

## 外電経済ニュースの英日機械翻訳

相 沢 輝 昭<sup>†</sup> 加 藤 直 人<sup>††,\*</sup> 鎌 田 雅 子<sup>†††</sup>

広範なニュース文を精度よく翻訳するための分野別機械翻訳システムを目指して、まず経済分野を取りあげ、外電経済ニュースにチューニングした英日機械翻訳システムをインプリメントした。チューニングには、経済記事に頻出する定型的な表現パターンを利用している。すなわち、一方ではそれらを翻訳テンプレートの形で直接的に活用し、それを用いて定型文に対する100%の翻訳率を得た。他方では、経済記事専用の解析文法の構成に間接的に活用して、分野限定のない一般文法に比べて約5分の1の規則数ですむ、解析精度のよい経済分野専用の文法を構築した。翻訳には、原文の定型性の程度に応じて、翻訳テンプレート、経済専用文法、一般文法を切り替えて実行する。翻訳実験の結果、一般文法のみを用いた場合に比べて大幅に高い、70%近い翻訳率が達成できた。

### English-to-Japanese Machine Translation for Wire-Service Economic News

TERUAKI AIZAWA,<sup>†</sup> NAOTO KATOH<sup>††,\*</sup> and MASAKO KAMATA<sup>†††</sup>

We have implemented an English-to-Japanese machine translation system tuned to AP wire-service economic news stories. The tuning is done by using expression patterns frequently observed in economic news stories. This system directly uses these patterns to construct bilingual sentence templates, which allow the system to achieve perfect translation for many fixed sentences. The patterns can also be used indirectly enabling a grammar writer to make a grammar compact. In fact, an economy-specific grammar thus compiled is only 1/5 the size of a general (nonspecific) grammar. Experimental results show that the system can achieve good translation for about 70% of source sentences.

#### 1. はじめに

機械翻訳は完成された技術とはいえないが、使えるところには使って逆にそこから具体的問題を引き出し、それを解決することで一歩前進する、そんな進み方もあるのではないか。このような考え方から、ニュース文を対象とした英日機械翻訳システムの試用ならびに改善研究を重ねてきた<sup>1)~3)</sup>。

取りあげた対象とその課題は、英語のTVニュースとその日本語字幕の作成、および外電ニュースとその高速粗翻訳の2つであったが、本稿では後者について、特に経済ニュースへの機械翻訳システムのチューニングの試みとその結果について報告する。

いうまでもなくニュースの内容は多岐にわたってい

るので、政治、経済、社会、文化、科学、スポーツといったニュースの全分野を1つの包括システムで翻訳させるのはやはり難しい。このことは話し言葉主体のTVニュースよりも書き言葉主体の外電ニュースの場合に顕著で、とりわけ経済ニュースとスポーツニュースに対する翻訳結果が思わしくなかった。これを解決する1つの現実的手法は、ニュース分野ごとに専用システムを用意し、それらを切り替えて使うやり方であろう。そこで手はじめに、経済分野向き専用システムの構築を目指して、外電経済ニュースへの機械翻訳システムのチューニングを行い、その可能性を実験的に検討した。

チューニングの基礎になるのは経済ニュースに頻発する定型表現である。それらは、約1年分の外電経済ニュースを統計的に調査することで抽出した。得られた定型表現は、一方では翻訳テンプレートの作成に直接的に活用し、他方では経済専用解析文法の構成に間接的に活用した。このように、この翻訳システムは必然的に統計的<sup>4)</sup>、実例的<sup>5)~8)</sup>側面を持つことになる。

翻訳テンプレートは定型的な文を単位に構成した。

<sup>†</sup> 広島市立大学

Hiroshima City University

<sup>††</sup> NHK 放送技術研究所

NHK Science and Technical Reserch Laboratories

<sup>\*</sup> 現在、ATR 音声翻訳通信研究所

<sup>†††</sup> 漢字情報サービス(株)

