

## 翻訳という行為と翻訳支援

「翻訳者を支援するオンライン多言語レファレンス・ツールの構築」の基本理念をめぐって

語彙資源の深化とNLP新時代  
科研・合同シンポジウム  
2006年9月14日  
名古屋大学

影浦 映  
東京大学大学院教育学研究科

## 発表の概要

- 翻訳とは何でないか？ 翻訳とは何か？
  - 自然言語処理・言語学と翻訳との、大きな断絶
    - 扱う対象／言語学者・NLP研究者と翻訳者の位置／「言語」とは…
  - 翻訳とはどのような行為か？
  - 翻訳にはどんな情報ツールが有効で、ツールの要件は何か？
- 翻訳者が求める基本的ツールの現状と要請
  - 翻訳者が使っているツール
    - 各種の辞書と各種のテキストの集積(図書館)／TRADOS
  - 翻訳者が使いたいツール
    - 各種の辞書と各種のテキストの集積を、翻訳者の意向に添って統合したもの
  - 翻訳者が使わないツール
    - 機械翻訳システム(サンプルに基づく)用例検索・翻訳システム

## 翻訳とは何でないか？

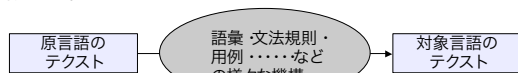
- 自然言語処理・言語学と翻訳との、大きなギャップ
  - 扱う対象
    - 自然言語処理(NL)・言語学(L)は「与えられた言語」を扱う。
    - 翻訳は「与えられた／自ら選んだテキスト」を扱う。
  - 立っている位置
    - NL・Lは、言語が自らの言語活動とは別に、その外側にあると考え(注:コーパスを用いた場合だけでなく「内観」法のときもそうである)、それを「観察」し「モデル化」する。外から見る・言語にコミットしない。確率的でよい。
    - 翻訳者は、自らの言語活動により、与えられた一つのテキストから対象言語のテキストを生成し、それを介して言語を創成する(ベンヤミン・ルター・福澤・二葉亭・ジェレニスキー)。非確率的な「100%」の決断行為。
  - それぞれにとって、言語とは何か
    - NL・Lにとって言語は特定共時態の存在を仮定した上での計算・運用。
    - 翻訳者にとって言語は歴史的に蓄積されたモニュマンへの介入。

## 翻訳とは何でなく、何であるか？

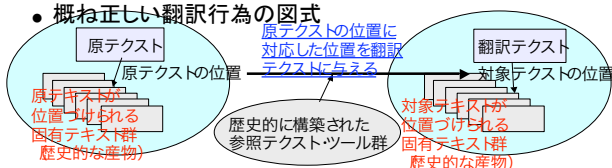
- 扱う対象の点からは…
  - 翻訳は、自然言語処理や言語研究者が想定するような「言語」を、いさかも扱いはしない。「言語」を扱うのは「X文Y訳」のレベル。
  - 翻訳は、テキストを扱う。
- 立っている位置の点からは…
  - 翻訳者は「言語一般」を扱うのではなく、それを外から観察してモデル化するわけでもない。
  - 翻訳者は意志決定を行い、正解を世に問う。人はそれを参照する。
- 言語とは何かという点からは…
  - 翻訳者にとって、テキストは言語学者や自然言語処理研究者が想定する「言語」の「母集団」からのサンプルではない。
  - 翻訳者は、歴史的に固有のテキスト群を「言語」として参照する。

## 翻訳とは、どんな行為か？

### ● 誤った翻訳行為の図式



### ● 概ね正しい翻訳行為の図式



## 翻訳とはどんな行為か？

- 翻訳は、テキストの変換行為ではなく、テキストの位置を対応づける行為である。
- 原テキストは歴史的に集積された固有のテキスト群を背景に生まれ、その中に固有の位置を占める。
- 翻訳は、その位置に対応する位置を、対象言語のテキスト群において求め、翻訳テキストをそこに位置づける。
- 背景テキスト群は原言語と対象言語で同じではないから、位置決めは、位置を作り出す決断行為でもある。
- それらは、サンプルに基づきばらされた「言語」の対応関係(X文Y訳)を自明の前提とし、その上になされる行為である。
- 決断は、翻訳者においては100%正しい。

## この枠の中での、翻訳の多様性

- 多様性の範囲
  - 技術文書翻訳の場合、位置づけ行為は専門語彙のような「言語パーツ」を中心に行われる。
  - 文学テキストの場合、位置づけ行為はテキストの大枠、文体と引用の網の目によって行われる。
  - その間に多岐にわたる階層がある。
- ただし、技術文書翻訳も含め、基本的な枠組みは変わらない。
  - この枠組みを踏まえないテキスト変換行為は翻訳ではなく「X文Y訳」であり、本研究のターゲットではない。
- 蛇足：既訳文書の少ない言語間の翻訳がうまくいかないのは、翻訳がX文Y訳ではなく、翻訳であるから。

