

日本語と日本手話の対訳コーパス開発のための 手話動作解析および対応調査

丸田 要¹ 久留 脩真¹¹ 都城工業高等専門学校 電気情報工学科

marutak@cc.miyakonojo-nct.ac.jp

概要

近年は高度情報化社会と言えるが、聴覚障害者にとっては未だに聴者と同様に情報を得ることが難しい状況にある。例えば、手話放送や字幕放送も放送されているが、手話放送の普及率はNHKでも2.8%である。字幕放送は、普及率がキー局で100%、地方局で86.6%であるが、聴覚障害者にとっては理解しやすい表現ではない。これは、特に先天的難聴者は学習進度にバラつきがあり、普段から使う日本手話には無い日本語の表現などが多々存在していたためである。つまり、現在の字幕放送の表現では情報格差の解消には不十分である。また、現状は、顔の表現や手の位置、動作の方向、感情によって意味が異なるという特性を考慮した日本語と日本手話の対訳コーパスデータが存在しない。そこで、日本語と日本手話の対訳コーパスの開発が必要となってくる。そのために、日本手話を普段から使用している聴覚障害者の方に協力してもらい日本手話の動作を撮影する。そこで、日本語と日本手話の対訳コーパスを開発するために必要な顔表情を含む上半身での手話動作の3次元データの計測を行う。これにより、聴覚障害者の情報格差解消に寄与すると考えられる。

1 はじめに

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を通じて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害者の死亡率は障害の無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¹⁾では約4.3倍の多さであると報告している。この原因は様々なことが考えられるが、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²⁾が令和4年5月25日に公布・施行された。この法律は、全ての障害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・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³⁾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⁴⁾である。しかし、手話放送の普及率はNHK(教育)で2.87%であり、在京キー5局に至っては0.1%しか普及していない。さらに、聴覚障害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。情報格差を解消するシステムとして、CS放送局の「目で聴くテレビ」[1]や浅井らの環境音可視化システム[2]、村松らのマルチモーダル環境音認識[3]などがある。しかし、日本手話を用いた情報提供の拡充を行うためには、日本語から日本手話への自動翻訳技術が必要となる。本著者は、手話辞書[6]をコーパスとして使用して自動翻訳を研究[4]したがコーパスの精度が圧倒的に不十分であった。現状は、顔の表情や手の位置、動作の方向、感情によって意味が異なるという特性[5]を考慮した日本語と日本手話の対訳コーパスが存在しない。そこで、その対訳コーパス開発への第1ステップとして日本手話の動作解析と、日本語と日本手話の対応調査が喫緊の課

1) <https://www.dinf.ne.jp/doc/JDF/20120323.miyagi/miyagiken.hisai.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>

題となっている。

1.1 研究目的

前節の通り、現状は日本語と日本手話の対訳コーパスは存在しない。そこで、その対訳コーパスを開発するための第1ステップとして、日本語に対する日本手話の構造を3次元データを含めて比較する行程が必要である。そのため、本論文では、日本語に対応した日本手話の3次元動作データの収集と、日本語と日本手話の対応を調査することが目的である。特に、日本語と日本手話の対応を調査する際には、日本手話を扱う上で重要な空間言語としての特徴に着目して調査する。

また、日本手話はノンバーバルな言語であり、感情により意味が大きく変化する言語である。さらに、手の位置や動作する方向によっても意味が変化する特性がある。そのため、最終的には手・腕・指の形だけではなく、顔の表情・動作などの情報も抽出や推定してコーパス開発をする必要がある。

2 提案手法

本論文では、顔の表情を含めた上半身の手話動作を対象にカメラでの撮影を行う。撮影の構図は図1に示す。撮影において日本手話で表現する日本語は「Japanese Daily Dialogue」コーパスを活用する。この際、日本手話への翻訳は、手話通訳士1名及び聴覚障がい者4名が協議してなるべく個人の癖などを排除した一般的な手話となるように翻訳する。特に、聴覚障がい者は、学習進度にバラつきがあり、聴者が一般的と考える日本語の単語も聴覚障がい者にとっては見慣れない単語がある。そのため、手話通訳士の助言や聴覚障がい者間での協議が必須となる。

日本手話に翻訳する日本語文のデータにおいては、日本手話が会話形式をメインに発展してきたという背景があるため、対話形式のコーパスを利用する。特に、日常生活で使うことが多い単語や文で構成されたコーパスを対象とする。

撮影されたデータに対して、3次元動作解析を行う。3次元動作解析には、「MediaPipe」⁵⁾というGoogle社が開発したオープンソースの機械学習フレームワークを活用する。MediaPipeには、FaceやPoseやHandsやHolistic等のさまざまな体の部位に特化した解析機能がある。本論文では、顔表情、腕、

