

産業日本語研究会活動報告

2016年2月29日

「日本人のための日本語マニュアル」 の紹介

横井俊夫

一般財団法人日本特許情報機構特許情報研究所顧問

東京工科大学名誉教授

日本語マニュアルの制作

『日本人**ビジネスマン**のための**ビジネス文章ライティング**日本語マニュアル

- 言葉の仕組を学び、外国語との対照を通じて日本語スキルを磨く - 』

制作グループ：日本語マニュアルの会

横井俊夫（Japio特許情報研究所顧問、東京工科大学名誉教授）

石崎 俊（慶応大学名誉教授、一般財団法人SFCフォーラム理事）

佐野 洋（東京外国語大学教授）

石黒 圭（国立国語研究所准教授、一橋大学連携教授）

猪野真理枝（翻訳家、語学教材作家）

鳥日哲（ウリジャ）（一橋大学国際教育センター非常勤講師）

日本語マニュアルの特徴

- ①ライティング全体のモデルプロセスに沿っている
- ②ライティングルールが言葉の仕組に裏付けられている
- ③他言語への翻訳プロセスが組み込まれている
- ④校正支援や機械翻訳などのテキスト処理ソフトの利用が組み込まれている。

日本語マニュアルの構成

はじめに

1章 文書・文章ライティングのモデルプロセスを学ぶ

1.1 言葉の役割

1.2 ビジネス文書とビジネス文章

1.3 文章特性がライティングを特徴付ける

1.4 ライティングのモデルプロセス

1.5 マニュアルの対象範囲と利用手順

2章 情報を表わし伝える言葉の仕組みを学ぶ - 日本語と外国語とを照らし合わせ -

2.1 文章技術のための言葉の仕組み

2.2 事例から学ぶ

2.3 仕組みを学ぶ

2.4 言葉の仕組みとライティングルール

3章 「表す日本語」で書き、「伝える日本語」へと言い換える

3.1 「表わす日本語」と「伝える日本語」の役割

3.2 「表わす日本語」で書く - 日本語パラグラフライティング

3.3 「伝える日本語」への言い換えルール

4章 「訳せる日本語」へ言い換える

4.1 「訳せる日本語」の役割

4.2 「訳せる日本語」への言い換えルール

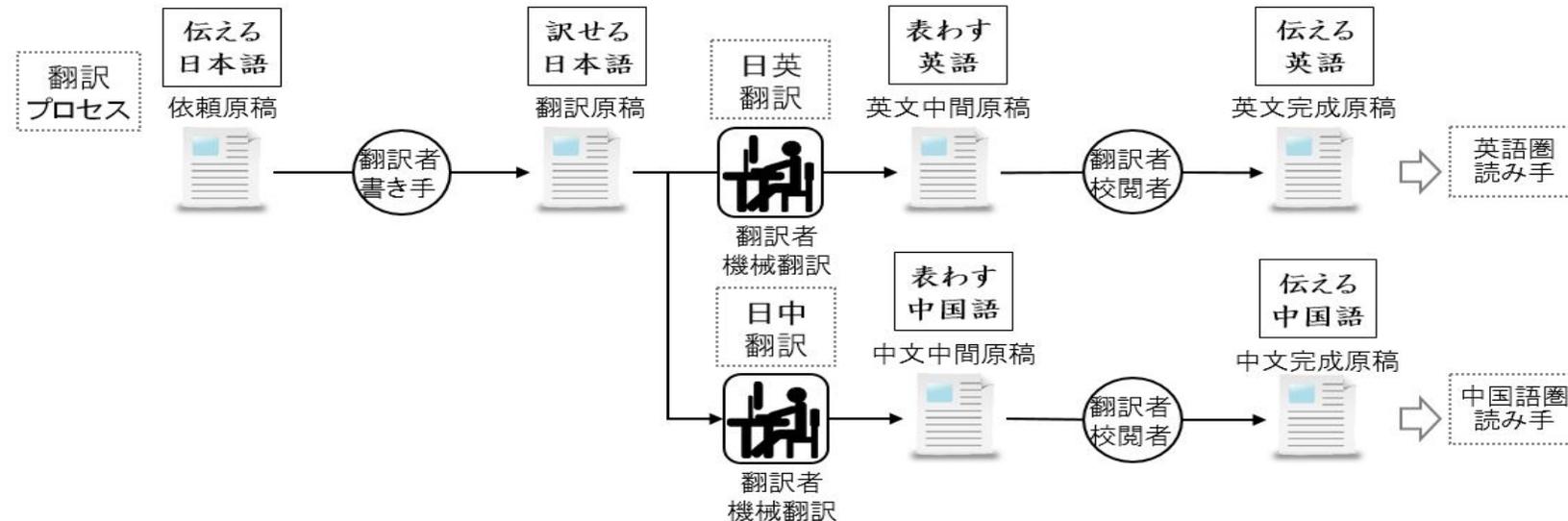
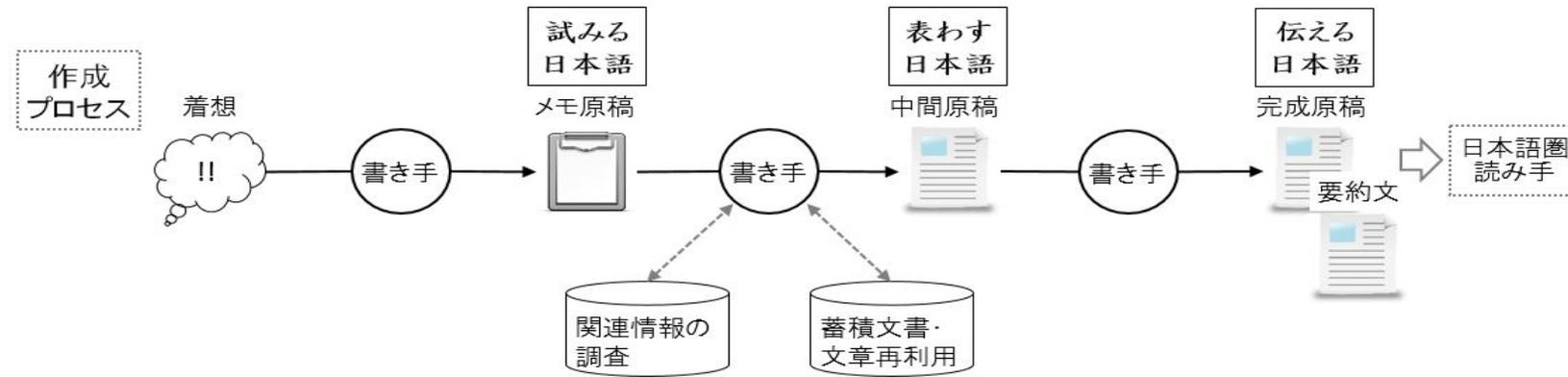
5章 コンピュータの支援機能を活用する - 文章校正ソフトと機械翻訳ソフト