Kanji Infomation Services

定型文としては、それを構成する語の 80% 以上が定型表現によって重複なく被覆されるものを考えた。これは佐藤<sup>5)</sup>の複数の事例による照合表現と似ているが、被覆しきれない部分を変数として残す点、入力文に対して動的に作成するのではなく前もって翻訳テンプレートの形にまとめておく点で異なっている。事例翻訳では入力文に類似の実例の検索が難しく種々の手法が提案されているが<sup>6),7)</sup>、我々は翻訳テンプレートを出現頻度の高い約 400 の定型文に絞ることでこの問題を現実的に回避した。その限りで 100% の翻訳率が得られている。また翻訳テンプレートの作成も、対象とした定型原文に対する日本語訳文を与えるだけで DP マッチング手法により自動化でき、この点で、Kaji ら<sup>8)</sup>の原文と訳文の構文解析結果を利用するテンプレート自動作成手法より簡単なものになっている。

以下、まず 2 章で、外電経済ニュース記事はその表現の定型性に依拠して 3 種類に大別できることを示す。それに基づいて 3 章で、4 種類の手法を切り替える翻訳方式を提案する。次に 4 章で、この翻訳方式にとって重要な、翻訳テンプレートの作成と経済専用文法の構成について述べる。最後に 5 章で、翻訳結果とその評価を述べる。

## 2. 経済ニュース記事の 3 タイプ

AP 通信社からの外電（いわゆる AP 電）は一日約 350 記事（約 700 KB）入電するが、そのうちの約 50 記事が本稿で機械翻訳の対象とした経済ニュースである。それらは以下の特徴を持つ。

- 1) 経済記事は毎日ほぼ一定の順序で送られてくる。
- 2) それぞれの記事にはヘッダーがついていて、しかもヘッダーセットはあらかじめ定まっている。
- 3) そのヘッダーによって、経済記事は次の 3 種類に分類できる。

### (1) 定型文記事（約 10 記事）

数字、曜日等の少部分のみが変化する定型文で構成されるもの。

[例] ヘッダー：Yen-Dollar Opening

“The U.S. dollar opened at 144.58 yen on the Tokyo foreign exchange market Tuesday, up from Monday's close of 144.28 yen.”

ここでは下線を施した部分のみが変化する。このタイプの記事は非常に短く、ほとんどは 1 文だけから成り、長くても 3 文を超えることはない。

### (2) 表形式記事（約 30 記事）

表とそれに対する定型的な説明文のみで構成されるもの。

[例（一部）] ヘッダー：World Gold

“Selected world gold prices Monday in dollars:  
 Hong Kong late: holiday.  
 London morning fixing: holiday.  
 Paris afternoon fixing: holiday.  
 Frankfurt fixing: 360.15, off 0.86.  
 NY Handy and Harman: 360.20, off 0.25.”

ここでも下線を施した部分のみが変化する。通常は数字の変化であるが、この例のように、数表現全体が“holiday”に置き換わるようなこともある。

### (3) 一般文記事（約 10 記事）

特定のスタイルや定型的パターンを持たない文から構成されるもの。ただしまったくの自由記述とはいええず、そのことは次の例で示されよう。

[例] ヘッダー：Japan Market

この記事の 3 日分の冒頭文は次のとおりである。

“The U.S. dollar opened slightly lower against the Japanese yen Monday, while stock prices rebounded.”

“The U.S. dollar edged up in early trading against the Japanese yen Tuesday, while share prices on the Tokyo Stock Exchange rose moderately.”

“The U.S. dollar fell sharply in early trading against the Japanese yen Wednesday, while share prices on the Tokyo Stock Exchange moved slightly higher.”

これらは確かに定型文とはいえないが、相互に似通っている。さらにこの記事を読み進めていくと、以下のような表現が、必ずこの順に、現れることが分かる。

“The dollar started (or began) ...”

“Currency dealers said ...”

“On the stock market ...”

