

科学研究費補助金 基盤研究(C)

(課題番号 17K02926)

日本語母語学習者データに基づく  
ロシア語学習者コーパス構築の  
基盤研究

2017 年度研究成果報告

2018 年 3 月 31 日

# 学習者言語を探る

— 科研費プロジェクト「日本語母語学習者データに基づくロシア語学習者コーパス構築の基盤研究」  
2017年度研究成果報告総括 —

林田 理恵

(大阪大学大学院言語文化研究科)

## I. 研究の概要

本研究は、国内外で初となる日本語を母語とする学習者データに基づくロシア語学習者コーパス構築に向け、理論的・技術的な基盤整備を行うことを目的としている。作業はロシア・コーパス開発研究チームとの協同研究体制を前提に、以下の3点の最終目標に向け、最初の1年の研究調査作業を終了したところである。

- 1) 信頼性の高いアノテーション済み学習者コーパスを作成するための、付加する情報タグの有効な分類・構成のあり方を検討、新たな付加情報タグ・ガイドラインを提案。
- 2) 収集されている作文試験結果の電子データに基づき、アノテーション、すなわち情報タグの付加を試行、公開されている品詞タグ付けシステムを利用し、日本語母語学習者データに基づくロシア語学習者コーパスのパイロット版を作成。
- 3) パイロット版により試験的にコーパス分析を行い、試作したコーパスによってロシア語学習者の言語使用特徴がどの程度、明示化されるのか、提案した付加情報タグ・ガイドラインの有効性も含めて検証。

## II. 研究の背景

学習者コーパスは、非誘導型・大規模データという特性によって、学習者言語における特定言語素材の過剰・過少使用や母語転移、回避方略などの客観的特徴を明らかにし得る有効な手段として注目されており、国内でも英語・日本語の学習者コーパスを中心に近年、整備が進んでいる（石川 2012:215-249, 望月 2012:112-115）。

ロシア語学習者コーパスは、現時点ではモスクワの National Research University, Higher School of Economics (= 以下, HSE) 人文学部言語学コースのスタッフが中心となって2013年に公開した Русский учебный корпус (Russian Learner Corpus = RLC, <http://web-corpora.net/RLC>) が唯一のものとなっている。

筆者の所属する機関では、2000年より15年間に亘って収集した学習者のロシア語作文データが存在し、それを利用したロシア語学習者コーパス (= Japanese RFL Learner Corpus 以下, JRFLC Corpus) 構築が可能なことに着目、今回、科研プロジェクトとして上記の研究調査作業に取り掛かることとした。日本語母語学習者のデータを対象としたロシア語コーパス

が構築できれば、学習者のロシア語使用における多面的な言語特性が明らかになり、その分析に基づいて、日本での現行ロシア語教育における教材や指導案、カリキュラムについて抜本的な見直し、改良を図っていく可能性が広がる点が期待される。

### III. 研究の特色・独創性

本科研プロジェクトの特色としてまず第 1 に挙げることができるのは、教育現場への実証的データ提供に道を開くことができるという点である。国内のロシア語教育現場では、これまでは主として経験値に依拠したカリキュラムや教材開発等がなされているが、本科研プロジェクトによるロシア語学習者コーパス開発に向けた基盤整備が進めば、

- 1) ロシア語学習における日本語を母語とする学習者の学習者言語の傾向や特性に関する、より客観的・実証的データに基づいた分析
- 2) 教育現場における教材開発や指導案作成等に対する有効かつ信頼性に足る情報の還元

といったことが展望できるようになる。

第 2 には、収録を予定しているデータそのものの優位性である。コーパスで収録対象とするデータは、筆者所属機関が実施している TORFL 試験（The Test of Russian as a Foreign Language = ロシア教育科学省主催「外国人のためのロシア語検定試験」）のうち、1, 2 年次における A2 (Basic level), B1 (Intermediate level) レベルの作文試験結果、約 1400 人分、約 2800 編の作文であるが、次のような特色を持つ。

