

# 学習者はどのように日本語を学ぶのか —ポストコロナ時代における日本語運用データの収集法—

石黒 圭 (国立国語研究所)

## 要旨

日本語教育の目的が学習者による日本語運用力の獲得にあり、日本語教育学の目的がその獲得を支援する日本語習得支援研究であると考えれば、日本語教育学では、学習者が日本語という言語をどのように身につけていくのか、その習得過程を記述・分析する基礎資料、すなわち学習者コーパスの構築が必要になる。ところが、新型コロナウイルス感染症の世界的流行により、JFL 環境で学ぶ海外の学習者のもとを訪れての現地調査も、JSL 環境で学ぶ国内の留学生との対面調査も困難になってしまった。そこで、本稿では、現地調査や対面調査を行うかわりに、オンライン環境を活用して収集する作文コーパス、会話コーパス、ゼミ談話コーパスの収集法を紹介した。たとえコロナが終息したとしても、パンデミックの状況下で急速に発展したオンライン・コミュニケーションが今後衰退化することは考えにくく、むしろポストコロナ時代にあっては、オンライン・コミュニケーションにおける学習者の日本語運用のデータ蓄積が重要になる。その意味でも、本稿で示したようなオンライン環境を活用した調査法の試行錯誤と研究者間での情報共有が、日本語教育学の発展のカギとなると見込まれる。

**キーワード：** オンライン・コミュニケーション、フィールドワーク、学習者コーパス、日本語運用力、第二言語習得

## 1. 日本語教育学と学習者コーパス

### 1.1 日本語教育学と日本語教育関連

日本語教育学は道半ばである。日本語教育学という学問を存在しないとまで公言する人はまれであり、日本語教育にかんする学問の存在を否定する人は少数派であるかもしれない。しかし、「学」として日本語教育学の成立を認める社会的コンセンサスには乏しいように見える。事実、「日本語学・日本語教育」のように、「学」がつく学問とともに「学」のつかない形で併記されることは少なくない。

たとえば、2023 年度より適用される科学研究費助成事業「審査区分表」によると、「中

区分2：文学、言語学およびその関連分野」の小区分は下記のとおりである。

**表1 「中区分2：文学、言語学およびその関連分野」の小区分**

02010	日本文学関連
02020	中国文学関連
02030	英文学および英語圏文学関連
02040	ヨーロッパ文学関連
02050	文学一般関連
02060	言語学関連
02070	日本語学関連
02080	英語学関連
02090	日本語教育関連
02100	外国語教育関連
02010	日本文学関連

これは、小区分として日本語教育の重要性が学術的コミュニティのなかで認められている一方、日本文学をはじめとする文学、日本語学をはじめとする言語学は「学」と称することが認められているものの、日本語教育をはじめとする言語教育学は「学」と称することは認められていない、あるいは少なくとも機は熟していないと見なされている一つの証左であると見ることができる。

日本語教育学という「学」が成立していると考えた場合でも、「日本語」を重視する立場と「教育」を重視する立場とがありうる。1962年6月に「外国人のための日本語教育学会」という名称で発足した日本語教育学会は、70年代から80年代にかけて、学習者から出される日本語についての疑問や、学習者が日本語で表現する過程で産出する誤用をきっかけに日本語にかんする構造的な記述が急速な深まりを見せた。西口（2012）は、学会誌『日本語教育』のこれまでの歩みを3期に分け、そのうちの1990年までの日本語教育学の草創・確立期を第1期と呼び、庵（2012）はその時期を「日本語学と日本語教育の蜜月期」（庵2012:2）と呼んでいる。そのことからわかるように、当時は日本語教育の研究は日本語の研究が中心であり、「日本語」を重視する面が強かったと考えられる。

一方、西口（2012）が第2期と呼ぶ、複数の分野が並行して進む時期を経て、2005年以降の第3期に入ると、日本語教育学の確立が再び叫ばれるようになる。そこで中心となったのは、教室という日本語教育の現場を対象にする実践研究（市島2009）であり、学習者の内面を対象にするピリフ（阿部2014）や学習動機（岡2017）、学習不安（元田2005）などの心理学的な研究や、学習者の声に耳を傾けるライフストーリー研究（三代2014）などの社会学的な研究であった。こうした一連の多様化は、日本語教育研究の「教育」とし

この側面に光を当て、学際的な広がりをもたらした半面、日本語教育学が何を扱う学問か、その焦点をあいまいにしてしまい、「日本語教育学」ではなく「日本語教育関連」となる現状を招いてしまった側面もある。

## 1.2 日本語教育学の基礎資料としての学習者コーパス

このように日本語教育学には、「日本語」を重視する立場と「教育」を重視する立場があるが、筆者の立場としては、日本語教育学の目的は、学習者が目標言語である日本語の運用力を獲得することの支援にあると考え、二つの立場の融合を目指す、学習者の日本語習得支援研究として位置づける。もちろん、多くの場合、日本語教育の教室活動を設計する教師の役割も重要ではあるが、日本語を自然に習得する者も独学で身につける者もあり、教師の存在は絶対的なものではない。日本語を学ぶ主体である学習者の日本語習得研究をその中心に置くのが、あるべき姿であると考えられる。

