

SignWriting 表記の手話文を対象とした 簡単なアメリカ手話-日本手話翻訳の検討

杉山 真也[†] 藤垣 俊也[†]
[†] 岐阜大学大学院工学研究科

松本 忠博[†] 加藤 三保子[‡]
[‡] 豊橋技術科学大学総合教育院

1 はじめに

SignWriting[1] (以下, SW) は手話文を読み書き可能にするために考案された手話の書記体系 (文字体系) であり, 現在複数の国で, ろう児教育などへの利用が試みられている. SW コミュニティの Web サイトでは, SW で記述された児童文学や手話に関する文書なども公開されている. SW の考案者がアメリカ人であることもあり, アメリカ手話 (American Sign Language, ASL) によるものが最も多い. アメリカ手話は主に北米英語圏で使用されているが, 母語を共有しない話者間の会話で共通語としても利用され, 国際手話にも影響を与えている.

手話は国や地域によって異なるが, 顔の表情などの非手指動作による文法標識や, 複雑な空間の利用, topic-comment 構造など, 文法的には共通する点も多い. 音声言語では全く知らない言語による話を理解することは困難だが, 視覚言語である手話では異なる言語間でも意思の疎通がある程度可能といわれている. 本研究では, ASL 文中の非手指文法標識などを保存したまま, 単語を日本手話 (JSL) に置き換えるだけでも, 日本の手話使用者が ASL 文を理解する助けになるのではないかと考え, SW で書かれた ASL 文を対象にした簡単な ASL-JSL 機械翻訳について検討した. 機械翻訳や外国語の学習には辞書が欠かせないが, 手話には標準的な文字がないこともあり, 異なる手話言語間の対訳辞書は見当たらない. そこで我々は既存の複数の辞書を組み合わせて, SW 表記の ASL の単語をそれに相当する JSL の単語に変換する ASL-JSL 対訳辞書 [2] を試作しており, 本研究では単語の置き換えにこの辞書機能を利用する.

2 SignWriting による単語表現

ここでは手話の表記に用いる SW について述べる. 言語を活用する上で文字は重要な役割をもつが, 手話

には文字がないため, これまでにいくつかの表記法が提案されている. SW の特徴は, 日常生活における利用を想定している点にある. 手話単語を構成する手の形, 動き, 位置, および, 顔の表情などの非手指要素を表す図像的な記号 (International SignWritingAlphabet, ISWA) を 2 次元的に配置して単語を表現することで, 人にとって比較的分かりやすい表現になっている.

内部的には, SW における手話単語は, 単語を構成する ISWA 記号の集合 $\{s_1, s_2, \dots, s_n\}$ であり, 各記号 s_i は, 記号コードと単語内での位置の情報を持つ. 記号コードは, 図 1 に示すような記号のカテゴリ (手の形, 動き, 頭・顔, 体など), カテゴリ内グループ, 基本記号, 手のひらの向き・視点, 指先の向きなどを表す 6 つの値からなる.

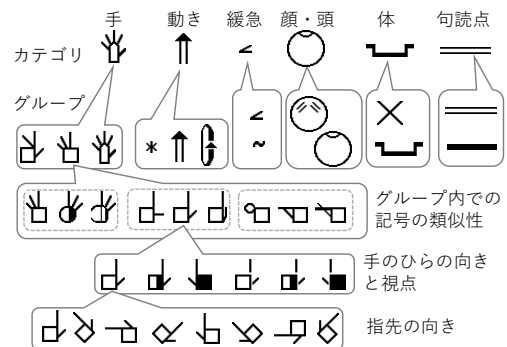


図 1: SW の表記に用いる記号 (文字) の種類

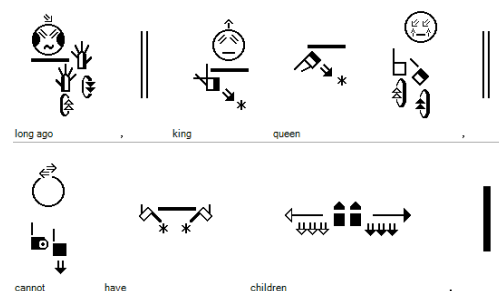


図 2: SignWriting による ASL 文の例 (SignPuddle[3], “Sleeping Beauty” より)

3 ASL-JSL 翻訳の検討

SW 表記の ASL 文から、それに相当する JSL 文を出力する処理の流れを図 3 に示す。処理は、SW 形式の ASL 単語の同定と英訳の取得 (図中①→②→③)、英語の語句から日本語ラベルで表された JSL 単語の取得 (③→④)、日本語ラベル形式から SW 形式の JSL への変換 (④→⑤)、文法標識の取得 (①→⑥) の各処理に分けられる。

