

日本語 - 手話機械翻訳システム jaw/SL による翻訳実験

吉村康寛 松本忠博 池田尚志

岐阜大学工学部

1 はじめに

手話はろう者にとっての第一言語である．これまで行われてきた手話に関する工学的な研究は，手話動画の合成や手話動作の認識といった画像処理の観点からの研究が中心となっており，言語処理の観点からの研究はまだ少ない．これは，手話には一般的に認知された文字による表現がないことが要因として挙げられる．我々の研究室ではこれまでに，日本手話を書き記すために，日本語援用手話表記法 (jjs-notation) を提案し [1]，jjs-notation による手話ビデオ教材 [6] 映像の書き取り実験 [4]，jjs 表記を出力とする日本語 - 手話機械翻訳システム jaw/SL の構築 [3] を行ってきた．本論文では，jjs-notation での書き取り実験で得られたデータを出力目標とする，513 例文に対する jaw/SL による翻訳実験と分析・評価について述べる．

2 jjs-notation について

手話単語は手の形，位置，動きなど手指要素と，顔の表情や視線，頭や上体の動きなどの非手指要素で構成される．これらの要素をパラメータのように変化させることで様々な意味を表現している．jjs-notation では，日本語を援用した手話単語ラベル (日本語 - 手話辞典 [5] による) を基本形とし，語形変化をパラメータとして下記のように記述する．表 1 に例を示す．

手話単語ラベル [手型] (空間要素; 修飾要素)
語形変化パラメータ

日本語	jjs-notation での表記
[1] 行く	行く。 (基本形)
[2] 2 人で行く	行く [2]。 (手形: 主体の数を表す)
[3] あなたが私を見る	見る (2 1)。 (方向変化: 主体と対象を表す)
[4] 妻の仕事はパートです	{<t> 妻 仕事 } , 例 - 仕事。 (<t> は話題化を表す非手指文法標識)

表 1: jjs-notation による手話表記の例

3 日本語 - 手話機械翻訳システム jaw/SL の概要

3.1 機械翻訳エンジン jaw

jaw は，日本語から多言語への翻訳を行うパターン変換型の翻訳エンジンであり，目的言語毎に翻訳辞書と生成関数の DLL を入れ替えることで，任意の言語への翻訳が可能なシステムとなっている．翻訳システムは，日本語の表現構造を目的言語の表現構造へ変換するための翻訳規則 (パターン翻訳規則・機能語翻訳規則)，目的言語の表現構造から目的言語のテキストを生成する線状化関数等を目的言語毎に記述することで構成される．パターン翻訳規則とは日本語の表現パターンと，それにマッチした目的言語の表現構造の対であり，機能語翻訳規則は，パターン翻訳規則で作られた目的言語の表現構造に，モダリティ等の情報を設定する規則である．目的言語の表現構造は，VC++ のオブジェクトとして実装されている．

jaw を用いての翻訳は，入力日本語文の解析を行い，翻訳規則を用いて目的言語の表現構造へと変換し，目的言語の語順への並べ替え，語形変化を加えて目的言語テキストの出力を行うという流れで行われる (図 1)．

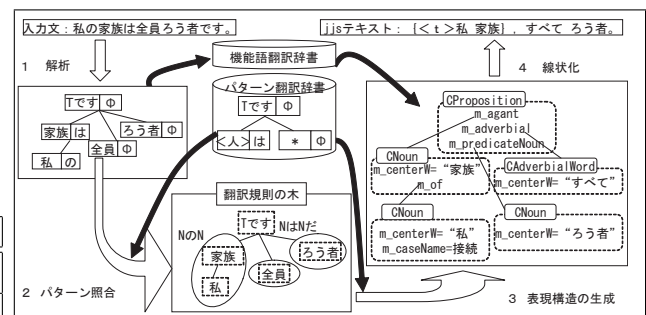


図 1: jaw/SL の処理の流れ

3.2 日本語 - 手話機械翻訳システム jaw/SL

jaw/SL は日本語文から jjs-notation テキストへの翻訳を行うシステムである．

日本語パターン	翻訳規則
見る <N1: 体言>が <N2: 体言>を	述語オブジェクト {訳語←“見る”, 主体←N1, 対象←N2, 一致動詞フラグ←TRUE}
犬	名詞オブジェクト {訳語←“犬”}
病院	名詞オブジェクト {訳語←“ビル-脈”}
美しい <N: 体言>が	述語オブジェクト {訳語←“美しい”, 対象←N}