日本語教育学を学習者の日本語習得を支援する日本語習得支援研究として考える場合、学習者が日本語と初めて出逢ってから、日本語という言語をどのように身につけていくのか、その過程と向きあうことが必要になる。そのときの基礎資料として役に立つのが学習者コーパスである。学習者コーパスは、学習者が話したり書いたりした日本語を大量に収集してデータベース化した言語資料のことであり、学習者の日本語使用の実態を反映した記録である。日本語教育学が科学である以上、学習者が産出したデータを用いて客観的な分析手法に基づき研究を遂行する必要がある、学習者がどのように日本語を習得していくのかを分析するのに優れたデータベースとなる。

## 2. コロナ時代以前の学習者コーパス

### 2.1 コーパス以前のデータベース

日本語教育学が学習者の日本語運用を対象とする以上、コーパスという概念が普及する以前からコンピュータによる処理を前提としないデータベースが個人レベルで多数作られ、利用されていた。なかでも、寺村（1990）は日本語学の記述文法研究に大きな影響を与えたデータベースとして著名であり、国立国語研究所で冊子の pdf 版を公開しているためか、現在でも筆者のもとに問合せが来ることがある。日本語教育研究における初期の関心は、日本語を使う人ではなく、日本語そのものに偏っており、学習者が産出する誤用が日本語としてなぜ不自然なのかを考えることによって、日本語文法の諸現象を観察・分析するさいのヒントとして用いられていた。寺村（1990）は現代的な観点から見直してみても日本語研究への多様な示唆を与えてくれるものであるが、学習者の運用力の向上よりも、母語話者の運用力の基盤となる仕組みの解明に直結するもので、日本語教育学に貢献するものとしては、学習者コーパスの出現を待つ必要があった。

その後の学習者コーパスの展開のなかで、学習者という集団をひとくくりにするこゝへの教育上の課題が共有され、学習者一人ひとりの背景によって習得の仕方が変わるため、それに応じて教え方も変える必要があるという考え方が定着した。そのため、母語別、レベル別、タスク別という三つの観点が、学習者の日本語運用の実態や発達を解明するうえで重要な意味を持つようになり、それが学習者コーパスの設計にも反映されるようになった。以降では、母語別学習者コーパス、レベル別学習者コーパス、タスク別学習者コーパス、それぞれについて説明する。

## 2.2 母語別学習者コーパス

先に述べたとおり、記述文法の時代は、一人ひとりの学習者の背景にまで考えが及ぶことは少なく、学習者として一括されることが多かった。しかし、教育文法という考え方が浸透し（小林 2002、野田編 2005）、学習者の言語運用に関心が向けられるようになってくると、学習者の背景も視野に入るようになり、学習者の母語との関連で誤用が分析されるようになった。また、中間言語研究や対照言語研究の考え方が浸透し、母語でどのような言い方がなされるのかを示す、対訳を付すという発想も生まれた。

国立国語研究所（2009）では、学習者の書いた作文に加え、学習者自身の母語訳も加えた『作文対訳データベース』を公開している。宇佐美洋氏が中心となって収集したもので、収録されている作文は 1754 本であり、その内訳は国によってまちまちであるものの、収録先の国名は、オーストリア、ベルギー、中国、ブラジル、ドイツ、フィンランド、フランス、ハンガリー、インドネシア、インド、カンボジア、韓国、スリランカ、マレーシア、モンゴル、ポーランド、スロベニア、アメリカ、シンガポール、タイ、ベトナム、日本の計 22 ヶ国に及んでいる。

一方、話し言葉を中心としたコーパスでは、国立国語研究所（2020）の『多言語母語の日本語学習者横断コーパス』、通称 I-JAS という学習者コーパスがよく知られている。迫田久美子氏が中心となって世界各地に赴いて収集したもので、統一的な基準に基づいて、12 言語 17 ヶ国 20 地域の 1000 名の学習者のデータを収集した大規模横断学習者コーパスである。調査課題のバリエーションが豊富である（6 種類 12 タスク）点に特徴があり、国内外の学習者の発話データを中心に、作文データも備えている。学習者のデータとは別に、日本語母語話者の 50 名のデータもあるため、学習者一般の特徴も分析できるほか、各言語 50 名以上の学習者のデータがあり、〇〇語母語話者の日本語の特徴を定量的に分析することが可能になっている。これによって、ある誤用や使用傾向が、学習者一般の特徴なのか、ある言語話者に特有なものなのかの判定が可能になった点で大きな進歩と言える。

## 2.3 レベル別学習者コーパス

母語という学習者の背景による広がりにくわえ、日本語をどのくらいの期間学習し、ど

ここまで習熟しているかという日本語の習熟度による違いも日本語学習者の言語運用に大きな影響を及ぼすことは言うまでもない。現在の多くの学習者コーパスには、学習期間やテストの成績など、一人ひとりの学習者の日本語の習熟度のデータがつけられるようになっている。また、レベルごとにある程度の人数のデータを収集することで、レベル間の比較を容易にする工夫がなされている。

たとえば、鈴木智美氏らが中心となり、東京外国語大学留学生日本語教育センターが構築した『JLPTUFS 作文コーパス』は、異なるレベルの留学生を豊富に抱える外国語大学ならではの強みを生かし、55の国・地域出身の、入門～初級、初級後半から初中級、初中級、中級前半、中級後半、上級前半、上級後半、超級の8レベルにわたる学習者の1515本の作文を収録している。これによって、学習者の日本語能力のレベルによってどのような作文を書くのかが詳しくわかるようになっている。

