

共同翻訳を考慮した「翻訳教育用みんなの翻訳」システム ～みんなの翻訳第4報～

Anthony Hartley[†], 影浦峯[‡], Martin Thomas[‡], 内山将夫[‡]

[†] 東京外国語大学 [‡] 東京大学 [‡] University of Leeds [‡] NICT

1 はじめに

近年、協調的な翻訳が翻訳現場の主流になりつつある (O'Hagan, 2011)。また、協調的な環境での翻訳教育は学習効果を増すとされている (Király, 2000)。従って、翻訳教育において協調的な翻訳を前提とすることが重要であり、有用である。

ボランティア翻訳などでは協調は SNS 等を介して行われるが、翻訳支援機能は提供されていない。プロの翻訳環境では翻訳支援ツールが協調翻訳の機能を強化してきているが、こうした翻訳支援ツールを翻訳教育に利用することは、価格やライセンス、機能等の面で必ずしも容易／十分ではない。そのため、翻訳教育に適したシステムが望まれる。

翻訳教育の観点からは、もう一つ、適切なシステムが望まれる重要な理由がある。翻訳の授業において、素材の選択、学生の訳と修正及び修正理由の説明などは、毎年繰り返し行われるものであるが、そうしたデータの蓄積と共有は進んでいない。システム上で翻訳行動のログをアーカイブ化することで、翻訳知をデータに反映されたかたちで共有し活用することが可能になる。このようなデータは言語処理技術の向上にも資することが見込まれる。

こうした背景から、我々は翻訳教育向け協調翻訳システム「翻訳教育用みんなの翻訳」(以下 MNH-TT)を開発している。概念と仕様の一部は既にいくつかの場で発表しているが (Babych et al., 2012; Babych et al., 2012; Hartley et al., 2013)、試験公開準備が進んだため、システムの実装と機能、公開に関連する要件等を中心に紹介する。

2 基本的な考え方

2.1 翻訳プロセス

日本語で翻訳の技術を解説する書物は、与えられた起点言語の表現をどう訳すかという、狭義の翻訳を論

ずるものが多い¹。けれども、特に翻訳の大部分を占める産業翻訳では、翻訳プロセスは受注から担当の割当て、用語統制、翻訳、修正、レビュー、検証と納品、データ管理等を含むもので、共同で担われる (EU, 2006)。翻訳プロセスの基本単位は、このプロセスをカバーする「プロジェクト」であり、そこに複数の人が参加する。扱われる文書も複数に及ぶことが普通である。

2.2 翻訳知

協調的翻訳のプロセスで特に重要となる「翻訳知」としては、以下がある。

1. 翻訳プロセスで発生する対話／やり取りに反映される知識。
2. 目標言語文書作成の中で行われる辞書引き等の情報参照に反映される知識。
3. 目標言語文書の作成プロセスで加えられる修正に反映される知識。

この他に、翻訳プロセス全体を考えると、プロジェクト管理等をめぐる知識も重要となる。この知識を翻訳知の一つと考えることができるかどうかは判断の分かれるところであろうが、我々のシステムではプロジェクト定義の記録も蓄積される。

翻訳研究の領域では、近年、目の動きやキータイピングのログ、プロトコル分析を使うなど、人間の翻訳行動や認知プロセスを追う、いわゆる「翻訳プロセス研究」(Translation Process Research) が盛んになっているが (Alves et al., 2012; Carl, 2012; O'Brien, 2011)、MNH-TTで蓄積する「翻訳知」は、個々の翻訳者の認知プロセスから翻訳時における人間の認知プロセスへと向かうタイプのものではなく、むしろ翻訳テキストとそれに対する操作という外在化された知識の記述へと向かうタイプのものとして捉えられるものであり、

