

## ニュース速報の日本語-手話機械翻訳 Japanese to Japanese Sign Language Machine Translation of Breaking News

宮崎 太郎<sup>†</sup> 梶山 岳士<sup>‡</sup> 箱崎 浩平<sup>†</sup> 村上 智哉<sup>†</sup>  
Taro Miyazaki Takeshi Kajiyama Kohei Hakozaki Tomoya Murakami  
内田 翼<sup>†</sup> 木下 光太郎<sup>†</sup> 金子 浩之<sup>†</sup>  
Tsubasa Uchida Kotaro Kinoshita Hiroyuki Kaneko

### 1 はじめに

手話を母語とするろう者に重要な情報を伝えるために、日本語テキストを手話 CG (Computer Graphics) に変換するシステムの研究開発を進めている [1]. このシステムではまず入力 of 日本語を機械翻訳技術により手話単語列に翻訳する. それぞれの手話単語の動作データは、モーションキャプチャにより事前に作成してあり、この動作データを翻訳結果の単語列に従って接続、アバターに動作させることで手話 CG を生成する. 本稿では、このシステムで使用する機械翻訳について述べる.

我々が手話 CG に翻訳するターゲットの一つに、緊急時などにテレビの画面上に表示されるニュース速報がある. ニュース速報は、(1) 緊急時に送出するため手話通訳者による対応が難しい、(2) 迅速に伝える必要がある情報を含む、(3) 助詞などの省略が多ろう者を含む日本語非母語話者には理解が難しい場合がある、という理由により、テキストを自動で手話 CG に変換して手話でろう者に伝える意義があると考えられる.

ニュース速報は一般のニュース文と比べ短い文章で情報を伝えるために、「避難指示 ○○市 土砂災害のおそれ」のように助詞の省略や体言止めが頻出するなど、一般のニュースとは文体が異なる. そのため、一般のニュース文で学習した翻訳モデルでは翻訳が難しい. そこで、ニュース速報に特化した日本語-手話対訳コーパス (ニュース速報手話コーパス) を作成した. このコーパスを利用して翻訳モデルを学習することで、ニュース速報の翻訳精度が大きく向上したので報告する.

### 2 手話翻訳に用いるコーパス

NHK では手話 CG 制作のために日本語-手話の対訳コーパスを作成している. まず従来から作成している手話ニュースコーパスについて述べたあとに、今回新規に作成したニュース速報手話コーパスについて述べる.

#### 2.1 手話ニュースコーパス

手話ニュースコーパスは NHK で放送している「手話ニュース」を 2009 年 4 月から収集したもので、番組の日本語音声の書き起こしと手話映像、手話映像の書き起こしが含まれている. 手話映像の書き起こしは、手話ニュース中の手話映像を基にろう者や CoDA<sup>1</sup>が手作業で単語を書き起こすことで作成した. 手話の書き起こしは、手話 gloss 表記と呼ばれる、その手話単語が表す代表的な意味を表す日本語単語で表記している.

<sup>†</sup> NHK 放送技術研究所

<sup>‡</sup> NHK 財団

<sup>1</sup> Children of Deaf Adults: 聴覚障害者を親に持つ子どもで、聴者かつ第1言語が手話の人のことを指す.

表1 コーパスの諸元

	手話ニュース コーパス	ニュース速報 手話コーパス
文数	129,058	1,108
手話話者数	20	3
日本語語彙数	26,941	2,678
手話語彙数	3,880	1,590
日本語平均単語数	23.13	15.71
手話平均単語数	22.91	24.64

#### 2.2 ニュース速報手話コーパス

ニュース速報手話コーパスは、NHK が緊急時などに放送するニュース速報を収集して作成した. ニュース速報はテレビ画面上に文字情報として送出されるが、手話は付与されていない. そこで、文字情報を手話ニュースキャスター経験のあるろう者に提示し手話に翻訳したものを動画で撮影、この動画をろう者や CoDA が書き起こすことで対訳コーパスとした.

#### 2.3 コーパスの比較

表1にコーパスの諸元を示す. 日本語語彙数は形態素解析器 mecab<sup>2</sup>により単語分割した語彙の異なり数を、手話語彙数は手話 gloss の異なり数をそれぞれ表す. また、日本語、手話の平均単語数はそれぞれ1文あたりの日本語単語数、手話 gloss 数を表す.

ニュース速報手話コーパスは手話ニュースコーパスと比較して日本語の平均単語数が短い一方、手話の平均単語数は多くなっている. これは、ニュース速報の短い文章に込められた意味を正しく手話に翻訳するためには、テキストに含まれない情報を補完する必要があることが理由と考えられる.

### 3 評価実験

#### 3.1 実験条件

作成したニュース速報手話コーパスによるニュース速報文への翻訳性能の影響を調査するために評価実験を行った. 機械翻訳で一般に用いられる 6 層の Transformer encoder-decoder モデル [2] を用い、まずベースとなるモデルの学習をした後に、ニュース速報のデータのみを用いた fine-tuning によりモデルをニュース速報に特化させる. このモデルの学習データを変えた場合の翻訳性能を比較する. ベースとなるモデルの学習には (1) 手話ニュースコーパスのみを利用、(2) 手話ニュースコーパスとニュース速報手話コーパスを組み合わせ利用、(3) ニュース速報手話コーパスのみを利用、の3パターンを用意した. ベースモデルの学習時には手話ニュースコーパスの開発データで、fine-tuning 時に