しかも、“Currency dealers said”の後には様々な内容が予想されるにもかかわらず、実際には、日本とアメリカの市場比較の話題に限定されているのである。

以上のことから、このタイプの記事は、何らかのガイドラインもしくはスケルトンに従って書かれていると推察できる。したがって、それに特化した精度のよい解析文法の構成が期待できる。また、かなりの定型表現も含まれている可能性がある。

## 3. 翻訳処理の構成

経済ニュース記事に関する上記の 3 タイプを考慮して、我々は以下のようないくつかの翻訳処理方式を用意した。

### (1) 表テンプレート方式

表形式記事を翻訳するためのもの。表ごとに英日対

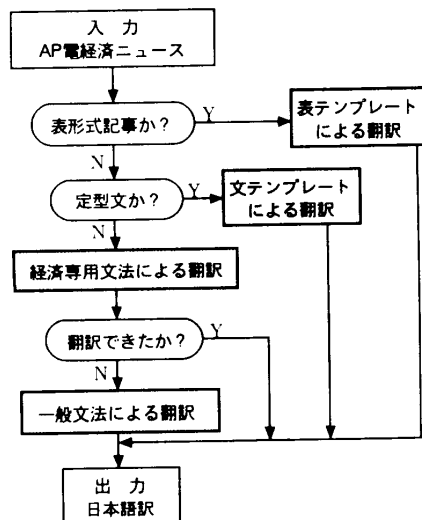


図1 外電経済ニュースの翻訳方式

Fig. 1 Proposed method of translating economic news.

訳テンプレートを作成しておき、変化部分（主として数値）を実際の記事に従って変えてやる。

#### (2) 文テンプレート方式

各定型文に対して英日対訳テンプレートを作成しておき、定型文記事のすべてと、一般文記事中の定型文の翻訳に適用する。ただし、一般文記事中の定型文としてはかなり拡張したものを考えた（4.3節参照）。

#### (3) 経済専用文法方式

一般文記事中の非定型文の翻訳に適用する。英語解析文法の作成や日本語訳語の絞り込みのために、一般文記事に見られる定型表現パターンを間接的に活用する。

#### (4) 一般文法方式

一般文記事中の非定型文で、経済専用文法では解析できなかったもの（たとえば、会話の中味等）については、一般文法によって再度解析し翻訳する。

以上の4方式を図1のように組み合わせて翻訳を実行する。すなわち、

- ヘッダーにより記事のタイプを決定する。
- 表形式記事に対しては、対応する英日対訳テンプレートを用いて翻訳する。（表テンプレート方式）
- 定型文記事の全文および一般文記事中の定型文に対しては、対応する英日対訳テンプレートを用いて翻訳する。（文テンプレート方式）
- 一般文記事中の非定型文に対しては、経済専用文法を用いて翻訳する。（経済専用文法方式）  
翻訳処理が成功すれば終了。原文解析が失敗したらその段階でe)へ。
- 一般文法を用いて再翻訳する。（一般文法方式）

## 4. 定型パターンの抽出とその利用

経済ニュースに現れる各種の定型パターンは、文テンプレート方式や経済専用文法方式において重要な役割を果たす。ここでは、それらパターンの原テキストからの抽出と、経済専用文法および翻訳テンプレート作成への利用について述べる。

### 4.1 定型パターンの抽出

AP電経済ニュースの一般文記事1年分（1989年4月～1990年3月）から、次の条件を満たす単語列を定型パターンとして機械的に抽出した<sup>9)</sup>。

- 1) 数字、年月日、曜日等の違いは無視
- 2) 出現頻度が4以上
- 3) 語長が2～44の範囲

その結果、約2,000のパターンが抽出された。いくつかの例を示すと、

“the price of the benchmark No.#-year Japanese government bonds”

“Silver traded in London at a bid price of # dollars a troy ounce”

“(# GMT), the Financial Times-Stock Exchange #-share index was”

“# billion Hong Kong dollars (# million U.S. dollars)”

“, with # up, # down and # unchanged”

ここで#は数字列を表す。これらは意味的にまとまったパターンであるが、機械的な抽出のため、下例の下線部に示すように、意味的にまとまりのない奇妙なパターンも得られることがある。

“#-# cents higher, with March at # dollars a bushel”

“, which shed # points Thursday, lost # points,”

“For instance, he noted, the dollar had traded about # West German marks”

これらについては、原テキストを参照しながら、前後に語を補ったり削除したりして整形を試みた。その結果、最終的に約1,700の定型パターンが得られた。

### 4.2 経済専用文法の構成

このようにして得られた整形済の定型パターンに対して、次に文法カテゴリーと日本語訳を与えた<sup>10),11)</sup>。使用したカテゴリーは次の16種類である。