図 2: jaw/SL における翻訳規則の例

3.2.1 パターン翻訳規則

jaw/SL では、例えば「N1 が N2 を見る」という日本語パターンに対しては、“N1 が動作主、N2 が対象”という情報を翻訳規則に記述する。また「見る」は動作主や対象が変化することによって動作の方向が変化する動詞であるため（表 1[3]），“方向変化する動詞である”という情報も併せて記述する（図 2）。

3.2.2 機能語翻訳規則

日本語で助詞や助動詞によって表現される意味内容は、手話では助動詞的な単語や、単語の語形変化（手の位置や運動の変化）、非手指要素 (NMS) などによって表現される（表 2）。

表 2: jaw/SL における機能語翻訳規則の例

日本語：構造	jjs-notation
飲まない：飲む + ない	飲む ない
もらえる：もらえる	もらう できる

3.2.3 線状化処理

線状化処理では、話題化などの NMS のテキスト化や動詞の方向変化を行いながら、手話の語順に従って、手話の表現構造から jjs 手話テキストを生成する。例えば、数と単位の語順について、手話では対象が変わることで語順が変わる（図 3）。

- (1) 単位を数に後置する-価格「400円=400-円」など
- (2) 単位を数に前置する-年齢「20歳=年齢 20」など
- (3) 数を手形、単位を動きで表す-年「20年=n年[20]」など
- (4) 片手で数、もう一方の手で単位を表す-人数「4人=4/人」など

図 3: 数と単位の語順について

4 翻訳実験

4.1 実験題材について

翻訳実験の題材として手話ビデオ教材 [6] を使用した。この教材には、ろう者がカメラに向かって日本手話を表

現している映像が記録されている。全文約 720 文に対しての自然な日本語訳と、手話の構造に即した日本語訳（以下、構造訳）を記した冊子が付属している（図 4）。今回の翻訳実験では、基本的にはこの構造訳を入力文とし、それに対応する手話文（手話映像を jjs-notation で書き取ったもの）を正訳として用いた。

日本語訳：私と父、母はろう者、
祖父と祖母は聴音です。
構造訳：私と、父、母の 3 人はろう者、
祖父と祖母の二人は聴者です。

図 4: 日本語訳と構造訳

4.2 実験内容

jjs-notation で書き取った手話ジャーナル約 720 文から、次に示す 4 項目の条件で 513 文を抽出し、翻訳対象文とした。

- 1 文を単独で見たときに、人間でも正しく翻訳が出来ない文は除外する。
- 現状で jjs-notation による表記が困難な表現を含む文は除外する。
- 文中に言い直し・補足挿入などを含むような、文として整っていないものは除外する。
- 複数の文が並列に長く連なった文については、手話文に合わせて分割可能ならば分割して実験対象とし、出来なければ除外する。

513 文中には、以下のような文があり、これらを翻訳するためのパターン翻訳規則、機能語翻訳規則、線状化関数をシステムに与えた。

- 動詞の方向変化・名詞の位置変化を含む文
例) そのあと、東京に引っ越ししました。
将来、東京 (x) 引っ越す (x)。
- 話題化、疑問文、分裂文、条件節の非手指要素の標識 (t, ?, cleft, cond) を含む文
例) いたら援助して下さい。
{<cond> いる}, 助ける (2 1) 頼む。
- 助動詞を含む文
例) けれども、最近、私はスポーツに目覚めつつあります。
しかし、最近、私 スポーツ 目覚める ~ 中。
- 手話単語の (半) 同時的な組み合わせによる表現を含む文
例) 前の家族は 4 人です。
{<t> 過去 家族}, 4 / 人。