<sup>2</sup> <https://taku910.github.io/mecab/>

表 2 学習/開発/評価データの数量

	学習	開発	評価
手話ニュース	128,058	1,000	1,000
ニュース速報	1,000	—	108

表 3 実験結果

ベースモデルの学習データ	Fine-tuning	BLEU (ニュース速報)	BLEU (手話ニュース)
手話ニュース	✓	17.77	25.83
手話ニュース		47.53	24.51
手話ニュース + ニュース速報	✓	46.88	26.01
ニュース速報		51.42	26.21
		28.89	1.10

はニュース速報手話コーパスの評価データでそれぞれ学習途中のモデルでの BLEU 値を算出し、最大 50 epoch の学習のうちの性能が最大となるものを最終的なモデルとした。ニュース速報手話コーパスは分量が少ないため、fine-tuning 用の開発データを用意せず、評価データでの BLEU 値が最大となるモデルを用いて性能を評価した。なお、ニュース速報手話コーパスでベースモデルを学習したモデルでは fine-tuning をせず、ニュース速報手話コーパスの評価データでの BLEU 値が最大となるモデルで性能を評価した。それぞれのモデルはランダムシードを変えて 3 回の実験を行い、その平均の BLEU 値により評価した。

学習データと開発データ、評価データのそれぞれの数量は表 2 に示す。以下の図表中では手話ニュースコーパスを「手話ニュース」、ニュース速報手話コーパスを「ニュース速報」と表記する。

### 3.2 実験結果

表 3 に実験結果を示す。BLEU(ニュース速報) はニュース速報文に対する BLEU 値を、BLEU(手話ニュース) は手話ニュースコーパスの一般的なニュース文に対する BLEU 値をそれぞれ表す。

ベースラインとなるのは、ベースモデルを手話ニュースコーパスのみで学習し、fine-tuning をしないモデルである。ニュース速報手話コーパスを学習に用いることで、ベースラインと比較してニュース速報文に対する翻訳性能が大きく向上した。ニュース速報文の翻訳性能が最大となったのは、ベースモデルの学習に両方のコーパスを用い、さらに fine-tuning をした場合で、ベースラインから 33.65 ポイントの BLEU 値の向上となった。一方で、これらのモデルの間で手話ニュースコーパスに対する翻訳性能には顕著な差は見られなかった。

### 3.3 考察

学習データによる翻訳性能への影響の詳細を調査するための分析を行った。以下では評価実験で最良の結果となった、両方のコーパスを用いてベースモデルを学習し、さらにモデルを fine-tuning する学習手法を用いる。

3.2 節の実験で最良の性能となったモデルを用い、ニュース速報、手話ニュースのそれぞれを翻訳した際の文ごとの BLEU 値の分布を図 1 に示す。図から、ニュース速報と手話ニュースでは分布に大きな差があることがわかる。BLEU が 70 を超えるような文は手話ニュース文にはほぼ存在しないが、ニュース速報文では全体の 3 割近くを占める。このような高精度で翻訳できた文は、

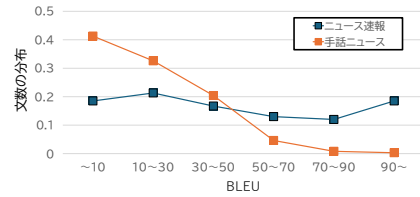


図 1 文ごとの BLEU 値の分布

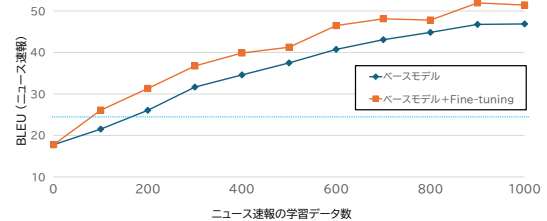


図 2 ニュース速報手話コーパスの学習データ数を変化させた場合の翻訳性能の変化

類似した文が学習データに含まれていたものと考えられる。ニュース速報では定型的な表現が使われることが多いため、翻訳対象の文と類似した文が学習データに含まれる可能性が高く、それらの文については高精度に翻訳できたものと考えられる。一方で、手話ニュースでは定型的な文があまり使われないため、翻訳対象の文と類似した文が学習データには含まれず、ニュース速報文に見られるような高い BLEU 値となる文がほぼ存在しなかったものと考えられる。

ニュース速報手話コーパスの学習データ数を 100 刻みで変化させた場合の翻訳性能の推移を図 2 に示す。図からは、概ね学習データが 700 程度までは線形に性能が向上しているが、それを超えたところで翻訳性能の向上がやや鈍化した。これは、学習データを単純に増やしただけでは今までの学習データにすでに類似の文があるような文章が多く追加されてしまい、翻訳できる情報の幅を広げることができなかったことが原因と考えられる。これらのことから、より幅広いニュース速報文を翻訳できるようにするためには、単純に学習データの数を増やすのではなく、より多様なニュース速報文を収集することが必要と考えられる。

### 4 おわりに

本稿ではニュース速報に特化した日本語-手話翻訳について述べた。一般のニュースとニュース速報では文体が大きく異なることから、ニュース速報を精度良く機械翻訳するためには、ニュース速報に特化したデータを収集する必要があることがわかった。今回作成した約 1,000 文のニュース速報に特化したコーパスではカバーできないニュース速報文もあるため、今後は翻訳できる情報の幅を広げるために、より多様なニュース速報文を収集することが必要である。

#### 参考文献

- [1] T. Uchida et al. "Systems for supporting deaf people in viewing sports programs by using sign language animation synthesis." ITE Transactions on Media Technology and Applications 7.3 (2019).
- [2] A. Vaswani et al., "Attention is all you need." Advances in neural information processing systems, 30 (2017)