また、話し言葉では『KY コーパス』(鎌田・山内 1999)がよく知られている。『KY コーパス』は鎌田修氏と山内博之氏が中心となって構築した会話コーパスで、OPI (Oral Proficiency Interview) の手法が用いられており、熟達度による能力レベルが明示されている点に特徴がある。90名分の会話データを文字化した言語資料であり、母語別に見ると、中国語、英語、韓国語がそれぞれ30名ずつ、OPIの判定結果別の内訳は、それぞれ初級5名、中級10名、上級10名、超級5名となっている。

ただし、レベル別コーパスを用いて学習者の言語習得の過程を明らかにしようとする、限界が生じる。それは、異なるレベルの学習者は別人であるため、厳密な意味での比較は困難であるからである。その意味で、同一の学習者の発達過程を追跡するのが理想であるが、同一の学習者を何年にもわたり追跡調査を行うのは困難をきわめる。しかし、そうした研究も少数存在する。大学等の教育機関で組織的に行われた一定規模以上のものとしては、管見のかぎり、台湾東呉大学の『LARP@SCU 語料庫』(東呉大学 2007)、中国湖南大学の『湖南大学学習者中間言語コーパス』(杉村 2013)、中国北京師範大学の『北京日本語学習者縦断コーパス (B-JAS)』(国立国語研究所 2023 予定)の三つに限られる。この分野の先駆的な取り組みであり、37名(最終的には26名)の学習者に3年半のあいだ、毎月作文と、作文についてのインタビューを行った『LARP@SCU 語料庫』、日本語学科の中国語話者の日本語会話(計7回)と日本語作文(計19回)を94名分、やはり3年半のあいだ収集・構築した『湖南大学学習者中間言語コーパス』、いずれも貴重な学習者縦断コーパスであるが、ここでは筆者も収集に携わったB-JASについて紹介する。

『北京日本語学習者縦断コーパス (B-JAS)』は、日本語がゼロの段階で日本語学科に入学し、4年間日本語を集中的に学んで卒業した北京師範大学の学部学生17名(うち1名のみ1年程度の既習歴)を対象に、8回のインタビュー調査、および11回の作文調査を継続的に行い収集したものである。横断コーパスであるI-JASと組み合わせて分析することを想定しているため、調査はI-JASに準拠して行っている。研究成果は林・徐・迫田監修(2023

予定)にまとめられ、2022年度中にプロジェクトのサイトで収集したデータの公開を行う予定で準備を進めている。

## 2.4 タスク別学習者コーパス

学習者の日本語運用を考えると、単に日本語が上手であるという一般的な能力を考えるのでは不適切である。なぜなら、流暢に話せるが、まったく書けない学習者もいるし、読み書きは得意だが、会話は苦手という学習者もいるからである。大学院生には学術日本語能力が求められ、ビジネスパーソンにはビジネス・コミュニケーション能力が求められる。患者さんと日常的な会話をする一方、同僚とは専門的なやりとりが求められる看護師には複数の能力が必要とされ、子どもを学校に通わせ、地域のコミュニティで生活する保護者にはまた別の能力が必要となる。

学習者コーパスを収集するさいは、個人情報保護や業務上の守秘義務などがあり、現実のコミュニケーションそのものを収録することは難しく、多くはタスクを指定することになる。タスク別の書き言葉コーパスとしてよく知られているのが、当時横浜国立大学に所属していた金澤裕之氏を中心となって収集した『YNU 書き言葉コーパス』（金澤編 2014）である。メールがもっとも多いが、メール以外でもレポート、新聞、広報誌などの媒体を想定し、説明、紹介、意見、依頼などを行う 12 種類のタスクが、日本語話者、韓国語話者、中国語話者各 30 名によって書かれており、計 1080 本の文章が収録されている。

また、金井勇人氏を中心となって収集を行った『JCK 作文コーパス』（金井 2017）は日本語話者、中国語話者、韓国語話者各 20 名によって書かれ、説明文、意見文、歴史文という三つのタイプの作文計 180 本が収録されている。収録本数は少ないものの、それぞれの作文の長さが 2000 字以上と相対的に長く、文章の全体構成や展開構造を見るのに適したデータベースであり、石黒編（2017）はそうした特性を生かした研究成果をまとめたものである。

一方、話し言葉としては、2.2 の「母語別学習者コーパス」で見た I-JAS や、2.3 の「レベル別学習者コーパス」で見た B-JAS のタスクが比較的豊富である。30 分程度のインタビューとの会話にくわえ、二つの場面別ロールプレイ、二つの 4 コマ（5 コマ）漫画の説明、絵カードの説明などを含んでいる。しかし、タスク別学習者コーパスは、書き言葉、話し言葉ともにまだまだ不十分であり、今後、レジスターやジャンルを考慮しつつ、より広範囲のデータを収集していく必要がある。

以上見てきたように、母語別、レベル別、タスク別という三つの観点から学習者コーパスの整備が進むことによって、学習者の日本語運用の実態を包括的に把握することが可能になりつつある。

