

広島大学学術情報リポジトリ  
Hiroshima University Institutional Repository

Title	初・中級学習者用8言語例文パラレルコーパスの開発
Author(s)	岩崎, 克己
Citation	広島外国語教育研究 , 25 : 181 - 198
Issue Date	2022-03-01
DOI	
Self DOI	<a href="https://doi.org/10.15027/51969">10.15027/51969</a>
URL	<a href="https://ir.lib.hiroshima-u.ac.jp/00051969">https://ir.lib.hiroshima-u.ac.jp/00051969</a>
Right	Copyright (c) 2022 広島大学外国語教育研究センター
Relation	



# 初・中級学習者用8言語例文パラレルコーパスの開発

岩崎 克己

広島大学外国語教育研究センター

## 0. はじめに

筆者をはじめとする広島大学の初修外国語担当教員グループでは、日・英・独・仏・西・露・中・韓、8言語の初・中級学習者を対象とした例文パラレルコーパスを開発中である。またこのコーパスとその検索結果を利用した教材作成システムを開発し、両者をインターネット上で公開し、広く提供することを計画している。本稿では、本計画の概要と過去2年間における開発状況について報告するとともに、パラレルコーパス作成の際に生じる様々な課題について論じる<sup>1)</sup>。

## 1. 計画の概要

### 1.1. 概要

グローバル化が進み、多言語・多文化が共生する現代社会においては、国際共通語としての英語の教育に加え、それ以外の言語の運用能力を持った学習者をも一定教育で、多様性のある社会の実現に貢献していくことが重要である。本計画では、そうした多様性を持った言語学習や多言語教育に資するため、CEFR（ヨーロッパ共通言語参照枠）におけるA1からB1レベルに相当する初・中級段階の学習者でも使用可能な日・英・独・仏・西・露・中・韓の8言語例文パラレルコーパスを開発し、これをインターネット上の学習支援用コーパスとして公開し、国内外の外国語学習者および外国語教育関係者に提供することを目指している。作成規模は、各言語同じ内容の例文を約14,000セット、8言語合わせて約112,000文、含有語数は約100万語である。8つのどの言語からでも検索でき、例文の出力言語は、使用者が必要とする言語の組みあわせに応じて最大8言語まで自由に変えられるようにする。また、学習者が本コーパスを自律的な外国語学習のツールとして使えるよう、検索結果を自由に加工して利用するための諸機能や、検索結果を基に教員が授業用の様々なワークシートを作成・印刷できる教材作成機能を加える計画である。図1は現在作成中の検索システム（学習者用のインターフェース）の検索結果のデモ画面の一部である<sup>2)</sup>。

これまで、個別言語のコーパスは、言語学習に利用される場合も、検索結果を十分に読解できる力を



図1：検索例 ドイツ語のBuchで検索 8言語出力

持った上級レベルの学習者用検索ツールか教員による教材作成用ツールとしてのみ考えられてきた(岩崎, 2010)。しかし, 今回開発中の8言語例文パラレルコーパスでは, 作成する例文の語彙や難易度をあらかじめ下げて初学者向けに調整するだけでなく, 多言語型教材の特性を生かして母語や既習言語の手がかりも利用可能にする。このため, 本コーパスを使うことで, 文脈から意味を読み取ったり, 作文のための表現例を搜したり, 単語の使い方を調べたりする自己発見型学習やデータ駆動型学習を, 初・中級段階の学習者においても, 多くの言語で実施できるようになる。特に, コミュニケーションで多用される高頻度の用例をコロケーション(イディオムを含めた語の配列)の形で列挙して提供することにより, オンライン学習においてこれまで, 重点が置かれてこなかった表現(作文・発話)能力の向上にも寄与することが期待される。

## 1.2. コーパスの検索システム

本計画で扱う言語には, 中国語のような孤立語だけでなく, 日本語, 韓国語のような膠着語や, ドイツ語, フランス語, ロシア語等の英語と比べると語形変化の豊かな屈折語が含まれる。そこで, キーワードのみにマッチングする単純な検索だけでなく, 語幹の前後の接辞や語中の任意の文字列, さらには活用語尾も含めて検索できるワイルドカード検索, 複数のキーワードを使ったAND/OR検索, 除外すべき単語をあらかじめ指定した条件付き検索, あるいは, ドイツ語の分離前つづりのように, 熟語を形成する単語群が非連続の形で現れる非連続構成素に対応した検索など, 高度な検索設定を可能とした<sup>3)</sup>。将来的には, 例文の作成に使ったキーワードについては, すべての言語に関し, CEFRのレベル(A1からC2)に関する情報, 見出し語(レマ)の形と全活用形, 簡単な意味等を記載した辞書部門を組み込み, 使用されている単語のレベル別強調表示も可能にする予定である。

なお, 各言語の例文データと検索部門は完全に分離し, 例文データとそれと関連づけられた辞書情報さえあれば, 今後も他言語への拡張が可能な仕様としている。

## 1.3. 検索データを利用した教材作成システム

検索データを利用した教材作成システムとしては, 以下の機能を構想している。

- 1) 出力言語の切り替え(8言語から必要な言語を自由な組み合わせで複数表示)
- 2) 検索語の表示・非表示切り替え
- 3) 検索結果のソート
- 4) 検索結果の一部削除
- 5) 先行する検索結果に新たな検索結果を付け加えていく出力例文データの蓄積
- 6) 出力例文内に含まれる登録キーワードの色分けによるレベル別強調表示
- 7) 検索結果の読み上げ
- 8) 検索語の左右の文脈を分割・シャッフルして作った並べ替え問題とその正解確認シートの作成
- 9) 出力された例文の文脈を手がかりにして非表示になった検索語を当てるクイズ問題とその正解確認シートの作成
- 10) 検索結果の中の任意の単語(群)を非表示にしたクローズ・テスト(穴埋め問題)とその正解確認シートの作成
- 11) 表計算ソフト等へのデータの書き出しや印刷などの形によるデータの出力・保存

