

# 聴覚障がい者の情報格差解消のための文書要約モデル ～分散モデルと階層モデルの比較～

丸田 要<sup>1</sup>

## Document Summarization Model for Elimination the Information Gap with Deaf Person ～ Distributed Model vs. Hierarchical Model～

MARUTA Kaname<sup>1</sup>,

(Accepted October 30, 2024)

**概要** 近年は、インターネットや携帯端末の普及の影響から経済・社会・生活のあらゆる場面で情報化が進んでおり、老若男女の様々な人々が最新の情報を手軽に利用できる環境である。しかし、聴覚障がい者にとっては未だに情報を障害の無い人と同様に情報を得ることは難しい状況である。特に、手話放送や字幕放送も一部では実施されているが全ての放送に対応しておらず、災害情報といった速報情報や、駅や空港での遅延・運休情報などはまだまだ情報格差が大きい状況である。そこで、本研究では、T5 モデルによる聴覚障がい者の学習レベルに応じた多様な表現の複数文を生成する要約モデルを 2 つ提案する。これにより、聴覚障がい者の情報格差解消に寄与すると考えられる。

**キーワード** [要約、翻訳、情報格差、聴覚障がい者、ろう者、手話、字幕、T5、Beam Search]

## 1 はじめに

### 1.1 研究背景

現在、世界及び日本は高度な情報化社会となっており、多くの情報が展開されている。私たちは、その多様な情報をインターネットから能動的に自ら収集することが出来る。さらに、緊急性のある情報は J アラートや L アラートをエリアメールやテレビの速報などから受動的に得られるようにシステム化されてきている。

しかし、東日本大震災では障がい者の死亡率は障害のない人よりも多いという調査結果がある。NHK が 2011 年 9 月 11 日に報道した内容では約 2 倍の多さであり、宮城県が 2012 年 3 月 23 日に発表した内

容<sup>注1</sup>では約 4.3 倍の多さであると報告している。この原因は様々な事が考えられるが、初期避難を促す危険性や重大性を伝える情報の格差も原因の 1 つであると考えられる。

また、障害者情報アクセシビリティ・コミュニケーション施策推進法<sup>注2</sup>が令和 4 年 5 月 25 日に公布・施行された。この法律は、全ての障がい者があらゆる分野の活動に参加するために、情報の十分な取得利用・円滑な意思疎通を円滑にすることを推進するものである。この法律が施行されてから、地方局の字幕放送を 2027 年度までに 80%にすることが目標値<sup>注3</sup>とされている。現在は、字幕放送普及対象の放送番組に限定するとキー局では 100%であり、地方局では 86.6%と高い普及率<sup>注5</sup>である。しかし、手話放送

1 都城工業高等専門学校 電気情報工学科

National Institute of Technology, Miyakonojo College

の普及率は NHK（教育）で 2.87%であり、在京キー5局に至っては 0.17%しか普及していない。さらに、聴覚障がい者にとっては障害の度合いや文字学習進度によって文章認識に多様性があり、一律の表現での字幕放送では解消できない情報格差が依然存在している。

このように、障がい者の情報格差が大きな社会問題の1つである事が分かる。

### 1.2 研究目的

既存の字幕放送、緊急情報のテロップや公共交通機関の電光掲示板での情報提示では、聴覚障がい者の障害の度合いや文字学習進度による文章認識の多様性を考慮せずに情報を平叙文で表現している。

そのため、本研究では、元のテキストを2つの要約・翻訳モデルに適用することで表現方法の異なる複数のテキストを生成し、多様な字幕情報を提供する。それにより、聴覚障がい者への情報格差を解消することに寄与する事が基本的な目的である。

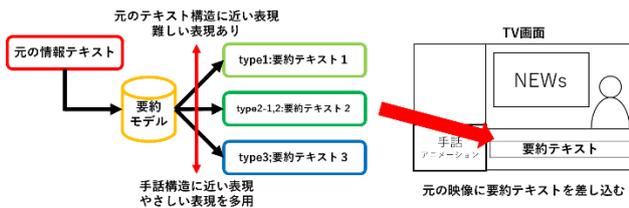


図 1. 多様な表現の要約文による自動字幕

本研究では、生成した多様性のある複数のテキストを図 1 のように差し込む形式での情報提供を構想している。この時、多様な表現の文で生成した複数のテキストの中から各聴覚障がい者の方にとって理解しやすいタイプのテキストを選択することで理解度の多様性に対応出来ると考えている。

