

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

福島 拓^{1,a)} 吉野 孝^{2,b)} 重野 亜久里³

受付日 2012年4月20日, 採録日 2012年7月2日

概要: 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は増加している. 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, 病名や症状を含む文の正確性と, 状況説明などの付与情報を含む文の網羅性を確保した多言語問診票の記入の可能性を示した.

キーワード: 問診票, 多言語, 用例対訳, 機械翻訳

Proposal and Evaluation of an Inputting Method for Multilingual Interview Sheet Using Parallel-text and Machine Translation

TAKU FUKUSHIMA^{1,a)} TAKASHI YOSHINO^{2,b)} AGURI SHIGENO³

Received: April 20, 2012, Accepted: July 2, 2012

Abstract: The number of foreign residents and foreign visitors in Japan has been increasing every year. The opportunities for communication amongst people whose native languages differ are increasing. In healthcare facilities, a paper-based multilingual interview sheet is used to facilitate communication between medical workers and foreign patients. However, this interview sheet has been found to be inadequate for such purposes. Moreover, Japanese medical workers find it difficult to understand the different languages written on a paper-based interview sheet. To resolve this problem, we have developed a multilingual interview-sheet composition system that uses parallel texts and machine translation. This system can convey essential patient information to a medical worker during consultation. Moreover, we have evaluated an inputting method for multilingual interview sheet. The contributions of this study are as follows: (1) We have proposed a multilingual interview-sheet composition system that can be used for communication between medical workers and foreign patients. We have developed this system using both parallel texts and machine translation. (2) This system uses both a parallel corpus and machine translation. As a result, we showed that a patient can write a multilingual interview sheet which has both the accuracy of sentences including disease names and symptoms, and the completeness of sentences including complementary information.

Keywords: interview sheet, multilingual, parallel text, machine translation

¹ 和歌山大学大学院システム工学研究科
Graduate School of Systems Engineering, Wakayama University, Wakayama 640-8510, Japan

² 和歌山大学システム工学部
Faculty of Systems Engineering, Wakayama University, Wakayama 640-8510, Japan

³ 特定非営利活動法人多文化共生センターきょうと
Center for Multicultural Society Kyoto, Kyoto 600-8191, Japan

a) fukushima@yoslab.net

b) yoshino@sys.wakayama-u.ac.jp

1. はじめに

現在, 在日外国人や訪日外国人は増加傾向にあり [1], [2], 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. しかし, 在日外国人や訪日外国人の中には, 日本語を理解できない人が多数存在している [3], [4], [5]. 一般に多言語を十分に習得することは非常に難しく, 母語以外の言語によるコミュニケーションは困難なこともあ

り [6], [7], [8], 日本語を理解できない外国人と日本人とのコミュニケーションは十分に行うことができない。このため、用例対訳や機械翻訳などの言語資源を組み合わせることで利用できる仕組みである言語グリッドの活動が広がるなど [9], [10], 言語の壁を越える活動が活発化している。

日本語を理解できないことの影響が顕著に現れる分野の1つに医療がある。医療分野では、わずかなコミュニケーション不足で医療ミスが発生する恐れがあり、正確な意思疎通が必要である。しかし、日本語が通じない外国人と日本人の医療従事者間でのやりとりは、意思疎通を十分に行うことができない。

初診の患者が日本の医療機関を訪れたとき、患者の症状を知るために問診票の記入を求めることが一般的である。また、多くの医療機関の問診票には自由記述の項目が存在している。これは患者の詳細な症状を医療従事者が得るために用意されている。しかし、多くの医療機関では日本人患者用の問診票のみを用意しており、外国人患者用の多言語問診票を用意している医療機関は少ない。このため、多言語に翻訳したPDF形式の問診票がWeb上で公開されている*1。しかし、PDF形式の多言語問診票には次の課題が存在していると考えられる。

- (1) あらかじめ用意された症状から選ぶ形式のため、患者が詳細な症状を伝えることが難しい。
- (2) 外国人患者が母語で自由記述項目に記入を行った場合、日本人医療従事者が問診票の内容を読み取ることが難しい。

そこで我々は、多言語問診票をWeb上で作成可能とした多言語問診票作成システムの開発を行った。本システムは、正確性の高い用例対訳と網羅性の高い機械翻訳の性質を考慮して併用し、患者の正確かつ詳細な症状伝達を可能とした。このことにより、上記の課題の解決を目指している。なお、用例対訳とは用例を多言語に翻訳した多言語コーパスのことを指す。

本論文では、まず関連研究と多言語問診票作成システムの設計について述べた後、外国人患者が使用する問診票入力機能の実験とその結果、考察、まとめの順に述べる。

2. 関連研究

正確性が求められる分野の多言語間コミュニケーション支援として、正確な翻訳が可能な用例対訳が多く用いられている。医療分野で用例対訳を用いたシステムとして、宮部らが開発した多言語医療受付支援システム M^3 がある [11]。 M^3 はタッチパネルで操作可能としたシステムで、対話機能、外国人患者の受診支援機能（問診機能、受診科選択機能など）を有している。また、杉田らは用例対訳を用いて多言語問診を携帯電話上で実現している [12]。これ

らのシステムでは、用例対訳を用いて症状ごとにフローをあらかじめ作成している。このため、医療機関が想定した症状について正確な対応が可能である。しかし、これらのシステムでは選択可能な用例対訳はあらかじめ作成しておく必要があるため、医療機関が想定していない症状の伝達や、用例対訳が用意されていない具体的な症状の伝達を患者が行うことはできない。本論文では、患者が症状を自由に記述した内容をもとに用例対訳や機械翻訳の選択を行い、この問題の解決を目指す。

自由な文を多言語に翻訳する仕組みとして機械翻訳がある。機械翻訳は入力された文をすべて翻訳可能なため、子供向けの機械翻訳 [13] や多言語対面環境の討論支援 [14] など、様々な分野で利用されている。しかし、機械翻訳の精度は年々向上しているものの、正確性が求められる医療分野でそのまま利用可能な精度には達していない [15]。このため、本論文では、正確な用例対訳とすべての文を翻訳可能な機械翻訳を併用し、正確性が求められる多言語医療問診の支援を目指す。

