

**Sandra Birzer**, Otto-Friedrich-Universität Bamberg, Deutschland

**Hellik Mayer**, Universität Regensburg, Deutschland

**Anton Gomer**, Otto-Friedrich-Universität Bamberg, Deutschland

## Lernerurteile zu lexikalischen Schwierigkeiten in russischen Lesetexten

This study investigates factors that contribute to the perceived difficulty of the Russian vocabulary among learners at different CEFR levels. We tested four hypotheses concerning word formation processes, learners' familiarity with lexemes, morphological alternations, and internationalisms. The first hypothesis stated that learners on advanced CEFR levels are more familiar with the productive word formation processes and will therefore experience difficulties only with opaque word formations; this hypothesis was confirmed. The second hypothesis stated that learners on advanced CEFR levels were more likely to know lexemes from the same word family and to use this to make inferences about new words. This hypothesis was also verified. The third hypothesis was not supported: lexemes with morphological alternation are perceived as difficult only at lower levels. The fourth hypothesis posited that internationalisms relieve learners. This hypothesis was confirmed up to level B2, whereas on level C1 internationalisms are register-specific terms and are often perceived as difficult.

---

Keywords: CEFR levels, degree of perceived difficulty, reading assignment, lexicon

---

### 1. Einleitung

Die Daten, die im Rahmen dieses Beitrags analysiert werden, stammen aus einer experimentellen Studie mit russischen Lesetexten, die aus Lehrbüchern für Russisch als Fremdsprache entnommen wurden. Das Experiment ist Teil eines größeren Projekts zur automatischen Klassifikation von russischen Lesetexten in sechs Schwierigkeitsstufen, in dessen Rahmen das "Levelled study corpus of Russian" (LeStCoR) erstellt wurde (vgl. Batinić & Zinsmeister, 2017 und Lemmenmeier-Batinić, 2019).<sup>1</sup> Das Korpus enthält knapp 2500 Lesetexte aus über 50

---

<sup>1</sup> Das Korpus ist u.a. aus Gründen des Copyrights bisher noch nicht öffentlich zugänglich.



verschiedenen Lehrwerken für Russisch als Fremdsprache, die in unterschiedlichen Ländern erschienen sind und alle GERS-Niveaus repräsentieren. Alle verwendeten Lehrwerke enthielten einen Hinweis, auf welchem GERS-Niveau ihr Einsatz vorgesehen ist; bei Einspeisung in LeStCoR wurden die Texte zudem morphologisch und syntaktisch annotiert.

Bei dem automatischen Klassifikationsversuch von Lemmenmeier-Batinić (2019) zeigte sich, dass sich bei gut 30% aller Texte aus dem LeStCoR-Korpus die Einstufung des GERS-Niveaus durch den im Rahmen des LeStCoR-Projekts erarbeiteten automatischen Klassifikators von jener in den Lehrwerken unterschied. Es stellte sich also die Frage, welche der beiden GERS-Einstufungen – jene der Expert\*innen, welche die Lehrwerke verfasst hatten, oder die des Klassifikators – besser widerspiegelt, wie die Lernenden selbst den Schwierigkeitsgrad der jeweiligen Texte wahrnehmen. Die Einschätzungen der Lernenden ermöglichen so eine empirische Validierung der Einstufung durch Expert\*innen bzw. den Klassifikator. Die Reflexion der durch die Lernenden wahrgenommenen Schwierigkeiten stellt ein Desiderat nicht nur im Sinne der Lernorientierung dar, sondern auch für die automatische Textklassifikation: Reynolds (2021) fordert die Operationalisierung der durch die Lernenden wahrgenommenen Schwierigkeiten als Klassifikationsmerkmale (*features*) für die automatische Klassifikation von Lesetexten in Schwierigkeitsgrade.

Ziel unseres Experiments war deshalb festzustellen, welche lexikalischen Einheiten, morphologischen Phänomene und morphosyntaktischen Strukturen den Lernenden unterschiedlicher GERS-Niveaus Schwierigkeiten beim Textverständnis bereiten und wie dies mit den Einstufungen der Expert\*innen bzw. des Klassifikators korreliert. Eine erste Analyse der Daten aus diesem Experiment zeigte, dass die Einstufungen des Klassifikators häufiger mit den Einschätzungen der Lernenden (der Kürze halber im folgenden als Lernerurteile bezeichnet) korrelierte als jene der Expert\*innen (Birzer & Mayer, 2022).

Dieser Beitrag bietet eine Analyse der lexikalischen Schwierigkeiten, welche die Teilnehmer\*innen unseres Experiments wahrnahmen (zu den morphologischen und morphosyntaktischen Schwierigkeiten vgl. Birzer & Mayer, 2022). Zunächst werden wir die Hypothesen vorstellen, die unserer Analyse zu Grunde liegen (Abschnitt 2). Dem schließt sich eine Beschreibung des Experiments (Abschnitt 3.1), der erhobenen Daten (Abschnitt 3.2) sowie ihre Analyse (Abschnitt 3.3.) an; im Fazit (Abschnitt 4) diskutieren wir, inwiefern die Ergebnisse unserer Analyse die Hypothesen verifizieren bzw. falsifizieren.

## 2. Ziel der Studie und Fragestellung / Hypothesen

Ziel dieser Studie ist es, Lernerurteile zu „schwierigen“, d.h. von den Teilnehmer\*innen am Experiment als schwierig markierten lexikalischen Elementen zu nutzen, um

1. Faktoren zu identifizieren, welche die Wahrnehmung der besagten lexikalischen Einheiten als „schwierig“ erklären können und
2. diese Faktoren für die einzelnen GERS-Stufen zu modellieren.

Dazu formulieren wir folgende Hypothesen bezüglich möglicher Faktoren:

- **H1: Transparenz bzw. Opakheit der Wortbildungsverfahren**

Transparente Wortbildungsverfahren sind in der Regel produktive Verfahren, sodass bei einer gewissen Typenfrequenz die jeweiligen Verfahren aus dem bekannten Wortschatz abgeleitet und transferiert werden können (vgl. Krasil'nikova, 2015; Tafel, 2009, 233–245 und Tažev, 1997, 14–17 speziell zu den Substantiven und Verben). Opake Verfahren hingegen lassen solche Regelmäßigkeiten nicht erkennen, sodass die Größe des bekannten Wortschatzes und daraus abgeleitete Wortbildungsmuster bei der Erschließung nicht hilfreich sind. Wir gehen deshalb davon aus, dass mit steigendem GERS-Niveau umso mehr Lexeme mit opakem Wortbildungsverfahren als schwierig wahrgenommen werden, als die Lernenden viele produktive Wortbildungsverfahren bereits kennen und bei der Erschließung der Wortbedeutung darauf zurückgreifen können.

