

SignWritingのための語彙辞書の構築

松本忠博[†] 原田大樹[†] 吉村康寛[†] 加藤三保子^{††} 池田尚志[†]
[†]岐阜大学工学部 ^{††}豊橋技術科学大学

1. はじめに

SignWriting[5] は日常生活での手話の読み書きを想定して Sutton により考案された手話の表記法の一つである。海外では SignWriting による手話教育や新聞の発行なども試みられている。本稿では、日本語から SignWriting で記述された手話への機械翻訳システムを実現する上で必要となる SignWriting 用手話語彙辞書について述べる。

日本語から SignWriting への変換は2段階に分けて行う。我々は言語処理に適した手話表記法を提案した [4]。第1の段階はこの表記法で書かれた手話テキスト（以下、単に「手話テキスト」と記す）を目的言語とする日本語から手話言語への機械翻訳であり [3]，第2の段階は手話テキストから SignWriting への同一言語内での表現の変換である。我々の表記法が日本語を援用し、手話の動作そのものよりもそれによって表される語彙や文法情報の記述に重点を置いているのに対し、SignWriting では手の形や向き、動き、顔の表情などを表す図像的な記号を2次元的に配置して手話単語を視覚的に分かりやすく表記する。

本稿で述べる SignWriting 用辞書は、手話テキストとして記述された手話の語彙や文法情報を、SignWriting の視覚的な文字表現に変換するため

に構築しているものである。なお、SignWriting の計算機内部での表現には SWML (SignWriting Markup Language) [1] を用いている。

2. 手話の特徴

国や地域によって使われる手話は異なるが、空間の利用の仕方など共通した特徴も見られる。手話単語は手の形・位置・動き・顔の表情など複数の要素で構成され、各要素をパラメータのように変化させることによってさまざまな意味が表現される。例えば日本手話の場合、「1日」を表す手話単語の（動きはそのまま）手形だけを2に変化させると「2日間」になる。

話者の周りの空間も文法的に利用される。次の例は「友達からもらう」という意味の日本手話に対する手話テキスト表現である。

友達 (3) もらう (3 1)。

これは動作的には、名詞 友達 が3人称の位置（話し手の斜め前、または、話し手と聞き手の間）で表現され、次に、動詞 もらう が 友達 の位置から1人称（話し手）の位置に向かって表現されることを表している。動詞 もらう はその方向によって「誰が誰からもらう」という格関係を表すため、この手話文は「友達 から もらう」

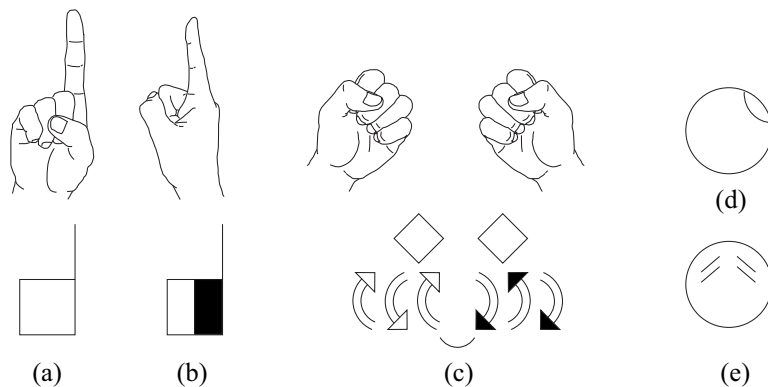


図1 SignWriting で使われる記号の例。(a)(b)手の形と指先・掌の向き,(c)手とその動き,(d)頭部の位置,(e)眉の状態。

という意味になる。

顔の表情や頭の動きなどの非手指要素も言語的に重要な役割を持っている。次の例は「妻の仕事はパートです」に対する手話テキストである。

{ <t> 妻 仕事 } 例 - 仕事。

ここで <t> は、眉上げなどの表情によって表される「話題化」の文法標識を示し、中括弧はその有効範囲を示している（「妻の仕事」の部分が話題化されている）。なお、例 - 仕事 は複合語である。

