

多言語自動翻訳の実力

—日韓翻訳—

Automatic Multilingual Translation - Japanese - Korean Translation -



大阪大学名誉教授

成田 一

英日対照構造論・機械翻訳・言語教育/外国語習得論専攻。大阪大学功績賞受賞。著書『パソコン翻訳の世界』（講談社）、『日本人に相応しい英語教育』（松柏社）ほか、編著『こうすれば使える機械翻訳』（バベルプレス）、『英語リフレッシュ講座』（大阪大学出版会）ほか、共著『英語教育徹底リフレッシュ』（開拓社）ほか。論文・新聞・雑誌記事多数。英語教育総合学会会長。情報通信技術研究交流会運営委員。

✉ narita.hazime@wave.plala.or.jp

TEL 0797-81-7061

2016年度の拙稿「多言語自動翻訳の実力—文法処理能力を検証—」では、英日、英仏、英独翻訳を構文構造を中心に「文法処理能力」について検証した。これは英語が文法構造に大きく依存した構造維持型の言語であり、文型ならびに主要な文法構造について検証することが最も正確なそして改善に向けてのヒントを得られる方策であるためだ。本稿では日韓翻訳について検証するが、英欧翻訳と違い、日韓翻訳では（文型など）構文構造は解析/生成の主要な因子ではないので、構文構造を中心に文法処理能力について検証するということはあまり意味がない。このことから、一般的な文（筆者の著書の一節以外は、最近の読売新聞の記事）を検証データに使う。今回はロゴヴィスタの「コリヤ英和！—発翻訳2017マルチリンガル」と高電社の「J-SERVER Office Biz.クラウド」を使って、翻訳能力を見て行く。本稿での評価は大阪大学大学院言語文化研究科修士の白姫恩氏¹と開発企業²に依頼し、筆者がまとめた。

韓国語の語彙/文法的な特徴 韓国では1960年代から漢字の使用が極端に制限され表音文字ハングル一色になったことが障壁となって、多くの日本人には韓国語（朝鮮語）は日本語と「かなり違った言語」という印象を持

- 1 白姫恩（ベク・ヒウン）：大阪商業大学リエゾンセンター講師、韓国語弁論大会審査委員（2011年～）。主な著書に『口を鍛える韓国語作文—語尾習得メソッド—初級編』ほか。
- 2 ロゴヴィスタ評価者の劉承桓氏：韓国人プログラマー。日本語能力試験（JLPT）2級。

たれているが、実は韓国語と日本語は世界の言語の中でも極めて良く似た言語なのだ。かつてのように、韓国語を「漢字ハングル混じり」で表記したら、日本語の「漢字仮名混じり」とそっくりだ。日本語と韓国語は、どちらも格助詞で文法関係を表し、述部の動詞と助動詞の接続構造がほとんど同じで、連体修飾構造や文成分の語順も共通だ。（敬語法³など表現面で若干の違いはあるものの）文法構造がほぼ同じなので、構造解析は問題にならない。日本語の活用形態の解析がうまく行けば、語彙レベルの置き換えだけでも、意味の理解できる韓国語に翻訳できる。言わば、日本語の文章の文法をそのまま踏襲して韓国語に換えられるのである。ただし、「助詞や時制を選択する情報に不備がある」などの要因が訳文の不具合を生むこともある。特に助詞は日本語よりかなり多く、用法にズレがあるので誤訳が生じやすい。なお、文化的な背景が関わる表現は、言語的には正しい翻訳であっても、理解できないことがある。たとえば、「懐が少し深くなって」、「おひとりさま」（下記表中の例文）などは、直訳では、日本人のような理解にはならない。

日韓翻訳ソフトは95%±3%ほどの翻訳率になる。だが韓日翻訳ソフトはそれより5~15%ほど翻訳率が低い（「確認翻訳」参照）。漢字語を音声表記したハングル語には同音異義語が多く、複数の可能性を文脈的に選択できないため語彙認定の誤りがみられるのに対し、日

- 3 日本語では「相対敬語」なので、身内について言及する際には敬語を使わないが、「絶対敬語」の韓国語では身内にも敬語を使う。

本語は漢字仮名混じりで表記された段階で「人間の判断に依る同音異義語の選択」が済んでいるためだ。漢字国である日本や中国、台湾との経済・文化的関係を睨み、韓国における漢字使用が増えれば翻訳精度も上がる。ただし、中国は簡体字なので繁体字との対応が要る。

以下の表が、今回実施した翻訳例の一部（約半分）の評価⁴（正訳レベルを◎、○、△などで示すほか、流暢さ、適切性については、「J SERVER」の評価を斜線/の左に、「コリヤ英和！」の評価を右に示す）だが、翻訳結