### 3. ポストコロナ時代の学習者コーパス

#### 3.1 オンラインというフィールド

日本語教育学を日本語習得支援研究として考え、そのために学習者が日本語を習得する過程を記録した学習者コーパスの構築が必須であると考えた場合、学習者の日本語の習得記録を収集する必要がある。そうした記録が可能となるのは、学習者が学ぶ空間としての教室をはじめとする日本語学習の現場である。その意味で、研究者は学習者のいるところに出かけていって学習記録を収集するフィールドワーカーであり、日本語教育学はフィールドワークとしての側面を持つ。しかし、新型コロナウイルス感染症の世界的流行により、他の諸分野と同様、フィールドワーク自体が困難になってしまった。これまで日本語教育学のフィールドワークの有力な現場の一つであり、筆者が深く関わってきた国内外の大学等の高等教育機関を考えてみても、従来のような学習者と研究者の対面によるフィールドワークの場は損なわれてしまったと言わざるをえない。

まず、日本国内の大学・大学院では、海外から日本への渡航が困難になってしまったため、留学生数は激減した。また、少なくなった留学生と対面して調査しようとしても、コロナの感染リスクがあるため、対面による調査は控えざるをえなくなった。また、海外の大学・大学院では、対面授業が一時的に中止され、オンライン授業に切り替わった地域も、コロナの感染状況に配慮しながら、可能なかぎり対面の授業を継続しようとした地域もあるが、いずれの場合でも、筆者のような国内にいる研究者が海外に渡航して学習データの収集をすることはできなくなった。また、海外の大学で教えている現場の研究者もネコの目のように方針が変わるコロナ対応に追われ、学習データの収集どころではないという現実が生じてしまい、学習者を対象とした日本語教育研究は事実上ストップするという研究の危機にさらされることとなった。

しかし、そうした厳しい環境のなかでも日本語教育研究者は学習者の記録を収集しつづける必要がある。国内外の学習者はけっして日本語学習をやめたわけではなく、変わりはてた環境のなかでも従来と変わることなく地道に日本語学習を継続しているからである。これはパンデミック下でもそうであったし、きたるポストコロナ時代にあっても変わらずに続く状況であろう。

今後私たちは、コロナが爪痕を残す環境のなかで生きていかなければならない。新型コロナウイルス感染症の流行が近い将来終息すれば、対面でのコミュニケーションもある程度回復するだろうが、いつ再流行するともかぎらない状況のなかで、私たちのコミュニケーションにおけるオンライン環境への依存は減ることはないと予想される。事実、私たちは対面の会議ができるような環境にあっても、移動時間を節約できるオンライン会議を選択するようになっており、対面での簡単な打ち合わせが必要な場合でも、記録の残るビジネスチャットツールを使う機会が増えている。そうしたオンライン化の時代の波に学習者

私たちはすでに巻きこまれており、今後は一層激しさを増すだろう。しかし、オンライン・コミュニケーションには対面コミュニケーションとは異なる難しさが存在する。そのため、学習者が行うオンライン・コミュニケーションの実態を記録することで、学習者にとってのオンライン・コミュニケーションの困難点を明らかにし、そうした困難点に対応できる教材や教授法の開発が早急に求められている。そう考えると、新型コロナウイルス感染症の流行がなかなか終息しない状況だからからこそ、学習データの収集を一時的に休止してよいわけではなく、むしろ学習者が否応なく向きあっているオンライン・コミュニケーションの実態把握に邁進する必要がある。

そこで、以降では、筆者自身が新型コロナウイルス感染症の流行下で行ってきた、オンライン環境を活用した学習データ収集の実例と、そこでの試行錯誤を共有することで、今後の学習データ収集の一助とすることを目指したい。

### 3.2 オンラインで収集する作文コーパス

2.3「レベル別学習者コーパス」で述べたように、学習者一人ひとりを追いかける経年的なデータは、学習者の習得過程を知るうえで有力な手がかりとなる。しかし、経年的なデータの収集は手間も時間もかかるためデータの収集自体が難しく、そうした経年的な学習データは質の面でも量の面でも十分とは言えない現状がある。

そこで、筆者は、「海外縦断作文コーパスの構築に基づく文章産出能力の発達過程の実証的研究」というテーマで科研費の基盤研究(A)を取得し、中国の7大学、台湾の2大学、韓国の2大学、ベトナムの3大学、タイの1大学と連携し、開始時で約600名の学習者を対象に、大学の日本語学科に入学して卒業するまでの4年間にわたる執筆能力獲得過程の追跡調査に着手している。調査の開始時期は早い大学で2020年、多くの大学で2021年、遅い大学で2022年であり、この調査の開始時期はいずれも、新型コロナウイルス感染症の世界的流行のときとほぼ重なっており、そのため、調査のすべての過程をオンラインで行っている。ここではそれを、(1)一連の調査管理、(2)作文の執筆、(3)作文の提出、(4)背景情報の収集、(5)学習者との連絡の五つに分けて紹介する。

(1)「一連の調査管理」というのは、学習者が調査に参加するさいの調査の一連の流れの管理のことである。学習者は次の表2のような流れで調査に参加している。



表 2 学習者の調査への参加手順

①Moodle へのアクセス	国立国語研究所が作成した URL から Moodle にアクセスする (年 3 回)
②レベル判定テストの受験	Moodle 上で 3 種類の日本語のテストを受ける (年 1 回)
③作文の執筆	作文ソフトをダウンロードし、その回のテーマに沿った作文を 書く (年 3 回)
④情報シートの入力	作文執筆に関する情報入力シートに記入する (年 3 回)
⑤アンケートの記入	作文執筆に関するアンケートに記入する (年 1 回)
⑥フェイスシートの記入	学習者の基本情報をフェイスシートに記入する (初回のみ)
⑦作文のフィードバック	作文のフィードバックを受ける (年 3 回)