1. 文
2. 文から主語の抜けたもの
3. 文から目的語の抜けたもの
4. 文から補語の抜けたもの
5. 名詞句

6. 前置詞句
7. 形容詞句
8. 副詞句

9. 1~8に前置詞の付いたもの

各パターンに付与したカテゴリー番号は1から8までのいずれかである。9にあたる前置詞付きのパターンには、1から8のいずれかの番号を与え、前置詞付きであるというマークを別に付けた。文法カテゴリーを付与している途中で、構文・意味上区別したほうがよいと判断したものに対してもマークをふった。

各文法カテゴリーの数は次のようになった。

1. 441 2. 215 3. 56 4. 34
5. 413 6. 208 4. 14 8. 123

文と名詞句がそれぞれ400パターン余り、続いて文から主語の抜けたもの、前置詞句、副詞句の順となっている。

同じカテゴリーでも、文を構成する際に異なる振舞いをするものに対してはマークを付与した。たとえば、前置詞句のなかで、大文字で始まるものは「文頭に置かれる」ことを意味するマークをふった。その他に与えたマークの種類には、独立文、結果のto、挿入、開始、引用導入、等がある。

定型パターンに対する日本語訳は、経済ニュース文であることから、専門用語を正しく訳出すること、簡潔な表現を用いることを心がけた。前後関係がはっきりしないと訳語がふれないパターンもあり、そのような場合には原テキストにあたって再度パターンを整形し直し訳語を付与した。たとえば、

原パターン：“¥¥¥ dollars to ¥¥¥”

整形パターン：“rose ¥¥¥ dollars to ¥¥¥”

訳語：[数] ドル上げて [数] になった

このような文法カテゴリー付き対訳パターンを取り入れながら経済記事解析用文法規則を作成した。規則作成にあたって参照した記事は、定型パターンの抽出に用いたうちの89年8月21~25日、89年8月28日~9月1日、89年11月20~25日、90年2月5~10日のものである。その後、91年4月15~30日の記事によって見直しを行ったが、新構文はなく主として経済用語の補充で足りた。

文法規則を作成に際しては、経済記事に類出する構文に重点を置き、文法規則数を抑え、解析の曖昧性を少なくするよう努めた。そのため、解析可能構文を以下のように限定した。これらに該当しない構文は一般文法で解析されることになる。

- 1) 動詞の用法別の、いわゆる5文型の基本構文、分詞句、不定詞句、that節も含む。

2) 並列は、名詞句、前置詞句、動詞句、文など、同じ要素のものを扱う。

3) 従属接続詞による複文構造。省略を伴う比較構文は扱わない。

4) 関係節は、主語関係節のみを文法化。関係詞の省略、目的語関係節は扱わない。

5) かっこによる挿入、文末の同格の構文を文法化。文中のカンマによる挿入、同格文は扱わない。

この結果、規則数は一般文法の約5分の1(約500)と小さくなり、修正・改良の行いやすい、見通しのよいものになった。

辞書としては、名詞を中心とする経済専門用語のほか、対訳パターンのうち変数を持たないもの約400も含めて約3万語から成るものを作成した。これら対訳パターンには次のような長く複雑なものも含まれている。

“In Tokyo, where trading ends before Europe's business day begins,”

“against all major foreign currencies except the British pound”

経済専用文法がコンパクトになった理由は、一義的には対象が経済分野に限定されたことにあるが、このような対訳パターン(約400)や次節で述べる翻訳テンプレート(約400)によって約800の文や句に対する解析規則が不要になった点もあげられる。

#### 4.3 翻訳テンプレートの作成

表形式記事と定型文記事に対しては、記事全体を対象として、いくつかの変数を含む対訳テンプレートを用意した。

一般文記事に対しても、1文を構成する80%以上の語が、4.1節で抽出した定型パターンによって重複なく被覆される文を定型文と定義し、その抽出とそれらに対する対訳テンプレートの作成を試みた。その結果、約8千の定型文が抽出されたが、そのうち、出現頻度に関して上位約400の文に対して対訳テンプレートを作成した。

定型文抽出は自動化でき、さらに対訳テンプレートの作成も、抽出された定型文に対する日本語訳を与えるだけで、DPマッチング手法により自動化できた。これらの詳細は文献12)に譲る。

特徴的なことは、この場合の対訳テンプレートを、対訳付きの単純なCFGルールの形にまとめたことである。これにより、必ずしも定型パターンとはとらえられていない固有名詞の変化や、“up”に対する“down”等への対応が、辞書(局所解析用辞書と呼ぶ)の拡張だけで行える利点が生まれ、結果的に対訳テンプレ-

トの適応範囲を高めることができる。たとえば、

“In Kuala Lumpur, Malaysian tin closed at 17.76 dollars per kilo, up 5 cents.”

