

自動翻訳を用いた、日本人の英語能力の不利性の実証

A proof of handicap of Japanese in English language capability using translation tools

藤田雅之[†] 枚田哲也[†] 朝枝優子[†] 平山敏弘[†] 陳建和[†]
Masayuki Fujita[†] Tetsuya Hirata[†] Yuhko Asaeda[†] Toshihiro Hirayama[†] Kenwa Chin[†]

[†] 日本アイビーエム株式会社

[†] IBM Japan Ltd.

要旨

国際的な情報氾濫が加速される中、大量の英語情報から価値のある情報をすばやく識別・摂取できる能力が重要となっている。しかし日本人は、英語に対する不利性が大きく、このような能力を獲得することが困難である。筆者のグループは、その最大の原因は日本語と英語の言語差異であると考え、しかしこのことは、外国人のみならず、日本人にも本当のところはあまり理解されていない。この点をもっとアピールして、抜本的な対策を講ずることが必要である。従来は、このような日英の言語差異を客観的に証明し、わかりやすく説明できた例は無かった。そこで、自動翻訳を用いて言語間の差異を示す方法を考案し、実際に検証を行った。その結果、日本語と韓国語は、他のヨーロッパ系言語や中国語とは明らかな差異があることを示すことができた。

1. はじめに

インターネット・テクノロジーの進化が加速される中、情報社会はますますコラボレーションに基づく情報流通に依存するようになってきた。従来、マスコミや組織のチャネルを通じて公式に伝達されてきた情報が、現在ではコミュニティの非公式な情報チャネルを通じて拡散・伝播する形態が増えている。その結果、個人に届く情報は大量である上に玉石混交となり、それらの大量情報の中から価値のある情報をすばやく識別して摂取できる能力が、非常に重要となっている。

このような情報氾濫の動きは、海外で顕著である。しかし、海外から流入する情報のほとんどは英語であり、日本語母国語者（以下「日本人」と表記）の英語摂取力が、国際適応力や国際競争力を維持する上で、重篤な課題となっている。

現状は、日本人がこのような能力を発揮できる状況からは程遠い。著者のグループは、日本人が英語をうまく摂取できない原因は日本語と英語の言語的な差異によると考える。このことに関しては、対照言語学などがすでに明らかにしているが、非日本語母国語者（以下「外国人」と表記）はおろか日本人でさえ、その差異の大きさを本当に理解しているとはいえない。これを客観的に示し、日本人にも外国人にもわかりやすく納得できる形で説明できた例があまり無かったからではないだろうか。

そこでこの課題を訴求するため、自動翻訳によって様々な言語と日本語・英語を比較し、その差異を明らかにする方法を考案し、検証を行った。

2. 検証の背景

言語の差異に関しては、言語の普遍的本質を探るため、あるいは外国語学習に役立てるために、主に言語学の領域で、音声・統語・意味・用法などの幅広い側面で分析されてきた[1][2]。そして、日本語の言語特性から、英語などのヨーロッパ言語との差異性が大きいことが指摘されてきた[3][4]。

しかし、個々の議論はそれぞれの分野内で行われることが多く、言語間の包括的な差異性や類似性を示すことは困難であった。また、日本の英語教育の問題や、日常的英語環境の不足など、環境側面が強調されてきたため、言語そのものの差異が日本人の英語能力に致命的な不利性を与えることについて、声高には論じられなかった。さらに、主な応用分野として英語教育が想定されているため、どちらかといえば、初級・中級者の英語習得で問題となるようなテーマとして取り扱われてきた。

本発表では、大量の英語情報の処理に必要な、以下の能力を問題としている。これらの能力は、日本人では英語上級者であっても能力発揮が困難である。十分に発揮できなければ、英語摂取の生産性が低くなり、国際環境において様々な不利が生ずる。また、英語情報摂取に偏りや誤謬が生じ、情報をただしく理解・応用できないと懸念される。

