

日本語 - 手話の単語アライメントによる 非手指動作の検出

加藤 直人 宮崎 太郎

NHK放送技術研究所

E-mail: {katou.n-ga, miyazaki.t-jw}@nhk.or.jp

1 はじめに

手話サービスの拡充をめざし、テキストから手話CGに自動的に翻訳（手話CG翻訳）する研究を進めている [1]。手話は、先天的あるいは幼少時に失聴した聾者にとって母語であり、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増えてはいるものの、まだまだ少ない。これは、日本語から手話への“翻訳”という作業を伴うからであろう。手話は日本語とは異なる言語であるので、英語への翻訳のように、手話への翻訳が必要となる。

手話CG翻訳実現への第一歩として、我々は気象ニュースの手話CG翻訳システムを開発している [2]。本システムにニュース映像と日本語字幕を入力すると、日本語を手話CGに自動翻訳し、映像と同期がとれるように手話CGの時間長を自動調整する。図1にシステムの出力例を示す。しかしながら、まだ翻訳精度は十分とは言えない。特に問題となっているのは非手指動作である。手話では手や指の動きなどの手指動作とともに、顔の表情や頭部の動きなどの非手指動作によって言

語的情報を表出しているが、我々のシステムでは非手指動作はあまり扱っていない。これは、機械翻訳にコーパス方式を採用しているが、その対訳コーパスに非手指動作の情報がほとんど入っていないためである。したがって、対訳コーパスに非手指動作の情報を人手で付加していけばよいと考えられるが、実際には難しい。非手指動作は言語的役割とともに、音声言語でいうプロソディの役割も持っており、その区別をつけることが人間でも難しいからである。

本報告では、日本語 - 手話対訳コーパスに非手指動作の挿入箇所を検出する手法について述べる。本手法では、まず対訳コーパスに対して単語アライメントを行う。単語アライメントの際には日本語と手話では単語の切れ目が異なることが問題となるが、本手法では日英対訳辞書を利用することにより、日本語単語を自動的に分割している。非手指動作の挿入箇所は、単語アライメントされなかった箇所を発見することで、自動検出する手法を提案する。また、提案手法による単語アライメントと非手指動作検出の評価実験を行ったので報告する。

気象ニュースの手話CG翻訳システム



図1 気象ニュースの手話CG翻訳

2 日本語 - 手話対訳コーパス

2.1 手話ニュースコーパス

手話 CG 翻訳システムではコーパスベースの機械翻訳を行っている [3]。手話の文法がまだ十分に解明されていないので、規則翻訳（規則ベース翻訳）が困難だからである。具体的には、用例翻訳（用例ベース翻訳）と統計翻訳（統計的機械翻訳）の 2 つの手法を併用している。我々が当面の対象としている気象ニュースは定型表現が多いので、句や節単位で定型表現を用例翻訳し、用例翻訳できなかった箇所は統計翻訳を行う。

コーパスベースの翻訳には大規模な対訳コーパスが必要となる。我々も日本語 - 手話対訳コーパスとして「手話ニュースコーパス」を構築している [4]。これは NHK・Eテレの「手話ニュース」[5] を書き起こしたものである。手話ニュースコーパスは現在、約 62,000 文の日本語と手話の対訳で構成されている。機械翻訳に利用する用例はこの手話ニュースコーパスから気象ニュース 5,300 文を抽出し、人手で作成した。その結果、約 3 万個の用例が得られている。

手話を書き起こす際に問題となるのは表記法である。手話ではいくつかの表記法が提案されているものの、確立された表記法はない [6]。手話ニュースコーパスでは、単語ごとの表記法である日本語ラベルを使って書き起こした。日本語ラベルでは手話単語は、“{雨}” のように、手話単語と意味的に近い日本語単語を括弧“{}”で括って記述する。手話ニュースコーパスでは日本語ラベルの定義は、全日本聾啞連盟が発行している日本語 - 手話辞典 [7] を参考にしている。日本語ラベルは単語ごとの書き起こしであるので、機械翻訳で扱いやすいという利点がある。

2.2 非手指動作

手話ニュースコーパスは手話の書き起こしではあるが、すべての手話情報を書き起こしているわけではない。主に手指動作（手や指などの動作）であり、非手指動作（顔の表情や頭部の動きなど）は一部のみしか書き起こされていない。これは、手指動作は単語の同定をしやすいが、非手指動作は難しいからである。非手指動作は言語的役割や音声言語のプロソディの役割もあり、その区別をつけて書き起こすことは人間でも容易ではない。したがって、現在、非手指動作の書き起こしの対象としているのは、比較的同定がしやすい、『頷き』（図 2 中の“N”）と『首振り』のみである。

しかしながら、コーパスに依存している我々の機械翻訳手法では、言語的役割をもつ非手指動作を扱わなければ、翻訳精度の低下をまねく。例えば、図 2 の例では、日本語単語「強い」は、顔の表情で（険しい表情で）表出しているので手話では書き起こされていない。したがって、図 2 の翻訳例を使ってしまうと日本語単語「強い」は手話では翻訳されないので、正しく翻訳できないという問題が生じる。

