

正確かつ自由度を高めた多言語間対話支援を目的とした 応答用例対構築モデル

福島 拓^{1,a)} 吉野 孝^{2,b)}

受付日 2014年4月11日, 採録日 2014年10月8日

概要: 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. 正確な多言語支援が求められる場では, 用例対訳が多く用いられている. しかし, 既存の用例対訳利用システムでは, 文脈が異なる用例対訳が提示される場合があった. また, 用例対訳は定型文のため, 完全一致の文のみの利用しか行うことができなかった. そこで本論文では, これらの解決を目指した応答用例対構築モデルを提案する. また, モデルを適用したシステムを用いて実験を行い, モデルの有用性検証を行った. 本研究の貢献は次の3つである. (1) 従来の対話型用例対訳利用システムにおける問題解決を目的とした, 応答用例対を用いたモデルを提案した. (2) 応答用例対の自動収集により, 適切な回答選択の支援の可能性を示した. (3) 応答用例対への類似文の付与により, 用例対訳を用いた自由な対話の可能性を示した.

キーワード: 応答用例対, 用例対訳, 多言語, コミュニケーション支援

Paired Dialogue Parallel Text Composition Model for Reliable and Unrestricted Multilingual Communication Support

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

Received: April 11, 2014, Accepted: October 8, 2014

Abstract: Recently, worldwide globalization has helped to improve communication among people with different native languages. However, sharing accurate information in many languages is not sufficient. In general, multilingual support systems for applications that require high accuracy use parallel texts, but the context may be incorrect in systems that use existing parallel text systems. Moreover, parallel texts are fixed phrases, i.e., a system can only select a perfectly matching sentence. In this study, we propose a paired dialogue parallel text composition model to resolve the existing problems with multilingual communication support. We also developed this system and conducted an experimental verification of the usefulness of the model. The main contributions of this study are as follows. (1) We propose a model that uses dialogue parallel texts to solve the problems of existing multilingual dialogue support system by using parallel texts. (2) This model collects dialogue parallel texts automatically. We also demonstrate that the model can support the selection of the most appropriate answers. (3) We show that the model can support unrestricted dialogue by connecting similar text with parallel dialogue texts.

Keywords: paired dialogue parallel text, parallel text, multilingual, communication support

1. はじめに

近年の世界的なグローバル化により多言語間コミュニケーションの機会が増加している. 日本国内でも訪日外国人が増加傾向にあり [1], 今後, 外国人住民の増加も予想されている [2]. しかし, 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

¹ 静岡大学大学院工学研究科
Graduate School of Engineering, Shizuoka University,
Hamamatsu, Shizuoka 432-8561, Japan

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

a) fukushima@sys.eng.shizuoka.ac.jp

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

ションは困難なこともあり [3], [4], [5], 日本語を理解できない外国人と日本人との間で正確な情報共有を十分に行うことはできない。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る。現在、日本語を理解できない外国人の支援は医療通訳者が行っているが、医療通訳者の慢性的な人員不足や医療通訳者の身分保障などの問題が存在している [6]。

そこで、多言語対応の医療支援システムの開発が多く行われている [7], [8], [9], [10]。これらのシステムでは、正確な多言語変換が可能な用例対訳が用いられている。用例対訳は、用例を多言語に翻訳した多言語コーパスのことを指す。この用例対訳を用いて、医療従事者が質問したり患者が回答したりすることで、正確な多言語対話が可能となる。

また、これらのシステムへ用例対訳の提供を目的として、多言語用例対訳共有システムの開発が行われている [11], [12]。しかし、収集された用例対訳を自由対話的に使用するためには、(1) 質問に対して文脈が異なる回答候補が提供される場合がある、(2) 用例対訳と完全一致の文しか即時利用ができない、の2点の課題解決が必要である。

そこで本論文では、これらの問題を解決する応答用例対構築モデルの設計を行う。なお、応答用例対は質問と回答の用例対訳を1つずつ含んだものとする。詳しい定義は4.1節で行う。本モデルでは応答用例対を用いることで、質問に対する適切な回答の割合を高める。また、用例の類似文の収集により、用例対訳を用いた対話の自由度を高める。

なお、本論文では日本国内の外国人患者支援および海外の日本人患者支援の例を用いて議論を行う。ただし、応答用例対は用例対訳をもとに作成するため、対象となる言語が異なっても提案モデルの性能に影響しない。

2. 対話場面における用例対訳利用時の課題

本論文では、文献 [9], [10] のように、ある質問に対して利用者が自由文で回答を記入し、記入内容をもとに用例対訳を利用するシステムをPT利用システムとする。従来のPT利用システムの利用の流れを図1に示す。従来は図1(1), (2)で用例対訳検索、機械翻訳による翻訳を行い、結果を得ることで用例対訳などを利用している。また、図1(3)の不足用例対訳の作成依頼により、適切な用例対訳の収集が行われている。しかし、対話場面における用例対訳の利用時には2つの課題が存在している。課題の例を図2に、課題の詳細を以下にそれぞれ示す。本論文ではこれらの課題解決を行い、入力文と完全一致のものがあつた場合は機械翻訳のようにそのまま入力文を翻訳し、完全一致のものがあつない場合も検索結果に不適切な文の提示を減少させる、応答用例対構築モデルについて述べる。

課題1 文脈が異なる回答候補が提供される場合がある。

用例対訳の検索には、N-gram や tf-idf を用いて類似

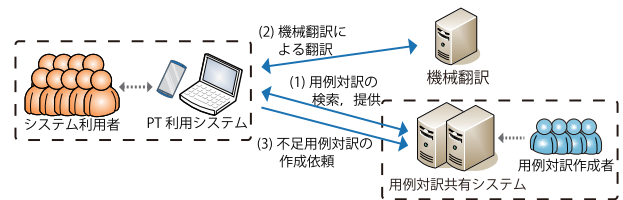


図1 従来の用例対訳の利用の流れ

Fig. 1 Flow of the use in existing parallel texts.