表 1.分散的要約モデル

要約文タイプ	特徴
type1	やさしい日本語表現
type2-1	小学校の教科書で使用される単語を多用した表現
type3	日本手話構造に近づけた表現

表 2.階層的的要約モデル

要約文タイプ	特徴
type1	やさしい日本語表現
type2-2	日本語の読解に重要な語彙を重視した表現
type3	日本手話構造に近づけた表現

また、2つの要約モデルを提案するが、各要約モデルにおける要約文の特徴を表 1 と表 2 に示す。

分散的要約モデルと階層的的要約モデルにおいて3つのタイプの要約文を生成するが、両モデルにおいて、type1 と type3 の特徴は同じである。

それぞれ特徴が異なるこの3つのタイプのテキスト文を生成する自動翻訳モデルを作成し、聴覚障がい者の学習タイプに応じた多様な表現や文タイプでの情報伝達手法を提供することで聴覚障がい者の情報格差を解消することが具体的な研究目的である。

## 2 関連研究

### 2.1 やさしい日本語

「やさしい日本語」とは、1995年の阪神・淡路大震災を経て多くの外国人にも素早く的確に情報を伝えることを目的に考案されたものである。その「やさしい日本語」での表現でニュースを作成するシステム<sup>3)</sup>を田中らが提案している。この「やさしい日本語」では、やさしい漢字やルビを多用することで理解しやすい表現をしている。本研究では、この「やさしい日本語」データセットを type1 表現での教師データとして活用する。

### 2.2 日本語を読むための語彙データベース

松下言語学習ラボ<sup>5)</sup>で公開されている「日本語を読むための語彙データベース」<sup>4)</sup>は読解に必要な語彙力や、読解テキストの語彙の難易度を測るための基礎的なデータベースである。

このデータベースは書籍やYahoo知恵袋で使われる重要な60894語の語彙を、使用頻度と使用範囲・分散度のデータから算出した日本語を読解能力における重要度が付与されている。本研究ではこの各語彙の重要度におけるランキングを用いている。また、このデータベースでは、分野別の統計情報や旧日本語能力試験出題基準レベルを1から4までの4段階もされている。

### 2.3 日本手話と日本語対応手話

日本における手話は日本手話、中間型手話、日本語対応手話の3種類が存在している。長南<sup>9)</sup>がこの3種類の手話について整理している。

日本手話は、日本語とは異なる独自の統語構造を持っている。これは、日本手話が聴覚障がい者のコミュニティの中で自然発生的に成長してきた手話だからであり、そのため聴覚障がい者の中で日常的に用いられる1番の手話である。

中間型手話は、日本語の統語構造に従って手話単

語が表現されている。そして、日本語対応手話は、中間型手話に指文字で日本語の格助詞などの付属語を付加した手話である。中間型手話と日本語対応手話は、手話を学習している聴者が用いることが多いとされている。

この3種類の手話は成り立ちや使用されるコミュニティの違いや統語構造の異なりがあり、聴覚障がい者にとっては基本的に日本手話が最も理解しやすい手話であることが研究結果<sup>9)</sup>で判明している。

## 2.4 情報提示システム

CS放送局である「目で聴くテレビ」では、テレビ放送へ「字幕と手話」を画面合成により付与することを可能にした「アイ・ドラゴン II」<sup>5)</sup>の開発及び利用をしている。

浅井らは、聴覚障がい者の音声だけではなく、周囲の環境音が聞こえないという困難を抱えていることに着目している。この問題を解決するために、周囲の音を自動認識して必要な情報のみを聴覚障がい者に通知する環境音可視化システム<sup>6)</sup>を提案している。

また、松村らはマルチモーダル環境音認識に基づく擬音語変換と音声認識を用いた聴覚障がい者支援システム<sup>7)</sup>を提案している。具体的には、周囲の音声や環境音を自動認識し、発話音声や環境音を文字列や擬音語へ変換する。変換された文字列や擬音語はARヘッドセットを通じて提示される。これにより、発話音声や環境音を視覚的に認識できるようになることを目的としていた。

このように聴覚障がい者も障害がない人と同じように情報を提供することは重要であり社会的関心度も高い。

## 2.5 T5モデル

T5モデル<sup>2,8)</sup>とは、Text-to-Textモデルであり、要約、翻訳や分類など異なる自然言語処理タスクの問題を統一的なフレームワークで扱うことが出来る言語モデルである。Text-to-Textモデルであり、入力文にタスク命令を表すPrefixを付与した「タスクPrefix:入力文」の形式で入力することでタスクに応じた出力が出来るモデルである。図2にT5モデルの入力出力の形式を示す。

本研究では、元のニュース文を多様なタイプに変換して表示することが目的となる。この変換が要約でもあり翻訳でもあるため、要約特化や翻訳特化のモデルではなく、要約と翻訳の両タスクを学習したT5モデルをベースとしてファインチューニング<sup>1)</sup>することが本研究目的に沿っていると考えた。

