

クラウドソーシング上の単言語話者を対象とした翻訳パズルの提案

Proposal of Translation Puzzle for Monolingual Speakers on Crowdsourcing

福島 拓^{*1}

Taku Fukushima

吉野 孝^{*2}

Takashi Yoshino

^{*1}静岡大学大学院総合科学技術研究科

Graduate School of Integrated Science and Technology, Shizuoka University

^{*2}和歌山大学システム工学部

Faculty of Systems Engineering, Wakayama University

Generally, translation tasks have been translated by multilingual speakers. However, multilingual speakers as well as many monolingual speakers belong to crowdsourcing services. It is difficult to reject such monolingual speakers to participate the multilingual translation tasks. That is, we have allowed that monolingual speakers have gotten improper wages. In this study, we propose a multilingual translation method for monolingual speakers. We aim to support monolingual workers' motivation using a task of puzzle type.

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。しかし、すべての訪日外国人が日本語を理解しているとは言い難い。また、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 [Takano 93, Aiken 94], 非母語によるコミュニケーションは十分に行うことができない。非母語によるコミュニケーションで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ため、適切な支援が求められている。

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 [宮部 09, 尾崎 11]。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。用例対訳とは、用例を多言語に正確に翻訳したコーパスのことを指し、「保険証はお持ちですか?」「はい」「いいえ」などの利用現場で使用される言葉を多言語で提供することができる。このように使用される用例対訳の収集・共有を目的として、我々は多言語用例対訳共有システム TackPad の開発を行っている [福島 12b]。しかし、収集済みの用例対訳数は十分ではなく、現在の約数十倍の用例が必要であると考えられるが、対訳作成を行う翻訳者への負担が非常に大きくなるという課題を抱えている。

そこで本稿では、クラウドソーシング [Howe 08, Doan 11] を活用した用例対訳の収集について述べる。クラウドソーシングとは、人々 (群衆) への作業や業務の委託を指す。本稿では、クラウドソーシング上の労働者を Worker と呼ぶ。クラウドソーシングでは大量の用例に対して安価で評価依頼を行うことができる利点がある。しかし、クラウドソーシング上で多言語が関係する作業委託を行った場合、特に不正確なものが多く含まれることが分かっている [Callison-Burch 09, Negri 10, 福島 12a]。このため、本稿では多言語による悪影響を減らすために、クラウドソーシングへの作業委託を単言語で行うこととする。単言語話者を対象とした多言語翻訳は機械翻訳を用いることが多いが [福島 13, 福島 14], 機械翻訳の翻訳精度に対訳の精度が依

存する問題がある。そこで本稿では、文の構造と正確に翻訳された単語を使用した「翻訳パズル」を提案する。本手法では、多言語辞書を用いて単語を正確に変換し、文の構造をグラフ化することで正確な多言語用例対訳候補の取得を目指す。

2. 関連研究

クラウドソーシングを用いた多言語データの収集は多く行われている。Callison-Burch はクラウドソーシングを用いた多言語対の正確性評価を [Callison-Burch 09], Negri らはクラウドソーシングを用いた多言語対の作成 [Negri 10] をそれぞれ行っている。これらの研究では、多言語話者を対象として行われている。また、クラウドソーシング上の不適切に対価を得ようとする Worker を考慮した手法がそれぞれ提案されている。我々も翻訳者の作業特徴を考慮し、作業時間をもとにした多言語対の正確性評価手法を提案した [福島 12a]。

しかし、これらの研究では、クラウドソーシング上に少ないと考えられる多言語話者を対象とした作業委託を行っている。2言語が関係するこれらの作業は比較的難解な作業となる。このため、単純な作業で対価を得ようとする Worker が多いと考えられるクラウドソーシング上では、不適切に対価を得ようとする Worker が存在しており、その対策が求められている。

このため、我々は単言語話者のみを用いて多言語用例対訳の作成を可能とする手法の提案を行った [福島 13, 福島 14]。これらの手法では、画像と機械翻訳、複数の機械翻訳をそれぞれ用いることで、正確な多言語用例対訳の作成を目指している。しかし、これらの手法は機械翻訳の翻訳精度に依存するため、精度の低い機械翻訳文からは正しい対訳文を取得できる可能性が低かった。また、作業タスクが比較的単純であるため、不適切に対価を得ようとする Worker が比較的多く存在していた。