なお、これらの詳細については、稿を改めて別途論じる予定なのでここでは扱わない。

## 2. 計画初年度の作業の進捗状況

### 2.1. 各言語の例文作成方法

本計画の出発点としては、過去に筆者らが開発した「ドイツ語学習者用独日例文パラレルコーパス DJPD」(岩崎, 2005; Iwasaki, 2005)がある。本計画では、このコーパスの日独の例文約14000個を出発点に、対象言語を2言語から8言語に拡張する形で約112,000個の例文作成作業を、2年掛けて行ってきた。例文作成者のプロフィールは以下の通りである<sup>4)</sup>。

英 語：ドイツ語がA2レベルのイギリスで生まれ育った日本人学部生1名、英語とドイツ語がネイティブレベルのドイツ人学生1名  
中 国 語：日本語がA2からB1で英語がB1からB2レベルの中国人民大学院生2名  
韓 国 語：日本語B2からC1で、英語がB1からB2レベルの韓国人大学院生3名  
スペイン語：英語がB2からC1レベルの日本で生まれ育ったスペイン人学生2名  
フランス語：英語とドイツ語を理解し、日本在住20年以上のフランス人教員1名  
ロシア語：英語を理解し、日本在住20年以上のロシア人教員1名、

各言語の例文は、彼らの協力を経て、主に以下の手順で作成した。

#### 第1段階

日本語とドイツ語：ドイツ語学習者用独日例文パラレルコーパス DJPD の例文を再チェックし一部修正する<sup>5)</sup> 形で利用



#### 第2段階

英語：日本語とドイツ語の両方の例文を参考に、他言語に先行して作成



#### 第3段階

中国語：主に日本語と、副次的には英語の例文を基に作成  
韓国語：主に日本語と、副次的には英語の例文を基に作成  
スペイン語：主に英語・日本語と、副次的にはドイツ語の例文を基に作成  
フランス語：英語・日本語・ドイツ語の例文を基に作成  
ロシア語：主に英語と、副次的に日本語の例文を基に作成

例文作成のために、まず、作成時に参考とする英語・日本語・ドイツ語の例文データを各々300個ずつに区切って300行1列ずつにまとめたシート47枚からなるエクセルファイルを作成した。これを例文作成協力者に渡し、その例文データの横の列に、コミュニケーション上ほぼ同じ機能を持つそれぞれの言語の例文を作成し書きむよう依頼した。

## 2.2. 作成した例文の検討・調整状況

次に、それぞれの例文のミスを訂正すると共に、8言語の例文間に内容的に大きなずれや誤りが無いかをチェックし、相互調整していく作業を行った。この作業は、もともとは複数の作成者が顔を合わせて一文ずつ問題点を吟味していく対面による会議方式を進めることを考えていたが、2020年春より始まったコロナ禍のために対面での作業を行うことがむずかしくなった。そこで、こうした作業をオンライン上で行うための編集専用のインターフェースを、学習者用のインターフェースとは別に、急遽作成した。

具体的には、作成例文をひとまずデータベースに登録し、図2のような例文一覧表示機能と図3のような各例文セットの表示・編集機能を付けた。



図2：例文一覧表示画面 独日2言語出力の場合



図3：例文セット表示画面 問題番号36番の編集画面

これを使うことで、リモートによる以下の4種の編集作業が可能になった。

第1に、各言語の個々の例文右横に修正ボタンを付け、例文作成者が提案する修正案を書き込む欄とその理由を必要に応じて書き込めるコメント欄を作ることで、個々の例文作成者による言語ごとのオンラインでの修正提案を可能にした。

第2に、各例文セットに通しのID番号を振り、例文セット全体に関してコメントする機能を付けることで、例文セットごとに言語間を越えて必要な指示を出したり、問題点を共有したりすることを可能にした。もともと、ドイツ語と日本語の2言語の例文が先行して存在し、それを基に英語を作り、その後、この3言語の例文を基に、さらに他の言語の例文を作成するという方式を取ったため、最初の英独日の例文にミスや誤訳等があった場合は、他の言語の例文すべに問題が及んだ。この機能は、特にそうした事態に対処するために必要となった。

第3に、スーパーユーザである責任者だけが利用可能な機能として、各言語の例文作成者が行った修正提案に対し、最終的に「採用」あるいは「不採用」の判断を下し、必要に応じて、データベースの内容を更新する機能を付けた。これにより、例文データの一元管理を可能にした

第4に、すべての変更は履歴を残し、各言語の例文作成者がいつでも見られるようにすると共に、スーパーユーザは履歴の残った過去の変更段階にいつでもデータベースの内容を戻すことができるようにした。

上記の機能を備えた編集専用のインターフェースを使い、オンライン会議で画面共有しながら、各言語の例文作成者や編集者<sup>6)</sup>があらかじめ書き込んだ編集提案を吟味したり、その場の議論をもとにリアルタイムで例文を変更したりする形で、オンラインによる同期型および非同期型の例文編集を行った。この編集専用のインターフェースは、学習者用のインターフェースとは独立しているため、今後、公開後も、データベースに含まれる例文の品質向上作業は並行して進めることが可能である。

### 3. 8言語例文パラレルコーパス作成上の課題

言語表現を、文脈から切り離された一例文として抽出する以上、多義性や曖昧性の問題が生じることは避けて通れない。あるひとつの名詞を扱う場合でも、それが何を指しているかという指示対象のレベルにとどまらず、名詞として成立するために最低限表示しなければいけない情報量のレベル、それがどのようなアプローチで外的世界を分節しようとしているかという概念化のレベル、どのような意図や行為と結びつくのかという言語機能のレベル等が有り、そのいずれにおいても言語ごとに相違があるため多義性や曖昧性が生じる。作成の出発点となった独日例文コーパス DJPD の場合には、両言語の間で調整するだけで良かったが、ほぼ同じ「内容」あるいは「機能」を持つべき8言語の例文の間の相互調整は、収束させるべき意味の中心をどこに置くかという観点から見た場合、8言語すべてを等しく満足させるようなひとつの正解がない場合が多い。また、各言語の例文作成に関わったメンバーが担当言語以外に共通して理解できる言語が多くの場合、英語や日本語に限られたことも、新たな問題を生んだ。具体的には、オリジナルのドイツ語例文をいったん英語や日本語にした際に、失われた情報が他の言語での例文作成の際に再び必要になるケースが多数生じたからである。この1年間の例文の調整作業を通して見えてきた課題は、以下の点である。