3. SignWriting と SWML

3.1 SignWriting

SignWriting では図 1 のような記号を 2 次元的に配置して手話単語を表す。この記号は IMWA (International Movement Writing Alphabet) と呼ばれ、SSS (SignSymbolSequence) によりその分類と順序が規定されている。例えば、図 1 (a) の記号は 01-01-001-01-01-01 というシーケンスで表される。SSS-2004 には 8 つのカテゴリ（手、動作、顔の表情、頭、上体、体、空間、句読点）、50 のグループ、500 の基本記号が含まれる [6]。表 1 に記号とその SSS の値の例を示す。




3.2 SWML

SWML は SignWriting を計算機上でテキストとして扱うための XML ベースの言語である。Sign-

```
<?xml version="1.0"?>
<swml version="1.0" file-type="text" xmlns="http://www. ...
:
<line x="0" y="0" height="150" width="750">
<cell x="0" y="0" height="150" width="150" ...
<signbox>
<!--セル内での座標-->
<symp;x="69" y="77";x-flop="0" y-flop="0" color="0,0,0">
<category>01</category>
<group>06</group>
<sybnum>007</sybnum>
<variation>01</variation>
<fill>01</fill>
<rotation>01</rotation>
</symp>
<!--記号の水平/垂直反転-->
<symp x="97" y="56";x-flop="0" y-flop="1";color="0,0,0">
<category>02</category>
<group>05</group>
<sybnum>001</sybnum>
<variation>01</variation>
<fill>01</fill>
<rotation>02</rotation>
</symp>
:
</signbox>
</cell>
:
```

図 2 SWML による単語の記述例

表 1 記号とその SSS の値の例

カテ ゴリ	グル ープ	記号	変種	塗り	回転
	01	01	001	01	01
	01	02	001	01	02
	02	03	001	01	01

Writing 用のエディタ SW-Edit を使って生成した SWML テキストの一部を図 2 に示す。SSS では記号の回転は 16 とおり（8 方向の回転と左右の手の区別）用意されているが、SWML ではそれらを 4 つの回転とその水平/垂直反転で表現している。

4. 語彙辞書と SignWriting の生成

手話テキストにおける手話単語名からその SignWriting (SWML) 表現を検索できる手話語彙辞書をデータベース上に作成している。前述のように、手話単語は手の形、位置、動きの変化や非手指要素が言語的な意味をもつ場合があり、手話テキストではこれらが語形変化パラメータや句読点等として記述されている。このような変化を施した SignWriting 表現をあらかじめすべて辞書に用意するのは効率的でないため、辞書に登録された単語の基本形から必要に応じて、単語を構成する記号の種類や表示位置・塗り・回転を変更、あるいは、表情を追加することにより、可能な限り変化形を自動生成する。そのための情報も辞書に持たせている。

4.1 辞書の構成

辞書は Sign テーブル (表 2) と SWMLData テーブル (表 3) で構成される。Sign テーブルには IMWA 記号で構成された手話単語に関する情報を、SWMLData テーブルには手話単語を構成する個々の記号に関する情報を持たせている。主なフィールドに格納される情報は次のとおりである。

【Sign テーブルのフィールド】

- Sign: 手話単語名
- ID: 手話単語番号
- #symbols: 単語を構成する記号の個数
- head: 頭部記号を含むならその記号番号
- width, height: 単語の矩形サイズ

【SWMLData テーブルのフィールド】

- ID: 手話単語番号
- sym: 単語内の記号番号

表 2 Sign テーブルの例 (概略)

Sign	yomi	ID	#symbols	head	width	height	fixed
。		1	1	0	36	48	1
,		2	1	0	36	48	1
t		101	1	1	43	43	1
妻	つま	1001	3	0	79	76	0
仕事	しごと	1002	5	0	75	71	0
例	れい	1003	3	0	51	53	0

表 3 SWMLData テーブルの例 (概略)