本稿では、正確に翻訳が可能な多言語辞書を用いることで正確性の向上を目指す。また、クラウドソーシングのタスクにおいてゲーム性を持たせた場合、不適切に対価を得ようとする Worker が減少する可能性が示されている [Eickhoff 12]。そこで本稿では、グラフ化した文の構造を Worker に提示し、Worker にパズルを解くような感覚を与えることで、不適切に対価を得ようとする Worker 数の減少を目指す。

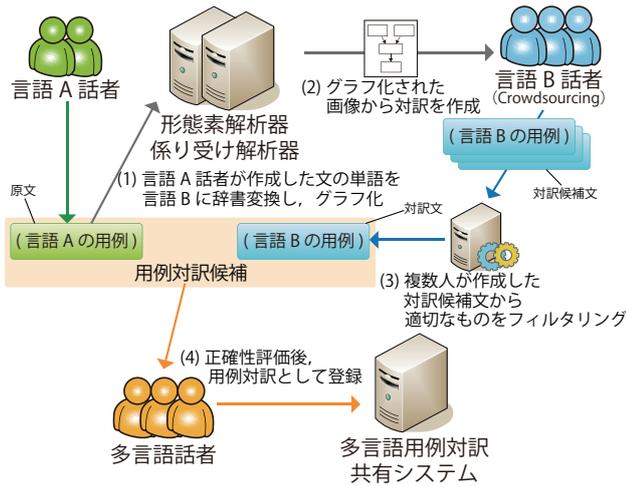


図 1: 本手法のコンセプト

3. 翻訳パズル

本手法のコンセプトを図 1 に示す。本手法は、形態素解析器と係り受け解析器、多言語辞書を用いたグラフ化、クラウドソーシングでの対訳作成、フィルタリング、正確性評価で構成されている。本稿では言語 A を日本語、言語 B を英語とし、以降の各節で日本語の用例を英語へ翻訳する流れの詳細について述べる。なお、本手法は文献 [福島 15] に基づいているが、文献 [福島 15] の考察をもとにして“◇”記号を付与した部分を新たに追加している。

3.1 形態素解析器と係り受け解析器を用いたグラフ化

本節では、形態素解析器と係り受け解析器を用いたグラフ化について述べる。本節は図 1(1) の内容にあたる。以降の各項を順に行うことでグラフ化を行う。

3.1.1 形態素解析器を用いた形態素分割

日本語の用例を形態素解析器を用いて単語に分割する。本手法では、以下の 4 品詞のみ次項以降で用いることとした。

- 名詞（非自立を除く、非自立の副詞可能は含める）
- 動詞（非自立を除く）
- 形容詞
- 副詞

また、否定形や疑問形、上記の 4 品詞以外で欠落する情報を表現するために以下の処理を行った。

- 終助詞「か」の場合、“?”に置き換え
- 動詞が未然形かつ、対象動詞の次の単語が助動詞の場合、“NOT”を付与
- 助動詞が未然形かつ、対象動詞の前の単語が動詞の場合、“NOT”に置き換え
- ◇ 動詞「する」の場合、“DO”に置き換え
- ◇ 名詞「よう」の場合、“FEEL LIKE”に置き換え
- ◇ 助動詞「たい」の場合、“WANT TO”に置き換え

3.1.2 多言語辞書を用いた多言語変換

前項で取得した 4 品詞を、以下の手順で多言語変換を行った。

1. 多言語辞書を用いて日本語の単語を完全一致で検索し、英語の単語へ辞書変換を行う。その際、動詞と形容詞は原形を用いた。また、以下の処理を行っている。
 - 名詞が連続しているときは複合語として別途検索
 - ◇ 形容詞の後に名詞（接尾語、特殊）がある場合は、両者を結合して名詞として利用*1。
2. 1. で変換失敗した名詞（複合語を除く）は機械翻訳で翻訳したものを使用する。
3. 1. で変換失敗した動詞は以下の手順を繰り返して辞書に存在する単語（動詞）となるまで変換し、使用する。
 - (a) 単語の最後の 1 文字を削る
 - (b) 語尾を「ウ列」にして前方一致検索
 - (c) 語尾を元に戻して前方一致検索