#### 3.1. 2人称代名詞における親称と敬称の区別

英語以外の欧米系言語には、英語の you に当たる2人称の代名詞に、親族・友人・恋人等の親しい間柄で使う親称と疎遠あるいはフォーマルな間柄で使う敬称との違いがある(表1参照)。ドイツ語の授業等で日本語に訳す際は、便宜的に親称を「きみ」、敬称を「あなた」などとする慣習があり、また訳文全体で親しさや、よそよそしさを表現したりしているが、今回は、フランス

語、スペイン語、ロシア語の例文を、英語や日本語の例文を媒介として作成したため、もともとの親称のドイツ語例文が敬称で表現されたケースや、その逆のケースもあり、逐一チェックが必要であった。これは、単に主語と動詞の活用形の変更だけにとどまらず、言語使用の社会的場面が異なるため、文全体のニュアンスにも影響を与え、共起する副詞の選定や文の丁寧さを表示する語彙や文法形式等も含む全般的な調整が必要となった。

表 1：各言語の2人称の代名詞

主格\言語	英語	ドイツ語	フランス語	スペイン語	ロシア語
親称 (単数)	you	du	tu	tú	ТЫ
親称 (複数)		ihr	vous	vosotros / vosotras	ВЫ
敬称 (単数)		Sie	vous	usted	
敬称 (複数)				ustedes	

さらに、以下の ID 番号1089の例文が示すように duzen (= du で話す) という動詞を使って、親称や敬称の使用自体をテーマとして取り扱った例文も存在した。こうした例に対しては、親称や敬称の使い分けがない言語の場合にも、英語・日本語・韓国語のようにそのままドイツ語の du を使って表現しつつ英語の例のように du についての解説を付け足す方法や、あるいは中国語のように当該言語内部での2人称表現の使い分け<sup>7)</sup>を援用する等の方法が考えられる。しかし、学習という観点では、敬語とは異なって主に親疎を表す親称・敬称の使い方の情報や、それらの違いが、ファーストネームとファミリーームの使い分け等と関連することなども言及したいところである。そのような判断から、たとえば、日本語訳文を「どうぞファーストネームで呼んでください。」と意識する方法も考えられた。しかし現状では、追加的措置として解説用の補助ページを別途作り、そこへ跳ぶリンクを設けることを前提に、以下の ID 番号1089番の例文で示すような対応に留めた。

言語	ID 番号1088の例文
日本語	どうぞ du で話し掛けてください。
ドイツ語	Sie dürfen mich gerne duzen.
英語	Feel free to use “du” with me. (du: the familiar form of “you” in German)
フランス語	Tu peux me dire “tu”.
スペイン語	Por favor, llámame de “tú”.
中国語	请用‘你’开始说话。
韓国語	부디 ‘du’ 로 말을 걸어 주세요.
ロシア語	Пожалуйста, обращайтесь на “ты”.

### 3.2. 時制やアスペクトにおける差異

各言語で時制やアスペクトとそれを表す文法手段に相違があるだけでなく明示される情報量にも差がある。たとえば、表2の時制の比較表<sup>8)</sup>が示すように、英語の現在完了形の形式は、「経験、完了、結果、継続」など多様の機能を持つが、英語では、現在完了で表現される「(過去から現在までの)継続」は、ドイツ語やフランス語では現在形で、また、ドイツ語でも現在完了形を用

表2：時制の比較表

機能\言語	英語	ドイツ語	フランス語	スペイン語	ロシア語
時間的前後関係	過去完了形	過去完了形	大過去 (複合過去の過去)、 前過去 (単純過去の過去)	過去完了	反復：不完了体過去、 その他：完了体過去
過去 (話者から切り離された過去の事実の提示)	過去形	過去形	単純過去	点過去	反復：不完了体過去、 その他：完了体過去
過去進行中の動作 (同上)	過去進行形	過去形	半過去	線過去, <i>estar</i> 動詞の線過去 + 現在分詞	不完了体過去
過去の一時点における状況 (同上)	過去形	過去形	半過去	線過去	不完了体過去
過去 (話者と心理的に結びついた過去の説明)	過去形	現在完了形	複合過去	現在完了 (スペイン語の多くのバリエーションでは点過去が使われることもある)	反復：不完了体過去、 その他：完了体過去
過去進行中の動作 (同上)	過去進行形	現在完了形	半過去	線過去, <i>estar</i> 動詞の線過去 + 現在分詞	不完了体過去
過去の一時点における状況 (同上)	過去形	現在完了形	半過去	線過去	不完了体過去
完了	現在完了形	現在完了形	現在	現在完了	完了体過去
結果	現在完了形	現在完了形	現在	動作：現在完了, 状況：現在	完了体過去
経験	現在完了形	現在完了形	複合過去	現在完了	不完了体過去
継続	現在完了形	現在形	現在	現在	不完了体過去
現在の事実	現在形	現在形	現在	現在	不完了体現在
普遍的真理	現在形	現在形	現在	現在	不完了体現在
現在における反復する行為・事象	現在形	現在形	現在	現在	不完了体現在
現在進行中の動作	現在進行形	現在形	現在	<i>estar</i> 動詞の現在 + 現在分詞	不完了体現在
近未来	未来形	現在形	<i>aller</i> + 動詞の原形	現在, 未来, <i>ira</i> + 不定詞	反復：不完了体未来、 その他：完了体未来
未来	未来形	未来形	単純未来	未来, 確定的：現在	反復：不完了体未来、 その他：完了体未来
未来進行中の動作	未来進行形	未来形	単純未来	<i>estar</i> 動詞の未来 + 現在分詞	不完了体未来
未来の一時点での完了	未来完了形	未来完了形	前未来	未来完了	完了体未来
過去から見た未来一時点	未来形	仮定法過去 ( <i>would</i> )	条件法現在	過去未来	過去形 + <i>был</i>
過去から見た未来の一時点での完了	未来完了形	仮定法過去完了 ( <i>would</i> )	条件法過去	過去未来完了	過去形 + <i>был</i>

