

## 市況速報記事に対するテンプレート型日英翻訳の効果

5 J-6

白井 論\*1 細野明美\*2 野沢弥生\*2 木村知子\*2 阿部さつき\*2 内野 一\*1

\*1NTTコミュニケーション科学研究所

\*2NTTアドバンステクノロジー(株)

## 1 はじめに

経済記事は類似の表現が繰り返し使用されるため、用例系の機械翻訳の適用が有利と考えられる。特に定型性が高い場合は、あらかじめ翻訳テンプレートを作っておくことにより、翻訳処理の効率化・高精度化を図ることが考えられる。例えば、英語の経済ニュースの場合、3割の文が100%正しく高速に翻訳されることが報告されている[加藤 95]。

日本語の市況速報記事も定型性が高いため、同様の考え方が適用できる。テンプレート方式では、テンプレートの適用を厳密にすると表現の揺らぎに弱くなり、逆に柔軟性を持たせると誤適用の恐れが生じるという問題がある。

本稿では、翻訳テンプレートを用いる日英翻訳システムの構成を示すとともに、翻訳テンプレートの概要とヒット率を向上させるための改良項目について述べる。また、3種類の市況速報記事の数カ月分に対する適用実験の結果を報告する。

## 2 市況速報記事の形式

本稿では、市況速報記事14週分(1995年6月~9月)のうち、東証外国部(1日1記事)、大証(1日4記事)、東証CB(1日3記事)を対象とした。これらの記事は日本経済新聞社のテレコンデータベースから取り出し、[高橋 97]の方法により日英記事を対応付けた。記事の例を図1に示す。

市況速報記事を数カ月分を通してみると次のような特徴があることがわかる。

3種類の記事とも日本語・英語とも構成は同じで、見出しの後、市場の総括が1文、概況や背景が1~3文、個別銘柄の様子が2~3文書かれている。

日本語記事では、市場の総括には類似の表現が頻

(1995年7月6日)

◇大証大引け・ハイテクなど買われ3日続伸

後場寄りあとにかけて一時伸び悩む場面があったが、大引けにかけてハイテク株や関西系の小型材料株中心に買い進まれ、大証修正は大幅続伸。村田製、ローム、任天堂が一段高。オムロン、森精機、青山商、住友林、きんでん、ダイダグ、参天薬も堅調。ホシデンはストップ高。兼松日産農、ルシアン、住友シチック、堺化学、紀州紙、松村組、明星工も人気付く。半面、日精化、サンスター、ロート、キムラタンは安い。

Osaka Stocks Cls: Surge to close higher

Stocks ended sharply higher Thursday. The morning advance stalled in early afternoon trading. But high-tech issues and small-cap, incentive-backed stocks based in Osaka drew active buying just ahead of the market close. The Adjusted Stock Price Average gained 264.55 points to finish at 16,485.18. Among the many gainers were Murata Mfg., Omron and Sumitomo Forestry. Nippon Fine Chemical, Sunstar and Rohto Pharmaceutical slipped.

(1995年7月7日)

◇大証大引け・急騰し1万7000円台回復

大証修正は前日比944円高と昨年1月31日(1255円高)以来の上幅を記録し、1カ月ぶりに1万7000円台を回復。日米同時の金融緩和が素直に好感され、主力のハイテク株から材料株までほぼ全面高。ローム、村田製が買われ、ホシデンは連日のストップ高。富士通ゼ、ルシアン、ダイワボウ、住友シチック、松村組、中外炉がにぎわい、小野薬、オートバックス、島精機、シマノ、青山商も上げた。半面、オムロン、堺化学は一服。

Osaka Stocks Cls: Soar to 17,000 level

Stocks on the OSE gained steadily throughout Friday, climbing above the 17,000 level for the first time in month. The easing of U.S. and Japanese monetary policies stimulated the buying of a wide variety of stocks, from high-tech issues to small-cap issues with specific incentives. The Adjusted Stock Price Average closed the day at 17,430.01 points, up 944.83. Rohm and Murata Mfg. were bought. Hosiden went limit-up. Meanwhile, Omron and Sakai Chemical Industry fell.

図1 日英の市況速報記事の例

出し、個別銘柄の様子は極めて定型的であり、共にテンプレート化が期待される。これに対して、概況や背景は文が長く表現も変化に富むため、テンプレート化は難しい。また、類似表現には助詞の省略や揺らぎ、繰り返しや語尾の違いなどがあり、これらに対応することにより適用性を高めることができる。一方、英語記事には日本語記事にはない情報の付加や表現のバリエーションの多さが目立つ。機械翻訳では不足情報の補充は基本的に不可能である。

また、日英の記事は直訳的には対応していない。しかし、意識的なものを含め日本文と対応しそうな英文を集めて眺めてみると、標準的な訳し方が推定できることがわかった。そこで、当面は英語表現のバリエーションへの対応を諦め、標準訳を決定した後、それを翻訳テンプレート化した。

Effects of Template-based Japanese-to-English Machine Translation for Stock Market Reports  
Satoshi SHIRAI\*1, Akemi HOSONO\*2, Yayoi NOZAWA\*2, Tomoko KIMURA\*2, Satsuki ABE\*2 and Hajime UCHINO\*1

\*1NTT Communication Science Laboratories and \*2NTT Advanced Technology Corporation