調査を開始した当初は試行錯誤があったものの、2022 年現在では①～⑦の作業がすべて Moodle 上で完結するようになっている。それによって、学習者は海外の大学の教室や寮にいても、帰省やロックアウトで自宅にいても、留学していて日本で生活していても、どこからでも決まった時期に調査に参加することができる。学習データを収集するにはフィールドが必要であるが、そのフィールドを仮想空間に置くことでコロナ禍でも調査の安定した継続が可能になっている。

(2)「作文の執筆」は、学習者はパソコンで入力する方式を採っている。手書きには、文字の正確な字形把握という面や、機械に頼らない自力での日本語の思考という面では一日の長はあるが、現実の世界では手書きをする場面はきわめて限られており、学習者に無理に手書きをさせる意義は乏しくなっている。また、遠隔地での学習データ収集という面から考えても、手書きは現実的ではない。

そこで、この調査では、EssayLogger という入力支援システムを国立国語研究所のスタッフが開発し、それをを用いている。将来的には一般に公開し、無料でダウンロードできるようにする予定ではあるが、特徴としては、母語での入力画面やメモ欄を備えている点(図 1 を参照)と、文字を入力したさいのパソコンのキーストロークが時間も含めてすべて記録できるシステムとなっている(図 2 を参照) 点の二つが挙げられる。とくに後者については、文字入力のさいの一連のプロセスが明確になり、学習者が入力や修正に苦労した点が解明できるため、プロダクトとしての作文だけでなく、プロセスとしての作文も分析対象にできるため、今後の新たな研究分野の創出が見こまれる(田中・石黒(2018)や布施・石黒(2018)も参照)。