いる「完了」や「結果」は、フランス語では現在形で表現される。他方、英語ではこれから起きる事象の記述には原則としてすべて未来形が使われるが、ドイツ語では「近未来」であれば現在形と適切な副詞で問題なく表現できる。またフランス語でも「近未来」は *aller* + 原形という形で語彙的な補助手段を使って表現される。このように、ひとつの言語では多義性が許されても、その解釈により他の言語では、採用する時制や表現形式自体がそれぞれ大きく変わってしまうことはよくある。そのため、副詞句などを加えて、もともとのドイツ語や英語例文の解釈の可能性を限定し、それに合わせて他言語の訳を調整することも必要になった<sup>9)</sup>。

このほか、日本語のタ形 (例：食べた) は過去の事象を記述する場合にも使われるが、完了相を表す場合には現時点の事象やこれから起こる事象の記述にも使用可能である。そのため、英語例文を作る際に完了相のタ形を間違えて過去形で訳してしまい、他言語の表現を含めてひとつの例文セット全体の例文の修正が必要となるケースも見られた。

### 3.3. 意味表示としての抽象的表現の許容度における言語間の差異

文レベルの表現には、その文が可能性としてどのような命題内容を表示し得るかという意味表示のレベルと、実際にその意味表示を特定の文脈や場面と結び付けて理解する意味解釈のレベルがある。人は通常、ひとつの文を見るとき、単独の例文であっても、それに自然な意味を付与し得る文脈や場面を半ば無意識に想定して意味解釈まで行っているが、時々例文だけからは、すぐには、適切な文脈を思い浮かべられず、意味を指し示す可能性を持った意味表示としてしか、あるいはせいぜい状況記述としてしか理解できないケースがある。特に現在形の状況記述の文においては、日本語やドイツ語に比べて、英語の場合は文としての許容度が低いようである。以下に

例文データベースの中から ID 番号と共にいくつか例を挙げる<sup>10)</sup>が、これらに関しては、英語の母語話者の一部から、使用できる自然な文脈が見当たらないので、このままでは違和感があるとの指摘があった。これらのいくつかは、英語の現在形の使用範囲がドイツ語などの他言語に比べて狭いことに起因する。たとえば、ID 番号168/284/1063などの例文は、ドイツ語でも進行形の

表3：問題となったケースの一例（部分）

168	He hangs on the roof with only one hand. → He is hanging on the roof with only one hand.	進行形に変更
	彼は片手で屋根にぶら下がっている。	
	Er hängt mit einer Hand am Dach.	
185	I pack another bottle of wine, just in case. → I'll pack another bottle of wine, just in case.	助動詞 will を付加
	念のため、さらにワインを1本入れておこう。	
	Ich packe vorsichtshalber noch eine Flasche Wein ein.	
244	My husband and my son argue about politics the entire night. → My husband and my son often argue about politics the entire night.	全言語に関し、頻度の副詞を付加
	私の夫と息子は、一晩中政治について議論する。 →私の夫と息子は、よく一晩中政治について議論する。	
	Mein Mann und mein Sohn diskutieren die ganze Nacht über Politik. → Mein Mann und mein Sohn diskutieren oft die ganze Nacht über Politik.	
284	The kids shake the tree and apples fall to the ground. → The kids are shaking the tree, and apples are falling to the ground.	進行形に変更
	子供たちが木をゆらして、リンゴが落ちる。	
	Die Kinder schütteln den Baum und Äpfel fallen herunter.	
388	The students protest against tuition fees. 学生たちは授業料に対し抗議している。	
	Die Studenten protestieren gegen Studiengebühren.	
496	The mailman runs away from the dog. 郵便配達人はその犬が怖くて走って逃げる。	
	Der Postbote rennt vor dem Hund davon.	
650	The people stand up and applaud vigorously. 人々は立ち上がり、大きく拍手喝采する。	
	Die Leute stehen auf und applaudieren lautstark.	
885	Do you build your house by yourself? → Will you build your house by yourself?	助動詞 will を付加
	あなたは自分で家を建てるのですか。	
	Bauen Sie Ihr Haus selbst?	
1034	Do you visit Spain again for a vacation? → Will you visit Spain again for a vacation?	助動詞 will を付加
	あなたはまたもやスペインで休暇を過ごすのですか。	
	Machen Sie schon wieder in Spanien Urlaub?	
1063	Karl wears a fancy shirt today. → Karl is wearing a fancy shirt today.	進行形に変更
	カールは今日しゃれたシャツを着ている。	
	Karl hat heute ein schickes Hemd an.	

な解釈が普通なので、英語ではそれを明示してやれば良い。また、ID 番号185/885/1034などの例文では1人称や2人称の現在形が使われており、ドイツ語でも主語の意志を含んだ解釈が普通なので、英語ではそれを明確にするため、willなどを付け加えた。その他、「反復する行為・事象」という解釈を明確にして文の許容度を上げるため、英語だけでなく他言語を含めて、頻度の副詞（英語の例で言うと always/often 等）を補う扱いもした。ID 番号244などがその例である。しかしながら、状況記述として理解できるケースに関しては、ID 番号388/496/650などのように、英語話者の一部が違和感を感じるケースでもそのままとした。

### 3.4. 名詞的表現において義務的に明示される情報量の差異

名詞的表現においては、言語により義務的に明示される情報とその量に差がある。その典型は、以下に挙げるような親族表現である。

表4：各言語の親族表現（部分）

言語	対応する表現	対応する表現
日本語	兄・弟	おじさん、叔父さん（父母の弟または父母の妹の夫）、 伯父さん（父母の兄または父母の姉の夫）
ドイツ語	Bruder	Onkel
英語	brother	uncle
フランス語	frère	oncle
スペイン語	hermano	tío
中国語	哥哥・弟弟	親族名称：叔父（父の弟）、伯父（父の兄）、姑父（父の姉・ 妹の夫）、舅父（母の兄・弟）、姨父（母の姉・妹の夫）、 親族の呼称：叔叔（父の弟）、伯伯（父の兄）、姐夫（父の姉・ 妹の夫）、舅舅（母の兄・弟）、姨夫（母の姉・妹の夫）
韓国語	형（弟から見た兄）・ 오빠（妹から見た兄）・ 동생（弟）	삼촌（父の兄・弟）、큰아버지（父の兄）、작은아버지（父 の弟）、고모부（父の姉・妹の夫）、외삼촌（母の兄・弟）、 이모부（母の姉・妹の夫）
ロシア語	брат	дядя

