The University of Osaka
Institutional Knowledge Archive

Title	ロシア語学習者コーパス構築の可能性と意義
Author(s)	林田, 理恵
Citation	言語文化研究. 2017, 43, p. 119-134
Version Type	VoR
URL	https://doi.org/10.18910/61277
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

ロシア語学習者コーパス構築の可能性と意義

林　田　理　恵

Возможность и значение создания нового учебного корпуса JRFLL Corpus

HAYASHIDA Rie

Цель настоящей статьи—показать возможность и значение создания нового учебного корпуса JRFLL Corpus, представляющего собой набор текстов на русском языке, написанных студентами нашего университета в течение 15 лет.

В статье рассматриваются следующие основные планы корпусного проекта: 1）план разработки перечня метатекстовой и морфологической разметки, 2）план выработки принципов построения многоуровневой системы разметки по ошибкам в текстах.

В заключение статьи, основываясь на данных планируемого корпуса, автор приводит примерный лингвометодический анализ ошибок студентов в видо-временных формах, а также анализ их залоговых конструкционных нарушений.

キーワード：学習者コーパス，学習者言語，ロシア語教育

1．はじめに

近年，目覚ましい発展を遂げてきたコーパス言語学の応用研究領域として，1990年代以降，英語を中心に学習者コーパスが相次いで構築され，日本語学習者コーパスについても2009年あたりから整備が進んでいる。日本国内でこれまでに公開されている英語，日本語学習者コーパスに関しては望月（2012）に詳しい（望月 2012: 112–115）。

一方，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。

本稿では，筆者の所属する機関で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を利用した，ロシア語学習者コーパス（=Japanese RFL Learner Corpus 以下，JRFLL Corpus）

構築の可能性とその意義を考察する。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を構築することで，日本語母語学習者のロシア語使用における多面的な言語特性を明らかにし，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と考える。

以下，ロシアで開発されている学習者コーパスについて2節で概観し，3節ではJRFLL Corpus構築に向けた，収集データ，アノテーション，比較コーパス等に関する基本構想を紹介する。また，JRFLL Corpusに収録を予定しているデータの一部を使って，アスペクト，ヴォイスの2領域でパイロット調査を実施したが，その分析結果を4節で明らかにし，将来的展望を述べる。なお，コーパス構築におけるデータ電子化等，技術的側面については別の機会にふれることとし，本稿では扱わない。

2．現存するロシア語学習者コーパス

現存する唯一のロシア語学習者コーパスであるRLCは，常時そのデータ数を拡充，またタグ付け追加作業を行っているが，現行（2016年8月）の概要は以下のとおりである。

- 収録データ：1）ロシア語を外国語として学習する（RFL）約2000人の学習者及び 2）約1500人の継承ロシア語話者（RHL）によるロシア語発話，作文データ
- 収録データ数：約72万4千語
- RFL学習者母語または継承ロシア語話者の第2言語[1)]：英語，フランス語，ドイツ語，イタリア語，韓国・朝鮮語，ノルウェー語，セルビア語
- 習熟度情報：ACTFL能力基準のIM，IH（Intermediateレベルにおける下位レベルMid，High：CEFR A2，B1に相当），AL，AM，AH（AdvancedレベルにおけるLow，Mid，Highの3下位レベル：CEFR B2レベルに相当），CEFR能力基準のC1，C2の7レベルの区分データ

収録されているテクストについては，発話／作文データの区別に加え，簡単なタスク情報（レポートか宿題か，あるいは授業内活動としての課題か，統制／非統制の区別と制限時間，テクスト内容のジャンル等）についても検索画面に表示されるようになっている。またFL/HL区分に加え，母語（FLの場合）または第2言語（HLの場合）情報も表示される。

このコーパスの特徴は，上記の概要からも分かるようにRFLのみならずRHL，すなわち継承ロシア語のデータを含む点である。国内外の多数の継承ロシア語研究者がプロジェクトの企画立案に参画し，母語やRFLと比較した継承語の言語的特徴分析や，継承語話者の言語使用ストラテジーなどの分析研究への貢献が期待されている。

また，RLCのサブコーパスとして，アメリカ国内における特定大学の学生（英語母語学習者

1）サイト上では，掲載されている言語が母語，継承語のいずれであるかを区別する情報はない。

及びロシア語継承語話者）のロシア語作文データによるRULEC（Russian Learner Corpus of Academic Writing, http://www.web-corpora.net/RLC/rulec）が作成されている。RULECは同一学習者の4年間に亘る継時的変化を観察できる縦断的コーパスとなっている。

