

みんなの翻訳

ボランティアの翻訳者を支援するサイト

独立行政法人情報通信研究機構 知識創成コミュニケーション研究センター MASTAR プロジェクト言語翻訳グループ主任研究員 内山 将夫

PROFILE

1992年筑波大学卒業、1997年同大学院工学研究科修了、博士（工学）。現在、情報通信研究機構主任研究員。機械翻訳と人手翻訳を融合した翻訳手法を研究している。

✉ mutiyama@nict.go.jp



1 ボランティアによる翻訳

ボランティアの翻訳者は、様々な文書を翻訳している。そのなかには、たとえば、オープンソースのマニュアルの翻訳や、ブログの翻訳や、NPO/NGOの文書などの翻訳がある。

ボランティアの翻訳者は、翻訳により、世の中に貢献していると言える。たとえば、オープンソースのマニュアルの日本語訳は、日本人のユーザーにとっては、大変有難いものであるし、ブログの翻訳は、他のメディアが注目しない場所や人々について光を当てるものと言える。

したがって、ボランティアの翻訳者を支援することは、世の中に貢献することになる。

また、日本にいるボランティアの翻訳者は、現在、数千人程度であるが、外国語特に英語を翻訳できる潜在的なボランティア翻訳者の数は、数十万人程度ではないかと推定される。

そのため、翻訳をしたい人が、簡単に翻訳ができる環境を提供すれば、現状よりも、もっと多くの人が翻訳をするようになり、より多くの外国の情報を取り込めるとともに、日本の情報を発信することもできるようになると思われる。

このような動機から、情報通信研究機構言語翻訳グループと東京大学図書館情報学研究室は、共同で、「みんなの翻訳」というWebサイト (trans-aid.jp、図1) を開発した。



図1 「みんなの翻訳」サイト

みんなの翻訳の特徴は、(1) 高機能な翻訳支援エディタ QRedit を誰もが利用できることと、(2) みんなの翻訳で公開されている翻訳には、「2次著作物を作成し、それを公開しても良い」というライセンスが付与されているため、適切な使用であれば、翻訳を利用できるということと、(3) 三省堂の協力により「グランドコンサイズ英和辞典(36万項目収録)」が翻訳支援に利用できることである。

また、みんなの翻訳で翻訳され公開された対訳データや、その他大量の対訳データの検索サービスもある。

2 高機能な翻訳支援エディタ QRedit

翻訳支援エディタ QRedit の基本設計理念は、以下の4点に集約される。(1) 新たな情報・機能を提供するのではなく、翻訳者が現に行っている作業の手間を省く、

(2) システムが決めるのではなく翻訳者が決めるのに必要な情報を提供する、(3) 翻訳者の発想を豊かにする情報を表示する、(4) できるだけシンプルにする。これらの方針は、翻訳者へのインタビューおよび現状の翻訳支援技術の水準に基づいて決めたものである。

QRedit では、入力された原文に対し、複数の辞書や翻訳者が登録した用語を対象に辞書引きを行ない、翻訳者は単語をクリックすることで簡単にその訳語を把握することができる。また、高度なイディオム検出機能を備えていて、語間に挿入を許した異形イディオムを検出できる。

これらのイディオムに対して、QRedit では、図 2 のように、下線を引いて示すなどして、翻訳者が見落さないようにしている。イディオムは、熟練翻訳者でも誤訳する可能性があるので、このように、エディタ側から何らかの警告を与えることは、有用である。



図 2 翻訳支援エディタ QRedit

3 翻訳の共有

翻訳結果を共有するためには、原文と翻訳文の使用許諾について考慮する必要がある。たとえば、当然であるが、原文の著者が翻訳文の公開を許可していない場合には、翻訳文は公開できないので、翻訳結果を共有することはできない。

そのため、みんなの翻訳の利用者には、原文と翻訳文の

使用許諾について確認を求めている。また、みんなの翻訳の利用者には、各自が翻訳した文は、2 次的利用ができるように許可することを求めている。そのために、システムは、みんなの翻訳の利用者が翻訳文を保存するときに、以下のようにして、使用許諾などを確認している。

(1) まず、システムは、「あなたが翻訳の対象とした文書(原文)は、原文著者が明示的に許可している場合を除いて、私的な利用その他など、著作権法で認められている範囲でしか利用できません。原文著者は、その翻訳を公開しても良いと(あなたや他の人に)許可をしていますか?」と確認する。(2) それが「はい」の場合には、システムは、クリエイティブ・コモンズ等の「あなたの文書から 2 次的著作物を作成し、それを公開しても良い」という条件に矛盾しない使用許諾条件を設定してもらうようにしている。

このようにして、みんなの翻訳では、原著者や翻訳者の著作権を尊重しつつ、翻訳を共有できる仕組みを準備している。

4 今後の展開

みんなの翻訳は 2009 年 4 月 8 日に一般公開した、それから 1 年半ほどで、約 1350 人のユーザが登録している。また、アムネスティインターナショナル日本を含むボランティアの翻訳グループにも使っていただいている。

また、みんなの翻訳をつかって、公立大学法人神戸市外国語大学と英国リーズ大学が、共同翻訳プロジェクトをしている。このプロジェクトでは、両大学の学生ボランティアが、英国の小説家ブロンテ姉妹ゆかりのブロンテ博物館の英語の展示情報を日本語に翻訳している。

現状のみんなの翻訳は、英日、日英、日中、中日の翻訳をサポートしている。さらに、英—カタロニア、日韓の翻訳のサポートも実装中である。今後、さらに多言語に展開するとともに、より多くの方に使っていただくよう努力したい。