

「翻訳部品の配置による翻訳作業」を目指した 翻訳統合環境の提案

熊野 正

田中 英輝

E-mail: {kumano.t-eq, tanaka.h-ja}@nhk.or.jp

NHK 放送技術研究所

〒157-8510 東京都世田谷区砧1-10-11

概要

我々は従来より、人間の翻訳過程は、翻訳元文書中の各部分がどのように翻訳されるかを検討する段階と、それらを取捨選択し翻訳先文書を構成する段階とに大別され、前者のコンピュータによる支援によって、翻訳者の創意を妨げずに翻訳の時間的効率や品質の向上を図ることができると考えてきた。我々が先に提案した「翻訳パレット」は、この前者（「翻訳部品」準備作業）を、コンピュータと人間が協調して行うためのツールであったが、本稿では、ここで準備した翻訳部品の使用に焦点を当て、ユーザが翻訳部品を取捨選択し配列することで翻訳先文書を構成することができるエディタを統合した、新しい翻訳統合環境を提案する。

1. はじめに

人間の翻訳者が行う翻訳には、多かれ少なかれ「意識」が含まれる。意識とはすなわち、翻訳元文書を出発点として、翻訳者が「何を書くか」を（ある意味元文書を逸脱して）「決定」した結果である。翻訳者がどのような内容をどのように翻訳先文書として表現するかは自由であり、自らのセンスを働かせ、自らの責任において判断する。

一方、執筆文書が翻訳である以上、いくら表現方法は自由であるといっても、当然のことながらその文書の内容は翻訳元文書に支配されている。例えばニュース記事の翻訳において、「誰がどこで何をした」という事実は、この事実の一部もしくは全部を省略したり、大幅な補足を加えたりして

翻訳されることはあっても、『誰』『どこ』『何』をまったく別のものにしてしまうことはないであろう。そして『誰』『どこ』の翻訳にはそれほど創意工夫の介在の余地はない。

このように、人間の翻訳作業は、翻訳元文書中の各部分をどのように翻訳するかという、翻訳者や意識の度合の違いによってそれほどぶれがない問題と、それらを取捨選択し翻訳先文書を構成するという、各人の技芸に属する問題とに大別できると考えられる。

我々は先に、この2つの次元の作業を分離する翻訳作業の進め方を提唱し、放送ニュース翻訳者向けに、前者の作業、すなわち翻訳元文書中の各部分の翻訳である「翻訳部品」の準備をコンピュータによって支援する翻訳支援ツール「翻訳パレット」を提案した [3]。ユーザである翻訳者は翻訳先文書の執筆に先立って翻訳パレットを利用する。翻訳パレットは翻訳元文書を解析して自動的に調査を行い、翻訳部品を提案する（ユーザの指示に応じて調査を行い翻訳部品を準備することもできる）。この準備作業の後、ユーザは準備された翻訳部品を適宜使いながら翻訳先文書を執筆する。システムによる翻訳部品の提案を採用するかどうかは完全にユーザの自由であるから、システムはユーザが翻訳先文書の内容を決定する自由を妨げることがない。ユーザは翻訳にかけられる時間が少ない時には提案を積極的に導入し、時間に余裕がある時には提案内容をより吟味するなど、システムと柔軟な協調を行うことができ、結果として、システムが翻訳作業の時間的効率化や翻訳品質の向上に寄与することが期待できると考える。