さらに，同機関が進めているプロジェクトでは，ロシア国内大学生の作文データを収録したCoRST（Corpus of Russian Student Texts, https://www.hse.ru/org/hse/cfi/corpora/krut_，http://web-corpora.net/CoRST/search/index.php?interface_language=ru），ロシア語母語学習者のデータを収録した2つの英語学習者コーパス，RusLTC（Russian Learner Translator Corpus, https://www.hse.ru/org/hse/cfi/corpora/rusltc_，http://rus-ltc.org/），REALEC（Russian Error-Annotated Learner English Corpus, http://www.realec.org/）の開発が行われている。

そのうちCoRSTはJRFLL Corpusの参照コーパスとしてその情報の活用が期待できる内容を持ち，その点については3–3でもふれる。

現時点で公開されているCoRSTには，ロシア語を母語とするロシア国内複数大学所属の学生，約3700人による作文データが収録されている。データ数は約311万5千語である。収録作文は学期末・学年末レポート，研究会等での報告レジュメ，授業レポート，エッセーなどで，テクスト内容の分野情報も付加されている。学習者の属性情報として所属学部，性別，年齢，学年以外に，大学入学以前の居住地，継承語，第2言語の有無とその言語名といった内容が含まれる。

2013年に開発が開始されたCoRSTであるが，アノテーションは続行中で，現時点ではデータの35%で作業が終了，さらに誤用タグの付加は13%とされている。

ペレストロイカ以降，ロシア語にはドラスティックな変化が観察されると言われる。現代ロシア社会におけるグローバル化の進行，ボーダレスな人や物の移動はロシア語にも顕著な影響を与えている。特に若い世代においてどのような言語変容が起こっているのか，CoRSTはそうした変容の実態解明に向け，オーセンティックな大規模言語資料の入手と分析を目的に開発されたものである。

3．JRFLL Corpus構築に向けた構想の概要

以下，JRFLL Corpus構築に向けた基本構想について，その概要を見ていく。

3-1．収録対象とするデータと学習者属性情報

筆者が所属する機関では，ロシア語を専攻とする学生を対象に，2000年度よりロシア教育科学省主催の「外国人のためのロシア語検定試験（Типовые тесты по русскому языку как иностранному = 以下ТРКИ）」に準拠した形で1，2年次で3レベル（ТЭУ（Элементарный уровень）–A1レベル，ТБУ（Базовый уровень）–A2レベル，ТРКИ–1（Первый уровень）–B1レベ

ル）のロシア語総合試験を行っている[2)]。

ТРКИ では，CEFR-ALTE（Common European Framework of Reference for Languages: Learning, teaching, assessment; Association of Language Testers in Europe）基準による 6 段階のロシア語能力記述・学習プログラムに基づいて各レベル試験が作成されているが，今回の企画では A2，B1 レベルの作文試験結果を収録データとして予定している。現在，約 1400 人分，約 2800 編の作文がデータとして収録可能な状態にある。

ТБУ（A2），ТРКИ−1（B1）における作文試験は，年度によって多少のばらつきはあるが，おおよそ以下のようなタスクから構成されている。

【ТБУ（A2）】

時間制限 50 分　辞書の使用可

第 1 タスク

約 400 ワードの文章の概要を，与えられた質問を手がかりに作成。

第 2 タスク

与えられたテーマについて，質問を手がかりに 10 センテンス以上の文章を作成。

【ТРКИ−1（B1）】

時間制限 60 分　辞書の使用可

第 1 タスク

約 600−800 ワードのテクストの概要作成。

第 2 タスク

設定された項目を手がかりに，与えられた課題達成のための文章を作成。18−20 センテンス以上。

第 1 タスクでは「女性の職業選択」，「ソーシャルネットワーク」といったトピックが出題されている。第 2 タスクは「ロシア人の文通相手を求めて自己紹介の手紙を書く」，「ロシア留学を希望する友人のために現地の知人に紹介の手紙を書く」，「希望する企業宛にエントリーシートを作成する」といった課題となっている。

ТБУ（A2）は学習時間約 260 時間の，ТРКИ−1（B1）は学習時間約 520 時間の学習者が対象である。ТБУ（A2）を受験した同じ学習者が，約 1 年後[3)]に ТРКИ−1（B1）を受験しているので，それぞれの収録作文に学習時間数・試験レベルと合わせて受験年月情報を入れることで，個別学習

2）2011−2013 年度の 3 年間は，モスクワ大学国際教育センター・ТРКИ 試験センターとの協定によりセンターから試験問題の送付を受け，正規試験を実施。それ以外の年度はロシア国内で出版されたオープンテストに準拠して所属機関で独自試験を作成，実施している。

