

クラウドソーシング上の単言語話者による 用例対訳作成手法への折り返し翻訳利用の提案

山本 里美^{1,a)} 福島 拓^{2,b)} 吉野 孝^{1,c)}

概要：現在、グローバル化によって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の共有は困難である。医療分野などの正確な情報の共有が求められる現場では、正確性が確保された多言語の対である用例対訳が多く用いられている。しかし、医療現場で必要とされる用例対訳の数は多く、十分な評価を得た用例対訳を収集することは困難であると考えられる。我々は、現在、機械翻訳文の評価と訂正をクラウドソーシングにおいて依頼することで、単言語話者であっても用例対訳作成を行うことができる手法の研究を行っている。これまでの実験より、クラウドソーシング上の単言語話者と機械翻訳を組み合わせることによって、正確性の高い用例対訳作成が行えることが分かった。しかし、正確性の高い対訳が作成できた用例は一部であり、多くの用例では対訳の作成や正確性評価が正しく行われなかった。そこで本稿では、機械翻訳文の訂正精度向上のために、従来のクラウドソーシング上の単言語話者による用例対訳作成に、折り返し翻訳を用いた機械翻訳適応文作成タスクを追加した手法を提案し、機械翻訳適応文作成タスクにおける評価実験を行った。

1. はじめに

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は増加している。それに伴い、多言語間コミュニケーション支援のための研究やシステムの開発が多く行われている。正確な情報の共有が重要となる医療分野などでは、正確性の確保された多言語の対である用例対訳が多く用いられている。しかし、正確な情報の共有のために必要とされる用例の数は多く、十分な数の用例対訳の収集は困難である [1]。

我々は、現在、機械翻訳文の評価と訂正を行うことで、単言語話者であっても用例対訳を作成できる手法の研究を行っている。これまで、クラウドソーシング上の作業者に会話文の形式で機械翻訳文を提示することで、用例対訳や応答用例対^{*1}を作成できることが分かった [2]。しかし、文献 [2] の手法では、クラウドソーシングの作業者によって作成された機械翻訳文の訂正文 380 文のうち、専門家が正確な対訳と評価した文は 110 文 (約 29%) だった。また、110 文のうち、クラウドソーシングを用い

た正確性評価で正確だと判断された文は 14 文 (約 13%) だった。このことから、クラウドソーシング上の単言語話者によって用例対訳を作成することは可能だが、多くの用例に対して対訳を作成することはできず、また、正確性評価ではほとんどの正確な対訳を抽出することができないことが分かった。

本稿では、クラウドソーシング上の単言語話者による用例対訳作成手法において、折り返し翻訳を用いた機械翻訳適応文作成タスクを追加した手法を提案し、評価実験を行った。これは、折り返し翻訳が適切に行える文であれば、機械翻訳文の評価と訂正を行う際に、作業者が文の意図を推測しやすくなり、正しい訂正が行える可能性があると考えたためである。これにより、単言語話者によって作成される訂正文の訂正精度向上を目指す。

2. 関連研究

現在、クラウドソーシングを用いて多言語データを収集する研究が多く行われている。多言語テキストの正確性評価をクラウドソーシングの作業者に依頼する研究や [3]、翻訳対象の文やクラウドソーシング作業者の特徴をもとに、翻訳文を分析し、品質の良い翻訳文を作成する研究 [4]、クラウドソーシングと機械翻訳を併用してより品質の高い対訳コーパスや翻訳結果を取得する研究 [5]、[6] などがある。これらの研究では、クラウドソーシングの作業者を 2 言語以上を理解することができることを前提としている。しかし、クラウドソーシング上の