3 単語アライメントによる非手指動作の検出

提案手法では逆に日本語単語に対応する手話単語がない箇所を発見することによって、非手指動作の挿入箇所を検出する。すなわち、日本語と手話の単語アライメントを行い、手話単語とアライメントが取れなかった日本語単語（主に内容語）を見つけ出すことにより、非手指動作の挿入箇所とする。例えば、図 2 のような単語アライメントを行うと、日本語単語「強い」は手話単語とはアライメントが取れないので、非手指動作と推定さ

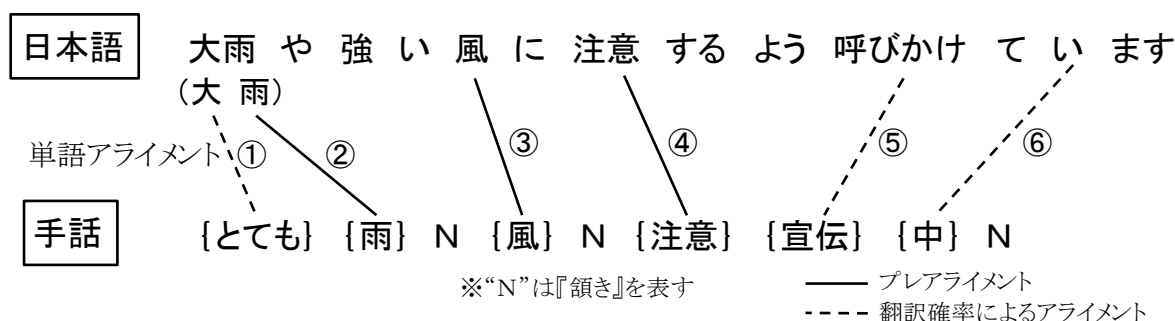


図 2 単語アライメントの例

れる。また、非手指動作の表出は、日本語では「強い」は「風」に係るので、手話では“{風}”という手指動作と同時にする。したがって、日本語単語「強い」に対応する非手指動作を「険しい顔表情」と定義とし、同時に表出する動作を記号+で定義しておく、手話のコーパスでは“{風} + [険しい顔表情]”と記述することができる。手話CGの生成では、“{風}”のモーションデータ上に「険しい顔表情」のデータを付加すればよい。

非手指動作を自動検出するためには、まず手指動作の単語アライメントを精度よく行う必要がある。そのために、提案手法では、プレアライメントと、日英対訳辞書を利用した単語分割を行っている。以下で説明する。

3.1 プレアライメント

プレアライメントでは、日本語単語が手話単語と表層的に一致した場合にはアライメントする。これは手話の書き起こしは意味的に近い日本語単語で定義しているために可能となる。

例えば、図2の場合では、次の2つの単語がアライメントできる。

風 ⇔ {風} (単語アライメント③)

注意 ⇔ {注意} (単語アライメント④)

さらに、日本語単語「大雨」では、それを構成している文字「雨」に注目すると、手話単語“{雨}”とアライメントできることがわかる。プレアライメントでは、この例のように日本語単語中の一部の文字が手話単語と一致するならば、日本語単語を分割してアライメントする。「大雨」の例では、「雨」が手話の“{雨}”と一致するので、「大」と「雨」と分割してアライメントする。

雨 ⇔ {雨} (単語アライメント②)

ただし、この段階では日本語文字「大」のアライメント先は決まらない。

上述したアライメントは、例えば、文中に同じ単語が複数出現した場合には、誤ってアライメントする可能性がある。そこで、その可能性を低減するために文中の出現位置による制約を加えた。手話ニュースコーパスの日本語と手話では、語順が大きく変わらない場合が多いので、文中での出現位置が近い単語同士をアライメントするほうがよい。具体的には次式で計算される値が閾値以下

の単語のみをプレアライメントした。

[文内出現位置制約]

$$\text{dist}(j_p, s_q) = |p/N_j - q/N_s|$$

ここで、 j_p, s_q はそれぞれ日本語文中 p 番目の単語、手話文中 q 番目の単語を表す。また、 N_j, N_s はそれぞれ、日本語文の単語数、手話文の単語数を表す。

3.2 日英対訳辞書を利用した単語分割

プレアライメントでは、手話側の情報を使うことにより、日本語単語「大雨」を「大」と「雨」に分割できたが、そのような情報がない場合もある。例えば、日本語単語「猛暑」は手話では“{とても} {暑い}”と翻訳される。したがって、次のように単語アライメントしたいが、表層の一致がないためにできない。

猛 ⇔ {とても}

暑 ⇔ {暑い}

あるいは、すべての漢字文字列は文字に分割する手法も考えられる。しかし、日本語単語「強風」は手話では“{台風}”と一単語で翻訳されることもあるので、必ずしも2文字（「強」と「風」）に分割すべきではない。

漢字に分割するのがよい場合とそうではない場合があるので、提案手法では、日英対訳辞書を使い、英訳の単語数で判断した。例えば、「猛暑」の英訳は“fierce heat”と2単語であるので分割する。一方、「強風」の英訳は“gale”であるので分割しない。ただし、「強風」はプレアライメントで分割される場合はある。