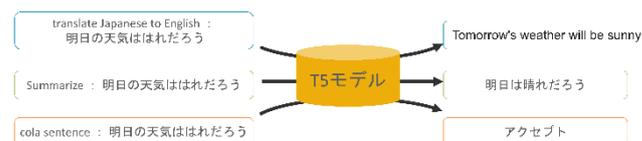


図2. T5モデル

## 3 提案手法

第1.1節の研究背景で述べたように字幕放送の普及率はキー局及び地方局共に比較的高くなってきている。しかし、生放送や緊急性の高い速報情報への普及率は未だ低く対応が難しいという状況である。さらに、聴覚障がい者の日本語の読み能力の学習進捗にはバラつきがある。このバラつきの原因としていくつか考えられている。1つ目は、どうしても聞こえる人よりも聴覚障がい者にとっては情報が少なく学習難易度が比較的高くなるからである。2つ目は、第2.3節に明記した通り、日本手話と日本語とでは文法が異なり、極端な事を言えば別言語となり手話を学習すれば手話を学習した人同士では問題なく会話やコミュニケーションを取ることが可能だからである。このような理由から、聴覚障がい者の方は1人1人で認識できるテキストの語彙数が異なる。これは、日本語の語彙数は約25万語であるのに対して手話の語彙数は約8000語であり、語彙数に大きな差が存在している。さらに、日常会話に必要な単語数で考えても日本語は約5000語であるのに対して手話は約500語とされている<sup>9)</sup>。

そのため、障害がない人と同じように豊富な語彙数を用いたテキスト表現では情報格差を生じてしまう。そこで、図1のように情報を3種類のテキストタイプに要約や翻訳を行い、生成した3タイプのテキストを字幕形式で情報の表示を行う。サイズを縮小して挿入したTV画面との余白に生成したテキストを出力することで情報を獲得しやすくする。

### 3.1 分散的要約モデル

本論文では、計算コスト削減のために分散的要約モデルを提案する。この分散的要約モデルでは、type1, type2-1とtype2の3タイプの表現へ分散処理にて要約を実施する。図3に分散要約モデルの構成を示す。分散的要約モデルでは、要約タスクと翻訳タスクを同時に処理するモデルである。これにより、第3.2節で述べる階層的な要約モデルよりも計算コストが少ないモデルを構築することが出来る。

各タイプへの要約・翻訳手法について述べる。まず、type1への要約では、元の情報テキストからやさしい日本語表現へファインチューニングしたT5モデルにより要約・翻訳処理を適用して要約文を生

成する。

type2-1 への要約では、元の情報テキストから小学校・中学校・高校の教科書における単語の使用頻度を活用する。高校で多用されており、小学校では多用されていない単語は難しい表現であり、小学校や中学校で多用されている単語は易しい表現であると仮定して、その易しいと判断された単語を多用するように要約・翻訳を行った。特に、使用した学習済みの T5 モデルの途中にある BeamSearch を改良することで実施している。

type3 の要約では、元の情報テキストから日本語の構造へ近づけるように要約・翻訳を実施している。ここでは T5 モデルを日本語から日本語へ翻訳できるようにファインチューニングして日本語構造の要約文を生成している。

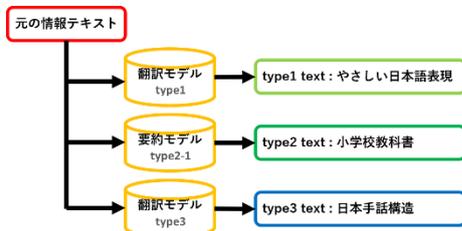


図 3. 分散的要約モデル

### 3.2 階層的要約モデル

本論文では、多様な表現の要約テキストを生成するために3タイプのテキスト生成に図4のような階層的な要約モデルを提案する。

第3.1節で述べた分散的要約モデルのように要約処理と各表現への翻訳処理を一度に適用するのではなく、要約部分から各表現への翻訳での階層的な要約モデルにする。これにより、長文に対する精度の高い要約が期待出来る。特に、階層的な要約モデルは4段階目の手話構造への翻訳処理と相性が良いと考えられる。なぜなら、難しく普段からあまり使用されない単語に対する手話のコーパスは作成が困難であるが、2段階目や3段階目のやさしい日本語表現への翻訳を前提処理とすることで、手話コーパスに含まれる単語を対象として4段階目の翻訳処理が可能となることが期待出来るからである。

各階層的な要約・翻訳手法について述べる。まず、1段階目として元の情報テキスト $S_0$ から既存の T5 モデルを用いて要約テキスト $S_1$ を生成する。

そして、2段階目では要約テキスト $S_1$ を入力テキストとしてやさしい日本語でファインチューニングした翻訳モデルで要約テキスト $S_2$ を生成する。この要約テキスト $S_2$ を本手法の type1 とする。