に対して（抽出済の定型パターンのみを使って）対訳テンプレートを作成すると，“Kuala Lumpur”も“up”も固定項となってしまうが、あらかじめ局所解析用辞書に、

CITY → “Kuala Lumpur”, “Tokyo”, ……

UPDW → “up”, “down”

のような辞書項目を追加しておくことで、この文における都市名や語“up”の変化にも適応可能な、柔軟な対訳テンプレートが作成できる。

以上、本章では定型パターンを利用した機械翻訳システムの対象分野へのチューニングの方法を具体的に述べた。このうち、定型パターンの抽出と定型文の抽出は自動化されている。対訳テンプレートも、抽出された定型文（ここでは約400）に対する日本語訳を与えることで自動的に作成できる。人手に頼らざるをえないのが専用文法の作成であるが、文法記述の経験を持つ一人の専門家がそれにあたり、定型パターンのカテゴリー付けと日本語訳の付与に延べ約3カ月（実働約14分/パターン）、文法規則の作成にやはり延べ約3カ月（実働約55分/規則）を要した。チューニングのコストは主にここにかかっている。これに対して翻訳精度の向上はどの程度になったかを次章で述べる。

### 5. 翻訳実験とその評価

翻訳精度を評価するために、91年4月23～25日のAP電3日分の経済記事をサンプルとして翻訳実験を行った。ただし表形式記事は文章部分が短く、表テンプレート方式の成功率も100%であることが容易に観察されたので、評価対象から除いた。その結果、このサンプルには短いもので3語、長いもので45語の長短472文が含まれている。1文の平均の長さは、AP電全記事の1年分（1989年4月～1990年3月）の平均値の約23.7語に近いと思われる。

各翻訳処理方式の処理速度の目安と諸元は以下に示すとおりである。

☆：文テンプレート方式 …… 5秒/文

対訳テンプレート数： 約400  
局所解析用辞書語数： 約4,000\*

\* このうち約6割が人名と地名である。

表1 一般文法のみによる翻訳の評価

Table 1 Evaluation of translation using only the general grammar.

評価	5	4	3	2	1	翻訳率
翻訳文数	2	92	82	198	98	19.9%

表2 3方式併用による翻訳の評価

Table 2 Evaluation of translation by integrating the propose three methods.

評価	5	4	3	2	1	翻訳率
翻訳文数	209	139	62	52	10	73.7%

表3 3方式別の翻訳評価

Table 3 Evaluation of translation of each of three methods.

評価	5	4	3	2	1	翻訳率
方式☆	147	0	0	0	0	100.0%
方式◎	62	136	59	27	2	69.2%
方式★	0	3	3	25	8	7.7%

◎：経済専用文法方式 …… 10秒/文

書き換え規則数： 約500

基本語数： 約57,000

経済専門語数： 約31,000

★：一般文法方式 …… 20秒/文

書き換え規則数： 約2,500

基本語数： 約57,000

専門語数： 約111,000

翻訳評価の5ランクと翻訳率を次のように定義する。  
評価5：入力文の文意は出力文に忠実に再現されている。

評価4：文意は再現されているが細部に問題がある。

評価3：文意はほぼ再現されているが誤解を受けやすい部分がある。

評価2：文意はほとんど再現されていない。

評価1：文意はまったく再現されていない。

翻訳率 = (評価5の文数+評価4の文数)/全文数  
このとき、各方式の翻訳精度は表1～表3のように評価できる。

表1は、全文を一般文法方式（ニュース用包括システム）のみで翻訳した場合の評価である。これは政治、社会、科学等のニュース全般の平均値に比べてきわめて低い値になっている。

表2は、3種の方式を切り替えて適用した場合の評価である。翻訳率は約50%改善されている。

表4 非学習サンプルによる翻訳の評価  
Table 4 Evaluation of translation for non-trained samples.