<課題1>

質問:「どのような症状がありますか?」



<課題2>

質問:「眠れないです」



図2 対話場面における用例対訳利用時の課題例

Fig. 2 Examples of problems that use parallel texts in dialogue scenarios.

度計算を行う手法が提案されている [13]。しかし、この手法では質問に対して不適切な用例対訳を提示する場合がある。図2では、質問「どのような症状がありますか?」に対して、検索文「気持ちが悪いです」を用いて回答を検索している。このとき、従来手法は文脈非依存のため、質問には不適切な「空気が悪いです」という用例対訳が提示される場合がある。

課題2 用例対訳と完全一致の文しか即時利用ができない。 用例対訳は静的に言葉を保持している性質上、1文字でも異なる文は多言語に変換できない。たとえば、図2のように用例対訳「よく眠れませんでした」を用いて「眠れないです」という言葉は多言語変換できない。このため、用例対訳の利用者は回答文を入力して用例対訳を検索した後、回答に適した用例対訳を選択するという複雑な作業を強いられている。用例対訳「眠れないです」を作成すれば解決できるが、用例対訳の作成には正確な翻訳が必要であるため、すべての表現の揺れに対応するためには非常に大きなコストがかかる。

3. 関連研究

入力された自由文をすべて多言語に翻訳が可能な機械翻訳を用いて、子供向けの機械翻訳 [14] やビジネスでの多言語支援 [15] など、様々な分野で多言語間コミュニケーション支援が行われている。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 [16]。また、機械翻訳はルールや統計データに基づいて動的な翻訳を行うため [17]、すべての対訳の正確性を確保することはできない。

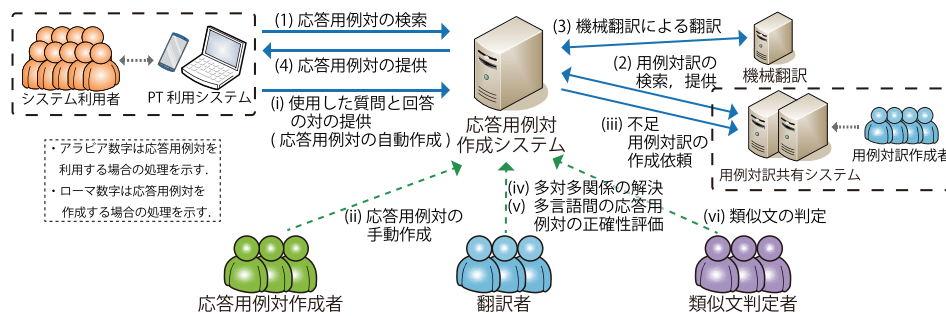


図 3 応答用例対構築モデル

Fig. 3 Paired dialogue parallel text composition model.

そこで現在、正確性が求められる分野では用例対訳による支援が多く行われている。用例対訳は、あらかじめ用例対訳を用いてフローを作成したシステム [7], [8] で多く用いられてきた。しかし、フロー型では選択の自由度が低いため、自由文をもとに用例対訳を検索可能としたシステム (PT 利用システム) [9], [10] が開発されている。用例対訳はあらかじめ翻訳された文であるため、入力文に適した用例が存在しない可能性がある。このため、PT 利用システムでは用例対訳と機械翻訳を併用し、入力文が用例対訳に存在しない場合は機械翻訳を使用することで入力文の多言語化を行っている。また、N-gram などを用いて検索性を高めているが [13], 前述の 2 つの課題解決を行えていない。本論文では、応答用例対と類似文の収集をそれぞれ行い、翻訳精度の低い機械翻訳の利用頻度を下げつつ、正確かつ自由度を高めた対話支援により、これらの問題解決を行う。

なお、応答用例対に近い隣接応答対という概念がすでに言語グリッドプロジェクト [18], [19] で利用されている。隣接応答対は 1 つの質問の用例対訳に対して、回答の用例対訳が複数結合されたものを指し、フロー型のシステム (文献 [7], [8] など) における用例対訳の管理を目的としている。また、隣接応答対を作成するシステムの構築も行われている [20]*1。しかし、隣接応答対は PT 利用システムには対応していないため、前述の課題解決が行われていない。このため、本研究では質問と回答が 1 つずつ含まれる応答用例対を用いて課題の解決を行う。また、応答用例対に類似文を付与することで、より自由度の高い対話を目指す。

本研究では人を計算資源の 1 つととらえ、人と計算機を組み合わせることで課題解決を目指す考え方であるヒューマンコンピュータシオン (以下 HC) [21] を活用する。従来の HC として、クラウドソーシング [22] を活用した翻訳手法 [23] や多言語対の正確性評価手法 [24] がある。これらの研究では各課題の解決手法が提案されているが、本論文では HC を活用するための枠組みを提案し有用性確認を行う。このことで、応答用例対の収集を行う際に必要な人員やその知識、技術的な問題点の確認などを行う。提案モデルの実装

時には従来研究の各手法が応用可能であると考えられる。

なお、本研究は検索や質問応答などと類似しているが、多様性を求める点と高い正確性を担保する点が相違点として存在する。病気などの表現は様々なニュアンスなど多様性が重要であるが、検索などは 1 つ (もしくは数個) の正解発見が課題として設定されることが多い。また、正確性の担保に重点を置いているものは少ない。このため本研究では正確性を担保しつつ、多様性を確保することを目指す。