なお、用例対訳と機械翻訳の併用は、旅行中の会話支援を目的として行われているが [16]、正確性が求められる医療の分野での利用は行われていない。本論文では、正確性が求められる分野での利用を考慮して用例対訳と機械翻訳の併用を目指す。その際、機械翻訳を使用したときは、機械翻訳を利用している旨の提示と折り返し翻訳の提示を行うことで、医療問診での適用を目指す。なお、これらについては3章で詳しく述べる。また、文献 [16] では機械翻訳を使用した場合の正確性についての議論を行っていない。本研究では正確性が求められる分野での利用を想定しているため、機械翻訳を併用した場合の正確性についての議論を行うこととする。

最後に、関連研究と本研究との比較を表1に示す。正確性が求められる分野では、文献 [11], [12] のように用例対訳のみでの支援が多く行われてきている。しかし、用例対訳のみで十分な網羅性を確保することは難しい。そこで本論文では用例対訳のほかに機械翻訳を併用することでこれらの問題解決を目指す。機械翻訳を用いた場合、正確性を考慮する必要がある。文献 [16] では用例対訳と機械翻訳の併用が行われているものの、適用対象が旅行会話分野であることもあり、併用したことに関する正確性の検証が行われていない。本システムの適用対象は正確性が求められる医療分野であるため、本論文では用例対訳と機械翻訳を併用した場合の正確性に関して議論を行う。本論文では、表1に示すように、病名や症状の内容は用例対訳を用いて正確性を確保し、症状が発生したときの状況説明などは機械翻訳を用いて網羅性を高め、用例対訳と機械翻訳の併用の可能性を示す。

*1 <http://www.k-i-a.or.jp/medical/>

表 1 関連研究と本研究との比較

Table 1 Comparison between related works and our system.

| 対象分野 | システム名 | 利用技術 | 網羅性 | 正確性 |
|------------|--------------------|--------|----------------------|----------------------|
| 医療 | 本システム | PT, MT | ○ | 病名, 症状: ○ 状況説明: △ |
| | 多言語医療受付支援システム [11] | PT | 病名, 症状: ○ 状況説明: — | ○ |
| | ケータイ多言語対話システム [12] | | | |
| 子供向け 討論 | 子ども向け機械翻訳システム [13] | MT | ○ | △ |
| 旅行 | 旅行会話自動通訳システム [16] | PT, MT | ○ | 未検証 |

・PT は用例対訳を, MT は機械翻訳をそれぞれ示す。
 ※医療分野の用例対訳を利用した既存システムは, 状況説明をサポートしていない。

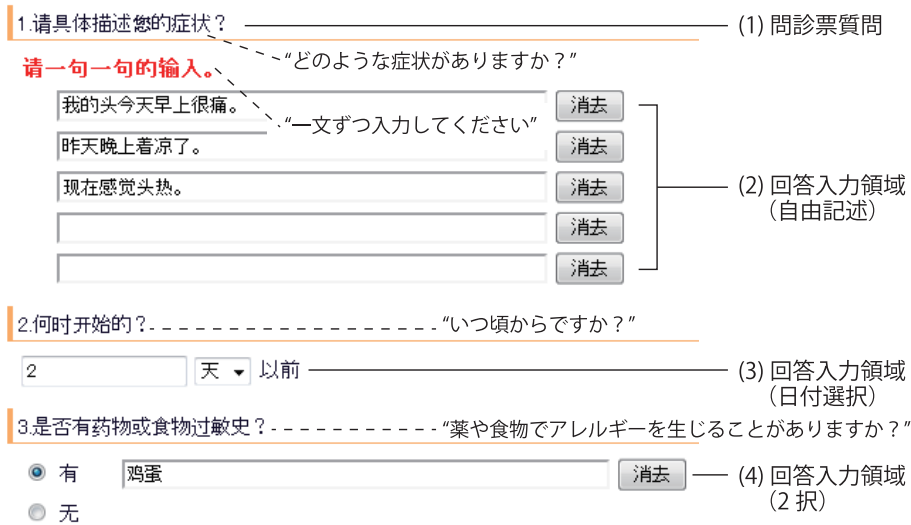


図 1 問診票入力機能の画面例

Fig. 1 Screenshot of function for inputting data into interview sheet.

3. システムの設計

本章では多言語問診票作成システムの設計について述べる。以降の各節で問診票入力機能について述べた後, 本システムで利用している用例対訳, 機械翻訳の各サービスについて述べる。なお, 本システムは Web 上で動作する。

3.1 問診票入力機能

本機能は, 多言語問診票に患者が症状を入力する機能である。利用者は日本語が理解できない外国人患者を想定している。なお, 利用に関してアカウント登録は不要とした。これは, 患者名と個人情報である問診票の内容との関連づけを行わないようにするためである。また, 作成した問診票の情報も個人情報保護のために記録していない。

本機能の画面例を図 1 に示す。なお, 患者の母語は図 1 の前の画面で選択できるようにしている。また, 現時点で公開しているシステムは日本国内での利用を想定しているため, 医療従事者の言語は日本語としている。

問診票を作成する患者は図 1(1) の質問項目を見て入力領域に症状を入力する。図 1(2) や図 1(4) などの自由記述

用テキストエリアをクリックすると, 翻訳画面が現れる。図 2 に翻訳画面例を示す。図 2 は図 2(1) のテキストボックスをクリックしたときの画面例となっている。翻訳画面では, 図 2(2) に入力した文字列をもとに用例対訳の検索と機械翻訳での翻訳を行う。検索した結果は, 図 2(3) に用例対訳が, 図 2(4) に機械翻訳を使用した翻訳がそれぞれ表示される。

図 2(3) の用例対訳の検索は, N-gram に基づく用例対訳検索手法を利用している [17]。文献 [17] では, 用例を 2-gram もしくは 4-gram に分割し, 検索文字列との共起を調べることで多言語の類似文検索を実現している。なお, 用例対訳は最大 3 件表示する。類似した用例対訳が存在しなかった場合は, 類似用例対訳がない旨のみが表示される。