3）2007 年度までは ТБУ（A2）は 2 年次第 1 期末（総学習時間約 400 時間）に実施。その後，組織改編等に伴い上記のように 1 年次学年末に実施するようになった。

者の継時的変化も観察可能となる。

3-2. アノテーション

3-2-1. 品詞情報タグ

学習者コーパス構築においては，アノテーション，すなわち収録データについてどのような情報をタグとして付加していくかという点が中心的課題となる。

タグには3-1で見た学習者属性情報やテクスト作成年，トピック，テクスト語数などの情報を付与したメタ・データタグと並んで，品詞情報や誤用情報に関する言語情報タグがある。

ロシア語関係のコーパスにおいては，Russian National Corpus（Национальный корпус русского языка = НКРЯ, http://www.ruscorpora.ru/index.html）における品詞タグ分類が信頼度が高いものとして広く使用されている。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，表1はその分類表である。JRFLL Corpusにおいても公開されているMystemの利用が可能である。

表1　Russian National Corpus　品詞タグ分類

Часть речи ☐ существительное ☐ прилагательное ☐ числительное ☐ числ-прил ☐ глагол ☐ наречие ☐ предикатив ☐ вводное слово ☐ мест-сущ ☐ мест-прил ☐ мест-предикатив ☐ местоименное наречие ☐ предлог ☐ союз ☐ частица ☐ междометие	**Падеж** ☐ именительный ☐ звательный* ☐ родительный ☐ родительный 2 ☐ дательный ☐ винительный ☐ винительный 2* ☐ творительный ☐ предложный ☐ предложный 2 ☐ счётная форма	**Наклонение / Форма** ☐ изъявительное ☐ повелительное ☐ повелительное 2 ☐ инфинитив ☐ причастие ☐ деепричастие	**Степень / Краткость** ☐ сравнительная ☐ сравнительная 2 ☐ превосходная ☐ полная форма ☐ краткая форма
		Время ☐ настоящее ☐ будущее ☐ прошедшее	**Переходность** ☐ переходный* ☐ непереходный*
	Число ☐ единственное ☐ множественное	**Лицо** ☐ первое ☐ второе ☐ третье	**Прочее** ☐ цифровая запись ☐ аномальная форма* ☐ искаженная форма* ☐ инициал* ☐ сокращение* ☐ несклоняемое* ☐ топоним**
Имена собственные ☐ фамилия ☐ имя ☐ отчество	**Род** ☐ мужской ☐ женский ☐ средний ☐ общий*	**Залог** ☐ действительный ☐ страдательный ☐ медиальный	
	Одушевленность ☐ одушевленное ☐ неодушевленное	**Вид** ☐ совершенный ☐ несовершенный	

3-2-2. 誤用情報タグ

学習者コーパスでは，とりわけ重要な作業として誤用情報タグの付加がある。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。

JRFLL Corpusで収録対象とするA2，B1レベル作文試験結果はすべて，教員による添削済みデータである。添削はТБУ(A2)，ТРКИ-1(B1)で定められている評価項目（表2）に従って行われている。

表2 ТРКИ・ライティング評価項目

1）	Адекватность создаваемого тестируемым текста цели, поставленной в задании コミュニケーション課題をこなしているか
2）	Соответствие заданному объему высказывания　必要文章量
3）	Полнота представления информации　基本情報が漏れなく伝えられているか
4）	Точность передачи информации 情報（状況等の客観的情報と著者の意図，構想等についての情報）が正確に伝えられているか
5）	Логичность и связность изложения　テクストの論理性，結束性，首尾一貫性
6）	Коммуникативно значимые ошибки（КЗО）　コミュニケーションに支障のある誤り
7）	Потенциально коммуникативно значимые ошибки　КЗОになる危険性をもつ誤り
8）	Коммуникативно незначимые ошибки　コミュニケーションに支障のない誤り

（Косарева и др. 2010: 5,45）

今回のコーパスでの誤用タグ付け対象となるのは表2，(5)-(8)の項目での添削部分である。(5)はディスコースレベルでの誤用，(6)-(8)については語彙レベル，形態レベル，構文レベルなど，多様な誤用タイプが混在している。したがって，さらに誤用表現のカテゴリ分類を行い，その情報をタグ情報として付加していくことが必要となる。以下，信頼性の高い言語資料の提供を目標とした場合，どのような分類が可能かつ有効であるか，試案を提示しつつ考察する。