14 学習者はどのように日本語を学ぶのか（寄稿論文）  
 —ポストコロナ時代における日本語運用データの収集法—

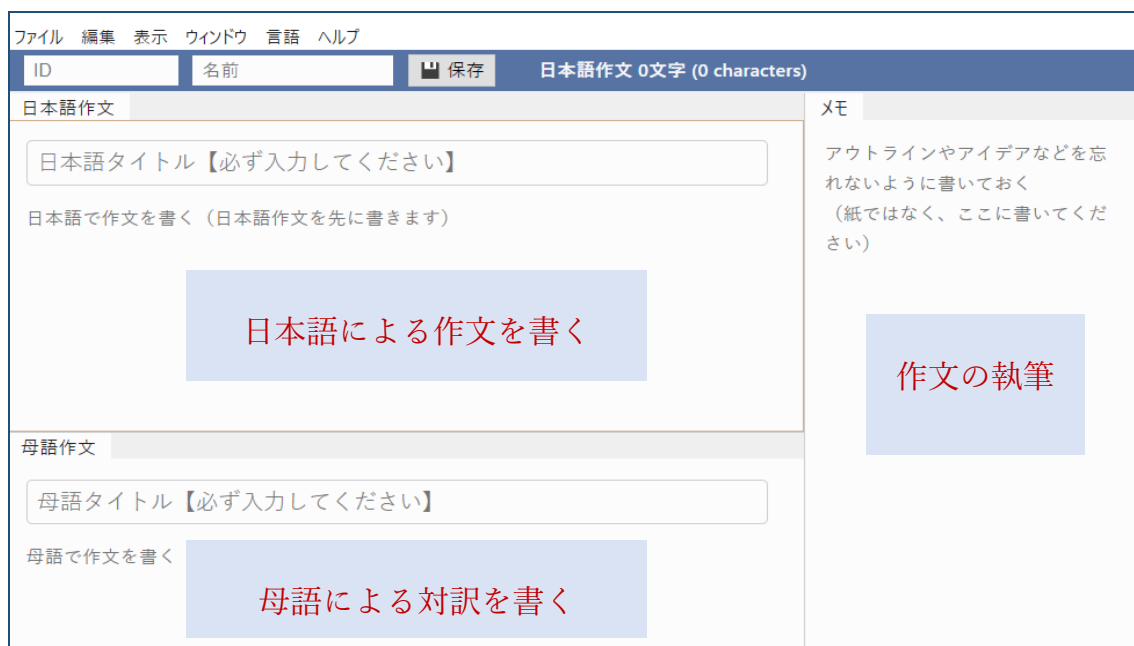


図1 EssayLogger の学習者入力画面

ID番号	分類	日付	時間	入力間隔 (秒)	操作カテゴリ	操作内容1	操作内容2	テキスト全文
0	essay-title	2022/6/14	9:22:49	0	mouse	マウス操作	LeftButton	
1	essay-title	2022/6/14	9:23:08	18.705	key	キー入力	MetaLeft	
2	essay-title	2022/6/14	9:23:27	19.625	key	キー入力	MetaLeft	
3	essay-title	2022/6/14	9:23:27	0.001	key	キー入力	KeyH	h
4	essay-title	2022/6/14	9:23:28	0.125	key	キー入力	KeyA	は
5	essay-title	2022/6/14	9:23:28	0.17	key	キー入力	KeyS	は s
6	essay-title	2022/6/14	9:23:28	0.081	key	キー入力	KeyI	はし
7	essay-title	2022/6/14	9:23:28	0.127	key	キー入力	KeyR	はし r
8	essay-title	2022/6/14	9:23:28	0.105	key	キー入力	KeyE	はしれ
9	essay-title	2022/6/14	9:23:28	0.224	key	キー入力	KeyM	はしれm
10	essay-title	2022/6/14	9:23:28	0.096	key	キー入力	KeyE	はしれめ
11	essay-title	2022/6/14	9:23:29	0.343	key	キー入力	変換	走れメロス
12	essay-title	2022/6/14	9:23:29	0.16	key	キー入力	変換確定	走れメロス
0	essay	2022/6/14	9:23:30	0	key	キー入力	KeyM	m
1	essay	2022/6/14	9:23:30	0.071	key	キー入力	KeyE	め
2	essay	2022/6/14	9:23:30	0.177	key	キー入力	KeyR	め r
3	essay	2022/6/14	9:23:30	0.079	key	キー入力	KeyO	めろ
4	essay	2022/6/14	9:23:30	0.16	key	キー入力	KeyS	めろ s
5	essay	2022/6/14	9:23:31	0.353	key	キー入力	KeyU	めろす
6	essay	2022/6/14	9:23:31	0.160	key	キー入力	KeyU	めろす

図2 EssayLogger のキーストローク記録データ

(3)「作文の提出」は、2022年現在では Moodle を介して OneDrive に保存するようにしているが、それまでは学習者に書いた作文を直接 OneDrive の個人フォルダに提出してもらう形を取っていた。OneDrive は国立国語研究所のアカウントが存在するため、セキュリティ上もリスクが低く、また、他のクラウドストレージよりも中国国内からのアクセスがしやすいため、OneDrive を用いた。すべての作文データを OneDrive に集約することで、デ

ータの管理・整理がより容易になっている。

(4)「背景情報の収集」は、Moodle を介して行っており、中国以外の国・地域では Google フォームに、中国では問巻星に質問内容等を記入してもらっている。すでによく知られていることではあるが、アンケートの記入・集計に適したこうしたフォームを使うことで、学習者にとってはスマホ等からの記入も容易になる一方、調査者側も自動集計が可能になるという点で重宝する。

(5)「学習者との連絡」は、スマホで使えるメッセージング SNS が確実である。学習者がふだん使っているもので連絡すると、もっとも見る可能性が高いからである。日本や台湾では LINE の普及率が高いが、中国では WeChat (微信)、韓国では Kakao Talk (カカオトーク)、ベトナムやタイでは Facebook の Messenger で連絡を取るようになっている。また、大事なことは学習者の母語で連絡を取ることである。日本語だと意味がわからないこともあり、意味がわかっていても面倒なので処理が後回しになりがちだからである。

このように、ポストコロナ時代においては仮想空間を活用した調査が主流になる。移動の手間もなく、コロナの感染リスクもなく、コストも比較的安く済む。(1)～(5)の方法は、2022 年現在では必然の感はあるが、こうした方法を確立するまでは、長い時間をかけ、試行錯誤の繰り返しが必要であった。今後、同種の調査を計画する方は他山の石としていただければさいわいである。

### 3.3 オンラインで収集する会話コーパス

2.3「レベル別学習者コーパス」で紹介した B-JAS というコーパスは、中国語話者のコーパスであった。横断コーパスである I-JAS が横糸、縦断コーパスである B-JAS が縦糸であり、この二つが組み合わさることで学習者の日本語習得の実態が明らかとなり、日本語教育研究は進む。しかし、B-JAS は中国語話者に限られるので、中国語話者以外の学習者のデータも集めて初めて、縦糸と横糸は組み合わせり布を織りなすことになる。そこで、2022 年に着手した新たな調査では、母語の違いによる経年的な比較ができるよう、中国語話者にくわえ、ベトナム語話者、タイ語話者のデータ収集を行い、母語の転移も含めた学習者の習得過程を体系的に比較できる環境を整えることを目指している。

B-JAS では、半年に 1 度、インタビュアーが北京に足を運んでデータを取りつづけており、今回の調査でも同様の体制を継続する予定でいた。ところが、コロナの影響で海外出張が不可能となり、オンラインでのデータ収集を選択せざるをえなくなってしまった。基本的には、3.2「オンラインで収集する作文コーパス」の方法を踏襲したが、こちらの調査の場合、会話のインタビュー調査があるため、オンラインでインタビューデータを収集する方法を考えた。Zoom が普及しているベトナムとタイでは Zoom、Zoom が使いにくい中国では Teams といったオンライン会議システムを用いてインタビュー調査を行うことにしたが、その場合、音質の問題がある。そこで、インタビュー担当者 (インタビュアー) と学

習者（インタビューイ）の双方にリニア PCM レコーダーを用意し、イヤホンやヘッドホン装着して調査を行った。これにより、それぞれのリニア PCM レコーダーには自分の声だけが入り、調査終了後に二つの音声を組み合わせることで、クリアな会話が再生できるという仕組みを作り出すことができた。このようにして、現地に足を運ばなくても、対面での調査と同等の音質のインタビューデータを収集することができるようになっている。

### 3.4 オンラインで収集するゼミ談話コーパス

3.2「オンラインで収集する作文コーパス」および3.3「オンラインで収集する会話コーパス」では、プレコロナ時代であれば対面環境を活用して収集できたデータを、コロナ時代、そしてポストコロナ時代にどのように収集するかという関連からの記述が中心であった。

