

ISSN 2686-7249

ВЕСТНИК РГГУ

Серия

«Литературоведение.
Языкознание. Культурология»

Научный журнал

RSUH/RGGU BULLETIN

“Literary Theory.
Linguistics. Cultural Studies”
Series

Academic Journal



Московский лингвистический журнал

Том 22

Moscow Journal of Linguistics

Volume 22

Основан в 1996 г.
Founded in 1996

7
2020

VESTNIK RGGU. Seriya "Literaturovedenie. Yazykoznanie. Kul'turologiya"
RSUH/RGGU BULLETIN. "Literary Theory. Linguistics. Cultural Studies" Series
Academic Journal

There are 10 issues of the journal a year.

Founder and Publisher: Russian State University for the Humanities (RSUH)

RSUH/RGGU BULLETIN. "Literary Theory. Linguistics. Cultural Studies" Series is included: in the Russian Science Citation Index; in the List of leading scientific journals and other editions for publishing PhD research findings.

Peer-reviewed publications fall within the following research area:

10.01.00 Philology:

10.01.01 Russian literature

10.01.03 Foreign literature

10.01.08 Literary Theory. Textology

10.01.09 Folkloristics

10.02.00 Linguistics:

10.02.14 Classical philology, Byzantine and Modern Greek Studies

10.02.01 Russian language

10.02.02 Languages of the Russian Federation

10.02.19 Theoretical linguistics

10.02.20 Historical-comparative, typological and contrastive linguistics

24.00.00 Culturology:

24.00.01 Cultural history and theory

24.00.03 Museology, conservation and restoration of historical and cultural objects

GOALS OF THE JOURNAL: Presentation of the results of the latest research in the field of philology, linguistics and cultural studies which have an unquestionable theoretical and practical value and are promising for the development of research in these fields of knowledge. Advancement of empirically oriented linguistic research and high-quality studies of Russian, languages of the Russian Federation, and languages of the world within a variety of theoretical frameworks and in comparative, historical and typological perspectives.

OBJECTIVES OF THE JOURNAL: implementation and development of expertise of scientific articles taking into account the dominance of modern interdisciplinary and integrated approaches; presentation of the most significant achievements important for the development of science and capable of being introduced into the educational process as examples of correct scientific work; attraction of new authors, researchers, showing high theoretical culture and undeniable scientific achievements; strengthening the interaction of academic and university science; translation of scientific experience between generations and between institutions. The journal invites original research articles by Russian and international scholars.

RSUH/RGGU BULLETIN. "Literary Theory. Linguistics. Culturology" Series is registered by Federal Service for Supervision of Communications, Information Technology and Mass Media. Certificate on registration: PI No. FS77-61883 of 25.05.2015

Changes were made to the record of media registration in connection with the name change, renaming of the founder, clarification of the subject – registration number FS77-74270 of 09.11.2018

Editorial staff office: 6, Miusskaya Sq., Moscow, 125993

e-mail: msk.ling.j@gmail.com

ВЕСТНИК РГГУ. Серия «Литературоведение. Языкознание. Культурология»
Научный журнал

Выходит 10 номеров печатной версии журнала в год.

Учредитель и издатель – Российский государственный гуманитарный университет
(РГГУ)

ВЕСТНИК РГГУ. Серия «Литературоведение. Языкознание. Культурология» включен в систему Российского индекса научного цитирования (РИНЦ); в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук по следующим научным специальностям и соответствующим им отраслям науки:

10.01.00 Литературоведение:

10.01.01 Русская литература

10.01.03 Литература народов стран зарубежья (с указанием конкретной литературы)

10.01.08 Теория литературы. Текстология

10.01.09 Фольклористика

10.02.00 Языкознание:

10.02.14 Классическая филология, византийская и новогреческая филология

10.02.01 Русский язык

10.02.02 Языки народов Российской Федерации (с указанием конкретного языка или языковой семьи)

10.02.19 Теория языка

10.02.20 Сравнительно-историческое типологическое и сопоставительное языкознание

24.00.00 Культурология:

24.00.01 Теория и история культуры

24.00.03 Музееведение, консервация и реставрация историко-культурных объектов

ЦЕЛЬ ЖУРНАЛА: представление результатов новейших исследований в области литературоведения, языкознания и культурологии, имеющих несомненное теоретическое и практическое значение и перспективных для развития исследований в этих областях знания.

Продвижение эмпирически-ориентированных исследований по русскому языку, языкам Российской Федерации и языкам мира в рамках разнообразных теоретических подходов и в сопоставительной, исторической и типологической перспективе.

ЗАДАЧИ ЖУРНАЛА: осуществление и развитие экспертизы научных статей с учетом господства современных междисциплинарных и комплексных подходов; представление наиболее значимых достижений, важных для развития науки и способных быть внедренными в образовательный процесс как примеры правильной научной работы; привлечение новых авторов, исследователей, показывающих высокую теоретическую культуру и неоспоримые научные достижения; усиление взаимодействия академической и университетской науки; трансляция научного опыта между поколениями и между институтами.

Журнал принимает к публикации оригинальные статьи, комплексные исследования российских и зарубежных авторов, ранее не публиковавшиеся научные доклады.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций, свидетельство о регистрации ПИ № ФС77-61883 от 25.05.2015 г. В запись о регистрации СМИ внесены изменения в связи с изменением названия, переименованием учредителя, уточнением тематики – регистрационный номер ПИ № ФС77-74270 от 09.11.2018 г.

Адрес редакции: 125993, Москва, Миусская пл., 6
электронный адрес: msk.ling.j@gmail.com

Founder and Publisher
Russian State University for the Humanities (RSUH)

Editor-in-chief

P.P. Shkarenkov, Dr. of Sci. (History), Professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation

Editorial Board

D.I. Antonov, Dr. of Sci. (History), associate professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation

P.M. Arkadiey, Dr. of Sci. (Philology), Institute of Slavic Studies RAS/Russian State University for the Humanities (RSUH), Moscow, Russian Federation (*deputy editor*)

O.L. Akhunova, Dr. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation

S.I. Baranova, Dr. of Sci. (History), The Moscow State Integrated Art and Historical Architectural and Natural Landscape Museum-Reserve, Moscow, Russian Federation

L.V. Belovinskiy, Dr. of Sci. (History), professor, Moscow State Art and Cultural University, Moscow, Russian Federation

V.V. Gudkova, Dr. of Sci. (art studies), State Institute for Art Studies, Moscow, Russian Federation

N.P. Grintser, Dr. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation

Yu.V. Domanskiy, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation

A.V. Dybo, RAS corr. memb., Dr of Sci. (Philology), professor, Institute of Linguistics RAS, Moscow, Russian Federation

I. Rzepnikowska, Dr. of Sci. (Philology), Nicolaus Copernicus University, Toruń, Poland

G.I. Zvereva, Dr. of Sci. (History), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation (*deputy editor*)

I.I. Isaev, Cand. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation

G.I. Kabakova, Dr. of Sci. (Philology), Université de Paris-Sorbonne, Paris, France

N.V. Kapustin, Dr. of Sci. (Philology), professor, Ivanovo State University, Ivanovo, Russian Federation

V.I. Kimmelman, PhD, Bergen University, Bergen, Norway

J.D. Clayton, PhD, University of Ottawa, Ottawa, Canada

I.V. Kondakov, Dr. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation

G.Ye. Kreidlin, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation

L.I. Kulikov, Cand. of Sci. (Philology), PhD, Ghent University, Ghent, Belgium

M.N. Lipovetskiy, Dr. of Sci. (Philology), professor, University of Colorado, Boulder, USA