3段階目も要約テキスト $S_1$ を入力テキストとする。

3段階目では、2段階目と同様のやさしい日本語でファインチューニングした翻訳モデルに日本語を読解するのに重要な単語がより生成されるように改良した BeamSearch を適用する。これにより要約テキスト $S_3$ を生成する。この要約テキスト $S_3$ を本手法の type2-2 とする。

4段階目では、要約テキスト $S_3$ を入力テキストとして、手話辞書をファインチューニングにて学習した翻訳モデルで要約テキスト $S_4$ を生成する。この要約テキスト $S_4$ を本手法の type3 とする。

これにより、平叙文での要約テキスト $S_1$ の他に3タイプの要約テキスト $S_2, S_3, S_4$ を生成する

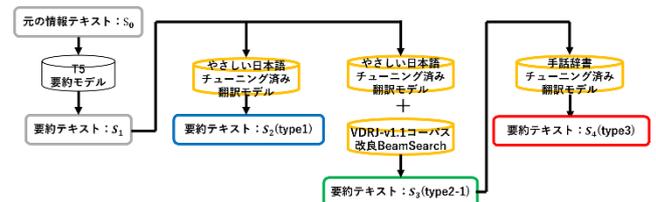


図 4. 階層的な要約モデル

### 3.3 type1 表現 : やさしい日本語

多くの言語に対応している MT5 を数種類の言語に絞ってファインチューニングすることで各言語での精度を向上させた K024/mt5-zh-ja-en-trimmed モデル(以下 K024 モデル<sup>注6</sup>)を使用する。日本語を第2.1節で述べたやさしい日本語に翻訳した対訳コーパス T15 を使ってファインチューニングする。コーパスの一部を図5に示す。

原文 日本語	: この事実を心に留めておいて下さい。
やさしい日本語	: この事実を覚えておいてください。
原文 日本語	: 申告する物は何もありません。
やさしい日本語	: 言っておく物は何もありません。
原文 日本語	: 渋滞にひっかかった。
やさしい日本語	: 混んでいる道路に入った。

図 5. やさしい日本語と日本語の対訳コーパス例

### 3.4 type2-1 表現 : 小学校教科書

type2-1 では、「現代日本語書き言葉均衡コーパス」の中の教科書コーパス語彙表を用いる。この教科書コーパスには各語彙に対して、小学校、中学校及び高等学校のそれぞれでの出現頻度が格納されている。そこで、次単語 $\omega_{i+1}$ の各学校での出現頻度を Transformer の Decoder 部分における BeamSearch での文生成の次単語予測のスコアにバイアスを足す。これにより、小学校や中学校でよく使用される単語をより生成しやすくし、反対に高等学校でよく使用される単語は生成されにくくなる効果が期待できる。式(1)において、 $score()$ は BeamSearch の次単語予

測スコアである。また、 $freq\_j(\omega)$ は小学校の教科書、 $freq\_m(\omega)$ は中学校の教科書及び、 $freq\_h(\omega)$ は高等学校の教科書における単語 $\omega$ の出現頻度である。

$$score(\omega_i, \omega_{i+1}) = score(\omega_i, \omega_{i+1}) + freq_j(\omega_{i+1}) + freq_m(\omega_{i+1}) - freq_h(\omega_{i+1}) \quad (1)$$

### 3.5 type2-2 表現：日本語の読解に必要な重要度

type2-2 では、第 2.2 節で示したデータベースを用いる。このコーパスで整理されている 60894 語の語彙に対する日本語を読解する上で重要とされる順位から BeamSearch を改良するためのバイアスを算出する。

まず、各単語の重要度での順位の値を 0 - 1 への範囲に式(2)を適用し正規化する。この時、順位が低い単語が 0 に近く、順位が高い単語は 1 に近づくように正規化を行う。式(2)における  $Rank\_max$  は順位の最大値であり、 $rank()$  は単語  $\omega_i$  の重要度による順位の値を求める関数である。

さらに、正規化した値を式(5)の正則化不完全ベータ関数を用いてバイアス値を算出する。式(3)はベータ関数であり、式(4)は不完全ベータ関数である。また、式(5)のグラフが図 6 である。図 6 の横軸が  $r\_norm(\omega_i)$  であり、縦軸が  $I_{\omega_i}(\alpha, \beta)$  である。式(5)により、順位が高い単語ほどバイアス値が高くなり、順位が低い単語ほどバイアス値が低くなるように算出される。