- 4 白姫恩氏の評価：IWSLT (International Workshop on Spoken Language Translation) で採用の下記の基準に基づき数値化したものと個人的なコメント。
 「流暢さ」：翻訳文が、どの程度「自然な表現か」を評価する。
 5－まったく問題ない、4－良い、3－日本語的、
 2－不自然、1－理解不能
 「適切性」：翻訳文が、どの程度「原文の情報を含むか」を評価する。
 5－すべての情報、4－ほとんどの情報、3－多くの情報、
 2－少しの情報、1－情報なし。

果はほぼ同じなので、(翻訳結果が貼り付けできた) ネット版 J SERVER の画面を表示する⁵。韓訳のほか、①原文の日本語を韓訳し、②その韓訳を和訳した「確認翻訳」の結果を併記している。韓国語の知識がなくても原文の日本語と確認翻訳の日本語を対照すれば、韓訳の翻訳精度がほぼレベルが推定できるだろう。「確認翻訳」は「逆翻訳」とも言うが、日韓翻訳と韓日翻訳という2回の翻訳を経ているだけに誤訳率は理論的には2倍になる。だが、韓日翻訳の段階では、入力文の韓文はハングル表記になるので、同音異義語の誤訳がかなり増える。このため、2倍以上の誤訳率となる。また、ハングルの音をそのままカタカナの日本語に変えるようなケース（「廃炉技術」で→「ペファロギス リラ」）も見られる。それでも確認翻訳の日本語はほとんどが十分に意味理解可能だ。

- 5 「コリヤ英和！」の翻訳結果は、ソフト上でないと対照表示できない。

	原文	確認翻訳	訳文	評価	コメント	流暢さ	適切性	コメント
	男子体操団体、12年ぶり優勝	男体操団体、12年だけ優勝	남자체조단체, 12년만 우승	△	「ぶり」の訳が正しくない。	4/4	2/2	
	新黄金時代に	新しい黄金時代に	새황금시대에	◎	文句なし	3/3	4/4	
1	リオデジャネイロ五輪第4日の8日、体操の男子団体総合決勝で日本が2004年アテネ大会以来3大会ぶりに優勝した。	リオデジャネイロオリンピック第4日の8日、体操の男子団体総合決勝で日本が2004年アテネ大会以来3大会ぶりに優勝した。	리오데자네이루 올림픽 제4일의 8일, 체조의 남자단체종합 결승으로 일본이 2004년 아테네 대회 이래 3 대회만에 우승했다.	◎	「で」の訳が正しくない。	2/2	3/3	「第4日の8日」は、普通の韓国人には、具体的な意味が全く分からない。
	政府はロシアに対し、エネルギー分野で包括的な協力策を打ち出す方針だ。	政府はロシアに対し、エネルギー分野で包括的な協力策を前に出す方針だ。	정부는 러시아에 대하여, 에너지 분야에서 포괄적인 협력책을 내세울 방침이다.	◎	文句なし	5/5	5/5	自然。原文の情報の伝達性も非常に高い。
2	極東や東シベリア地域で石油や天然ガスの資源開発を検討するほか、原子力発電所の廃炉技術でも協力する。	極東や東シベリア地域で石油や天然ガスの資源開発を検討すること以外に、原子力発電所のペファロギスリラも協力する。	극동이나 동시베리아 지역에서 석유나 천연가스의 자원개발을 검토하는 것 이외에, 원자력발전소의 폐화로 기술이라도 협력한다.	○	「廃炉」の訳が正規ではない。「でも」の訳が正しくない。	4/3	3/3	「～でも」の使い方の誤りによって、文の意味やニュアンスもだいたい変わる。
	本がなくても生きられる。	本がなくても暮らせる。	책이 없어도 살 수 있다.	◎	文句なし	5/5	5/5	自然な訳。精度も高い。
	けれども本を読めば「自分以外のものに関する想像力」がつく。	ところが本を読めば「自身以外の物に関する想像力」が立つ。	그렇지만 책을 읽으면 「자신 이외의 물건에 관한 상상력」이 선다.	○	「つく」の訳が正しくない。	3/3	4/4	「물건」は物体的な対象だけを意味する。
3	懐が少し深くなって、世に流通する「安易な物語」が恥ずかしくなる…	ポケットが少し深くなって、世の中に流通する「安易な話」が恥ずかしくなる…	호주머니가 조금 깊어져서, 세상에 유통하는 「안이한 이야기」가 부끄러워진다…	△	文化的意味合い濃い原文表現のため、韓国語的表現としては分かりにくい。	1/1	1/1	理解不能、不自然。