D.M. Magomedova, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation

- I.V. Morozova*, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- V.G. Mostovaya*, Cand. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- S.Yu. Neklyudov*, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- V.I. Podlesskaya*, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- O.I. Polovinkina*, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- E.Yu. Protasova*, Dr. of Sci. (pedagogy), University of Helsinki, Helsinki, Finland
- R.I. Rozina*, Dr. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- J. Sadowski*, Dr. of Sci. (History), Jagellonian University, Krakow, Poland
- A.Yu. Sorochan*, Dr. of Sci. (Philology), associate professor, Tver State University, Tver, Russian Federation
- Ya.G. Testeleets*, Dr. of Sci. (Philology), Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- O.I. Togoyeva*, Dr. of Sci. (History), Institute of General History RAS, Moscow, Russian Federation
- V.I. Tyupa*, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- A.A. Kholikov*, Dr. of Sci. (Philology), Moscow State Lomonosov University, Moscow, Russian Federation
- O.B. Khristoforova*, Dr. of Sci. (Philology), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- I.O. Shaytanov*, Dr. of Sci. (History), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation
- S.A. Yatsenko*, Dr. of Sci. (History), professor, Russian State University for the Humanities (RSUH), Moscow, Russian Federation

Executive editor of the issue

P.M. Arkadiev, Dr. of Sci. (Philology) (RSUH)

Учредитель и издатель:
Российский государственный гуманитарный университет (РГГУ)

Главный редактор

П.П. Шкаренков, доктор исторических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Редакционная коллегия:

Д.И. Антонов, доктор исторических наук, доцент, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

П.М. Аркадьев, доктор филологических наук, Институт славяноведения РАН, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация (*заместитель главного редактора*)

О.Л. Ахунова, доктор филологических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

С.И. Баранова, доктор исторических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Л.В. Беловинский, доктор исторических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

В.В. Гудкова, доктор искусствоведения, Государственный институт искусствознания, Москва, Российская Федерация

Н.П. Грицлер, доктор филологических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Ю.В. Доманский, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

А.В. Дыбо, член-корреспондент РАН, доктор филологических наук, профессор, Институт языкознания РАН, Москва, Российская Федерация

И. Жетниковска, доктор филологических наук, Университет Николая Коперника, Торунь, Республика Польша

Г.И. Зверева, доктор исторических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация (*заместитель главного редактора*)

И.И. Исаев, кандидат филологических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Г.И. Кабакова, доктор филологических наук, университет Сорбонны, Париж, Франция

Н.В. Капустин, доктор филологических наук, профессор, Ивановский государственный университет, Иваново, Российская Федерация

В.И. Киммельман, PhD., Берген, Королевство Норвегия

Д.Д. Клейтон, доктор филологических наук, Оттавский университет, Оттава, Канада

И.В. Кондаков, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Г.Е. Крейдлин, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Л.И. Куликов, кандидат филологических наук, PhD., Гентский университет, Гент, Королевство Бельгия

- М.Н. Липовецкий*, доктор филологических наук, профессор, Университет Колорадо Болдер, США
- Д.М. Магомедова*, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- И.В. Морозова*, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- В.Г. Мостовая*, кандидат филологических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- С.Ю. Неклюдов*, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- В.И. Подлеская*, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- О.И. Половинкина*, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- Е.Ю. Протасова*, доктор педагогических наук, Хельсинкский университет, Хельсинки, Финляндская Республика
- Р.И. Розина*, доктор филологических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- Я. Садовский*, доктор исторических наук, Ягеллонский университет, Краков, Республика Польша
- А.Ю. Сорочан*, доктор филологических наук, доцент, Тверской государственный университет, Тверь, Российская Федерация
- Я.Г. Тестелец*, доктор филологических наук, доцент, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- О.И. Тогоева*, доктор исторических наук, Институт всеобщей истории РАН, Москва, Российская Федерация
- В.И. Тюпа*, доктор филологических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- А.А. Холиков*, доктор филологических наук, Московский государственный университет им. М.В. Ломоносова (МГУ), Москва, Российская Федерация
- О.Б. Христофорова*, доктор филологических наук, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- И.О. Шайтанов*, доктор исторических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация
- С.А. Яценко*, доктор исторических наук, профессор, Российский государственный гуманитарный университет (РГГУ), Москва, Российская Федерация

Ответственный за выпуск:

П.М. Аркадьев, доктор филологических наук (РГГУ)

Contents

| | |
|---------------------------|----|
| Editorial statement | 10 |
|---------------------------|----|

Chinese linguistics

Taras V. Ivchenko

| | |
|---|----|
| Chinese language: issues of standardization and terminology. Based on materials from the International Conference “Chinese Linguistics and Sinology”, Russian State University for the Humanities, October 3–5, 2019 | 11 |
|---|----|

Natalya Yu. Chechina

| | |
|--|----|
| Features of terminologization of common vocabulary in modern Chinese (on the example of economic discourse of Chinese-language media) | 16 |
|--|----|

Kirill I. Semenov

| | |
|---|----|
| Adaptation strategies of Russian phonetic loanwords in Chinese. Phonetic and graphic aspects | 30 |
|---|----|

Aleksandr N. Sboev

| | |
|--|----|
| Chinese language “latinization”. (On the basis of Chinese Internet communication) | 64 |
|--|----|

Research articles

I.G. Bagirokova, A.I. Baranova, Yu.A. Lander

| | |
|--|----|
| Data for the typology of associative plurals: Kuban Kabardian | 84 |
|--|----|

*Olga I. Vinogradova, Elizaveta A. Smirnova,
Anna V. Viklova, Irina M. Panteleeva*

| | |
|---|-----|
| Syntactic complexity of academic text. A Corpus Study of Written Production by Learners of English with Russian L1 in Comparison with Expert Texts of English Authors | 107 |
|---|-----|

Oksana A. Kanerva

| | |
|---|-----|
| Russian onomatopoeic verbal interjections. Why use ‘non-words’ instead of ordinary ones? | 130 |
|---|-----|

In memoriam

Leonid P. Krysin, Leonid L. Iomdin

| | |
|---|-----|
| Vladimir Zinovyevich Sannikov (1931–2019) | 147 |
|---|-----|

Содержание

От редакции «Московского лингвистического журнала» 10

Китайская лингвистика

Т.В. Ивченко

Китайский язык: вопросы стандартизации и терминологичности:
по материалам Международной конференции «Китайская лингвистика
и синология», РГГУ, 3–5 октября 2019 г. 11

Н.Ю. Чечина

Особенности терминологизации общеупотребительной лексики
современного китайского языка: на примере экономического
дискурса СМИ КНР 16

К.И. Семенов

Стратегии преобразования русских фонетических заимствований
в китайском языке: фонетические и графические аспекты 30

А.Н. Сбоев

«Латинизация» китайского языка:
на примере китаеязычной интернет-коммуникации 64

Описательные и теоретические исследования

И.Г. Багироква, А.И. Баранова, Ю.А. Ландер

Материалы к типологии ассоциативной множественности:
кубанский диалект кабардино-черкесского языка 84

О.И. Виноградова, Е.А. Смирнова, А.В. Виклова, И.М. Пантелеева

Синтаксическая сложность академического текста:
корпусное исследование работ русскоязычных студентов,
изучающих английский язык, в сравнении с текстами
профессиональных авторов 107

Оксана А. Канерва

Звукоподражательные глагольные междометия русского языка:
зачем использовать «неполноценные» слова вместо обычных? 130

In memoriam

Л.П. Крысин, Л.Л. Иомдин

Владимир Зиновьевич Санников (1931–2019) 147

От редакции «Московского лингвистического журнала»

В 22-й номер «Московского лингвистического журнала» помимо регулярных статей вошли три статьи по материалам Международной конференции «Китайская лингвистика и синология», состоявшейся в РГГУ в октябре 2019 г., предваренные вступительной заметкой Т.В. Ивченко. Редакция «МЛЖ» выражает глубокую признательность Т.В. Ивченко, М.Б. Рукодельниковой и Л.С. Холкиной за идею тематической секции, посвященной китайской лингвистике, и за помощь в рецензировании и подготовке этих статей к публикации и рассчитывает на сотрудничество в будущем.