4. 応答用例対構築モデル

4.1 本モデルの概要

図 3 に、2 章の課題解決を目指した応答用例対構築モデルを示す。本モデルでは、知識を持つ各ユーザをモデルに組み込む HC を用いることで応答用例対の構築を行う。以降、PT 利用システムからの応答用例対依頼時に行う処理をリアルタイム (図中の実線)、それ以外の処理を非リアルタイム (図中の破線) とする。

本論文で扱う応答用例対は、1 個の質問 (用例対訳)、1 個の回答 (用例対訳)、0 個以上の類似文から構成されるものとする。類似文については 4.4 節で詳しく述べる。ただし、本モデルでは多言語化されていなかったり、多対多関係が解決されていなかったりするものも取り扱う。本論文では質問と回答が対となっているものを「質問と回答の対」と表記する。これらの詳細は 4.5 節で述べる。また、本モデルで使用する質問と回答の対の例を図 4 に示す。なお、図 4 (4) は隣接応答対 (文献 [18] などで使用) の例である。

本モデルで用いる応答用例対は、フロー型のシステムで使用されていた隣接応答対を、PT 利用システムで利用可能な構造に改良し、知識を蓄積可能とした。質問に適切な回答を蓄積、利用することで、仮に別の質問に対して同じ検索を行った際にも文脈に依存した回答を提示可能とし、相対的に文脈非依存の回答候補を減らすことで課題 1 の解決を目指す。また、従来の隣接応答対には存在していなかった類似文の付与を可能とした。擬似的な用例対訳を増やすことで課題 2 の解決を目指す。

なお、本モデルでは応答用例対内の用例対訳を ID 管理するものとする。これは、用例対訳共有システムで用例対

*1 文献 [20] では「応答用例対」の作成を行っているが、作成されたものは隣接応答対と同じ構造である。

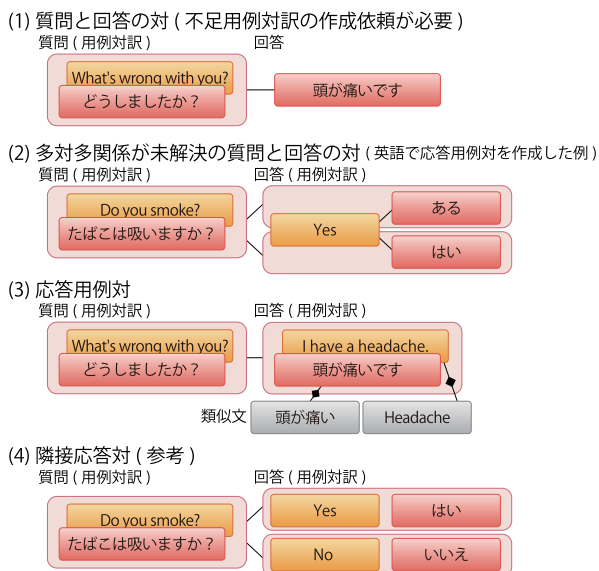


図 4 モデルで扱う質問と回答の対の例

Fig. 4 Examples of a question and answer combination in the model.

訳の修正が行われる可能性があるためである。

4.2 応答用例対の作成

本モデルでは、図 3(i) (自動作成) と図 3(ii) (手動作成) の 2つの方法で図 4(3) に示す応答用例対の作成を行う。ただし、本節の流れで図 4(1) や (2) が作成される場合もある。以下に両方法の作成の流れについて述べる。

自動作成 (リアルタイム) 図 3(i) では、利用システムで使用された質問と回答の対から、応答用例対を自動的に作成する。実際に使用された質問と回答の対を用いて適切な応答用例対の提供を目指す。なお、同じ応答用例対が作成済みの場合は利用回数のみ記録する。

手動作成 (非リアルタイム) 図 3(ii) では、利用が想定される応答用例対をあらかじめ手動で作成する。その際、複数の用例対共有システムを横断的に利用することで、多様な応答用例対の作成を可能とする。主に PT 利用システムの管理者の利用を想定している。

自動作成が行われた際の応答用例対作成の流れを図 5 に示す。図 5 に示すとおり、本モデルでは作成された質問と回答の対によって行う処理が異なる。応答用例対 (図 4(3)) が作成された場合は図 3(v) で多言語間の正確性評価を行い、PT 利用システムへ提供されるが、図 4(1) や (2) の場合は 4.5 節の処理が行われる。また、類似文の判定 (図 5(vi), 4.4 節) は、図 5(iii)~(v) と同時に行われる場合もある。なお、手動作成の場合は図 5(1)~(4) の処理がなくなり、図 5(i) の代わりに手動作成が行われることとなる。

4.3 応答用例対の提供 (リアルタイム)

本モデルでは図 3(4) で応答用例対の提供を行う。その

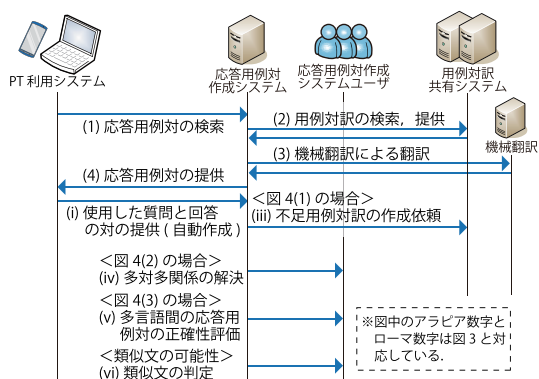


図 5 応答用例対作成の流れ

Fig. 5 Flow of creating dialogue parallel texts.