しかし、3.1「オンラインというフィールド」でも述べたように、学習者はオンライン環境のなかで授業を受講し、日常生活を送っている。そのため、日本語教育の教材やカリキュラム自体は、プレコロナ時代の対面コミュニケーションを前提にしたものになっており、ポストコロナ時代において学習者が生きていかなければならない環境からはかけ離れたものになっている。ポストコロナ時代にふさわしい教材やカリキュラムを作成するには、学習者たちがコロナ時代、ポストコロナ時代においてどのようなコミュニケーションを行っており、オンライン環境で生きていくなかでどのような困難点を抱えているのか、その実態を把握することが求められるだろう。

そこで、筆者とその同僚である鳥日哲氏が中心となって研究しているゼミ談話コーパスを紹介する。ゼミ談話コーパスは、日本語学や日本語教育学の大学院生たちが参加している大学院ゼミナールの談話を録画したものである。録画は、新型コロナウイルス感染症が猛威をふるいはじめ、ゼミナールがオンラインに移行した2020年4月から収録を続けているものである。なかには、大学院に入学後、日本に入国できなくなってしまい、母国でゼミナールに出席しつづけた修士課程の大学院生もおり、高度に専門的なコミュニケーションが行われ、新参加者にとっては対面でもコミュニケーションに苦勞するゼミナールという環境に当初からオンラインで参加し、一層の苦勞を強いられた学習者のデータも含まれ、貴重なデータとなっている。

分析自体は十分には進んでいないが、興味深い事実も明らかになっている。一般に、オンライン・コミュニケーションは、対面コミュニケーションに比べて話し合いへの参加者の積極的関与が失われやすいと考えられがちであるが、実際には、参加者同士の協力によって話者交替が積極的に行われ、沈黙による気まずさが回避される多様な方略が用いられていることがわかった。また、オンライン・コミュニケーションでは予期せぬトラブルに見舞われがちであるが、困っている参加者を別の参加者が自発的にサポートする様子も観察された。さらに、脱線が少なく、決まった内容以外は話題に上らないと思われがちなお

ンライン・コミュニケーションであるが、実際には、オンライン会議システムで共有されるビデオの映像や音声に積極的に言及することで自然な雑談が行われ、リラックスした雰囲気を作っていることがわかった。

このように、私たちの思い込みや固定観点で捉えられがちなオンライン・コミュニケーションも、きちんとデータを収集し、仔細に観察することで、実態が正確に捉えられることが期待される。石黒（2022）は『社会言語科学』で組まれた特集『『コロナ禍』と社会言語科学』に掲載されたものであるが、ポストコロナ時代にあってはこうした研究が質量ともに増加することが望まれる。

## おわりに

日本語教育学の目的を、学習者の目標言語である日本語の運用力を獲得することであると考え、学習者の日本語習得支援研究として位置づけた場合、母語別・レベル別・タスク別の学習者の言語運用の実態を記録した学習者コーパスの構築が必要になる。しかし、たとえコロナが終息したとしても、パンデミックの状況下で急速に発展したオンライン・コミュニケーションが今後衰退することは考えにくい。むしろ、ポストコロナ時代にあっては、学習者の日本語運用を対面コミュニケーションのみに求めるのは現実的ではなく、オンライン・コミュニケーションの日本語運用の記録を収集することは重要な課題となろう。

コロナ禍にあって、対面の教室という教育現場を奪われた学習者も教師も、双方の協力のもと、試行錯誤を重ねながら、オンラインという仮想空間に新たな教室を作りあげることが求められてきた。同様に、日本語教育学の研究者も、学習者から言語運用のデータを得られる教室という対面のフィールドを失った以上、オンライン環境を最大限に活用して、試行錯誤を重ねながら、仮想空間に新たなフィールドを構築せざるをえなくなった。本稿はその試行錯誤の記録である。

しかし、試行錯誤の結果、オンラインによる調査は負の側面ばかりではなく、正の側面も少なくないことがわかってきた。データ収集という観点から見て、オンラインのフィールドはむしろ効率的であり、統一的な基準に基づいて大量のデータを収集・整理するのに向いている。また、学習者はすでにオンライン環境のなかで生きており、オンラインを活用した日本語コミュニケーション能力を獲得することが社会的に求められている以上、オンライン・コミュニケーションにおける学習者の困難点を整理し、それを教育実践や教材開発に役立てる必要もある。

筆者が示した三つの研究は、オンラインを活用したささやかな試みにすぎない。異なるアプローチからの優れた研究も存在するはずである。重要なことは、こうしたデータの収集・分析法を研究者間で共有することである。それによって、学習者がどのように日本語を学ぶのかという日本語教育学の中心的テーマの解明に一步近づくことができると考える。

## 付記

本研究は科研費 JP21H04417 および JP22K00655、国立国語研究所機関拠点型基盤研究プロジェクト「多様な言語資源に基づく日本語非母語話者の言語運用の応用的研究」のサブプロジェクト「日本語学習者の作文の縦断コーパス研究」および「日本語学習者の談話の縦断コーパス研究」の研究成果である。本研究の遂行にあたり、国立国語研究所の上記プロジェクトのスタッフ各位の協力を得た。記して感謝申し上げる。