Также редакция «МЛЖ» выражает глубокую признательность внешним рецензентам статей последних двух выпусков: В.М. Алпатову, А.В. Архипову, А.Л. Полян, А.Н. Баранову, М.Д. Воейковой, Е.В. Горбовой, П.В. Гращенкову, Д.О. Добровольскому, Л.Л. Иомдину, Ю.Б. Корякову, Л.О. Наний, Е.В. Перехвальской, С.Ю. Толдовой, К.А. Ульяновой.

Редакционная коллегия выпуска:

П.М. Аркадьев (отв. ред., Ин-т славяноведения РАН/РГГУ), А.М. Ивойлова (секретарь, РГГУ), О.И. Беляев (МГУ/Ин-т языкознания РАН), А.В. Дыбо (Ин-т языкознания РАН/НИУ ВШЭ), И.И. Исаев (РГГУ), В.И. Киммельман (Бергенский университет, Норвегия), М.Б. Коношенко (РГГУ), Н.А. Коротаев (РГГУ), Г.Е. Крейдлин (РГГУ), М.А. Кронгауз (РГГУ/НИУ ВШЭ), Л.И. Куликов (Гентский университет, Бельгия), С.А. Оскольская (Ин-т лингвистических исследований РАН), Б.Х. Парти (Университет Амхерста, США), О.Е. Пекелис (РГГУ), А.Ч. Пиперски (РГГУ/НИУ ВШЭ), В.И. Подлеская (РГГУ), Е.Ю. Протасова (Хельсинкский университет, Финляндия), Р.И. Розина (Ин-т русского языка РАН/РГГУ), А.В. Сидельцев (Ин-т языкознания РАН/РГГУ), Н.А. Слюсарь (НИУ ВШЭ/СПбГУ), Н.Р. Сумбатова (РГГУ), Я.Г. Тестелец (РГГУ/Ин-т языкознания РАН), А.Б. Шлуинский (Ин-т языкознания РАН)

Китайская лингвистика

УДК 811.581

DOI: 10.28995/2686-7249-2020-7-11-15

Китайский язык:
вопросы стандартизации и терминологичности:
по материалам Международной конференции
«Китайская лингвистика и синология»,
РГГУ, 3–5 октября 2019 г.

Тарас В. Ивченко

*Российский государственный гуманитарный университет
Москва, Россия, tarasivchenko@yandex.ru*

Аннотация. Данная публикация предвещает издание трех статей по материалам докладов, представленных на Международной конференции «Китайская лингвистика и синология» (РГГУ, 3–5 октября 2019 г.). Конференция была посвящена памяти выдающегося советского и российского лингвиста Тянь Аошуан (1931–2017) и собрала вместе многих лингвистов и сиологов. Для этого выпуска отобраны близкие по теме статьи, а настоящие вступительные замечания призваны «задать фон» для погружения в соответствующую проблематику.

Ключевые слова: китайский язык, путунхуа, терминология, заимствования

Для цитирования: Ивченко Т.В. Китайский язык: вопросы стандартизации и терминологичности: по материалам Международной конференции «Китайская лингвистика и синология», РГГУ, 3–5 октября 2019 г. // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 11–15. DOI: 10.28995/2686-7249-2020-7-11-15

Chinese language: issues of standardization and terminology.
Based on materials from the International Conference
“Chinese Linguistics and Sinology”,
Russian State University for the Humanities, October 3–5, 2019

Taras V. Ivchenko

Russian State University for the Humanities, Moscow, Russia, tarasivchenko@yandex.ru

Abstract. This publication serves as an introduction to the section comprising three articles based on presentations at the International Conference “Chinese Linguistics and Sinology” (Russian State University for the Hu-

© Ивченко Т.В., 2020

manities, October 3–5, 2019). The conference was dedicated to the memory of the outstanding Soviet and Russian linguist Tan Aoshuang (1931–2017) and brought together many linguists and sinologists. Related articles were selected for this issue, and the introductory remarks provided are designed to set the background for the issues discussed therein.

Keywords: Chinese, putonghua, terminology, loan words

For citation: Ivchenko, T.V. (2020), “Chinese language: issues of standardization and terminology. Based on the materials of the International Conference “Chinese Linguistics and Sinology”, Russian State University for the Humanities, October 3–5, 2019”, *RSUH/RGGU Bulletin. “Literary Theory. Linguistics. Cultural Studies” Series*, no. 7, pp. 11-15. DOI: 10.28995/2686-7249-2020-7-11-15

Китайский язык в последние 30–40 лет находится в процессе постоянной эволюции и трансформации. Под влиянием извне и по внутренним причинам он стремительно меняет свой облик, что затрагивает практически все уровни языка. Появляются новые словообразовательные модели, новые синтаксические конструкции, меняется управление глаголов, изменяется представление о грамматической правильности, а новые пласты лексики захлестывают средства массовой информации и Интернет.

За внешней очевидностью этих явлений иногда не замечают принципиально важного процесса: в КНР в настоящее время формируется новый китайский язык с реальным статусом «общий». Другими словами, многодиалектный Китай постепенно переходит от диглоссии (диалект (разговорный язык в быту) – литературный язык (язык СМИ, литературы и образования)) к моноглоссии, в результате чего появляется большое количество носителей китайского языка, владеющих только литературным вариантом.

Необходимость пользоваться общелитературным вариантом связана, в частности, с неспособностью диалектов решить вопрос создания собственной системы терминов, которая бы охватывала все аспекты жизни современного общества и была бы понятна носителям других диалектов, то есть имела бы наддиалектный характер. Статьи, представленные в «китайской» части этого выпуска, посвящены как раз актуальным вопросам формирования терминологии в общелитературном китайском языке «путунхуа».

Нельзя сказать, чтобы эта тема была абсолютно нова. Предлагаемые статьи интересны не столько новизной проблематики, сколько попыткой увидеть новые закономерности давно идущих процессов. Но сначала необходимо пояснить общий контекст данной проблемы.

Речь прежде всего идет о формировании новой терминологии в языке с иероглифической письменностью, который практически лишен возможности фонетического заимствования. Новые термины привносят в язык новые концепты и новые «способы мышления», а вместе с ними и новые способы построения дискурса. Синтаксис предложения и особенности построения нарратива в китайском языке претерпели существенные изменения после появления большого количества текстов (художественных, философских, технических), переведенных с других языков. Эта проблема еще ожидает своего более целостного анализа.

Особенностью заимствований в китайском языке, как удалось показать Наталье Чечиной в статье «Особенности терминологизации общеупотребительной лексики современного китайского языка (на примере экономического дискурса СМИ КНР)», является необходимость решения целого набора задач разного уровня. Какие это задачи?

Во-первых, практически каждый термин в китайском языке с необходимостью получает свою внутреннюю форму: он не может «скрыть» свое значение, например, за латинскими или греческими корнями. С этой точки зрения новый термин должен соответствовать правилам внутреннего синтаксиса сложного слова в китайском языке, то есть правилам словосложения.

Во-вторых, новый термин должен быть сравнительно легко вычленим в предложении, то есть его внутренняя форма должна отличаться от «внешнего» синтаксиса предложения или, как это подчеркивается в статье, каждый термин должен обладать достаточной степенью «терминологичности».

В-третьих, китайский язык меняет механизм построения терминологии той или иной предметной области в зависимости от общего объема терминов и концептов, который необходимо «переварить». В связи с колоссальным объемом новой терминологии, заимствуемой прежде всего из английского языка, постепенно выработался способ поморфемного калькирования с частичным сохранением сочетаемости и модели управления языка-оригинала. Это, в свою очередь, приводит к появлению в китайском языке все большего количества терминов с «непонятной» внутренней структурой и новых синтаксических конструкций, а внутренний синтаксис термина в результате меньше влияет на его синтаксическую сочетаемость.