- 3) データが大規模であり、作文のトピックや作成年、習熟度等、テキストや学習者属性情報の内容が明確であるという特徴を持つ。そのことで、学習者言語の特性についてより実証的で正確な観察、記述、分析を可能とするコーパス構築が志向できる。
- 4) 習得レベル別の定量比較分析や同一学習者の継時的変化を観察できる縦断的コーパスとしての活用も視野に入れることができる。
- 5) 収録予定の作文試験結果はすべて、教員による添削済みデータであり、学習者コーパスの成否を左右する誤用情報タグの付加についても、その基礎となる情報がすでに準備された状態にある。加えて、添削は TORFL で定められた客観的評価項目に基づいて実施されたものであり、それらは、誤用情報タグの分類・構成を考察する際にも有効な情報として活用できる。

### IV. 2017 年度 研究の流れ

2017 年度は、まず、他言語における既存学習者コーパスに関する情報収集と先行研究の整理を行い、その上で次の 3 点について研究調査を進めた。

- 1) HSE コーパス開発部門（ロシア連邦・モスクワ）における研究調査

現存する唯一のロシア語学習者コーパス RLC を構築した HSE コーパス開発部門（ロシア連邦・モスクワ）に滞在し、研究協力者である部門長 E.V. Raxilina 氏をはじめ、開発に関わった研究チームから開発の経緯、具体的な技術的問題、現行の稼働状況等、詳細についてヒ

アリングを行った。また、日本語母語学習者のデータ利用に基づくロシア語学習者コーパス・パイロット版構築に向けた協力体制、作業分担等の打ち合わせを行い、RLC システム上に下位コーパスとして JRFL Corpus を開設することになった。HSE コーパス開発部門での打ち合わせは研究協力者である佐山豪太氏、恒任翔吾氏が担当、詳細については佐山氏、恒任氏の報告を参照されたい。

## 2) データの電子化作業の開始

収集されている TORFL 作文試験結果について、データの電子化作業を HSE 側の助力を得て開始した。現在は一部がすでに RLC 本体のデータとしてアップロードされている。詳細は佐山氏の報告を参照のこと。

## 3) HSE コーパス開発部門でのアノテーション作業に関する研修に参加

この点の詳細についても研修に参加した恒任氏の報告を参照されたい。

## V. 今後の研究計画・課題

本科研プロジェクトは最終年度の目標として 1) コーパス・パイロット版の作成、2) コーパス・パイロット版による試験分析、3) コーパス・パイロット版のオンライン公開を予定しているが、それに向け、次年度の課題としてアノテーション・ガイドライン試案の設計がある。これについては今年度行った HSE でのヒアリング及び研修で得た知見を基に、データテキスト及び学習者の属性情報タグ、品詞情報タグ、誤用情報タグの分類・構成について試案設計を行わなければならない。

品詞情報の分類は、ロシア語関係のコーパスにおいて信頼度が高いものとして広く使用されている Russian National Corpus (<http://www.ruscorpora.ru/index.html>) における品詞タグ分類<sup>(1)</sup>を採用することを予定している。

一方、誤用情報タグの付加は一連の作業の中でもとりわけ重要な作業となる。学習者の言語使用特徴の言語学的分析、ひいてはカリキュラム・指導案開発や教材作成などに有効なデータを提供し得るかどうかが、アノテーション済みコーパスとしての信頼性に直結する。誤用情報タグの分類については、ロシア側の研究チームとの意見交換を踏まえ、RLC での基準を参考に、a) 日本語を母語とする学習者における特徴的な言語使用特性や b) 誤りの頻度が多く観察される言語領域 等を考慮した独自の分類・構成を検討し、試案を設計することが求められる。

さらに、現行 RLC はかなり緻密な誤用タグ分類、構成となっているが、あくまで個々の誤用項目についてのタグであり、1 度の検索では複合的な誤用情報が引き出せない。また、そもそも誤用情報のみに絞った検索ができず、そのままでは誤りの頻度データを得ることもできない。JRFL Corpus ではこれらの点に特に留意した設定を考慮することが必要となっている。

