

## みんなの翻訳第3報

影浦 峽<sup>3</sup>

2 国立情報学研究所    3 東京大学

## 1 はじめに

「みんなの翻訳」(<http://trans-aid.jp/>) は、世界中の文書をみんなで協力して翻訳するサイトである [2, 3, 4] . みんなの翻訳は 2009 年 4 月にオープンして、現在の登録翻訳者数は約 1500 人である .

本稿では、みんなの翻訳の概要を述べるとともに、2010年度に新たに追加された機能であるプロジェクト管理機能について述べる。また、「みんなの翻訳」の妹版として「留学生ネットワーク@みんなの翻訳」をオープンしたので、その概要を述べる。

みんなの翻訳は、プロジェクト管理機能の追加により、個人だけでなく、少人数のグループによる翻訳をサポートできるようになった。また、留学生ネットワーク@みんなの翻訳は、日本の商品を世界に展開するためにみんなで協力して英語や中国語に翻訳するサイトであり、留学生のアルバイトが翻訳をしている。

## 2 みんなの翻訳の概要

みんなの翻訳では、ボランティアの翻訳者を支援するための各種ツールがある。特に、翻訳支援工ディタ QRedit [1] (図 1) は、高品質な三省堂「グランドコンサイス英和辞典 (36 万項目収録)」の辞書引きに加えて、Wikipedia や Web サーチを翻訳リズムを崩さずに利用できることから、翻訳時間を 20 ~ 30% 削減することができる [3]。また、対訳検索 (図 2) をすることもできる。

次に、みんなの翻訳は、翻訳した文書を共有するための仕組みを備えている。そのためのツールとしては、Creative Commons <sup>1</sup>によるライセンス CCL (およびそれと類似するライ

センス)を利用した。CCLを利用することにより、著作者は、自分の著作物に対して、その使用許諾条件をあらかじめ宣言しておくことができるので、他人は、その使用許諾条件に従うかぎりは、自由に、その著作物を利用することができる。

みんなの翻訳では、翻訳者に対して、自分の翻訳文書を公開するときには「その文書を翻訳し、その文書を公開してもよい」という条件に矛盾しないライセンスを付けてもらうようにしているため、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。

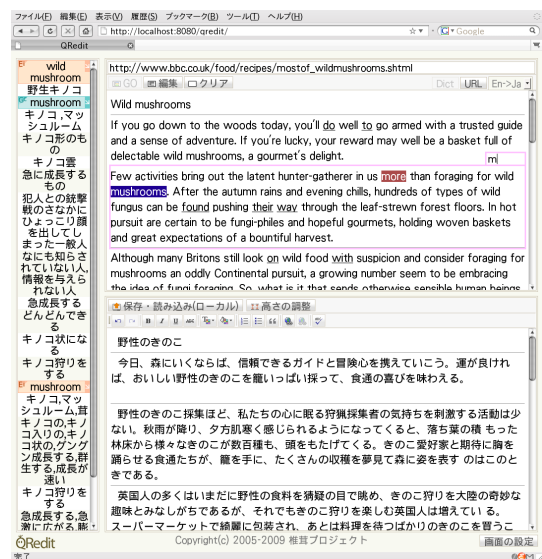


図 1: 翻訳支援エディタ: QRedit

### 3 2010年に追加された機能

### 3.1 グループ翻訳のサポート

みんなの翻訳は当初は個人の翻訳者をサポートすることを目的として設計していたが、サ

<sup>1</sup><http://creativecommons.org/>



図 2: 対訳検索画面

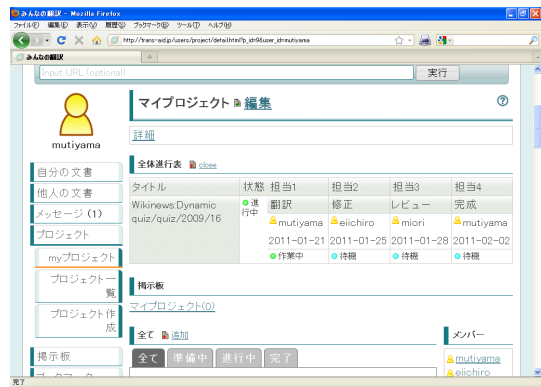


図 3: プロジェクト管理機能

イトをオープンした結果、グループ翻訳の需要が高いことが判った。

グループ翻訳としては、たとえば、大学の授業で翻訳をしたり、ボランティアのグループが翻訳したりすることが考えられる。それらをサポートするために、プロジェクト管理機能を付けた (図 3)。