- **H2: Bereits bekannte Lexeme aus demselben Wortnest (словообразовательное гнездо)**

Alle Lexeme eines Wortnests gehen auf dasselbe lexikalische Wurzelmorphem zurück. Kennen die Lernenden also bereits ein Lexem aus jenem Wortnest, zu dem auch das zu analysierende Lexem gehört, können die Lernenden für das neue Lexem semantische Inferenzen anstellen. Weil es auf den Niveaus A1–A2 aufgrund des begrenzten Wortschatzes eher unwahrscheinlich ist, dass weitere Lexeme desselben Wortnests bereits bekannt sind und für Inferenzen herangezogen werden können, gehen wir davon aus, dass die Wahrscheinlichkeit für semantische Inferenzen mit steigendem GERS-Niveau anwächst.

- **H3: Morphologische Alternationen**

Morphologische Alternationen sind regelhaft und treten häufig als Ergebnis derivationaler Wortbildung auf (z. B. книга > книжный oder книжка, siehe dazu wenig konkret Diuf, 2011). Deshalb ist anzunehmen, dass die Regelmäßigkeit morphologischer Alternationen umso bewusster ist, je größer der bereits

bekannte Wortschatz und damit die Möglichkeit zur Ableitung dieser Regeln ist. Wir gehen deshalb davon aus, dass Lexeme mit morphologischer Alternation umso häufiger als „schwierig“ wahrgenommen werden, je niedriger das GERS-Niveau der Lernenden ist.

- **H4: Internationalismen**

Aufgrund ihres sprachübergreifenden Auftretens werden Internationalismen häufig im Fremdsprachenunterricht – gerade auf den Anfangsstufen – genutzt, um die Lernenden auf der Wortschatzebene zu entlasten (vgl. Baryšnikova & Kažuro, 2019; Kažuro, 2019; Krasnokutskaja, 2019; Tafel, 2009, 61–65). Deshalb gehen wir davon aus, dass Internationalismen von den Lernenden nicht als „schwierige“ Wörter wahrgenommen werden.

### 3. Methodik

#### 3.1 Design des Experiments

Die Textausschnitte, die im Rahmen des Experiments zu bewerten waren (im Folgenden als Stimuli bezeichnet), sind LeStCoR entnommen und erfüllen folgende Kriterien:

- a) Ihre Satzlänge, gemessen in Wortformen, entsprechen annähernd der durchschnittlichen Satzlänge für das jeweilige GERS-Niveau in LeStCoR (siehe Tab. 1).

Niveau	Ø Satzlänge LeStCoR	SD Satzlänge LeStCoR
A1	11.26	7.07
A2	16.13	10.62
B1	20.18	14.46
B2	23.54	17.29
C1	23.68	18.73
C2	28.94	22.93

Tabelle 1: Satzlänge mit Standardabweichung (SD) auf den unterschiedlichen Niveaus

- b) Der Inhalt der Textausschnitte musste ohne die kontextuelle Einbettung im ursprünglichen Gesamttext verständlich sein, d.h. sie sollten u.a. keine ana- oder kataphorischen Verweise enthalten, die sich über den

## Lernerurteile zu lexikalischen Schwierigkeiten

jeweiligen als Stimulus verwendeten Ausschnitt hinaus auf den Gesamttext erstreckten.

Für das Experiment wurden für jedes GERS-Niveau je acht Stimuli ausgewählt, deren Klassifikationsergebnisse sich folgendermaßen in Gruppen unterteilen lässt:

- Vier Stimuli, die von Lehrwerken und automatischem Klassifikator mit demselben Schwierigkeitsgrad versehen worden waren,
- Zwei Stimuli, die vom Klassifikator bzw. in den Lehrwerken eine Schwierigkeitsstufe niedriger eingestuft worden waren, sowie
- Zwei Stimuli, die vom Klassifikator bzw. in den Lehrwerken eine Schwierigkeitsstufe höher eingestuft worden waren.

Aus diesen acht Stimuli wurden für das Experiment zwei Blöcke gebildet, die beide je zwei Stimuli mit gleichlautender Einstufung enthielten; im ersten Block kamen dann noch zwei Stimuli mit niedrigerer Einstufung bzw. im zweiten Block zwei Stimuli mit höherer Einstufung hinzu.

Die nachfolgend beschriebene Implementation des Experiments wurde im online-Tool SoSciSurvey angelegt.

Die erste Aufgabe für die Teilnehmer\*innen bestand darin, die vier Stimuli eines Blocks hinsichtlich der subjektiv empfundenen Schwierigkeiten beim Leseverständnis zu rangieren. Danach sollten die Teilnehmer\*innen in jenen beiden Stimuli-Texten, die sie innerhalb eines Viererblocks als den einfachsten bzw. den schwierigsten eingestuft hatten, jene Einheiten markieren, die zu ihrer Einschätzung geführt hatten, und angeben, ob die Schwierigkeiten auf der lexikalischen, morphologischen oder (morpho)syntaktischen Ebene angesiedelt waren.

Technisch wurde dies so realisiert, dass für jeden Stimulus eine Bilddatei angelegt wurde, in der jedes manuell definierte Feld einer Wortform entsprach (vgl. Abb. 1). So konnten die Teilnehmer\*innen in jeder Bilddatei beliebig viele Wortformen markieren.



Abbildung 1. Interne, für Teilnehmende nicht sichtbare Definition der Felder am Beispiel eines Stimulus auf dem Niveau A2.<sup>2</sup>

Die Einteilung der Felder war für die Teilnehmer\*innen nicht ersichtlich (vgl. Abb. 2). Da die Zielgruppe unseres Experiments neben Lehrpersonen vor allem Fremdsprachenlernende, die Russisch in unterschiedlichen Kontexten erlernen, sowie Herkunftssprecher\*innen umfasste, konnten wir die Vertrautheit mit linguistischer Terminologie nicht voraussetzen. Deshalb verwendeten wir zur Klassifikation in den Instruktionen die Laientermini „непонятное слово / unklare Wortbedeutung“ für Schwierigkeiten auf der lexikalischen Ebene, „сложная грамматика / schwierige Grammatik“ für Schwierigkeiten auf der morphologischen Ebene und „непонятная зависимость в тексте / unklarer Bezug im Satz“ für Schwierigkeiten auf der (morpho)syntaktischen Ebene.