- 大量の英語情報を速読、あるいは斜め読みして、重要な情報をすばやく探し出す能力
- 速読によって大量のドキュメントの内容をすばやく的確に把握する能力
- まとまった英語ドキュメントを生産性高く短時間に日本語に翻訳する能力
- 英語情報の内容を日本語の日常環境に実感を持つて的確に適用する能力

英語上級者は、英語情報を英語環境において活用することは容易だが、英語を通じて受け取った情報を日本語の環境に的確に適用することが難しい。「翻訳せずに英語で考えよ」という英語習得時の訓練により、英語で考えてしまっ、日本語環境での実感を伴わないからである。

ここにあげた能力は、母国語と外国語の言語情報処理がどれくらい類似しているかということに依存する。日本語と英語だけでなく、他の言語との比較において、それぞれの差異性を明らかにできれば、日本人にとって、ここに上げた能力の獲得が非常に困難であることを示すことが可能と考える。

3. 日本語と英語の差異の検証

3.1. 自動翻訳を用いる基本的なアイデア

英日・日英翻訳ツールが実用に程遠いことは、誰もが認めることである。一方、韓国語と日本語、あるいは、ヨーロッパ言語同士などでは、自動翻訳の結果がよいこともよく知られている。

差異の大きな言語間では、多義性が構造にまで及び、解釈によって翻訳後の構造が変わる。また、語彙体系が一致しておらず、一つの単語が複数の訳語に写像される。さらに、発想が離れているため、言い回しに対応していない。これらのあいまい性が、自動翻訳の困難さをもたらしている。

これに対して、類似する言語は、多義性のある文であっても同じ構文に写像されるので、意味の解釈まで踏み込む必要がない。また、一つの単語が一つの訳語に写像されることが多い。さらに、類似する言い回しを持っていることも多い。自動翻訳にとってのあいまい性は非常に低い。

したがって、複数の言語から自動翻訳された日本語を日本人が評価し、どれくらい日本語として正しいかを調べることで、もとの言語と日本語との差異性を測ることができるのではないかと考えられる。

しかしこの考え方には以下のような問題点がある。

1. 言語間の差異というより、単にツール間の翻訳性能の比較となる可能性がある。
2. 翻訳前のオリジナルテキストが異なる言語のテキストなので、その翻訳テキストを比較しても、翻訳結果の厳密な比較とならない。また、複数の言語のテキストを準備する必要がある。

1.については、複数のツールで同じ結果が得られれば、個々のツールの性能でなく、もっと本質的な差異のある可能性が高い。そこで、テクノロジーの異なる複数のツールを用いることとした。

2.に対しては、同一のテキストを用いて検証ができるよう、以下のステップの評価方法を採用した。

- ① オリジナルテキストを日本語で準備する
- ② 自動翻訳を行って、比較する諸言語へ翻訳する（たとえば日英翻訳など）
- ③ そこで得られたテキストを、自動翻訳を行って日本語に逆翻訳する（たとえば英日翻訳など）
- ④ その結果を評価する

自動翻訳は不可逆なので元のテキストに戻ることはない。この方法ならば、オリジナルテキストは日本語で準備すればよく、オリジナルも評価対象も日本語なので、日本人だけで評価が可能となる。

3.2. アンケートによる検証

前項のアイデアがどれくらい有効なのかを検証した。まず、ニュース、日本国憲法、ビジネス記事の3つの日本語テキストを準備した。それらを、英語・韓国語・中国語を媒介とし、翻訳ツールとして Google 翻訳（統計的翻訳）と Yahoo 翻訳（構文解析翻訳）を用いて、それぞれ双方向2回の翻訳を行い、合計 3 （素材数） $\times 3$ （媒介言語数） $\times 2$ （翻訳ツール数） $= 18$ のテキストを準備した。

これらの18のテキストとオリジナルテキストを示し、一般の日本人（30代から60代までの男女）8人に、日本語としてどれくらい正しいかを評価してもらった。設問は以下とし、4段階評価の悪いほうから、それぞれ1~4点の評点を与えて、カテゴリーごとの平均を算出した。