プロジェクト管理機能の役割は、特定の文章を翻訳する際に、翻訳の役割毎に担当者を割り当てて締切を設定し、メッセージにより、順番や締切などを通知することである。プロジェクト管理者は、一つないし複数の文書をプロジェクトに登録し、翻訳者の割り当て、翻訳期限の設定などを行なうことができる。プロジェクトにおける翻訳のステップは、デフォルトで、下訳、修正、レビュー、完成という、EUの翻訳サービス基準に従ったものを指定できるが、プロジェクトの性格に応じて柔軟にステップを定義することもできる。この機能により、翻訳のスケジューリングが容易になった。

### 3.2 多言語化

みんなの翻訳は当初日英・英日の翻訳のみをサポートしていたが、現在は、日英中の3言語間の翻訳をサポートしている。また、英＝カタロニアの翻訳もまもなくサポートされる。

### 3.3 ブックマークレット

ブログなどで関心のある文章を部分的に翻訳している個人を潜在的な対象として、翻訳したい文章を指定してからみんなの翻訳を呼び出せるブックマークレットを作成した。

利用者はまず、QRedit ブックマークレットを自分のブラウザのブックマークツールバーに登録する。ブックマークレットに登録すると、例えばニュースサイトを読んでいてテキストの一部を翻訳しようと思った場合、ツールバーのブックマークレットをクリックし、翻訳したいテキスト範囲を指定し、ブラウザ内に現れる「Translate」をクリックすることで、(そのとき「みんなの翻訳」にログインしているならば直接、ログインしていないならば自動的に現れるログイン画面でのログインを経て) QRedit が起動され、翻訳作業に入ることができる。

これにより、これまで、(1) 翻訳対象テキストの認定 → (2) 「みんなの翻訳」(transaid.jp) の画面を別途開く → (3) QRedit を立ち上げ翻訳テキストをコピーペースト → (4) 翻訳開始、という手続きで行なっていた Web 文書の翻訳作業のうち、2 と 3 は基本的にスキップすることが可能となった。

ブックマークレットは技術的には比較的単純な機能であるが、社会的な認識の観点からは、これまでは「みんなの翻訳」サーバにアクセスして翻訳支援機能を使うというイメージであったものが、ブックマークレットにより、

自分の手元に翻訳支援機能が置かれたイメージに変換され（あくまでイメージであるが）、翻訳支援機能のユビキタスな受容を促す第一歩となる。

## 4 今後の展開

文献 [4] で述べた課題のうちで、未実施のものは以下の通りである。

まず、翻訳者のための翻訳支援機能は実装したが、読者のために読みたいものが読める機能の実装は不完全である。みんなの翻訳で公開されている文書は既に 2500 以上なので、これらの文書に簡単にアクセスできるような機能が必要である。

また、みんなの翻訳の利用者の拡大という観点について、

- NPO や NGO、現在もいくつかのグループが利用しているが一層の拡大をはかる。
- 共訳による出版翻訳、及び一定の主題を継続的に翻訳出版している出版社。有用性はすでに実証されている。
- 大学のゼミにおける文献購読と翻訳。現在、複数の大学・大学院のゼミで活用されている。
- 高校レベルでの英文和訳学習。翻訳と英文和訳は異なるし、高校英語には到達目標があるので、みんなの翻訳を活用できるかどうかはさらなる検討が必要である。

来年度は、利用者の拡大を主要な目標としたい。また、「みんなの翻訳」とは何であるのかを調べる必要がある。最後に、みんなの翻訳の運営という観点からは、永続的な運営を可能とする枠組を確立したいと考えている。

## 5 留学生ネットワーク@みんなの翻訳

留学生ネットワーク@みんなの翻訳（留みん翻）は、日本の商品を世界に展開するためにみ

んなで協力して英語や中国語に翻訳するサイトであり、留学生のアルバイトが翻訳をしている。留みん翻は現在日英・日中・日韓・中日がオープンしている (<http://ecom.trans-aid.jp/>)。また、留学生のアルバイトを常時募集中である。

留みん翻が翻訳対象とする分野は、E コマースサイトの商品など、多量で安価な翻訳が必要なものである。加えて、最高レベルの翻訳でなくても可能なものが対象である。

留みん翻では「人と機械の協調翻訳」(Human Machine Translation, HMT) を実現している。HMT は、留学生による効率的な人手翻訳と対訳データに基づく機械翻訳から構成される。これら 2 つが結びつくことにより、時間の経過とともに人手翻訳が増えると同時に、機械翻訳の精度が向上する。

まず、留学生による人手翻訳では、図 4 に示すように、高度翻訳支援による効率的な翻訳が可能である。基本的に 1 文を単位として翻訳するのであるが、そのときに、機械翻訳を候補文として閲覧できる。また、類似文検索も可能である。

更に、翻訳済みの文はグレーでハイライトすることにより、修正が不要な場合には、そのままにしておくことができる。また、翻訳が難しい文は、スキップすることが可能である。

次に、あらかじめ大規模な辞書をもっていて、それを辞書引きすることができし、また、アルバイトの留学生が登録した用語を全員が共有することもできる。用語は、QReddit 上から登録できるようになっている。用語は誰かしらにより毎日登録されているようである。