---

<sup>2</sup> Da aus technischen Gründen maximal 32 Felder definiert werden können, auf den höheren GERS-Niveaus die Textausschnitte jedoch teilweise aus mehr als 32 Wortformen bestanden, sahen wir uns in diesen Fällen gezwungen, auch einige Felder zu definieren, die ein Syn- und ein Autosemantikon enthielten. Bei den Synsemantika handelte es sich entweder um Präpositionen oder Adverbien, die bereits auf einem niedrigeren GERS-Niveau eingeführt worden waren, oder um Partikeln wie *не*, *же* oder die Konjunktionen *и*, *а*, *или* oder *но*.

## Lernerurteile zu lexikalischen Schwierigkeiten

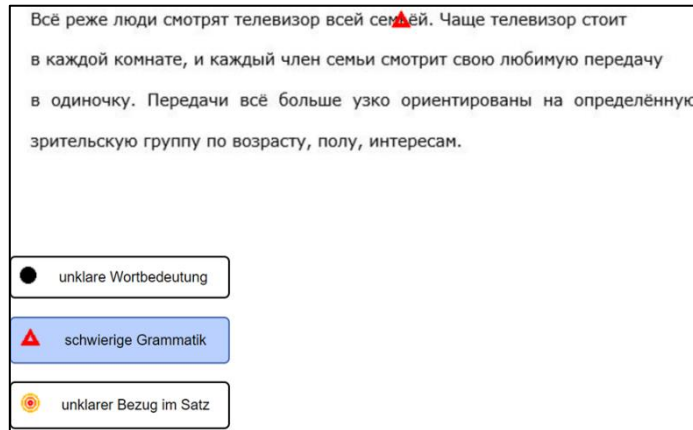


Abbildung 2: Design des Experiments aus Sicht der Teilnehmer\*innen mit Deutsch als Instruktionssprache.

Unsere Zielgruppe umfasste sowohl Lernende des Russischen und Herkunftssprecher\*innen als auch Lehrpersonen für Russisch als Fremdsprache. In beiden Gruppen wurden folgende soziodemografische Daten abgefragt:

- Alter
- Gender
- Momentanes Aufenthaltsland
- Geburtsland
- Migrationsjahr, falls Aufenthalts- und Geburtsland nicht identisch waren
- Status des Russischen (Fremdsprache, Herkunftssprache, Primärsprache)
- Höchster Bildungsabschluss
- Vorliegen linguistischer Kenntnisse
- Kenntnisse weiterer (Fremd)Sprachen
- Domänen, in denen Russisch genutzt wird, und Angabe der Nutzungsfrequenz

Bei Lernenden wurde zusätzlich Folgendes erhoben:

- Selbsteinschätzung des GERS-Niveaus für Leseverständnis
- Wo wurde Russisch gelernt bzw. erworben

Bei Lehrpersonen wurde zusätzlich Folgendes erhoben:

- Institution(en), an denen die Lehrperson unterrichtet
- GERS-Niveaus, welche die Lehrperson selbst unterrichtet

- Möglichkeit zur Auswahl eines GERS-Niveaus für das Experiment mit der Option, später noch ein weiteres GERS-Niveau zu wählen

### 3.2 Durchführung und Rücklauf

Das Experiment wurde in Form einer online-Umfrage auf der Plattform SoSciSurvey durchgeführt. Die Erhebungsperiode lief von Mai bis September 2022. Die Probandenakquise erfolgte über die persönliche Kontaktaufnahme der Autor\*innen mit Lehrpersonen und Lektor\*innen v. a. im deutschsprachigen Raum und Russland, sowie durch Teilnahmeaufrufe über verschiedene Mailing-Listen (LinguistList, SlavList) und Facebook-Gruppen. Die Teilnehmer\*innen konnten zwischen den Instruktionssprachen Russisch, Deutsch und Englisch wählen. Der Rücklauf vollständiger Datensätze verteilte sich wie folgt über die GERS-Niveaus (Tab. 2):

Niveau	Anzahl Fremdsprachenlernende und Herkunftssprecher*innen	Anzahl Bewertungen durch Lehrpersonen
A1	20	9
A2	15	3
B1	21	10
B2	27	5
C1	31	5
C2	83	13

Tabelle 2: Teilnehmer\*innenzahl (Lernende) und der Einschätzungen durch Lehrpersonen pro Niveau

Insgesamt stammten 45 Einschätzungen von 27 Lehrpersonen, da es jeder Lehrperson möglich war, Einschätzungen zu mehreren Niveaus abzugeben. Alle Teilnehmer\*innen waren im Erwachsenenalter. Aufgrund der geringen Gruppengrößen hätte ein Splitting in Untergruppen wie z. B. Lernende vs. Lehrpersonen oder Fremdsprachenlernende vs. Herkunftssprecher\*innen bei der Analyse zu nicht reliablen Aussagen geführt. Bei den Lehrpersonen gehen wir davon aus, dass sie Wortformen markieren, die ihrer Erfahrung nach den Lernenden Probleme bereiten könnten.

Eine erste Analyse der morphologischen und morphosyntaktischen Schwierigkeiten zeigte im Übrigen eine übereinstimmende Einschätzung von Lernenden und Lehrpersonen (Birzer & Mayer, 2022). Zudem haben wir bei unserer Analyse die Ergebnisse für Niveau C2 nicht ausgewertet, da anzunehmen war, dass bei



diesem Grad der Sprachbeherrschung sich die Schwierigkeiten von Lernenden und erstsprachlichen Sprecher\*innen angleichen.

