

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 3 日現在

機関番号：14701  
 研究種目：基盤研究（B）  
 研究期間：2010～2012  
 課題番号：22300044  
 研究課題名（和文） 多言語間コミュニケーション高信頼化のためのデータセントリック価値創出に関する研究  
 研究課題名（英文） Study on data centric value creation for high-reliability multilingual communication  
 研究代表者  
 吉野 孝（YOSHINO TAKASHI）  
 和歌山大学・システム工学部・准教授  
 研究者番号：90274860

## 研究成果の概要（和文）：

多言語間コミュニケーションの方法として、機械翻訳を用いる方法と用例対訳を用いる方法がある。本研究では、多言語間コミュニケーションの高信頼化のために、折り返し翻訳の高信頼化に関する研究、多言語用例対訳の高信頼化に関する研究を行った。さらに、上記の研究成果を応用した、医療分野および教育分野を対象とした多言語支援ツールの構築に関する研究を行った。

## 研究成果の概要（英文）：

There are two methods for multilingual communication: using machine translation and using parallel texts. We studied the high-reliability of Back translation and parallel texts for high-reliability multilingual communication. We also developed multilingual support software for medical field and education field.

## 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010 年度	4,900,000	1,470,000	6,370,000
2011 年度	4,400,000	1,320,000	5,720,000
2012 年度	4,700,000	1,410,000	6,110,000
年度			
年度			
総計	14,000,000	4,200,000	18,200,000

研究分野：総合領域

科研費の分科・細目：メディア情報学・データベース

キーワード：多言語，機械翻訳，コミュニケーション

## 1. 研究開始当初の背景

研究代表者らは、HCI (Human Computer Interaction) の技術をもとに、機械翻訳技術を組み込んだシステムを構築し、機械翻訳の多言語間コミュニケーション支援への効果を確認してきた。これまでの研究成果から、翻訳精度の確認や修正作業の促進のために「折り返し翻訳」を提示し、利用者が「翻訳リペア作業」(翻訳結果が正しくなるように原文を修正する作業)を行う

ことにより、翻訳結果の精度を向上させることが可能であることを明らかにした。

しかし、同時に、「翻訳リペア」には非常に時間がかかり、その効率や成功率は個人差が大きいことも明らかとなった。これらの問題は、機械翻訳の能力を十分に引き出せず、機械翻訳を介したコミュニケーションを阻害する。また、研究代表者は、入力したチャットの文に対して、注釈(アノテーション)を付与する仕組みを構築した。こ

の仕組みを用いることにより、対話者の理解向上の効果は確認できたが、チャット中に注釈を付与する作業の負担は非常に大きく、実利用に大きな困難があった。

研究代表者らは、機械翻訳を用いた多言語間コミュニケーションツールを、国際的に活動している NPO へ提供した。実際のコミュニケーションを観察すると、機械翻訳の精度は不十分であるが、対話そのものは継続できることがわかった。伝わっていないことに気付いた場合、補足説明や言い換えにより対話を継続していた。しかし実際には、相手に伝わったと思っているだけであり、誤解が発生していた。同言語間のコミュニケーションと異なり、機械翻訳を用いた多言語間コミュニケーションの問題は、「異なる言語を用いている相手にどの程度正しく伝わるかが分からない」ことである。

## 2. 研究の目的

現在、Web 上には、大量の多言語のテキストデータ、画像データ、人手により整理された多言語百科事典データ等が存在している。研究代表者らは、これらのデータに、人を介在させ、知識として蓄積し、適切に利用することで、高い信頼性をもつ「多言語間コミュニケーション」を実現できると考えた。

本申請の「データセントリック価値創出」とは、人の能力を適切に介在させることにより、多言語間コミュニケーションの高信頼化に利用可能な、新しい価値を創出することを意味する。本申請の目的は、多言語間コミュニケーションの高信頼化の手法を提案し、その実現方法と効果を、実証を通じて明らかにすることである。