先に提案したツールは、もっぱら翻訳部品の準備に着目して実装したため、準備された翻訳部品

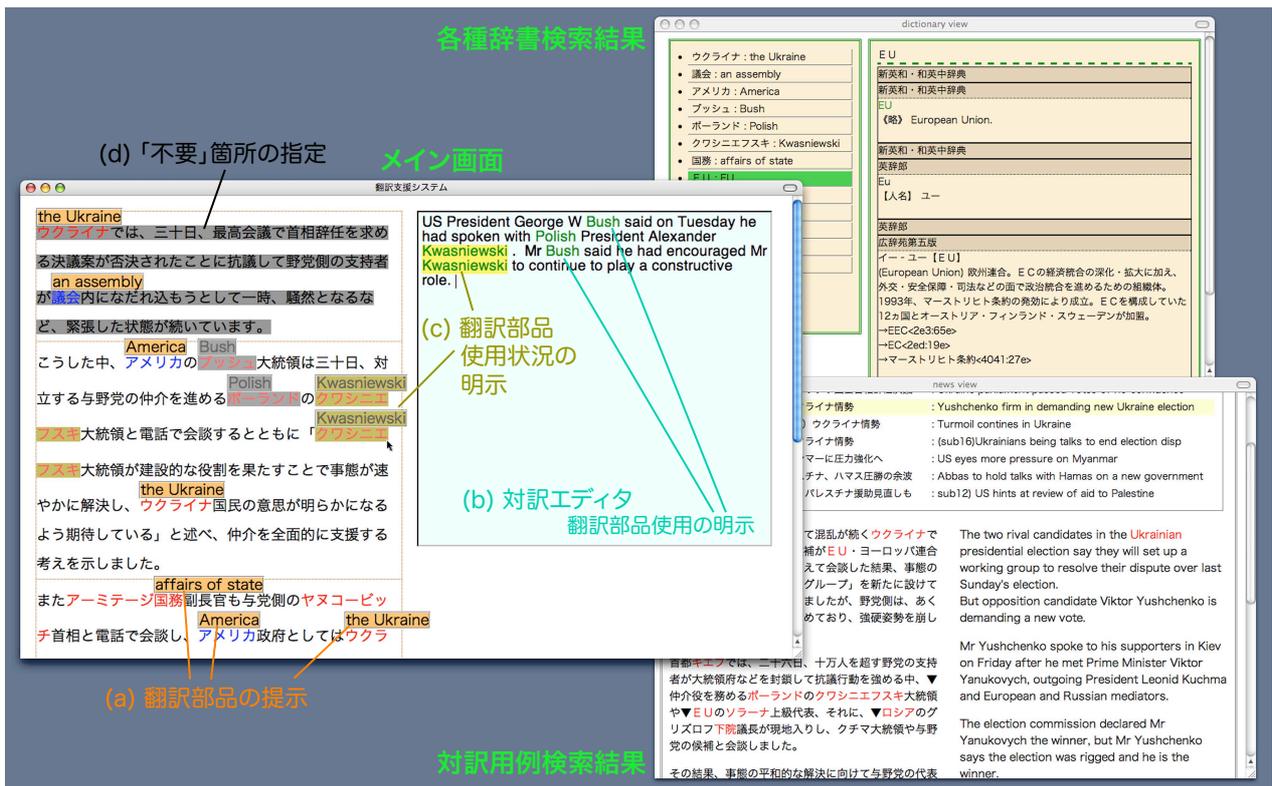


図1 統合翻訳環境・新「翻訳パレット」プロトタイプ

を使用して翻訳先文書を執筆することを支援する観点に欠けていた。そこで本稿では、ユーザが翻訳部品を適宜選択し配列することで翻訳先文書の執筆を進めることができる翻訳エディタを統合した翻訳統合環境・新「翻訳パレット」を提案する。本システム上で、ユーザは選択・使用した翻訳部品の使用状況やその他翻訳元文書中の内容の取捨を容易に把握することができ、ユーザが翻訳先文書の内容を決定していくのに役立てることができる。本稿では、今回実装した新「翻訳パレット」プロトタイプの機能を説明し、今後の展望を述べる。

2. 翻訳エディタを統合した統合翻訳環境・新「翻訳パレット」

本稿で提案する統合翻訳環境・新「翻訳パレット」(以後、単に新「翻訳パレット」と呼ぶ)のプロトタイプ画面を図1に示す。

新「翻訳パレット」は、先に提案した「翻訳パレット」(以後、単に旧「翻訳パレット」と呼ぶ)と同様に、翻訳元文書を読み込むとユーザに先立つ

て以下の処理を行い、結果を「翻訳部品」としてユーザに提示する¹。

1. 日本語文書中から調査したい表現を発見
(現状は固有表現抽出器を用い、めずらしい固有表現を調査対象に認定)
2. 1.をさまざまな情報資源(電子辞書、対訳用例、Web、etc.)から検索
(現状は各種電子辞書、NHK日英対訳ニュースデータベース、GoogleによるWeb検索が対象)
3. 2.の結果中から有用な英語表現を抽出
(現状は各種電子辞書からのみ対訳表現を抽出)