図 2(4) に表示される機械翻訳の翻訳結果は精度が不十分な場合がある。このため, 機械翻訳の翻訳文には図 2(5) のように, 間違いが含まれている可能性を示す注意文を赤字で明記している。また, 図 2(4) に表示している言葉は機械翻訳の結果と, 再度元言語に翻訳し直した結果を併記して表示している (図 2(6))。この手法は折り返し翻訳と呼ばれ, 翻訳結果を利用者が理解可能にする目的で用いら

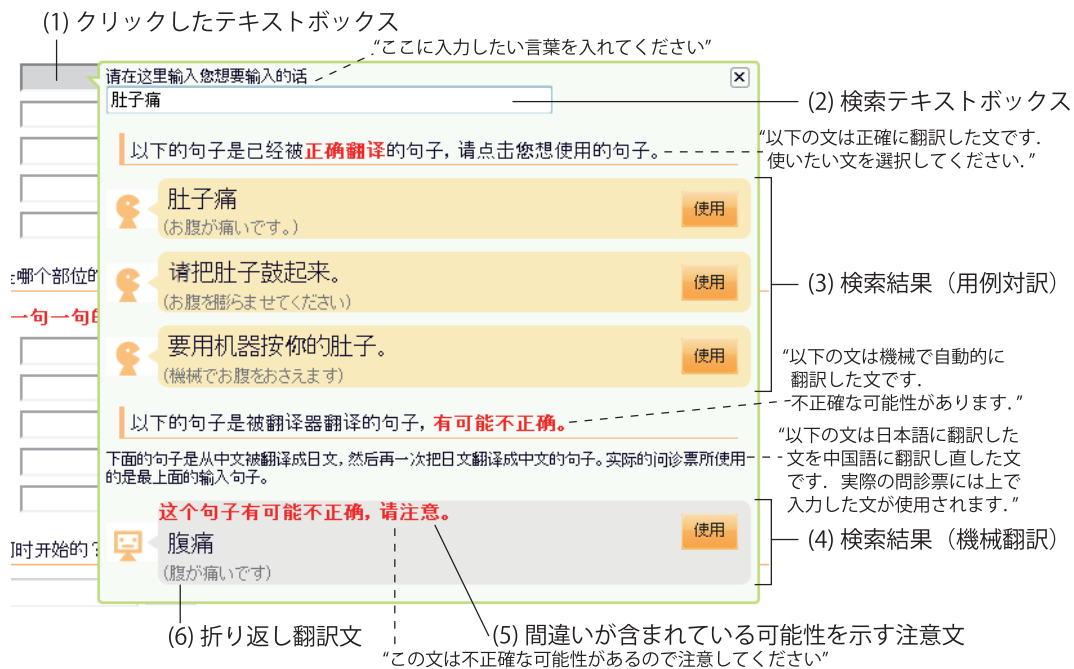


図 2 翻訳入力画面例

Fig. 2 Screenshot of function for selecting translation.

れている [15], [18]. 折り返し翻訳を用いることで, 文の入力者である患者は文の精度を確認し, 必要に応じて文の修正による翻訳精度向上が可能である. なお, 機械翻訳はどのような文でも翻訳することができるため, 翻訳結果は必ず表示される. このため, 図 2 の翻訳画面には, 最低 1 文 (機械翻訳 1 文のみ), 最大 4 文 (用例対訳 3 文, 機械翻訳 1 文) が表示される.

最後に, 図 2(3) もしくは図 2(4) の右にある「使用」ボタンをクリックすることで図 2(1) のクリックしたテキストボックスに用例対訳または機械翻訳の内容が入力される. このようにすることで, 入力された元言語と翻訳先の言語の 2 カ国語 (図 1, 図 2 の例の場合日本語と中国語) に対応した問診票を作成することができる.

患者が問診票の記入を行い, 「問診票作成」ボタンをクリックすると PDF 形式の問診票が生成される. 生成される問診票は医療機関の言語と患者言語の併記を行っている. また, 機械翻訳を使用した文には, 機械翻訳を使用している旨を 2 言語併記している. このことで, 医療従事者が用例対訳か機械翻訳かを判断することができる. 問診票の大きさは, 日本の医療機関で多く使用されている A4 サイズとした. このため, 生成した PDF 形式の問診票は, 印刷して医療従事者が確認, 保存することができる.

3.2 用例対訳および機械翻訳サービス

本節では, 本システムで利用している用例対訳サービスと機械翻訳サービスについて述べる. 本システムで利用している用例対訳は, 用例対訳共有システム TackPad [19] で収集したものを利用している. TackPad は様々な母語の利

用者が協力して多言語用例対訳を収集するシステムである. TackPad は医療分野で使用される用例対訳を包括的に収集することを目的としている. なお, TackPad は用例対訳の収集, 共有が目的のシステムのため, 収集した用例対訳の利用は TackPad が用例対訳の提供を行う多言語対応システムでのみ可能となっている. TackPad では, (i) 医療従事者や患者などの利用者が必要な用例をシステムに登録, (ii) 翻訳者が (i) で登録された用例を各言語に翻訳, という流れで用例対訳を収集している. 収集言語は日本語, 英語, 中国語, 韓国朝鮮語, ポルトガル語, スペイン語, ベトナム語, タイ語, インドネシア語の 9 言語である. また, 9 言語合計で約 14,500 件*2の用例を収集している. TackPad で収集した用例対訳は, 複数の評価軸を用意した評価手法を用いて, 利用者相互で正確性評価を行うことで精度の確保を行っている [20]. また, 機械翻訳は言語グリッド [9] が提供している機械翻訳の中で, 最も精度の良い傾向のある J-Server を利用した [21]. 今回は, 上記の用例対訳および機械翻訳サービスを利用したが, 本システムは医療分野を対象とした用例対訳や, 言語グリッド以外の機械翻訳器も利用可能である.

4. 試用実験

開発した多言語問診票作成システムを用いて実験を行った. 実験の目的は, 3.1 節で述べた問診票入力機能についての有用性確認である. 具体的には, 用例対訳と機械翻訳の併用利用の可能性調査, 作成された多言語問診票の正確性調査, 従来の紙の問診票との比較を実験の目的とする.