### 3.3 Analyse und Auswertung der erhobenen Daten

Bei unserer Analyse berücksichtigt sind Lexeme, die im Experiment von mindestens zwei Teilnehmer\*innen als „schwierig“ für das jeweilige GERS-Niveau eingeschätzt wurden. Diese Lexeme wurden in eine Tabelle überführt und mit folgenden Kategorien annotiert:

1. GERS-Niveau des jeweiligen Textes, in dem das Lexem im Experiment auftrat;
2. GERS-Niveau des Lexems laut Listen des *Leksičeskij minimum* (2023);
3. falls das Lexem nicht in den Listen des *Leksičeskij minimum* geführt wird, wurde überprüft, ob andere Lexem(e) mit dem- bzw. denselben lexikalischen Morphem(en) (im folgenden als „Mitglieder desselben Wortnests“ bezeichnet) im *Leksičeskij minimum* gelistet sind und wenn ja, auf welchem GERS-Niveau.<sup>3</sup>
4. Wie häufig das jeweilige Lexem als „schwierig“ für dieses GERS-Niveau markiert wurde;
5. Wie häufig insgesamt unterschiedliche Lexeme als „schwierig“ für dieses GERS-Niveau markiert wurden.

Da die Datenbank als Birzer et al. (2023) publiziert ist, werden wir im Folgenden keine Wortlisten anführen, sondern lediglich repräsentative Beispiele diskutieren.<sup>4</sup>

Auf ihrer Grundlage können wir zwei Kategorien von „schwierigen“ Lexemen identifizieren:

---

<sup>3</sup> So ist z. B. das Verb *свидетельствовать* nicht im *Leksičeskij minimum* verzeichnet, jedoch wird das Substantiv *свидетель* auf Niveau B2 des *Leksičeskij minimum* angeführt.

<sup>4</sup> So können zwar keine Aussagen über alle Lexeme getroffen werden, die im *Leksičeskij minimum* (2023) als Teil des Mindestwortschatzes für das jeweilige Niveau definiert sind. Dies wäre zum einen im Rahmen eines einzelnen Experiments auch kaum leistbar und zum anderen muss der Wortschatz verschiedener Lehrbücher nicht zwangsläufig vollständig mit dem *Leksičeskij minimum* übereinstimmen (vgl. Lapošina & Lebedeva, 2019; Lapošina & Lebedeva, 2021 zu einigen ausgewählten Gründen). Sehr wohl können auf Grundlage der Datenauswahl jedoch Aussagen getroffen werden über die angenommenen Faktoren, die wiederum verallgemeinerbar sind.

- a) Lexeme, die auf dem jeweiligen GERS-Niveau noch nicht eingeführt sind und aus deren Wortnest auch keine anderen Mitglieder bereits bekannt sind;
- b) Lexeme, die auf dem jeweiligen GERS-Niveau selbst noch nicht eingeführt sind, andere Mitglieder aus ihrem Wortnest jedoch schon. Internationalismen bilden eine spezielle Untergruppe von b), da davon auszugehen ist, dass die Bedeutung ihrer Wurzelmorpheme aus der jeweiligen Erstsprache der Lernenden bekannt ist. Bei Lexemen der Gruppe b) kann angenommen werden, dass sie als „schwierig“ Lexeme eingeschätzt werden, weil morphologische Alternationen die Rückanbindung an bereits bekannte Lexeme des Wortnestes und/oder unbekannte Derivationsmorphologie die Erschließung der Wortgesamtbedeutung erschweren.

### 3.3.1 Niveau A1

Unsere Stimuli auf Niveau A1 enthielten insgesamt 82 Wortformen, die auf 58 Lexeme zurückführbar sind. 10 der Lexeme sind Eigennamen, weitere vier gehören zu den Internationalismen und 14 zur Gruppe der Synsemantika. Lediglich ein Lexem, nämlich *учительница*, enthält Derivationsmorphologie.

Als „schwierig“ empfunden wurden lediglich die beiden Lexeme *самый* und *известный*, die gemeinsam als Determiner im Stimulus *Эрмитаж – это самый известный музей в Петербурге*. auftraten. *Самый* und damit die lexikalische Funktion des Superlativs sind laut *Leksičeskij minimum* für Niveau A2 vorgesehen; *известный* wird bereits auf Niveau A1 eingeführt. Die Nennung beider Elemente ist wohl auf ihre verschränkte Funktion rückführbar.

### 3.3.2 Niveau A2

Unsere Stimuli auf Niveau A2 enthielten insgesamt 301 Wortformen, die auf 192 Lexeme zurückführbar sind, darunter 19 Eigennamen oder von Eigennamen abgeleitete Adjektiva, 27 Internationalismen und 30 Synsemantika. 20 Lexeme waren durch Derivation gebildet.

Fünf Wortformen wurden als „schwierig“ wahrgenommen; drei von ihnen waren jedoch auf Niveau A2 bereits eingeführt. Bei den anderen beiden handelt es sich um *в одиночку* und *достижение*. Das motivierende Lexem *один* wurde laut *Leksičeskij minimum* bereits auf Niveau A1 eingeführt, allerdings ist das Wortbildungsverfahren von *один* („eins“) hin zur adverbialen Präpositionalphrase *в одиночку* („einsam, allein“, wörtl. ‚in Einsamkeit‘) intransparent, da sich die Gesamtbedeutung nicht additiv aus *один* ‚eins‘ und dem üblicherweise

## Lernerurteile zu lexikalischen Schwierigkeiten

diminuierenden Suffix *-очка* ergibt. *Достижение* (‚Errungenschaft‘) ist im *Leksičeskij minimum* erst für Niveau B2 angeführt, auch sind laut *Leksičeskij minimum* auf A2 noch keine anderen Lexeme desselben Wortnests bekannt,<sup>5</sup> sodass trotz des produktiven Wortbildungsverfahrens die Wortbedeutung nicht inferiert werden kann.

Von den 20 Lexemen, welche durch Derivation gebildet sind, handelt es sich bei 12 um Adjektive. Darunter finden sich Beispiele, die von Eigennamen abgeleitet sind, z. B. *индийский* (‚indisch‘), *итальянский* (‚italienisch‘), und solche, die von internationaler Lexik abgeleitet sind, z. B. *популярный* (‚populär, berühmt‘), *нестабильный* (‚instabil‘), *ориентированный* (‚orientiert‘). Da diese von den Lernenden nicht als schwierig markiert wurden, können wir davon ausgehen, dass die Derivation bereits für Lernende auf dem A2 Niveau transparent ist. Das gilt auch für das Kompositum *десятилетний* (‚zehnjährig‘), in dem zwei Wurzelmorpheme zu einem Lexem verbunden werden und eine kompositionale semantische Einheit bilden. Ebenfalls nicht als schwierig markiert wurde das von *улица* (‚Straße‘) abgeleitete Adjektiv *уличный* (‚straßen-‘), bei dem eine morphologische Alternation auftritt.