単語アライメントは対訳コーパスから推定される翻訳確率を使って行う。その推定はプレアライメント、日英対訳辞書を利用した単語分割を行った後、単語アライメントが取れなかった単語のみを対象にし、EMアルゴリズム [8] を適用して行った。ただし、その適用の際にも文内出現位置制約を用いた。

我々のコーパスからは、例えば、次のように翻訳確率を推定でき、図2のようにアライメント①、⑤、⑥が得られる。

大 ⇔ {とても} (翻訳確率 = 0.54)

呼びかけ ⇔ {宣伝} (翻訳確率 = 0.65)

い ⇔ {中} (翻訳確率 = 0.16)

3.3 評価実験

単語アライメントと非手指動作検出の評価実験を行った。ただし、今回は、単語分割は2文字漢字のみを対象とし、評価は日本語単語「猛暑」と「強風」で行った。単語アライメントの評価ではこれらの単語（あるいは文字）が文中で正しくアライメントされているか否かで行った。「猛暑 (fierce heat)」はプレアライメントでは単語分割されないが、日英対訳辞書で分割される。分割されると、猛⇔ {強力}、暑⇔ {暑い} のようなアライメントが可能となる。また、「強風 (gale)」はプレアライメントで分割される場合があるが、日英対訳辞書では分割されない。分割されないと、強風⇔ {台風} のようなアライメントが可能となる。一方、非手指動作検出の評価は、「猛暑」の「猛」や「強風」の「強」が顔の表情で表出されると検出されたときに、実際の映像で顔表情が表出されているか否かで行った。

また、比較のために、2文字漢字を分割する場合と分割しない場合も実験した。すなわち、

Char : プレアライメント
+ 2文字漢字を分割する

Word : プレアライメント
+ 2文字漢字を分割しない

Proposed: プレアライメント
+ 日英対訳辞書による分割

ただし、日英対訳辞書には文献 [9] の語彙拡張で用いた辞書を使った。

単語アライメントと非手指動作検出の評価をそれぞれ、表1、表2に示す。表1を見ると、Char手法やProposed手法がWord手法よりも精度が高い。これは「猛暑」では単語分割をしたほうが有効であるためであろう。一方、「強風」ではWord手法やProposed手法がChar手法より精度が高い。これは単語分割をしないほうがよい場合も少なくないからである。Proposed手法は「猛暑」でも「強風」でも精度がよいので、日英対訳辞書が有効に機能していると考えられる。表2から、今回非手指動作を検出した箇所はすべて非手指動作であった。しかし、Char手法とProposed手法には、検出された非手指動作の個数にほとんど差がなかった。ただし、今回は検出個数が少ないので、今後は多くの日本語単語で評価実験する必要がある。

表1 アライメント精度

	Char	Word	Proposed
「猛暑」	57.3%(47/82)	26.8%(22/82)	52.4%(43/82)
「強風」	78.9%(30/38)	97.4%(37/38)	92.1%(35/38)

表2 非手指動作検出精度

	Char	Word	Proposed
「猛暑」	100%(15/15)	0%(0/0)	100%(14/14)
「強風」	100%(3/3)	100%(5/5)	100%(3/3)

4 おわりに

対訳コーパスで日本語-手話の単語アライメントを行い、非手指動作を検出する手法について述べた。提案手法では、日英対訳辞書を利用することにより、日本語の漢字2文字単語を自動的に分割することにより、検出精度の向上を評価実験により確認した。

提案手法では自動的に非手指動作を挿入できるが、実際には自動検出した箇所を手で確認してもらうことで、挿入精度の向上が可能である。今後は手話ニュースコーパスに非手指動作も加え、手話CG翻訳の精度向上をめざす。

参考文献

- [1] 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-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, I-1, 2011.
- [2] 加藤直人, 宮崎太郎, 井上誠喜, 梅田修一, 清水俊宏比留間伸行, 長嶋祐二, “気象ニュースを対象とした手話CG翻訳システム,” 電子情報通信学会 HCG シンポジウム 2013, B-10-1, 2013.
- [3] 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, “気象情報の日本語-手話CG翻訳,” 言語処理学会年次大会, PA1-21, pp. 275-278, 2012.
- [4] 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.
- [5] NHK手話ニュース : <http://www.nhk.or.jp/str/tvml/index.html>
- [6] 加藤直人, “手話における言語資源の研究動向,” NHK技研 R&D, No. 139, pp. 10-19, 2013.
- [7] 米川明彦監修, “新 日本語-手話辞典,” 中央法規出版, 2011.
- [8] Peter F. Brown, Stephen Della Pietra, Vincent J. Della Pietra, Robert L. Mercer, “The Mathematics of Statistical Machine Translation: Parameter Estimation,” Computational Linguistics Vol. 19, No. 2, pp. 263-311, 1993.
- [9] 加藤直人, 金子浩之, 井上誠喜, 清水俊宏, 長嶋祐二, “日本語-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, I-3, 2009.