## 3. 研究の方法

本研究は、人を介したシステムであるため、提案手法の評価のために、実際に人に利用してもらい評価を行い、提案手法の改善を行うという方法で、研究を行う。

## 4. 研究成果

研究成果として、以下の知見を得た。

(1) 機械翻訳を介したコミュニケーションでは、翻訳精度が低い場合、十分な相互理解ができない可能性が高い。現在、母語のみを用いて自分の発言がどのように伝わっているのかを把握するための手法として、折り返し翻訳が用いられている。対象言語翻訳文と折り返し翻訳文の精度の同等性に関する検証の結果、対象言語翻訳文が不正確であるにもかかわらず、折り返し翻訳文が正確であるという状況(精度不一致)が発生する場合があることが明らかになっている。このような精度不一致が発生した場合、ユーザの確認する折り返し翻訳文には問題がないため、対象言語翻訳文が低精度であることに気づくことができず、大きな問題となる。そこで本研究では、複数翻訳機を利用した精度不一致の検出

手法について検討を行う。検証の結果、提案手法とあわせて対象言語翻訳文中に原言語表現が残っているかどうかを検証する処理を行うことにより、約 71%の精度不一致を検出できることを示した。

(2) 機械翻訳を利用した多言語間コミュニケーションにおいて不正確な翻訳文を用いると、意思疎通が困難となる。しかし、ユーザは翻訳精度を正しく判断できない場合があるため、翻訳精度に関する気付きを与える必要がある。本研究では、他者やシステムによって判定された精度を提示した場合、ユーザ自身の翻訳精度判定にどのような影響があるかを検証する。本研究の貢献は次の点にまとめられる。(1) 誤った精度を既判定精度として提示した場合でも、既判定精度の影響を受ける場合があることを示した。また、正しい精度(本来の精度)を提示した場合、提示した精度との一致率が最も高くなることを示した。(2) 本来の精度と隣接した評価値を既判定精度として提示した場合、本来よりも高い隣接評価値との一致率の方が、本来よりも低い隣接評価値との一致率よりも高い傾向があることを示した。

(3) 機械翻訳を介したコミュニケーションでは、翻訳精度が低い場合、十分な相互理解ができない可能性が高い。現在、母語のみを用いて自分の発言がどのように伝わっているのかを把握するための手法として、折り返し翻訳が用いられている。折り返し翻訳を精度確認手法として用いる場合、対象言語翻訳と折り返し翻訳の精度に正の相関関係があり、対象言語翻訳と折り返し翻訳の精度が大きく異なることが求められる。これまでに、折り返し翻訳の精度確認手法としての妥当性の検証を行った結果、折り返し翻訳文と対象言語翻訳文の精度には正の相関がみられた。また、対象言語翻訳文が不正確であるにもかかわらず、折り返し翻訳文が正確であるという状況(第 1 種の精度不一致)の発生率は低いことがわかった。しかし、発生数は少ないものの、第 1 種の精度不一致が発生する場合があることがわかった。第 1 種の精度不一致は、コミュニケーションにおいて意思疎通を困難にする可能性が高く、不一致が発生した際の対応が必要となる。そこで本論文では、第 1 種の精度不一致が発生し得る入力を行った場合に対応できるようにするために、第 1 種の精度不一致が発生した文についての分析を行った。分析の結果、第 1 種の精度不一致の発生要因として、5 種類の要因があることを示した。

(4) 世界的なグローバル化を背景に、我々は多言語間コミュニケーション支援を目的とした多言語用例対訳共有システムの開発を行っている。共有対象の用例対訳は、各言語間で一対一に対応する必要があるが、用例対訳には各言語間