В-четвертых, иероглифы, используемые для записи нового термина, должны обладать нейтральной семантикой и не вызывать двусмысленных ассоциаций. Для китайского языка иероглифика далеко не всегда является просто «способом фиксации устного

сообщения». При подборе того или иного иероглифа учитывается весь комплекс передаваемых им значений, которые «аккумулируются» в результате длительного употребления в разных контекстах.

В-пятых, китайский язык, конечно, не полностью лишен возможности фонетического заимствования терминов из других языков. Правда, это нередко ограничивается именами собственными, географическими названиями, единицами измерения и некоторыми специальными терминами. В случае специальных терминов обращение к передаче звучания интересна как «крайняя и последняя» мера: в таких случаях имеет смысл задаваться вопросом, почему не была использована другая стратегия.

Для фонетических заимствований должны быть решены задачи выделения множества «фонетических» иероглифов, стандартизации этого набора и формулирования единых правил «фонетико-иероглифической» передачи имен собственных. Это в чем-то сходно с процессом формирования набора стандартных фонетиков для иероглифов типа фоноидеограмм, который происходил в истории развития китайской письменности. Фонетическим и графическим аспектам заимствований посвящена статья Кирилла Семенова «Стратегия преобразования русских фонетических заимствований в китайском языке: фонетические и графические аспекты». Русские фонетические заимствования интересны прежде всего своей трудностью для передачи средствами китайского языка и вариативностью в истории, что объясняется длительными языковыми контактами.

К графическому аспекту заимствований можно отнести явление сочетания разных графических систем (в первую очередь иероглифики и латиницы) в контексте одного термина. Об этом аспекте рассказывается в статье Александра Сбоева «“Латинизация” китайского языка (на примере китаеязычной интернет-коммуникации)». Интернет, который является источником огромного количества неологизмов, интересен как зона свободного и неконтролируемого творчества, на которую не накладываются практически никакие ограничения. Для китайского языка она представляет особый интерес, так как языковое творчество всегда считалось прерогативой лишь высокообразованной части населения.

В результате подобранные статьи позволяют показать формирование новой (большой частью заимствованной) терминологии как неотъемлемую часть процесса постоянного обновления современного китайского языка и частично объясняют то, как формируется современный путунхуа, который сами китайские лингвисты называют «европеизированным вариантом китайского языка».

Информация об авторе

Тарас В. Ивченко, доктор филологических наук, доцент, Российский государственный гуманитарный университет, Москва, Россия; 125993, Россия, Москва, Миусская пл., д. 6; tarasivchenko@yandex.ru

Information about the author

Taras V. Ivchenko, Dr. of Sci. (Philology), associate professor, Russian State University for the Humanities, Moscow, Russia; bld. 6, Miusskaya Sq., Moscow, Russia, 125993; tarasivchenko@yandex.ru

Особенности терминологизации общеупотребительной лексики современного китайского языка: на примере экономического дискурса СМИ КНР

Наталья Ю. Чечина

*Московский государственный лингвистический университет
Москва, Россия, nataliachchina@yandex.ru*

Аннотация. Статья посвящена особенностям функционирования экономической лексики в материалах китайскоязычных СМИ. Медиа-материалы отражают основные изменения, которые происходят в жизни общества, прежде всего в экономической сфере. Для Китая большое значение имеет развитие рыночной экономики, что также отражено в текстах средств массовой информации. Медиа-материалы являются источниками новой информации, их получателями в основном выступают широкие слои населения. Тем не менее экономика – это профессиональная сфера деятельности со своей терминологической системой. Основная задача медийных материалов – передача информации от специалистов в данной области неспециалистам. Таким образом, с учетом обилия терминов тексты деловых СМИ обогащают лексический запас слушателей или читателей. В то же время анализ текстов средств массовой информации на китайском языке показал, что в настоящее время идет активный процесс терминологизации (переход общеупотребительной лексики в категорию терминов). Этот процесс оказывает влияние на повседневное общение носителей языка и недостаточно изучен на материале современного китайского языка.

Ключевые слова: китайский язык, СМИ, экономические термины, терминологизация, общеупотребительная лексика

Для цитирования: Чечина Н.Ю. Особенности терминологизации общеупотребительной лексики современного китайского языка: на примере экономического дискурса СМИ КНР // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 16–29. DOI: 10.28995/2686-7249-2020-7-16-29

Features of terminologization of common vocabulary in modern Chinese (on the example of economic discourse of Chinese-language media)

Natalya Yu. Chechina

Moscow State Linguistic University, Moscow, Russia, nataliachchina@yandex.ru

Abstract. The paper is devoted to the features of the economic vocabulary functioning in the materials of the Chinese-language media. Media materials reflect the major changes that are occurring in the life of society, primarily in the economic sphere. For China, the development of a market economy is of great importance, which is also reflected in the texts of the mass media. Media materials are sources of new information, their recipients are mainly the general public. Nevertheless, economics is a professional field of activity with its terminological system. The main task of media materials is to transfer information from specialists in this field to non-specialists. Thus, taking into account the abundance of terms, the texts of the economic media enrich the lexicon of listeners or readers. At the same time, an analysis of the texts of the Chinese-language media showed that an active process of terminologization (the transition of commonly used words to the category of terms) is currently underway. This process has an impact on the daily communication of native speakers and is not enough studied on the material of modern Chinese.

Keywords: Chinese language, mass media, economic terms, terminologization, common vocabulary

For citation: Chechina, N.Y. (2020), "Features of terminologization of common vocabulary in modern Chinese (on the example of economic discourse of Chinese-language media)", *RSUH/RGGU Bulletin. "Literary Theory. Linguistics. Cultural Studies" Series*, no. 7, pp. 16-29, DOI: 10.28995/2686-7249-2020-7-16-29

На современном этапе материалы СМИ являются источником новых значений слов и контекстов их употребления. Кроме этого, язык СМИ оказывает огромное влияние на формирование массового сознания.

Язык СМИ позволяет наблюдать за эволюционными процессами в лексическом составе языка. По мере того как происходят эволюционные процессы в обществе, изменяется словарный состав языка.

Ввиду тематического многообразия лексика газеты также должна быть очень богатой. В публицистике широко представлена общепотребительная, нейтральная лексика и фразеология, книжная и разговорная речь. Зачастую в языке газеты используются термины как в прямом значении в соответствующем контексте, так и в образном, переносном значении.

Так как газетный язык призван информировать читателей о политической или экономической ситуации в стране, объяснять важнейшие социально-экономические и политические проблемы и задачи, формировать у читателя определенное видение данных проблем, в публицистических текстах встречается большое количество специальной лексики (политическая, экономическая, социологическая терминология и др.). Таким образом, язык газеты сближается с научным стилем. В научном стиле терминологическая лексика составляет в среднем 15–25% от общей лексики. Термины составляют ядро научного стиля.

В публицистическом тексте термины выполняют особые функции. Они являются неотъемлемой частью языка газеты, отражают многообразие содержания информации, помогают передавать ее и воздействовать на читателя.

Термины в прямом значении выполняют функцию передачи информации. Они употребляются в своем терминологическом значении для передачи объективной информации, служат свидетельством ее достоверности.

Термин в публицистическом тексте выражает основные признаки понятия, носит формальный характер. Читатель, обладая определенным багажом знаний, приблизительно понимает суть термина. Постепенно термин становится доступным для читателя-неспециалиста, повышая, таким образом, его уровень знаний.

Термины в переносном значении в основном выполняют функцию воздействия. Несмотря на то что понятийно-предметная соотнесенность термина и его строгое содержание исключают экспрессивность, функционируя в неспециальных контекстах, в том числе в материалах СМИ, термины приобретают оттенок экспрессивности. Метафоризация терминов широко используется как источник экспрессии в газетном тексте.

