

## iTesting チャンネルによる HIV 検査体制の構築と確立のための研究

研究分担者：吉田理加（愛知県立大学外国語学部ヨーロッパ学科）

### 研究要旨

本研究は様々な異なる形態の HIV 検査を実践するための基礎資料（マニュアル）を作成することを目的としている。研究分担者が属するアウトリーチンググループでは、外国人など多様な人々のコミュニティにアウトリーチをする方法を模索した。例えば、外国人コミュニティへのアウトリーチのために iTesting@Nagoya ウェブサイトと予約サイト、そして検査通知サイトを多言語化した。検査通知サイトは、システム上の都合で機械翻訳を使用せざるを得ないのだが、原文の日本語に前編集を施し、より精度の高い機械翻訳となるように工夫した。その前編集の詳細について本項で説明した。結論として、機械翻訳に依存することの危険性並びに前編集や後編集など、機械翻訳に人手をかけることが訳文の精度の向上に結び付いていることが明らかになった。

### A. 研究目的

本研究は、現行の HIV 検査体制に加えて様々な HIV 検査を実践し、より受検者にとっても医療従事者にとっても受検/施行可能な検査の導入を進めるための基礎資料（マニュアル）を作成することを目的とする。その中で、外国人や学生、LGBTQ など多様なコミュニティへのアウトリーチのよりよい実践方法を考察する。

### B. 研究方法

アウトリーチンググループでは、iTesting@Nagoya の HIV 検査の重要性を説明した啓発資料の作成および設置、検査の広報（SNSでの発信・動画配信・多言語対応）、検査を評価するためのアンケートの作成（多言語対応）・解析、検査施行可能な場の模索、各コミュニティでの啓発活動、多言語対応した iTesting@Nagoya ウェブサイトの作成を行う。

本年度はまず、iTesting@Nagoya ウェブサイトと予約サイト並びに検査結果通知サイトの多言語化を行い、日本語、英語、スペイン語、ポルトガル語、ベトナム語の多言語サイトを作成した。多言語化のためにどの言語への翻訳を実施するかに関しては、名古屋医療センターにおけるよく用いられた通訳言語を参考にした。9月の検査会用には日本語、英語、スペイン語、ポルトガル語で作成し、12月の検査会用にはさらにベトナム語を追加した。多くの言語に翻訳すること、並びに何語に翻訳するかは、特定の言語コミュニティに iTesting@Nagoya 検査会で扱っている性感染症の負のイメージを結び付けた偏見を生じさせないためにも重要である。

ウェブサイトと予約サイトの翻訳の手順としては、人間の翻訳者が翻訳した後、分担研究者が理解できる言語（スペイン語）以外は、機械翻訳を用いて逆翻訳をした。それを起点テキスト（ST）と見比べ、時制、主語、名詞などに齟齬がみられた箇所や訳抜けや余剰と思われる箇所を翻訳者に再検討を依頼した。このような逆翻訳の工程を入れることで、確実に正確な翻訳であることを担保するシステムとなっていると考える。

検査結果通知サイトに関しては、システムの都合上、人間の翻訳者が行った翻訳を用いることができず、機械翻訳（Google）を用いざるを得なかった。そこで、正確性に劣る日本語からベトナム語への翻訳が他の言語ペアに比べて翻訳精度が低い状況があるため、検査結果通知並びに結果通知に先立って行われるアンケート調査で、誤った翻訳によって誤った結果・情報が通知されないように、起点テキスト（原文の日本語テキスト）の「前編集（プリエディット）」を実施し、サイト提供会社の協力を得ながら、機械翻訳であれども大きな誤訳が生じないように最大限の工夫を施した。

（倫理面への配慮）

検査は、予約からすべて匿名検査で行っている。またアンケート調査も匿名で行い、名古屋市立大学の倫理審査を受審の上実施している。本年度は、多言語でアンケートを実施するには、システム上機械翻訳を使用せざるを得なかったため、人間の翻訳によるアンケートの多言語化は見合わせた。