誤用タイプ

1）語彙レベルの誤用

1. 1. 意味的誤用

1. 2. 語結合上の誤用

1. 3. 形態的誤用（綴りにおける誤り，語形成的誤用 etc.）

2）文法レベルの誤用

2. 1. 語形変化における誤用（名辞類 性・数・格変化，形容詞 短語尾／長語尾区別，形容詞 比較級形成，動詞人称変化，分詞形態 etc. における誤用）

2. 2. 時制使用上の誤用

2. 3. 条件法，命令法使用上の誤用

2. 4. アスペクト表現における誤用

2. 5. ヴォイスにおける誤用（受動文使用における誤用 etc.）

2. 6. モダリティ表現における誤用

2. 7. 統語上の誤用

2. 7. 1. 単文レベルにおける一致，支配，付加関係の誤用

2. 7. 2. 文タイプ構造（人称文／無人称文／不定人称文／普遍人称文，平叙文／疑問文／命令文／感嘆文）使用における誤用

2. 7. 3. 複文構造における誤用（接続詞，関係詞使用上の誤用 etc.）

2. 7. 4. 直接／間接話法使用上の誤用

2. 7. 5. 語順における誤用

3）ディスコースレベルの誤用

3. 1. 結束性における誤用

3. 2. 首尾一貫性における誤用

3. 3. 提題機能における誤用

3. 4. 指示，照応，省略 etc. における誤用

4）文体上の誤用

以上の分類はあくまで現段階における試案であり，実際のタグ付け作業過程を経ることでより有効なカテゴリ化の整備がなされ，修正が施されていくことが予定される。

例えば表 2 の（1）は，テクストが全体として設定されたタスクに対応した内容になっているかどうかを判断する項目であり，タグ化が難しく，現段階では誤用タグの対象とはしていない。ただ，学習者の習得状態を示す情報として重要な内容を示しており，それらをどのように扱うかは今後の課題となる。

また当然のことながら，誤りの各々は多くの場合，複数の項目にまたがる内容を持つ。例えば，再帰動詞を使用すべきところを再帰接辞 -ся のない他動詞が誤って使われた場合，語彙レベルの形態的誤用であると同時に，統語上の支配関係（自／他動詞の区別）の理解不足が原因であるとも考えられる。また，不完了体動詞・現在時制の使用が適切な場合に，完了体人称変化形が使われていれば，アスペクト上の誤用であると同時に時制の誤用でもあり，さらに語彙の形態的誤用である可能性もある。したがって，誤り情報のタグ付けは重層的な構造を想定したものが準備されなければならないだろう。

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。JRFLL 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。

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。TRKI 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。

表2の（6)-(8）の区分はそうした基準を反映した項目となっている。これらの情報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，今回の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，語学教師にいまだ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」[4)]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。

3-3. 比較コーパス

JRFLL Corpusの開発・構築によって，ロシア語学習における日本語母語話者特有の傾向を分析することが可能になるが，さらにロシア語母語話者，日本語母語話者の双方から収集された作文を比較コーパスとして使用することで，学習者の学習言語特性に母語や学習環境など，どのようなファクターが反映しているのか，あるいはしていないのかなどについて，実証データによってそのヒントが与えられる可能性も出てくる。

JRFLL Corpus開発プロジェクトでは，日本語を母語とするロシア語学習者に，収録予定の作文データと同じトピックで，かつ同条件で母語による作文作成を依頼し，収集されたデータを比較コーパスとする予定である。また，参照コーパスとして，収録作文の作成者と同世代のロシア国内大学生のデータを扱っており，トピックも類似している上記のCoRSTを利用することも可能である。

4．JRFLL Corpus分析の可能性と意義

この節では，アスペクト，ヴォイスという2つの領域を例に，JRFLL Corpus分析を通してロシア語学習における日本語母語学習者の言語使用特徴をどのように明示化し得るのか，その可能性と意義を検討する。JRFLL Corpusは先にも見たようにCEFR，A2-B1レベル学習者の作文データをベースとしたものを予定しているが，アスペクト，ヴォイスは当該レベルではとりわけ習得に負荷がかかるとされる2領域であり，有効な指導法や教材作成に向け，Corpus分析結果によって得られる情報の応用が期待される。

4）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の中ー上レベル）であったという報告がなされている。

4-1. アスペクト領域における Japanese RFL 学習者言語の特徴

ロシア語においては，大多数の動詞が完了体／不完了体というアスペクトの文法的表示を担う 2 項対立システムに組み込まれている。そういったアスペクト体系を持たない言語を母語とする学習者にとって，この現象の存在は習得上の高い壁になることが一般的には予想される。

それでは実際に学習者の言語使用において，アスペクト表現上のどのような傾向が観察されるのか，JRFLL Corpus 分析を通して得られることが期待される情報について，以下，収録データとなる言語資料を実際に検討することで，その可能性について見ていきたい。

（1） В свободное время я смотрю фильм, и люблю покупках. Я буду возврашаться к Эхиме на летних каникулах.