Материалы СМИ отражают важнейшие изменения, которые происходят в жизни общества, в первую очередь в экономической сфере. Для КНР развитие рыночной экономики с китайской спецификой является первостепенной задачей, что находит свое отражение и в текстах масс-медиа. Материалы СМИ являются источниками новой информации, их реципиентом в основном выступают широкие слои общественности. Тем не менее экономика – профессиональная сфера деятельности со своей терминосистемой. Основной задачей материалов масс-медиа является передача информации от специалистов в данной области неспециалистам. Таким образом, учитывая обилие терминов, тексты СМИ экономической тематики пополняют лексический запас слушателей или читателей. Вместе с тем анализ текстов китайскоязычных СМИ показал,

что в настоящее время происходит активный процесс терминологизации – перехода общеупотребительных слов в разряд терминов. Этот процесс оказывает влияние на повседневную коммуникацию носителей языка и является недостаточно изученным на материале современного китайского языка.

Политика реформ и открытости (改革开放 gǎigé kāifàng), курс на которую был взят в КНР в конце 70-х гг. XX в. по инициативе Дэн Сяопина, привела к значительным изменениям во всех областях китайского общества и в первую очередь в экономике. В настоящее время Китай активно строит социалистическую рыночную экономику и проводит «политику открытых дверей». Благодаря усилиям китайского руководства Китай уже стал второй экономикой мира по номинальному ВВП и первой – по ВВП по паритету покупательной способности (с 2014 г.).

Экономические процессы постоянно находятся в центре внимания общественности и, естественно, оказывают влияние на содержание текстов масс-медиа. А материалы средств массовой информации, которые являются источником новой информации, направленной на широкую целевую аудиторию, обогащают лексический запас читателей и слушателей.

Материалы СМИ чаще всего выступают в качестве посредника при передаче справочной или аналитической информации от специалистов широкой аудитории, большую часть которой составляют неспециалисты в данной области знаний. Таким образом, лексический состав неспециалистов обогащается экономической терминологией и профессионализмами. В то же время происходит упрощение терминологической лексики для более адекватного восприятия новой информации неподготовленным реципиентом. Тем не менее подобное устное или письменное сообщение не может не содержать специальной экономической лексики.

Экономический термин играет существенную роль в медийном тексте – информирует и ориентирует получателя сообщения, что отвечает его познавательным потребностям и интересам [Голованова 2008, с. 79].

Термины не противоречат природе публицистики, напротив, они близки газетно-публицистическому стилю точностью номинации, глубиной обобщения понятия, однозначностью, лаконизмом выражения, составляющим один из важных компонентов эстетического идеала газетно-публицистического стиля. Вообще термины не выглядят чужеродными в газете именно вследствие книжно-письменной природы газетного языка [Солганик 1981, с. 73].

Специальная лексика формируется на базе общенародного языка и в то же время постоянно пополняет и обогащает его систему. Термин становится элементом общелитературного языка тогда, когда его содержание становится широко известным и выходит за пределы системы специальных понятий. Между общелитературным языком и терминологией происходит постоянный обмен. Помимо этого, некоторые слова общелитературного языка становятся терминами в результате семантического преобразования. Такие процессы называются детерминологизацией и терминологизацией. Они свидетельствуют о взаимопроникновении терминологической и нетерминологической лексики.

В ходе терминологизации общеупотребительной лексики реализуются процессы сужения или расширения семантики лексической единицы, различные типы переноса, т. е. процессы, приводящие в итоге к переходу лексической единицы в терминосферу.

Как утверждает китайский лингвист Лян Юн в своей статье «Терминологизация и перевод терминов» (梁镛 «术语化和术语翻译»), лексическая единица может обладать потенциальным, первоначальным и специальным значениями.

Потенциальное значение отражает все семантические особенности лексической единицы, которые она имеет в рамках данной языковой системы. Оно имплицитно заложено в значении слова и актуализируется в процессе общения. Между объектом описания и потенциальным значением слова не существует четко фиксированной связи.

Первоначальное значение – основное значение, которым обладает слово в процессе практического употребления. Является компонентом потенциального значения слова. Взаимосвязь между первоначальным значением слова и его морфологической формой произвольная, однако если она устанавливается, практическое употребление данной лексической единицы в определенной степени ограничивается привычками в общении. Процесс терминологизации обычно начинается с изменением первоначального значения слова.

Специальное значение – это своего рода продукт процесса терминологизации. В отличие от первоначального значения слова, существует фиксированная связь между специальным значением и морфологической формой слова [Лян Юн, 1986].

Термин «терминологизация» в настоящее время не имеет строгой понятийной закреплённости. В широком понимании терминологизацией называют все процессы образования новых терминов в результате семантического переосмысления общеупотребительной лексики, способ языковой номинации, при котором основным

средством создания нового термина является семантическое развитие слова.

В узком смысле под терминологизацией понимаются отдельные изменения в семантической структуре лексической единицы в процессе ее перехода в разряд термина (при таком понимании терминологизация оказывается одним из подвидов семантического способа) [Думитру 2008, с. 8].

Ряд ученых (О.Б. Алексеева, И.Д. Бесценная, О.П. Исаева, Г.П. Снетова и др.) понимают под терминологизацией процесс сужения значения общеупотребительного слова. «Основной путь проникновения общеупотребительной лексики в состав специальной терминологии – это сужение общеупотребительного широкого значения, сопровождающегося его специализацией» [Снетова 1984, с. 61].

Лингвисты расходятся во мнении относительно отношений между исходным словом и термином, который образовался в результате изменений в его семантической структуре. Некоторые исследователи полагают, что при терминологизации используется один из дифференциальных признаков общеупотребительного слова, который берется в качестве «общей идеи» для термина, которая впоследствии конкретизируется исходя из содержания научного понятия [Даниленко 1977, с. 17].

Приведем некоторые примеры, взятые из материалов СМИ КНР.

| | | | | | | | |
|-----|-------------------|---------------|----------|---------|-----------------------|----------|---------|
| (1) | 23日, | 两市 | 小幅 | 低 | 开, | 开盘 | 之后 |
| | 23 rì | liǎngshì | xiǎofú | dī | kāi | kāipán | zhī hòu |
| | 23 число | две биржи | немного | низкий | открывать | начинать | после |
| | | | | | | торги | |
| | 在 | 热点 | 板块 | 的 | 杀跌 | | |
| | zài | rèdiǎn | bǎnkuài | de | shādiē | | |
| | LOC | горячая точка | плита | ATR | играть на понижение | | |
| | 之 | 下, | 指数 | 迅速 | 走 | 弱, | |
| | zhī | xià | zhǐshù | xùnsù | zǒu | ruò | |
| | ATR | DIR | индекс | быстро | идти | RES | |
| | 沪指 | | 一度 | 跌破 | 3200 点 ¹ . | | |
| | hùzhǐ | | yīdù | diē pò | 3200 diǎn. | | |
| | индекс Шанхайской | | один раз | разбить | 3200 пунктов | | |
| | фондовой биржи | | | | | | |

¹ Worldstreetcn. Com [Электронный ресурс]. URL: <https://wallstreetcn.com/articles/3516279> (дата обращения 01 августа 2019).

‘23 числа торги на Шанхайской и Шэньчжэньской биржах открылись в небольшом минусе. Из-за продажи акций основных секторов фондового рынка после начала торгов индексы начали стремительно падать, сводный индекс Шанхайской фондовой биржи Shanghai Composite снизился до 3200 пунктов’.

