

# 多言語自動翻訳掲示板の利活用に関する基礎的 研究

早稲田大学人間科学部人間情報科学科  
西村研究室

学籍番号：1J20F037

名前：奥村飛悠

2023 年 7 月 17 日

# 目次

第 1 章	はじめに	2
1.1	現状 . . . . .	2
1.2	翻訳技術の進歩 . . . . .	3
1.3	既存サービスと先行研究 . . . . .	3
1.4	本研究の目的と方法 . . . . .	4
参考文献		5

# 第 1 章

## はじめに

### 1.1 現状

インターネットの普及やソーシャルメディアの台頭により、オンラインでのコミュニケーションが一般的になっている（大向, 2006）。その中でも、誰でも気軽に参加することのできる掲示板は重要なコミュニケーションの場となっている。日本では、「5ちゃんねる（サイト名を2ちゃんねるから5ちゃんねるへ2017年10月に変更）」が広く知られている一方で、米国発の「reddit」は国際的に認知度が高く、日別のアクティブユーザー数は5700万人、総投稿数は130億投稿を超えている（Reddit Inc, 2023）。これらの大規模な掲示板は情報の集積場所として、またユーザー間の活発な議論の場として重要な役割を果たしている。

しかしながら、掲示板は誰もが利用できるコミュニケーションの場であるにも関わらず、現状ではそのコミュニケーションは主に同一言語間で行われている。具体的には、「5ちゃんねる」では主に日本語、「reddit」では主に英語が使用されている。そのため、異なる言語を使用するユーザーは、翻訳ツールや外部の翻訳サービスを頼るか、専用のスレッドや言語コミュニティを探すことが一般的となっている。しかし、これらの方法にはデメリットが存在する。例えば、翻訳ツールを利用すると時間と手間がかかるため、掲示板の持つ即時性や気軽さというメリットを十分に享受することが難しくなる。

同一言語間でのコミュニケーションが主流となっている背後には複数の要因が考えられるが、その一つとして翻訳技術の品質が十分でなかったことが考えられる。2004年には「コミュニケーションツールとして使用する場合に十分な品質を持っているとはいえない」（船越ほか, 2004）との指摘があり、さらに2009年にも「近年、翻訳技術は急速に進展しているが、高精度な翻訳を行うことは困難である。コミュニケーションにおいて、不適切な翻訳箇所を含む文章は話者間の相互理解を困難にし、円滑のコミュニケーションの妨げとなる」（宮部ほか, 2009）とも指摘されている。また、多言語間でのコミュニケーションにおいては、翻訳の品質が極めて大きな影響力を持つことも確認されている（船越ほか, 2004）。このように、コミュニケーションに大きな影響を

与える翻訳技術の品質が十分でなかったため、ユーザーは翻訳技術を利用して会話の中心となっている言語以外を使用してまで会話を試みなかったのではないだろうか。

## 1.2 翻訳技術の進歩

しかしながら、翻訳サービスの精度は日々向上している。これについて、「近年、Google 翻訳や DeepL、そしてページ全体翻訳機能の進化が著しい」(村本, 2022) との報告があり、その背景には機械学習の進歩が影響を与えている。「Google 英日翻訳が NMT (ニューラル機械翻訳) を採用したことで、目標言語の流暢さが格段に向上した」(影浦, 2017) との報告がある。同様に NMT を採用している DeepL は、2017 年にサービスを開始し、その高品質な翻訳サービスが評価されている(亀田, 2022)。さらに「2020 年と 2021 年には、文章の意味をより正確に伝えられ、業界特有の専門用語もうまく処理できる新たなモデルを発表」(DeepL, 2023) している。これらのことから、翻訳サービスの精度は日々向上されていることが分かる。

また、多くの翻訳サービスが開発者向けに API を提供している。その代表例としては、Google Cloud の Translation AI (Google Cloud, 2023) や DeepL API (DeepL, 2023) がある。これらのサービスを開発者が利用するための便利なライブラリも存在している。具体的には、Google Translate API (現在の Translation AI) を利用するための Python のライブラリである googletrans (PyPI, 2023a) や deepl (PyPI, 2023b) がある。このような API やライブラリの存在により、開発者は翻訳機能を自身のサービスに容易に組み込むことが可能となっている。

## 1.3 既存サービスと先行研究

過去には、「enjoy Korea」という日本語と韓国語の翻訳機能を持つ掲示板サービスが存在していたが、利用率の低下を理由に 2009 年 6 月 8 日にサービスを終了している(野津, 2009)。また、小川ら(2009)は日本語とウイグル語間の翻訳掲示板システムを開発している。しかし、彼らの研究は主にシステムの開発に焦点を当てており、システムを使用するユーザーのデータ収集やその分析までには至っていない。

一方、藤井ら(2005)はアノテーションや折り返し翻訳に着目し、中国語、韓国語、日本語間の翻訳 BBS である「AnnoChat」を開発した。翻訳の精度がコミュニケーションの理解度に影響を与える可能性を示しているが、ユーザー同士の具体的なコミュニケーションの内容までは調査していない。また、吉野ら(2006)はユーザインタフェースのカスタマイズ性に焦点を当てた研究を行い、「CustomChat」というシステムを開発したが、これも具体的なチャットの内容などについては触れられていない。