で一対一に対応しない言葉の組み合わせが存在している。このような多言語間の言葉の多様性への対応が求められているが、単純な用例間の意味のつながり情報のみでは対応することができていなかった。そこで、本論文では、一対多、多対多の関係にある用例対訳グラフを提供可能な用例対訳とするために、メタノード作成手法を提案する。また、メタノード作成手法適用により顕在化したリンク不足の問題を解決するために、新たな用例間リンクを発見する方法についても提案する。

(5) 現在、グローバル化による多言語間コミュニケーションの機会が増加している。しかし、多言語間での正確な情報共有は十分に行われていない。この問題は、正確性が求められる医療分野や災害時において顕著に現れ、解決が求められている。このため、正確な情報共有を可能にする一技術である用例対訳を用いた支援が行われており、用例対訳の作成も多く行われている。しかし、従来の用例対訳作成においては、(1) 新たな用例対訳の追加が難しい、(2) 翻訳者 1 人あたりの負担が大きい、(3) 用例対訳の利用現場で求められている用例を収集することが難しい、という問題が存在していた。そこで本論文では、用例対訳の作成の基盤となる場を用例対訳作成環境に提供し、円滑に正確な用例対訳の収集を行うことを目的とした、多言語用例対訳共有システムを提案し、実装を行った。また、実際のシステム構築後に顕在化した問題点とその解決策について述べる。

(6) 現在、在日外国人数や訪日外国人数は増加傾向にあり、多言語によるコミュニケーションの機会が増加している。多言語環境支援の一方法として、用例を正確に多言語に翻訳した用例対訳が用いられている。用例対訳は正確な多言語間コミュニケーション支援が可能のため、医療分野などの正確性が求められる分野で多く利用されており、動的な用例対訳収集も行われている。しかし、これまで用例対訳を含めた用例収集の取り組みにおいて用例の正確性評価が行われてこなかった。用例の正確性評価がなされていない用例は、医療などの正確性を求められる分野で使用することはできない。そこで我々は、用例の正確性評価手法の確立を目指して複数の評価手法の比較評価を行った。本論文では、比較実験の結果から次の知見を得た。(1) 用例の正確性評価実験の結果、評価軸を提示していない評価手法と提示した評価手法の間に相関関係がみられなかった。用例の正確性評価には、評価軸を明確に提示した評価手法が必要である。(2) 用例の評価において、評価者は詳細な評価を行うことを好んだ。このため、2 値の評価段階よりも複数の評価段階を選択可能とした評価手法が評価者から支持された。

(7) 現在、在日外国人数は年々増加しており、多言語によるコミュニケーションの機会が増加している。医療の分野では医療機関を訪れる外国人患者とのコミュニケーションのために多言語問診票が使用されている。しかし、多言語対応の紙の問診票は種類が少ないため各医療機関の要求を満たすことができていない。また、日本人医療従事者が外国人患者の母語で書かれた紙の問診票を理解することは困難である。そこで我々は、用例対訳や機械翻訳を利用して診察に必要な基本情報である症状の伝達を支援する、多言語問診票作成システムの開発を行い、問診票入力機能についての評価を行った。本研究の貢献は次の 2 つである。(1) 日本人医療従事者と外国人患者との間のコミュニケーションを支援する、多言語問診票作成システムの提案を行い、実現した。(2) 用例対訳と機械翻訳を併用することで、多言語問診票の記入ができることを示した。

(8) 近年、在日外国人が増加しており、外国人と対面でコミュニケーションを行う機会が増加している。しかし、異文化間コミュニケーションにおいては、それぞれの人間が持つ文化的な背景が異なるため、相互理解が困難である。そこで本研究では、文化的な背景がもたらす知識の違いを補うために、会話中の名詞の関連情報を用いた対面型異文化間コミュニケーション支援システム iGengo を開発した。対面環境における異文化間コミュニケーションにおいて、iGengo の評価実験を行い、次の知見を得た。(1) 会話中に発言された名詞に関する画像、母語の情報、説明の文章、関連名詞、関連画像をユーザに提示することは対話を効果的に支援する可能性がある。(2) 発言された名詞に関連している名詞を表示する機能は、話題に関する情報をユーザに与え、会話の話題を広げられる可能性がある。(3) 発言された名詞の画像を複数同時に表示する機能は、多方面、多角度から話題を理解することを支援できる。(4) タッチパネルディスプレイを用いたクリック操作インタフェースと音声入力は会話をしながら手軽に操作を行える。