¹ 和歌山大学
Wakayama University, Wakayama 640-8510, Japan

² 静岡大学
Shizuoka University Johoku 3-5-1, Naka-ku, Hamamatsu, 432-8561 Japan

a) s165058@center.wakayama-u.ac.jp

b) fukushima.taku@shizuoka.ac.jp

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

*1 質問とその回答の対、またその類似文からなる用例対訳のことであり、会話の支援に用いられる。

用例対訳作成の手順

日本語から英語の対訳を作成する場合の例

タスク画面の例

Step1 機械翻訳適応文の作成

翻訳対象文と同じ意味を表し、折り返し翻訳文との一致度が高い文を作成する



Step2 翻訳対象言語へ機械翻訳

機械翻訳適応文を翻訳対象言語へ翻訳する



Step3 機械翻訳文の評価と訂正

機械翻訳文を5段階で評価し、訂正文を作成する



Step4 用例対訳準候補文の評価

用例対訳準候補文(機械翻訳やその訂正文)から、提示された条件に適切な文を選択する



Step5 翻訳対象文の言語へ機械翻訳

Step4で選択された文を翻訳対象文の言語へ翻訳する



Step6 用例対訳候補文の正確性評価

翻訳対象文とStep5の機械翻訳文を比較し、同じ意味を示すものを選択する



翻訳対象文

ちょっと熱が出てまして…

機械翻訳適応文

私はすこし熱があります。

機械翻訳文

I have a little fever.

訂正文

I have a slight fever.

訂正文

It has a little fever.

用例対訳準候補文

✗ I have a little fever.

○ I have a slight fever.

✗ It has a little fever.

用例対訳候補文

I have a slight fever.

機械翻訳文

私は微熱があります。

翻訳対象文

ちょっと熱が出てまして…

機械翻訳文

私は微熱があります。

用例対訳として利用

Step1

機械翻訳しやすい日本語を作成するタスク

タスクの説明

【重要】ボタンをクリックすると「入力文」を再度別の言語へ機械翻訳し、翻訳結果と元の日本語を比較することができます。

「原文」をそのまま、入力文と同じ機械翻訳結果の一致度を評価する日本語を作成してください。

「入力文」をそのままの形で作成してはいけません。原文を分割して、長い文章を複数入力することは禁止です。例として日本語ですが、例文は機械翻訳可能な英語とさせていただきます。

※今日の天気は晴ですが、明日は晴れると予想されています。しかし、明日の天気は晴れるようではありません。

【原文】病気の可能性があるかどうかわかります。

入力文

病気の可能性があるかどうかわかります。

再翻訳

再翻訳の結果が表示されるまで数秒ほどお待ちください

現在の入力文と折り返し翻訳文の一致率 - 熱度 (とくくしい位) になっているのは5段階評価の5です。

タスクを完了する

Step3

Please evaluate the sentences of the following conversation.

I assume the conversation at the hospital.

Health care provider Hospital user

(1) What kind of pain is it?

(2) I feel pressure in my head.

(3) I will take a CT scan of your brain.

How do you judge the fluency of the sentence (2)?

- Incomprehensible
- Disfluent English
- Non-native English
- Good English
- Flawless English

If you find any error, please correct the sentence (2).

Step4

Health care provider Hospital user

How's your appetite?

Please select flawless sentences.

Options

- 1 It's not very good.
- 2 It is not too great.
- 3 I don't have much of one.
- 4 It is not too much.
- 5 I do not have much of an appetite.
- 6 It's lower than usual.
- 7 It's not much of an appetite.
- 8 I don't have much of one.
- 9 I don't have much of one.
- 10 My appetite is not too great.

Step6

2番目の文と入れ替えても意味が通じるような文を選択してください。(複数選択)