ID	sign	sym	category	group	symbolnum	variation	fill	rotation	x-flop	y-flop	width	height	pos.x	pos.y	fixed	varshape
1	。	1	10	01	001	01	01	01	0	0	36	48	0	0	0	0
2	,	1	10	02	001	01	01	01	0	0	36	48	0	0	0	0
101	t	1	3	01	002	01	01	01	0	0	43	43	0	0	0	0
1001	妻	1	01	06	007	01	01	01	0	0	36	36	20	19	0	0
1001		2	02	05	001	01	01	02	0	1	36	36	0	40	0	0
1001		3	05	06	001	01	01	01	0	0	36	36	43	0	1	0
1002	仕事	1	01	05	008	01	04	03	0	0	36	36	39	15	0	0
1002		2	01	05	008	01	04	03	0	0	36	36	39	15	0	0
1002		3	02	01	001	01	02	01	0	0	43	43	13	0	0	0
1002		4	02	05	002	01	02	03	0	1	36	36	26	35	0	0
1002		5	02	05	002	01	01	03	0	1	36	36	7	35	0	0
1003	例	1	01	09	001	01	02	01	0	0	36	36	0	1	0	0
1003		2	01	05	009	01	01	02	0	1	36	36	7	0	0	0
1003		3	02	01	001	01	01	01	0	0	43	43	8	10	0	0

- category ~ y-flop: SignSymbolSequence
- width, height: 記号の矩形サイズ
- pos_x, pos_y: 単語矩形内での相対位置
- fixed: 頭部など位置固定記号なら 1
- varshape: 手形が変化可能なら 1

4.2 手話テキストから SignWriting への変換

手話テキストから SignWriting (SWML 形式) への変換は次のような手順で行う。

1. (単語への分割) 手話テキストでは一連の単語に同じ非手指動作が伴うような場合, それらを 1 つのブロックにまとめて記述するが, SignWriting では非手指要素等の情報も単語ごとに記述する必要がある。そこでまず, 手話テキストのもつ情報を単語ごとに分割する (句読点も単語として扱う)。以下は分割した単語ごとの処理である。
2. (空間上の位置の決定) 手話テキストでは空間上の位置を, 人称を表す記号で表現する。1 と 2 はそれぞれ話し手 (1 人称) と聞き手 (2 人称) の位置で固定だが, その他の記号は話し手の斜め前方, または, 話し手と聞き手の中間 (3 人称の位置) のどこかを表しており, 具体的な位置は特定していない。SignWriting では視覚的に表記するため具体的な位置 (現状では L (左), C (中央), R (右) のいずれか) を決定する。
3. (単語情報の取得) 手話単語名をキーに Sign テーブルを検索して単語基本形の矩形サイズ等の情報を得る。取得した単語 ID をキーに SWMLData テーブルから, 単語を構成する記号の情報を得る。
4. (手形の変更) 単語に手形が指定されている場合は, 手形 (単語名) を辞書から検索し, 元の単語の手形変更可能 (varshape=1) な記号を, 検索された記号で置換する。
5. (動詞の方向変化) 格関係によって変化する動詞の方向については, 辞書には 1 人称から 2 人称, またはその逆 (動詞によって異なる) を基本形として登録している。動詞の方向が指定されている場合は, 手およびその動作を表す記号の位置と回転を変更する。図 3 は, 記号の回転と位置のシフトを行って変化形を