名詞レベルの解釈の多様性に関しては、一般的な対応として、言語ごとに必要に応じて複数の可能性を列挙する方式を取った。しかし、上記の uncle の場合、欧米系言語の1つの表現に対し、中国語では10個の表現が対応し、すべて挙げるのは事実上不可能である。そこで、複数の言語の表現を連携させて、解釈の可能性を数個に絞って列挙する方針を取った。たとえば、以下の ID 番号4766の例文がその例であるが、まず日本語例文であえて「おじ」ではなく漢字の「叔父」（父母の弟または父母の妹の夫）を使うとともに、この例では、中国語の「親族の呼称」に限ることで、中国語での表現を4個、韓国語での表現を5個に絞った。ここで、除外された「伯父さん」（父母の兄または父母の姉の夫）の例や、「親族の呼称」ではなく「親族名称」の方を使う例に関しては、この例文コーパス全体の中での uncle の出現数を調べ、他の例文では、意図的にそちらの解釈を採用することで、全体としてすべての例がまんべんなく登場するよう配慮した。なお、アジア系言語における親族名称等の理解がむずかしいケースについては、既に述べた親称・敬称に対する対応と同じく、わかりやすく説明する解説用の補助ページを作り、そのページへのリンク

を設ける予定である。

言語	ID 番号4766の例文
日本語	私の叔父は物理学の教授だ。
ドイツ語	Mein Onkel ist Professor für Physik.
英語	My uncle is a professor of physics.
フランス語	Mon oncle est professeur de physique.
スペイン語	Mi tío es profesor de física.
中国語	我叔叔(姐夫・舅舅・姨夫)是物理学教授。
韓国語	작은아버지(삼촌・고모부・외삼촌・이모부)는 물리학 교사이다.
ロシア語	Мой дядя - профессор физики.

この他にも、性別や数に関わる情報など、例文作成において問題となるケースは多い。たとえば、「同僚・従業員」という概念は英語でも他のアジア系言語でも性を意識せずに使えるが、ドイツ語をはじめとする他の欧米系言語では、Mitarbeiter ([男性の] 同僚) / Mitarbeiterin ([女性の] 同僚) のように、たとえそこに文の焦点が無くても性の情報が自然に入ってしまう表現がある。そのため、他言語に訳す際、欠けていた性の情報を形容詞などで補うと、わざわざ加えたことで、意図せず文の焦点がそこに移ってしまうことがある。ID 番号1200の例文がその例だが、ドイツ語例文では、「新しく来た同僚に仕事を教えてほしい」ということが伝達すべき情報の中心にあり、その対象がたまたま女性であったに過ぎないと解釈するのが普通だが、英語の場合、女性という性別情報にもかなりの重点が置かれているように見える。日本語ではそれを避けるため、直訳調の「女性の同僚」ではなく「同僚の女性」と訳して「同僚」の方に焦点が行くように工夫しているが、それがどの程度功を奏したかは定かではない。これと同じ問題は、形容詞型名詞でも見られる。ドイツ語の「大人」は過去分詞 erwachsen から派生してできた形容詞型名詞であるが、定冠詞を付けたケースに限っても、冠詞と語尾の組み合わせで die Erwachsene (大人の女性) / der Erwachsene (大人の男性) / die Erwachsene (大人たち) の3通りがある。adult/adults という単複の違いしかない英語や、「大人」という表現しかない日本語でこれらの付随情報をどの程度再現すべきかは正解のない問題である。

言語	ID 番号1200の例文
日本語	新しい同僚の女性に仕事を教えてもらえますか。
ドイツ語	Können Sie bitte die neue Mitarbeiterin einschulen?
英語	Could you train the new female employee, please?
フランス語	Pouvez-vous former la nouvelle collègue, s'il vous plaît ?
スペイン語	¿Puedes enseñar el trabajo a la nueva empleada?
中国語	你能教新来的女同事关于工作上的事情吗？
韓国語	새로운 동료 여성에게 일을 가르쳐 주시겠습니까？
ロシア語	Вы не можете дать пояснения по нашей работе нашей новой сотруднице.

### 3.5. 使用される文型の言語による差異

ドイツ語の man などの不定代名詞で表現される欧米言語の文は、日本語でもよく受け身的に訳されるが、ひとつの言語で受動態で表現されても他の言語では、能動態で表現されるなど、言語ごとに、好まれる形式にばらつきが多い。たとえば、英語以外の言語では、通常、すべて受動態で表現される文も、英語では一般的な人々を表す they を主語として表現することが多い。ID 番号386の例文は、その一例だが、このような事例で文型の統一を図るべきか、表現の自然さや多様性を考えて違う扱いをするべきか、また例文コーパス全体で似たような事例が数多くある場合、多様なバリエーションがまんべんなくコーパスに含まれるようどう調整するかなど、考えるべきことが多い。

言語	ID 番号386の例文
日本語	彼には癌摘出の手術がなされた。
ドイツ語	Ihm wurde der Krebs operiert.
英語	They operated on his cancer.
フランス語	Il a été opéré d'un cancer.
スペイン語	Le hicieron una cirugía para extraer el cáncer.
中国語	他做了癌症肿瘤的摘除手术。
韓国語	그에게는 암 절제 수술이 이루어졌다.
ロシア語	Ему прооперировали опухоль.