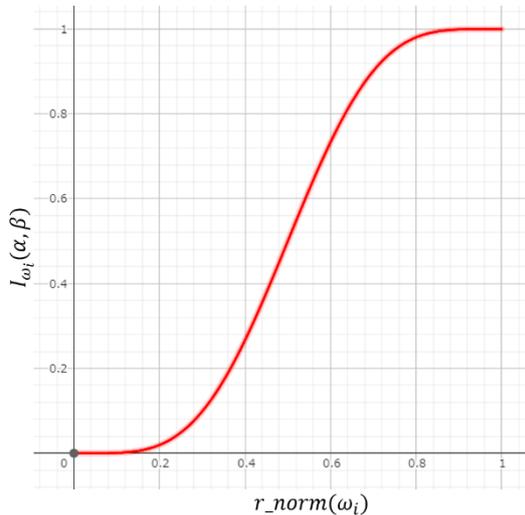


図 6. 正規化不完全ベータ関数  $\alpha = 5, \beta = 5$

式(6)のように Transformer の Decoder 部分における BeamSearch での文生成の次単語予測のスコアに式(5)で算出したバイアスの値を足す。これにより、日本語を読解するのに重要な単語をより生成しやすくし、日常で使用する機会が少なく重要度が低い単

語は生成されにくくなる効果が期待できる。式(6)において、 $score()$  は BeamSearch の次単語予測スコアである。また、 $\mu$  は重みである。

$$r_{norm(\omega_i)} = \frac{Rank\_max - rank(\omega_i)}{Rank\_max} \quad (2)$$

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt \quad (3)$$

$$B(\omega_i, \alpha, \beta) = \int_0^{r_{norm(\omega_i)}} t^{\alpha-1} (1-t)^{\beta-1} dt \quad (4)$$

$$I_{\omega_i}(\alpha, \beta) = \frac{B(\omega_i, \alpha, \beta)}{B(\alpha, \beta)} \quad (5)$$

$$score(\omega_i, \omega_{i+1}) = score(\omega_i, \omega_{i+1}) + \mu * I_{\omega_i}(\alpha, \beta) \quad (6)$$

### 3.6 type3 表現：日本手話構造

type1 で表現したテキストを生成するのに使用した K024 モデルを type3 でも使用する。type3 では、「新・日本語-手話辞典」<sup>10)</sup>における例文を日本語-手話文の対訳コーパスとして K024 モデルをファインチューニングする。コーパスの一部を図 7 に示す。図 7 において、<と> は手話に対応したトークンである。手話の単語を日本語で表現すると、日本語の単語と表面上同じになり T5 モデル内で同じトークンとして学習してしまい、間違った学習が進む危険性がある。その危険性を防ぐために、<や>を用いて別の単語として扱う。

原文 日本語：	親が娘に愛情を注ぐ
日本手話	：<両親><娘><かわいがる>
原文 日本語：	あそこまで言わなくてもいいのに
日本手話	：<言う②><いらぬ>
原文 日本語：	必ずしも晴れるとはかぎらない
日本手話	：<明るい②><定まる><ない②>

図 7. 日本語と手話の対訳コーパス例

## 4 実験

分散的要約モデルでは LivedoorNew を 3 タイプに要約し検証する。階層的な要約モデルは要約前のテキスト  $S_0$  と、事前要約及び 3 つのタイプに要約及び翻訳して生成した各テキスト  $S_1, S_2, S_3, S_4$  を比較し効果を検討する。実験で要約するテキストは LivedoorNews のニューステキストと、テレビ放送の音声コーパスである LaboroTVSpeech2<sup>注7)</sup> の字幕データを使用する。LivedoorNews は文語体であり、LaboroTVSpeech2 のデータは口語体である。本実験では、式(5)において  $\alpha = 5, \beta = 5, \mu = 0.01$  とした。

### 4.1 データセット

type1 ではやさしい日本語の T15 コーパスを, type2-1 では BCCWJ コーパスを, type2-2 では VDRJ-v1.1 コーパスを, type3 では新・日本語-手話辞典の用例を使用した. 表 3 に type1~type3 で使用したコーパスにおける文数やトークン数を示す.

表 3. データセットの詳細

コーパス	個数
T15 の文数(train)	40000
T15 の文数(validate)	10000
BCCVJ	9255
VDRJ-v1.1	60894
手話辞典の語彙数	4283

### 4.2 ファインチューニング

type1 と type3 では, 学習済みのモデルをそれぞれファインチューニングしている. 分散的要約モデルにおいてファインチューニングした際の学習曲線を図 8 と図 9 に示す. 同様に, 階層的的要約モデルの学習曲線を図 10 と図 11 に示す

学習が進むにつれ, loss 値が下がっており Train データに対しては学習が来ていることが分かる. また, 構造が日本語とは異なる手話の方は学習曲線の収束が遅く loss 値の収束値も高いことが分かる.