際、応答用例対を利用可能とするために、PT 利用システムと用例対共有システムとの間を応答用例対作成システムが仲介している (図 3(1), (2))。また、単純な類似テキスト検索であった従来手法とは異なり、本モデルでは、もとなつた質問文を含む応答用例対の情報と類似テキスト検索結果の情報の両者を用いる。文脈に依存した検索を可能とすることで、課題 1 の解決を目指す。

なお、本モデルでは検索文が未入力の場合でも検索を可能とする。その際は、もとなつた質問文が含まれる応答用例対のうち、過去の利用回数の降順で提示する。また、検索文が入力されていた場合は下記の順に提示する。

- (1) 応答用例対が作成済み、かつ類似用例対検索で類似している用例
- (2) 類似用例対検索で類似している用例
- (3) 応答用例対が作成済みの用例

4.4 類似文の収集 (非リアルタイム)

本論文で述べる類似文は、回答の用例対訳に類似した文とする。類似文は課題 2 の解決を目指している。図 3 のモデルでは (1), (i), (vi) が関係する。既存の PT 利用システム [9], [10] では、入力文をもとに用例対訳の検索を行っている。このとき、入力文と完全に一致する用例対訳が存在しない場合、利用者は入力文と類似した用例対訳を使用することとなる。本モデルでは、入力文を用例対訳の類似文として保存し、後の応答用例対の利用時に使用する。

類似文の収集例を図 6 に示す。また、詳細を以下に示す。

- (1) PT 利用システム利用者が「どうしましたか?」の回答として「頭が痛い」という文を入力する。その後、類似用例対訳として提示された「頭が痛いです - I have a headache.」を利用者が選択した場合、本システムにそれらの情報が提供される。この時点で、質問と回答は応答用例対として自動登録される。
- (2) 類似文判定者は「頭が痛い」が「頭が痛いです」と類似しているか、かつ質問文である「どうしましたか?」の回答として適切かを判定する。類似かつ適切と判定

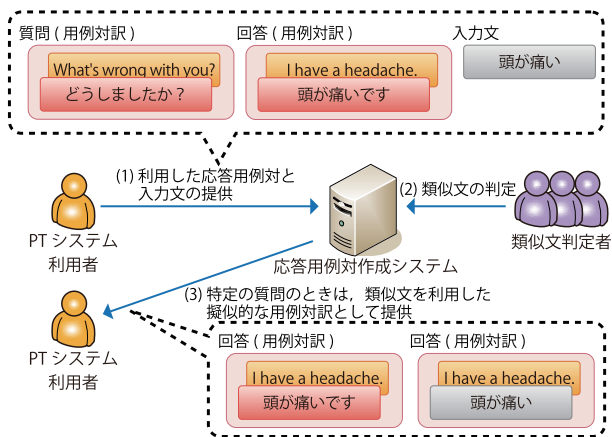


図 6 類似文の収集の流れ
Fig. 6 Flow of collecting similar texts.

された場合、類似文として応答用例対に追加される。
(3) 次回以降、「どうしましたか?」の回答として「頭が痛い」が入力されたとする。このとき、もとの用例対訳の「頭が痛いです」と類似文「頭が痛い」とを入れ替えて「頭が痛い - I have a headache.」という擬似的な用例対訳が利用可能となる。

類似文を使うことで、1つの用例対訳が複数の文から利用可能となる。このため、従来手法よりも用例対訳の自由度が高くなる。類似文は、次節で述べる用例対訳の多対多関係とは異なり、用例対訳とは別に文を保持している。なお、類似文は完全一致の入力文のときのみ利用可能とする。

4.5 質問と回答の対の応答用例対化 (非リアルタイム)

前節まで、各言語が1対1に対応した用例対訳を用いた応答用例対について述べた。しかし実際には、用例対訳が存在しない文や、多対多関係にある用例対訳を利用する場合が考えられる。本節ではこれらの解決について述べる。

4.5.1 不足用例対訳の作成依頼

PT 利用システムでは用例対訳のみではなく、対訳が正確と担保できない機械翻訳 (図 3(3)) を用いたものが存在している。このため、図 3(i) で機械翻訳を使用した文が提供された場合、用例対訳のときのように応答用例対を作成することができない。図 4(1) は、対訳が不足した状態の質問と回答の対の例である。本モデルでは図 4(1) のような質問と回答の対が作成された場合、用例対訳共有システムに不足用例対訳の作成依頼を行う (図 3(iii))。用例対訳共有システム内で翻訳者が用例を翻訳し、正確性評価を行うことで、応答用例対 (図 4(3)) として利用可能となる。なお、従来の用例対訳利用環境でも同様の機能を有している (図 1(2))。しかし、本モデルではあらかじめ質問と回答の対として保存しているため、用例対訳が作成された直後に応答用例対として提供可能となるという特徴がある。

4.5.2 多対多関係の解決

用例対訳には1対多、多対多関係のものが含まれている

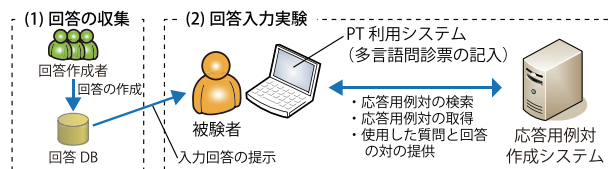


図 7 本実験の流れ
Fig. 7 Flow of the experiment.