例として「押さえつける」の場合、(a) で「押さえつけ」とし、(b) で「け」をウ列にして「押さえつく」で前方一致検索を行う。単語が存在しない場合は (c) で語尾を戻して「押さえつけ」で前方一致検索を行う。それでも存在しない場合はさらに (a) に戻って「押さえつ」とし、単語（動詞）が見つかるまで検索を行う形を取った。

3.1.3 係り受け解析器を用いたグラフ化

係り受け解析器を用いて原文の文の構造を得る。その後、以下の手順でグラフ化を行った。作成例を図 2 に示す。

1. 原文を係り受け解析し、係り受け解析結果中の単語を 3.1.2 項の辞書変換結果と置き換える。なお、3.1.2 項で複合名詞が存在した場合は複合名詞を優先的に使用した。
2. 3.1.1 項の“?” “NOT”などを配置する。なお、“?”以外は一つ前の品詞の前に表示する形を取った。
3. 4 品詞以外の単語を除去し、空白となったチャンクも除去する。
4. “NOT”が 2 つ連続する場合、強い肯定と捉えて両方の“NOT”を除去する。
5. 係り受け関係をもとにチャンク間に矢印（係り先から係り元*2）を引き、グラフ化する。また、下記の処理を行う。
 - ◇ 助詞「が」が係り元の場合、係り先への矢印へ変更
 - ◇ 助詞「の」「と」の前後が名詞の場合、“of” “+”で結合
 - ◇ 形容詞の基本形の後に名詞が存在した場合は、矢印を用いずにチャンクのブロックを結合

3.2 クラウドソーシングでの対訳作成

本節では、3.1 節で作成したグラフをクラウドソーシング上の複数の Worker（英語話者）に提示し、対訳作成を行う方法について述べる。本節は図 1(2) の内容にあたる。

本節では、グラフを Worker に提示し、以下に示すタスクを依頼する。本節の内容により、日本語を翻訳した英文（対訳）が取得できる。以降、Worker が作成した文を「対訳候補文」とする。

1. グラフをもとに英語の正しい文の記入を依頼。
2. 1. で入力した文の確信度を 5 段階で依頼*3。
3. タスクに関するコメント（任意項目）。

*1 例として「寒」+「さ」は「寒さ」という名詞として利用した。
 *2 文献 [福島 15] では係り元から係り先への矢印を用いていた。
 *3 評価段階は、1: Poor, 2: Fair, 3: Average, 4: Good, 5: Excellent, である。

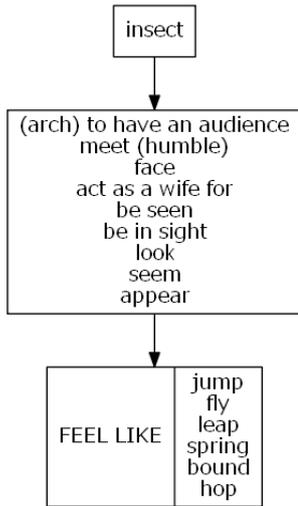


図 2: グラフの例

3.3 対訳候補文のフィルタリング

本節では、3.2 節で作成された対訳候補文のフィルタリングについて述べる。本節は図 1(3) の内容にあたる。フィルタリングでは、不真面目な Worker により作成された対訳候補文を除去した後、正確性の高い対訳候補文の抽出を行うものとする。なお、本稿では不適切な対訳候補文の除去についてのみ述べる。

関連研究でも述べたとおり、クラウドソーシング上には不適切に評価を得ようとする Worker が存在している。本手法では単言語話者が作業可能なタスクへ分解することで作業の難易度を下げ、不真面目な Worker の減少を目指している。しかし、不真面目な Worker を完全になくすことは難しいと考えられるため、本手法では不真面目な Worker が作成した対訳候補文の除去を行う。

本稿で不適切な対訳候補文と判定したものを以下に示す。

1. 同一 Worker が別の翻訳文に対して同一の対訳候補文を記入した場合。
2. 3 単語以下の対訳候補文で、“not good” “none” “jkljlkj” などの明らかに不適切な文。