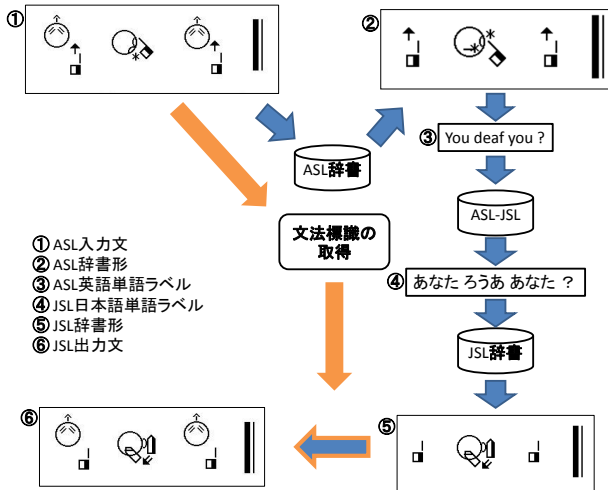


図 3: ASL-JSL 翻訳処理の流れ

3.1 ASL 単語の同定

SW による手話単語表記は視覚的に分かりやすい反面、自由度が高く、書き手による表記のゆれが大きいため、単語の同定が機械的にに行いにくい。そこで、文献 [4] の JSL の単語間類似度算出アルゴリズムを用いて、目的の単語を ASL 辞書から検索する。ASL 辞書は SignPuddle から取得した XML 形式のものを SQLite データベースに変換して用いる。

単語間類似度の算出にはまず、2 つの単語を構成する各記号間の類似度を求める。記号の類似度は、記号コードから得られる記号のカテゴリ、グループ、基本記号番号などの値と、各単語内での記号の位置のずれを元に算出する (図 4)。これにより、図 5 のような記号間類似度の表を作成する。なお、同じグループ番号を持つ記号でも類似していないものも存在するため、個別の記号の類似性も考慮している。

単語間の類似度は、単語を構成する記号の数の差が閾値 k より大きいなら 0 であり、そうでなければ、2 つの単語を構成する各記号の対応関係 (図 4) のうち、

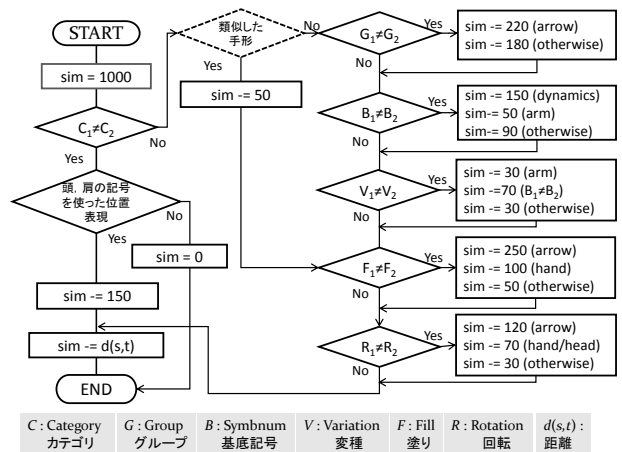


図 4: 記号間類似度の算出

	☞	☞	☞☞	☞☞	☞
☞	972	876	0	0	0
☞☞	884	970	0	0	0
☞☞☞	0	0	880	750	0
☞☞☞☞	0	0	630	1000	0
☞☞☞☞☞	0	0	0	0	934

図 5: 2 つの単語を構成する記号間の類似度の表 (類似度の最大値は 1,000)

最も記号間類似度の平均値が高くなるような組み合わせにおける記号間類似度の平均値と定義している。

3.2 ASL-JSL

英語ラベル形式の ASL 単語からそれに相当する JSL 単語を求める処理の流れを図 6 に示す。英語の日本語訳を取得する処理 (図中①)、日本語から日本語ラベル形式の JSL 単語を取得する処理 (②) を行う。



図 6: ASL-JSL 翻訳処理の流れ

(英語から日本語へ) ASL 辞書から得られた手話単語の英訳 (英語の語句) から, JSL 単語を表す日本語の語句 (日本語ラベル) を取得するため, まず, 英和辞書を参照して英語の語句の日本語訳を取得する. 参照する英和辞書の見出し語句の数は 973,710 である.

(日本語から JSL 単語名へ) 日本語に対応する手話単語を検索するため, 『わたしたちの手話 学習辞典』 (全日本ろうあ連盟, 全国手話検定試験出題単語), 『日本語-手話辞典』 (全日本ろうあ連盟) など複数の手話辞典を参考にして, 手話単語名 (単語の日本語ラベル) とその日本語訳の対応テーブル (SQLite データベース) を作成した. 登録手話単語の日本語ラベルによる異なり数は 9,366 語, 訳語の異なり数は 9,412 語である. ただし, 同じ手話単語が辞典によって異なる日本語ラベルで記載されていることも多いが, それによる単語の重複については考慮してないため, 実際の単語数はこれより少ない.

(シソーラスの利用) JSL の語彙は (辞書に掲載されたものを比較すると) 日本語の語彙よりかなり少ないため, ASL 単語の日本語訳が上記のテーブル上に存在しない可能性がある. その場合は, 日本語大シソーラスを参照して日本語訳の類義語を取得し, 再度 JSL 単語を検索する.

3.3 日本語ラベル形式から SW 形式の JSL への変換

手話文書エディタ JSPad[5] 用に構築中の SW 形式の JSL 辞書を使用して変換を行う. 図 7 に単語登録画面を示す. 現状での SW 形式の JSL 辞書の登録単語数は約 1,490 語である.



