

格構造に注目した日本語-日本語手話の並び替えと 述語項構造に注目した語義曖昧性解消

三澤 賢祐 酒井 啓道 吉川 友也 水本 智也 松本 裕治
Kensuke Mitsuzawa Hiromichi Sakai Yuya Yoshikawa Tomoya Mizumoto Yuji Matsumoto

奈良先端科学技術大学院大学
Nara Institute of Science and Technology

Some deaf people have difficulties in reading Japanese Texts. Therefore, it is preferable to convert written text into animation of sign language automatically for better communication. To generate sign language animation, parameter information which defines movement of animation is needed. Generally, this parameter is extracted by machine translation method. But there is two problems for translation. The first problem is the difference of word order. The second problem is the difference of the meaning of a word. In this paper, for the first problem, we propose a translation method using the predicate-argument analysis. For the second problem, we translate each Japanese word to relevant sign language expression using an existing paper dictionary.

1. はじめに

手話は主にろう者により利用される言語であり、市田らにの推定によれば平成 11 年時点ではおよそ 5 万 7 千人の話者人口がいる [1].

ろう者の中にはしばしば日本語読解に困難を持つケースが報告されている [2]. そのため、テレビの字幕のように単純にアナウンサーの読み上げ原稿を字幕表示しても日本語読解に難を持つろう者は情報を正しく理解できないという問題が発生する. このように情報が正しく伝わらないと、日常的に情報の公平性を失うばかりか、災害時において最悪の場合は命の危険につながることも考えられる. したがって、この読解の障害は解決すべき問題である.

入力された日本語に対し、それに対応する手話のアニメーションを表示できれば、先の問題は解決できる. 一般的に、アニメーションを生成するためには、まず各単語の語義情報と構文情報を得て、手話文への変換を行い、最後に手話文からアニメーションの動作を定義するパラメータを得る.

そこで本研究では、このパラメータを得る前段階として、述語項構造解析を用いた手話構文への並び替えと既存の辞書知識を用いて日本語単語を適切な手話表現に置き換える手法を提案する. 具体的には、述語項構造解析から得られた述語とそれに対応する格の情報を用い、翻訳ルールにより手話文への翻訳を行う. また、日本語単語と手話単語では、同じ表層形の日本語単語が手話では異なる表現を行うことがある. この訳し分けを行うために、辞書の例文を用いる.

以下、2 節で既存手法の紹介と提案手法との違いについて述べ、3 節で手話の統語的な特徴を記述し、4 節で述語項構造解析を用いた並び替えの手法を紹介する. 最後に 5 節で、辞書知識を利用した日本語単語から手話単語への置き換え手法について述べる.

2. 既存手法

松本ら [3] による jaw/SL は、ルールベースの日本語-日本語手話の機械翻訳エンジンである. これはパターン変換方式を採用

した多言語翻訳エンジンの jaw[4] を手話にも拡張したシステムであり、ルールの追加記述が容易であるという利点がある. しかし、日本語の表層形に対するルールを定義しているため、ルールの数が膨大になってしまうという問題点がある. 本研究では表層形でなく、述語と格の関係にルールを定義し、この問題に対処した.

一方、近年の機械翻訳の主流である統計的機械翻訳を利用した翻訳も提案されている. 世界の手話の中には、アメリカ手話のように大規模な音声言語と手話の対訳コーパスが利用できるケースが報告されている [5] が、現状では日本語にはそのような大規模な対訳コーパスは存在していない. NHK 放送技術研究所では NHK 手話ニュースの放送記録を元にした 3 万文程度と比較的小さいサイズの日本語-日本語手話コーパスを構築し、用例翻訳と統計的機械翻訳を組み合わせた手法を提案している [2]. 対訳データを元にした手法であるため、手話に関する文法知識を必要としないが、その一方で対訳データが気象情報に特化したドメインであり、対象が気象情報に限られてしまう. 本研究では、ルールベース翻訳により、またドメインにも依存しない日本語-日本語手話の翻訳手法を目指した.

3. 手話の統語的特徴について

日本語手話の文法的な特徴について、概略を統語面から述べる. なお、手話の文法ははまだ研究途中にあり、研究者により諸説ある. 本研究では [6] および [7] の文法体系に従った.