医療従事者「何語を話しますか？」
病院利用者「英語を話します。」
医療従事者「英語を話せるスタッフを連れてきます。」

●選択肢●

- 私は英語を話します。
- 私は英語に堪能です。
- 私は英語を話します
- あてはまるものがない

エントリーする

図 1 提案手法を用いた用例対訳作成の流れとタスク画面例

多言語話者の数は少なく、特に、話者の少ない言語を理解できる多言語話者をクラウドソーシング上で確保することは困難である。

クラウドソーシング上の単言語話者を対象にした多言語データ収集の研究では、児童向け図書を対象にした、Chan らによる“Crowdsourced Monolingual Translation”がある [7]。これは、翻訳対象文の言語を母語とする作業者と、翻訳対象言語を母語とする作業者が、文への注釈付けや、それぞれが作成した文の機械翻訳文を修正することで行う。複数の単言語話者によって、繰り返し注釈付けと機械翻訳文の修正を行うことで、1人の多言語話者による修正と同程度の質の翻訳文を取得することができる。しかし、この手法では、翻訳対象文と同じ意味を翻訳文が示すかどうかの評価をしておらず、正確性の確保が必要とされる用例対訳作成に用いることは困難である。そこで我々は、クラウドソーシング上の単言語話者と機械翻訳を用いることで、用例対訳候補文の作成、正確性評価を行う手法について研究を行っている。

3. 提案手法の概要

本章では、文献 [2] の用例対訳作成手法に折り返し翻訳を用いた手法について述べる。文献 [2] の手法を従来手法、従来手法に折り返し翻訳を用いた手法を提案手法とする。

3.1 用例対訳作成の流れ

本節では、提案手法を用いた用例対訳作成の流れについて述べる。図 1 に提案手法による用例対訳作成の流れと、クラウドソーシングにおけるタスクの画面例を示す。本手法は、6 ステップで構成されており、Step 2~6 は文献 [2] における用例対訳作成の手順である。本稿では、機械翻訳文の訂正精度向上のために、Step 1 の処理を追加する手法を提案している。以下に、本手法による用例対訳作成手順について示す。

Step 1 機械翻訳適応文の作成

機械翻訳対象文の言語を母語とするクラウドソーシングの作業者が行う。翻訳対象文をもとに、同じ意味の文で、折り返し翻訳文との一致度が高くなるような文を作成する。本稿では、この手順において作成する、折り返し翻訳文と一致度が高い文を機械翻訳適応文と定義する。詳細は、後述の 3.2 節において述べる。

Step 2 翻訳対象言語へ機械翻訳

Step 1 で作成した機械翻訳適応文を、機械翻訳によって翻訳対象言語へ翻訳する。

Step 3 機械翻訳文の評価と訂正

翻訳対象言語を母語とする作業者に、Step 2 で取得した機械翻訳文の流暢性評価と訂正を依頼する。評

価は 5 段階で、文献 [8] の評価基準 *2 を用いる。この際、図 1 Step 3 のタスク画面のように、作業者には会話文の形式で、文の使用者の属性情報 *3 を付与した機械翻訳文を提示する。これは、文の使用者の属性情報を付与した会話文を提示することで、作業者が用例の使用される状況を推測することができ、訂正精度が向上すると考えたためである。本稿では、この手順において取得した機械翻訳文や訂正文を用例対訳候補文と定義する。

Step 4 用例対訳候補文の評価

翻訳対象言語を母語とする作業者が、用例対訳候補文（機械翻訳文や Step 3 で作成された訂正文）から流暢な文を選択する。この際、作業者には、Step 3 のタスクと同様、文の使用者の属性情報を付与した会話文が提示する。10 名の作業者による多数決を用い、5 名以上が流暢と判断した文を用例対訳候補文とする。

Step 5 翻訳対象文の言語へ機械翻訳

Step 4 で取得した用例対訳候補文を、機械翻訳によって翻訳対象言語へ翻訳する。

Step 6 用例対訳候補文の正確性評価

翻訳対象言語を母語とする作業者は、Step 5 で作成された機械翻訳文から、翻訳対象文と同様の意味を示す文を選択する。この際、作業者には、Step 3 と同様、文の使用者の属性情報を付与した会話文を提示する。

3.2 機械翻訳適応文作成タスク

本節では、3.1 節において述べた Step 1 の機械翻訳適応文作成タスクについて述べる。図 2 に、タスクの手順と画面例を示す。作業者は、提示された原文を、意図を変えない程度に修正し、折り返し翻訳文との一致度が高くなるような文の作成を行う。タスク画面では、機械翻訳適応文作成の手順のほかに、折り返し翻訳を行う際のコツ [9] をいくつか提示した *4。また、作業者には以下の手順で行うように指示文を提示した。