上記の処理の具体的内容については、旧「翻訳パレット」に関する文献[3]を参考にされたい。

以後では、新「翻訳パレット」で追加・変更された各機能を説明する。

¹ 以後日英翻訳の支援環境を例に説明するが、手法は言語に依存しない。

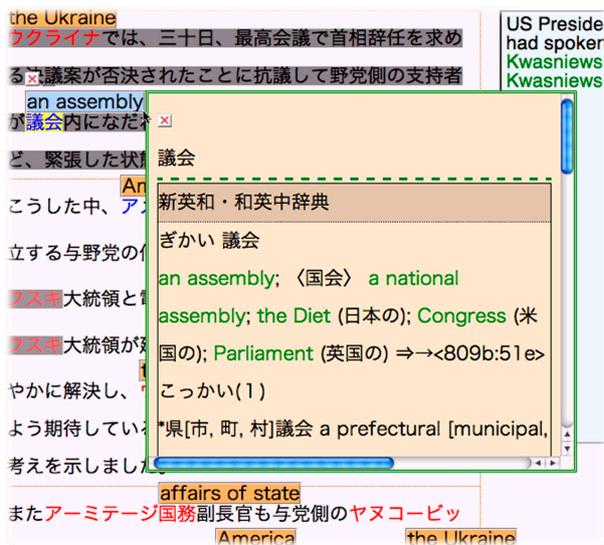


図 2 翻訳部品の選択肢一覧

2.1 翻訳部品の提示方法

システムが自動的に提案する翻訳部品、および、ユーザの指示に応じて調査を行い準備した翻訳部品の提示は、旧「翻訳パレット」では一括してパレットエリアに表示されていたのとは異なり、翻訳部品が対応する日本語文書中の各表現上に結びつけて表示される (図 1(a))。これは、ユーザが使用したい翻訳部品を探し出しやすくするため、そして、翻訳部品の使用状況 (2.4 節参照) を対応する日本語文書中の表現の出現と結びつけて把握しやすくするためである。

一方、直接日本語文書中の表現と結びつかない翻訳部品をユーザが作成することもできる。これらはウィンドウ上の任意の場所に配置することができ、自由に動かすことができる。

2.2 提案された翻訳部品の吟味

ユーザは提案された各翻訳部品に対して、その根拠 (現状では表現の出自である辞書エントリ) を参照することができる。また、辞書エントリから他の候補となる訳語が得られる場合やその他の辞書等の情報源からも候補を得られる場合、それらの選択肢を一覧してより気に入ったものを選択することができる (図 2)。旧「翻訳パレット」においては、システムが提案した翻訳部品に対し、他の選択肢を参照することはできなかった。

2.3 翻訳エディタの導入

ユーザは、日本語文書に併置されている翻訳エディタ (図 1(b)) を用いて英語文書を執筆する。ユーザは表示されている任意の翻訳部品をマウスでドラッグすることで、執筆中の英語文書中に導入することができる。導入された各翻訳部品は色によって翻訳部品であることが明示され、また各翻訳部品はマウスでドラッグすることで、英語文書中の任意の位置に動かしたり、文書外に取り出したりすることができる (2.4 節参照)。

2.4 翻訳部品の使用状況等の把握

表示されている各翻訳部品は、翻訳エディタ中に導入することで「使用済」となる。使用済の翻訳部品は対応する日本語表現とともに背景が灰色となって使用済であることを明示する。また、日本語文書中の使用済翻訳部品および英語文書に導入された翻訳部品上にマウスを重ねると、どの翻訳部品が英語文書中のどこに使用されているかの対応が色で把握できるようになっている (図 1(c))。一方、使用中の翻訳部品を文書外に取り出すと、翻訳部品の使用状況は「未使用」に戻り、灰色の使用済表示が解除される。