（2） На летних каникулах я буду ехать в Токио туризмом.

（3） На летних каникулах, я поеду в Иркутске. Буду поехать в Байкал, идти за покупками, поболтать с русским друзьями.

（4） На летних каникулах, я буду поехать на Англие.

（5） В свободное время, давай вместе. Я живу в городе Осака. Приходите в гости ко мне кагда ты ездешь в Японии.

（6） Куда ты поедешь на летних каниклах? Если ты приедешь в Японии, встречаем !!

ロシア語の未来時制は，一般には不完了体動詞では「быть 人称変化形 ＋ 不完了体 Inf.」という合成形を，完了体動詞の場合は人称変化形そのものが未来を示すとされ，学習の早い段階（A1–A2 レベル）で重要な習得項目として同様の文法説明がなされる。

（1）–（6）は収録データとして予定している ТБУ（A2）での作文タスクからの例であるが，いずれの文例においても未来表現部分（　部分）において動詞アスペクトの理解が関係する誤用が観察される。

（1），（2）と（3）の「идти」部分は完了体動詞人称変化形が使われるべきところに不完了体動詞の未来形式が用いられている。また，（3）の「Буду поехать」，「поболтать」，（4）の「буду поехать」では，完了体動詞人称変化形が使われるべきところに「быть 人称変化形 ＋ 完了体動詞 Inf.」が使用されている。さらに，（5），（6）の文脈においても，本来ならば完了体動詞人称変化形による未来表現が求められるところであるが，現在を表現する不完了体動詞人称変化形の

誤用が観察される。

これらのデータだけを個別に分析すれば，(1)，(2) 及び (3) の「идти」部分は，時制におけるアスペクト対立関係が正しく理解できていないと判断されることになる。(3) の「Буду поехать」，「поболтать」と (4) からは，未来表現における文法形態のみならず，動詞・アスペクト形式そのものが未習得である可能性が示唆され，(5)，(6) に対しては，時制における誤用という判断が下されるであろう。

ところで，今回，試験的に分析の対象とした単年度分，31編の作文例において，未来表現におけるアスペクト領域の誤用がみられる学習者は実に全体の 3 分の 1，11名に上っているが，そのすべてが完了体人称変化形に対する誤用であるという点は興味深い。

ロシア語学習では，動詞・未来形は「быть 人称変化形 + 不完了体 Inf.」という不完了体合成形が最初に導入され，完了体人称変化形による未来表現は後の単元で出てくる場合が多い。ロシアで出版されている教科書などでも通常，この順番が採用されている。恐らく，完了体人称変化形には子音交替や力点移動が頻繁に起こる第 2 変化や不規則変化が多く，習得に負荷がかかると判断されるためであろう。その一方で，不完了体合成未来形は形態的に習得が容易であるという判断がある。

しかしながら，実際の言語データでは，ロシア語未来時制において完了体人称変化形使用が圧倒的優位を示すのである。表 3 は林田（2007）で分析したロシア語中・長編 8 作品での完了体人称変化形[5)]と不完了体未来形の出現度数を比較したものである。完了体人称変化形：不完了体未来形の割合が約 8：1となっていることが分かる。現代ロシア語において未来の出来事表現では，不完了体合成未来形は基本的には途中経過と多回の強調表現として機能しており，限定された使用領域しかもたないのである。

表 3　未来時制における完了体人称変化形／不完了体未来形の出現度数

	完了体人称変化形	不完了体未来形
«Игра на чужом поле»	373	91
«Первый учитель»	126	31
«Убийца поневоле»	646	93
«Деньги для Марии»	432	55
«Свой круг»	39	5
«Прощание с Матерой»	498	34
«Джамиля»	100	11
«Белое облако Чингисхана»	191	20
計	2405	340

ロシア語教育の現場では，こうした言語事実が考慮されることなく，両形式の差は単なるア

5）完了体人称変化形のうち，1，3 人称命令法での使用は除いている。作品名については作品出典参照。

スペクト機能上の異なり―プロセス／完了，反復／1回 etc.―として説明されるにとどまる。

未来時制表現において，完了体人称変化形の誤用が全体の3分の1の学習者に観察される一方，不完了体合成未来形の誤用は見られないという上記の現象は，教育現場での言語事実に立脚しない指導にその原因がある可能性が高い。単に習得に負荷がかからないという理由のみで不完了体合成未来形が学習の初期段階で導入されることで，その後，完了体人称変化形が導入された後も，未来表現において学習者はすでに習得ずみの，よりたやすく再生できる不完了体合成未来形をもっぱら選択したと考えられるのである。