---

(1) この分類はロシアで開発された MyStem (<https://tech.yandex.ru/mystem/>) という品詞タグ付けシステムを使ったものである。このシステムについては恒任氏の報告に詳しい。

上記のカテゴリ分類に加え、誤り情報のタグ付けで重要な役割を担うのが、誤りの程度を示す情報の付加である。TORFL の評価基準にある「コミュニケーションに障害をもたらすかどうか」という項目は、学習者コーパス作成に際しても重要なタグ情報として採用すべきである。これらの情報は今回のデータにはすでに教員の添削内容に情報として盛り込まれており、タグ情報として利用が可能な状態にある。

誤りの程度を示す情報の付加という点は、学習者言語をどのような視点から評価していくのかという、今回のコーパス開発の基本理念にも関わってくる内容である。すなわち「母語話者」基準から逸脱するものをすべて一律の誤りとして括るのではなく、「コミュニケーションに障害をもたらすかどうか」というレベル設定を付加することで、学習者の国際語としてのロシア語能力を分析データとすることが可能になる。そこには、語学教師に根強い意識としてある「ネイティブ信仰」(鎌田 2005: 322)からの脱却、「母語話者レベル」<sup>(2)</sup>を 100% として、そこから学習者の能力を測定していくのではなく、あくまで複数言語・複数文化が共存する社会で生きていく力として、学習者が何をどこまでできるのかを見ていこうという姿勢がある。

こういった点を考慮した独自のアノテーション・ガイドラインを作成できるかどうか、次年度の課題は本科研プロジェクトの成否を決める内容になるとも言えよう。

(はやしだ りえ)

#### 参考文献

- 石川 慎一郎 (2012) 『ベーシックコーパス言語学』, ひつじ書房.  
鎌田 修 (2005) 「OPI の意義と異議—接触場面研究の必要性—」, 『言語教育の新展開』, ひつじ書房, 311-332.  
駒井 裕子 (2003) 「日本語母語話者の OPI—より明確な超級判定のために—」 ソウル OPI 国際シンポジウム (於韓国ソウル建国大学校) ハンドアウト.  
望月 通子 (2012) 「日本語教育における学習者コーパスの構築と ICLEAJ」, 『関西大学外国語学部紀要』 7号, 111-119.

#### 参照サイト

- <https://ling.hse.ru/> (最終閲覧日 : 2018 年 3 月 25 日)  
<http://web-corpora.net/RLC> (最終閲覧日 : 2018 年 3 月 25 日)  
<http://web-corpora.net/RussianLearnerCorpus/search/> (最終閲覧日 : 2018 年 3 月 25 日)  
<https://tech.yandex.ru/mystem/> (最終閲覧日 : 2018 年 3 月 25 日)  
<http://www.ruscorpora.ru/search-main.html> (最終閲覧日 : 2018 年 3 月 25 日)

---

(2) そもそも「母語話者」と言っても、それぞれのネイティブによって言語能力には、当然、差があり、「母語話者レベル」がどのようなレベルを指すのかという点にも疑問が残る。駒井 (2003) で、日本語母語話者 11 名に OPI (Oral Proficiency Interview) を行った結果、ACTFL ガイドライン (アメリカ外国語教育協会会話能力評価基準) で Advanced-Mid/-High 程度の能力 (4 段階中 3 の中—上レベル) であったという報告がなされている。

# RLC の仕組みと誤用タグ付けの実際

恒任 翔吾

(東京大学大学院人文社会系研究科博士後期課程)

Russian Learner Corpus がどのように実装されているかを概観し、誤用タグ付けがどのようになされているかをお伝えする。

## コーパスの仕組み

### データを文法素によってタグ付けする

誤用コーパスを構築する上で、単語レベルに必要なタグは2種類ある。文法上の性質を表す文法素タグと、誤用タグである。

テキストは単語ごとに文法上の性質によって検索できるようタグ付けされていなければならない。品詞、性・数・格、体・時制などで検索できなければ、コーパスの有用性を著しく減じることになる。