4	<p>岩下久美子（故人）さんは、「おひとりさま」の概念として2001年に出版した著書「おひとりさま」（中央公論新社）の中で、「個が確立できた大人の女性」、「仕事も恋もサクセスする為の、身につけるべき生き方の哲学」、「individual（個々の、個人のという意）」、「自他共生していく為のひとつの知恵」としています。</p>	<p>岩下欧米鼻（故人）さんは、「一人様」の概念として2001年に出版した著書「一人様」（中央空論（公論）鳥社）案（中）で、「犬を確立できたおとなの女性」、「ことも恋愛もサクセスするための、実らせなければならない人生の態度の哲学」、「individual（それぞれの、大人のという意）」、「自他全部産んで行くための一つ知恵」とっています。</p>	<p>이와시타 구미코（故人）씨 는, 「혼자님」의 개념으로서 2001년에 출판한 저서 「혼자님」（중앙공론 새사）안（속）에서, 「개를 확립할 수 있었던 어른의 여성」, 「일도 연애도 석세스 하기 위한, 익혀야 할 삶의 태도의 철학」, 「individual（각각의, 개인의라고 하는 뜻）」, 「자타 모두 낳아 가기 위한 하나 지혜」라고 하고 있습니다.</p>	△	<p>「できた」の訳が正しくない。 「共生」の訳が正しくない。</p>	1/1	2/1	<p>「個」の翻訳が不自然で意味が分かりにくい。「共生」の「共」と「生」がそれぞれ別に訳され、意味が変わっている。漢字の「生」だけに反応し、「産む」という意味に変換。 ※韓国語では「～の」を省く方が自然な場合が多い。</p>
5	<p>厚生労働白書によると、このような日本の生涯未婚率（生涯未婚者の割合で通常は50歳時点の未婚率で代替している。40代後半と50代前半の未婚率を平均して算出）は年々高くなっており、厚生労働白書2015では男性は24.2%、女性は14.9%と推計されています。</p>	<p>厚生労働白書によれば、このような日本の生涯未婚率（生涯未婚者の割合で通常は50才時点の未婚率に変えている。40台後半と50台全般（前半）の未婚率を平均して産出）は毎年高まっていて、厚生労働白書2015では男性は24.2%、女性は14.9%だと推計されています。</p>	<p>후생노동 백서에 의하면, 이러한 일본의 생애 미혼율（생애 미혼자의 비율로 통상은 50세 시점의 미혼율로 대체하고 있다. 40대 후반과 50대 전반의 미혼율을 평균하여 산출）은 해마다 높아지고 있어, 후생노동 백서 2015에서는 남성은 24.2%, 여성은 14.9%이라고 추계되고 있습니다.</p>	◎		3/3	4/3	<p>助詞が連続するとその自然さが低下する。 例) 백서 2015에서는 남성은</p>
6	<p>これを分かりやすく言うと、日本人男性の4人に1人、女性の7人に1人は生涯未婚、おひとり様ということになります。</p>	<p>これを理解しやすく話せば、日本人男性に4人に1人、女性に7人に1人は生涯未婚、一人様ということになります。</p>	<p>이것을 이해하기 쉽게 말하면, 일본인 남성에게 4명에게 1명, 여성에게 7명에게 1명은 생애 미혼, 혼자님이라고 하게 됩니다.</p>	○	<p>「に」の訳が不自然。</p>	3/4	4/4	<p>助詞の翻訳が不適切。不自然。</p>
7	<p>経済界や世の中が「日本の英語教育は失敗だ」と断罪して、「使用できる英語」を要求する声に押されて、文部科学省は文法を軽視して形式だけの「会話に偏った英語教育」を行ってきたが、その結果、「融通性教育」による「授業時間と教育内容が大幅な削減」と一緒に、急激な学力低下を招いた。</p>	<p>経済界や世の中が「日本の英語教育は失敗だ」と断罪して、「使用できる英語」を要求する声に押されて、文部科学省は文法を軽視して形式だけの「会話に偏った英語教育」をしてきたが、その結果、「融通性教育」に従う「授業時間と教育内容が大幅な削減」と一緒に、急激な学力低下を招いた。</p>	<p>경제계나 세상이 「일본의 영어교육은 실패다」라고 단죄하고, 「사용할 수 있는 영어」를 요구하는 목소리에 밀려서, 문부 과학성은 문법을 경시해서 형식만의 「회화에 치우친 영어교육」을 해 왔지만, 그 결과, 「융통성 교육」에 따른다 「수업 시간과 교육 내용이 대폭적인 삭감」이라고 같이, 급격한 학력저하를 초래했다</p>	○	<p>「ゆとり」の訳が正しくない。 「招く」の訳が正しくない。</p>	3/2	3/3	<p>「融通性教育」をどこで切るかによって意味が全く変わる。「性教育」の意味にもなる。 「～と」の助詞の翻訳は不自然・不適切。</p>