留みん翻では、これまでに健康食品関連の E コマースサイトの日英翻訳をした。そして、その対訳データを利用することにより、そのサイトの翻訳についていえば、機械翻訳のみで、十分に理解可能な翻訳文が生成されるようになった。

具体的には、当初の対訳データ 13.5 万文による統計的機械翻訳では、人手による評価の結果として、48% が理解可能であったが、対訳データを 29.3 万文に増やすことにより、74% が理解可能になった。この結果、「人と機械の協調翻訳」の効果が実証された。

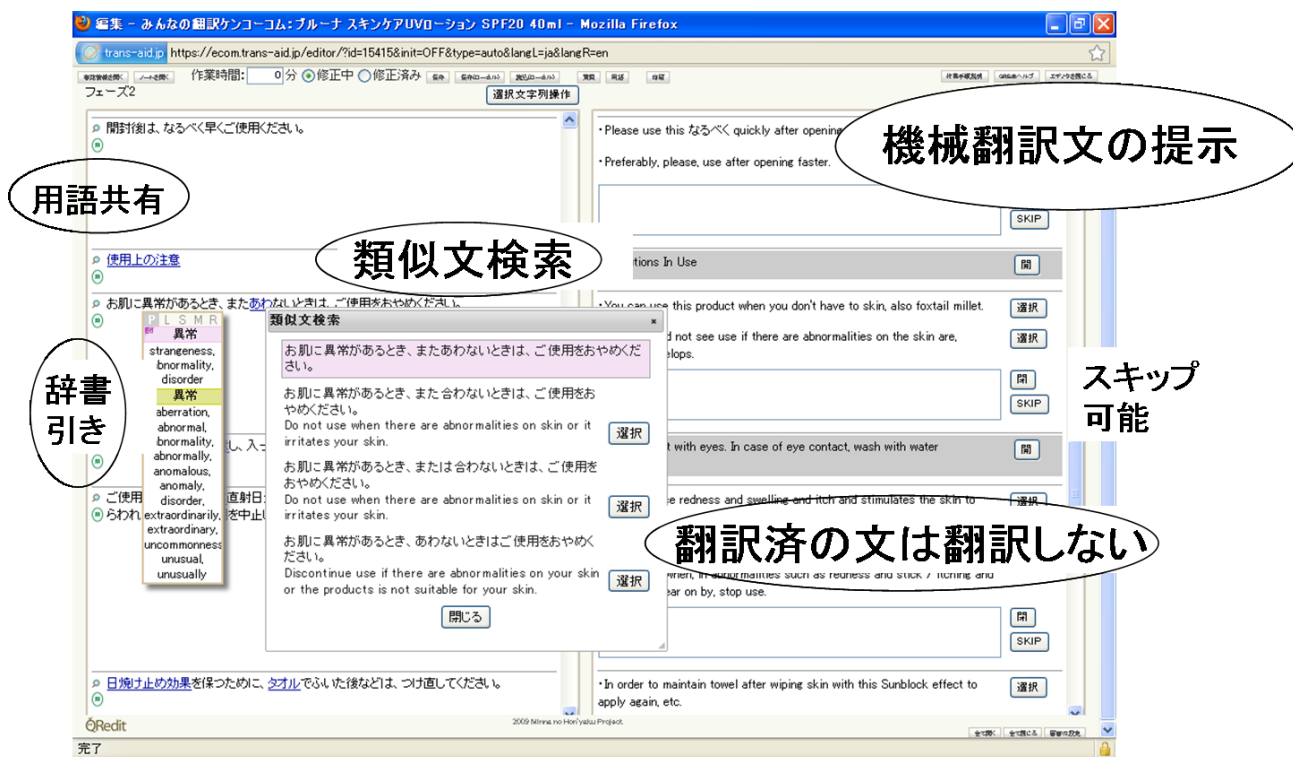


図 4: 翻訳支援による効率的な翻訳

もちろん、対訳データが増えれば、統計的機械翻訳の精度は一般には向上する。そのため、対訳データを継続的に増す仕組みを作ることが重要であり、留みん翻の意義の一つとしてそれがある。

[4] 内山将夫, 阿辺川武, 隅田英一郎, 影浦峯. みんなの翻訳第2報. 言語処理学会第16回年次大会, 2010.

## 参考文献

- [1] Takeshi Abekawa and Kyo Kageura. QRedit: An integrated editor system to support online volunteer translators. In *Digital humanities*, pp. 3–5, 2007.
- [2] Masao Utiyama, Takeshi Abekawa, Ei-ichiro Sumita, and Kyo Kageura. Hosting volunteer translators. In *MT Summit*, 2009.
- [3] 内山将夫, 阿辺川武, 隅田英一郎, 影浦峯. みんなの翻訳. 言語処理学会第15回年次大会, 2009.