- (1) 入力欄に「原文」が入力されているので、そのまま再翻訳ボタンをクリックする
- (2) 折り返し翻訳文（再翻訳結果）をみて入力文と同じような文になっているか確認する
- (3) 同じような文でなければ入力文を修正して再翻訳をする
- (4) 2 つの文の一致度が高くなるまで手順 (2) と (3) を

*2 評価段階は、1: Incomprehensible, 2: Disfluent English, 3: Non-native English, 4: Good English, 5: Flawless English

*3 文の使用者の属性情報とは、「医療従事者」または「病院利用者」のことである。

*4 作業者には、「主語・述語を省かない」「できるだけ漢字で書く」「曖昧な表現・難しい表現（比喩など）は避ける」「（意味が変わらない程度に）単語を書き換える」「慣用表現を避ける」の 5 つを提示する。これは、文献 [9] を参考にした。

表 1 機械翻訳適応文作成タスクの設定

設定項目	設定値
対象文数	38 文
1 文あたりのタスク実行回数	10 回
1 人あたりのタスク回数上限	1 回
総設問数	380 問
総タスク数	190 件
1 タスクあたりの報酬	10 ポイント *6

*6 作業者に支払われる報酬は T-ポイントとして支払われる。T-ポイントは、カルチャア・コンビニエンス・クラブが展開するポイントサービスである。

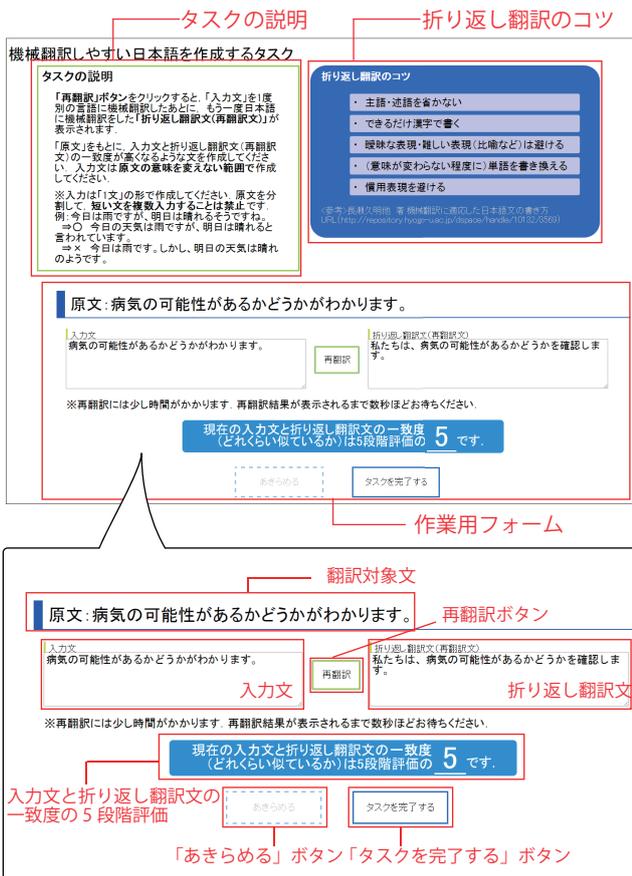


図 2 機械翻訳適応文作成タスク画面例

繰り返す

指示文中の「同じような文」というのは、同じ文や類似文、同様の意味を示している文などのことであり、作業者が主観的に判断を行う。機械翻訳適応文タスクにおいて作業者に提示される文は、入力文と折り返し翻訳結果のみであり、作業には母語で記述された文のみ提示される。提示された 2 文が「同じような文」かどうかの主観的な評価は、提示される文の言語を母語とする作業であれば、容易に行える可能性があると考えている。