手の動作を解析するために上半身全体の解析を行う「Holistic」モデルを用いる。Holisticモデルでの骨格解析結果のサンプルを図2に示す。

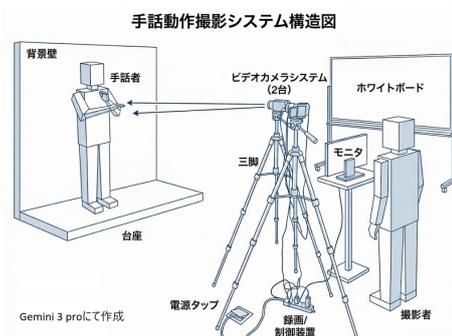


図1 日本手話の撮影方法



図2 Mediapipeでの動作解析例

3 実験

日本手話の動作解析実験を行う。聴覚障がい者の4名に678文の日本語を日本手話で表現してもらい。特に、感情を含める必要のある文では、顔の表情も意識して表現をしてもらう。カメラは被験者から2m程度の距離に2台設置する。被験者の4名は、20歳代～70歳代で、男性1名、女性3名の老若男女と多様である。また、身長も150cm台～180cm台と幅広いため、画角を人に合わせて調整する。2台のカメラで撮影を行い、1日の撮影時間は8時間とした。

表1 各部位のランドマーク数

部位	ランドマーク数
pose	33
face	468
left hand	21
right hand	21
total	543

MediaPipeで動作解析したデータは、x軸y軸z軸の座標点群として、csvデータを作成する。検出する各部位のランドマークは表1の通りである。

5) <https://github.com/google-ai-edge/mediapipe>

表2 日本語から日本手話への翻訳例

No	元の日本語文
1	おすしの話をしていたら、おすしが食べたくなくなってきましたね。
2	クラゲが泳いでいる姿は、フワフワしていて幻想的ですね。
3	今回の修学旅行は、行き先を3か所から選択するそうです。東京、大阪、沖縄から選べます。
4	パパ、見て。あそこにゾウさんがいます。
5	ライオンは、怖いです。このように近づいて、大丈夫なのですか？
No	翻訳した日本手話文
1	<お寿司> <話> <食べたい> <私>
2	<クラゲ> <泳ぐ>
3	<今度> <修学旅行> <場所> <3> <東京> <大阪> <沖縄> <選ぶ>
4	<パパ> <見て> <あそこ> <ゾウ> <いる>
5	<ライオン> <怖い> <私> <近づく> <大丈夫?>

4 実験結果

日本手話で表現する日本語の例を表2に示す。さらに、表2には、原文の日本語を日本手話に翻訳した文を示す。表2内の日本手話の単語[6]は、日本語の単語と表面上の字面は同じであるが、異なるものである。対訳コーパスとして、構築する際に、日本語の単語と日本手話の単語を同じトークンとして学習する危険性を防ぐために、<や>を用いて別の単語として扱い表現している。

8時間撮影した結果、撮影ペースは19.75[文/h]であった。但し、被験者が実験に慣れておらず、1文に対して3～4回の撮影失敗が発生したり、日本手話への翻訳表現が難しく協議時間が長くなる文もあった。この章では、撮影した動画からMediaPipeで動作解析した結果と、日本語文と日本手話文の構造の比較結果を示す。まず、MediaPipeによる手話動作解析における3次元データの例を図3に示す

図3では、「同意」と「午後」を表す手話動作の解析に成功していることが分かる。

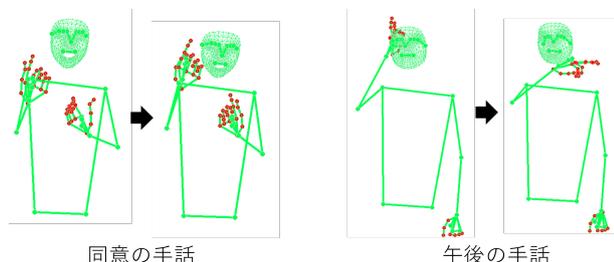


図3 手話動作解析例

撮影データから動作解析したところ、手話の動作抽出には成功した。但し、手話によっては、いくつか正常に動作を抽出することが出来なかった。その

例を図4に示す。図4は、「コピー」を表す手話である。手話の「コピー」では、右手が左手に重なるタイミングがある。そのタイミングの右手のランドマーク検出に失敗していることが分かる。図4で示したような、ランドマーク検出の失敗が他の手話でも確認出来た。