(9) 日本国内の大学などで行われる講義や会議において、会議に参加している複数の日本人が少しずつ助力することにより外国人の内容理解を支援する、All for one 型の多言語対面会議支援システムが提案されている。しかし、現在の課題として、All for one 型対面会議支援システムにおいて発生する作業競合などの問題への対応がなされておらず、また All for one 型の支援を行う際の、支援者の内容理解への影響に関する議論がなされていないことが挙げられる。そこで、本研究では、All for one 型対面会議支援システムにおける入力内容や移動作業のような作業競合の発生を減少させるために、ワークスペースアウェアネス機能を適用し、発表要約実験により、

その効果について検証を行った。実験の結果、以下の知見を得た。(1) リアルタイム入力状況提示を行うことにより、従来困難であった入力内容の競合の軽減に貢献できる。(2) テレポインタ機能および色による支援者識別機能により、移動作業の競合防止に貢献できる。(3) ワークスペースアウェアネスによる内容理解度への影響はないが、All for one 型支援による要約作業は支援者への負荷が大きく、支援者自身の内容理解に影響を及ぼす。

(10)日本国内の大学などで行われる講義や会議において、外国人留学生の内容理解には限界がある。外国人の内容理解を支援するために、会議に参加している複数の支援者が少しずつ助力することによって支援を行う、All for one 型の多言語対面会議支援システムが提案されている。All for one 型支援では、本来 1 名の支援者のみにかかる作業負荷を軽減できる可能性がある。しかし、これまでに作業を分担して支援を行う All for one 型支援における、支援者側の効果についての検証は行われていない。そこで、本論文では、発表要約実験を行い、単一ユーザによる支援と 4 名の支援者による All for one 型の支援との比較を行った。実験の結果、以下の知見を得た。(1) 4 名の支援者による All for one 型支援においては、要約の平均総文字数が単一ユーザによる支援の約 2 倍となっていた。また、All for one 型支援におけるスライドカバー率は単一ユーザによる支援よりも有意に高く、All for one 型支援は、単一ユーザによる支援よりも多くの情報を入力できていた。(2) All for one 型支援における支援者個人の操作数は単一ユーザの操作数よりも有意に少なく、4 名の支援者により作業を分担することにより、作業量を軽減できていた。(3) All for one 型支援により作業負担は軽減できたものの、支援者の内容理解度に悪影響を及ぼす可能性がある。

## 5. 主な発表論文等

[雑誌論文] (計 10 件)

[1] 福島 拓, 吉野 孝, 重野亜久里: 用例対訳と機械翻訳を併用した多言語問診票入力手法の提案と評価, 情報処理学会論文誌, 査読有, Vol.54, No.1, pp.256-265 (2013).

[2] 宮部真衣, 吉野 孝: 折り返し翻訳を用いた高精度なコミュニケーションのための複数翻訳機による精度不一致検出サービスの提案, 情報処理学会論文誌, 査読有, Vol.53, No.8, pp.2036-2043 (2012).

[3] 福島 拓, 吉野 孝, 重野亜久里: 正確な情報共有のための多言語用例対訳共有システム, 情報処理学会論文誌 コンシューマ・デバイス & システム (CDS4), 査読有, Vol.2 No.3, pp.23-33 (2012).

[4] 宮部真衣, 吉野 孝: 折り返し翻訳文と対象言語翻訳文の精度不一致要因, 電子情報通信学会. 電子情報通信学会論文誌 D, 査読有, Vol.95-D, No.1, pp.11-18 (2012).