図 2 における「原文」が翻訳対象文である。作業者は「再翻訳」ボタンをクリックすることで、「入力文」欄に入力した文の折り返し翻訳結果を見ることが出来る。入力文と折り返し翻訳文を比較し、同じ文または類似した文であると判断した場合には、「タスクを完了する」ボタンをクリックし、タスクを完了する。また、そうでないと判断した場合には、類似していると判断できるまで折り返し翻訳を繰り返す *5。

文が一致しているかどうかの最終的な判断は作業者に委ねるが、文の一致度の目安として、作業者に、折り返し翻訳との RIBES[10] の値をもとに、一致度を 5 段階で提示する。5 段階の基準は、RIBES の値を 0.2 ずつに区

*5 5 回以上折り返し翻訳を行って一致度が向上しなかった場合は、「あきらめる」ボタンを選択してタスクを完了することができる。

切ったものを提示した。RIBES は、機械翻訳の結果を自動的に評価するための尺度の一つであり、人間による評価と相関の高い評価値を算出することが可能である。折り返し翻訳文との RIBES の値が高い場合、その文は機械翻訳に適応した文である可能性があると考えられる。

4. 評価実験

4.1 実験用データセット

本節では、実験に用いるデータセットについて述べる。本実験では、文献 [2] の評価実験で用いた 38 組の日本語の会話文を用いた。この 38 組の会話文は、これまでの実験においてクラウドソーシングを用いて作成したものである。会話文は 3 文からなり、1 文目は疑問文、2 文目は 1 文目に対する回答、3 文目は 2 文目に続く応答となっている。また、会話文は医療現場における会話を想定して作成している。これは、本研究の目的が医療現場で利用可能な用例対訳を作成することとしているためである。なお、各機械翻訳文を区別するために、本実験で用いる 38 文にはそれぞれ文 ID を付与した。

4.2 実験概要

本稿では、クラウドソーシング上の単言語話者による機械翻訳適用タスクについて評価実験を行う。4.1 節で述べた 38 文について、表 1 に示す条件で実験を行った。

本実験では、38 組の会話文の 2 文目を翻訳対象文とした。各翻訳対象文について 10 回ずつ機械翻訳適応文作成タスクを行ったため、本実験における総設問数は 380 問である。本実験では機械翻訳サービスとして Google 翻訳 *7、クラウドソーシングサービスとして Yahoo!クラウドソーシング *8 を用いた。本実験で用いた Yahoo!クラウドソーシング上のタスク画面例を図 3 に示す。Yahoo!クラウドソーシングが提供するタスク画面において、折り返し翻訳を行うことは不可能である。そのため、本実験では作業者に、折り返し翻訳機能を持つ作業用 Web ページへ移動し、タスクを行うように指示した。

作業者は 1 タスクあたり 2 文の翻訳対象文について、

*7 <https://translate.google.com/>

*8 <http://crowdsourcing.yahoo.co.jp/>



図 3 Yahoo!クラウドソーシング上のタスク画面例

機械翻訳適応文の作成を行なう。設問は 1 タスクあたり 2 問となるが、そのうちの 1 問はチェック設問として設定した。チェック設問とは、既に正解が用意されている設問のことであり、不真面目な作業者の存在への対策として設定されるものである。本タスクでは、各設問終了後にランダムな 9 文字の文字列を提示する。この文字列をタスク完了パスワードとし、図 3 のタスク完了パスワード入力欄に入力するよう指示した。2 問とも作業内容は同じだが、各設問の回答後に表示される、どちらかの設問のタスク完了パスワードが、正解（それ以外の文字を入力した場合は不真面目な作業者と判断する）として設定されている。本実験では、チェック設問において取得した回答も実験結果として含める。