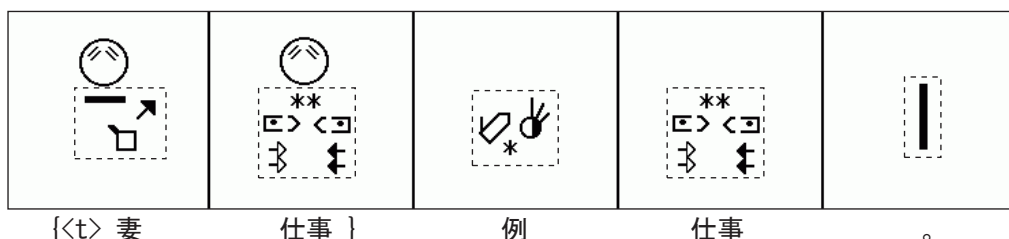


図5 SignWritingの生成例．下部は対応する手話テキスト．

生成した例である．

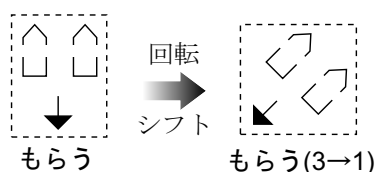


図3 動詞の方向変化

6. (非手指要素の追加) 単語に話題化や疑問を表す顔の表情が伴う場合、顔の記号を追加する。ただし、単語の基本形に頭部が含まれている場合は、同じ座標位置に表情を重ねる。
7. (セルへの単語の配置) 表現位置が指定されている場合は、対応する位置に単語を配置する(図4)。位置指定がない場合は基本的に中央に配置するが、肩のラインや頭部など位置が固定した記号を含む場合はその記号を基準に水平配置を決める。単語が頭部を含むか、手だけか、句読点かによって垂直位置を調整する。図5は「妻の仕事はパートです」を表す手話テキストに対するSignWriting表現をセルに配置した結果である。

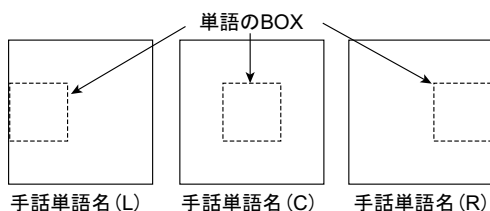


図4 セルへの単語の配置

5. おわりに

日本語からSignWritingで記述された手話への機械翻訳を目的とした手話の語彙辞書について述べた。辞書には単語の基本形と、語形変化の

生成に必要な情報を登録している。この辞書を用いて、手話テキストに記述された情報からSignWriting(SWML)を生成するプログラムを作成し、基本形からの手の形・位置・方向の変化や非手指要素を含むSWMLの生成を試みた。一部の処理はまだ手作業で行っている。また、語形変化のうち、記号の回転とシフトだけで対応できない方向変化には対応できていない。単語の動きの変化(大小・強弱・緩急など)によって表現される副詞等の情報や同時的な表現などのSignWritingへの変換についても今後検討していきたい。

謝辞 日頃手話に関してご教示・ご議論いただく手話言語研究会の皆様へ深く感謝します。なお、本研究の一部は、日本学術振興会科学研究費補助金(基盤C:18500111)により行われた。

参考文献

- [1] Costa, A.C.R. and Dimuro, G.P., "SignWriting and SWML: Paving the Way to Sign Language Processing," *TALN 2003*, pp.193-202 (2003).
- [2] 原田大樹, 松本忠博, 加藤三保子, 原大介, 池田尚志: 日本語を援用した手話表記およびSignWritingによる手話ビデオの書き取り実験, 言語処理学会第12回年次大会発表論文集, pp.719-722 (2006).
- [3] Matsumoto, T., Taniguchi, M., Yoshida, A., Tanaka, N. and Ikeda, T., "A proposal of a notation system for Japanese sign language and machine translation from Japanese text to sign language text," *PACLING2005*, pp.218-225 (2005).
- [4] 松本忠博, 原田大樹, 原大介, 池田尚志: 日本語を援用した手話表記の試み, 自然言語処理, Vol.13, No.3, pp.177-200 (2006).
- [5] Sutton, V., "Lessons in SignWriting" (<http://www.SignWriting.org/lessons/lessonsw/>), The Deaf Action Committee For SignWriting (2002).
- [6] Sutton, V., "Sutton Movement Writing: SSS-2004, The International Movement Writing Alphabet (IMWA)" (<http://www.SignWriting.org/archive/docs2/sw0144-SSS-Manual-2004>), The Center For Sutton Movement Writing, Inc. (2004).