次章で述べる実験では、2. の内容を著者の一人が調査して除去した。今後は自動化する必要があると考えられる。

3.4 対訳文の正確性評価

前節までの内容は単言語話者のみが関わっているため、正確性を期すために日本語と英語を理解できる多言語話者が最終的な評価を行う。評価対象は、原文（日本語）と対訳文（英語）の正確性である。

4. 対訳作成実験

本稿では、前章で述べた手法を提案手法、文献 [福島 14] で行われた、1 つの機械翻訳器を用いて翻訳文を生成し、Worker に提示する手法を“MT1”、複数の機械翻訳器を用いて翻訳文を生成し、Worker に提示する手法を“MT2”とする。また、文献 [福島 15] で行われたグラフ化手法を“IMG”とする。本章では、上記の 4 手法を比較する実験について述べる。本実験のデータを表 1 に示す。

表 1: 実験データ

手法	MT1	MT2	IMG	提案
提示内容	MT: 1 文	MT: 4 文		グラフ
原文数	50 文			
1 文あたりの Worker 数	20 人			
Worker の報酬 (5 文につき)	10¢		25¢	10¢

・ MT は機械翻訳によって翻訳された文を示す。

表 2: 不真面目な Worker と不適切な対訳候補文の数

	Worker 人数				対訳候補文数			
	MT1	MT2	IMG	提案	MT1	MT2	IMG	提案
適切	16	17	36	25	694	686	855	887
不適切	13	16	8	10	306	314	145	113
総数	29	33	44	35	1000			

- ・ Worker 人数の列の単位は人、対訳候補文数の列の単位は文である。
- ・ 不適切の行の Worker は、1 文でも不適切な対訳候補文を作成した人数である。
- ・ 適切の行の Worker は、すべての対訳候補文が適切であった人数である。

使用する原文（日本語）は、多言語用例対訳共有システム TackPad[福島 12b] から、医療従事者に対して患者が使用する文を 50 文抽出した。なお、文献 [福島 14, 福島 15] と同じ原文群を用いている。平均文字数は 14.0 文字、標準偏差は 6.1 文字であった。なお、50 文中 10 文は 20 文字以上の文を選択している。

本稿では形態素解析器として ChaSen^{*4} を、多言語辞書として「デ辞蔵^{*5}」の EDICT 和英辞典を、機械翻訳器として言語グリッド [Ishida 11] が提供する J-Server を、係り受け解析器として CaboCha^{*6} を、グラフ化には Graphviz^{*7} をそれぞれ使用した。

なお、一部の単語^{*8}の形態素解析に失敗したが、形態素解析器の辞書への追加により改善が可能であるため、本稿では正しく形態素解析ができたものとして扱う。また、「デ辞蔵」から取得した動詞が“to V”の形となっていたため、“to”を取り除く処理を加えている。

クラウドソーシングサービスとして CrowdFlower を利用し、アメリカ合衆国在住の Worker を対象として、原文 1 文につき 20 名分を作業依頼した。その際、グラフは患者が医療従事者に対して使用する言葉をもとにして作成したことを伝えている。Worker の報酬は 5 文につき 10 セントを支払った。このため、原文 1 文に対する対訳候補文の取得には手数料を含めて 54 セントの費用がかかった。

5. 実験結果と考察

5.1 不適切な対訳候補文のフィルタリング

3.3 節のフィルタリングの結果である、不真面目な Worker と不適切な対訳候補文の数を表 2 に示す。なお、表 2 では 1 文でも不適切な対訳候補文を作成した Worker の人数を不適切

*4 <http://chasen-legacy.sourceforge.jp/>

*5 <http://dejizo.jp/dev/>

*6 <https://code.google.com/p/cabocha/>

*7 <http://www.graphviz.org/>

*8 ふくらはぎ、できもの、けがしました

に分類している。表 2 の MT1 手法および MT2 手法と提案手法との比較から、提案手法を用いることで同一 Worker が別の類似文に対して同一の対訳候補文を作成したり、明らかに不適切な文を入力したりする割合は減少していることが分かる。提案手法と金額のみが異なっている IMG 手法とも同傾向であったため、グラフ化することで Worker にパズルを解くような感覚を与えたことが良い影響を与えていると考えられる。