*2 2012 年 4 月現在

実験の被験者は中国人留学生 10 名である。また、実験後に実験で使用した用例対訳や機械翻訳の評価を学生 6 名（中国人留学生 3 名、日本人学生 3 名）が行った。なお、中国人留学生は全員日本語を理解することができる。このため、実験時のシステムインタフェースはすべて中国語とし、日本語の併記を行っていない。これは、本システムが日本語を記入、入力できない外国人を想定しており、今回の被験者と異なっているため、より実利用の状況に近づけるために行った。また、有用性確認実験と評価実験の中国人留学生は別の被験者が行っている。4.1 節で有用性確認実験の概要を、4.2 節で入力内容の評価実験の概要を述べる。有用性確認実験では、問診票入力機能の有用性確認を行う。また、入力内容の評価実験では、有用性確認実験で入力された多言語の文の正確性を検証する。

なお、著者らと実験被験者は同一大学の所属であるため、下記の配慮を行っている。

- (1) 被験者自身の症状を依頼し、幅広い症状入力を可能とする。
 - (2) 4.1 節と 4.2 節の被験者を変え、さらに主観の影響が大きい 4.1 節の被験者はすべて研究室外の学生とする。
- ただし、実験者効果を完全には排除しきれていない可能性がある。なお、(1) については 4.1 節で詳しく述べる。

4.1 有用性確認実験

問診票入力機能の有用性確認の実験を、中国人留学生 10 名に対して行った。本実験では、医療機関で多く用いられている紙の問診票と、多言語変換可能な本システムでの問診票入力機能の比較を行った。紙の問診票は翻訳が行われないため実際の医療現場で簡単に使用できないが、入力時間の比較のために記入を依頼している。実験では、被験者自身の症状を記入するように依頼した。これは、著者らが想定した問診内容の記入を依頼した場合、症状に偏りが発生する可能性があるからである。

実験の流れを以下に示す。

(1) 実験の概要説明

自分の病気の症状を書いてもらうため、プライバシーポリシーを用意し、承諾を得た被験者のみの参加とした。なお、伝えたくない症状については書かなくてもよいものとした。

(2) 紙の問診票およびシステムの間診票への記入

中国語で紙とシステム両方の問診票の記入を依頼した。なお、紙とシステムの間診票に書く内容については同一の症状を書くように依頼した。また、紙とシステムの記入については順序効果を考慮している。なお、被験者にシステムの使用方法は説明せず、操作方法が分からないときに質問するように依頼した。

(3) 中間アンケートの記入

紙の問診票およびシステムの間診票の記入に関するア

表 2 問診票の質問項目

Table 2 Questions used in the interview sheet for the experiment.

| | 質問項目 | 質問形式 |
|---|-------------------------|--------------|
| 1 | どのような症状がありますか？ | 自由記述 |
| 2 | どの部位の症状ですか？ | |
| 3 | いつごろからですか？ | 数字, 単位の選択 |
| 4 | 薬や食物でアレルギーを生じることがありますか？ | ラジオボタン, 自由記述 |
| 5 | 現在飲んでいる薬はありますか？ | |

・被験者へ質問項目の提示はすべて中国語で行っている。

ンケートの記入を依頼した。

(4) 紙の間診票およびシステムの間診票への記入

(2) の 1 回目の問診票記入と同様に、紙とシステムの間診票に中国語で記入を依頼した。なお、1 回目の症状とは別の症状の記入を依頼している。また、紙とシステムの記入順序は 1 回目の問診票記入と同様の順序で行っている。

(5) 最終アンケートの記入

紙の間診票およびシステムの間診票の記入に関するアンケートの記入を再度依頼した。

なお、本実験では問診票の記入を 2 回行っている。これは、1 回目の問診票記入は初めてシステムを操作する患者を、2 回目の問診票の記入はシステムに慣れた患者をそれぞれ想定している。

実際に使用した問診票の質問項目を表 2 に示す。この内容と同一のものを紙とシステムで構築している。なお、実験時の質問の提示はすべて中国語で行っている。また、医療機関で用いられている既存の間診票は、症状をチェックボックス形式で並べ、選択する形式が多く用いられている。しかし、本実験では自由記述を用いた問診票入力機能の有用性確認実験のため、表 2-1、表 2-2 の質問項目は自由記述のみとしている。

また、本システムで利用する TackPad の用例対訳には質問項目が多く含まれている。このため、本実験では「？」を含む用例は実験中に表示しないようにした。本実験で利用可能な日中の用例対訳数は 314 対であった。

4.2 入力内容の評価実験

本実験では、4.1 節で入力された文の評価を、中国人留学生、日本人学生それぞれ 3 名に依頼した。なお、中国人留学生は、日本語検定 1 級、もしくは日本語検定 1 級と同等の日本語能力を持つ人を選定している。

本実験で評価を依頼した文は下記の内容である。

中国語の文：4.1 節の被験者が実験中に使用した文

- 用例対訳の場合：被験者が選択した文
- 機械翻訳の場合：被験者が入力した文

表 3 評価実験の調査項目

Table 3 Investigation items of the evaluation experiment.

| | 調査項目 | 評価段階 |
|-----|-------------------|------|
| 中国人 | 日本語と中国語の意味の比較 | 5 段階 |
| | 中国語の質問と回答の組合せの妥当性 | 2 値 |
| 日本人 | 日本語の質問と回答の組合せの妥当性 | 2 値 |

日本語の文：上記の中国語の文を変換もしくは翻訳した文 (4.1 節の被験者には提示されていない)

- 用例対訳の場合：被験者が選択した文の対訳文
- 機械翻訳の場合：被験者が入力した文を機械翻訳が翻訳した文

評価項目を表 3 に示す。また、表 3 中の 2 値、5 段階の評価は、それぞれ次に示す基準で評価を依頼した。

2 値：あてはまるかあてはまらないか

5 段階：1：まったく違う意味，2：雰囲気は残っているが元の意味は分からない，3：意味はだいたいつかめる，4：文法などに多少問題があるがだいたい同じ意味，5：同じ意味