5. 実験結果と考察

本章では、4 章で述べた評価実験の結果と考察について述べる。

5.1 取得した回答数

本節では、クラウドソーシングの作業員から取得した回答の数について述べる。

なお、実験中にシステムに不具合があり、図 2 の原文が正しく表示されない期間があった。そのため、一部結果を取得できなかった用例がある。4.2 節で述べたタスクの設定より、計算上は 1 文あたり 10 件の回答があり、回答の総数は 380 件となる予定である。しかし、前述のシステムの不具合や、1 問のみ回答しタスクを中断した作業員の存在があったため、取得した回答の総数は 714 件となった。この 714 件から、不正な回答を除いた 655 件を考察に用いる。不正な回答とは、“edgfhfhrtret” のような文字列や、翻訳対象文「あまり眠れていません。」

表 2 翻訳対象文と機械翻訳適応文における折り返し翻訳文の RIBES の値の比較

	平均	標準偏差
翻訳対象文 (38 文)	0.756	0.256
機械翻訳適応文 (672 文)	0.861	0.167

に対する「お腹へった」のような翻訳対象文と大きく意味の外れた回答である。また、翻訳対象文 38 文のそれぞれにおいて 10 件の回答が収集できる予定だが、前述の理由（システムの不具合や作業を中断した作業員の存在）により、翻訳対象文によって回答数に偏りがあった。10 件以上の回答を取得した翻訳対象文は 38 文中 34 文、回答が 10 件未満だった翻訳対象文は 4 文であった。4 文に対する回答数は 6 件～9 件であり、10 件以上回答があった翻訳対象文のうち最も回答が多かった文は、33 件の回答を得ていた。なお、各文における回答数の平均は 17.6 文、標準偏差は 6.9 文だった。回答が 10 件未満だった 4 文はチェック設問として設定されている翻訳対象文だった。このことから、Yahoo!クラウドソーシングにおけるチェック設問が、チェック設問に設定された設問から、ランダムに作業員へ出題する仕様のため、回答数が 10 件未満の翻訳対象文があったと考えられる。

5.2 作成された機械翻訳適応文

本節では、クラウドソーシングの作業員によって作成された機械翻訳適応文について述べる。翻訳対象文と機械翻訳適応文における、折り返し翻訳文との RIBES の値の比較を表 2 に示す。表 2 は、各翻訳対象文とその折り返し翻訳文の RIBES の値と、機械翻訳適応文とその折り返し翻訳文の RIBES の値それぞれの平均と標準偏差を示している。655 件中 318 件の回答において、折り返し翻訳文との RIBES の値が翻訳対象文よりも機械翻訳適応文の方が向上していた。また、RIBES の値に変化がなかった回答は 276 件であり、これは、作業員が翻訳対象文をそのまま回答として用いたためである。なお、38 文中 35 文において、折り返し翻訳文との RIBES の値が、翻訳対象文を用いた場合よりも高くなる機械翻訳訂正文を取得することが出来た。

本実験において、折り返し翻訳文との RIBES の値が、翻訳対象文の場合より向上した機械翻訳適応文を取得できなかった文は 3 文である。翻訳対象文よりも機械翻訳に適切な文を取得することができなかった 3 文を表 3 に示す。RIBES の値が向上しなかったのは、この 3 文の翻訳対象文とその折り返し翻訳文との RIBES の値が 1 であり、これ以上は値を向上させることができなかったためである。なお、従来手法を用いた用例対訳作成評価実験において、この 3 文から取得できた訂正文は、専門家による正確性評価で高い評価を得た。これは、機械翻訳文の訂正文が用例対訳として使用可能であるというこ

表 3 折り返し翻訳文との RIBES の値が向上しなかった機械翻訳適応文の例

文 ID	翻訳対象文		機械翻訳適応文	
	文	RIBES	文	RIBES
12	副作用が心配です	1	副作用が心配です	1
			副作用も心配です	0.809
31	咳が止まりません。	1	咳が止まりません。	1
			咳が酷いです。	0.946
34	お腹が痛いです	1	お腹が痛いです	1
			私のお腹には痛みがあります。	0.926