(以下、白姫恩氏のコメント) 結論から言うと、日韓翻訳の精度は非常に高い。しかし、より精度を高めるためには、漢字系の語彙を多く使用する必要がある。例えば、日本語と韓国語に同じ単語が存在するとしても、その一つの単語がいくつかの意味として使用される場合は、不自然な翻訳になる可能性が高くなる。例えば、「招く」の場合、辞書には①손짓하여 부르다 (手招きする)、②불러오다 (呼んでくる)、③초대하다 (招待する) の3つの意味が掲載されている。③を直訳すると「招待する」という意味になり、3つの中で最も多く使用されるが、自動翻訳はこれを使う設定の可能性が高い。文脈的に②の方が自然な場合でも、自動翻訳では文脈に合わせた適切な選択はできない。また、「行」の場合、辞書には①가다 ②한 곳에서만 곳으로 움직여가다 ③떠나 (가)다とあり、主に「行く」の意味しか載せていない。日本語には「行く」の意味もあるのだが、今回の自動翻訳では漢字の「行」だけに反応してしまった。その単語が持つイメージをどのように分類するかという問題がある。例えば、日本語には「～のせい」と「～のため」という表現がある。「のせい」は主に否定的なイメージを持ち、「～のため」はそうとは言えない。一方、韓国語には「～때문에」という表現があり、どちらの意味合いでも使用されている。

なお、ロゴヴィスタ評価者の劉承桓氏も、「12年ぶり優勝」が「12년 만 우승」と訳される点については、「12年ぶり」で文が終わっていると認識されているようで「12년 만」となっているが、「優勝」に続ける場合は「12년 만에」とするべき。原文が「12年ぶり優勝」ではなく「12年ぶりに優勝」となっていれば「12년 만에」と訳出される。」と指摘している。そもそも、日本語の原文自体が「新聞記事の見出し」なので、普通の表現とは違うのだから、韓訳が不自然になるのも致し方ない。翻訳ソフトに非はない。

劉氏は、さらに、「おひとりさま」は書籍のタイトルなので訳出が難しいと思うが、「한명님」だと「一人の人」の意味になり、「한 사람」(独身者)の方が理解しやすい。また「サクセス」にサ変の「する」を付けて「サクセスする」という使い方が認識されず、「する」がそのまま訳出されてしまっている。「する」はなくても意味は通じる。「サクセスする」を「成功する」にした場合には「する」が訳出されない。」と指摘する。やはり、「サク

セスする」という言い方は、日本語として不自然な表現だ。「サクセス」は英語のカタカナ表記といった代物で、まだ一般に使われる日本語になっていないのだ。日韓翻訳で「サクセスする」という使い方が認識されないのは当然だ。劉氏は、このほか、「原文の「個人のという意」の「のという」の解釈が難しかったのか、「개인의고 말하는 뜻」は意味が通じない。原文が「「個人の」という意」であれば「「개인의」라고 말하는 뜻」という訳出になり意味が通じる。」と指摘する。実際、「個人のという意」は表記が拙い。原文の著者は「「個人の」という意」に改めるべきだろう。日韓翻訳ソフト側の問題ではない。

以上、二人の韓国人による翻訳結果の評価を見ると、評価の低い翻訳は、日本語の原文の特徴や不自然さ、表現法や表記の拙さに起因するもの、そして文化の違いによる表現の解釈や含意の違いに因るものが大半を占める。こうしたものは「翻訳ソフトの性能や守備範囲から外れる」ものだ。それ以外の言語的なレベルの過ちは限られる。日韓翻訳は、文法や語彙が極めて近い⁶だけに、日本文の特徴を良い部分も悪い部分も直接的に反映した翻訳にならざるを得ないのである。「日本語の装いをした韓国語」と言っても良い。かなり多様な日本語による検証だが、文化社会的な思考や慣習の違いが、韓訳の理解を妨げる表現以外は、ほとんどが理解可能なのである。日韓翻訳ソフトは使用を強く推奨すべき実用的で信頼に足る水準にあると言える。

6 文法語は活用や形態は異なるものの接続構造などは同じで、語彙は、基本的日常語は違うにしても、漢字を背景としたものは(中国からの語彙の導入時期や韓国語の音韻特性により読み方には規則的に若干の違いがあるが、)同じである。