今回は予定しているデータの一部についてのみの試験的分析結果であるが，データを計量的に分析することで初めて，1）未来表現のアスペクト領域での誤用出現に極端な偏りがあるという事実が明らかになり，データが示す誤用に関して，単にこの時制でのアスペクト形態・機能の未習得という点にとどまらず，2）不完了体合成未来形の過剰使用，さらには 3）学習者の完了体人称変化形に対する回避方略の可能性が見えてきたのである。

4-2. ヴォイス領域における Japanese RFL 学習者言語の特徴

A2-B1レベルのロシア語学習において，ヴォイス関係，特に受動表現並びに再帰動詞を用いた中動相表現は，アスペクトと並んで重要な習得項目である。

ロシア語では受動文は動詞アスペクトによって2種類の表現形式をもつ。完了体動詞は受動過去分詞を形成し，「コピュラ＋分詞」形―以下，分詞形―によって，不完了体動詞の場合は再帰動詞によって―以下，再帰形―受動が表現される。ただし，再帰形受動文で表現される受動の意味は，再帰動詞が本来持つ中動的意義の個別・特殊な変異として存在する。再帰受動文では主語の習慣的，一般的状態，恒常的属性の表現という中動的意義の延長線上の意味が主に表現される。一方，不完了体動詞の基本的なアスペクト意味である特定時点における現実的持続相は観察されないのである。

今回，受動表現についても3-1の場合と同様に，収録データとして予定しているТРКИ-1（B1）作文タスクでの学習者の作文から，2年度分，計63編を抽出し，受動表現に関わる言語使用の傾向を試験的に分析した。

この2年度分のタスク課題は，あらかじめ3000字程度のロシア語文テクストが与えられ，テクスト内容の概要とそれについての自らの考えを記述するというものである。与えられたロシア語テクストは「現代社会における女性」，「ソーシャルネットワーク」というテーマである。

63編の作文中，不完了体動詞から派生した再帰形受動文が17例，完了体動詞派生の分詞形受動文は15例の使用が観察された。そのうち誤用は以下の1例のみが認められる。

（7） Но если женщины работают, как мужчины, воспитание детей должно поразделиться, поэтому мужчинам надо понять важность своего участия в воспитании детей.

（7）はどの点が強調されるかによって再帰形受動文，分詞形受動文のいずれもが選択可能な文脈である。習慣的行為という点が強調される場合は不完了体派生の再帰形受動文が選択される。もし「必要」というモダリティーに焦点があたれば，完了体派生の分詞形受動文が選択され得る。

ただ，学習者の作文例では完了体動詞に接辞 -ся が付加され，動詞形態における誤用―不完了体動詞の形態的な誤用―，あるいは完了体派生受動文表現の未習得―分詞を使うべきところを再帰接辞を付加―といういずれかの可能性が考えられる。

上記以外はすべて正しい受動文表現が用いられているが，1 例以外はすべて元々，与えられたテクスト内で使用されていた表現をそのまま書き写したもので，結果として学習者自身が自律的に産出した受動文表現は，上の誤用例を加えてもたった 2 例しかないということになる。

学習者コーパスでの未出現項目については，迫田（2011）や石川（2012）で指摘されているように，それらの項目が未習得なのか，それともコーパスサイズやトピックの制約によるものかを特定することは，コーパスからだけではできない（迫田 2011: 8，石川 2012: 220）。

問題となっている受動文表現の場合，あるトピックに関する発話・作文の際に，そもそも話者のイメージ形成において受動を選択する動機が存在するかどうかがまず問題になる。その検証には，同様のトピックでのロシア語母語話者作文との比較が一つの手がかりとなる。本稿では暫定的に，上で紹介した CoRST で類似トピックデータが抽出されるような検索条件を設定し，その下位コーパスで分詞形受動文使用の観察を試みた[6]。以下はその結果である。

テクスト種類：エッセー，学年レポート，学期末レポート，文献要約
分野：社会学，文化学
下位コーパスデータ数：1850 編（全データ数 3677 編の 50.31%）
分詞形受動文使用データ数：9051 例（1 編平均，約 5 例の使用）

もちろん，トピックやテクスト種類，テクスト分量（CoRST では 1 編の語数情報等は表示されない）等で制約条件が異なっており，あくまで目安としての比較であるが，類似トピックにおいて母語話者の分詞形受動文使用頻度は相当高いと言え，日本語母語学習者データとの差異が際立つ。