- 手型の変化による語義の変化を含む文
例) どこかという、世田谷区役所で、2年間勤めました。

{<cleft> 何} 世田谷-区役所, 仕事 年間 [2]

- 語形変化による副詞の表現を含む文
例) でも、少し違いがあります。
しかし、違う (; ちょっと) ある

- 複文
例) 学校が同じで、結婚して引っ越して、横浜に住んでいます。
学校 同じ; 結婚 引っ越す (両手で場所を移動), 横浜 いる。

- 指文字によって単語を表現する文
例) ペットは、だいが前に飼っていました。
{<t> 'ペ・ツ・ト'}, 過去 (; とても) ある。

具体例として、非手指文法標識を含む文の中から「分裂文」「条件節」「話題化・焦点化」の処理について以下に述べる。

4.3 分裂文を含む文の処理

分裂文は、文の焦点を明示・強調する構文である。手話における分裂文は、疑問詞を文中に持っているの、分裂文を翻訳するために疑問詞 (例文中では“どこ”) を含めた日本語パターンを作成し、翻訳規則を記述することで分裂文の翻訳を行う。

入力文: 母の生まれは どこかという、東京です。

- N1< 人名/人間 > が N2< 普通名詞 > だ (N1 は省略可) N1 N2
- N1< 体言 > の N2< 体言 > N1 N2
- N< 普通名詞 > は どこかという P< 事象 > {<cleft> N 何}, P

翻訳結果: {<cleft> 母 生まれる 何} 東京。

4.4 条件節を含む文の処理

条件節を表す非手指文法標識 (cond) を含む文は、「~の場合は~、~がなければ~」という特徴を持っており、条件節を翻訳するための特徴として用いることが出来る。

入力文: 仕事が終わった後、用事がない場合は、早く帰ります。

- N1< 人名 / 人間 > が N2< 地名 / 場所 > に帰る (N1・N2 は省略可) N1 N2 帰る
- P1< 事象 > 場合 P2< 事象 > {<cond> P1 とき}, P2
- N< 普通名詞 > は 用事がない (N は省略可) N 必要 ない
- P1< 事象 > 後 P2< 事象 > P1 以降 P2
- 早い P< 事象 > はやく P
- N< 人間活動 > が終わる N 終わる

翻訳結果: {<cond> 仕事 終わる 以降 必要 ない とき}, はやく 帰る。

4.5 話題化・焦点化の処理

手話において話題化は、眉上げや顎引きという非手指要素により表現される [7]。日本手話では、主格 (動作主) に付随する副助詞「は」に伴って話題化されることが多いが、翻訳対象文では助詞が「は」の場合でも動作主が1人称の場合には、話題化されないことが多いため、人称によって訳し分けを行う。

5 実験結果

5.1 望ましい翻訳結果を得た文例

望ましい翻訳結果を得た文の例を以下に示す。

例1 - 入力文: 私の家族は全部で4人です。
正訳 & jaw/SL: {<t> 私 家族}, すべて 4/ 人。

例2 - 入力文: 生まれは どこかという、東京あたりだと思っています。
正訳 & jaw/SL: {<cleft> 生まれる 何}, 東京 あたり 思う。

例 3 - 入力文：私は夜、家に帰る時間は、まちまちです。

正訳：{<t> 私 夜 } {<t> 家 / 帰ってくる 時間 } , まちまち。

jaw/SL：{<t> 私 夜 , 家 / 帰る 時間 } , まちまち。

5.2 望ましい翻訳結果を得られなかった文例

望ましい翻訳結果を得られなかった文の例を以下に示す。

例 4 - 入力文：6月20日と12月10日です。

正訳：6-月-20(L), 12-月-10(R)。

jaw/SL：6-月-20, 12-月-10。

例 5 - 入力文：私が今まで飼ったペットは、カニ、カエル、鳥、亀、犬といろいろ飼いました。

正訳：{<t> 私 今 まで 育てる 何 } , カニ Enum[1] , カエル Enum[2] , 鳥 Enum[3] , 亀 Enum[4] , 犬 いろいろ 育てる 終わる。