## 参考文献

- 阿部新（2014）「世界各地の日本語学習者の文法学習・語彙学習についてのビリーフ —ノンネイティブ日本語教師・日本人大学生・日本人教師と比較して—」『国立国語研究所論集』8、1-13 頁、国立国語研究所。
- 庵功雄（2012）「日本語教育文法の現状と課題」『一橋日本語教育研究』1号、1-12 頁、一橋日本語教育研究会。
- 石黒圭（2022）「コロナ禍におけるオンライン・ゼミナールの可能性 —オンラインのゼミ談話に見るコミュニケーション活動の豊かさ—」『社会言語科学』25-1、39-54 頁。
- 石黒圭編（2017）『わかりやすく書ける作文シラバス』くろしお出版。
- 市嶋典子（2009）「日本語教育における『実践研究』論文の質的变化 —学会誌『日本語教育』をてがかりに—」『日本語教育論集』25、3-17 頁、国立国語研究所。
- 岡葉子（2017）「日本語教育学における「学習動機」の概念について —motivation の訳語をめぐる問題—」『東京外国語大学留学生日本語教育センター論集』43、19-32 頁。
- 金井勇人（2016）『JCK 作文コーパス』  
(<http://nihongosakubun.sakura.ne.jp/corpus/> 2022 年 8 月 15 日取得)。
- 金沢裕之編（2014）『日本語教育のためのタスク別書き言葉コーパス』ひつじ書房。
- 鎌田修・山内博之（1999）「KY コーパス」  
(現行版は、李在鎬氏らによって形態素解析と誤用タグが施され、検索システムを備えた「タグ付き KY コーパス」<http://jhlee.sakura.ne.jp/kyc/> 2022 年 8 月 15 日取得)。
- 国立国語研究所（2009）『作文対訳データベース』(<https://mmsrv.ninjal.ac.jp/essay/> 2022 年 8 月 15 日取得)。
- 国立国語研究所（2020）『多言語母語の日本語学習者横断コーパス (I-JAS)』  
(<https://www2.ninjal.ac.jp/j11/lsaj/> 2022 年 8 月 15 日取得)。
- 国立国語研究所（2023 予定）『北京日本語学習者縦断コーパス (B-JAS)』  
(<https://www2.ninjal.ac.jp/j11/>内に 2022 年度中に公開予定)。
- 小林ミナ（2002）「日本語教育のための教育文法」『日本語文法』2-1、153-170 頁。
- 杉村泰（2013）『湖南大学学習者中間言語コーパス』  
(<https://www.hum.nagoya-u.ac.jp/about/about-sub2/japanese/sugimura/2-3.html> 2022 年 8 月 15 日取

- 得)。
- 田中啓行・石黒圭「日本語学習者の作文執筆修正過程 —中国人学習者と韓国人学習者の修正の位置と種類の分析から—」『国立国語研究所論集』14、国立国語研究所、2018. 1、255-274 頁。
- 寺村秀夫 (1990)『外国人学習者の日本語誤用例集』平成 2 年度科学研究費特別推進研究報告書  
(<http://doi.org/10.15084/00003533>)。
- 東呉大学日本語文学系 (2007)『LARP at SCU 語料庫』  
([https://web-ch.scu.edu.tw/japanese/web\\_page/3936](https://web-ch.scu.edu.tw/japanese/web_page/3936) 2022 年 8 月 15 日取得)。
- 西口光一 (2012)『教育』分野 —日本語教育研究の回顧と展望—『日本語教育』153、8-24 頁、日本語教育学会。
- 野田尚史編 (2005)『コミュニケーションのための日本語教育文法』くろしお出版。
- 布施悠子・石黒圭「日本語学習者の作文執筆過程における自己修正理由 —上級中国人学習者、上級韓国人学習者、日本語母語話者の作文の比較から—」『国立国語研究所論集』15、国立国語研究所、2018. 7、17-42 頁。
- 三代純平 (2014)「日本語教育におけるライフストーリー研究の現在 —その課題と可能性について—」『リテラシーズ』14、1-10 頁、くろしお出版。
- 元田静 (2005)『第二言語不安の理論と実態』溪水社。
- 林洪・徐一平・迫田久美子監修、張林・野山広・石黒圭・岩崎拓也編集 (2023 予定)『北京日本語学習者縦断コーパス「B-JAS」と日本語の教育研究 (北京日語学習者历时語料庫 (B-JAS) 与日語教学)』中国国際広播出版社。

## Title

### **How learners learn Japanese: The method for collecting data on Japanese language usage in the post-corona era**

ISHIGURO, Kei

## Abstract

Considering that the purpose of Japanese language pedagogy is to help learners acquire and support their Japanese proficiency, Japanese language pedagogy requires learner corpora, which is the material for analyzing the acquisition process learners acquire Japanese language skills. However, it became impossible to collect the learner materials necessary for constructing the learner corpora. This is because the global outbreak of the new coronavirus has made it difficult to conduct field surveys of overseas learners in JFL environments and face-to-face surveys of international students in Japan who study in JSL environments. Therefore, instead of conducting field and face-to-face surveys, this paper introduces the method for collecting composition corpora,

conversation corpora, and seminar discourse corpora, which are collected using the online environment. Even if the corona comes to an end, it is unlikely that online communication will decline in the future. Rather, in the post-Corona era, it will be important to accumulate data on learners' use of Japanese in online communication. In this sense, trial and error in survey methods utilizing the online environment and the sharing of information on survey methods among researchers are expected to be keys to the development of Japanese language pedagogy.

**Keywords** : online communication, fieldwork, learner corpora, Japanese language proficiency, second language acquisition