この基準は、調査項目の種類によって分けている。5 段階評価を適用した調査項目は 2 言語間の意味の確認を行っている。このため、2 言語間の意味比較に用いられる Walkerらの適合性評価を利用した [22]。本項目は、3 名の平均が 4 以下だった場合、不正確と判断している。4 は「文法などに多少問題があるがだいたい同じ意味」であるため、厳しめに判定を行っている。この結果は、5.2 節で行う考察で使用する。また、質問と回答の組合せの意味が通じるかどうかの評価項目については、上記の 5 段階調査の補助的な位置づけで行った。本項目は、5.2 節で行う考察の中で、上記の 5 段階評価で不正確と判定された項目において詳細な理由調査でのみ用いる。このため、細かい分類は不要と判断して 2 値での評価とした。本項目は、3 名のうち 1 名でもあてはまらないと判断した場合、不正確と判断している。

5. 実験結果と考察

5.1 用例対訳と機械翻訳の利用割合

本節では実験で使用された用例対訳と機械翻訳について考察する。実験中に使用された用例対訳と機械翻訳の利用回数を表 4 に示す。表 4 より、入力された文の 8 割以上が機械翻訳で作成されていたことが分かる。今回の結果は、用例対訳の提供元である TackPad は医療従事者からの質問項目が多かったためであると考えられる。アンケートからも、「(用例対訳では) 欲しい文が出てこなかった」という意見があった。このため、他の用例対訳サービスを併用するなど、今後用例対訳を増やしていく必要があると考えられる。なお、文献 [11] の問診機能では、1 言語あたり約 400 文の用例*3が使用されている。このため、最低でも問診関連の用例を 400 文以上用意する必要があると考え

表 4 用例対訳と機械翻訳の使用回数

Table 4 The number of parallel texts and machine-translated sentences used.

| | 1 回目 | 2 回目 | 合計 | 割合 |
|------|------|------|-----|--------|
| 機械翻訳 | 39 | 47 | 86 | 81.1% |
| 用例対訳 | 5 | 15 | 20 | 18.9% |
| 合計 | 44 | 62 | 106 | 100.0% |

・単位は文である。

表 5 機械翻訳を使用した文の分類

Table 5 The detailed content of the sentences being used machine translation.

| | 症状の説明 | | 状況の説明 | 合計 |
|-------|-------|-----|-------|-----|
| | 単語 | 文 | | |
| 文数 | 42 | 39 | 5 | 86 |
| 平均文字長 | 2.4 | 7.1 | 12.2 | 5.1 |

・文数の単位は文、平均文字長の単位は単語である。

られる。

また、問診票に入力され、機械翻訳で使用された文は、他の利用者が利用する可能性があると考えられる。このため、入力された文は用例対訳にすべき文であると考えられるため、多言語用例対訳共有システム TackPad [19] との連携を検討する。(i) 機械翻訳に入力された文のみ TackPad に提供する (翻訳された文は不正確な可能性があるため利用しない)、(ii) TackPad で文を多言語に翻訳し、正確性評価を行った用例対訳を作成する、という方法をとることで、実際に必要な用例対訳を効率良く収集可能になると考えられる。

なお、表 4 から 2 回目の方が多くの文を入力していることが分かるが、これは被験者の実験に対する慣れが影響していると考えられる。

機械翻訳を使用した文の分類を表 5 に示す。表 5 より、機械翻訳を使用した文の半数以上は病名や症状の単語であった。単語の利用頻度が高かったため、本システムで利用可能な単語の用例対訳を増やす必要があると考えられる。

また、「昨日バスケットボールをしたときに転んで腕が骨折した」など、症状が発生したときの状況を説明する文が 5 文含まれていた。これらの文は医療用語ではないため、患者が使用する文としてあらかじめ用意しておくことが難しい。ただし、機械翻訳は翻訳ルールや学習データをもとに翻訳を行っている [23]。このため、専門用語よりも一般的な用語の方が精度良く翻訳できる可能性が高いと考えられ、病状を説明する文は機械翻訳で対応可能である可能性が高い。ただし、これらの文は含まれる文字数が多くなっている。機械翻訳は文字数が短い方が精度良く翻訳できるため [24]、翻訳しやすい文を患者が入力するように誘導する必要があると考えられる。

*3 重複なしでの用例数である。同じ症状が複数の部位で利用されることがあるため、重複を許すと約 2,000 文の用例が存在している。

実際の利用時には、用例対訳と機械翻訳の特性を考慮してシステムインタフェースを設計する必要があると考えられる。具体的には、まず正確性が求められる症状伝達を用例対訳を用いて確認した後、それ以外の状況伝達を機械翻訳を用いることで、正確性と網羅性の両立が可能であると考えられる。

5.2 使用された文の正確性

本節では実験で使用された用例対訳と機械翻訳の正確性について考察する。なお、本節では4.2節で評価した結果を用いている。用例対訳、機械翻訳の日中間の意味が不正確な文(表3の5段階評価において、平均が4以下であった文)の数とその割合を表6に示す。表6より、機械翻訳の14%が不正確であったことが分かる。今回の精度評価は正確性が求められる医療分野への適用のため、大半の利用者が文を理解できるようにするために、厳しめに判定を行っている。機械翻訳の正確な割合は比較的高かったが、間違いも含まれていたことが分かる。このため、正確性が求められる用語はできるだけ用例対訳を利用する必要があると考えられる。

使用された文の一例を表7に示す。表7中の入力文の列は、被験者が入力した文である。また、翻訳文は用例対訳の場合は日本語に変換した文、機械翻訳の場合は日本語

表6 日中間の意味が不正確な文の数とその割合

Table 6 The percentage of inaccurate sentences between Japanese and Chinese translations.

| | 利用文数 | 不正確 | 不正確な割合 |
|------|------|-----|--------|
| 機械翻訳 | 86 | 12 | 14.0% |
| 用例対訳 | 20 | 1 | 5.0% |
| 合計 | 106 | 13 | 12.3% |

- ・単位は文である。
- ・4.2節の5段階による評価を行い、平均が4以下の場合不正確と判定している。

表7 実験で使用された文(一例)

Table 7 Examples of used sentences in the experiment.