5.1 「伝える日本語」への言い換えを支援する文章校正ソフト

5.2 「訳せる日本語」の翻訳を支援する機械翻訳ソフト

おわりに

文書・文章ライティングのモデルプロセス

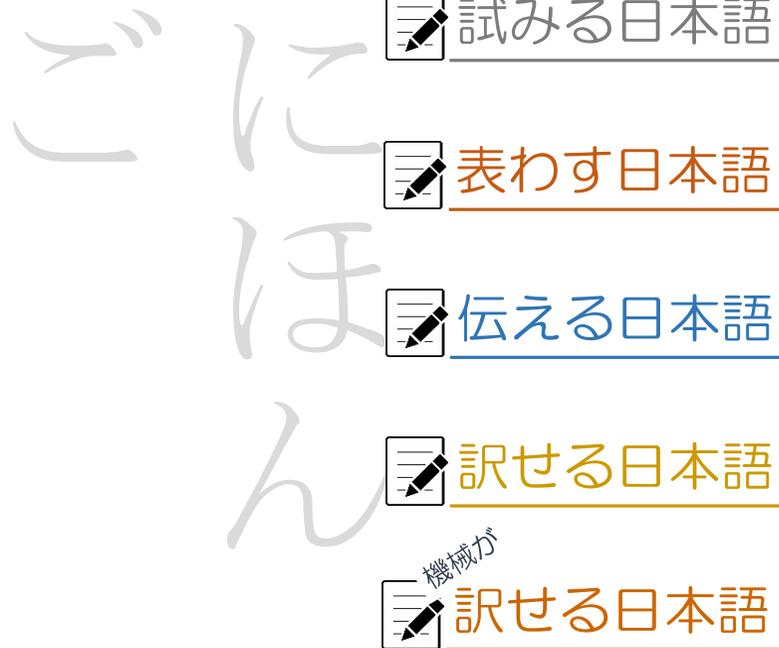


日本人のための日本語マニュアル（試作版）

ビジネスマン

ビジネス文章

言葉の仕組を学び、外国語との対照を通じて日本語スキルを磨く



試作版を公開します

- 産業日本語研究会HPからアクセスできます
- みなさんのご意見をお待ちしています
- 詳しくは配布資料をご覧ください



DOWNLOAD

<http://www.tech-jpn.jp/>



制作グループ 日本語マニュアルの会

<http://ngc2068.tufs.ac.jp/nihongo/htdocs/>

Project Next NLP (2014年春から2015年春)

