

気象情報の日本語一手話 CG 翻訳

加藤 直人[†] 宮崎 太郎[†] 金子 浩之[†] 井上 誠喜[†]
梅田 修一[†] 比留間 伸行[†] 長嶋 祐二[‡]

[†] N H K 放送技術研究所

[‡] 工学院大学情報学部

E-mail: [†] {katou.n-ga, kaneko.h-dk, inoue.s-li, umeda.s-hg, hiruma.n-dy}@nhk.or.jp

[‡] nagasima@cc.kogakuin.ac.jp

1 はじめに

手話は、聴覚障害者にとって重要なコミュニケーション手段である。特に、先天的あるいは幼少時に失聴した聾者にとって手話は母語であり、日本語より理解しやすい。しかしながら、手話による情報提供は少ない。

N H K でも手話による放送を行っているものの（例えば、「手話ニュース」）、健聴者向けの番組と比べるとその放送時間は多くない。さらなる手話放送の拡充が望まれるが、その実現のためには増大する手話翻訳に対処しなければならない、しかしながら、手話通訳者の数が少なく、翻訳量を増やすことは容易ではない。

手話翻訳の理想的な解決策は日本語から手話への自動翻訳 [1] の実現である。計算機に日本語を入力すると、手話を映像で出力するというものである。これまでも手話翻訳の研究がいくつか行われてきた [2] [3] [4]。しかしながら、そこで扱われている語彙規模は小さく、実験的なシステムである場合も少なくない。

我々は実用的なシステムをめざし、日本語テキストから手話 CG に翻訳する研究を進めている [5]。特に、天気や気象災害などの気象情報を対象として翻訳することを当面の目標としている。これは、地震や台風などの気象災害が深夜や早朝に起きた場合に、手話通訳者を確保することは困難であると考えられるからである。

本稿では、現在開発を進めている、気象情報の日本語一手話 CG 翻訳システムについて述べる。本システムでは、入力された日本語を節（句）単位で手話単語列に言語変換し、手話単語列から手話 CG を自動生成している。

2 日本語から手話への言語翻訳

2.1 基本的な考え方

言語翻訳の手法は、規則翻訳（規則ベース翻訳）、用例翻訳（用例ベース翻訳）、統計翻訳（統計的機械翻訳）に大別される。日本語から手話へ言語翻訳をするにあたっては、以下のように考えて、用例翻訳と統計翻訳を用いた。ここで、言語翻訳とは文字列から文字列への変換をいう。手話は視覚言語であるので、最終的には映像への変換が必要となる。しかし、日本語単文字列から手話映像へ直接変換することは難しいので、まずは手話単語列（手話動作を単語列に書き起こしたもの）に変換する。そして、変換された手話単語列から手話映像（CG）に変換する。

規則翻訳では人手によって構築された翻訳知識が必要となる。翻訳知識の構築には、翻訳対象の言語的知見が必要となるが、手話（日本手話）に関しては機械翻訳を開発するまでにはその解明が進んでいない。したがって、規則翻訳で早期に翻訳システムを開発することは現実的ではない。一方、統計翻訳では対象言語の知見が必要ないという利点がある。しかしながら、統計翻訳には大規模な対訳コーパスが必要であり、現在 100 万文～1,000 万文規模のものが使われている。日本語一手話の対訳コーパスにはそれほど大きなものではなく、比較的大きな我々の対訳コーパス [6] でも 2 万～3 万文程度と遥かに小さい。今後も対訳コーパスを増やしていく予定であるので、将来的には統計翻訳の適用も考えられるが、現時点では難しい。その点、同じコーパスベースの翻訳手法である用例翻訳は対訳文を直接利用しているため、統計翻訳ほど大きなサイズでなくとも、ある程度の

翻訳精度を得ることが可能である。さらに、我々は当面の翻訳対象を気象情報に絞っているのも、そこに出現する言語現象はある程度限られたものとなる。特に、節や句単位でみると気象情報は定型的な表現が多い。もちろん用例翻訳では、統計翻訳と異なり、翻訳対象に対する言語的知見が必要となるが、規則翻訳ほど詳細なものでなくともよい。