より精密な分析のためには同じ制約条件で収集したデータとの比較が必要であるが，ロシア語母語話者データのみならず，同じトピック，テクスト分量，制約時間で作成した学習者の母語による作文データも重要な検討対象となるだろう。受動表現において学習言語と母語が持つ機能上の異同点が使用頻度の差異に関与しているかどうかが，それらのデータによって明らか

6）接辞 -ся をもつ再帰動詞は多義的であり，それが受動の意味で使用されているかどうかは文脈で判断するしかなく，コーパスの検索条件では受動用法のみを抽出することができないので，分詞形受動文のみを対象としている。

になる。

さらに，コーパスでデータが出現しなかった項目については，その理由を特定するための補完的な実験調査が不可欠であろう。

今回対象とした受動文使用については，別の機会に学習時間約350時間（A2レベルテスト合格者）の学習者32名を対象に，日本語受動文7例を含む20文のロシア語文への書き換えを求めるテストを実施した。表4がその結果である。実施時期は受動文表現導入の半月後で，テスト制限時間は15分，辞書使用は認めていない。

表4　受動表現文法性判断テストにおける誤用数と内訳

誤用総数／使用総数	128/224	57.1%
未解答／誤用総数	50/128	39.1%
形態的誤用／誤用総数	35/128	27.3%
分詞形→再帰形／誤用総数	22/128	17.2%
語彙選択における誤用／誤用総数	21/128	16.4%

テスト実施までに文型練習や文法練習課題，受動文を使用したテクストの読解・聴解練習，受動文表現を誘発するようなテーマでの作文課題やディベートなど，4技能すべてを対象とした授業活動が繰り返し行われていた。にもかかわらず，表からも分かるように誤用率はかなり高い。形態的誤用の多く―35例中27例―は単なる語形変化的なものではなく，分詞形の形成自体ができていない，あるいは能動動詞がそのまま使われているといった誤りである。未解答が50例に上っていることを重ね合わせると，ロシア語受動文の体系的理解ができていない学習者が，この時点でも4割近く存在していることを示している。

また，完了体派生分詞形受動文を使用すべきところを，完了体動詞に再帰接辞 -ся を付加したものを使用した学習者も22名に上り，受動分詞形の形成が学習者にとって高い障壁となっていることが分かる。

こうした文法性判断テスト結果や先のCoRSTの数値を踏まえれば，上記TPKИ-1（B1）作文タスク・データで受動文の自律的使用がほとんど観察されないという事実が新たな様相を帯びる。つまり，問題となる現象は単なるコーパスサイズやトピックの制約によるものではなく，原因として学習者の回避方略の可能性が示唆されるのである。

誘導型の文法性判断テストでは，学習者の半数以上が未解答や誤用，さらには再帰形の過剰一般化といった結果を出すが，B1レベルテスト―学習時間540時間の学習者対象―における非統制の作文タスクでは，ほとんどの学習者があまり自信のない受動文表現を回避し，能動文などで置き換えて表現するという行動をとっていると考えられる。

受動文表現は事実伝達という側面からは，一見，能動文や不定人称文と同内容の意義を持つと考えられ，それらで代替するという回避方略がとりやすい項目とされる。ただ，受動文表現

にはテクスト結束性や主題選択にかかわる重要な機能があり，さらには分詞形受動文の場合は，伝達される事実内容そのものも能動文や不定人称文とは異なるケースが多い。その点を考慮すると，受動文表現を射程したアカデミック・ライティングの特別な訓練など，学習者の受動文回避克服のための方策の必要性が浮上してくる。こうした方向性は，文法性判断テスト結果の分析だけでは導き出しにくいものであり，コーパス・データ分析と統合することで見出すことが可能になる。

4-3. まとめと将来的展望

以上，アスペクト，ヴォイス領域における未来表現と受動表現という 2 つの項目について，今回のコーパス構築で予定しているデータの一部を抽出・分析し，この領域における学習者言語の傾向，特性の一端を明らかにしてきた。これらの分析結果はあくまでパイロット調査として，暫定的性格を持つものであるが，今後，JRFLL Corpus が構築されれば，大規模な実証データに基づいたさらに精度の高い，有益な情報を得る可能性が開けると予想される。

ロシア語初級レベル（A1–A2）における習得困難項目である名辞類格変化や動詞人称変化などは，学習者コーパスを使用すれば瞬時にしてその誤用の傾向，また学習者がその時点で作り上げている学習者言語体系の実像が明らかになるであろう。語彙についても，CoRST 等，ロシア語母語話者コーパスとの比較により，日本語母語学習者における過剰・過小使用といった特徴的傾向を明らかにしていくことができる。