図4 ランドマーク検出の失敗例 コピーの手話



図5 感情を表現した顔の表情解析例

次に、感情が顔に現れているかを検証した。3つの例を図5に示す。結果を見ると、感情が顔に現れており、それを動作解析で抽出することが出来ている。しかし、被験者によっては、感情を顔に出すのが苦手な人もいる。その人の顔の動作分析では、感情を捉えるのが難しいと予想される。

また、顔の表情と手の動きを組み合わせる例を図6に示す。図6は、「コーヒーはいかかでしょうか?」という日本語文に対して、手の動きは「コーヒー」を、顔の動きは「どうですか?」という意味を表現している。

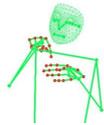


図6 顔の表情で文のニュアンスが変化する例

5 考察

実験結果から日本手話の構造を対訳コーパスとして組み込む際に注意すべき9点を挙げる。

1. 日本語（1万語）と日本手話（5000語）の日常会話での語彙数 [7] の違い
2. 顔の表情で文の強弱や意味が変わる場合がある
3. 1つの手話で複数の意味があるなど多義性
4. 1つの手話で名詞と動詞を表現可能
5. 固有名詞や動物の名前の多くは指文字で表現
6. 地域や個人で異なる手話がある
7. 「が」「の」「と」などの接続詞が無いので、同じ単語（同じ形の手話）を1つの文で2回使用することを避ける
8. 左右の手で意味の異なる手話を組み合わせて表現する場合がある（空間言語の特徴）
9. 文中に登場する人や場所や建物等の空間的な位置関係で手話の動きが変化する（空間言語の特徴）

この中でも、特に項目8は翻訳AIモデルを開発する上で学習が難しい文となることが予想される。図7に例を示す。図7は表2におけるNo3の文を日本手話で表現した時の解析データの一部である。図7では、左手で選択肢が3つあることを常に示して、右手で各選択肢の説明を表現している。このように、空間言語ならではの表現方法を考慮したコーパス作成が必要となる。

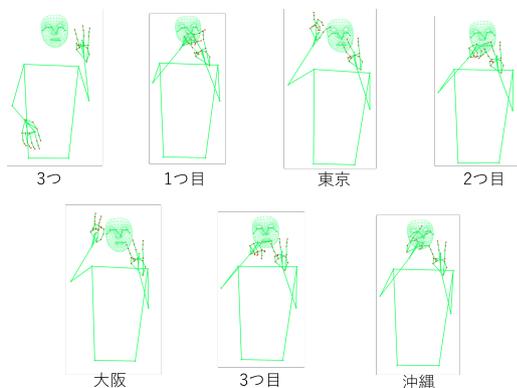


図7 左手と右手で異なる手話の例

6 おわりに

本論文では、日本手話の動作をカメラで撮影して、動作解析した。その撮影実験では、678文の日本語を日本手話で表現した動作を撮影した。実験結果は、ほとんどのデータにおいて顔、腕、手の動作解析に成功していた。但し、ランドマーク検出に失敗しているケースもあった。そして、日本語日本手話の対訳コーパスを開発するにはもっと多くの動作解析データが必要となる。また、解析データは、日本手話の1語1語が連続したデータである。そのため、日本手話の動作に対しての形態素解析が必要となる。そして、顔の表情と日本語文の意味から感情分析も必要である。それらの解析データを組み合わせ、かつ、注意事項の9点を考慮して、日本語日本手話の対訳コーパスを開発していく予定である。

謝辞

本研究は競輪の補助を受けて実施しました。

本研究の実施および聴覚障害者の現状調査にあたり、ご助言を賜りましたNPO法人手話ランゲージの大橋正敏氏、宮城県聴覚障害者情報センターみみサボみやぎの松本隆一氏、庄子陽子氏に深く感謝申し上げます。

参考文献

- [1] 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機「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.
- [2] 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. アクセシビリティ研究会研究報告 (AAC), Vol. 2019-AAC-9, No. 5, pp. 1–8, 2019.
- [3] 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.
- [4] 米川明彦. 新・日本語一手話辞書. 中央法規出版, 2011. 編集: 社会福祉法人全国手話研修センター.
- [5] 丸田要. 聴覚障がい者の情報格差解消のための多様な文書要約による情報伝達. IS-167, Vol. 2024-IS-167, No. 12, pp. 1–5, 2024.
- [6] 長南浩人. 日本手話, 中間型手話, 日本語対应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. 教育心理学研究, Vol. 49, No. 4, pp. 417–426, 2001.
- [7] 松下達彦. 日本語を読むために必要な語彙とは? : 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. 日本語教育学会春季大会, 2010.