『エラー分析を通じて自然言語処理技術の方向性を考える』

- ボランティアベースで200名を超える研究者が参加
- 自然言語処理で重要な基礎技術、要素技術、応用技術の18タスク



精度「60%」の壁

精度「60%」の壁を越えるには

- 新たな方法論の研究開発

- 人がコンピュータに歩み寄る

- 人とコンピュータの協調

 - ブラックボックスNLPからインタラクティブNLPへ

↓

「日本語マニュアル」の役割

ブラックボックス機械翻訳からインタラクティブ機械翻訳へ

□ 機械翻訳の方式（安定した基幹技術）

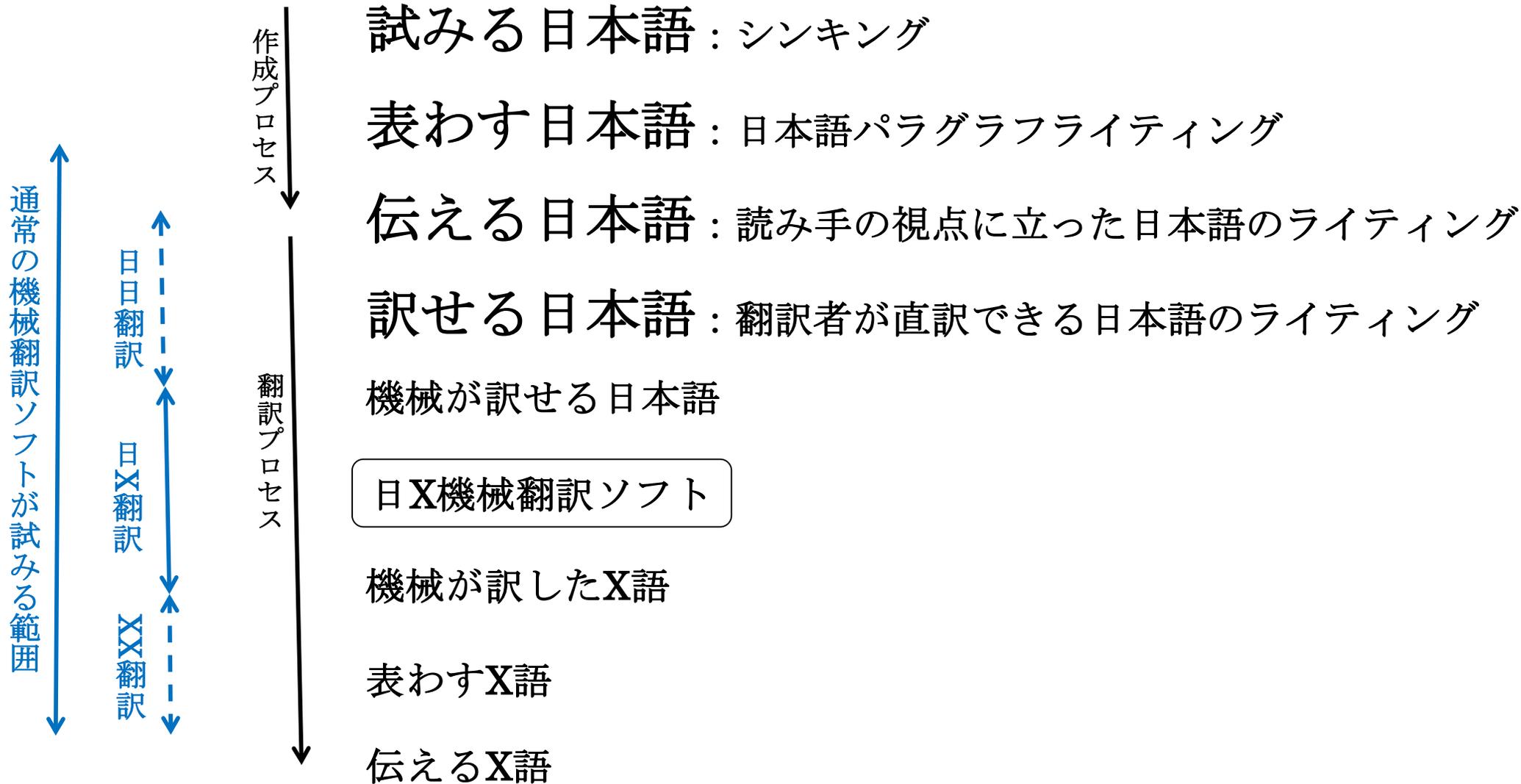
- 規則ベース
- 事例ベース
- 統計ベース

□ 機械翻訳のシステム構成

- ブラックボックス化の呪縛からインタラクティブ化へ



「伝える日本語」 「訳せる日本語」 「機械が訳せる日本語」



日日翻訳とXX翻訳は
インタラクティブ化

「伝える日本語」

「英語に訳せる日本語」

「機械が英語に訳せる日本語」が実現すること

- 市販されている機械翻訳ソフトを活用する
- 新たな活用法を提案する
 - 安定した翻訳結果が得られるようにする
 - 文脈処理を含むより高度な翻訳に対応できるようにする

市販されている機械翻訳ソフトとして、東芝ソリューション社の
The翻訳プロフェッショナルV15（日英）を使用

[文1]

伝える日本語：

「日本人は、外国語が苦手である。」

The翻訳：

“Japanese people are weak at a foreign language.”

訳せる日本語：

「日本人が外国語に苦手である。」

：「伝える日本語」に（訳せる文 - 1 - 3）を適用、二重主語文を主語ひとつ文に言い換える。

翻訳者：

“Japanese people are weak at foreign languages.”