| ID | 質問 | 翻訳 | 入力文(中国語) | 翻訳文(日本語) | 正確 |
|-----|----|----|--------------|--------------------------|----|
| 7 | 2 | MT | 皮膚又疼又痒 | 皮膚はかゆいがありましたり、かゆいでかたりします | |
| 10 | 2 | MT | 我的头今天早上很痛。 | 私の頭は今朝とても痛いです。 | ○ |
| 19 | 2 | MT | 昨天踢球时,小腿骨折了。 | 昨日ボールをキックする時、下腿骨は折れました。 | ○ |
| 20 | 2 | MT | 摔倒了,手腕划破了。 | 転んで、計略は傷つきました | |
| 24 | 1 | MT | 腿 | 足 | |
| 46 | 1 | PT | 上下方向投影。 | 頭尾方向撮影 | |
| 63 | 2 | PT | 没有食欲 | 食欲がありません | ○ |
| 82 | 1 | MT | 看书看久了眼睛也疼 | 本を読んで長くなった目もかわいがることを見ます | |
| 91 | 2 | PT | 发烧 | 熱があります | ○ |
| 102 | 2 | MT | 脸部流血 | 顔は血を流します | ○ |

- ・IDは実験中に使用された文の連番である。
- ・質問の1は、「どのような症状がありますか?」への回答、2は「どの部位の症状ですか?」への回答である。
- ・翻訳のPTは、用例対訳で変換した文、MTは機械翻訳で翻訳した文である。
- ・正確の○は、その文が正確であると判断されたことを示す。

に翻訳した文である。

機械翻訳で翻訳された文で不正確と判定されたものの例として、中国語で「腿」、日本語で「足」というものがあつた(表7のID:24)。中国語で「腿」は足の付け根から足首までのことを指す。また、日本語の場合、足の付け根からつま先までのことを指す。このため、不正確と判定された。このように、体の部位は言語によって異なることがあり[25]、誤診の原因となる可能性がある。このことから、文字情報だけで誤診を防ぐことは難しいと考えられる。

解決策として、人体図を用いる例をあげる。患者が人体図に印をつけることで、患部の情報を視覚的に伝えることができるため、誤診を減らすことができると考えられる。なお、人体図については有用性検証実験を行っていないため、今後行う必要がある。

なお、他の11件の不正確な文は機械翻訳の誤訳により意味が通じないものとなっていた。このように部位以外の誤訳も含まれているため、5.1節で述べたように機械翻訳で使用された文を用例対訳にする必要があると考えられる。

また、用例対訳の1文も不正確であった。これは、「どの部位の症状ですか?」という質問に対して「頭尾方向撮影(日本語)-上下方向投影(中国語)」*4という用例対訳を用いて回答を作成していた(表7のID:46)。この用例対訳の場合、日本語と中国語の意味(表3の5段階)は同じであった。しかし、日本語、中国語ともに、質問と回答の組合せ(表3の2値)を評価した結果は不正確という判定になっていた。このことから、この用例対訳自体は正確であったが、今回の用例対訳の使用場面が適切でなかったことが分かる。今回の場合、用例対訳を選択した被験者が誤用した可能性が考えられる。ただし、出現した用例対訳は一般的な問診票でありあまり利用されないものであると考えられるため、用例対訳に利用頻度による重み付けを行い、検索性能を向上させる必要があると考えられる。

5.3 用例対訳と機械翻訳の併用の効果

本節では、用例対訳と機械翻訳を併用したことによる効果を、(1)用例対訳のみを利用したシステムの場合との比較、(2)機械翻訳のみを利用したシステムの場合との比較により検証する。

5.3.1 用例対訳のみを利用したシステムとの比較

5.1節の結果より、本システムは不足していた用例対訳を機械翻訳で補っている。このことにより、本システムは患者の詳細な症状伝達支援を可能にしたと考えられる。これは、用例対訳のみを使用した従来の医療支援システムよりも優位な点である。本システムは利用可能な用例対訳の数が十分ではないという問題は存在していた。しかし、用例対訳を十分に用意した場合でも、新語への対応や幅広い

*4 レントゲン撮影で用いられる用語

表 8 実験で使用された用例対訳を機械翻訳で翻訳した例

Table 8 Example of translated texts for used parallel texts in the experiment.

| ID | 原文(中国語) | 用例対訳(日本語) | 機械翻訳(日本語) | 正確性 |
|----|---------|-----------|-----------|-----|
| 1 | 发烧 | 熱があります | 熱が出る | × |
| 3 | 头疼 | 頭が痛い | 頭痛 | ○ |
| 4 | 肩膀酸疼 | 肩が凝っている | 肩はだるくて痛い | × |
| 6 | 感冒 | 風邪 | 風邪 | ○ |
| 21 | 肚子痛 | お腹が痛いです。 | 腹が痛い | ○ |
| 47 | 食物过敏 | 食物アレルギー | 食品は過敏だ | × |
| 55 | 头晕 | めまい | くらくらする | ○ |
| 60 | 没有食欲 | 食欲がありません | 食欲がない | ○ |
| 69 | 呼吸困难 | 息苦しいです | 呼吸困難だ | ○ |
| 81 | 眼睛疼 | 目が痛い | 目のかわいがること | × |

・ID は実験中に使用された文の連番である。

表現への対応などは、用例対訳のみでは対応が難しいと考えられる。たとえば、表 7 の ID: 19 (日本語文: 昨日ボールをキックする時、下腿骨は折れました) は症状のみでなく症状が発生したときの状況を説明しているため、用例対訳として用意することが難しいが、本システムでは機械翻訳によって意思の伝達を可能としている。これらのことから、用例対訳のみの場合と比較し、機械翻訳との併用を行う方が効果的に患者の詳細な症状伝達支援を可能にすると考えられる。

5.3.2 機械翻訳のみを利用したシステムとの比較

本実験では用例対訳と機械翻訳の併用を行ったが、本項では機械翻訳のみの利用を想定し、用例対訳が使用された文を機械翻訳で翻訳した。用例対訳を機械翻訳で翻訳した例を表 8 に示す。表中の「原文 (中国語)」と「用例対訳 (日本語)」が実験中で使用された用例対訳である。表 8 では、「原文」を機械翻訳で翻訳したものを「機械翻訳 (日本語)」に示している。また、機械翻訳 (日本語) の正確性に関しては、用例対訳 (日本語) と機械翻訳 (日本語) を比較して判定した。なお、実験では 20 文の用例対訳が使用されたが、重複した文と 5.2 節で不正確と判定された文を除いた 10 文に関して調査した。表 8 より、10 文中 4 文が不正確であったことが分かる。なお、表 8 の ID: 1 の機械翻訳は、中国語との対としては正確である。しかし、この文に対応する質問文が「どのような症状がありますか?」であったため、不正確であると判定した。