また，アスペクトやヴォイスに関係する事項は，それぞれがいくつもの領域にまたがる複層的機能を持つが故に，大規模コーパスの使用が有効な項目である。限定的事例だけを扱う文法性判断テストや質問紙調査，インタビューなどでは得られない情報獲得の可能性が高まる。アスペクトやヴォイスの複層的機能について，学習者はどのような理解を形成しているのか。それらの項目が持つ多岐にわたる使用場面で，実際に学習者はどのような言語使用を行っているのか。学習者言語の詳細な実態解明に向け，コーパスという大規模実証データの使用は確かな手掛かりを与えてくれるはずだ。

これまでロシア語教育の現場では，学習者言語の傾向や特性に関して指導者の経験値に依拠した議論が中心であり，そのような主観的判断に基づいて教材作成やカリキュラム，指導案作りがなされてきた。今後，ロシア語学習者コーパスの開発・整備が進めば，より客観的，実証的データに基づいた分析が可能になり，それらの情報と各種の実験調査結果とを組み合わせることで，教育現場への有効な還元を図っていくことができよう。

今回は収集された TPKИ 作文タスクのデータを使ったコーパスの企画であるが，同じ TPKИ の口頭能力試験における発話データについても，現在，15 年間分が収集されており，モノローグ，会話，テクスト要約などの音声データが利用可能である。将来的に作文データでの運用が軌道に乗った時点で，JRFLL Corpus に音声データ及び文字化データを追加していくことが考え

られる。作文データと並行して発話データを収録することで，「書く」活動，「話す」活動それぞれにおいて学習者の習得プロセスにどのような差が観察されるかなど，教育面のみならず言語学的にも興味深い実証研究を行うことができるようになるであろう。

参考文献

林田理惠（2007）「第 2 部ロシア語のアスペクト」，『ロシア語のアスペクト』，南雲堂フェニックス.

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.

Косарева, Е. В., М. Ф. Лазарева, Т. А. Шутова и Е. Е. Юрков 2010. *Лингводидактическое тестирование: процедура и методика проведения в рамках российской государственной системы тестирования: Субтесты «Письмо», «Говорение»: ТЭУ, ТБУ, ТРКИ–1.* СПб., «МИРС».

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.

迫田久美子（2011）「第二言語習得研究の深さと広がり ―学習者の学び方から教師の教え方へ―」，*Japanese Studies Journal*, 53/54, 1–18.

参照サイト

https://ling.hse.ru/（最終閲覧日：2016 年 9 月 13 日）

http://web-corpora.net/RLC（最終閲覧日：2016 年 9 月 13 日）

http://web-corpora.net/RussianLearnerCorpus/search/（最終閲覧日：2016 年 9 月 13 日）

http://www.web-corpora.net/RLC/rulec（最終閲覧日：2016 年 9 月 13 日）

https://www.hse.ru/org/hse/cfi/corpora/krut_（最終閲覧日：2016 年 9 月 13 日）

http://web-corpora.net/CoRST/search/index.php?interface_language=ru
（最終閲覧日：2016 年 9 月 13 日）

https://www.hse.ru/org/hse/cfi/corpora/rusltc_（最終閲覧日：2016 年 9 月 13 日）

http://rus-ltc.org/（最終閲覧日：2016 年 9 月 13 日）

http://www.realec.org/（最終閲覧日：2016 年 9 月 13 日）

http://www.ruscorpora.ru/index.html（最終閲覧日：2016 年 9 月 13 日）

https://tech.yandex.ru/mystem/（最終閲覧日：2016年9月13日）

データ用作品出典一覧

Айтматов, Ч. *Джамиля.* Киев, 1976. «Джамиля»

Айтматов, Ч. *Первый учитель.* Киев, 1976. «Первый учитель»

Айтматов, Ч. *Белое облако Чингисхана.* Знамя. 1990. № 8. «Белое облако»

Маринина, А. *Убийца поневоле.* М., 1998. «Убийца поневоле»

Маринина, А. *Игра на чужом поле.* http://www.allbest.ru/ «Игра на чужом поле»

Петрушевская, Л. С. *Свой круг.* М., 1995. «Свой круг»

Распутин, В. *Деньги для Марии.* М., 1984. «Деньги для Марии»

Распутин, В. *Прощание с Матерой.* М., 1984. «Прощание»

（行末の《 》内は本文中に記載した出典表示を示す。）