以下、手話文を文字表記する際には、意味が日本語単語に相当する単語を表記する. これは「日本語ラベル」と呼ばれる. このようなラベルを使用するのは、日本語の単語と日本語手話の単語が持つ意味が異なることを明示的に示すためである. 本論文では、図 1 の日本語手話文のように、日本語ラベルは単語を山カッコ { < > } で囲って表記する.

3.1 Non Manual Signals の利用

日本語手話では、日本語の助詞、接続詞といった文法情報を持つ語を、通常は手で表現しない. その代わりに、眉の上げ下げ、目の開き加減、うなずき、首の位置といった顔や全身を使った表現を行う. このように全身を使って表現する文法情報を NMS(Non Manual Signals) と呼ぶ. これは通常、単語を手で表現しながら同時に行われる. NMS は図 1 の N や jaw_up よ

日本語: 私は弟に弁当を作らせた。
日本語: <私> <弟> pt3 <弁当> <作る> jaw_up <わかる> N <弁当> <作る> <終わる> pt3 PAST
pt3:3人称の指差し N:うなずき jaw_up:顎あげ PAST:過去形の表現

図 1: 日本語と対訳の日本語手話の例

NMS	動作	主な働き
N	うなずき	句の切れ目
jaw_up	あごあげ	命令, 依頼
pt3	指差し	動作主体の明示 (3人称)

表 1: 代表的な NMS の例

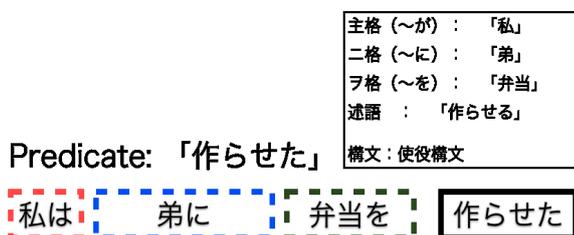


図 2: 述語項構造解析による結果

うに記号で記述される。本論文中で利用する NMS は表 1 に記載した。

3.2 手話の語順

日本語手話の語順は平叙文では、日本語と同じく SOV 型であるとする説が一般的である。しかし、「もし」のような条件節、「〜なので」、「〜だから」のような原因と結果を表す文、使役文、受け身文、疑問文等は NMS を併用し大きく語順変える。一例として、図 1 に使役文における日本語と対応する日本語手話の日本語表記を示す。

日本語では、一般的には修飾語と被修飾語の位置関係は修飾語が先行する。しかし、日本語手話においては一般的には被修飾語が先行する。

4. 述語項構造を用いた並び替え規則

前述のように、日本語手話では平叙文を除いて、翻訳元である日本語文から語順が大きく変化する。しかし、これらには定型的な構文が存在し、述語と述語に対応する格に注目するとルールによる翻訳が可能である。一例として、使役構文における述語と格の位置を図 2 に示す。

このように特殊な語順の構文は、述語と格の位置が予め決定しているため、日本語文の述語と格を特定できれば、文法レベルの翻訳が可能である。したがって、述語項構造を用いた翻訳を行う。図 1 のように対象文の述語と、述語に対応する格を特定し、図 3 のように格が挿入される場所と必要な NMS をルールによって配置し、翻訳を完了させる。

次節からはそれぞれの構文での並び替え規則について述べる。

4.1 使役構文

手話の使役文は以下の構文により構成される。

主格 二格 pt3 ヲ格 述語 jaw_up

わかる N ヲ格 述語 終わる pt3

図 3: 手話の使役構文に置ける格の配置場所

主格 二格 pt() ヲ格 述語 jaw_up <わかる> N ヲ格 述語 <終わる> pt()

jaw_up はこの構文において、命令を意味する NMS である。N はここで句が一度終了していることを意味する。pt() には 1, 2, 3 のいずれかが挿入される。これは二格の名詞の人称に対応する。したがって、手話の使役構文を直訳すると、「主格は二格に〜しると言い、二格は『わかった』と言い、〜した」という構成になっている。

この並び替えを行うために、まず二格の人称の判定を行い、対応する pt() を決定する。さらに上記の順番に格を配置し、NMS を挿入し、並び替えを完了する。

4.2 受け身構文

手話の受け身は以下の文法により構成される。

< 二格 主格 述語 pt() >

pt() は前述のように二格に対応する人称が挿入される。したがって、たとえば「彼は私に好かれている」という文を手話で表すと、「<彼> <私> <好き> pt1」となる。