[5] 宮部真衣, 吉野 孝: ユーザの翻訳精度判定に対する既判定精度提示の影響, 情報処理学会論文誌, 査読有, Vol.53, No.1, pp.39-48 (2012).

[6] 福島 拓, 吉野 孝: 多対多の関係性を持つ多言語用例対訳グラフにおけるメタノード作成手法, 情報処理学会論文誌データベース, 査読有, Vol.5, No.2, pp.1-10 (2012).

[7] 宮部真衣, 吉野 孝: All for one 型対面会議支援システムのためのワークスペースアウェアネスの効果, 電子情報通信学会論文誌 D, 査読有, Vol.94, No. 2, pp.27-36 (2011).

[8] 宮部真衣, 吉野 孝: 多言語対面会議支援システムのための All for one 型支援の効果, 情報処理学会論文誌, 査読有, Vol.52, No.2, pp.90-96 (2011).

[9] 福島 拓, 吉野 孝: 用例の正確性評価を目的とした用例評価手法の比較, 情報処理学会論文誌, 査読有, Vol.52, No.2, pp.131-139 (2011).

[10] 岡本健吾, 吉野 孝: 会話中の名詞の関連情報を用いた対面型異文化間コミュニケーション支援システムの構築と評価, 情報処理学会論文誌, 査読有, Vol. 52, No. 2, pp.1213-1223 (2011).

[学会発表] (計 52 件)

[1] Taku Fukushima and Takashi Yoshino: Automatic Translation Repair Method for Improving Accuracy of Translated Sentences, IEEE Seventh International Conference on Knowledge, Information and Creativity Support Systems, 2012 年 11 月 10 日, メルボルン, オーストラリア

[2] Taku Fukushima and Takashi Yoshino: Metanode Composition Method for Multilingual Parallel-text Having Many-to-many Relationship, KES 2012, 2012 年 9 月 11 日, サンセバスチャン, スペイン

[3] Mai Miyabe and Takashi Yoshino: Features of Accuracy Mismatch between Back-translated Sentences and Target-translated sentences, IEEE 2011 Second International Conference on

Culture and Computing, 京都府京都市, 2011年10月21日

[4] Mai Miyabe and Takashi Yoshino: Can Indicating Translation Accuracy Encourage People to Rectify Inaccurate Translations?, HCI 2011, 2011年7月12日, オランダ市, アメリカ合衆国

[5] Takashi Yoshino, Katsuya Ikenobu, Development of Community Space for Multilingual Informal Communication Support Using Augmented Reality and Second Life, IEEE International Conference on Advanced Information Networking and Applications (AINA-2011), 2011年3月23日, シンガポール市, シンガポール

[6] Mai Miyabe, Takashi Yoshino, Development of Multilingual Medical Reception Support System with Text-to-speech Function to Combine Utterance Data with Voice Synthesis, ACM International Conference on Intercultural Collaboration (ICIC2010), 2010年8月19日, コペンハーゲン市, デンマーク

[7] Mai Miyabe, Takashi Yoshino, Influence of Detecting Inaccurate Messages in Real-time Remote Text-based Communication via Machine Translation, ACM International Conference on Intercultural Collaboration (ICIC2010), 2010年8月19日, コペンハーゲン市, デンマーク

[その他]

ホームページ等

<http://www.wakayama-u.ac.jp/~yoshino/lab/>

## 6. 研究組織

### (1)研究代表者

吉野 孝 (YOSHINO TAKASHI)  
和歌山大学・システム工学部・准教授  
研究者番号:90274860

### (2)研究分担者

林 良彦 (HAYASHI YOSHIHIKO)  
大阪大学・言語文化研究科・教授  
研究者番号:80379156

山下直美 (YAMASHITA NAOMI)  
日本電信電話株式会社 NTT コミュニケーション研究所・メディア情報研究部・社員  
研究者番号:70396141

## (3)連携研究者

該当なし