評価	5	4	3	2	1	翻訳率
全体	67	39	33	12	9	63.3%
方式☆	50	0	0	0	0	100.0%
方式◎	16	38	31	7	0	58.7%
方式★	1	1	2	5	9	22.2%

3方式の適用割合は、

- ☆: 文テンプレート方式 …… 31.1%
- ◎: 経済専用文法方式 …… 60.6%
- ★: 一般文法方式 …… 8.3%

であるが、これら方式別の評価は表3に示すとおりである。翻訳率の向上には、文テンプレート方式と経済専用文法方式が大きく貢献していることが分かる。一般文法方式へ渡されるのは、それらで翻訳できなかった難文(長文が多い)だけなので、その翻訳率は表1の結果をさらに下回ることになる。経済専用文法方式では評価3になったものが約20%あるが、これらに評価4に上げることで翻訳率をかなり向上させることができる。

以上の評価実験に用いたサンプルは、文テンプレートや経済専用文法の見直しに用いた、いわば学習サンプルともいえるものであったが、まったく無関係の非学習サンプル(91年3月7日の160文)による評価実験も行ってみた。

その結果を表4に示す。翻訳率は全体として約6%低下している。その主原因は経済専用文法方式の翻訳率の低下にあるが、これは評価4から評価3への移行による。3方式の適用割合は学習サンプルの場合と同程度であったから、この経済専用文法は、大局的には経済記事を広くカバーしているものの、評価4と評価3の割合に関して学習サンプルと非学習サンプルの間で差があることが分かる。いずれにしても、評価3となった翻訳結果を分析して、それを評価4に上げる方策を探ることが、翻訳率のさらなる向上にとって有望と考えられる。以下に評価3となった翻訳例をコメントとともに示す。

But the dollar remained in a narrow range in Tokyo in the absence of fresh market-moving news, dealers said.

◎ただドルはフレッシュな market-moving ニュースがないので東京で狭い範囲の中に残ったと、ディーラーが言った。

⇒ 未知語 market-moving が改善されれば評価4になる。用語の問題。

Tokyo share prices rebounded in thin trading Tuesday after a three-day decline, while the U.S. dollar slipped against the Japanese yen after a three-day rise.

◎米ドルが、日本円に対して下落した後3日が上昇する一方で、3日が下落したあと、東京株価は、火曜日に薄い取引で反発した。

⇒ 「3日が下落」/three-day decline, 「3日が上昇」/three-day rise が改善されれば評価4になる。用語の問題。

Advancing issues led declines by a ratio of about 12-to-5, with 771 up, 320 down, and 444 unchanged.

◎上げ銘柄は771上げて、そして320下げて、そして444変わらずだっておよそ12-to-5の比率で下落にまさった。

⇒ 列挙された数表現処理の問題。12-to-5のような数表現は“-to-”の形に着目すれば処理可能になろう。「444 変わらずだって」/444 unchanged に対しては日本語生成の改善が必要。

これら翻訳例からも分かるように、評価3のかなりのものは、経済用語の充実と数表現処理の強化によって評価4に上げられる可能性を含んでいる。

付録にさらに翻訳例を示した。

## 6. おわりに

AP 電経済ニュース記事にチューニングした英日機械翻訳システムをインプリメントした。システムは4つの翻訳方式を統合したもので、翻訳の実行にあたっては、原文の表現の定型性の程度に応じて4方式を切り替える仕組みになっている。

翻訳実験の結果から、次のようなことが結論できよう。

- 1) 経済記事には定型表現が多く、文テンプレート方式がきわめて有効に働く。
- 2) 分野を限定することで文法がコンパクトになる。実際、我々が構成した経済記事専用文法は一般文法に比べて約5分の1の約500の規則数ですんだ。これには約800の定型表現に対する解析規則が不要になったことも効いている。
- 3) チューニングのコストは主に専用文法作成にあるが、それに見合うだけの翻訳精度向上はあったといえよう。