このように、機械翻訳のみを使用した場合は、用例対訳と機械翻訳を併用したときと比べて不正確な文が多く発生することが分かる。これらのことから、用例対訳と機械翻訳を併用することで、特にあらかじめ想定可能な患者の症状伝達において正確性を向上させることが可能であると考えられる。

5.4 被験者の行動分析

本節では、被験者の行動について考察する。紙の問診票とシステムの間診票の作成時間を表 9 に示す。表 9 より、1 回目、2 回目ともに紙の間診票はシステムよりも作業時

表 9 紙とシステムの間診票の記入時間

Table 9 Time to write interview sheets on paper and using the system.

| | 1 回目 | | 2 回目 | |
|------|-------|--------|-------|-------|
| | 紙 | システム | 紙 | システム |
| 平均 | 154.7 | 293.4 | 112.0 | 201.8 |
| 標準偏差 | 60.04 | 203.25 | 48.75 | 82.78 |
| 最短 | 66 | 150 | 30 | 82 |
| 最長 | 258 | 820 | 197 | 341 |

・単位は秒である。

間が短いという結果になったことが分かる。これは、キーボード入力や入力文の翻訳に時間がかかるためであると考えられる。しかし、紙の間診票は患者の母語で書かれているため、医療従事者が理解するためには翻訳が必要である。翻訳者がいない場合は、問診票に書かれた文を翻訳することは困難である。また、システムには自由記述が平均 5.3 文 (最低 3 文, 最高 10 文) 入力されていた。翻訳者がこれらの文を翻訳し、医療従事者に伝えるために紙に再記入することをシステムの入力時間よりも早く行うことは困難であると考えられる。このため、システムの間診票が優位であると考えられる。

なお、20 回の実験のうち 3 回の実験では紙の間診票よりシステムの間診票の方が早く入力を完了していた。また、2 回目のシステム間診票の方は 1 回目よりも平均約 90 秒短くなっており、標準偏差も小さくなっている。このことから、1 度システムを操作し、慣れた利用者は比較的早く操作が可能であると考えられる。ただし、「普段パソコンを使っていますか?」というアンケートの質問項目に対して、被験者全員が「毎日パソコンを使っている」と回答していた。このため、計算機に慣れていない患者が利用可能かどうかについては今後検証する必要がある。

表 10 にアンケート結果を示す。表 10-1、「問診票作成システムで自分の伝えたいことを書くことができた」の結果は、中間アンケートと最終アンケートの中央値、最頻値が、ともに 4 という結果となった。このことから、用例対訳と機械翻訳とを併用することで患者側の意思を医療従事者側に伝えることができると考えられる。ただし、4 割の被験者は「自分の伝えたいことを書くことができなかった」もしくは「どちらとも言えない」という評価となった。これは、用例対訳の数の少なさと機械翻訳の翻訳精度が影響していたと考えられる。表 10-1 で 2 もしくは 3 を選択した被験者は、アンケートの自由記述で、用例対訳については「正しい文が出てこなかった (自分の使いたい用例対訳がなかった)」、機械翻訳については「(折り返し) 翻訳で間違えたものが出てきた」「使いたい文が出てこなかったときに修正が面倒だった」という回答を行っており、この内容がシステムで伝えたい内容を書くことができないという評価に至った理由であると考えられる。これらは、用例対

表 10 アンケート結果
Table 10 Result of questionnaire.

| | | 評価段階 | | | | | 中央値 | 最頻値 | |
|---|------------------------------|------|---|---|---|---|-----|-----|---|
| | | 1 | 2 | 3 | 4 | 5 | | | |
| 1 | 問診票作成システムで自分の伝えたいことを書くことができた | 中間 | 0 | 3 | 1 | 5 | 1 | 4 | 4 |
| | | 最終 | 0 | 2 | 2 | 5 | 1 | 4 | 4 |
| 2 | 問診票作成システムを使うことは難しかった | 中間 | 1 | 8 | 0 | 1 | 0 | 2 | 2 |
| | | 最終 | 2 | 5 | 2 | 1 | 0 | 2 | 2 |
| 3 | 病院にこのシステムがあれば使いたい | 最終 | 0 | 0 | 1 | 9 | 0 | 4 | 4 |

・評価段階：1：強く同意しない，2：同意しない，3：どちらとも言えない，4：同意する，5：強く同意する
 ・表中の評価段階の数字は人数を表す。
 ・「中間」は中間アンケートの回答結果，「最終」は最終アンケート結果を示す。

訳を増やすことで解決できるため，5.1 節で述べたように，使用可能な用例対訳を増やす必要があると考えられる。

また，表 10-2，「問診票作成システムを使うことは難しかった」の結果は中間アンケート，最終アンケートともに中央値，最頻値が2という結果となった。このため，操作は比較的分かりやすかったと考えられる。

6. おわりに

本論文では，多言語間コミュニケーションの機会の増加を背景に，多言語問診票作成システムの開発を行った。本システムは，医療従事者と患者との間の意思の伝達支援を目的としている。本論文では，従来の用例対訳のみのシステムに存在していた用例対訳の網羅性の問題と，機械翻訳のみのシステムに存在していた翻訳精度の問題を解決するため，用例対訳と機械翻訳の併用を行った。また，実験を通して患者が使用する問診票入力機能の評価を行った。

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) 正確性が担保された用例対訳と，すべての文の翻訳が可能な機械翻訳の併用を行った。このことで，単独の技術だけでは困難であった，言葉の正確性と利用可能な文の網羅性が確保された多言語問診票の記入の可能性を示した。具体的には，病名や症状伝達など，正確性が求められる内容に関しては用例対訳を，症状発生時の状況伝達などは機械翻訳を用いて網羅性を高めることで問題解決を行った。

今後，機械翻訳で入力された用例をもとにして多言語用例対訳の作成を行うため，多言語用例対訳共有システムとの連携を行う。また，今回構築した多言語問診票作成システムは Web 上で公開している*5。今後，利用者の行動ログや聞き取り調査から，計算機に慣れていない利用者の問診票作成の操作に関する評価を行う。