場合がある [25]。多対多関係にある用例対訳を含む質問と回答の対は、適切な言葉を自動的に選択することが難しい。図 4(2) は、多対多関係を持つ用例対訳を含む質問と回答の対の例である。この例の場合、1個の質問と2個の回答が存在していることとなる。適切な言葉を自動的に選択することは難しいため、本モデルでは翻訳者が適切な対訳の選択を行う (図 3(iv))。図 4(2) の場合、翻訳者が「たばこは吸いますか? - はい」を適切であると判断することで、応答用例対 (図 4(3)) の提供が可能となる。

なお、多対多関係が未解決の質問と回答の対は広義の応答用例対として「たばこは吸いますか? - はい/ある」のように複数の言葉を併記して利用することも可能である。

5. 応答用例対収集実験

前章で述べた、応答用例対構築モデルを適用したシステムによる応答用例対収集実験を行った。本実験は、用例対訳を用いた問診票の入力システムを想定して行った。目的は、本モデルを PT 利用システムに適用した場合の、応答用例対と類似文の効果の検証である。本実験の流れを図 7 に示す。本実験は、使用する回答を収集した後 (図 7(1))、それらを入力する (図 7(2)) という手順で行った。詳しくは後述するが、10名以上の利用者がランダムに症状を入力することを別条件で4回行うことを想定した実験を行った。また、本モデルでは利用者 (単言語話者) への適切な質問と回答の対の提供を目的としている。さらに、多言語に関する課題は文献 [11], [20] で検証済みであり、知見が利用可能であるため、本実験は単言語で行った。

5.1 回答の収集

本実験ではまず、問診票に記入する回答の収集 (図 7(1)) を行った。回答作成者は情報系の大学生、大学院生 11 名である。回答は問診票を想定し、質問文に対して回答する形で収集した。使用した質問文を表 1 に示す。なお、表 1 は医療機関で使用されている問診票を使用した、文献 [9] をもとに作成している。回答文作成の行いやすさを考慮し、多種の回答が存在している質問 ID1 と 2 は 10~20 件、それ以外は 5~10 件の入力をそれぞれ依頼した。なお、自身の症状だけで回答が作成できない場合は、Web 上での症状の調査を許可している。本実験では、5つの質問に対して合計 507 文の回答 (重複、類似文を含む) を収集した。

表 1 問診票の質問項目

Table 1 Question items on the interview sheet.

ID	質問文	入力依頼	回答
1	(内科) どのような症状がありますか？	10~20	151
2	(外科) どのような症状がありますか？	10~20	138
3	薬や食べ物のアレルギーは何ですか？	5~10	78
4	今飲んでいる薬は何ですか？	5~10	69
5	飲んでいるお酒は何ですか？	5~10	71

・単位は文である。
 ・入力依頼は、回答作成者に依頼した回答文の最低数と最大数を示す。
 ・回答は、回答作成者が入力した回答の数を示す。

5.2 回答入力実験

5.2.1 実験概要

本節では多言語問診票の回答入力実験 (図 7(2)) について述べる。本節では応答用例対を未使用の従来システムと、本モデルを適用した提案システムの比較実験を行う。従来手法は類似文検索による用例対訳の提示のみを行い、提案手法は応答用例対や類似文の作成や利用も合わせて行った。

本実験の被験者は情報系の大学院生で、5.1 節の回答作成者 3 名を含む 4 名で行った。被験者は患者を想定しており、図 3 の PT 利用システムの利用者にあたる。被験者には 5.1 節で収集した回答を提示し、その症状の患者であるという想定で入力を依頼した。なお、順序効果を考慮して被験者ごとにシステムの回答の順を入れ替えている。また、被験者は 1 つのデータセット中の同じ言葉を、従来システムと提案システムのそれぞれに入力を行っている。ただし、作成される応答用例対や提示内容はシステムや被験者ごとに独立している。

本実験では、図 3 中の用例対訳共有システムとして、TackPad [11] と多言語問診票作成システム [9] を用いた。実験では、合計 5,122 件 (TackPad : 4,658 件, 問診票 : 464 件) の日本語用例が利用可能であった。

5.2.2 実験手順

実験の手順を以下に示す。被験者には以下の手順を繰り返すことで、5.1 節で収集した回答に対して適切な用例対訳や機械翻訳の選択を依頼した。なお、5.1 節の回答の収集に参加した被験者に対しては、実験条件を統一するために各被験者自身が作成した回答を含まないデータセットを用いた。このため、被験者ごとに回答数が異なっている。

(1) ポップアップで 5.1 節で収集した回答を提示。

被験者には、ポップアップで提示された症状を持つ患者としてシステムに入力することを依頼した。

(2) 回答の入力画面を提示、被験者が回答を記入。

入力画面例を図 8 に示す。被験者は図 8(1) の質問文に対して図 8(2) に回答を記入する。また、症状を忘れたときのために、図 8(3) のボタンをクリックすると再度ポップアップで表示する機能を用意した。

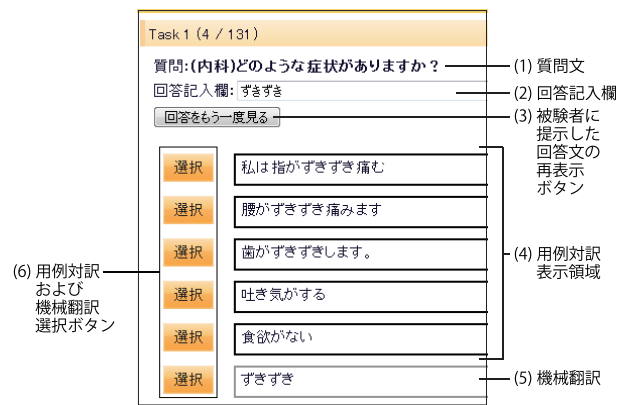


図 8 実験用入力インターフェース

Fig. 8 Screenshot of the input interface used in the experiment.