図 7: 単語登録画面

3.4 文法標識の取得

SW で記述された表情や頭の動きなどの非手指動作から, それが担う文法的役割の情報を取得する. ASL での非手指動作による文法標識の例を表 1 に示す.

表 1: 非手指動作による文法標識の例

文法標識	非手指動作	SignWriting
肯否疑問文	眉を上げ, 目を開き, 頭と体を前へ傾ける.	
WH 疑問文	眉をひそめる, 頭を前へ傾ける	
否定	首を左右に振る	

4 現状と課題

ASL 単語の英訳の取得 手話単語の同定精度評価としては, JSL 単語を対象として次のような実験を行っている. 手話単語のイラスト 20 語を用意し, 延べ 28 名の被験者が SW エディタでその単語を SW で記述する. 得られた各単語の SW 表現について, 明らかな誤り箇所は訂正した上で, 辞書に登録された目的の単語との類似度を求めた. 目的の単語との類似度の高さが 1 位となった場合を正解とする正解率の平均が 75.2%, 8 位以内を正解とする場合の正解率の平均が 90.6% となった. 辞書登録語数は約 1,100 語である.

ASL 単語を対象とした評価実験はまだ行っておらず, 検証はできていない. ASL 辞書の登録語数は 10,000 語余りと, JSL 辞書のおよそ 10 倍である. そのため, 形の類似した異なる単語が JSL 辞書より多く含まれている可能性があり, 同定の精度は低くなることが予想される. ただし, SW コミュニティで公開されている ASL 文書に現れる単語は, 全く同じ形で ASL 辞書に登録されている場合もあるため, そのような文書を対象とした評価では, 精度が向上する可能性もある. ASL 単語を対象とした評価と同定精度の向上は今後の課題である.

ASL-JSL 対訳辞書 ASL 辞書に載っている ASL 単語の英訳語句から JSL 単語に変換できた数と割合を表 2 に示す. 英語の和訳が得られないかかったものには

人名などの固有名詞のほか、句や文が多く含まれていた。手話では動詞に道具や対象、動作主体などの情報が埋め込まれるため、手話単語が音声言語では複数の単語で表されることとなり、辞書の見出しになりにくいことが原因と考えられる。

英訳語句から何らかの JSL 単語が取得できた割合は約 54% となった。しかし、シソーラスを介して JSL 単語を得られたものには、訳語としてふさわしくないものが多く混入している。より実用的な辞書を構築するには人手による修正・補完が必要である。

表 2: ASL から JSL に変換できた語句の数と割合

	ASL 単語	英訳語句
JSL 訳あり	2,810 (47.96 %)	3,119 (41.69 %)
JSL 訳あり (類)	709 (12.10 %)	898 (12.00 %)
JSL 訳なし	619 (10.56 %)	792 (10.59 %)
和訳なし	1,721 (29.37 %)	2,673 (35.73 %)
合計 (異なり数)	5,859 (100 %)	7,482 (100 %)

英語から日本語、または日本語から JSL 単語を取得する際に、複数の訳語が得られる場合がある。翻訳に用いるためには、複数の訳語から適切な訳語を選択する必要がある。

日本語ラベルから SW 表記の JSL への変換 現状では SW 形式の JSL 辞書の登録単語数が約 1,490 語程度とまだ少ないため、拡充を行う必要がある。

文法標識の取得 対応できている文法標識は、話題化、肯否疑問文、WH 疑問文、否定である。今後、SW で書かれた例文を調査し、他の文法標識に対応していく必要がある。

5 おわりに

SW で書かれた手話を対象として、非手指文法標識等を保存して、単語を置き換える簡易な ASL-JSL 翻訳について検討した。単語の置き換えには、既存の辞書を組み合わせて試作した ASL-JSL 対訳辞書を用いた。保存すべき文法要素としては、話題化などの一部の非手指動作のみを考慮している。その他の保存すべき文法要素への対応、単語の翻訳精度向上、生成された文の評価などを今後行っていく必要がある。

参考文献

- [1] Sutton, V. (2002). *Lessons in SignWriting, Textbook and Workbook*, 3rd ed., The deaf action committee for SignWriting.
- [2] 杉山真也, 松本忠博, 加藤三保子 (2014). SignWriting を用いたアメリカ手話-日本手話対訳辞書の検討, 言語処理学会第 20 回年次大会発表論文集, pp.31-34
- [3] Slevinski, S., Sutton, V., SignPuddle Online v2.0, <http://www.signbank.org/signpuddle/>.
- [4] 高瀬友宏, 松本忠博, 加藤三保子 (2012). SignWriting を利用した手話文書エディタ JSPad における手話-日本語辞書について, 言語処理学会第 18 回年次大会発表論文集, pp.1220-1223.
- [5] Matsumoto, T., Kato, M. and Ikeda, T. (2009). “JSPad—a sign language writing tool using SignWriting,” *Proc. 3rd Int’l. Universal Comm. Sym.*, pp.363-367. (<http://www.mat.info.gifu-u.ac.jp/jspad/>)