¹例えば安西 (1995)。光藤 & 田辺 (2008) は優れたテキストだが、それでも翻訳プロセス全体に費やされる紙幅は少ない。



図 1. MNH-TT プロジェクト一覧表示

もちろん、翻訳者の知識が一様ではない点は考慮されなくてはならないものの、テキストやタスクに帰属し、最終的にはそれらの一般化に応じて一般化されるようなものである。

2.3 足場の提供

翻訳プロセスで発生する行動や判断については、適切な制約を設けてユーザの足場とすることが教育上有用である (Babych et al., 2012; Hartley et al., 2013)。我々は、翻訳に関連する先行研究を参照し、プロジェクト参加者の役割とタスク、対話のタイプ、修正の際に用いられる誤りカテゴリーを設定している。これはまた、翻訳知をデータに反映させ、利用可能なかたちでアーカイブ化するためにも重要である。具体的なカテゴリーは既に報じているが (Babych et al., 2012; Babych et al., 2012)、3 節で簡単に紹介する。

3 MNH-TT システムの概要

システムはオンラインで提供され²、ユーザはブラウザ経由でシステムを利用する。

3.1 ユーザ

ユーザは、操作権限の順に、システム管理者、組織管理者、インストラクタ、学習者に分かれる。翻訳教育に関わる組織は、利用の際、組織管理者から組織管理者アカウントを取得する。組織管理者はインストラクタと学習者のアカウントを発行でき、インストラクタは学習者のアカウントを発行できる。組織に属するユーザをその組織のメンバと呼ぶ。

3.2 プロジェクト

プロジェクトは翻訳教育の基本ユニットであり³、機能はプロジェクトに集約される。各プロジェクトはプロジェクト説明、参加者、文書、用語集、翻訳メモリ、

²現在 <http://edu.ecom.trans-aid.jp> で試験運用されている。ベーシック認証をかけてある。

³文書が基本ユニットでないことに注意。

役割	リクエスタ、プロジェクトマネージャ、リサーチャ、ターミノロジスト、翻訳者、修正者、レビュー、校正者、TM マネージャ、MT マネージャ、MT 事前編集者、MT 事後編集者
タスク	概要作成、承認、作業選択、役割割当、文書登録、時間指定、スケジュール管理、最終成果物前チェック、用語抽出、用語翻訳、用語検証、用語集構築、翻訳調査、翻訳、修正、レビュー、校正、TM 選択、TM 更新、MT 事前編集、MT 翻訳、MT 事後編集、MT 更新
対話タイプ	[情報] (リクエスト、支持、代案)、[維持] (確認/感謝、動機付け、明確化、仲介/調整)、[状況] (確認、報告)、[役割] (リクエスト、申し出、辞退、受入)

表 1. 役割・タスク・対話タイプのカテゴリ

ワークフロー管理、掲示板から構成される。インストラクタ以上のユーザはプロジェクトを定義できる。図 1 に、組織管理者がプロジェクト一覧を表示している画面を示す。

参加者と役割 プロジェクト管理者が参加者を登録する。参加者は管理者が所属する組織のメンバが基本となるが組織横断的なプロジェクトの定義と運用もできる。管理者は、参加者の役割を指定する（管理者役割を通して参加者が翻訳プロジェクト管理を学習することもできる）。

文書と参照資源 プロジェクトには複数の文書、用語集、翻訳メモリを登録することができる。用語集の他にシステムは英日・日英・日独・英中・中英・日中・中日・日韓につき基本対訳辞書を提供しており、プロジェクト毎に利用するかどうかを選択できる。言語対によっては MT も利用できる。

ワークフローとタスク プロジェクト実行にあたり、管理者はタスクとスケジュールからなるワークフローを定義する。役割割当や文書登録等の上流工程までをプロジェクトのタスクに含めることもできるし、翻訳と修正といった狭義の翻訳タスクからなる部分プロセスのみからなるプロジェクトを定義することもできる。

[内容] (原文内容の欠落、原文にない訳の付加原文の歪曲、起点言語の侵蝕、目標言語の侵蝕、曖昧さ未解消)、[語] (不適切／一貫しない訳語、不自然な連語)、[文法] (構文誤り、前置詞や助詞の誤り、活用や一致の誤り、綴り誤り・誤変換、句読法誤り)、[テキスト] (原文レジスタからの逸脱、訳文レジスタ違反の用語や表現、不自然なスタイル、結束性違反)