4) したがって、広範な内容を持つニュースの機械翻訳にあたっては、ニュース分野を分類し、分野ごとに特化したシステムで対応することが有望かつ現実的と思われる。

今後は、経済専用文法方式による翻訳結果に評価3のものが多かったため、それらを評価4に上げるための分析検討を行う必要がある。部分的には、全章末尾にも示唆したように、比較的簡単に行えるように思われる。

ほかのニュース分野への特化を試みることも、ニュース分野別システム化の有効性を確認するうえで大切であろう。たとえばスポーツ記事においては、試合のスコアに関する記事と「ヒーローインタビュー」式のトピックス記事とが混在しているのが特徴であるが（およそ1対2の割合）前者の定型性は高い。後者については、定型性はそれほど高くないものの内容が一段と限定されるのでコンパクトな専用文法の構築が期待できる。

このような試みを反復することによって、適度に限定された対象に対する専用文法作成のノウハウのようなものが蓄積できれば、それも大きな成果になろう。

## 謝 辞

本研究を進めるにあたって種々ご討論いただいた先端制作技術研究部の自動翻訳研究グループに深謝する。また、元の表1と2にあった集計ミスをご指摘いただいたことに対し、査読者に感謝します。

## 参 考 文 献

- 1) 相沢, 浦谷, 田中: 放送ニュースへの機械翻訳システムの適用, 電子情報通信学会言語理解とコミュニケーション研究会資料, NLC91-20 (1991).
- 2) 相沢, 加藤, 鎌田: 外電経済ニュース用英日機械翻訳システム, 電子情報通信学会言語理解とコミュニケーション研究会資料, NLC92-39 (1992).
- 3) 江原, 浦谷, 田中, 畑田, 加藤, 金, 住吉: 放送用英日機械翻訳システムとその精度評価, NHK技研研究報告書 (1995).
- 4) Brown, P.F., Cocke, J., Pietra, S.A.D., Pietra, V.J.D., Jelinek, F., Lafferty, J.D., Mercer, R.L. and Roossin, P.S.: A Statistical Approach to Machine Translation, *Computational Linguistics*, Vol.16, No.2, pp.79-85 (1990).
- 5) 佐藤: MBT2: 実例に基づく翻訳における複数翻訳例の組合せ利用, 人工知能学会誌, Vol.6, No.6, pp.861-871 (1991).
- 6) Sumita, E. and Iida, H.: Example-Based Transfer of Japanese Adnominal Particles into English, *IEICE Trans. Inf. & Syst.*, Vol.E-75-D,

No.4, pp.585-594 (1992).

- 7) 古瀬, 飯田: 変換と解析の協調的処理による翻訳手法 - 変換主導型翻訳手法 -, 情処研報, NL84-4 (1992).
- 8) Kaji, H., Kida, Y. and Morimoto, Y.: Learning Translation Templates from Bilingual Text, *Proc. 14th International Conference on Computational Linguistics (COLING'92)*, Nantes, France, pp.672-678 (1992).
- 9) 浦谷, 加藤, 相沢: AP 電経済ニュースからの定型パターンの抽出, 第42回情報処理学会全国大会論文集, 6E-4 (1991).
- 10) 相沢, 浦谷, 鎌田: 外電経済ニュース英文の定型パターン, 第42回情報処理学会全国大会論文集, 2H-1 (1991).
- 11) 相沢, 鎌田: AP 電経済ニュースの英語解析用文法, 第45回情報処理学会全国大会論文集, 3E-6 (1992).
- 12) 加藤: 定型パターンを含む文の機械翻訳手法, 情報処理学会論文誌, Vol.36, No.9, pp.2081-2090 (1995).

## 付 録

### 1. 文テンプレート方式による翻訳例

Singapore rubber fetched 142.50 cents a kilo at noon Tuesday, up 0.25 cents.

☆火曜日正午, シンガポールゴムは0.25セント上昇し, 1キロ142.50セントとなった。[評価5]

In Kuala Lumpur, Malaysian tin closed at 14.93 dollars per kilo, up 3 cents.

☆クアラルンプールでは, マレーシア錫は3セント上昇し, 1キロ14.93で引けた。[評価5]

Gold in Hong Kong rose the equivalent of 0.96 U.S. dollars an ounce on Tuesday, to close at 358.20 U.S. dollars, compared to Monday's 357.24.

