

日中科学技術対訳コーパス作成における訳し分け用語・同形異義語の収集・分析

青木 雅子 蔣 葳 井上 聡子 菊池 俊一
 独立行政法人科学技術振興機構 研究基盤情報部

1 はじめに

日本と中国の間の経済交流および科学技術協力の進展に伴い、日中翻訳の需要が高まっている。科学技術振興機構は、科学技術振興調整費「日中・中日言語処理技術の開発研究」(代表機関:独立行政法人情報通信研究機構)に参画し、科学技術文献での用例ベース翻訳システムの開発研究の言語資源となる日中対訳コーパスを作成している。電子化された和文論文を学会の許諾を得て、それらを文単位に人手で中国語に翻訳している。

日本にも中国にも漢字語彙が大量に存在することは、翻訳において最大の「味方」でありながら、最大の「落とし穴」でもある。日中の漢字語彙は 1.同形同義、2.同形異義、3.異形同義、4.異形異義、の 4つのタイプに分けられ(王承伝¹⁾、翻訳における「同形異義語」の処理は訳質に大きく影響することはよく知られている。一般用語での日中同形異義語に関する書籍としては、基本的な 300 語を収録した『おぼえておきたい日中同形異義語 300』(光生館、1995 年、以下、同形異義語 300) や、1400 語を集めた『日中同形異義語辞典』(東方書店、2007 年、以下、同形異義語 1400) などがある。また、分野別で

は社会科学分野での日中同形異義語に関する研究(西谷まり²⁾)などが報告されている。

一方、科学技術分野における日中同形異義語に関する書籍および研究は少ない。我々は、対訳コーパスの作成過程において同形異義に起因する誤訳しやすい同形異義語を収集している。「同形同義語」と思われても、微妙な相違がある場合もあり、特に科学技術論文を翻訳する時に、分野によって訳し分ける必要がある例や、漢語以外に訳し分けが必要な外来語や和製漢語などの例もある。本稿では、これらの「同形異義語」と「訳し分け」に焦点をあてて分析した。

2 同形異義語の収集と分析

コーパス作成する際、「同形異義語」に対して、翻訳者は字面で日本語あるいは中国語の意味を憶測しがちであったり、共通の意味で解釈したりする傾向がよく見られる。我々は誤訳しやすい同形異義語を収集・分析し、4つのパターンに分類した。(1)日本語に他の意味がある同形語、(2)中国語に他の意味がある同形語、(3)意味の一部が共通しているが、それぞれ違う意味を持つ同形語、(4)意味・用法に共通点のない同形語。まず、この4種類の同形異義語の例を紹介する。

(1) **日本語に他の意味がある同形語** (「侵襲」と“侵袭”、「発生」と“发生”、「腐敗」と“腐败”など 42 語を収集。うち、同形異義語 300 にある同形語は 7 語、同形異義語 1400 には 16 語)。

例：侵襲

意味	訳語	分野	例文	訳文
【日・中】侵入し襲撃すること。	侵袭, 入侵	一般	外敵の 侵襲 を防ぐため、万里の長城を修築した。	为了防止外敌 入侵 ，修建了万里长城。
【日】医療において、生体内の恒常性を乱す可能性のある外部からの刺激のこと。外科手術、感染、中毒など。	创伤, 刺激, 发病, 发作	医学 生物	しかし、これらの方法で得られる麻酔域は、手術 侵襲 に耐えうる麻酔域を正確に反映しているとはいえない。	[正訳] 但是，并不能认为使用这些方法获得的麻酔区域是正确反映可承受手术 刺激 的麻酔区域。 [誤訳] 但是，并不能认为使用这些方法获得的麻酔区域是正确反映可承受手术 侵襲 的麻酔区域。

- (2) 中国語に他の意味がある同形語 (「失神」と“失神”、「研究」と“研究”、「傾向」と“傾向”など 5 語を収集。うち、同形異義語 300 にある同形語は 0 語、同形異義語 1400 には 4 語)。