・表中の RIBES の値は、各文とその折り返し翻訳文における RIBES の値である。

表 4 折り返し翻訳文との RIBES の値が向上した機械翻訳適応文の例

文 ID	翻訳対象文		機械翻訳適応文	
	文	RIBES	文	RIBES
9	ペットからだと考えられます。	0.740	ペットからであると思われます。	0.880
			原因は、ペットによるものだと考えられます。	0.912
10	例えば水虫の人のスリッパを履いたりすると感染します	0.454	例えば水虫の人のスリッパから感染します	0.904
			例えば、水虫に感染した人のスリッパを使うと感染します。	0.909
13	仕事中に発作が起きてしまわないかがとても心配です。	0.673	仕事中に発作が起きたらと心配です。	0.841
			私は仕事の途中で、症状悪化が起きるのではないかと大変心配です。	0.886

・表中の RIBES の値は、各文とその折り返し翻訳文における RIBES の値である。

とを示している。このことから、折り返し翻訳文と一致度の高い文の機械翻訳文は、クラウドソーシング上の単言語話者によって、適切な訂正が行なわれる可能性があると考えられる。

5.3 翻訳対象文と機械翻訳適応文の比較

本節では、翻訳対象文と機械翻訳適応文の比較について述べる。それぞれの折り返し翻訳文との RIBES の値を比較したところ、表 3 と表 4 に示すような結果が得られた。

表 3 は、機械翻訳適応文とその折り返し翻訳文の RIBES の値が翻訳対象文とその折り返し翻訳文の RIBES の値から向上しなかった例である。また、表中の 3 文は、5.2 節で述べた、翻訳対象文よりも機械翻訳に適切な文を取得することができなかった 3 文である。

また、表 4 に、機械翻訳適応文作成タスクにおいて作成した文とその折り返し翻訳文との RIBES の値が、翻訳対象文とその折り返し翻訳文との RIBES の値から向上した文の例を示す。作業者は、文 ID: 9 の機械翻訳適応文における「原因は」のように、翻訳対象文に単語を追加したり、文 ID: 10 において「スリッパから」を「スリッパを使うと」へ変更したように、表現を変更したりすることで、折り返し翻訳文との一致度の向上を行ったと考えられる。

しかし、文 ID: 13 の、「発作」が「症状悪化」に変更されたように、元の翻訳対象文と異なる意味を示す文も作成された。これは、「発作」の折り返し翻訳結果が「攻撃」となるために、「発作」以外の表現を用いて文の作成

を行った可能性がある。このように、折り返し翻訳文との RIBES の値は向上したが、翻訳対象文とは異なる意味を示す文を作成した回答も存在した。その例を表 5 に示す。表 5 のように、翻訳対象文とニュアンスが異なる機械翻訳適応文が作成される際、翻訳対象文に含まれていなかった主語を追加される場合が多かった。本実験において、作業者に提示した文は、3 文からなる会話文の 2 文目である。そのため、1 文目の内容によっては、主語の追加や表現の変更によって、会話文中の翻訳対象文が示す意味に近づく場合もあるが、多くの場合、別の意味を示す機械翻訳適応文となっていた。

これらのことから、機械翻訳適応文を用いて用例対訳作成を行うことで、機械翻訳文の訂正精度が向上する可能性があるが、翻訳対象文と同じ意味を示す対訳を作成するためには、翻訳対象文と機械翻訳適応文の意味が一致している必要があると考えられる。機械翻訳適応文作成を行う作業者は、翻訳対象文に単語を追加したり、表現を変更したりすることで折り返し翻訳文との一致度の向上を行う場合がある。よって、作業者が翻訳対象文の使用される状況を理解できるように支援することや、翻訳対象文と機械翻訳適応文の意味があっているかの確認を行う必要があると考えられる。

6. おわりに

本稿では、クラウドソーシング上の単言語話者による用例対訳作成手法において、折り返し翻訳を用いた機械翻訳適応文作成の手順を追加することで、単言語話者による機械翻訳文の訂正精度の向上を図る手法を提案し