しかし膨大なデータを人の手で処理するには大変な手間がかかるため、文法素タグについては自動タグ付けプログラム（以下「タガー」）を用いるのがふつうだ。

---

## TreeTagger

RLC や RNC で用いられているものではないが、個人にとって使いやすいであろうタガーとして、TreeTagger[<http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>]がある。個別にダウンロードできるパラメータファイルを読み込むことで、様々な言語のテキストを解析できる。ロシア語については Serge Sharoff 氏ら[<http://corpus.leeds.ac.uk/mocky/>]の作成したパラメータファイルが存在する。

このタガーの出力は後述の MyStem よりずっとシンプルだ。実際にタグとして使用されるのは複数の文法素タグを組み合わせた文字列（single-word MorphoSyntactic Descriptions = MSDs）であり、単語とタグの2列ないし見出し語を含めた3列の表として出力される。出力されたものを表計算ソフトにでも放り込めば自在に検索可能だ。

Vmps-snmfpl	perfective	Verb	locative	
Vmps-snpfel	perfective	Verb	locative	
Vmps-snpfpl	perfective	Verb	locative	
Vmps-snpse	perfective	Verb	-	
Vm---sf---n	-	Verb	nominative	
Vm---sm-f-g	-	Verb	genitive	
Vm---sm---g	-	Verb	genitive	
Vm---sm---l	-	Verb	locative	
Vm---sm---n	-	Verb	nominative	
Vm---sn---d	-	Verb	dative	-
Vm---sn---n	-	Verb	nominative	
Vm-s-p-a-e	perfective	Verb	-	
Vm-s-p-afei	perfective	Verb	instrumental	
Vm-s-p-afen	perfective	Verb	nominative	
Vm-s-p-afe	perfective	Verb	-	
Vm-s-n-afni	perfective	Verb	instrumental	

(画像 : TreeTagger タグセット (MSDs) )

## MyStem

RLC 等で使われているタグは MyStem[<https://tech.yandex.ru/mystem/>]である。出力形式は TXT、XML、JSON からの選択が可能で、タグ付けされたデータをさらに自動で処理したい場合に使い勝手が良い。しかし TreeTagger や MSDs のようなシンプルさはなく、そのまま人にとって読みやすいものではない。

```
$ mystem -ni sappir.txt
```

```
V{v=PR=|v=S,сокр=им,ед|=S,сокр=им,мн|=S,сокр=род,ед|=S,сокр=род,мн|=S,сокр=дат,ед|=S,сокр=дат,мн|=S,сокр=вин,ед|=S,сокр=вин,мн|=S,сокр=твор,ед|=S,сокр=твор,мн|=S,сокр=пр,ед|=S,сокр=пр,мн}
мурелки{мурелка?=S,жен,од=им,мн|?=S,жен,од=род,ед|мурелка?=S,жен,неод=им,мн|?=S,жен,неод=род,ед|?=S,жен,неод=вин,мн}
шлепают{шлепать=V,несов,пе=непрош,мн,изъяв,3-л}
пельсиски{пельсиск?=S,имя,муж,од=им,мн|пельсиска?=S,жен,неод=им,мн|?=S,жен,неод=род,ед|?=S,жен,неод=вин,мн}
V{v=PR=|v=S,сокр=им,ед|=S,сокр=им,мн|=S,сокр=род,ед|=S,сокр=род,мн|=S,сокр=дат,ед|=S,сокр=дат,мн|=S,сокр=вин,ед|=S,сокр=вин,мн|=S,сокр=твор,ед|=S,сокр=твор,мн|=S,сокр=пр,ед|=S,сокр=пр,мн}
стакелках{стакелка?=S,жен,неод=пр,мн|стакелка?=S,жен,од=пр,мн|стакелок?=S,муж,неод=пр,мн|стакелка?=S,имя,жен,од=пр,мн|стакелк?=S,имя,муж,од=пр,мн|стакелк?=S,муж,неод=пр,мн}
светится{светиться=V,несов,нп=непрош,ед,изъяв,3-л}
мычай{мычай?=S,муж,неод=им,ед|?=S,муж,неод=вин,ед}
```