В данном предложении используется общеупотребительное слово 板块 *bǎnkuài*, которое имеет следующие значения: 1) кусок плиты, плита; лист; доска; 2) блок, массив, часть; 3) геол. плита (тектоническая или литосферная). Здесь 板块 выступает в качестве финансового термина «ценные бумаги, акции определенного сектора фондового рынка».

| | | | | | | | | | |
|-----|---------------|---------------|----------|-------------------|-------------|------------|------------|--------|----------|
| (2) | 在 | 这 | 之前, | 亿达 | 中国 | 曾 | 于 | 4月 | 10日 |
| | Zài | zhè | zhīqián | Yídá | Zhōngguó | céng | yú | 4 yuè | 10 rì |
| | LOC | этот | ранее | Yida | Китай | когда-то | LOC | апрель | 10 число |
| | 公告 | | 称, | 由于 | 中 | 民 | 投 | | |
| | gōnggào | | chēng | yóuyú | Zhōng | Mín | tóu | | |
| | информировать | | заявлять | из-за | Китай | народ | инвестиции | | |
| | 被 | 实施 | | 资产 | 冻结, | | | | |
| | bèi | shíshī | | zīchǎn | dòngjié | | | | |
| | PASS | осуществлять | | активы | заморозка | | | | |
| | 在 | 技术 | 上 | 导致 | 亿达 | 中国 | | | |
| | zài | jìshù | shàng | dǎozhì | Yídá | Zhōngguó | | | |
| | LOC | техника | DIR | приводить | Yida | Китай | | | |
| | 未 | 偿还 | 本 | 金 | 总额 | 为 | | | |
| | wèi | chánghuán | běn | jīn | zǒng'ér | wèi | | | |
| | NEG | возвращать | данный | деньги | общая сумма | составлять | | | |
| | 42.77 | 亿 | | 元 | 的 | 贷款 | | | |
| | 42.77 | yì | | yuán | de | dàikuǎn | | | |
| | 42.77 | сто миллионов | | юань | ATR | кредит | | | |
| | 触发 | 提前 | 赎回 | 条款 ² . | | | | | |
| | chùfā | tíqián | shúhuí | tiáokuǎn. | | | | | |
| | возбуждать | заранее | выкуп | условия | | | | | |

‘До этого Yida China Hldgs 10 апреля объявила о том, что заморозка активов China Minsheng Investment Group технически создала условия для досрочного выкупа непогашенного займа по основному капиталу Yida China Hldgs в размере 4 млрд 277 млн юаней’.

² Sohu.com [Электронный ресурс]. URL: http://www.sohu.com/a/309766317_260616 (дата обращения 11 июля 2019).

В данном примере 资产冻结 *zīchǎn dòngjié* имеет значение экономического термина «заморозка активов». Тем не менее само по себе слово 冻结 *dòngjié* является общеупотребительным и имеет значение «замерзать; застывать; стать (о реке); замораживание, замерзание; мерзлота; мерзлый».

| | | | | | | | | |
|-----|-------------|--------------------------|------------------|--------------|-----------|------------------|-----------|-------------------|
| (3) | 周二, | 英国 | 央行 | 总裁 | 卡尼 | | | |
| | Zhōu`èr | Yīngguó | yāngháng | zǒngcái | kǎ ní | | | |
| | вторник | Великобритания | Центральный банк | председатель | Карни | | | |
| | 表示, | 随着 | 无 | 协议 | 脱 | 欧 | | |
| | biǎoshì | suízhe | wú | xiéyì | tuō | ōu | | |
| | заявлять | вслед за | NEG | соглашение | выходить | Европейский союз | | |
| | 可能性 | 的 | 提升, | 英国 | 经济 | 面临 | 的 | |
| | kěnéngxìng | de | tíshēng | Yīngguó | jīngjì | miànlin | de | |
| | вероятность | ATR | повышаться | В-британия | экономика | сталкиваться | ATR | |
| | 风险 | 与日俱增, | | 可能 | 需要 | | | |
| | fēngxiǎn | yǔrìjùzēng | | kěnéng | xūyào | | | |
| | риск | с каждым днем возрастать | | возможно | требуется | | | |
| | 更 | 多 | 的 | 帮助 | 来 | 应对 | 经济 | 滑坡 ³ . |
| | gèng | duō | de | bāngzhù | lái | yìngduì | jīngjì | huápō. |
| | более | много | ATR | помощь | чтобы | бороться | экономика | оползень |

‘Во вторник председатель ЦБ Великобритании Марк Карни заявил, что, поскольку растет вероятность «жесткого Brexit», возрастают риски, с которыми сталкивается экономика страны. Возможно, потребуется дополнительная помощь для борьбы с экономическим кризисом’.

В данном предложении используется словосочетание 经济滑坡 *jīngjì huápō* в значении «экономический кризис, снижение экономической активности, застой в экономике». Само по себе слово 滑坡 не является экономическим термином, это общеупотребительное слово со значением «оползень, обвал».

| | | | | | | |
|-----|---------|-------|------------------|--------------|--------|-------|
| (4) | 午 | 后 | 大盘 | 维持 | 低 | 位 |
| | Wǔ | hòu | dàpán | wéichí | dī | wèi |
| | полдень | после | основные индексы | поддерживать | низкий | место |

³ Reuters [Электронный ресурс]. URL: <https://www.reuters.com/article/britain-boe-carney-wrap-up-risks-0702-tue-idCNKCS1TX31M> (дата обращения 05 августа 2019).

| | | | | | |
|--------------|-----------|------------|--------------------|--------------------|-------------------|
| 弱 | 势 | 震荡, | 农业、 | 船舶 | |
| ruò | shì | zhèndàng | nóngyè | chuánbó | |
| слабый | положение | потрясение | сельское хозяйство | корабли | |
| 制造 | 等 | 板块 | 小幅 | 异动, | |
| zhìzào | děng | bǎnkuài | xiǎofú | yìdòng, | |
| производство | такие как | плита | немного | необычная динамика | |
| 但 | 市场 | 情绪 | 仍 | 十分 | 低迷 ⁴ 。 |
| dàn | shìchǎng | qíngxù | réng | shífēn | dīmí. |
| но | рынок | настроение | по-прежнему | очень | упадок |

‘Во второй половине дня наблюдалось снижение основных фондовых индексов, и, несмотря на незначительные изменения котировок в таких секторах, как сельское хозяйство и судостроение, рыночные настроения по-прежнему пессимистичные’.

Общепотребительное слово 大盘 dàpán имеет значения: 1) большое блюдо; 2) подставка со льдом под ложе умершего императора (рит.), однако в контексте рынка ценных бумаг имеет частотное употребление со значением «рыночные котировки, рыночный индекс».

| | | | | | | |
|----------------|-------------------------------|-----------|---------|-------|-----------|-------------------|
| (5) 根据 | 招股书 | 资料, | | | | |
| Gēnjù | zhāogǔ shū | zīliào | | | | |
| в соответствии | проспект эмиссии ценных бумаг | материалы | | | | |
| 中信 | 出版 | 的 | 存 | 货 | 周转 | 率 |
| zhōngxìn | chūbǎn | de | cún | huò | zhōuzhuǎn | lǜ |
| CITIC | издавать | ATR | хранить | товар | оборот | коэффициент |
| 明显 | 低 | 于 | 行业 | 的 | 平均 | 水平 ⁵ 。 |
| míngxiǎn | dī | yú | hángyè | de | píngjūn | shuǐpíng. |
| ясный | ниже | чем | отрасль | ATR | средний | уровень |

‘Согласно проспекту эмиссии ценных бумаг, оборот запасов Китайской международной трастовой и инвестиционной корпорации (CITIC) значительно ниже средних показателей в данной отрасли’.

周转 zhōuzhuǎn имеет значение «вращение, кружение, кругооборот». В экономическом дискурсе употребляется со значением «оборот». В данном примере можно увидеть словосочетание 存货周转率

⁵ Sina.cn. [Электронный ресурс]. URL: <https://finance.sina.cn/stock/gspl/2019-07-08/detailihycitm0515126.d.html?cre=tianyi&mod=wpage&loc=6&r=32&rfunc=79&tj=none&tr=32&cref=cj> (дата обращения 08 августа 2019).

cúnhuò zhōuzhuǎn lǜ – «оборот запасов, коэффициент оборачиваемости материальных запасов».