The翻訳：

“<想定外の翻訳結果>”

注）「苦手だ」に二格が共起すると解釈しない、ハガ構文の処理を優先するがための副作用

機械が訳せる日本語：

「日本人が外国語に苦手である。」

：「訳せる日本語」のまま

安定した翻訳結果が得られるようにする

The翻訳も、

「日本人は、外国語が苦手である。」は、翻訳できた。

「外国語は、日本人が苦手である。」を翻訳させると、

“A foreign language are weak at Japanese people.”となる。

注) 「AはBが苦手である」を“A is weak at B”に対応付けているのではないか

- 機械翻訳ソフトが「訳せる日本語」への言い換えルールを完全にカバー、あるいは、カバーしている範囲が明解なら、その言い換えルールの適用をスキップする
- 「訳せる日本語」への言い換えを画一的に適用した方が、安定した翻訳結果が得られる

文脈処理を含むより高度な翻訳に対応できるようにする

簡単な例を使う。

「象は、鼻が長い。」

日本語特有の構文である二重主語文やハガ構文の例として、言語学でよく取上げられる。意外に、ビジネス文章にも使用される。

「金属球は、表面がきれいに磨かれている。」

「我が社は、今年の目標が新興国における増産である。」

「象は、鼻が長い。」という文の意味は

- ① <象の鼻は長い>という簡単な「事柄」
- ② 「事柄」を伝える際に<「象」に焦点を当て前提とする>

「象は、鼻が長い。」と「鼻は、象が長い。」の2つ文

「事柄」に関しては、<象の鼻は長い>という同じこと伝える際に「象」を前提とするのか、「鼻」を前提とするのかが異なる

「象は、鼻が長い。」と「鼻は、象が長い。」をThe翻訳で翻訳すると、いずれも、

“An elephant’s trunk is long.” と翻訳され、伝える際の前提の違いが、翻訳されない。

A 「ここには、多くの動物がいる。動物について特徴的な身体部位を挙げてみる。麒麟は、首が長い。象は、鼻が長い。」

B 「ここには、多くの動物がいる。身体部位について特徴的な動物を挙げてみる。首は、麒麟が長い。鼻は、象が長い。」

A' 「ここには、多くの動物がいる。動物について特徴的な身体部位を挙げてみる。麒麟の首が長い。象の鼻が長い。」

B' 「ここには、多くの動物がいる。身体部位について特徴的な動物を挙げてみる。麒麟の首が長い。象の鼻が長い。」

A、Bは、「伝える日本語」で書かれています。まず、「訳せる日本語」に言い換えます。以下のA-訳、B-訳です。

A-訳「ここには、多くの動物がいる。動物について、我々は特徴的な身体部位を挙げてみる。麒麟が長い首を持つ。象が長い鼻を持つ。」

B-訳「ここには、多くの動物がいる。身体部位について、我々は特徴的な動物を挙げてみる。首については、麒麟の首が長い。鼻については、象の鼻が長い。」

「訳せる日本語」への言い換え操作は以下のようになります。

①主語なし文を、主語あり文に言い換える（訳せる文 - 1 - 1）

②二重主語文を、主語ひとつ文に言い換える（訳せる文 - 1 - 3）

主語ひとつ文への言い換え操作

① 「Xは、Yが長い」を「Xが長いYを持つ」と言い換える。

② 「Xは、Yが長い」を「Xに関しては、Yのそれが長い」と言い換える。

A-訳、B-訳をThe翻訳で翻訳すると、以下のようになる。
「訳せる日本語」と「機械が訳せる日本語」という仕組みを利用することによって、The翻訳は、文脈に渡る処理を行う上質な翻訳を行えるようになる。

A-訳 “Here, there are many animals. About an animal, we mention a characteristic body part. A giraffe has a long neck. An elephant has a long trunk.”

B-訳 “Here, there are many animals. About a body part, we mention a characteristic animal. About a neck, a giraffe’s neck is long. About a trunk, an elephant’s trunk is long.”

日本語マニュアルと人工知能・次世代言語処理

- 日本語マニュアルは、人工知能技術や次世代言語処理技術を駆使したソフトウェアシステムとして実現されるべきである。
- 日本語マニュアルによって、ブラックボックスAIをインタラクティブAIへと進化させることができる。
- 日本語マニュアルによって、言語関連ビッグデータをより良質なビッグデータへと改善し、ビッグデータの資産価値をたかめることができる。