例：失神

意味	訳語	分野	例文	訳文
【日・中】意識を失うこと。	失神	一般	心脏虽然保持正常的血压，但如果不能释放充分的血液，就会 失神 。	心臟が血圧を正常に保つのに十分な量の血液を送り出せなくなると、 失神 することがある。
【中】油断すること。 うっかりすること。	油断		作试验时仔细、精神集中非常重要，因为稍一 失神 可能就会失败。	[正訳] 実験をする際、慎重さと集中することが大事で、ちょっとした 油断する と、失敗するおそれがある。
				[誤訳] 実験をする際、慎重さと集中することが大事で、ちょっとした 失神する と、失敗するおそれがある。

- (3) 意味の一部が共通しているが、それぞれ違う意味を持つ同形語 (「処置」と“処置”、「装着」と“装着”、「伝達」と“传达”など 10 語を収集。うち、同形異義語 300 にある同形語は 4 語、同形異義語 1400 には 5 語)。

例：処置

意味	訳語	分野	例文	訳文
【日・中】状況などを勘案して扱いを決める。処理する	处理, 处置	一般	早急に 処置 する。	及早 处理 。
【日】けがや病気の治療・手当をすること。	治疗	医学	日常生活で起こりやすい事故やけが、あるいは突然の発熱や腹痛・めまい・嘔吐など、さまざまな病気に関わる症状が現れたときの応急 処置 をまとめた。	[正訳] 归纳了在日常生活中容易发生的事故及伤害、或者突发性的发烧以及腹痛、头晕、呕吐等各种各样的疾病症状出现时的急救 治疗 。
				[誤訳] 归纳了在日常生活中容易发生的事故及伤害、或者突发性的发烧以及腹痛、头晕、呕吐等各种各样的疾病症状出现时的急救 处置 。
【中】処罰すること。	処罰	一般	对于违反校规的人，必须严厉 处置 。	[正訳] 校則を破る者に対しては、 嚴重に処罰 すべきである。
				[誤訳] 校則を破る者に対しては、 嚴重に处置 すべきである。

- (4) 意味・用法に共通点のない同形語 (「發現」と“发现”、「温存」と“温存”、「難聴」と“难听”など 40 語を収集。うち、同形異義語 300 にある同形語は 10 語、同形異義語 1400 には 31 語)。

例：発現

意味	訳語	分野	例文	訳文
【日】表面に現れ出ること。 顕現。	呈现， 显现	一般	局所作用を目的とする医薬品の薬効を 発現 させるためには、それを循環血液中に移行することが必要である。	对于以局部作用为目的的医药用品，为了 显现 其药效，必须将其运送到循环血液中。
【日】遺伝子の情報が、細胞における構造および機能に変換される過程のこと。	表达	遺伝学	ゲノムの機能の基本は、複製による遺伝情報の伝達と、転写による遺伝情報の 発現 からなる。	[正訳] 基因组功能的基础由基于复制的遗传信息传递和基于转录的遗传信息 表达 所组成。 [誤訳] 基因组功能的基础由基于复制的遗传信息传递和 表现 基于转录的遗传信息所组成。
【中】見つける。発見する。気づく。	発見	一般	我 发现 他们并没有深刻地认识到问题的严重性。	彼らが問題の深刻さを肌で感じていないのに 気づいた 。

3 訳し分けの収集と分析

同じ用語であっても、一般文献と科学技術文献ではそのニュアンスが異なり、訳し分けをす

る場合がある。我々は訳し分けを必要とする用語を3種類に分けてみた。(1)同形同義語、(2)外来語、(3)中国語にはない和製漢語。

(1) 同形同義語（「挙動」と“举动”、「過渡」と“过度”、「結合」と“结合”など41語を収集）

一般文献を翻訳する際に、「同形同義語」の処理は極めて簡単で、漢字の字体を変換するだけで殆どの場合には問題が生じない。しかし、翻訳対象が科学技術文献になると、同じ方法では用語のニュアンスを正確に訳出できない恐れがある。以下に、「挙動」の訳し分け例を紹介する。