謝辞 本研究の一部は，総務省戦略的情報通信研究開発

推進制度 (SCOPE) の平成 22 年度採択課題「医療現場における利用者適応型多言語間コミュニケーション支援のための基盤技術の研究開発」および科研費基盤研究 (B) (22300044) の補助を受けた。

参考文献

- [1] 法務省：平成 22 年末現在における外国人登録者統計について，法務省 (オンライン)，入手先 (<http://www.moj.go.jp/nyuukokukanri/kouhou/nyuukantourokusyatoukei110603.html>) (参照 2012-04-19)。
- [2] 法務省：平成 23 年における外国人入国者数及び日本人出国者数について (確定値)，法務省 (オンライン)，入手先 (<http://www.moj.go.jp/nyuukokukanri/kouhou/nyuukokukanri04.00017.html>) (参照 2012-04-19)。
- [3] 田村太郎：多民族共生社会ニッポンとボランティア活動，明石書店 (2000)。
- [4] 文部科学省：「日本語指導が必要な外国人児童生徒の受入れ状況等に関する調査 (平成 20 年度)」の結果について，文部科学省 (オンライン)，入手先 (http://www.mext.go.jp/b_menu/houdou/21/07/_icsFiles/afieldfile/2009/07/03/1279262_1_1.pdf) (参照 2012-04-19)。
- [5] 犬飼 章：第 2 回多文化共生の推進に関する意見交換会 (宮城県の取り組み事例)，総務省 (オンライン)，入手先 (http://www.soumu.go.jp/main_sosiki/kenkyu/tabunka/21171_3.html) (参照 2012-04-19)。
- [6] Takano, Y. and Noda, A.: A Temporary Decline of Thinking Ability During Foreign Language Processing, *Journal of Cross-Cultural Psychology*, Vol.24, pp.445-462 (1993)。
- [7] Aiken, M., Hwang, C., Paolillo, J. and Lu, L.: A group decision support system for the Asian Pacific rim, *Journal of International Information Management*, Vol.3, No.2, pp.1-13 (1994)。
- [8] Kim, K.J. and Bonk, C.J.: Cross-Cultural Comparisons of Online Collaboration, *Journal of Computer Mediated Communication*, Vol.8, No.1 (2002)。
- [9] Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration, *IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06)*, pp.96-100 (2006)。
- [10] Sakai, S., Gotou, M., Tanaka, M., Inaba, R., Murakami, Y., Yoshino, T., Hayashi, Y., Kitamura, Y., Mori, Y., Takasaki, T., Naya, Y., Shigeno, A., Matsubara, S. and Ishida, T.: Language Grid Association: Action Research on Supporting the Multicultural Society, *International Conference on Informatics Education and Research for Knowledge-Circulating Society (ICKS-08)*,

*5 <http://is.tackpad.net/>

- pp.55-60 (2008).
- [11] 宮部真衣, 吉野 孝, 重野亜久里: 外国人患者のための用例対訳を用いた多言語医療受付支援システムの構築, 電子情報通信学会論文誌, Vol.J92-D, No.6, pp.708-718 (2009).
- [12] 杉田奈未穂, 丸田洋輔, 長谷川旭, 長谷川聡, 宮尾 克: ケータイ多言語対話システムとその応用, シンポジウム「モバイル'09」, pp.63-66 (2009).
- [13] Matsuda, M. and Kitamura, Y.: Development of Machine Translation System for Japanese Children, *Proc. 2009 ACM International Workshop on Intercultural Collaboration (IWIC'09)*, pp.269-271 (2009).
- [14] 福島 拓, 吉野 孝, 喜多千草: 共通言語を用いた対面型会議における非母語話者支援システム PaneLive の構築, 電子情報通信学会論文誌, Vol.J92-D, No.6, pp.719-728 (2009).
- [15] 林田尚子, 石田 亨: 翻訳エージェントによる自己主導型リペア支援の性能予測, 電子情報通信学会論文誌, Vol.J88-D1, No.9, pp.1459-1466 (2005).
- [16] Ikeda, T., Ando, S., Satoh, K., Okumura, A. and Watanabe, T.: Automatic Interpretation System Integrating Free-style Sentence Translation and Parallel Text Based Translation, *Proc. Workshop on Speech-to-Speech Translation*, pp.85-92 (2002).
- [17] 田淵裕章, 坂本 廣, 北村泰彦: N-gram に基づく用例対訳検索手法, 信学技報, 人工知能と知識処理研究会, Vol.108, No.441, pp.43-48 (2009).
- [18] 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).
- [19] 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, 情報処理学会研究報告, コンシューマ・デバイス&システム研究会, Vol.2012-CDS-4, No.5, pp.1-8 (2012).
- [20] 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, Vol.52, No.1, pp.131-139 (2011).
- [21] Miyabe, M. and Yoshino, T.: Features of Accuracy Mismatch between Back-translated Sentences and Target-translated sentences, *IEEE 2011 2nd International Conference on Culture and Computing*, pp.65-68 (2011).
- [22] Walker, K., Bamba, M., Miller, D., Ma, X., Cieri, C. and Doddington, G.: Multiple-Translation Arabic (MTA) Part 1, *Linguistic Data Consortium*, Philadelphia (2003).
- [23] 塚田 元, 渡辺太郎, 鈴木 潤, 永田昌明, 磯崎秀樹: 統計的機械翻訳, NTT 技術ジャーナル, Vol.19, No.6, pp.23-25 (2007).
- [24] Shigenobu, T.: Evaluation and Usability of Back Translation for Intercultural Communication, *Proc. 2nd International Conference on Usability and Internationalization (UI-HCII'07)*, pp.259-265 (2007).
- [25] 英語・英会話の情報ランド: 日本語と英語のズレ (3), 英語・英会話の情報ランド (オンライン), 入手先 (<http://www.eigo-eikaiwa.com/0236.html>) (参照 2012-04-19).



福島 拓 (学生会員)

言語間コミュニケーション支援に関する研究に従事。



吉野 孝 (正会員)

究に従事。



重野 亜久里

システム開発に従事。現在, 同法人代表理事長。