以上の考察により、我々は用例翻訳を主とし、統計翻訳を併用する方法をとることにした。本システムでは、まず、節（句）単位で完全一致による用例翻訳を行う。用例との完全一致なので、用例とは一致しない節（句）もある。用例と一致しない場合には統計翻訳で翻訳した。

概略を例を用いて説明する。次の日本語文を翻訳することを考える。

【日本語文 1】（入力）

「九州と沖縄は夕方から雷雨となりそうです」
ただし、節（句）単位の用例として、次が与えられているとする。

【用例 1】（節・句単位）

「九州と」⇔「九州 N」
「沖縄は」⇔「沖縄」
「九州と沖縄は」⇔「九州 N 沖縄」
「雷雨となりそうです」⇔「雷雨 夢 N」
「雷雨となり」⇔「雷雨」
「そうです」⇔「夢 N」

ここで、矢印（⇔）の左辺が日本語、右辺が手話である。手話の表記は、全日本ろうあ連盟が発行している日本語一手話辞典 [7] にしたがった。また、「N」は非手指動作である「頷き」を表し、意味をもつ単位の切れ目を表す。

本システムではまず用例翻訳を行う。すなわち、日本語文の中で節（句）用例を適用して翻訳可能な箇所は翻訳してしまう。例えば、日本語文 1 に用例 1 を適用すると、日本語の節「九州と沖縄は」と「雷雨となりそうです」はそれぞれ、「九州 N 沖縄」、「雷雨 夢 N」と翻訳することができる。しかし、用例にない「夕方から」は用例翻訳できない。そこで、用例翻訳できなかった単語列は統計翻訳で翻訳する。日本語一手話の翻訳では語順の入れ替えが少ないので、統計翻訳でも精度よく翻訳可能であることが期待できる。

2.2 節・句の用例翻訳

用例翻訳の用例と翻訳手法について説明する。

用例は節（句）単位に分割されているが、それは人手で行った。分割するにあたっては、日本語側の節（句）の切れ目と、手話側の意味的単位の切れ目を考慮した。手話側での切れ目は主に頷き（N）で判断した。また、手話側の切れ目の判断が難しいときもあったが、そのような場合には無理に切らなかった。

用例翻訳では、用例はなるべく長いもののほうが精度よい翻訳が期待できる。そこで、一文中で分割した節（句）から、そのすべての組み合わせを自動生成し、用例に追加した。例えば、次の日本語文 2 が人手により節（句）単位に分割されたとき、用例 2 が得られたとする。用例 2 のすべての組み合わせを自動生成すると、用例 3 が得られるので用例に追加する。

【日本語文 2】（文単位）

「きょうは雷雨となるでしょう」
⇔「今日 N 雷雨 夢 N」

【用例 2】（人手による節（句）分割）

「きょうは」⇔「今日 N」
「雷雨となる」⇔「雷雨」
「でしょう」⇔「夢 N」

【用例 3】（追加された用例）

「きょうは雷雨となる」⇔「今日 N 雷雨」
「雷雨となるでしょう」⇔「雷雨 夢 N」
「きょうは雷雨となるでしょう」
⇔「今日 N 雷雨 夢 N」

このようにして、気象情報の対訳約 3,500 文から約 10,000 個の節（句）単位の用例が得られた。

用例翻訳は CYK (Cocke-Younger-Kasami) テーブル [8] を埋めていくことによって行う。

今、入力された日本語文を形態素解析した結果、

$w_1 w_2 \dots w_i \dots w_j \dots w_n$

と得られたとする。本システムでは形態素解析に茶筌 [9] を用いた。構文解析は行っていない。

まず入力文に一致する用例を見つける。文中の日本語単語列 $w_i \dots w_j$ に対して節（句）単位の対訳コーパスと照合し、一致した場合にはその手話単語列を $t(j-i+1, j)$ に格納する。