この並び替えを行うために、まず二格の人称の判定を行い、対応する pt() を決定する。さらに上記の順番に格を配置し、NMS を挿入し、並び替えを完了する。

4.3 条件節構文

手話で条件節を含む文を表現する際には、眉の動きを利用する。まず条件節の開始部分で眉上げを行い、条件節を表現している間は眉を上げたままにする。条件節が終了すると条件節の最後の単語からほんの少しだけ遅れて眉上げを元に戻す。この眉上げを戻すことにより、条件節の終了を意味する。条件節の最後の単語から眉上げを戻すまでに時間には個人差があるとされている [8]。手話話者によっては、条件節の最初に「もし」という手話単語を表現するが、その場合でも眉上げと節の最後の眉戻しは必須である。例えば、条件節を含む文である「もし雨が降るなら、渋滞する。」は手話では「眉上げ <雨> 間 眉上げ解除 <渋滞>」となる。

条件を表す表層単語「もし」が日本語文中にあった場合に、上記のような並び替えを行う。

4.4 原因・結果節構文

手話で原因と結果を表す際には、大きなうなずきを節の間に用いる。原因を表す節の最後の単語から間を置き、そして大きくうなずく。この大きなうなずきをもって、原因を表す節の終了を意味する。例えば、日本語文「雨が降るから渋滞する。」は手話では「<雨> 間 N <渋滞>」と表現される。

構文解析を行い、日本語文中に原因・結果表現が検出された場合に、上記のような並び替えを行う。

4.5 並列語句

手話で並列を表現する際には、並列する語と語の間にうなずきを挟み表現する。うなずきは NMS で N を利用し表現される。例えば日本語並列句の「赤と紫」は手話では「<赤> N <紫>」で表現される。仮に「<赤> <紫>」と手話で表現すると、日本語では「赤紫」を意味する。

主辞の品詞	修飾語の品詞
名詞	形容詞
形容詞	副詞
動詞	副詞

表 2: 並び替えを行う主辞と修飾語

日本語表現	手話表現
旗があがる	胸の前あたりで 右手を横にして固定しておき、 左手をへその辺りから 体と右手の間の空間を通り抜けるように 上げる
花火があがる	へその辺りで 閉じている手を、 ひらきながら 顔の横あたりまであげる

表 3: 日本語と手話表現の違いの例

見出し語	あがる
例文	旗があがる
述語	上がる
格	ガ格:旗
意味カテゴリ	人工物-その他
意味ドメイン	null

表 4: 述語項を利用して語の翻訳を行う場合

並列語句の解析を行い、語または句が並列関係にある場合にはその間に N を挿入する。

4.6 修飾語の並び替え

手話文において、一般的には修飾語と被修飾語の関係は日本語文と逆である。つまり、被修飾語が先行し、その後修飾語が続く。

そこで、日本語文に係り受け解析を行い、係り受け関係にある修飾語と被修飾語を逆転させることで、手話文における修飾語句の並び替えを行う。主辞と修飾語の品詞が表 2 の場合に、この逆転操作を行う。

5. 単語単位での置き換え

5.1 辞書データベースの作成

単語の置き換えの際に問題になるのは、意味の細かな違いである。手話は情景を描写して表現するため、単語の持つ意味の違いをしっかりと反映する。したがって、日本語単語では同じ表層形でも、手話ではまったく違う手の動きをすることがしばしばある。例えば、表 3 のように、2 つの文は「あがる」が共通した述語であるが、「あがる」対象が「旗」と「花火」という違いがある。手話表現では、この「あがる」の違いを捉えた表現を行う。単語の置き換えの際には日本語の文字情報からこの違いを捉えて、違う表現であることを明示的にする必要がある。

我々は手話辞典の知識を利用して前述の違いを訳し分けるために、「新日本語-日本手話辞典」に則り、辞書データベースを作成した。この辞書データベースの例文は表 3 の例文のよ

見出し語	青い
例文	青い空
周辺語	空
意味カテゴリ	場所-自然
意味ドメイン	null

表 5: 周辺語を利用して語の翻訳を行う場合

うに、見出し語の意味を捉えた例文が掲載されている。意味カテゴリと意味ドメインは形態素解析ツール juman Ver.7.0^{*1} が利用する JUMAN 辞書に定義されていて、12 種類の意味ドメイン、22 種類の意味カテゴリが存在している。