さらに、日本語文書中のある範囲の表現をユーザが選択し、まとめて「翻訳済」もしくは「不要」として灰色に変更する機能も実装した (図 1(d))。

これらの機能は、ユーザが英語文書の内容や構成を考える際に、日本語文書のどの部分を取捨選択したかを把握するのに役立つことが期待される。

3. 考察

影浦らの調査によると、翻訳者が翻訳時に用いる情報参照ツールに求めるのは、「構造」が似通った表現の翻訳に関する情報ではなく、具体的な入力表現そのものの事例であり、かつ翻訳者が調べたいと思う表現の単位は語～熟語・慣用句～引用句・連語程度にとどまるという [1]。これはすなわち、翻訳者は、語や慣用句のような比較的小さな部分の翻訳の決定にあたってはさまざまな情報を参照したいが、それを越えた文章の組み立てレベルの決定にあたっては他人の翻訳例はあまり参考にならないと考えていると思われる。この調査結果は、我々が提案しているシステムのコンセプトである、翻訳部品の準備とその組み立てという 2 つの次元の作業を分離し、コンピュータは前者の

支援に注力するという方法論がうまく機能する可能性があることを示唆していると考ええる。

その一方で、これまでに提案されている翻訳支援システムの多くは構造の類似性に基づく類似用例を何らかの形で使用しているものが多い（一例として、[2,4]）。このことは、やはり多くの（特に熟練していない）翻訳者（我々自身も含めて）にとって、文書を組み立てる作業には困難が伴い、あるまじりの表現に類似した表現がどのように翻訳されたことがあるのかを知りたいという需要があることを示唆していると思われる。今後、例えば、非熟練翻訳者が学習用に用いる支援環境、ひいては非熟練翻訳者が熟練した翻訳者により近い翻訳を達成できる支援環境を実現する上で、類似した翻訳用例をどのように扱っていったらいいかについて、今後考えていきたい。

4. まとめ・今後の展望

本稿では、翻訳部品を適宜配置しながら翻訳先文書の執筆を行う、翻訳エディタを統合した新しい翻訳統合環境・新「翻訳パレット」について説明した。新「翻訳パレット」は、翻訳先文書の執筆に先立って準備された翻訳部品をストレスなく執筆中の翻訳先文書に導入し、あるいは移動・削除等の編集が行えることを意図して設計された。また翻訳部品等の使用状況を把握する機能の実装により、ユーザが翻訳先文書の内容や構成を組み立てることを支援できることが期待される。

今後は、このようなシステム設計の狙いが達成されているかどうかを検証し、またシステムをどのように改良していったらいいかを調査するために、実際のニュース翻訳現場の翻訳者に本システムを試用してもらう予定である。また、各種電子辞書以外の情報源から翻訳部品を自動提案するために、日英対訳ニュースデータベースからの表現対応の自動獲得の研究を進めており、次の機会に報告したい。

謝辞 翻訳統合環境・新「翻訳パレット」の開発にあたっては、(株)漢字情報サービスの小鶴友樹氏と東京工業大学の野口正樹氏より多大なる示唆と実装の協力をいただいています。ここに感謝いたします。

参考文献

- [1] 影浦峽, 佐藤理史, 竹内孔一, 宇津呂武仁, 辻慶太, 小山照夫. 翻訳者支援のための言語レファレンス・ツール高度化方針. 言語処理学会第12回年次大会発表論文集, pp. 707-714, 2006.
- [2] 熊野正, 後藤功雄, 田中英輝, 浦谷則好, 江原暉将. 翻訳用例提示システムの設計・開発・運用. 電子情報通信学会論文誌, Vol. J84-D-II, No. 6, pp. 1175-1184, 2001.
- [3] 熊野正, 西脇正通, 田中英輝. 「翻訳パレット」を用いた翻訳支援の提案. 言語処理学会第12回年次大会発表論文集, pp. 701-704, 2006.
- [4] 大蔵清司, 富士秀, 徐国偉, 長瀬友樹, 潮田明. Cliche: さらなる翻訳効率化のための翻訳支援インターフェース. 言語処理学会第12回年次大会発表論文集, pp. 699-700, 2006.