### 3.3.3 Niveau B1

Unsere Stimuli auf Niveau B1 enthielten insgesamt 351 Wortformen, die auf 224 Lexeme zurückführbar sind, darunter 14 Eigennamen, 37 Internationalismen und 33 Synsemantika. 49 Lexeme sind durch Derivation gebildet.

Insgesamt 22 Wortformen wurden als „schwierig“ wahrgenommen; 14 davon gehören zu Lexemen, denen Derivation als Wortbildungsverfahren zugrunde liegt.

In Bezug auf die Wortarten der Derivate unterscheiden sich die Niveaus: Waren es auf dem A2 Niveau noch hauptsächlich Adjektive, so treten auf Niveau B1 vermehrt auch derivierte Substantive (23) und präfigierte Verben (13) auf.

Bei den von den Teilnehmer\*innen als „schwierig“ empfundenen Substantiven handelt es sich fast ausschließlich um Abstrakta, die hauptsächlich von Adjektiven mit dem produktiven Suffix *-ость*, z. B. *внешность* (‚Aussehen‘), *закономерность* (‚Gesetzmäßigkeit‘), oder von Verben mittels *-тие*, z. B. *развитие* (‚Entwicklung‘), oder dem produktiven *-ание/-ение*, z. B.

---

<sup>5</sup> Mangels bereits bekannter Lexeme aus demselben Wortnest kann die Transparenz bzw. Opaikheit des Wortbildungsverfahrens bei *достичь* > *достижение* nicht als Erklärungsfaktor herangezogen werden kann. Dennoch sei aus derivationsmorphologischer Sicht angemerkt, dass das verwendete Wortbildungsverfahren transparent und auch produktiv ist, jedoch eine morphologische Alternation enthält.

*определение* (,Definition'), abgeleitet sind. Diese Verfahren können als transparent betrachtet werden. Allerdings fällt auf, dass bei der überwiegenden Mehrzahl der derivierten Substantive noch kein anderes Lexem (idealerweise das motivierende Adjektiv) desselben Wortnests eingeführt ist, sodass sich die Lernenden deshalb die Wortbedeutung weder über die Semantik des verwandten Lexems noch das produktive Wortbildungsverfahren erschließen können.

Die meisten „schwierigen“ Verben weisen eine oder sogar mehrfache Präfigierung auf, so z. B. *вы-делить* (,hervorheben'), *рас-про-странить* (,ausweiten') oder *вос-со-здать* (,wiederherstellen'). Auffällig ist, dass die Derivation per Suffigierung als morphologisches Verfahren transparent ist, aber dennoch Verblexeme, deren Basen auf Niveau B1 bereits eingeführt sind, z. B. *делить* (,teilen') > *выделить* (,hervorheben') und *создать* (,erschaffen') > *воссоздать* (,wiederherstellen'), gleichermaßen als „schwierig“ empfunden werden wie jene Lexeme, aus deren Wortnest noch keine weiteren Lexeme bekannt sind, wie z. B. *распространить* (,ausweiten'). Eine mögliche Erklärung ist, dass der semantische Anreicherungsprozess, der mit der Suffigierung einhergeht (vgl. u.a. Nessel, 2020 sowie Janda & Korba, 2008 zu präfigierten Verben in Lehrbüchern), opaker als das morphologische Verfahren ist und die Lernenden deshalb Schwierigkeiten haben, sich die Gesamtbedeutung des Verbs zu erschließen.

### 3.3.4 Niveau B2

Unsere Stimuli auf Niveau B2 enthielten insgesamt 341 Wortformen, die auf 253 Lexeme zurückführbar sind, darunter 5 Eigennamen, 28 Internationalismen und 47 Synsemantika. 83 Lexeme sind durch Derivation gebildet. Davon handelt es sich bei 39 um derivierte Substantive, bei 26 um Adjektive, bei 16 um Verben und bei zwei um Adverbien.

Bei den Verben manifestiert sich die „Schwierigkeit“ auf der Ebene der Flexionsform: Es ist augenfällig, dass viele der „schwierigen“ Verbformen Partizipien sind, die wie z. B. *заражённый* (,angesteckt') eine morphologische Alternation enthalten oder wie *изводивший* (,aufgebraucht') eine niedrige Verwendungsfrequenz (die hier sowohl das Ausgangslexem als auch die Form des Partizips Präteritum Aktiv selbst betrifft) haben. Dies betrifft auch bereits eingeführte Lexeme. Es ist anzunehmen, dass die Lernenden Probleme haben, die Flexionsform dem entsprechenden Lexem zuzuordnen und / oder die semantischen Prozesse zu identifizieren und zu interpretieren, die bei der Bildung von Transpositionsformen ablaufen (vgl. Lehmann, 2003), sodass wir es streng genommen nicht mit einem lexikalischen Problem, sondern mit einem Problem an der Schnittstelle von Semantik und Morphosyntax zu tun haben.

Die „schwierigen“ Substantive sind mehrheitlich deverbal und bilden zwei Gruppen: Die erste Gruppe umfasst Substantive, die mittels eines produktiven Verfahrens von einem infrequenten (und wohl deshalb noch nicht eingeführten) Wurzelmorphem gebildet werden wie z. B. *шелестение* („Rascheln“) oder *курлыканье* („Gurren“). Nach demselben Verfahren gebildete Substantive, aus deren Wortnest bereits andere Lexeme bekannt waren, wurden nicht als „schwierig“ markiert, was als Hinweis auf die grundsätzliche Transparenz der Wortbildungsverfahren zu bewerten ist. Zur zweiten Gruppe gehören Substantive auf der Basis von frequenten, bereits eingeführten Wurzelmorphemen, die jedoch mittels eines vergleichsweise wenig produktiven, z. B. *жил-ище* („Wohnung“), oder eines semantisch unspezifischen Suffixes, z. B. *догад-ка* („Annahme“), deriviert werden, sodass wir wiederum mit der Opakheit des Wortbildungsverfahrens konfrontiert sind.

Weiters fällt im nominalen Bereich auf, dass einige als „schwierig“ wahrgenommene Wortformen morphonologische Alternationen beinhalten, so z. B. der Komparativ *дороже* („teurer“) oder das deverbale Substantiv *пашня* („Ackerland“), abgeleitet vom Verb *пахать* („pflügen“) (das jedoch auf Niveau B2 ebenfalls noch nicht eingeführt ist).