## 翻訳に有効なツール

- 以下のものは有効でない
  - ランダムなサンプルに基づく情報を提供するツール
  - 平均してXパーセントの精度を提供するツール
  - 翻訳者に代わって決めてくれる(ふりをする)ツール
- 以下のものは有効
  - 社会的に、そこにあるものは「正解」で、そこになければ諦める場としての100%を提供するツール(良質で多くの人に使われる辞書・図書館)。
    - 「コーパス」ではなく「アーカイブ」
  - 翻訳するテキストが属するべき対象テキスト群の翻訳文書集合
    - 翻訳するテキストに応じた「部分アーカイブ」
  - 翻訳するテキストが与えられたときに、翻訳者の意志決定を補助するために集められる、関連する情報群(辣腕編集者の翻訳補佐作業)
    - 目的に応じた「アーカイブ」の部分的参照

## 翻訳者が求めないツール

- 機械翻訳
  - 翻訳者に代わって決めてくれる。
  - テキストの位置ではなく、言語の意味を変換する。
  - 読解・意味を取るには有効(辞書引きも有効だけど)
- ランダム・サンプルに基づく用例を提供するツール
  - トリストラム・シャンディーを知らないが、Timesは知っている。朱牟田夏雄は知らないが、毎日新聞は知っている。
  - つまり、翻訳者に必要な教養がないし、翻訳者の求めに応じてデータを集めてくれる編集力もない。アーカイブではなくコーパスに依拠する。
  - これらも、「X文Y訳」には有効。

## 翻訳者が使っているツール

- 辞書
  - 一般語・熟語等をカバーする汎用高品質大規模辞書
  - 固有名をカバーする辞書
  - 専門語をカバーする分野依存辞書
- 図書館
  - 公共図書館・大学図書館
  - 図書館まで行かないボランティア翻訳者の場合、Google。
- TRADOS(歴史的な文脈を保存する翻訳用例管理)
  - 自分のためのデータ管理。
  - 付加的に、辞書類のデータを提供。
  - 「翻訳管理者」に使われるという側面も大きい(マネージャに便利)

## 翻訳者が使っているツールと言語単位

単位	どんな情報を参照するか	ツール
一般語	対象言語の表現・原言語の意味	辞書
熟語	原言語の意味・対象言語の表現	辞書
固有名	対象言語の表現・事実情報	辞・図
専門語	事実情報・対象言語の表現	辞・図
引用句	既存の関連する訳	図書館
連語	対象言語の表現	???

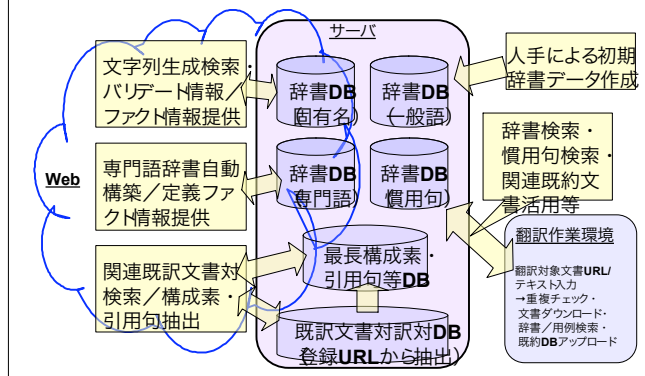
## 翻訳者が使っているツールの課題

単位	ツールの内容	ツールの活用機能
一般語	○	○
熟語	◦	×
固有名	△	△
専門語	△	△
引用句	×	×
連語	×	×

## 翻訳者が使いたいツール

- 現在使っているツールの基本的な問題をクリアすること
  - 熟語・慣用句の自動検索機能 (金平さんの発表)
  - 固有名・専門語のカバー範囲の充実(宇津呂さんの発表)
  - 引用句等を探すための関連既訳文書の提供
- 現在使われているツールの活用機能を組み入れた統一的な作業環境
  - 「翻訳はテキストを扱う作業である」ことに対応して、翻訳テキスト作成作業の流れを妨げない参照機能を備えた翻訳エディタ環境 (阿辺川さんの発表)

## 体系的な構成としては・・・



## 最後に、ご相談など

- 問題をこのように定義すると、「全体に対してXパーセント」という現在のNL業界の基準では、ペーパーが書きにくくなります。
  - 例えば、「イディオム検索」の場合……
  - 例えば、「専門語彙の拡充」の場合……
- これまで話してきたことは、翻訳者にとっては自明なことですが (自明すぎてしばしば翻訳関係の書物は、そのさらに先の技術を話すことになります)、自然言語処理関係者や言語関係者には残念ながら、なかなかわかってもらえないという印象を持っています。
  - 今回の発表でわかりにくいところがありましたら、ご質問・アドバイス等をお願いします。