220002, 2012-2016)の助成を受けたものです。

参考文献

- [1] NPO Pangaea : YMC Viet, 入手先 <http://www.pangaeaan.org/project/ymcprj/?page_id=37> (参照日 2015-03-10).
- [2] 独立行政法人情報通信研究機構: 言語グリッドプロジェクトポータルサイト, 入手先 <<http://langrid.org/jp/>> (参照 2015-03-10).
- [3] Kita, K., Takasaki, T., Lin, D., Nakajima, Y. and Ishida, T.: Case Study on Analyzing Multi-Language Knowledge Communication, *2012 International Conference on Culture and Computing (Culture and Computing 2012)* (2012).
- [4] Otani, M., Kita, K., Lin, D., and Ishida, T.: Analysis of Multi-language Knowledge Communication Service in Intercultural Agricultural Support. *The 2nd International Conference on Serviceology (IC-Serv2014)* (2014).
- [5] 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: 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).
- [6] 山下直美, 石田亨: 翻訳機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, No.2, pp.939-948 (2007).
- [7] 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, Vol.J92-D, No.6, pp.747-757 (2009).
- [8] Arendse, B. and Claudia, G.: MTranslatability, *Machine Translation*, Vol.16, No.3, pp.175-218 (2001).
- [9] Szulanski, G.: The Process of Knowledge Transfer: A Diachronic Analysis of Stickiness, *Organizational Behavior and Human Decision Processes*, Vol.82, No.1, pp.9-27 (2000).
- [10] 多言語工房-Language Grid: 入手先 <<http://langrid.org/developer/jp/index.html>> (参照日 2015-03-10).
- [11] 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, Vol. J90-D, No.12, pp.3141-3150 (2007)
- [12] 角田啓介, 菱山玲子: 異文化コラボレーションのための多言語参加型ゲーミングシステム”Langrid Gaming”の提案, ヒューマンインタフェース学会論文誌, Vol.13, No.1, pp.73-82 (2011).