1. 文法的に正しいか？ (4段階評価)
2. 意味が通るか？ (4段階評価)

また、ニュース・日本国憲法・ビジネス記事の素材ごとに、よいもの2つ・悪いもの2つを選んでもらい、「よい」に選ばれた場合 +1, 「悪い」に選ばれた場合 -1 として言語ごとに合計スコアを算出した。結果を表1~4に示す。

表1 素材別・言語別評点平均

言語	ニュース	憲法	ビジネス	合計
英語	1.53	1.34	1.64	1.51
韓国語	2.56	2.09	2.72	2.46
中国語	1.19	1.53	1.50	1.40
合計	1.76	1.66	1.95	1.79

表2 評価面別・言語別評点平均

言語	意味	文法	合計
英語	1.38	1.64	1.51
韓国語	2.50	2.42	2.46
中国語	1.39	1.42	1.40
合計	1.76	1.82	1.79

表3 ツール別・言語別評点平均

言語	Google 翻訳	Yahoo 翻訳	合計
英語	1.21	1.80	1.51
韓国語	2.40	2.52	2.46
中国語	1.52	1.29	1.40
合計	1.71	1.87	1.79

表4 言語別投票スコア

言語	合計
英語	-13
韓国語	32
中国語	-20
合計	-1

素材別・評価面（文法・意味）別・ツール別の、あらゆるカテゴリーの評価において、韓国語を媒介とする自動翻訳後のテキストの評点平均は、2点（「正しい部分もあるが破綻も目立つ」）から3点（「一部破綻もあるが日本語としてかなり正しい」）の間となり、英語および中国語を媒介とする自動翻訳後のテキストの評点平均は、1点（「全体に破綻しており日本語と認めがたい」）から2点（「正しい部分もあるが破綻も目立つ」）の間となった。また、韓国語媒介のテキストは、「よいもの」への投票が「悪いもの」への投票を大きく上回り、英語と中国語は、それぞれ「悪いもの」への投票が「よいもの」を上回った。これは、日本語と韓国語が SOV 語順であり英語と中国語が SVO 語順であること、歴史的に日本語と韓国語は近い言語であると認められることなどの言語学の知見と一致する。実際に処理済のテキストを筆者らも比較したが、ここに上げた数字以上に、その差を実感できた。

3.3. 検証対象の拡大

前項の調査検証により、言語の差異性を調べる方法として、自動翻訳を利用できることが確認できた。そこで、検証対象を他のヨーロッパ言語にも広げ、より多くの言語との差異性を調べることにした。また、日本語から日本語への変換だけでなく、英語から英語の変換に基づく検証も必要であると考えた。

しかしながら、アンケート調査は回答者の負担が大きく、言語数が増えたときに適切な回答を得られるかどうか懸念がある。また、英語を軸とした検証を行うためには、英語としての評価が必要であり、英語を母国語とする回答者を確保しなければならない。アンケートによる評価は困難と思われた。

この点を解決するため、アンケート検証のために作成した変換後テキストとオリジナルテキストとを、単語レベルで比較して、その比較結果とアンケートの回答とに何らかの相関があるかどうかを検証した。変換後テキストから付属語を除いた後、オリジナルテキストと同じ相対位置に現れている同一の単語の数を数え、その数をオリジナルテキストの全単語数で割って、18のテキストごとにマッチング率を算出した。その後、そのマッチング率と、アンケートの文法・意味評価得点の平均との相関係数を算出した。この結果、相関係数は+0.779 (f 値は 0.00013) であった。これは、マッチング率とアンケート評価に高い相関があることを示している。これによって、アンケート調査を行う代わりに、単語レベルでマッチング率を算出して比較すれば、翻訳結果のテキストの評価を行うことができることを確認した。