これらの事例や研究を見ると、多言語間のコミュニケーションを可能にする翻訳掲示板に関する研究やサービスは確かに存在している。しかし、それらは主にシステムの開

発や翻訳の精度と理解度の関係性、ユーザインタフェースの改良に焦点を当てており、異なる言語を使用するユーザーがシステムをどのように使うのか、どのようなコミュニケーションが起こるのかという点については、まだ十分に研究されていないと言える。

## 1.4 本研究の目的と方法

これらの背景から、本研究では、掲示板のグローバル化を進めるため、近年の高精度な翻訳サービスを利用した多言語自動翻訳掲示板の開発とその利活用について基礎的な研究を行う。我々が提案する多言語自動翻訳掲示板では、ユーザーは表示言語を選択することにより、選択した言語で掲示板の投稿を閲覧することを可能にする機能をつける。これにより、異なる言語を使用するユーザー間でも、自由なコミュニケーションが促進され、掲示板の持つ即時性や気軽さというメリットを維持することができる。

そして、この掲示板をインターネット上に公開し、使用者から得られるデータを収集する。その後、得られたデータを分析し、多言語自動翻訳掲示板がユーザーのコミュニケーションにどのような影響を与えるのか、多言語自動翻訳掲示板上で異なる言語を使用するユーザー同士がどのようなコミュニケーションをするのかを評価する。具体的には、ユーザー間のコミュニケーション量や内容、トピックの多様性、言語間のコミュニケーション方法などを指標として用いる。

我々の研究は、新たな掲示板の形を示すだけでなく、機械翻訳技術とその実用化の進歩に貢献することを期待している。本研究の結果が、ユーザーが自由に多言語コミュニケーションを享受できるインターネットの環境整備に向けた一歩となることを願っている。

# 参考文献

- [1] 大向一輝 (2006). SNS の現在と展望-コミュニケーションツールから情報流通の基盤へ-. 情報処理, 47(9): 993-1000.
- [2] 小川泰弘, 福田ムフタル, 外山勝彦 (2009). 日本語ーウイグル語翻訳掲示板システム. 言語処理学会 第 15 回年次大会発表論文集, 15: 212-215.
- [3] 影浦峽 (2017). 改めて、翻訳とは何か：Google NMT が使える時代に. 言語処理学会 第 23 回年次大会発表論文集, 23: 931-934.
- [4] 亀田倫史 (2022). 機械学習とバイオテクノロジー. 生物工学会誌, 100(11): 588.
- [5] 中澤敏明 (2017). 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, 60(5): 299-306.
- [6] 野津 誠 (2009). 日韓翻訳掲示板「enjoy Korea」終了へ、理由は利用率の低下. 株式会社インプレス, <https://internet.watch.impress.co.jp/cda/news/2009/02/12/22405.html#:~:text=会員数は非公開,にした>と説明する。(参照日 2023.07.17)
- [7] 藤井薫和, 重信智宏, 吉野孝 (2005). 異文化間コミュニケーションのための機械翻訳を用いたチャットシステム AnnoChat の開発と適用. 情報科学技術フォーラム一般講演論文集, 4(3): 437-438.
- [8] 船越要, 藤代祥之, 野村早恵子, 石田料亨 (2004). 機械翻訳を用いた協調作業支援ツールへの要求条件ー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, 45(1): 112-120.
- [9] 宮部真衣, 吉野孝 (2009). 折り返し翻訳を用いた翻訳リペアのチャットコミュニケーションへの影響.
- [10] 村本麻衣 (2022). 自動翻訳機能からの自立：学習者による気づきを通じて. ドイツ語教育 = Deutschunterricht in Japan / 日本独文学会ドイツ語教育部会 編, 26: 119-125.
- [11] 吉野孝, 藤井薫和, 重信智宏 (2006). 異文化間コミュニケーションのためのカスタマイズ可能なユーザインタフェースを持つチャットシステム CustomChat の開発. 情報処理学会研究報告 = IPSJ SIG technical reports, (60): 13-18.
- [12] DeepL (2023). DeepL. <https://jobs.deepl.com/> (参照日 2023.07.08)
- [13] Google Cloud (2023). Translation AI. <https://cloud.google.com/translate?hl=ja>

(参照日 2023.07.17)

- [14] Loki Technology, Inc (2023). 5ちゃんねる. <https://5ch.net/> (参照日 2023.07.17)
- [15] PyPI (2023a). googletrans 3.0.0. <https://pypi.org/project/googletrans/> (参照日 2023.07.17)
- [16] PyPI (2023b). deepl 1.15.0. <https://pypi.org/project/deepl/> (参照日 2023.07.17)
- [17] Reddit Inc (2023). reddit. <https://www.redditinc.com/> (参照日 2023.07.08)