表 5 翻訳対象文と機械翻訳適応文で表す意味が異なる例

文 ID	翻訳対象文	機械翻訳適応文
2	余りありません。	商品の残りが、余りありません。
4	よく眠れます。	一番に眠れました。 さてあなたは、睡眠状態になります。
7	家の中だと思えます	きっと彼は家の中だと思えます
26	まずは食生活です。	人間は食生活が重要です
36	あまりがありません。	お金がありません。

た。また、クラウドソーシング上の単言語話者を対象とした機械翻訳適応文作成タスクを作成し、実験を行い、以下のことを明らかにした。

- 折り返し翻訳とその一致度を作業者に提示することで、クラウドソーシング上の単言語話者による機械翻訳適応文作成が行える可能性がある。
- 機械翻訳適応文を用いて、クラウドソーシング上の単言語話者による機械翻訳の訂正を行う場合、適切な訂正が行われる可能性がある。
- 機械翻訳適応文作成の際、作業者による単語の追加や表現の変更が行われた場合、翻訳対象文と機械翻訳適応文の意味が一致しない場合がある。

今後は、機械翻訳適応文作成タスクにおいて作成された文を評価する手法の検討や、本実験で取得した機械翻訳適応文を 3.1 節で述べた用例対訳作成手法に用いて用例対訳作成実験を行う。

謝辞 本研究の一部は、JST A-STEP「多段クラウドソーシングを活用した多言語用例対訳プラットフォームの構築」、JSPS 科研費 24220002 および 26730105、和歌山大学システム工学部大学院プロジェクトの助成を受けた。

参考文献

- [1] 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, Vol. 2, No. 3, pp. 23-33 (2012).
- [2] 山本 里美, 福島 拓, 吉野 孝: クラウドソーシングにおける機械翻訳文の評価結果を活用した用例対訳作成手法の提案, 情報処理学会研究報告, グループウェアとネットワークサービス研究会, Vol. 2015-GN-93, No. 38, pp.1-8 (2015).
- [3] Chris Callison-Burch: Fast, Cheap, and Creative: Evaluating Translation Quality Using Amazon's Mechanical Turk, *Proceedings of EMNLP 2009*, pp. 286-295 (2009).
- [4] Omar F. Zaidan and Chris Callison-Burch: Crowdsourcing Translation: Professional Quality from on-Professionals, '11 Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Vol. 1, pp. 1220-1229 (2011).
- [5] Vamshi Ambati, Stephan Vogel: Can Crowds Build Parallel Corpora for Machine Translation Systems?, Proceedings of the NAACL HLT 2010 Workshop on Creating

Speech and Language Data with Amazon's Mechanical Turk, pp. 62-65 (2010).

- [6] Philip Resnik, Olivia Buzec, Yakov Kronrod, Chang Hu, Alexander J. Quinn, Benjamin B. Bederson: Using Targeted Paraphrasing and Monolingual Crowdsourcing to Improve Translation, *Transactions on Intelligent Systems and Technology (TIST)*, Vol. 4, No.3, Article No.38 (2013).
- [7] Chang Hu, Philip Resnik, and Benjamin B. Bederson.: Crowdsourced Monolingual Translation, *ACM Trans. Comput.-Hum. Interact.*, vol. 21, No. 4, pp. 1-35 (2014).
- [8] Kevin Walker, Moussa Bamba, David Miller, Xisoyi Ma, Chris Cieri, and George Doddington, Multiple-Translation Arabic (MTA) Part 1, In *Linguistic Data Consortium*, Philadelphia (2003).
- [9] 長瀬 久明, 橋 守亮: 機械翻訳に適応した日本語文の書き方, <http://hdl.handle.net/10132/3569> (2010).
- [10] 平尾 努, 磯崎 秀樹, 須藤 克仁, Kevin Duh, 塚田 元, 永田 昌明: 語順の相関に基づく機械翻訳の自動評価法, *自然言語処理*, Vol. 21, No. 3, pp. 421-444 (2014).