表 2. 誤りタイプ

掲示板 掲示板はプロジェクト毎に一つ作成される。翻訳プロジェクトにおける対話を導く足場として対話タイプが定義されており、掲示板への投稿時には、対話タイプを指定することが求められる。

表 1 に、参加者の役割、プロジェクトのタスク、対話のタイプを示す。

3.3 エディタ

文書の翻訳は、柔軟な辞書引きとウェブ検索機能を備えた翻訳支援エディタ QRedit (Abekawa & Kageura, 2007) に、添削モードと比較モードを追加した拡張 QRedit で行われる。添削モードでは、修正者は、訳文に対して修正範囲、誤りタイプ、修正後の訳を指定する。誤りタイプは MeLLange (Secară, 2005) に基づき一部修正を加えたものである (表 2)。比較モードでは、訳文のバージョンを指定し、修正による差分と修正理由を原文を参照しながら比較検討することができる。図 2 にエディタ上で修正者が訳文修正を行っている画面を示す。

4 翻訳知のアーカイブ化

4.1 翻訳知ログ

MNH-TT では、2.2 にあげた 3 種類の翻訳知に対応する翻訳行動ログを取得し、データベース化する。各翻訳知に対応した基本的なログ取得要素は、プロジェクト、タイムスタンプ、ユーザ・役割、タスクといった共通要素の他に、次のようなものがある。

1. 翻訳プロセスで発生する対話／やり取りに反映される知識：トリガテキスト要素、相手、対話タイプ。
2. 目標言語文書作成の中で行われる辞書引き等の情報参照に反映される知識：文書、対象テキスト要素と属性、参照情報タイプ、参照項目。
3. 目標言語文書の作成プロセスで加えられる修正に反映される知識：文書、対象テキスト要素、初期訳文、修正訳文 1、…、誤りタイプ、修正理由。

2.2 で述べたように、このようなログの集積として外在的にアーカイブ化された翻訳知は、人間の翻訳プロセスをモデル化するものとしてではなく、翻訳における言語表現と関与情報の一般的な操作に関する情報を記録したものと位置づけられる。従って、MNH-TT が構築する外在化された「翻訳知」のアーカイブは、言語処理を活用した翻訳支援メカニズムの開発や改善とも親和性が高い。

4.2 統計と可視化

それとは別に、MNH-TT では、取得されたログの統計情報を提供することで、翻訳教育の場で参加者が翻訳プロセスや翻訳行動を改善するために活用できるようになっている。対話の記録、辞書引きのログ、修正のログは、プロジェクト、利用者、役割、タスク、文書に対して集計できるようになっており、学習者やプロジェクトの診断に活用できるようになっている。また、これら統計情報とプロセスを視覚的に表示する機能も備えている。蓄積されたデータは、経年的に活用することもでき、データ公開をめぐる権利関係やプライバシーの側面をクリアすれば、相互に参照することもできるので、翻訳教育の診断に資するところは大きい。

5 利用に向けた準備

MNH-TT は共同翻訳を視野に入れ、産業翻訳に従事することも想定して翻訳教育／学習を、定義しているため、基本ユニットである「プロジェクト」では、与えられた起点言語の文書を対象言語に訳すことよりもはるかに広いタスクが定義されている。また、翻訳プロセスにおける「足場」を導入しており、利用者はいくつかの意思決定においてその制約に従うことを求められる。(日本の文脈でしばしば行われてきたような) 狭義の翻訳教育との間にはかなりのギャップがあるため、実利用に向けては単なるマニュアルの整備のほかにも、以下のような準備が必要となる。

- ワークフローの類型をいくつか定め、それに従った使い方のガイドを作成すること。とりわけ日本の文脈では、少なからぬ場合、利用初期には翻訳・修正・レビューと用語集構築程度をタスクとした利用が行われると予想される。このようなミニマムワークフローを含め、いくつかの定型的ワークフローを選択できるようにしておくことが重要である。
- 訳文修正における誤りカテゴリの例と修正例の作成。表 2 のカテゴリは、既に MeLLange プロジェクトで評価されたものに基づいているが (Secară,