(画像 : MyStem 出力サンプル)

## リレーショナルデータベースにデータを格納する

検索しやすい形でデータを保管するにはデータベースを用いる。MongoDB など JSON 形式をほぼそのまま保持できるものもあるが、一般的なのは単純なテーブルを複雑に組み合わせるようしてデータを格納するリレーショナルデータベースであり、RLC もリレーショナルデータベース (MySQL) を利用している。

Building a learner corpus for Russian (Ekaterina Rakhilina, 他. [http://www.aclweb.org/anthology/W16-6509] pp. 5, 6) によれば、RLC のデータベースにおいては、テキストのメタデータレイヤー、センテンスレイヤー、形態素アノテーションレイヤー、誤用アノテーションレイヤーについてそれぞれのテーブルがあり、後述の Django フレームワークによってテキストの処理のほとんどは自動化されているという。

## Web フレームワークは利用者とデータベースをつなぐ

利用しやすい形でデータベースアクセスするためには、仲介してくれるアプリケーションが必要になる。そのようなアプリケーションの開発を容易にするのがフレームワークである。特に Web フレームワークは、ブラウザを通して操作する Web アプリケーションの開発に特化した枠組みだ。HTML、CSS などの表側の見え方と、データベースの操作などの裏側の動作をひっくるめて管理・連携できる仕組みが整っている。

Web アプリケーションの具体的な動作を簡略化して示すと、フォーム等を設置して利用者の入力待ち構え、情報が入力されたらデータベース用のクエリになおして問い合わせ、得られた結果を利用者に読みやすいようになおして提示する、というようなものだ。

RLC で利用されているフレームワークは Python ベースの Django である。

## 誤用タグ付けは Web アプリケーション上から

自動でタグ付けが可能な文法素と違い、誤用タグは人が付けていく必要がある。人間が操作するならば、人間用のインターフェースが必要だ。

RLC ではタグ付けツールも Web アプリケーションとなっている。オープンソースの Annotator[http://annotatorjs.org]をもとにしたという。実際のタグ付けページの使用感は、確かに Annotator とほとんど変わらないものだった。

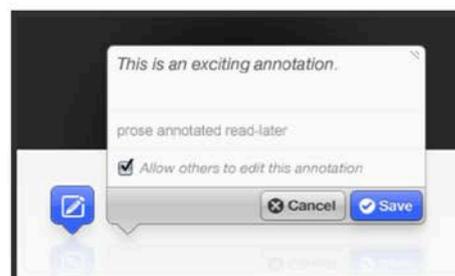
(画像 : Annotator)

# Annotator

[Code](#) / [Docs](#) / [Showcase](#) / [Plugins](#) / [News](#)

## The copy-editor of the web

Annotator is an open-source JavaScript library to easily add annotation functionality to any webpage. Annotations can have comments, tags, links, users, and more. Annotator is designed for easy extensibility so its a cinch to add a new feature or behaviour. Annotator also fosters an active developer community with contributors from four continents, building 3rd party plugins allowing the annotation of PDFs, EPUBs, videos, images, sound, and more.