☆香港の金は火曜日, 1オンス0.96米ドル相当上昇し, 月曜日の357.24に対し, 358.20米ドルで引けた。[評価5]

Silver traded in London at a bid price of 3.95 dollars a troy ounce, from Monday's 3.96 dollars.

☆ロンドンで銀は, 月曜日の3.96ドルから, 1トロイオンス3.95ドルの買値で取引された。[評価5]

### 2. 経済専用文法方式による翻訳例

比較のため, 一般文法による翻訳結果も示す。

◎: 経済専用文法による翻訳

★: 一般文法

The U.S. dollar opened slightly higher against the Japanese yen Tuesday morning in Tokyo, while share prices inched up.

◎株価がじり高になった一方で、米ドルは、東京で火曜日午前日本円に対して小高く寄り付いた。[評価 5]

★少しずつ動かれた株価の間アメリカ・ドルは、東京で日本の円火曜日の朝に反対してわずかにいっそう高く開いた。[評価 2]

The dollar started trading at 138.84 yen, up 0.14 yen from Monday's 138.70-yen close.

◎ドルは月曜日の 138.70-yen 引けからの 0.14 円高で 138.84 円で取引を開始した。[評価 5]

★月曜日の 138.70-yen からの上りの 0.14 円が、138.84 円におけるドル始まる取引を閉じる。[評価 2]

The Financial Times-Stock Exchange 100-share index closed 13 points, or 0.5 percent, higher at 2,503.8.

◎フィナンシャルタイムズ 100 種株価指数は 2,503.8 で 13 ポイントすなわち 0.5 パーセント高であった。[評価 5]

★フィナンシャルタイムズ 100 種株価指数は、2,503.8 でいっそう高い 13 ポイントすなわち、0.5 パーセントを閉じた。[評価 3]

He said the dollar was moving in a narrow range in the absence of fresh market-affecting news.

◎ドルは市場に影響を与える新材料がないので小動きであったと、彼が言った。[評価 5]

★ドルが市場に影響を与える新材料がない時狭い範囲で感動的だったと彼は言った。[評価 3]

Benchmark North Sea Brent crude oil for June delivery was traded at 19.50-19.60 dollars a barrel at midmorning, slightly higher than its closing level in New York.

◎6月引き渡しのための指標銘柄北海原油は、ニューヨークのその前夜の引けているレベルより小高く

全場中頃でバレル 19.50-19.60 なドルで取引した。[評価 4]

★6月配達のための水準点北部 Sea Brent 原油が、わずかにニューヨークでその夜通しの最後のレベルよりいっそう高い後場中頃でバレル 19.50-19.60 ドルで交換された。[評価 2]

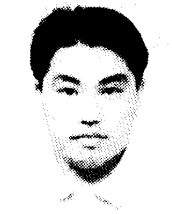
(平成 7 年 12 月 18 日受付)

(平成 8 年 3 月 12 日採録)



相沢 輝昭 (正会員)

1940 年生。1963 年京都大学工学部電気工学科卒業。同年日本放送協会入局。放送技術研究所においてオートマトン、カナ漢字変換、静止画像検索、放送ニュース英日機械翻訳の研究等に従事。1986~89 年 (株) ATR 自動翻訳研究所に出向。1995 年 4 月より広島市立大学情報科学部教授。電子情報通信学会、言語処理学会、ACL 各会員。



加藤 直人 (正会員)

1962 年生。1986 年早稲田大学理工学部電気工学科卒業。1988 年同大学大学院理工学研究科博士前期課程修了。同年日本放送協会入局。放送技術研究所にて機械翻訳の研究に従事。1994 年 (株) ATR 音声翻訳通信研究所に出向。音声言語処理、対話処理の研究に従事。電子情報通信学会会員。



鎌田 雅子 (正会員)

1954 年生。1976 年大阪大学文学部卒業 (仏文学専攻)。京都大学長尾研究室にて言語分析に従事後、科学技術庁機械翻訳 (Mu) プロジェクトで自動翻訳システムの開発、(株) ATR 自動翻訳電話研究所で、話者の発話意図を考慮した話し言語翻訳の研究に携わる。現在、漢字情報サービス株式会社勤務。ニュース文を中心とした言語処理研究に従事。言語処理学会会員。