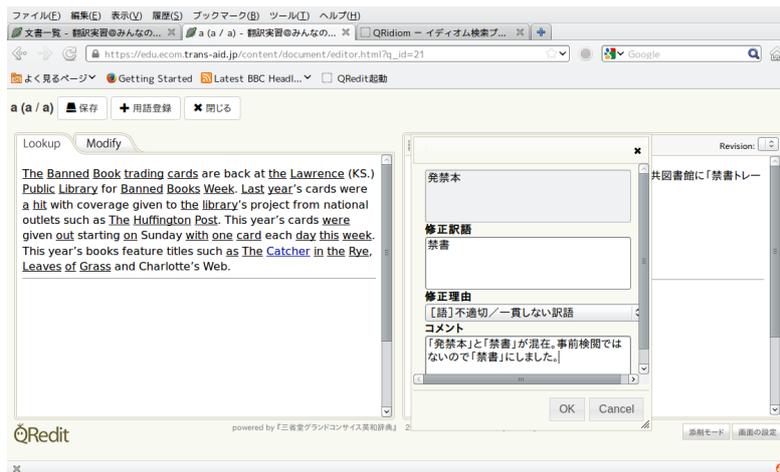


図 2. QRedit での訳文修正プロセス

2005)、運用にあたっては参照できる事例が必要となる。

これらの他に、インターフェースの多言語化、試験公開利用にあたっての権利関係の整備などを現在検討しているところである。

6 おわりに

「翻訳教育用みんなの翻訳」(MNH-TT)の基本的な考え方、機能と実装、公開に向けた要件について簡単に紹介した。発表ではシステムのデモも行う。

MNH-TTは2014年3月までに試験公開する予定であり、公開後、複数の大学で実証試験を開始し、翻訳知アーカイブの構築を進めるとともに、システムについてフィードバックを得てシステムの改善を進める予定である。

謝辞

本研究は2013-2017年度学術振興会科学研究費補助金基盤(A)「翻訳知のアーカイブ化を利用した協調・学習促進型翻訳支援プラットフォームの構築」(課題番号:25240051)の枠組みで行われている。英日・日英の基本対訳辞書として三省堂より高品質辞書『グランドコンサイス英和辞典』『グランドコンサイス和英辞典』の利用許可をいただいている。

参考文献

Abekawa T. & Kageura K. 2007. A translation aid system with a stratified lookup interface. *The 45th ACL Poster and Demo Sessions*, p. 5-8.
 Alves F. et al. 2012. Identifying instances of processing effort in translation through heat maps. *COLING*

Workshop on Eye-tracking and Natural Language Processing, p. 5-20.

安西徹雄. 1995. 英文翻訳術. ちくま学芸文庫.

Babych B. et al. 2012. Scaffolding, capturing, preserving interactions in collaborative translation. *LITC International Conference: The Making of a Translator*.

Babych B. et al. 2012. MNH-TT: a collaborative platform for translator training. *Translation and the Computer 34*.

Carl M. 2012. Translog-II: a program for recording user activity data for empirical reading and writing research *LREC12*, p. 4108-4112.

EU 2006. *EN 15038: European Quality Standard for Translation Services*.

Hartley, T. et al. 2013. MNH-TT to support collaborative translator training. *14th Annual Conference on the Japan Association for Interpreting and Translation Studies*.

Kiraly D. 2000. *A Social Constructivist Approach to Translator Education*. St. Jerome, Manchester.

光藤京子, 田辺希久子 2008. 英日英 プロが教える基礎からの翻訳スキル. 三修社.

O'Brien S. (ed.) 2011. *Cognitive Explorations of Translation*. Continuum, London.

O'Hagan M. 2011. Introduction: community translation: translation as a social activity and its possible consequences in the advent of Web 2.0 and beyond. *Linguistica Antwerpensia*, 10:1-19.

Secară A. 2005. Translation evaluation: a state of the art survey. *eCoLoRe/MeLLANGE Workshop*, <http://ecolore.leeds.ac.uk/downloads/workshop/>.