こうした差異は、もらう・されるのように視点が違うだけで同じ事態を表現しているケース(例：ID 番号98の例文)や、ある結果をもたらす決定をしたことと、その結果を情報として伝えられる状態にあることが、けっきょくは同じ事態を表現しているようなケース (ID 番号1253の例文) などでも見られる。ID 番号98の例文では、英語のみ「賞賛された」と動詞の受動態で表現されているのに対し、他の言語では、名詞的な表現を使って、「賞賛を受けた」と表現されている。同じく、ID 番号1253の例文では、ドイツ語以外の言語は、「どこに泊まるか決めた」と直接的に表現しているのに対し、ドイツ語のみ「どこに泊まるか情報を伝えられる」という間接的な形で同じことを伝えている。

言語	ID 番号98の例文
日本語	彼は、その出勤に対しては非常に多くの称賛を受けた。
ドイツ語	Er hat ein dickes Lob für seinen Einsatz bekommen.
英語	He was praised a lot for his commitment.
フランス語	Il a reçu beaucoup d'éloges pour son engagement.
スペイン語	El recibió muchos elogios por la ejecución.
中国語	他的参与得到了非常多的赞扬。
韓国語	그는 그 출동에서 큰 칭찬을 받았다.
ロシア語	Он получил много хвалебных отзывов по поводу этой миссии.

言語	ID 番号1253の例文
日本語	君は私たちがどこに泊まるかもう決めてるの。
ドイツ語	Weißt du schon, wo wir übernachten?
英語	Did you already decide where we will spend the night?
フランス語	Tu as décidé où on va dormir ?
スペイン語	¿Habéis decidido donde nos vamos a quedar a dormir?
中国語	你已经决定好我们的住宿安排了吗？
韓国語	넌 우리가 어디에 머무를지 벌써 정한거니?
ロシア語	Ты уже решил, где мы остановимся ночевать?

### 3.6. 言語による概念化の仕方の差異

同じ語場に属する単語の場合も、その概念の切り分け方は、言語によって異なる。たとえば、「他人の所有物を非合法的な形で取得する」ことに関して、日本語では、以下のような定義があり<sup>11)</sup>、ひそかに取るか、むりやり取るかでまず大きく切り分けられている。

盗む：ひそかに他人のものを取って自分のものにする事

奪う：他人の所有するものを無理に取り上げる事

窃盗：他人の財物をひそかに盗み取る事。

強盗：暴力や脅迫などで他人の金品を奪う事

それに対し、英語・ドイツ語・フランス語などでは、建物等への侵入によるか、路上等で直接奪うかで大きく切り分けられる。前者では、「戸口の錠や窓などを暴力的に壊して、あるいは暴力的な仕方直接盗みに入る」という意味の概念 burglary (英) / Einbruch (独) / cambriolage (仏)があり、そこに人がいなければ、「窃盗」や「泥棒に入る」という表現に対応するが、人がいれば、「(居直り)強盗」という意味にもなるし、そもそも銀行強盗のように人がいることがわかっているでもむりやり入る「(押し込み)強盗」も表すことができる。ID 番号1299の例文(調整前)が、その例であるが、一義的に解釈できない場合にどちらの解釈に寄せるかは、ケースバイケースで、簡単に答えの出ない問題である。

言語	ID 番号1299の例文 (現在調整中)
日本語	私は散歩の際に(家宅侵入による)泥棒の場面に出会い、ずっと見ることができた。
ドイツ語	Ich habe den Einbruch beim Spaziergang beobachten können.
英語	I observed the burglary during a walk.
フランス語	Pendant ma promenade, j'ai assisté à un cambriolage, j'ai tout vu.
スペイン語	Me encontré con un robo mientras paseaba y pude ver toda la escena.
中国語	我在散步的时候看到了小偷潜入别人家，然后观察了整个窃盗案件。
韓国語	나는 산책 도중, 강도하는 장면에 마주쳐 계속 보고 있을 수 있었다.
ロシア語	Во время прогулки я оказался на месте грабежа и смог все увидеть.

### 3.7. 比喩やことわざ等における文化的イメージの差

文化的な背景がわからないと理解できない比喩や、敷衍しないと理解できない表現、あるいは当該言語の文化圏特有のことわざを踏まえた表現などを、複数の言語間でどう調整して表現するかも、大きな問題である。たとえば、ID 番号1601のドイツ語で使われている *ein junger Gott* は、おそらくギリシア神話のアポロン神を指していて、神々しいまでの美的な美しさを強調していると思われるが、これに対する当初の日本語訳は、「彼は、神わざのように完璧に踊っている！」で、踊りの技術の方に引き寄せて理解していた。

言語	ID 番号1601の例文（現在調整中）
日本語	彼は、神々しいほど美しく完璧に踊っている！
ドイツ語	Er tanzt wie ein junger Gott!
英語	He dances splendidly!
フランス語	Il danse merveilleusement bien, comme un dieu !
スペイン語	Él está bailando bella y perfectamente.
中国語	他如同精灵般跳出美妙的舞蹈。
韓国語	그는 귀신 같은 솜씨로 아름답고 완벽하게 춤추고 있어!
ロシア語	Он танцует, как бог, идеально красиво.

同じく ID 番号4358のドイツ語で使われている *Pech im Spiel, Glück in der Liebe*. は、直訳すれば、「ギャンブルではついてないが、愛ではついてる」という意味だが、意図することは、*Kein Nachteil ohne Vorteil./ Jedes Ding hat zwei Seiten./ Wo Licht ist, ist auch Schatten*. 等と同じく、「何にでも両面が有り、幸と不幸は表裏一体だ」あるいは「良いときもあれば悪いときもある」という意味で、日本語のことわざに置き換えると「人間万事塞翁が馬」の意味に相当する。このように、対応する日本のことわざで表現するか、直訳に近い形にするかがまず問題である。もうひとつの問題は、欧米の文化圏には各国にほぼ同じことわざがあるが、英語やフランス語では、*Lucky at cards, unlucky in love./ Malheureux en amour, heureux au jeu*. とその方向が正反対であり、言語によっては両方のバージョンが共存する場合もあることである。今のところは、ほぼ同じことわざや比喩がある場合は、その各国語版を採用し、そうでない場合は訳す方式にした。

言語	ID 番号4358の例文
日本語	ことわざで言われている。ギャンブルではバッドラック、愛ではグッドラック。
ドイツ語	Ein Sprichwort sagt: Pech im Spiel, Glück in der Liebe.
英語	A proverb says: Lucky at cards, unlucky in love.
フランス語	Comme le dit le proverbe : “Heureux au jeu, malheureux en amour”.
スペイン語	Hay un dicho que dice: desgraciado en el juego, afortunado en amores.
中国語	俗话说：“赌场得意，情场失意”。
韓国語	속담으로 전해지고 있다. 장난 속의 불행, 사랑 속의 행복
ロシア語	Пословица гласит: неудача в игре, удача в любви.