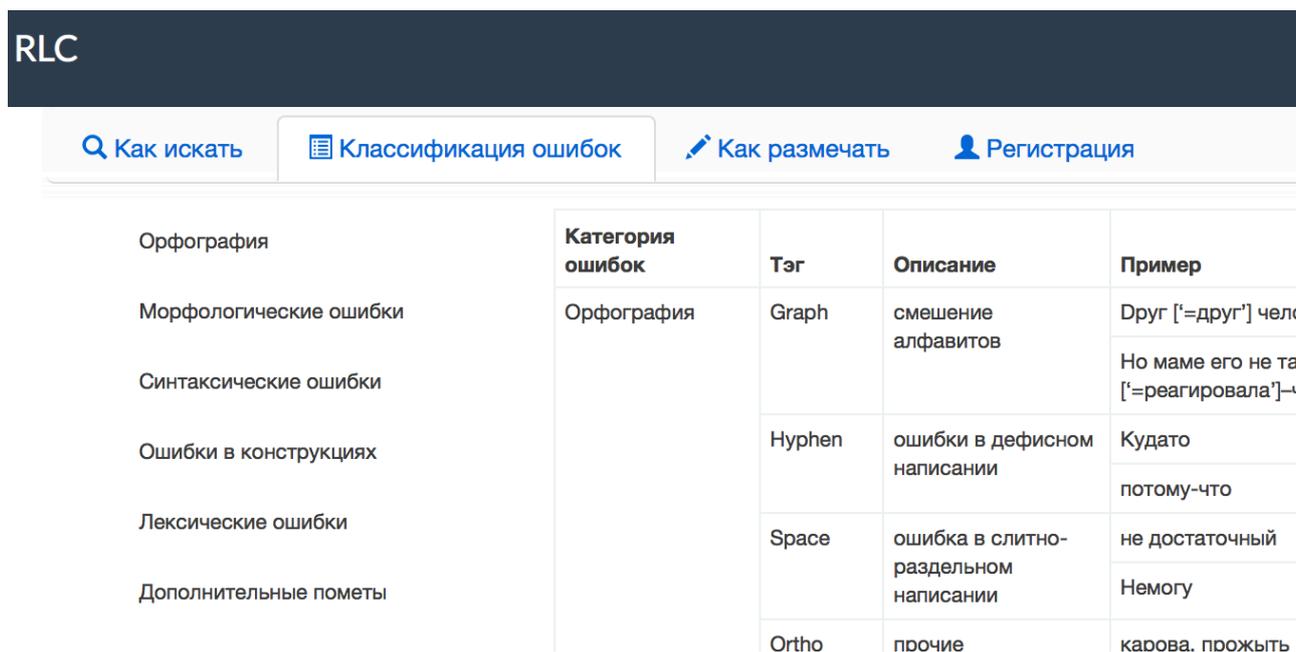
## 誤用タグ付けの実際

実際のところ誤用タグはどのように付けられるのか。筆者はモスクワでの研修に参加し、その様子をこの目で見てきた。

### 誤用タグの分類

モスクワに到着し、最初の面会で読むように指示されたのが誤用タグの分類表[\[http://web-corpora.net/RLC/help#classification\]](http://web-corpora.net/RLC/help#classification)である。シンプルながらタグ付けの唯一の基準であり、マニュアルである。

(画像: 分類スクリーンショット)



Category	Tag	Description	Example
Орфография	Graph	смешение алфавитов	Друг ["=друг"] челс
Морфологические ошибки			Но маме его не та ["=реагировала"]-ч
Синтаксические ошибки	Hyphen	ошибки в дефисном написании	Кудато
Ошибки в конструкциях			потому-что
Лексические ошибки	Space	ошибка в слитно-раздельном написании	не достаточный
Дополнительные пометы			Немогу
	Ortho	прочие	карова. прожить

少し内容を見てみよう。  
誤用はまず大きく 5 つに分類される。

1. 正書法的誤用
2. 形態論的誤用
3. 統語論的誤用
4. 構文上の誤用
5. 語彙的誤用

各大分類に具体的なタグが含まれるが、いわば誤用の文法的な「位置」を特定するものだ。基本的には、ひとつの間違いにひとつのタグが対応するよう分類されている。例えば、名詞の格にかかわる誤用でも、その要因によって排他的に異なるタグがつく。