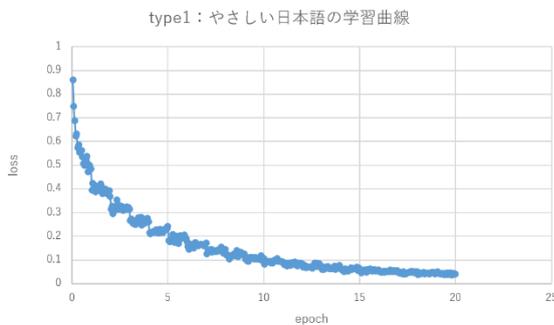


図 8. 分散的要約モデル type1 の学習曲線

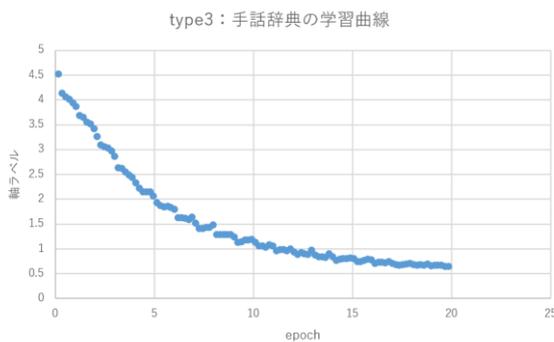


図 9. 分散的要約モデル type3 の学習曲線

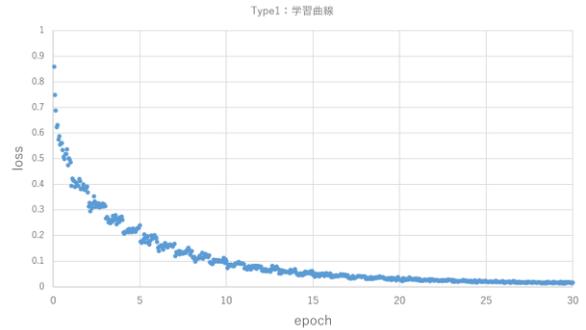


図 10. 階層的的要約モデル type1 の学習曲線

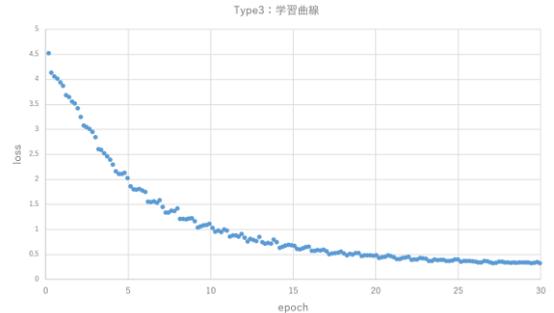


図 11. 階層的的要約モデル type3 の学習曲線

### 4.3 分散的要約モデルの結果

図 12 に Livedoor News コーパスにあるトピックニュースを分散的要約モデルによって 3 つのタイプに翻訳した結果を示す.

図 12 から, 原文を基に 3 つの異なったタイプの表現に要約及び翻訳されていることが分かる. 但し, 実際に聴覚障がい者にとって分かりやすいかを定量的に評価することができていない.

図 12 以外の結果では, type1 は難しいと思われる表現も出力された. type2 でも高等専門学校でよく使用するような難しい漢字が出力されていた. type3 においては, 手話辞典に収録されていた<>付きの単語以外が出力されることが見られる.

原文: 原子力基本法の改正に、韓国から「日本の核武装を阻止すべき」「我々も核武装だ」の声 20日、日本の原子力基本法に「安全保障」条項が追加されたが、韓国では「日本が核武装を企んでいる」と、メディアやネット掲示板などで波紋を呼んでいる。 21日、複数の韓国メディアは同日付け東京新聞の記事を引用し、…
type1: 原子力基本法の改正に、韓国から「日本の核武装を止めるべき」、「私たちも
type2: 韓国の藤村修官房長官は21日、日本の原子力基本法に「日本の核武装を阻止すべき」と述べた。
type3: <原子> <憲法> <規則> <決める①> <発表>

図 12. 分散的要約モデルの結果

### 4.4 階層的的要約モデルの結果

図 13 に LivedoorNews のニューステキストを対象として事前要約したテキストと 3 つのタイプに要約・翻訳したテキストの結果を示す. 同様に図 14 に LibroTVSpeech2 の字幕データを対象とした実験結

果を示す。

図 13・14 から、原文を基に 3 つの異なったタイプの表現に要約及び翻訳されていることが分かる。但し、実際に聴覚障がい者の人にとって分かりやすい表現を生成出来ているかを定量的に評価することができていない。