### 3.8. 社会構造の違いによって生じる問題

それぞれの言語の文化圏では、文化的な側面だけでなく、政治、経済、社会制度も大きく異なっていることが多く、そうしたテーマの例文の訳では、また固有の問題が生じる。たとえば、ID番号8303の例文がその例であるが、ドイツの場合、外国人の労働許可は、Ausländeramt（外国人局）での滞在許可を前提に，Arbeitsamt（労働局）で得ることになっている。仕事の斡旋もするArbeitsamtは日本でいうと公共職業安定所（通称：ハローワーク）に相当するが，Sie sollten sich beim Arbeitsamt eine Arbeitsgenehmigung besorgen. というドイツ語文を「公共職業安定所（ハローワーク）で労働許可を入手してください。」と訳すと誤解を招きかねない。独日の2カ国語間の調整だけで良ければ、「労働局で労働許可を入手してください。」というドイツ語からの直訳でかまわないが、この文を8言語で表現しようとする、国によって制度が違うので、どの国の制度を基に調整するかが問題となる。

言語	ID番号8303の例文（現在調整中）
日本語	公共職業安定所（ハローワーク）で労働許可を入手してください。
ドイツ語	Sie sollten sich beim Arbeitsamt eine Arbeitsgenehmigung besorgen.
英語	You should obtain a work permit from the employment office.
フランス語	Vous devez obtenir un permis de travail auprès de l'agence pour l'emploi.
スペイン語	Obtenga un permiso de trabajo en la Oficina Pública de Seguridad en el Empleo.
中国語	请在职安（公共职业安定所）取得工作许可。
韓国語	직안（공공직업안정소）에서 노동허가를 입수해 주세요.
ロシア語	Вы должны получить разрешение на работу в службе занятости.

同じことは、たとえば「国道」や「県道」などの社会的インフラを表す表現を含む文の際にも生じる。国によって、道路の管理システムやその呼称が違うので、直訳すべきか、その国の対応するシステムの名称を使うべきかの判断が問題となる。他にも、教育機関での数字や記号を使った成績評価をテーマとした文などでも、国により最高評価が1や5やAなど、異なった記号で表される場合、同じく直訳すべきか当該社会で通常使われている記号を使うかがやはり問題となる。ID番号539/3674の例文などはその例である。

言語	ID番号539の例文
日本語	国道8号線で2台のトラックが衝突した。
ドイツ語	Auf der A8 sind zwei Lastwagen zusammengestoßen.
英語	On Highway 8, two trucks crashed into each other.
フランス語	Deux camions sont entrés en collision sur la nationale 8.
スペイン語	Dos camiones chocaron en la Nacional 8.
中国語	两辆卡车在国道八号上相撞了。
韓国語	국도 8호선에서 2대의 트럭이 충돌했다.
ロシア語	На дороге №8 столкнулись два грузовых автомобиля.

言語	ID 番号3674の例文
日本語	彼女は数学ではきっかり5を取った。
ドイツ語	Sie hatte in Mathematik eine glatte Eins.
英語	She got an A for the mathematics course.
フランス語	Elle a eu la meilleure note en mathématiques.
スペイン語	Ella sacó un diez en matemáticas.
中国語	她数学成绩刚好拿了 A (等级)。
韓国語	그녀는 수학에서 정확히 1을 얻었다.
ロシア語	По математике она получила отличную оценку.

### 3.9. その他

上記の他にも、丁寧さに関わる言語の使用域 (Register) の考慮、性別情報の取り扱いを含むポリティカル・コレクトネスへの対応、例文コーパスを最低でも十数年以上使うと仮定するなら、FAX カセットなどの現在もすでに死語となりつつある語をどこまで残すかの判断等が例文の検討の際には問題となった。

## 4. 今後の展望

今後は、第3章全体で述べた問題点を考慮しながら8言語間での例文内容の本格的な調整を進め、例文データベース自体の品質向上に努めることが第1の課題である。他方、それと並行して、1.3. 節で述べた機能を持つ検索システムと教材作成システムを、遅くとも2022年前半までに完成させることが第2の課題である。

また、それらの作業を前提にして、来年度の後半以降を目処に本システムを利用した自己発見型学習やデータ駆動型学習を組み込んだアクティブ・ラーニングを1、2年生を対象とするドイツ語授業の中に取り入れる試みを開始する予定である。具体的には、1) ライティングや翻訳の授業におけるデータ駆動型学習のツールとして、2) 学習管理システムを使った作文プロジェクト実施の際の自習用のツールとして、3) 教員による授業用ワークシート作成ツールとして、本コーパスをそれぞれ半年間、利用する。それによって得られたフィードバックを基に評価改善を行うとともに、ライティングの際の語彙 (バリエーションと総数)、語法の正確さ、多言語学習方略の獲得、学習者の意欲等の複数の指標を使い、どのような変化が見られるかを調査・分析する予定である。

## 参考文献

- 岩崎克己 (2005) : 初級ドイツ語授業における自己発見型学習の手段としてのコーパスの可能性 - 日独例文コーパス DJPD を例にとりて -, 『ドイツ語教育』 10, pp.48-60. 日本独文学会ドイツ語教育部会.
- 岩崎克己 (2010) : コーパスとしてのインターネットと初級ドイツ語学習, 『広島外国語教育研究』 13, pp.53-64, 広島大学外国語教育研究センター.
- 宇多文雄 (2009) : 『ロシア語文法便覧』, 東洋書店.
- 大瀧敏夫 (1982) : 時制暗喩と Weinrich の『時制論』, 金沢大学文学部論集. 文学科篇巻2, pp.1-25.

ハラルト ヴァインリヒ (2003) : 『テキストからみたドイツ語文法』 訳 : 脇阪豊 他9名, 三修社.  
ハラルト ヴァインリヒ (1982) : 『時制論』 訳 : 脇阪豊 / 大瀧敏夫 / 竹島俊之 / 原野昇, 紀伊國屋書店.