翻訳処理では、ある日本語の部分単語列 $w_i \dots w_j$ の翻訳結果 $t(j-i+1, j)$ は、2 つの連続する部分

図 1 に日本語文 1 に対する CYK テーブルの様子を示す。この場合には $t(11, 11)$ が空であるので、用例翻訳のみでは一文全体は翻訳できていない。部分訳を得るために CYK テーブルを上から順に見ていくと、 $t(5, 11)$ が空でないことがわかる。すなわち、部分単語列「雷雨となりそうです」が「雷雨 夢 N」と手話単語列に翻訳されている。残りの単語列に同様な処理を繰り返すと、 $t(4, 4)$ が空でないので、「九州と沖縄は」は「九州 N 沖縄」と翻訳される。用例翻訳されなかった単語列（「夕方から」）は次の統計翻訳で翻訳される。

統計翻訳では、翻訳モデルの学習には用例翻訳で使った節（句）単位の用例を利用し、言語モデルの学習には文単位の手話文を利用した。翻訳モデルにはGIZA++ [10] を使い、デコードにはmosesを使った [11]。言語モデルはSRILM [12] を使用した。

ここで、手話 CG よる動作生成について簡単に説明する。詳細については文献 [13] [14] を参照されたい。動作生成は言語翻訳の結果から、各単語を日本語手話対訳辞書 [15] を使って変換し、各手話単語のモーションを線形補間で繋いでいくことにより文を生成する。我々の日本語一手話対訳辞書には日本語語彙約 86,600 語、手話語彙約 4,900 語が登録されている。手話語彙は、光学式モーションキャプチャによって取得した動作データと、骨格構造を持った人体 CG モデルで構成されている。これは映画やビデオゲームで用いられる手法であり、自然な動作を再現できる。また、CG の描画は映像コンテンツ記述言語 TVML [16] を使っているので、容易に動作編集が可能となる。



4 おわりに

現在開発を進めている、日本語一手話 CG 翻訳システムについて述べた。本システムの言語翻訳は、節（句）単位の用例翻訳と統計翻訳を用いているという特徴がある。

今後はさらにコーパス規模の拡大を行い、システムの改善をはかりたい。そのためにはシステムの評価が必要となる。評価には、言語翻訳の評価とともに手話 CG の評価があるが、後者は評価手法そのものが確立しておらず、評価手法も含めて考える必要がある。

参考文献

- [1] 市川熹, 神田和幸, 黒川隆夫, 佐川浩彦, 長嶋祐二, 神田和幸, “手話日本語間の自動翻訳とその周辺,” ヒューマンインターフェース学会誌, Vol. 3, No. 3, pp. 179–188, 2001.
- [2] 日立製作所, 手話アニメーションソフト MimehandII, <http://www.hitachi.co.jp/Prod/comp/app/shuwa/mimehand/index.html>
- [3] 田丸満枝, 猪木誠二, 山本泰則, 宇治谷恵, 杉田繁治, “モバイル環境での手話アニメーションによる情報提供,” 電子情報通信学会 HCG シンポジウム 2002, pp. 453–454, 2002.
- [4] 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.

[5] 加藤直人, 金子浩之, 井上誠喜, 梅田修一, 比留間伸行, 長嶋祐二, “用例利用による日本語一手話 CG 翻訳システム,” 電子情報通信学会 HCG シンポジウム 2011, 1–1, 2011.

[6] 加藤直人, “手話ニュースコーパスの構築,” 言語処理学会年次大会, PA2-5, pp. 494–497, 2010.

[7] 米川明彦 (監修), “日本語一手話辞典,” 日本手話研究所 (編), (財) 全日本聾唖連盟出版局, 2006.

[8] Alfred V. Aho, Jeffrey D. Ullman, “The Theory of Parsing, Translation, and Compiling,” vol. 1, Prentice-Hall, 1972.

[9] <http://chasen.naist.jp/>

[10] <http://www.statmt.org/moses/giza/GIZA++.html>

[11] <http://www.statmt.org/moses/>

[12] <http://www.speech.sri.com/projects/srilm/>

[13] 金子浩之, 浜口斉周, 道家守, 井上誠喜, “TVML による手話アニメーションの一検討,” 電子情報通信学会技術研究報告, WIT2008-82, pp. 79–83, 2009.