AgгCase : ошибка в согласовании по падежу (NB Tag AgгCase не используется в случаях нарушения падежного управления: ср. с этим значением vs стрчала с мою подрушку. Ошибка в управлении описывается тегом Gov) [\[http://web-corpora.net/RLC/help#classification より\]](http://web-corpora.net/RLC/help#classification)

もちろん同じ語のうちに複数の文法事項にわたる誤用が観察される場合もあり、そのときには複数のタグがつけられる。実際、筆者の参加したタグ付けでは、アスペクトの誤用タグと時制の誤用タグをつけたものがあった。

また、上記の5分類に加えて、付加的な第6の分類がある。

## 6. 補足的なタグ

ここに含まれるのは「Transfer」「Not-clear」「Miss」など。上記5分類のタグに加えることで具体的な誤用の性質を明らかにするものである。

## タグ付けの様子

タグ付けは上記の基準に準拠して個別の作業者が行う。しかしタグ付けされて、それが承認されるまでにはその内容を精査するプロセスがある。このプロセスに日本語話者として参加したのだが、精査というのは大げさではない。セミナー室に学生を含む7、8人ほどが集まり、つけられたタグに誤りや、そもそも見逃されている誤用がないか1文ずつ確認していくのである。7、8人がかりで2時間かけてチェックできるのが2つか3つの作文データといった調子で、ここが作業進捗のボトルネックになっているのは間違いのないところだろう。

しかしその作業自体は大変面白いものだった。ネイティブスピーカーの、個々人、異なる世代、異なる性別の間で活発に議論を交わす様子は、ロシア語に潜む細かなニュアンスを多層的に理解することを大いに助けるものであるように思われた。

筆者は主に Transfer の面で議論に参加した。その教室では、例えば性・数・格等の一致という特徴が日本語にはないことも知られていなかった。説明のたびに不審な表情、好奇の表情で反応が返ってきたのは気分が良いものでさえあったが、日本語について Transfer のタグを扱える人材が先方では確保できていないということでもあろう。

チェック作業のボトルネック。日本語専門家の不足。

この2点が目についた問題点であった。

(つねとう しょうご)

# 学習者コーパスの研究利用と 今後の方針

佐山 豪太

(東京外国語大学大学院総合国際学研究所博士後期課程)

前述の通り、本科研は HSE が運営する Russian Learner Corpus (以下, RLC) のシステムを借用し、自らの学習者コーパスを完成させる: 現在, 本科研が提供した日本人ロシア語学習者の作文は HSE の RLC に検索可能な状態でその一部がアップロードされている。以下に、現行の RLC の全体像を挙げる。

表 1. RLC に含まれるテキストの全体 (アクセス日: 2018 年 3 月 21 日)

	学習者の母語	テキスト数		学習者の母語	テキスト数
1	カザフ語	261	8	フランス語	495
2	スウェーデン語	154	10	オランダ語	18
3	ノルウェー語	28	11	フィンランド語	1,234
4	イタリア語	41	12	中国語	23
5	英語	3,143	13	ドイツ語	253
6	韓国語	197	14	セルビア語	16
7	日本語 <sup>(1)</sup>	182	15	不明	321

引き続き HSE 側の助力を得て日本人学習者による残りのテキストを電子化し, RLC に順次アップロードしていく。そして, 今後, 本科研は, RLC の検索機能やタグ情報を引き継いだ独自のページを作成し, 日本人ロシア語学習者の誤用研究に適するように改良を検討する。

ただし, 前述の通り, 現在 HSE で進行している作業はテキストの電子化である。本科研のテキストに対して, 誤用タグを付ける作業はまだ未着手である。誤用タグの付いたテキストは, 学習者がおかしやすいミス傾向を示してくれる。例えば, 中国語では 3 人称の「彼」, 「彼女」, 「それ」を区別せずに 1 語 (tā) で表現するが, 中国人ロシア語学習者と話していると, он「彼」と она「彼女」という代名詞を混同している場合が見受けられる。これは, 母語の特徴がロシア語の言語運用に影響している例であると予想される。他にも, 英語で「ジミーの猫」は, Jimmy's cat となり, 日本語と語順が似ていることから, その学習は容易であると推測される。このような母語の正の転移や負の転移を検討する上で, 学習者コーパスは