### 3.3.5 Niveau C1

Unsere Stimuli auf Niveau C1 enthielten insgesamt 146 Wortformen, die auf 106 Lexeme zurückführbar sind, darunter zwei Eigennamen, 20 Internationalismen und 19 Synsemantika. 50 Lexeme sind durch Derivation gebildet, 20 davon sind Substantive, 19 Adjektive, neun Verben und zwei Adverbien. Zehn Derivate sind auf der Basis von Internationalismen entstanden.

Die Beobachtungen zu den Hypothesen 1 bis 3 (H1–H3) decken sich mit jenen zu Niveau B2: Bei den Verben werden wiederum (adverbial)partizipiale Flexionsformen wie *заимствованный* („entlehnt“), *имевший* („gehabt habend“) und *творя* („bauend“) als schwierig wahrgenommen. Im nominalen Bereich werden abermals Derivationen als schwierig wahrgenommen, die zwar auf produktiven, transparenten Verfahren beruhen, jedoch morphonologische Alternationen nach sich ziehen, wie z. B. *вхождение* („Eintritt“),<sup>6</sup> oder aber niedrigfrequente Wurzelmorpheme aufweisen wie z. B. *ёмкость* („Kapazität“). Bei letzterem kommt noch hinzu, dass die Bildung auf *-ость* ein Konkretum bezeichnet, obwohl dieses Suffix üblicherweise bei der Derivation von Abstrakta zum Einsatz

---

<sup>6</sup> Die als „schwierig“ markierte Kurzform *чужд* des Adjektivs *чуждый*, eines kirchenslavischen Kognats von *чужой*, kann ebenfalls als Ergebnis einer morphonologischen Alternation im weiteren Sinne betrachtet werden.

kommt. *Ёмкость* ist also zu einem gewissen Grad semantisch opak. Selbiges trifft auch auf jene Substantive zu, die mittels semantisch unspezifischer Suffixe, z. B. *выработ-ка* („Produktion“), deriviert werden.

Erstaunlicherweise wird auf Niveau C1 erstmals eine beträchtliche Anzahl von Internationalismen, nämlich sieben von insgesamt 20, als „schwierig“ wahrgenommen. Konkret handelt es sich um die Lexeme *терминология* („Terminologie“), *изолированность* („Isoliertheit“), *регион / региональный* („Region / regional“), *посткоммунистический* („postkommunistisch“), *процесс* („Prozess“) und *параллельный* („parallel“) (in der Reihenfolge ihrer Nennung im Experiment, welche das Auftreten in den Stimuli und damit auch das Ranking der Stimuli nach Schwierigkeitsgrad widerspiegelt). Auffällig ist, dass sich sechs der sieben Internationalismen auf zwei Stimuli verteilen, weshalb es sinnvoll erscheint, die Internationalismen in ihrer kontextuellen Einbettung zu betrachten:<sup>7</sup>

- (1) В каждом национальном языке **терминология** формируется главным образом средствами своего национального языка. Но это не предполагает **изолированности** языка (как и изолированности общества), поэтому вхождение в язык, в специальные **терминологии** заимствованных слов процесс обычный для каждого национального языка, в том числе и русского.

Die Schwierigkeiten mit *терминология* („Terminologie“) erklären sich höchstwahrscheinlich dadurch, dass die beiden Kontexte (1) jeweils unterschiedliche Lesarten hervorrufen, nämlich ‚Fachwortschatz einer Sprache‘ (im Gegensatz zum Alltagswortschatz) im ersten Kontext und – hervorgerufen durch den Plural (*coercion* nach Fried & Östman, 2005, 27–28) im zweiten Kontext – ‚Terminologie einer Fachwissenschaft‘. Die Probleme bei der semantischen Interpretation sind also wiederum an einzelne Flexionsformen gekoppelt und treten an der Schnittstelle von Semantik und Morphosyntax auf. Auch bei *изолированность* (1) besteht grundsätzlich ein Polysemieproblem, da *изолированность* interpretiert werden kann als ‚Zustand der Isolation‘ oder als ‚Isoliertheit‘ (im sprachtypologischen Sinne) und damit als linguistischer Terminus. Da vermutlich nur wenige bis keine Teilnehmer\*innen den linguistischen Terminus kennen, ist es wahrscheinlicher, dass ein Problem an der Schnittstelle von semantischen Partizipanten und ihrer syntaktischen Realisierung vorliegt: **wovon** Sprache und Gesellschaft

---

<sup>7</sup> Da einige Lexeme in den Stimuli mehrfach auftreten, sind sie in den Kontexten, in denen sie als „schwierig“ wahrgenommen wurden, fett gedruckt.

## Lernerurteile zu lexikalischen Schwierigkeiten

isoliert sein könnten, nämlich von anderen Sprachen und Kulturen, wird syntaktisch nicht realisiert und ist nur aus dem erweiterten Kontext rekonstruierbar.<sup>8</sup>

- (2) Выработка политического курса в отношении **регионов** в **посткоммунистической** России коренным образом отличается от соответствующих **процессов**, имевших место в иные периоды российской истории. Принципиально новой является роль **региональных** факторов в процессе становления их отношений с центральной властью.

Bei *посткоммунистический* („postkommunistisch“) handelt es sich um einen geschichts- bzw. politikwissenschaftlichen Terminus; als Internationalismus tritt er zwar in vielen Sprachen auf, gehört jedoch einem gehobenen Register an, sodass seine Wahrnehmung als „schwierig“ nachvollziehbar ist. Bei *региональный* („regional“) scheint die Schwierigkeit abermals in der Flexionsform, genauer gesagt in der Pluralform, begründet zu liegen: wie eine Korpusrecherche im Zeitungskorpus des russischen Nationalkorpus zeigte,<sup>9</sup> tritt die Kollokation *региональный фактор* beinahe ausschließlich im Singular auf, die Pluralform impliziert jedoch nicht die Lesart ‚unterschiedliche regionale Faktoren‘, sondern ‚der(selbe) regionale Faktor in seiner jeweils spezifischen Ausprägung für unterschiedliche Regionen‘, sodass hier auch kein Fall von *coercion* sondern eine für diese Kollokation spezifische idiosynkratische Lesart vorliegt. Bei *регион* („Region“) und *процесс* („Prozess“) scheint es wahrscheinlich, dass die Lernenden keine Schwierigkeiten mit den Lexemen selbst hatten, sondern mit den

---

<sup>8</sup> Es handelt sich also um eine morphosyntaktische Leerstelle und damit eigentlich um eine morphosyntaktische Schwierigkeit. Im Experiment gab es keine Möglichkeit, Leerstellen zu markieren. Es ist daher konsequent, dass die Teilnehmer\*innen den semantischen und syntaktischen Kopf *изолированность* als „schwieriges“ Lexem markiert haben, da seine semantische Kontextualisierung durch die fehlende Konstituente erschwert wird.