- (6) 快速 爆发 到 泡沫 破裂,
 Kuàisù bàofā dào pàomò pòliè
 быстрый взрыв RES мыльный пузырь лопнуть
- 再 回归 到 稳步 增长,
 zài huíguī dào wěnbù zēngzhǎng
 снова возвращаться RES стабильный рост
- 是 几乎 所有 行业 发展 的
 shì jīhū suǒyǒu hángyè fāzhǎn de
 являться почти все отрасль развитие ATR
- 必 经 之 路, 泡沫 破裂
 bì jīng zhī lù, pàomò pòliè
 необходимо проходить ATR путь пузырь лопаться
- 也 意味 着 新 的 机会 即将 到来⁶.
 yě yìwèi zhe xīn de jīhuì jíjiāng dàolái.
 тоже означать PROG новый ATR шанс скоро приходит

‘Быстрый рост, взрыв «мыльного пузыря», за которым вновь следует стабильное развитие – вот путь, по которому проходит любая отрасль. Такой «взрыв пузыря» приносит новые возможности для развития’.

В данном примере используется общеупотребительное слово 泡沫 pàomò «пузырь» или «мыльный пузырь» в переносном значении. «Пузырь» (также называемый «экономический пузырь», «финансовый пузырь», «рыночный пузырь») – это торговля большими объемами товара или бумагами по цене, существенно выше справедливой. Учитывая огромный спрос на товар, цена возрастает, что приводит к еще большему спросу. Когда «пузырь» лопается, центральный банк вынужден пересматривать монетарную политику и понижать ликвидность в финансовой системе, чтобы минимизировать риск обвала своей валюты.

- (7) 监管 机构 已经 提出 了 一些 可行 的
 Jiānguǎn jīgòu yǐjīng tíchū le yīxiē kěxíng de
 контроль орган уже выдвинуть PRF некоторый осуществимый ATR

⁶ Sina.cn. [Электронный ресурс]. URL: https://finance.sina.cn/stock/ssgs/2019-07-05/detailihytcerm1579472.d.html?oid=3822017565952505&vt=4&p_os=17&wm=4007 (дата обращения 10 июля 2019).

| | | | | | | |
|----------|--------|----------------|----------|-------------------|--------------|-----|
| 政策 | 措施, | 重点 | 放 | 在 | 限制 | 不 |
| zhèngcè | cuòshī | zhòngdiǎn | fàng | zài | xiànzhì | bù |
| политика | меры | ключевой пункт | ставить | RES | ограничивать | NEG |
| 成熟 | 客户 | 的 | 交易 | 杠杆 ⁷ . | | |
| chéngshú | kèhù | de | jiāoyì | gànggǎn. | | |
| зрелый | клиент | ATR | операция | рычаг | | |

‘Регулирующий орган уже предложил ряд действенных мер, в первую очередь сделав акцент на ограничение финансового леввериджа для неопытных клиентов’.

Общеупотребительное слово 杠杆 *gànggǎn* имеет значение «рычаг». Часто используется в текстах экономического характера со значением «леверидж». В данном примере мы имеем дело со словосочетанием 交易杠杆 *jiāoyì gànggǎn*, что позволяет при переводе применить конкретизацию – «финансовый левверидж».

(8) 如果 这 样, 就 会 出现 经济 过热 现象,
 Rúguǒ zhè yàng jiù huì chūxiàn jīngjì guòrè xiànxàng
 если этот вид то FUT возникать эконо- перегрев признак
 мика

同 样 需要 引起 重视,
 tóng yàng xūyào yīnqǐ zhòngshì
 такой же вид необходимо повлечь обращать большое внимание

需要 通过 政策 进行 调节⁸.
 xūyào tōngguò zhèngcè jìnxíng tiáojié.
 необходимо посредством политика осуществлять регулирование

‘В этом случае можно столкнуться с признаками перегрева экономики, которые необходимо принимать во внимание и корректировать с помощью проведения определенной политики’.

В данном примере общеупотребительное слово «перегрев» 过热 *guòrè* используется как термин «перегрев в экономике» (经济过热), то есть такая ситуация, при которой темпы экономического роста достигают предельных значений, наблюдается нехватка производственных мощностей для удовлетворения быстро растущего спроса, из-за высоких темпов инфляции правительство и ЦБ вынуждены

⁷ Caijing.com.cn [Электронный ресурс]. URL: <http://m.caijing.com.cn/article/167805> (дата обращения 11 июля 2019).

⁸ Hiwpx [Электронный ресурс]. URL: <https://www.hiwpx.com/article/40352> (дата обращения 10 июля 2019).

повышать процентную ставку. В дальнейшем устойчивый рост исчерпывает себя и экономика страны переходит в стадию стагнации.

Экономические отношения – один из главных типов общественных отношений. Все люди в современном обществе так или иначе задействованы в процессе производства, распределения и потребления товаров и услуг. Каждый человек постоянно находится в рамках определенной системы, которая сформирована экономическими отношениями в данном обществе.

Развитие экономики, расширение внутреннего рынка, появление и развитие рынка ценных бумаг, интернет-торговли и другие процессы привели к тому, что появилась потребность в СМИ экономической направленности. Подобные издания ориентированы на определенный круг лиц, специалистов в области финансов и экономики, который постоянно расширяется.

Деловые масс-медиа призваны информировать аудиторию о событиях в экономической жизни страны, освещать актуальные вопросы экономики, бизнеса, промышленности, формировать целостную картину делового мира. Деловые СМИ освещают экономическую, биржевую, финансовую, статистическую, коммерческую информацию.

В ходе анализа новостных статей экономического характера было выявлено, что в текстах масс-медиа наблюдается процесс терминологизации – перехода общеупотребительного слова в разряд термина. Терминологизация – неизбежное явление в процессе развития любого языка, и китайский язык не является исключением. Это динамический процесс, оказывающий влияние на повседневную коммуникацию носителей языка, при этом на материале современного китайского языка малоизученный. Можно с уверенностью полагать, что такое лингвистическое явление, как терминологизация, будет продолжать демонстрировать тенденцию активного развития, поэтому оно заслуживает дальнейшего тщательного изучения и анализа.

Список сокращений, используемых в глоссах

| | |
|------|-------------------------|
| ATR | атрибутивный показатель |
| DIR | директив |
| FUT | будущее время |
| LOC | локатив |
| NEG | отрицание |
| PASS | пассив |
| RES | результатив |
| PROG | прогрессив |
| PRF | перфект |

Литература

- БКРС – Большой китайско-русский словарь [Электронный ресурс]. URL: <http://bkrs.info/index.php> (дата обращения 22 июня 2020).
- Голованова 2008 – Голованова Е.И. Когнитивное терминоведение: Учеб. пособие. Челябинск: Энциклопедия, 2008. 180 с.
- Даниленко 1977 – Даниленко В.П. Лексика языка науки. Терминология: Автореф. дис. ... д-ра филол. наук. М., 1977. 40 с.
- Думитру 2008 – Думитру Э. Особенности развития и современное состояние русской геоморфологической терминологии: Автореф. дис. ... канд. филол. наук. М., 2008.
- Мистюк 2013 – Мистюк Т.Л. Особенности современной терминологизации в языке газетной публицистики // Филологические науки. Вопросы теории и практики. 2013. № 12 (30), часть 2. С. 127–130.
- Снегова 1984 – Снегова Г.П. Русская историческая терминология: Учеб. пособие. Калинин: КГУ, 1984. 86 с.
- Солганик 1981 – Солганик Г.Я. Лексика газеты (функциональный аспект). М.: Высшая школа, 1981. 112 с.
- Liang Yong 1986 – Liang Yong [梁镛]. Shuyuhua he jishu fanyi [术语化和术语翻译. 上海科技翻译]. Shanghai, 1986.