[14] 金子浩之, 加藤直人, 井上誠喜, 清水俊宏, 長嶋祐二, “非手指動作を付加した手話映像生成に関する一検討,” 電子情報通信学会 HCG シンポジウム 2010, A5-2, pp. 89–93, 2010.

[15] 加藤直人, 金子浩之, 井上誠喜, 清水俊宏, 長嶋祐二, “日本語一手話対訳辞書の構築 ～日本語語彙の拡張～,” 電子情報通信学会 HCG シンポジウム 2009, 1–3, 2009.

[16] <http://www.nhk.or.jp/str1/tvml/index.html>

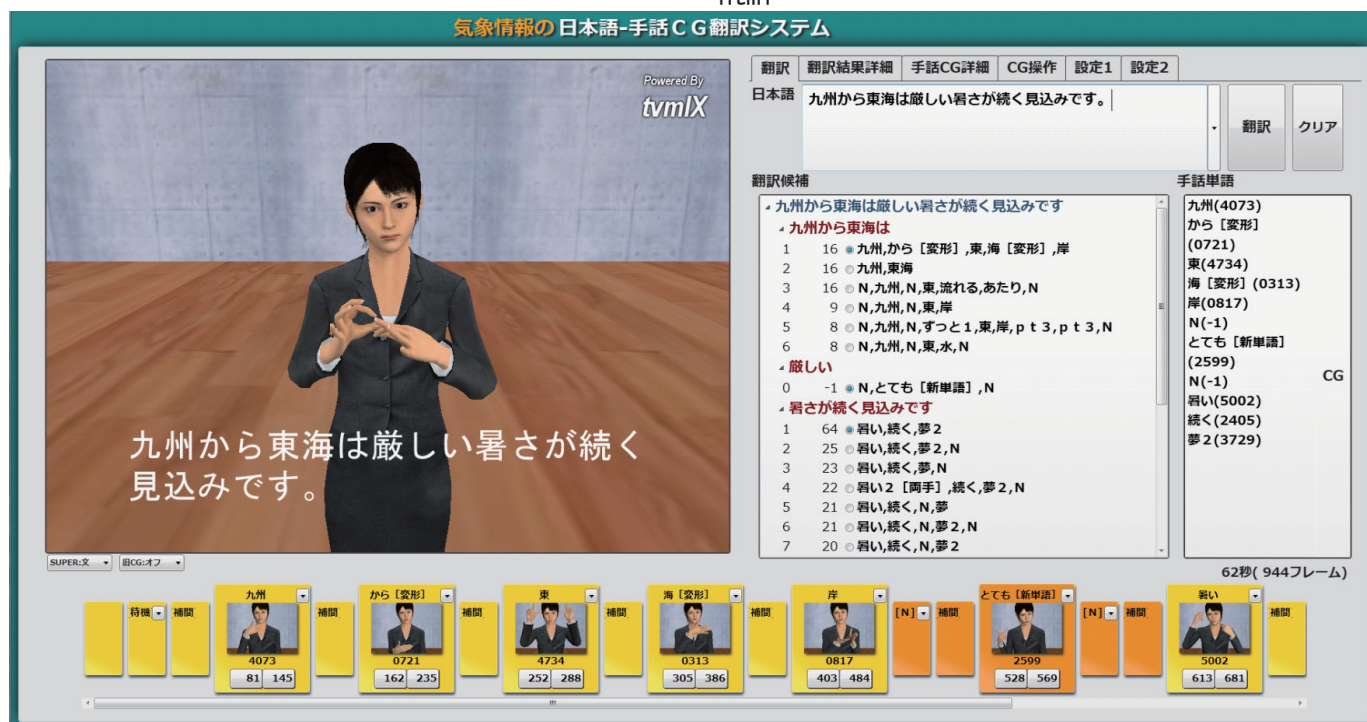


図2 気象情報の日本語一手話CG翻訳システム