Dodd, Bill: Exploiting a Corpus of Written German for Advanced Language Learning. In: Wichmann, Anne/ Fligelstone, Steven/ McEnergy, Tony/ Knowles, Gerry(ed.) (1997): Teaching and Language Corpora. London and New York: Longman.

Aston, Guy(ed.) (2001): Learning with Corpora. Houston: Athelstan.

Iwasaki, Katsumi (2005): Einsatzmöglichkeiten eines deutsch-japanischen Parallelkorpus als ein Mittel für exploratives Lernen im DaF-Unterricht, Neue Beiträge zur Germanistik Band 4 /Heft 4, pp.166-181, JGG.

Tribble, Ch./Jones, G. (1997): Concordances in the Classroom. Houston: Athelstan.

## 注

- 1) 本稿で論じているプロジェクトは、2020-2022年度科学研究費補助金基盤 (B) 一般「複言語・多言語教育を視野に入れた初・中級学習者用8言語例文パラレルコーパスの開発」(課題番号: 20H01285) の支援を受けたものである。当該プロジェクトの研究責任者は岩崎克己, 研究分担者は、青木利夫 (スペイン語), 荒見泰史 (中国語), Carlos García Ruiz-Castillo (スペイン語), 崔真碩 (韓国語), 平手友彦 (フランス語), 山本孝子 (中国語), 吉満たか子 (ドイツ語) (五十音順) である。
- 2) 図1は、ドイツ語の Buch をキーワードとして検索し、出力された例文を8言語すべてで単純出力した画面である。Buch で検索した場合の出力件数は、50件であった。なお、以下は、作成途中の段階で、機能を限って部分的に公開しているデモ版のアクセス用 URL である。  
<https://parc.lang.hiroshima-u.ac.jp/>
- 3) 同じ膠着語でも分かち書きの無い日本語と表記の際に文節レベルでスペースを入れる韓国語では、検索方法が異なる。後者の場合、欧米語と同じくスペースによる区切りがあるので、ワイルドカード検索をしても検索結果のハイライトの範囲は自ずと定まるのに対し、日本語の場合は分かち書きがないので、検索条件から検索者の意図を類推してハイライトの範囲を意識的に限定してやらないと実用にならない。この点は、分かち書きの無い中国語の場合も同様である。
- 4) 本計画の中核となった日独例文コーパスのドイツ語例文を作成したのは、当時広島大学に留学中であった3人のドイツ人学生で、それを基に日本語例文を作成したのは、同じく当時広島大学で非常勤講師をしていた20代の日本人ドイツ語教員2人および大学院生1人であった。例文作成時の彼らのドイツ語能力はいずれも B2から C1以上であった。
- 5) 例文のミスや誤訳を直すと共に、時代に合わない内容の例文を新しいものに一部差し替えた。
- 6) 編集段階では、例文作成者に加え、中国語教員、アメリカ人教員、スペイン人教員が各1名新たに編集者として加わった。
- 7) ID 番号1089の例文では、敬語としての「あなた」を表す「您」と対比させる形で通常の「きみ・あなた」を表す「你」を使っている。
- 8) この表の分類は完全なものではない。たとえば、英語の未来完了・現在完了・過去完了のそれぞれの時制における進行形は含まれていない。またフランス語教員の同僚によれば、フラン

ス語には、aller+動詞の原形で表される「近未来」だけでなく venir de+動詞の原形で表される「近過去」もあるそうだがそれも入っていない。さらに、ロシア語のアスペクトという観点では、同じ「過去の前後関係」でも1回限りであれば、完了体過去が、過去に反復していた行為では不完了体過去が使われるなど、更に下位分類が必要だが、それもしていない。ここでは、紙幅の関係で、欧米系の5言語に共通する主要な範疇のうち、初中級者用の例文に多く使われそうなものを挙げるに留めた。なお、表の作成に当たっては、広島大学の同僚の平手友彦氏、Carlos García Ruiz-Castillo 氏、青木利夫氏および広島大学客員教授 Sergey Tolstoguzov 氏から貴重な助言をいただいた。もし、誤り等があった場合の責任はすべて岩崎にある。

- 9) 説明と語りの時制（大滝 1982; 4）の相違も問題になった。これは、表2の時制の比較表の中でも、「話者からの心理的な距離」という表現で一部取り上げられているが、ドイツ語、フランス語、スペイン語においては、説明と語りの相違は現在完了形と（単純）過去形という時制の違いとしても明示される。今回のコーパスは、初中級用レベルの例文が多いというその性格上、過去表現の大部分は、もともとのドイツ語例文でも「説明」の時制である現在完了形（複合過去）が使われている。しかし、数は少ないが、「語り」の時制としての過去形を使っているケースもあり、両者の区別が明確でない英語を媒介として作られた、フランス語やスペイン語等の例文では、そのずれが見られるケースもあり、これについても適宜調整が必要であった。
- 10) 表3では、幅の関係で8言語ではなく英独日の3言語の例文を挙げるに留める。
- 11) いずれもデジタル大辞泉の定義より引用。

## ABSTRACT

### **A Parallel Corpus of Example Sentences in Eight Languages for Beginners and Intermediate Learners**

Katsumi IWASAKI

Institute for Foreign Language Research and Education  
Hiroshima University

This paper delineates a project concerning the development of a parallel corpus for language learners, named PARC8. The project is currently being developed by a group of foreign languages teachers at Hiroshima University. In order to contribute to multilingual education, the endeavor aims to build a parallel corpus of example sentences in eight languages — Japanese, English, German, French, Spanish, Russian, Chinese, and Korean — that can be used by learners at the beginner and intermediate levels, corresponding to A1 and B1 in the CEFR (Common European Framework of Reference for Languages). This corpus will be published on the Internet as a learning-support corpus, and will be widely available to foreign language learners and language educators in Japan and abroad. The corpus will consist of about 14,000 sets of example sentences of the same content in each language. The output language of the example sentences can be changed freely in up to eight languages, depending on the combination of languages required by the user. In this paper, I mainly discuss the challenges involved in creating high-quality example sentences in the eight languages that represent the same content in the development project.