見出し語が動詞の場合には、この動詞述語が取る格に対して意味カテゴリと意味ドメインを判定する。例を表 4 に示す。見出し語は「あがる」という動詞であり、この時、例文中の述語は「上がる」である。ここで述語項構造解析を行い「上がる」に対応する格はガ格で「旗」であるとわかっている。そこで、「旗」に対して意味カテゴリと意味ドメインを判定を行い、データベースに記録した。今回、意味ドメインは判定されなかったため、null となっている。

見出し語が名詞および形容詞の場合には、見出し語の周辺語に対して意味カテゴリと意味ドメインの判定を行った。例を表 5 に示す。見出し語は「青い」で、例文中には「青い」以外の単語として「空」が利用されている。そこで、「空」に対して意味カテゴリと意味ドメインの判定を行い、データベースに記録した。今回、意味ドメインは判定されなかったため、null となっている。

なお、例文の形態素解析と単語の意味カテゴリとドメインの判定には juman Ver.7.0、構文解析および述語項構造解析は KNP Ver.4.0.1 を利用した。

5.2 辞書データベースを利用した単語の置き換え

前述の辞書データベースを利用し、日本語文の単語を適切な手話単語に翻訳し置き換える。翻訳対象の日本語文にもデータベース中の例文と同じように形態素解析、構文解析、述語項構造解析、単語の意味ドメインと意味カテゴリの判定を行う。そして、解析された形態素を辞書データベース中の見出し語から検索する。この際、もし辞書データベース中に重複する見出し語がなければ、一意にその単語に置き換えを行う。しかし、表 3 のように見出し語「あがる」が重複する場合にはデータベースに記録された意味カテゴリと意味ドメインの情報を元に置き換えを行う。今回は意味カテゴリもしくは意味ドメインのどちらかが完全一致した場合に、その単語に置き換えを行った。もし、一致する意味カテゴリ、意味ドメインがない場合には置き換え不可とした。

6. おわりに

述語項構造を用いた語の並び替え手法の提案と辞書知識を利用した語の置き換え手法について述べた。この手法はルールベースながら、日本語の述語項構造解析が行えれば、手話の文法体系に従い並び替えが可能である。

現在は辞書データベースを整備中であり、小規模な翻訳しか行えない状態である。今後はこの辞書データベースの整備を進めていくと共に評価方法についても検討を行う。手話は未だ文字表記の統一規則が存在せず、各書籍が別々の方法で表記

*1 <http://nlp.ist.i.kyoto-u.ac.jp/>

を行っているため、評価手法そのものが確立していない。この点、評価手法も含めて考えていく必要がある。

7. 参考文献

参考文献

- [1] 市田泰弘, 難波友加, 伏原桃子, 三宅三枝子, 吉井美樹: 日本手話母語話者人口推計の試み, 日本手話学会第 27 回大会予稿集, pp.42-45 (2001).
- [2] 加藤直人, 宮崎太郎, 金子浩之, 井上誠喜, 清水俊宏, 長嶋祐二: 気象情報の日本語一手話 CG 翻訳, 言語処理学会第 18 回年次大会発表論文集, pp.275-278 (2012).
- [3] 松本忠博, 田中伸明, 吉田鑑地, 谷口真代, 池田尚志: 手話の表記法とテキストレベルの日本語-手話機械翻訳システムの試みについて, 電子情報通信学会 思考と言語研究会, Vol.104, Issue 316, pp.7-11 (2004).
- [4] 今井啓允, 池田尚志: オブジェクト指向言語のパラダイムを利用した機械翻訳エンジン jaw, 言語処理学会第 10 回年次大会発表論文集, pp.125-128 (2004).
- [5] Othman, A., Jemni, M., Hussein, A. T., Menara, B.: Statistical Sign Language Machine Translation : from English written text to American Sign Language Gloss, Prec. International Journal of Computer Science Issues, (IJCSI), Vol.8, Issue 5, pp.65-73 (2011).
- [6] 岡典栄, 赤堀仁美: <文法が基礎からわかる>日本手話のしくみ, 大修館書店 (2011).
- [7] 木村晴美, 市田泰弘: はじめての手話-初歩からやさしく学べる手話の本, 日本文芸社 (1995).
- [8] 市田泰弘: 手話の言語学 (第 9 回) 頭の動き・位置と顔の表情-日本手話の文法 (5) 文タイプと従属節, 月刊言語, 9 月号, pp.94-101 (2005).
- [9] 米川明彦, 全国手話研修センター-日本手話研究所: 新 日本語-日本手話辞典, 中法法規出版 (2011).