(3) システムが入力された文をもとに回答候補を検索。

1 秒おきに入力文をもとに用例対訳を検索し、最大 5 件の検索結果を図 8(4) に提示する。なお、提案システム (応答用例対利用) の場合は、文の入力前に回答候補を提示する場合がある。また、機械翻訳を想定し、入力文と同一の文を図 8(5) に提示した。被験者には図 8(5) は多言語変換の精度が低いことを伝えている。

(4) 適切な症状の選択。

図 8(6) の選択ボタンを用いて適切な回答を選択する。なお本モデルでは、用例対訳を選択した場合、応答用例対の自動作成と類似文の作成を、機械翻訳を選択した場合は、不足用例対訳の作成をそれぞれ行う。

5.2.3 モデルの適用条件

本実験でモデルを適用する際に条件を付けた項目を以下に示す。括弧内は図 3 中のローマ数字と対応している。

応答用例対の手動作成 (ii) 自動作成の効果を検証するため未適用。適用した場合は、あらかじめ応答用例対が存在している状態となるため、本実験の序盤を除いたものと同等の結果になると考えられる。

不足用例対訳の作成 (iii) 作成される対訳の正確性は本実験内容には影響しないためシミュレーションを用いた。機械翻訳が選択され、ポップアップでシステムが提示した文と被験者が入力した文が一致した場合、入力文が用例対訳化されたとした。

多対多関係の解決 (iv)、応答用例対の正確性評価 (v)

単言語での実験のため未実施。用例対訳の多言語に起因する問題は、文献 [11], [20] で検証済みであり、知見が利用可能である。

類似文の判定 (vi) 非リアルタイム処理であり本実験への適用が困難であったためシミュレーションを用いた。用例対訳が選択され、選択した用例対訳と被験者の入力文が一致せず、かつ、ポップアップでシステムが提示した文と入力した文が一致した場合、入力文を選択した用例対訳の類似文として登録した。類似文の可能性が高いもののみを採用する基準のため、実験で作成

表 2 応答用例対の提示割合

Table 2 Rate of displayed dialogue parallel texts.

ID	提案手法			従来手法			割合 の差
	DPT	PT	割合	DPT	PT	割合	
1	1,209	1,930	62.6%	369	1,506	24.5%	38.1
2	1,000	1,789	55.9%	268	1,475	18.2%	37.7
3	968	1,140	84.9%	193	524	36.8%	48.1
4	661	1,024	64.6%	141	484	29.1%	35.4
5	731	890	82.1%	190	276	68.8%	13.3
計	4,569	6,773	67.5%	1,161	4,265	27.2%	40.2

- ・PT は提示された用例対訳数を，DPT は PT のうち応答用例対であったものの数をそれぞれ示す。単位は対および文である。
- ・ID は表 1 と対応する。
- ・従来手法は提案手法と同様の手順で応答用例対を作成したと仮定して算出している。

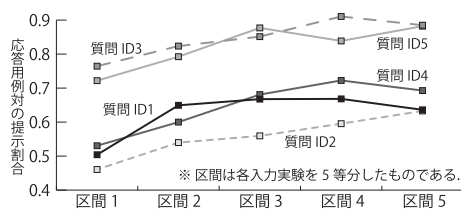


図 9 応答用例対提示割合の推移

Fig. 9 Process of rate of displayed dialogue parallel texts.

される類似文は実環境よりも少なくなると考えられる。

6. 実験結果と考察

本章では，実験の結果と考察について述べる。なお，本章で述べる正確性の評価は，前章の収集や実験に参加していない情報系の大学生・大学院生 3 名が行っている。

6.1 自動作成された応答用例対の利用効果

表 2 に，回答が選択された時点で提示されていた用例対訳のうち，応答用例対を利用したものの割合を示す。なお，従来手法は応答用例対を作成していないが，提案手法の応答用例対の自動作成と同様の応答用例対を仮定して算出している。表 2 より，提案手法は従来手法よりも約 40 ポイント多く応答用例対を提示していることが分かる。応答用例対は被験者が選択した用例対訳をもとに作成している。このため，応答用例対を用いて提示された回答候補は適切であると考えられ，提案手法は文脈が異なる可能性のある用例対訳の提示割合を減少可能であることが分かる。また，提案手法での応答用例対提示割合の推移を図 9 に示す。図 9 より，入力を進めると応答用例対の提示割合が増加することが分かる。なお，一部区間で減少しているが，応答用例対が未作成の回答入力，他の区間より多かったためであると考えられる。これらのことから，システムを繰り返し使用することで，課題 1（文脈が異なる回答候補の提供）の発生頻度の減少が可能になったと考えられる。なお，用例対訳数や応答用例対の提示割合が大きくなる要

表 3 登録類似文数と利用回数

Table 3 Numbers of similar sentences registered and those used.

被験者	A	B	C	D
入力回答文数	437 文	437 文	461 文	507 文
登録類似文数	46 文	19 文	20 文	17 文
類似文利用回数 (延べ)	10 文	9 文	9 文	13 文

因の 1 つとして，提案手法は検索前に応答用例対が提示されるなど，応答用例対を活用する設計としたことがあげられるが，この設計により，必ず検索が必要な従来手法に対し，提案手法は検索なしで回答を選択可能となっている*2。

作成された応答用例対の正確性を 3 名の評価者が確認したところ，内科の症状の回答に「鼻水が出る*3」「肩がこる*4」など，直接関係のない応答用例対が含まれていた。また，アレルギーを問う質問への回答に「アトピー性皮膚炎」が含まれていた。ただし，前者は内科の症状に付随した症状である場合があるため，後者は質問への回答としては不自然であるが，アレルギー名を記入しているため，大きな問題にはならないと考えられる。ただし，自動作成された応答用例対は正確性の担保を行っていないため，本モデルの手動作成 (図 3(ii)) を適用し，正確性の担保を行う必要があると考えられる。なお，診療科に直接関係のない応答用例対を除くと，すべての応答用例対は正確であった。