### C. 研究結果

上述した通り、検査ホームページおよび予約サイトの多言語化を実施し、日本語、英語、スペイン語、ポルトガル語、ベトナム語で情報にアクセスできるようにした。結果通知サイトは、システム全体に関わるため、機械翻訳を使用することになった。機械翻訳の精度を上げるために、原文の日本語に「前編集」と呼ばれる工程を設け、誤訳が表れにくくなるようにした。

### D. 考察

ここでは、「前編集」の工程について簡単に説明し、明らかになったことを述べる。機械翻訳を利用する際、日本語の原文の次の6つの基準を念頭に「前編集」を行うと英語への翻訳精度が改善することがこれまでの研究で明らかにされている。6つの基準とは、1. 文を短くすること、2. 主語を必ず入れること、3. 多義語の使用を控える、4. 名詞の単数・複数を明示的に示す、5. 限定する言葉（例. この、その、あの）を名詞の前におく、6. 固有名詞のスペルを確認する、である。

この「前編集」の手順は以下のように行った。まず、起点テキストを該当の外国語に機械翻訳し、英語、スペイン語、ポルトガル語は研究分担者が、ベトナム語は翻訳者が訳文を日本語に逆翻訳した。その上で、誤った訳が問題となりそうな箇所を抽出した。問題となり「前編集」を行った主な日本語表現と機械翻訳並びに修正は以下の表のとおりである。

原文（起点テキスト）	機械翻訳による目標テキスト（日本語で示したもの）	原文修正（前編集）
全ての言語に共通		
1) 自分が	「君が」	私が
2) コミュニティセンター rise	「コミュニティセンター増加」	rise をバナー画像にして機械翻訳されないようにした
3) (経験が) ない・ある	(存在が)「ない・ある」	いいえ、ありません/はい、あります
ベトナム語		
4) 自宅検査	「住宅検査」	自宅で検査
5) 採取日	「集金日」	採血日
6) 携帯の出会い系サイト	「可動式の出会い系サイト」	携帯電話の出会い系サイト

まず、起点言語と機械翻訳による目標言語を比べ、異なる意味になっている箇所に着目し、機械翻訳によって等価な目標言語となることを確認した上で、起点言語を修正するという「前編集」を実施した。riseのような固有名詞の場合は、機械翻訳のシステムにあらかじめ固有名詞として登録しておくことによって、自動的に訳されることを回避した。さて、「前編集」にかかる具体的な語彙・表現のいくつかを上表に記載して示した。1)から3)は全ての言語への機械翻訳で前編集が必要となった。4)から6)はベトナム語への機械翻訳で必要となった。このことから、機械翻訳と一言で言っても、何語への翻訳かによって訳の精度がかなり異なること、前編集という原文を修正・編集する工程を踏むことで、より精度の高い機械翻訳を提供できることが明らかになった。

### E. 結論

ここまで述べてきたことから、安易な機械翻訳の使用は控えることが重要であることが明らかになった。正しく機械翻訳されるかを事前に検証し、「前編集」を行うなど人手をかけることが機械翻訳の精度の向上に貢献することが明らかになった。

また、何語に翻訳するかが、その言語のコミュニティをステイグマ化することにもつながりかねない点にも配慮が必要であるため、複数の言語に翻訳をするアプローチが重要であることも確認できた。

### F. 研究発表

#### 1. 論文発表

〈研究ノート〉

松下佳世・古川典代・吉田理加 多言語通訳コーパスを活用した日英・日中・日西の訳出比較. 通訳翻訳研究 第22号 (印刷中)

#### 2. 学会発表

国内

1. 武田珂代子・辛島デイヴィッド・宮田玲・山田優・吉田理加 日本におけるトランスレーション・ポリシー研究事始め 日本通訳翻訳学会第33回年次大会(オンライン) 2022年9月
2. 松下佳世・古川典代・吉田理加 多言語通訳コーパスを活用した日英・日中・日西の訳出比較に基づく初期的考察 日本通訳翻訳学会第33回年次大会(オンライン) 2022年9月

3. 飯田奈美子・斎藤美野・坪井睦子・蓮池通子・水野真木子・吉田理加 通訳翻訳研究におけるデータセッションの有効性の検討 日本通訳翻訳学会第 33 回年次大会 (オンライン) 2022 年 9 月

G. 知的財産権の出願・登録状況 (予定を含む)

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他