5.2 作成された対訳候補文の正確性

すべての従来手法で正確に翻訳ができなかった文として、ID 9 の「虫が飛んでいるように見えます」がある。MT1 手法と MT2 手法では、機械翻訳によって翻訳された“A bug seems to be flying.”を、IMG 手法では「虫」「飛ぶ」「見える」のみをそれぞれ使用していたが、正確な翻訳作成が困難であった。本手法では比喩表現に対応し、図 2 のようにグラフ化することで、“I feel a insect flying in my sight.”という対訳候補を得ることができた。同様に、IMG 手法で正しい対訳を得ることができなかった ID 39 の「胸がえぐられるように痛いです」からも“I feel a pain in my chest like someone is gouge out a part of it.”という対訳候補を得ることができた。

また、IMG 手法では動詞「する」に対して“do”“try”“attempt”など 19 単語が提示されており、正確な翻訳作成に失敗する例が見られていた。本手法では、“DO”と表現することで、ID 16 「目が充血しています」から“eye”→“DO congetion (with blood)”というグラフ化を行い、“My eye is congested.”という対訳候補を得ることができた。ただし、“DO”を疑問詞であると捉え、“Do you experience congestion?”という疑問形の文が作成される例が一部存在していた。この点に関しては表現手法を再考する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を活用した対訳作成手法の提案を行った。本手法では、多言語辞書で変換した単語をグラフ化し、画像として提示することで正確性の向上を目指した。実験より、比喩表現などに対応することで従来のグラフ化手法よりも正確な対訳候補文が得られる傾向が見られた。今後は、従来手法と同様に翻訳者による評価を行い、本手法の有効性を確認する。

謝辞

本研究の一部は JSPS 科研費 (26730105) による。

参考文献

- [Aiken 94] 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)
- [Callison-Burch 09] Callison-Burch, C.: Fast, Cheap, and Creative: Evaluating Translation Quality Using Amazon’s Mechanical Turk, in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP 2009)*, pp. 286–295 (2009)
- [Doan 11] Doan, A., Ramakrishnan, R., and Halevy, A. Y.: Crowdsourcing systems on the World-Wide Web, *Communications of the ACM*, Vol. 54, No. 4, pp. 86–96 (2011)

- [Eickhoff 12] Eickhoff, C., Harris, C. G., Vries, de A. P., and Srinivasan, P.: Quality through Flow and Immersion: Gamifying Crowdsourced Relevance Assessments, in *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 871–880 (2012)
- [Howe 08] Howe, J.: Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business, Crown Business (2008)
- [Ishida 11] Ishida, T. ed.: *The Language Grid: Service-Oriented Collective Intelligence for Language Resource Interoperability*, Springer (2011)
- [Negri 10] Negri, M. and Mehdad, Y.: Creating a Bi-lingual Entailment Corpus through Translations with Mechanical Turk: \$100 for a 10-day Rush, in *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pp. 212–216 (2010)
- [Takano 93] 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)
- [宮部 09] 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)
- [尾崎 11] 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, 第 AI2010-47 巻, pp. 19–24 (2011)
- [福島 12a] 福島 拓, 吉野 孝: クラウドソーシング労働者の作業特徴に着目した多言語テキストペアの正確性評価手法, Web とデータベースに関するフォーラム (WebDB Forum 2012), No. B4-3, pp. 1–8 (2012)
- [福島 12b] 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. 22–33 (2012)
- [福島 13] 福島 拓, 吉野 孝: クラウドソーシングを用いた画像提示型多言語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, 第 2013-GN-86 巻, pp. 1–7 (2013)
- [福島 14] 福島 拓, 吉野 孝: クラウドソーシングを用いた高精度対訳作成のための低品質翻訳の活用, 情報処理学会, マルチメディア, 分散, 協調とモバイル (DICOMO2014) シンポジウム, pp. 1209–1215 (2014)
- [福島 15] 福島 拓, 吉野 孝: 翻訳パズル: クラウドソーシング上における不完全な翻訳を用いた対訳作成手法, 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, 第 2015-GN-93 巻, pp. 1–7 (2015)