jaw/SL：{ <t> 私 今 まで 育てる 何 } , カニ , カエル , 鳥 , 亀 , 犬 いろいろ 育てる 終わる。

例 6 - 入力文：ジョギングは職場の周りを1周走り終えたら、そのあと食事です。

正訳：{<t> 走る } 職場 1 周 走る 1 周 終わる , 将来 食べる。

jaw/SL：将来 , {<t> 走る } , 仕事 - 場所 1 周 走る 1 周 終わる 食べる。

5.3 分析と評価

表 3: jaw/SL による翻訳結果

翻訳結果	例文の数
A：正訳通りの結果を得る文	117 文
B：正訳とは異なるが、正しく意味を伝えている文	111 文
C：正訳とは異なり、正しく意味を伝えていない文	103 文
D：翻訳結果を得られない文	129 文
Z：未着手（今後実験を行う）の文	53 文
合計（= A + B + C + D + Z）	513 文

jaw/SL で得られた出力結果が、入力文の意味として通じるかどうかという観点で人手で評価を行った。正解と判断できる文（A + B）は 228 文であった（表 3）。条件節を含む文や、分裂文を含む文、2 文のみで構成される複文などに関しては、望ましい翻訳結果を得ることが出来た。望ましい翻訳結果を得ることが出来なかった

文としては、例 4 のように表現位置を変化させることで対比を行う文、例 5 のように数え上げを含む文のほか、手話には“時系列順に手話の表現を行う”という特徴があるために語順が不適切な文や、例 6 のような線状化処理の見直しが必要な文が見られる。望ましい翻訳結果を得ることが出来ていない文（C + D）は 232 文であった。

翻訳実験では、構造訳を入力文としているが、実際には自然な日本語訳を入力として翻訳結果を得ることが望ましい。しかし、現状の jaw/SL では、この自然な日本語訳から構造訳ないし“翻訳可能な文”に言い換える処理が出来ていない。これは意味理解に関わる問題であり、今後の大きな課題である。

自然な日本語訳：将来犬や猫を飼ってみたいとはあんまり思いません。

構造訳：将来犬や猫を飼ってみたいと思うかどうかということでは、あまりそんなふうには思っていない。

翻訳可能な文：私は将来犬や猫を飼うかという、そう思っていない。

6 おわりに

日本語 - 手話機械翻訳システム jaw/SL による翻訳実験を行った。今後の課題として、話題化、疑問文、分裂文以外の非手指文法標識を含む文の翻訳実験、辞書の拡充、自然な日本語文から構造訳への言い換え処理などが挙げられる。

参考文献

- [1] 松本忠博, 原田大樹, 原大介, 池田尚志「日本語を援用した日本手話表記法の試み」, 自然言語処理 Vol.13 No.3 pp.177-200, 2006
- [2] 松本忠博, 池田尚志「日本語テキストから手話テキストへの機械翻訳の試み」, 自然言語処理 Vol.15 No.1 pp.23-51, 2008
- [3] 谷口真代「日本語 - 手話機械翻訳システム jaw/SL に関する研究」岐阜大学工学部応用情報学科 知識情報講座 池田研究室 学士論文, 2005
- [4] 原田大樹「日本語テキストから手話テキストへの機械翻訳」岐阜大学工学部応用情報学科 知識情報講座 池田研究室 学士論文, 2006
- [5] 財団法人全日本聾唖連盟日本手話研究所 (編), 米川明彦 (監修). 日本語 - 手話辞典. 財団法人全日本聾唖連盟出版局. 1997.
- [6] サインファクトリー, 手話ジャーナル初級教材 No.1,2
- [7] 木村晴美・市田泰弘, はじめての手話, 日本文芸社, 1995