6.2 類似文の利用効果

表 3 に被験者別の登録類似文の数と利用回数を示す。本実験のシミュレーションでは類似文の可能性が高い文のみを採用するという基準を用いたが，類似文の登録や利用が行われていたことが確認された。実際には入力文の大半が類似文の候補となるため，より大きな効果が生まれると考えられる。類似文を利用することで，利用可能な用例対訳が擬似的に増えるため，本モデルにより課題 2（用例対訳と完全一致の文しか即時利用ができない）の発生頻度を下げることが可能であると考えられる。また，本モデルで作成された類似文は，従来手法とは異なり多言語翻訳が不要であるため，用例対訳作成のコスト軽減や対訳作成作業の時間を待たずに多言語支援が可能となるという利点も存在している。

本実験で登録された類似文を表 4 に示す。表 4(1)，(2) は正確に登録された類似文の例である。3 名の評価者が調べた結果，本実験では 102 文のうち 83.3% の類似文は正確であった。なお，前述の割合は過半数が正確と判定した割合である。表 4(3)，(4) は不適切な例である。表 4(3) の類似文には，用例対訳に含まれている部位の情報が欠落している。また，表 4(4) は被験者の勘違いに起因した例で

*2 本実験では，全体の 17.2% の回答が検索なしで選択されている。

*3 単独の症状の場合，耳鼻咽喉科の症状

*4 一般的には整形外科や神経内科の症状

表 4 実験で登録された類似文の例

Table 4 Examples of the similar sentences used in the experiment.

	元の用例対訳	登録された類似文
(1)	立ちくらみがします	立ちくらみが起きる
(2)	せきが出る	咳
(3)	皮膚が乾燥する	乾燥する
(4)	脊髄損傷	脊椎損傷
(5)	ピーナッツアレルギー	ピーナッツ

ある*5。このように、類似文候補のすべてを正確に自動判定することは困難であるため、提案モデルのように人手による正確性判定(図3(vi))が必要であると考えられる。

また、表4(5)は、一般的には類似文とはいえないが、質問が「薬や食べ物のアレルギーは何ですか?」の応答用例対として登録されていたため、適切な類似文となる。しかし、別の質問(「昨日は何を食べましたか?」など)に対する回答としては不適切となることが考えられる。このため、類似文の登録は用例対訳の形*6ではなく、応答用例対の形*7で行う必要があることが分かる。

なお、類似文の判定は単純な作業であるため、クラウドソーシング上の複数のWorkerによる判定でも正確性の担保は可能である[21]。Workerの判断が分かれた場合は用例対訳共有システムの作業員などによる判定が必要となるが、既存の作業員の負担を軽減しつつ、大量の類似文の判定が可能になると考えられる。

6.3 応答用例対の自動作成の効果

本節では、既存の応答用例対として多言語問診票作成システム[9]の応答用例対を用いて、自動作成効果の検証を行う。なお、本実験では文献[9]の用例対訳のみを使用しているため、文献[9]の応答用例対は実験に影響を与えていない。また、本実験で使用した質問文は、Yes/Noなどの単純な回答となる質問や、回答欄が自由記述のみのもの、診療科で異なる症状を尋ねる質問などを除くと、実際の間診票をもとに作成されている文献[9]の質問文をほぼ網羅している*8。

比較結果を表5に示す。なお、質問IDの1, 3, 5のみ比較が可能であった*9。表5より、既存応答用例対(表5(A))の大半が実験中に作成されており(表5(B))、自動作成

*5 脊椎はいわゆる背骨を指す。また、脊髄は背骨を通る神経を指す。
 *6 例として、用例対訳「ピーナッツアレルギー」に対して、類似文「ピーナッツ」を保存。
 *7 例として、応答用例対「薬や食べ物のアレルギーは何ですか? - ピーナッツアレルギー」に対して、類似文「ピーナッツ」を保存。
 *8 文献[9]の過去の病歴を問う質問は用いていない。これは、医療機関が患者に求める病歴はほぼ決まっており、動的に候補が増える自動作成は適さないと考えられたためである。
 *9 文献[9]では詳細な診療科名(整形外科)が用いられていたが、質問2は5.1節の回答作成者の症状記述を容易にするために広い意味の外科を用いた。また、質問4と対応する文献[9]の回答は自由記述のみであった。

表 5 既存の応答用例対との比較

Table 5 Comparison between existing dialogue parallel texts and those created.

質問 ID	1	3	5
既存応答用例対 (A)	23	4	4
実験中に作成された既存応答用例対 (類似している場合も含む) (B)	19	3	4
自動作成応答用例対の網羅率 (B/A)	82.6%	75.0%	100.0%
収集された応答用例対 (類似文含む)	102	32	20
既存応答用例対に存在していなかった 収集応答用例対 (類似文除く) (C)	24	22	12

・質問IDは表1と対応する。
 ・単位は対である。
 ・既存応答用例対は、多言語問診票作成システム[9]を用いた。

応答用例対の網羅率も高かったことが分かる(表5(B/A))。既存応答用例対は、実際に医療機関で使用されている問診票をもとに作成しており、実験中に使用された用例対訳は適切なものが含まれていたことが分かる。また、表5(C)より、既存応答用例対(表5(A))以外のもものが多く作成されていたことが分かる。このことから、本モデルを使用することで多様な応答用例対の自動的な収集が可能であり、応答用例対の自動作成のみでも既存応答用例対と同等のもものが作成できたことが分かる。なお、自動的に収集した応答用例対は、図3(4)の応答用例対の提供に用いることができるほかに、応答用例対を使用している別のシステムで、同一の質問を持つ応答用例対の利用も可能となるという利点がある。

6.4 本モデルの限界

本モデルでは、不正確な応答用例対が作成され、さらにその応答用例対が多く利用された場合*10は候補として出現しやすくなる。また、本モデルはPT利用システムが利用されることを前提としている。これらのことから、応答用例対の手動作成や類似文の収集をあらかじめ一定程度行ってから実運用を行う必要があると考えられる。また、似た文で両方とも適した応答用例対については、アノテーションを付けるなど別手法が必要になると考えられる。