---

(1) 本科研は, HSE 側に日本人ロシア語学習者 (約 1,700 名分) の作文のテキストを提供しているが, 現在, 電子化されて RLC 上にアップロードされているものは 182 にとどまる (2018 年 3 月 21 日)。手書きの作文を人力で電子テキストに変換しているため, 残りの全てが RLC にアップロードされるにはまだ時間がかかると思われる。

自らの主張を裏付ける根拠として機能する。他にも、学習者コーパスは、学習者が間違えやすい文法事項や語彙を浮かび上がらせてくれる（後述）。

本来であれば、誤用タグを用いて日本人ロシア語学習者のミス傾向を掴むことが理想であるが、現在、それを可能にするタグ付けは完了していない。そのため、ここでは教員が経験則から感じている学習者の典型的な誤用をいくつか提示し、その客観的な確認方法として RLC のデータが活用できることを参考までに示す。

### ● 合成未来の作り方

ロシア語では、不完了体動詞の未来を表す際に、*быть* と不完了体動詞の不定形という分析的な形を用いる。日本人学習者は、不完了体動詞の未来形を作る際、誤って *быть* と完了体動詞を用いてしまう場合が観察される。以下に、RLC で確認された例を挙げる。

- (1) \* Я не буду сказать нечего плохого о Сколково<sup>(2)</sup>.  
I not be-FUTURE. say-PFV.INF. nothing bad about Skolkovo-NAME.

「私はスコルコヴァについて何も悪いことは言わないだろう」

(RLC より引用：アクセス日 2018 年 3 月 21 日)

正しくは、完了体ではなく不完了体の動詞でなければならないが (*буду говорить*)、学習者は *быть* と完了体動詞の不定形を用いた(1)のような例を産出してしまう場合がある。

### ● 複数形の作り方

子音で終わる男性名詞の複数形は、末尾に *-ы* を付加することで作られる（例：「電話」*телефон* → *телефоны*）。だが、子音終わりの名詞の中には、一般的・規則的な複数形の作り方が適応されないものも存在する。例えば、*друг*「友人」は子音終わりの男性名詞であり、上記の作り方を適応すると *\*други* となるが、正しい複数形の形態は *другья* である。

ただし、子音終わりの男性名詞の複数形の作り方を *друг* にも用いた *други* という形態は、RLC 内において数多く確認される。

- (2) \* Но по-английскому, они - други.  
But in\_English they friend-PL.

「でも、英語で彼らは友人である。」

(RLC より引用：アクセス日 2018 年 3 月 21 日)

他にも、*брат*「兄弟」が *братья* となっている例も確認される。一般的・規則的な複数形の作り方が適用されない語は、基本語の中に多く含まれているため（例：*брат, учитель, стул* な

---

(2) 本報告書では必要な語にのみ次のグロスを用いる：FUTURE「未来形」、INF「不定形」、NAME「名前」、PFV「完了体」、PL「複数形」。

ど) , このようなミスが散見されるようであれば, 教材や授業でその点に言及する必要性を上げなければならないと言えよう.

- 母音の弱化

ロシア語ではアクセントのない母音 *o* と *e* は, それぞれ *a* と *u* に近い音として発音されるが, 学習者はこれらの母音を書く際に混同しがちである.

(3) \* В первом сценарий главная героиня на *карабле* сидит за писательным столом <...>

in first scenario main heroine in ship sit on desk

「最初のシナリオでは, 船においてメインヒロインが机に向かっている」

(RLC より引用 : アクセス日 2018 年 3 月 21 日)

ここでは, 「船」の綴りは *корабле* とすべきであるが, 母音の弱化の影響で *карабле* となって現れている. 基本語の中にもこのようなミスは散見されるため, 発音だけでなく, しっかりと書くことが重要であることがわかる.

(さやま ごうた)