<sup>9</sup> Suchanfrage *региональный* und mit Abstand 1 *фактор*; bei *фактор* wurde außerdem die Wiederholung von Kasus, Numerus und Genus der jeweils vorhergehenden Flexionsform von *региональный* festgesetzt, um das Vorliegen der Kollokation sicherzustellen. Die Suchanfrage ergab insgesamt 15 Treffer, von denen nur einer die Pluralform der Kollokation enthielt:

(1) Поэтому, с учетом **региональных факторов** Вьетнам, Куба и Монголии могли бы стать своего рода опорными пунктами совместно го предпринимательства, обеспечивающими активное продвижение продукции стран СЭВ на рынки стран соответствующих регионов. (На СРВ, МНР и Кубу приходится 16% населения и 2,3% национального дохода стран СЭВ. Помощь - дело взаимовыгодное // Аргументы и факты, 1987.10)

Konstituentenstrukturen, und fälschlicherweise das Label „unklare Wortbedeutung“ anstatt „unklarer Bezug im Satz“ wählten.

Für unsere Hypothese zur Rolle von Internationalismen lässt sich festhalten, dass – anders als auf den vorhergehenden GERS-Niveaus – auf Niveau C1 das Auftreten von registerspezifischen Fachtermini und von kontextabhängigen spezifischen Lesarten zu beobachten ist. Der Übergang von Niveau B2 zu C1 scheint also einen Wendepunkt darzustellen weg von einer entlastenden Funktion der Internationalismen und hin zu registerspezifischen Verwendungsarten mit teilweise idiosynkratischen Lesarten.

#### 4. Fazit

Die vorliegende Studie untersuchte Faktoren, die zur wahrgenommenen Schwierigkeit des russischen Wortschatzes bei Lernenden auf verschiedenen GERS-Stufen beitragen. Die Studie testete vier Hypothesen zu Wortbildungsprozessen, Bekanntheit von Lexemen, morphonologischen Alternationen und Internationalismen.

Die erste Hypothese (H1) besagte, dass mit höherem GERS-Niveau mehr Lexeme mit opaken Wortbildungsprozessen als „schwierig“ empfunden werden. Lernende auf höheren Niveaus kennen bereits viele produktive Wortbildungsprozesse und können auf diese zurückgreifen, um sich die Bedeutung von Lexemen zu erschließen. Die Ergebnisse zeigten, dass diese Hypothese im Allgemeinen zutraf. Ab Niveau B1 waren jedoch Unterschiede zwischen Substantiven und Verben erkennbar. Bei den Substantiven auf Niveau B1 wurden Abstrakta generell als opaker und „schwieriger“ wahrgenommen als Konkrete. Die meisten „schwierigen“ Verben auf Niveau B1 waren durch eine oder sogar mehrfache Präfigierung gekennzeichnet. Auf der Stufe B2 korrelierte bei den Verben die Wahrnehmung als „schwierig“ mit bestimmten Flexionsformen, insbesondere (Adverbial)Partizipien. Die meisten „schwierigen Substantive“ waren deverbale und bildeten zwei Gruppen: Die erste Gruppe umfasste Substantive, die durch produktive Prozesse mit einem infrequenten Wurzelmorphem gebildet wurden. Die zweite Gruppe bestand aus Substantiven, die auf frequenten, bereits eingeführten Wurzelmorphemen basierten, aber durch weniger produktive Prozesse oder mittels semantisch unspezifischer Suffixe gebildet wurden. Auf der Stufe C1 wurden ähnliche Tendenzen wie auf den Niveaus A1–B2 beobachtet, jedoch gab es zusätzlich eine Verschiebung hin zu registerspezifischen Fachtermini und kontextabhängigen spezifischen Lesarten, die als „schwierig“ wahrgenommen wurden.



## Lernerurteile zu lexikalischen Schwierigkeiten

Die zweite Hypothese (H2) nahm an, dass die Wahrscheinlichkeit, bereits Lexeme aus dem gleichen Wortnest wie das zu erschließende Wort zu kennen, mit steigendem GERS-Niveau zunimmt und deshalb semantische Inferenzen für die Erschließung des unbekanntes Worts genutzt werden können. Diese Hypothese konnte für die Niveaus A2–B2 bestätigt werden. Auf Niveau C1 erwies sich die Kenntnis von Lexemen aus demselben Wortschatz als weniger hilfreich, was mit der oben erwähnten Verlagerung hin zu registerspezifischen Fachtermini und kontextabhängigen spezifischen Lesarten zusammenhängen dürfte.

Die dritte Hypothese (H3) besagte, dass Lexeme mit morphologischer Alternation umso häufiger als „schwierig“ wahrgenommen werden, je niedriger das GERS-Niveau ist. Weil morphologische Alternationen regelhaft auftreten, u.a. bei derivationaler Wortbildung, sollten sie mit steigender Sprachbeherrschung zuverlässig erkannt werden und die Erschließung der Wortbedeutung nicht mehr behindern. Die Ergebnisse der Studie konnten diese Hypothese jedoch nicht stützen, da auch auf höheren GERS-Stufen Wortformen mit morphologischer Alternation als „schwierig“ empfunden wurden.

Die vierte Hypothese (H4) nahm an, dass Internationalismen von Lernenden nicht als „schwierig“ wahrgenommen werden, sondern die Lernenden entlasten, da sie sprachübergreifend auftreten. Die Ergebnisse der Studie zeigen, dass diese Hypothese für die Niveaustufen A1–B2 zutrifft. Mit steigendem GERS-Niveau nahm die Anzahl der Internationalismen in den Stimuli ab. Auf Niveau C1 traten Internationalismen als registerspezifische Fachausdrücke auf, die oft kontextabhängige spezifische Lesarten nach sich zogen. Niveau C1 stellt somit einen Wendepunkt dar weg von der entlastenden Funktion von Internationalismen hin zu ihrer registerspezifischen Verwendung als Termini mit teilweise idiosynkratischen Lesarten.