また、 $S_1$  からやさしい日本語表現に翻訳した  $S_2$  を見ると「創設者」から「作った者」や「温帯低気圧」から「非常に低い気圧」へ変換されていることが分かる。さらに、 $S_1$  からやさしい日本語表現に翻訳する際に日本語の読解に重要な単語をより生成するように改良した BeamSearch で生成した  $S_3$  を見ると「共有」から「一緒」や「広範な」から「広い」、「不確定」から「不決定」に変化している。そして、 $S_3$  から手話構造へ翻訳した  $S_4$  を見ると「作った」から「<作る>」や「気圧」が「<気圧>」へ翻訳されている事が確認出来る。

図 13・14 以外の結果では、type1 は「ライフセーバー」が「命を確実に守る」に翻訳されるなど元の単語の方が分かりやすいと思われる表現への翻訳も見られた。type2-2 では、type1 と同じテキストが生成されたケースが確認された。 $S_2$  と  $S_3$  を比較すると、LivedoorNews において 74.8%、LaboroTVSpeech2 の字幕データでは 67.1% のテキストで変化が見られた。type3 においては、手話辞典に収録されていた <> 付きの単語以外が出力されることが見られた。

原文 $S_0$ : 2ちゃんねるの創設者ひろゆきの発言で波紋 2ちゃんねるの創設者として知られる西村博之氏がtwitter上の発言が話題を呼んでいる。問題の発言は10月12日につぶやかれた内容である。アニメキャラやタレントの写真をアイコンにしてる入って…
T5要約モデル $S_1$ : 動画共有サイト「2ch」の創設者、広範な発言が話題になっている。
Type1 $S_2$ : 動画を共有するサイト「2ch」を作った者、広範な話が話になっている。
Type2 $S_3$ : 動画を一緒にするサイト「2ch」を作った者、広い話が話になっている。
Type3 $S_4$ : <テープ②> <一纏①> <作る> <ある①>

図 13. 階層的な要約モデルの結果(LivedoorNews)

原文 $S_0$ : まだ予報円がだいぶ大きいのでどこまで陸地に近づくか不確定ですしまこの辺りに来るとだんだん温帯低気圧化して行くことにはなるかと思えます
T5要約モデル $S_1$ : 温帯低気圧で陸地に近づくか不確定だ。
Type1 $S_2$ : 非常に低い気圧で陸に近づくか不確定だ。
Type2 $S_3$ : 熱帯の低い気圧で陸に近づくか不決定だ
Type3 $S_4$ : <暑い①> <気圧> <あたり> <知らない>

図 14. 階層的な要約モデルの結果(Laboro)

## 5 考察

### 5.1 分散的要約モデルについて

type1～type3 の共通した問題のある箇所は、固有名詞に対してやさしい表現であることや、手話構造

で表現することが出来ていないことである。これは、そもそも固有名詞のやさしい表現がコーパス化されておらず、手話としても存在しない固有名詞があるからだと考えられる。

type1 では、図 12 以外の結果も併せて考えると、よりやさしい表現へ翻訳するのが望ましいと考えられる。type2-1 においては、式(1)の小学校、中学校及び高等学校の教科書での出現頻度が同等の重みで計算されているため、難しい漢字のスコアを下げる事が出来なかったと考えられる。そのため、各学校に重みパラメータを付与して調整する必要がある。type3 は、学習させた語彙数が未だ少ないと考えられる。さらに、学習させた辞典の用例文は短文しかないため、長文の翻訳における精度が低くなってしまったと考えられる。

### 5.2 階層的な要約モデルについて

提案手法の階層的な要約モデルの結果を調査した所、原文の  $S_0$  と最後段の  $S_4$  において文の意味が大きく変化してしまっているケースも見られた。これは、各階層の処理を経る度にニュアンスが変化していき処理が積み重なったことが原因と考えられる。ただし、手話辞書のコーパスは難しい表現に対するデータが不足しており、階層的に手話構造へ翻訳することで手話辞書のコーパスが対応出来るケースが増加している。特に、階層的な要約モデルにおける type3 表現では、分散的要約モデルよりも翻訳出来ないケースが減少している。

type1 では、図 13・14 以外の結果も併せて考えると翻訳する前の方がやさしい表現の場合や難しい表現がそのままである事もあるためよりやさしい表現へ翻訳出来る方が望ましいと考えられる。

type2-2 では式(6)によるバイアスの計算によって Beam-Search での生成文を調整している。この式(6)の重み  $\mu$  やパラメータの  $\alpha$  や  $\beta$  のパラメータチューニングを行うことで、より分かりやすい表現のテキストを生成できると考えられる。

type3 は、特に LivedoorNews のような長文のデータに対して短いテキストを生成している。日本手話のコミュニケーションでは、あまり長文はなく短文でやり取りすることが多い。しかし、本手法による手話構造の  $S_4$  は原文  $S_0$  よりもあまりにも短いため意味が十全に伝わらない懸念がある。これは、学習させたコーパスの語彙数や対訳文数が未だ少ないためと考えられる。さらに、学習させた辞典の用例文は短文しかないため、長文の翻訳における精度が低くなってしまったと考えられる。