そこで、比較言語として、日本語・英語・フランス語・ドイツ語・スペイン語・イタリア語・韓国語・中国語を選び、日本語と英語のオリジナルテキストから、Google 翻訳を用いて双方向2回の翻訳を行い、その結果とオリジナルテキストとのマッチング率をそれぞれ測定した。結果を表5・6に示す。

表5 日本語テキストから各言語を媒介とする自動翻訳を行った結果のマッチング率

媒介言語	イタリア語	スペイン語	ドイツ語	フランス語	韓国語	中国語	英語
マッチング率	43.2%	38.8%	38.9%	40.1%	79.8%	33.3%	52.0%

表6 英語テキストから各言語を媒介とする自動翻訳を行った結果のマッチング率

媒介言語	イタリア語	スペイン語	ドイツ語	フランス語	韓国語	中国語	日本語
マッチング率	80.1%	83.0%	74.3%	85.4%	42.7%	68.6%	35.0%

日本語との差異を調べる比較(表5)では、韓国語が約80%のマッチング率を示している反面、ヨーロッパ系の言語のほとんどが、40%前後にとどまっている。一方、英語との差異の比較(表6)では、ヨーロッパ系言語のマッチング率が80%内外にあるのに対し、韓国語は40%強、日本語は30%台となっている。中国語は、日本語とは離れているが、英語とは日本語よりも近いように見受けられる。

この検証によって、英語と他のヨーロッパ言語、および日本語と韓国語が類似性を持つグループ言語であることが数字で示された。

まとめ

多義性が構造を変えるような、あるいは語彙体系や発想が異なるような言語の相違性が、自動翻訳を困難にする。逆に、これらが少ない言語間では、機械的な翻訳が可能である。上で「英語で考える」ことに関して言及したが、母国語と統語が大きく変わらない外国語の場合は、母国語と同じ思考の流れを用いることができる。そのような外国語ならば、わざわざその言語で「考える」必要はない。母国語で普段用いている言語能力をそのまま使用できるため、高度な能力を発揮することは容易である。

ヨーロッパ言語圏では、多くの言語が同じ思考の流れを持っていて、母国語と思考の流れが異なる外国語とはどういうものであるのかを理解することが難しいと思われる。一方、ほとんどの日本人は、最初の外国語が英語であるため、外国語とは、思考の流れの異なる言語なのだという認識を持っているだろう。日本人とヨーロッパ語圏の人々とは、外国語に抱くイメージが異なる可能性が高い。

このように、日本語と英語の基本的な差異性の相互認識にはギャップがある。そしてこれまで、ビジネスに求められる日本人の英語能力に致命的な不利性があることについて、正しく認識されることはなかったといつてよい。そしてその解決は、情報の受け手である日本人個人個人の努力にゆだねられてきた。しかしこの問題は、言語の本質にかかわることであって、個人の能力や努力に帰せられるものではない。日本人の英語情報の摂取能力に関する問題とは、単なる習得の困難さではなく、あるレベル以上の能力発揮が本質的に困難なのだということである。この課題に対しては、社会・組織を挙げた対策とともに、情報を発信する側の協力・対応が不可欠である。そのためにも、日本人の言語事情をヨーロッパ語圏の人々に広く知らせ、彼らが認識する外国語観を改めさせる必要があると考える。

本発表は、日頃自動翻訳を行っている人が感じていた事を実証したに過ぎない。しかし、実証の過程で得られるテキストは、誰でも比較でき、言語の相違性を実感することができる。検証は簡単に再現できるので、読者もぜひ比較してみられたい。言語の違いに関する理解促進の一助となれば幸いである。

参考文献

- [1] 風間喜代三, 上野善道, 松村一登, 町田健, 言語学 (第2版), 東京大学出版会, 2004
- [2] 石綿敏雄, 高田誠, 対照言語学, おうふう, 1990.
- [3] 角田太作, 世界の言語と日本語, くろしお出版, 1991.
- [4] 三枝 幸夫, “TOEIC テストガイダンス”, <http://www.nullarbor.co.jp/guidance/index.html>