類似した用例対訳*11が多数作成された場合、利用者に提示される用例対訳の多くが類似した文となる可能性がある*12。類似用例対訳の提示は利用者にとって必ずしも適切とはいえないため、用例対訳間の類似性を保持し、利用可能とする仕組みを今後検討する必要があると考えられる。

7. おわりに

本論文では、正確な用例対訳を用いた自由度の高い対話

*10 「脊髄」と「脊椎」など、違いが分かりにくい回答の場合、誤用が多発する可能性がある。
 *11 類似文とは異なり、用例対訳(対訳作成済み)どうしの類似。
 *12 類似文は完全一致のときのみ使用されるため、多く存在していてもこのような問題は発生しない。

支援のための、応答用例対構築モデルの提案を行った。本モデルでは、従来の対話形式の用例対訳利用環境に存在していた、(1) 文脈が異なる回答候補が提供される場合がある、(2) 用例対訳と完全一致の文しか即時利用ができない、の各課題解決を目的としている。

本研究の貢献は次の3つにまとめられる。

- (1) 従来の対話型用例対訳利用システムにおける問題解決を目的とした、応答用例対を用いたモデルを提案した。
- (2) 応答用例対の自動収集により、適切な回答選択の支援の可能性を示した。
- (3) 応答用例対への類似文の付与により、用例対訳を用いた自由度の高い対話の可能性を示した。

今後は、実運用システムへの適用を行うことで、応答用例対や類似文の収集や提供が円滑に行えるか調査を行う。

謝辞 本研究は JSPS 科研費 26730105 の助成を受けた。

参考文献

- [1] 法務省：平成 25 年における外国人入国者数及び日本人入国者数について（確定値），法務省（オンライン），入手先 <http://www.moj.go.jp/nyuukokukanri/kouhou/nyuukokukanri04.00039.html>（参照 2014-04-01）。
- [2] 総務省：多文化共生の推進に関する研究会報告書，総務省（オンライン），入手先 http://www.soumu.go.jp/kokusai/pdf/sonota_b5.pdf（参照 2014-04-01）。
- [3] 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).
- [4] 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).
- [5] Kim, K.J. and Bonk, C.J.: Cross-Cultural Comparisons of Online Collaboration, *Journal of Computer Mediated Communication*, Vol.8, No.1 (2002).
- [6] 高嶋愛里：在日外国人支援活動：京都における「医療通訳システムモデル事業」，国際保健支援会 2 (2005)。
- [7] 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)。
- [8] 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)。
- [9] 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，Vol.54, No.1, pp.256-265 (2013)。
- [10] 尾崎 俊，松延拓生，吉野 孝，重野亜久里：携帯型多言語間医療対話支援システムの開発と評価，電子情報通信学会技術研究報告，Vol.AI2010-47, pp.19-24 (2011)。
- [11] 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，CDS, Vol.2, No.3, pp.22-33 (2012)。
- [12] Bond, F., Nichols, E., Appling, D.S., et al.: Improving Statistical Machine Translation by Paraphrasing the Training Data, *Proc. IWSLT 2008*, pp.150-157 (2008)。
- [13] 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，電子情報通信学会技術研究報告，Vol.AI2008-52, pp.43-48 (2009)。
- [14] Matsuda, M. and Kitamura, Y.: Development of Machine Translation System for Japanese Children, *Proc. IWIC'09*, pp.269-271 (2009)。
- [15] 照井賢治，菱山玲子：多言語ケースメソッドシステムによる異文化分析，電子情報通信学会技術研究報告，Vol.AI2012-35, pp.79-84 (2013)。
- [16] 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)。
- [17] 塚田 元，渡辺太郎，鈴木 潤ほか：統計的機械翻訳，NTT 技術ジャーナル，Vol.19, No.6, pp.23-25 (2007)。
- [18] Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration, *Proc. SAINT 2006*, pp.96-100 (2006)。
- [19] Sakai, S., Gotou, M., Tanaka, M., et al.: Language Grid Association: Action Research on Supporting the Multi-cultural Society, *Proc. ICKS 2008*, pp.55-60 (2008)。
- [20] 福島 拓，吉野 孝，田淵裕章，北村泰彦：多言語用例対訳を用いたコミュニケーションのための応答用例対作成システムの開発，情報処理学会，DICOMO2009 シンポジウム，pp.1612-1618 (2009)。
- [21] 鹿島久嗣，馬場雪乃：ヒューマンコンピューテーション概説，人工知能，Vol.29, No.1, pp.4-11 (2014)。
- [22] Doan, A., Ramakrishnan, R. and Halevy, A.Y.: Crowdsourcing systems on the World-Wide Web, *Comm. ACM*, Vol.54, No.4, pp.86-96 (2011)。
- [23] Negri, M. and Mehdad, Y.: Creating a Bi-lingual Entailment Corpus through Translations with Mechanical Turk: \$100 for a 10-day Rush, *Proc. NAACL HLT 2010*, pp.212-216 (2010)。
- [24] Callison-Burch, C.: Fast, Cheap, and Creative: Evaluating Translation Quality Using Amazon's Mechanical Turk, *Proc. EMNLP 2009*, pp.286-295 (2009)。
- [25] 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，TOD, Vol.5, No.2, pp.1-10 (2012)。



福島 拓 (正会員)

1986 年生。2008 年和歌山大学システム工学部中退。2013 年同大学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，静岡大学大学院工学研究科助教。CSCW の研究に従事。



吉野 孝 (正会員)

1969 年生。1992 年鹿児島大学工学部卒業。1994 年同大学大学院工学研究科修士課程修了。博士（情報科学）。現在，和歌山大学システム工学部教授。CSCW, HCI の研究に従事。