意味	訳語	分野	例文	訳文
立ち居振る舞いのこと。	举动，行为，举止	一般	今朝、近所で 挙動 不審の男を見かけました。	今天早上，在家附近发现了 举动 可疑的男人。
反応する。 変化する。	反应，变化	化学・物理 など	通常のウイルス感染で発現する遺伝子産物の 挙動 と、単独発現系で認められた遺伝子産物の 挙動 を、単純には比較できないが、両者で得られた結果のギャップを埋めるために、現在、ウイルスゲノムの 挙動 についても各魚類細胞で解析を進めている。	[正訳] 虽然无法将在一般的病毒感染中所表现的基因产物的 反应 与在单独表达系统中发现的基因产物的 反应 作简单比较，但是为了填补两者所得结果中的差距，目前，正使用各种鱼类细胞，对有关病毒基因组的 反应 进行分析。 [誤訳] 虽然无法将在一般的病毒感染中所表现的基因产物的 举动 与在单独表达系统中发现的基因产物的 举动 作简单比较，但是为了填补两者所得结果中的差距，目前，正使用各种鱼类细胞，对有关病毒基因组的 举动 进行分析。

(2) 外来語

日本語には欧米諸国から入ってきた語、即ち「外来語」もたくさん存在する。科学技術文献中に出現する用語において、物質名・医学用語・生物用語などに、「外来語」が占める割合が極めて高い。通常、英語を経由することで、訳語を簡単に探し出せる。しかし、分野によっては、訳し分けを必要とする場合もある。以下に、「マトリックス」の訳し分け例を紹介する。

意味	訳語	分野	例文	訳文
行列	矩阵, 队列	数学など	現在は、伝達 マトリックス 法を用いて、棒内の任意位置における応力を算出する解析方法を研究している。	目前采用传递 矩阵 法, 对计算棒内任意位置的应力解析法进行研究。
母体、基盤、鋳型	基质, 基体	生物・医学など	癌細胞が湿润・転移を起こす場合も、マトリックス成分を分解する必要があります。	[正訳] 在发生癌細胞湿润、转移时, 也必须对其 基质 成分进行分解。
				[誤訳] 在发生癌細胞湿润、转移时, 也必须对其 矩阵 成分进行分解。

(3) 中国語にはない和製漢語

「和製漢語」とは、漢字を利用し、漢語らしく造った日本発の漢語である。既に中国に逆輸出されたものも少なくない。一方、中国語には存在しない「和製漢語」もある。一般文献と科学技術文献で、その意味が異なる場合には訳し分けをする必要がある。「嫌気」の訳し分け例を紹介する。

意味	訳語	分野	例文	訳文
何となく嫌だという気持ち。	厌烦, 厌倦	一般	互いの未熟さに 嫌気 がさす。	对相互之间的不成熟感到 厌烦 。
無酸素条件のこと。	厌気	環境・エネルギーなど	前処理設備で異物が除去された生ゴミは、可溶化設備ではほぼ同量の水を加えてスラリー化され、 嫌気 的雰囲気のメタン発酵槽内へ送られる。	[正訳] 对于在前处理设备中去除了异物的厨房垃圾, 在可溶化设备中加入大致等量的水分, 制成浆液, 然后送入 厌気 性环境的甲烷发酵槽内。
				[誤訳] 对于在前处理设备中去除了异物的厨房垃圾, 在可溶化设备中加入大致等量的水分, 制成浆液, 然后送入 厌気 性环境的甲烷发酵槽内。

4 結論

「同形異義語」を正確に処理できること、及び「訳し分け」ができることは、科学技術分野の翻訳品質の要といっても過言ではない。「日中・中日言語処理技術の開発研究」プロジェクトで、我々は、収集・分析している同形異義語と訳し分けが必要な用語を翻訳者に提供し、注意喚起を促すことで、翻訳品質の向上をはかっている。また、これらの情報を同プロジェクトで日中・中日翻訳辞書を作成している東京大学および静

岡大学に提供するとともに、翻訳システムを構築中の独立行政法人情報通信研究機構および京都大学にも提供することで、翻訳システムの性能評価への活用も期待できる。

参考文献

- (1) 王承伝 (1998) 「同形異義語における中国語と日本語の対照研究」『人文科学教育研究』25, p.143-152.
- (2) 西谷まり (2004) 「社会科学分野で使われる日中同形異義語」『一橋大学留学生センター紀要』7, p.15-28.