Die Merkmale aus den verifizierten Hypothesen zeigen abgestufte Ausprägungen über die GERS-Niveaus hinweg, sodass es lohnenswert ist, sie als *feature* für die automatische Textklassifikation zu operationalisieren und dann die Eignung der *features* zu testen. Zudem zeigen die Ergebnisse des Experiments deutlich, dass im Bereich des Wortschatzes bei der Gestaltung von Sprachlehrplänen und -materialien mehrere Faktoren zu berücksichtigen sind. Die Studie unterstreicht noch einmal, wie sinnvoll das seit langem etablierte Verfahren ist, bereits bekannten Wortschatz und bekannte Wortbildungsverfahren für die Erschließung neuer Lexeme zu nutzen. Die Ergebnisse legen zudem nahe, dass produktive Wortbildungsverfahren auf niedrigeren GERS-Niveaus explizit thematisiert und somit Sprachbewusstheit geschaffen werden sollte und andererseits weniger produktive Verfahren und semantisch unspezifische Suffixe auf höheren Niveaus thematisiert werden sollten. Ebenso könnten Lernende von der aktiven

Auseinandersetzung mit morphologischen Alternationen in der Form- und Wortbildung profitieren.

## Literatur

- Baryšnikova, E.N. & Kažuro D.V. (2019). = Барышникова Е.Н. & Кажуро Д.В. (2019). Дидактическая многофункциональность интернационализмов в практике преподавания русского языка. *Известия Саратовского университета. Новая серия. Серия: Филология. Журналистика*, 19(2), 164–169. <https://doi.org/10.18500/1817-7115-2019-19-2-164-169>
- Batinić, D., Birzer, S. & Zinsmeister, H. (2017). Automatic classification of Russian texts for didactic purposes. *Proceedings of the international conference "Corpus Linguistics – 2017"* (9–15). St. Petersburg.
- Birzer, S. & Mayer, H. (2022). *Satzbewertung durch Lernende als Validierungsinstrument für die morphosyntaktische und lexikalische Komplexität von russischen Lesetexten*. Vortrag auf dem 14. Deutschen Slavistiktag in Bochum am 24.09.2022.
- Birzer, S., Mayer, H., & Gomer, A. (2023). Datensatz zu lexikalischen Schwierigkeiten beim Textverständnis russischer Lesetexte (Version\_14\_05\_2023) [Data set]. Otto-Friedrich-Universität Bamberg. <https://doi.org/10.48564/unibaf-d-0zz9s-yth17> (14.05.2023)
- Diuf, A. (2011) = Диуф, А. (2011). Морфологические чередования при суффиксальном словообразовании в обучении РКИ. *Вестник Российского университета дружбы народов. Серия "Русский и иностранные языки и методика их преподавания"*, 2011(1), 92–97.
- Janda, L. & Korba, J. (2008). *Cluster Types for Russian Verbs*. <http://ansatte.uit.no/laura.janda/clusters/clusterfrontpage.html> (07.03.2023)
- Kažuro, D.V. (2019) = Кажуро, Д.В. (2019). Границы понятия «Интернациональная лексика» в лингводидактике РКИ. In В.В. Богуславская (ред.), *Язык. Коммуникация. Культура*. 2019. *Альманах по итогам III Международной научно-практической конференции молодых ученых – магистрантов и аспирантов – «Язык. Коммуникация. Культура» (20 апреля 2019 г.)* (78–80).
- Krasil'nikova, L.V. (2015) = Красильникова, Л.В. (2015). Лингводидактические аспекты описания русского языка иностранным филологам на современном этапе преподавания РКИ. *Вестник РУДН, серия Русский и иностранные языки и методика их преподавания*, 2015(4), 69–82.
- Krasnokutskaja, N.V. (2019) = Краснокутская, Н.В. (2019). *Обучение заимствованной лексике в русском языке как лингводидактическая проблема (на материале спецкурса для иностранных учащихся)*. Автореферат диссертации на соискание ученой степени кандидата педагогических наук. <https://www.dissercat.com/content/obuchenie-zaimstvovannoi-leksike-v-russkom-zazyke-kak-lingvodidakticheskaya-problema-na-mate> (10.05.2023)

## Lernerurteile zu lexikalischen Schwierigkeiten

- Laposhina, A. & Lebedeva, M. (2021). *Compiling CEFR graded vocabulary lists for Russian L2 learners based on 4 sources of data*. Vortrag auf dem GR4L2 Workshop an der UC Louvain am 07.12.2021.
- Ларошина, А. & Лебедева, М. (2019) = Лапошина, А. & Лебедева, М. (2019). Корпусный подход к решению проблемы отбора лексики в обучении РКИ. In A. Nikunlassi & E. Protassova (Hrsg.), *Russian language in the multilingual world* (359–368).
- Lehmann, V. (2003). Grammaticische Derivation (Aspekt, Genus verbi, Komparation, Partizip und andere Phänomene zwischen Flexion und Wortbildung). In T. Berger & K. Gutschmidt (Hrsg.), *Funktionale Beschreibung slavischer Sprachen* (139–162). München.
- Leksičeskij minimum RKI (2023) = Лексический минимум РКИ. (2023). <https://www.ros-edu.ru/basic-dictionary> (07.03.2023)
- Lemmenmeier-Batinić, D. (2019). *Klassifikation der Texte des LeStCoR-Korpus' in sechs Stufen mittels SVM Classifier*. Vortrag am Institut für Slawistik der HU Berlin am 16.12.2019.
- Neset, T. (2020). What's in a Russian Aspectual Prefix? A Cognitive Linguistics Approach to Prefix Meanings. *Journal of Slavic Linguistics*, 28(2), 141–162.
- Reynolds, R. (2021). *Comparing the validity of human-annotated CEFR labels from different sources using machine-learning classifiers: Examples from Russian*. Vortrag auf dem GR4L2 Workshop an der UC Louvain am 07.12.2021.
- Tafel, K. (2009). *Slavische Interkomprehension. Eine Einführung*. Gunter Narr.
- Тажев, В.Р. (1997) = Тажев, Б.П. (1997). *Структурное изучение русского языка иностранными студентами-нефилологами (1-2 курсы) на материале имени существительного*. Автореферат диссертации на соискание ученой степени кандидата педагогических наук. [https://new-disser.ru/\\_avtoreferats/01000157689.pdf](https://new-disser.ru/_avtoreferats/01000157689.pdf) (10.05.2023)