## 6 おわりに

本提案手法の分散的要約モデルと階層的的要約モデルにより異なる表現のテキストを生成することが出来た。しかし、各表現における手法には改良の余地が見られた。

今後の展望としては、まず、type1では、ファインチューニングの再調整を行う。次にtype2-2は改良BeamSearchにおけるパラメータチューニングを実施する。そして、type3においては手話のコーパスを改良することが必要である。

また、要約及び翻訳して生成した3タイプのテキストが、聴覚障がい者にとって理解しやすい表現になっているか調査を行う。

さらに、手話放送に至ってはキー局及び地方局共に普及率はかなり低い状況である。そのため、聴覚障がい者によってはテキストよりも普段から利用している手話での情報提供を望まれる方も居る。これは、実際のコミュニケーションに利用している話者の方がテキストよりも認識速度が早いという理由もある。このように可能なら手話での情報提供がある方が望ましい状況である。そのため、図1のように本手法で生成したtype1, type2-1, type2-2の要約テキストと一緒にtype3から手話のアニメーションを生成し表示する事を検討していく。

## 注

- 1) <https://www.dinf.ne.jp/doc/JDF/20120323miyagi/miyagikenhisai.html>
- 2) <https://www8.cao.go.jp/shougai/suishin/jouhousyutoku.html>
- 3) 一般財団法人全日本ろうあ連盟「放送分野における情報アクセシビリティに関する指針」見直し検討に向けた要望
- 4) <https://barrierfree.nict.go.jp/nict/program/ratejimaku.html>
- 5) <http://www17408ui.sakura.ne.jp/tatsum/>
- 6) <https://huggingface.co/K024/mt5-zh-ja-en-trimmed>
- 7) <https://laboro.ai/activity/column/engineer/laborotvspeech2/>

## 謝辞

NPO 法人「手話ランゲージ」の代表である大橋正敏様、宮城県聴覚障害者情報センターの松本隆一様と庄子陽子様には聴覚障がい者の実情調査及び本研

究へのご助言をいただき深く感謝しております。

## 参考文献

- 1) A.Radford, K.Narasimhan, T.Salimans, I.Sutskever, “Improving Language Understanding by Generative Pre-Training.”, OpenAI Blog, 2018.
- 2) T.Goyal, J.J.Li, G.Durrett, “News Summarization and Evaluation in the Era of GPT-3”, Computation and Language, arXiv.2209.12356, 2022.
- 3) 田中英輝, 熊野正, 後藤功雄, 美野秀弥, “やさしい日本語ニュースの製作支援システム”, 自然言語処理, Vol.25, No.1, pp.81-117, 2018.
- 4) 松下達彦, “日本語を読むために必要な語彙とは? 一書籍とインターネットの大規模コーパスに基づく語彙リストの作成”, 日本語教育学会, 2010 年度 日本語教育学会春季大会 予稿集, pp.335-336, 2010.5.
- 5) 佐藤至, “聴覚障害者を対象とした字幕・手話配信サービスによるアクセシビリティ向上の活動～「目で聴くテレビ」の配信サービスと専用受信器「アイ・ドラゴン」～”, 映像情報メディア学会誌, Vol.69, No.9, pp.682-688, 2015.
- 6) 浅井研哉, 綱川隆司, 西田昌史, 西村雅史, “聴覚障害者支援のための環境音可視化システムの開発”, 研究報告アクセシビリティ (AAC), 2019(5), pp.1-8, 2019.
- 7) 村松遼, 石津龍真, 岡村一矢, 田口創, 栗原佑弥, 峯千瑛, 北風裕教, “マルチモーダル環境音認識に基づく擬音語変換と音声認識を用いた聴覚障がい者支援システムの試作”, 産業応用工学会論文誌, Vol.10, No.2, pp.98-106, 2022.
- 8) C.Raffel, N.Shazeer, A.Roberts, K.Lee, S.Narang, M.Matena, Y.Zhou, W.Li, P.J.Liu, “Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer” Journal of Machine Learning Research 21, 2020.
- 9) 長南浩人, “日本手話, 中間型手話, 日本語対応手話の構造の違いが聴覚障害者の手話の理解に与える影響”, The Japanese Journal of Educational Psychology, No-49, pp.417-426, 2001.
- 10) 米川明彦, “新・日本語-手話辞典”, 編集: 社会副詞法人全国手話研修センター 日本手話研究所, 製作: 中央法規出版, 2011.