References

- Bol'shoy kitaysko-russkiy slovar'* (BKRS), [Online], available at: <http://bkrs.info/index.php> (Accessed 22 June 2020).
- Danilenko, V.P. (1977), *Leksika yazyka nauki. Terminologiya* [Vocabulary of the language of science. Terminology], Abstract of PhD dissertation, Moscow, Russia.
- Dumitru, E. (2008), *Osobennosti razvitiya i sovremennoye sostoyaniye russkoy geomorfologicheskoy terminologii* [Features of development and the current state of Russian geomorphological terminology], Abstract of Ph.D. dissertation, Moscow, Russia.
- Golovanova, E.I. (2008), *Kognitivnoye terminovedeniye: ucheb. posobiye* [Cognitive terminology: study guide], Entsiklopediya, Chelyabinsk, Russia.
- Mistyuk, T.L. (2013), "Features of modern terminologization in the language of newspaper journalism", *Filologicheskiye nauki, Voprosy teorii i praktiki*, vol. 12 (30), chast' 2, pp. 127–130.
- Snetova, G.P. (1984), *Russkaya istoricheskaya terminologiya. Uchebnoye posobiye* [Russian historical terminology. Study guide], KGU, Kalinin, Russia.
- Solganik, G.Y. (1981), *Leksika gazety (funktsional'nyy aspekt)* [Newspaper vocabulary (functional aspect)], Vysshaya shkola, Moscow, Russia.
- Liang, Yong (1986), *Shuyuhua he jishu fanyi* [Terminology and translation of terms], Shanghai, China.

Информация об авторе

Наталья Ю. Чечина, Московский государственный лингвистический университет, Москва, Россия; 119034, Россия, Москва, ул. Остоженка, д. 38; nataliachechina@yandex.ru

Information about the author

Natalya Y. Chechina, Moscow State Linguistic University, Moscow, Russia; bld. 38, Ostozhenka Str., Moscow, Russia, 119034; nataliachechina@yandex.ru

Стратегии преобразования русских фонетических заимствований в китайском языке: фонетические и графические аспекты

Кирилл И. Семенов

*Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия
kir.semenov@yandex.ru*

Аннотация. Статья посвящена рассмотрению фонетических и графических трансформаций, которые претерпевают звуковые заимствования из русского языка. Исследование включает в себя анализ имен нарицательных и имен собственных, кодифицированных в словарях и справочниках и употребляемых в Интернете. Рассмотренные данные позволяют выявить основные тенденции в адаптации согласных, а также локализовать предположительное влияние русско-китайского пиджина на современное употребление русских заимствований в путунхуа. В сфере графики выявлено существенное несоответствие норм транслитерации, предписываемых государственными СМИ КНР, и узуса в Интернете. Кроме того, обнаружена значительная специфичность наиболее частотных иероглифических N-грамм в заимствованиях по сравнению с контрольным корпусом китайских текстов. Мы рассчитываем на то, что работа будет полезна как для фундаментальных лингвистов, занимающихся языковыми контактами и фонетикой, так и для специалистов по автоматической обработке естественного языка.

Ключевые слова: китайский язык, русский язык, языковые контакты, фонетические заимствования, автоматическая обработка естественного языка, N-граммы

Для цитирования: Семенов К.И. Стратегии преобразования русских фонетических заимствований в китайском языке: фонетические и графические аспекты // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 30–63. DOI: 10.28995/2686-7249-2020-7-30-63

Adaptation strategies of Russian phonetic loanwords in Chinese. Phonetic and graphic aspects

Kirill I. Semenov

*National Research University "Higher School of Economics", Moscow, Russia
kir.semenov@yandex.ru*

Abstract. This article considers phonetic and graphic transformations of Russian loanwords in Chinese. The study comprises an analysis of both proper and common nouns, as well as both loanwords included in dictionaries and those used in the Internet. The data considered make it possible to detect the main trends in the adaptation of Russian consonants in Chinese, as well as to localize the hypothetical influence of the Russian-Chinese pidgin on current loanword adaptation in Mandarin Chinese. It is noted that there is a dramatic discrepancy between the norms of transliteration prescribed by the PRC media and the usage in the Internet. Furthermore, a significant level of specificity of the hieroglyphic N-grams in the loanwords is revealed, compared to the reference corpus of the Chinese texts. The author expects that the results of the work will be useful for specialists both in phonetic typology and in NLP.

Keywords: Chinese, Russian, language contacts, phonetic loanwords, NLP, N-grams

For citation: Semenov, K.I. (2020), "Adaptation strategies of Russian phonetic loanwords in Chinese. Phonetic and graphic aspects", *RSUH/RGGU Bulletin. "Literary Theory, Linguistics, Cultural Studies" Series*, no. 7, pp. 30-63, DOI: 10.28995/2686-7249-2020-7-30-63

1. Введение

На протяжении всей истории китайского языка (как в нормативном варианте, так и в диалектах) можно увидеть, что он испытывал влияние других языков, типологически и генеалогически очень сильно от него отстоящих. За последние две тысячи лет, согласно устоявшемуся взгляду в западной лингвистике, можно выделить несколько основных волн заимствований в китайский язык из других языковых семей. Первая – массовые заимствования буддийской терминологии при адаптации текстов пали на китайский язык (начиная с I в. н. э.) [Vervaeet 2017, p. 23]. Вторая – заимствования из языков Средней Азии в эпоху существования Великого шелкового пути. Третья – христианская лексика, пришедшая в Китай вместе с миссионерами-иезуитами в XVI в. [Vervaeet 2017, p. 26]. Четвертая – заимствования в науке и технике из быстро модернизировавшейся Японии, иногда – относящиеся к категории обратных (заимство-

ванных из китайского в японский, затем – наоборот). Это происходило во второй половине XIX в. Последняя волна заимствований, начавшаяся в период правления Дэна Сяопина и продолжающаяся по сей день, включает в себя заимствования из европейских языков, в первую очередь из английского.

Весьма разнообразны стратегии, которые используют носители китайского языка, заимствуя те или иные слова. Традиционным считается разделение заимствований в китайском на фонетические, семантические (кальки) и гибридные (различные комбинации фонетических и семантических элементов). Такую классификацию можно увидеть, например, в [Семенов 2005, с. 210] или в [Verwaet 2017]. Впрочем, западные исследователи нового поколения предлагают более подробное и исчерпывающее деление всех заимствований на 15 типов, включающее все возможные вариации с передачей смысла, звучания и графического облика заимствования [Cook 2018, р. 20]¹. Примечательно, что даже если базовой стратегией адаптации слова стала фонетическая, то семантическая интерпретация все равно будет влиять на финальный фонетический облик этого заимствования, то есть звуковой, тоновый и морфемный состав заимствования будет зависеть от семантически интерпретируемых китайских морфем [Семенов 2005, с. 216], [Nelson 2013, р. 500].

В настоящий момент большинство ученых находят более перспективной сферой анализ графических и семантических заимствований в современном китайском. Это связано с тем, что основные языки, из которых продолжают приходить заимствования в путунхуа, – английский и японский – дают по большей части именно графические заимствования или семантические кальки.

На фоне повышенного интереса к изучению вышеописанных типов заимствований в китайском языке крайне слабо затронутым (как в фундаментальном, так и в прикладном языкознании) остается исследование фонетических заимствований. Те же работы, которые посвящены этой сфере, отличаются либо отсутствием системного взгляда на весь процесс фонетической адаптации слов, либо узкой областью заимствованных слов (ограниченных либо конкретным языком-донором, либо конкретным семантическим полем).

¹ Отдельно стоит отметить феномен «буквенных слов» (lettered words, 字母词 – zìmǔ cí). Пример такого слова – 3G手机 (sāngē shǒujī) – обозначает «мобильный телефон третьего поколения», при этом G здесь – сокращение от английского generation. Большая часть данного лексического класса образует пересечение с заимствованиями из европейских языков.

Между тем, согласно [Cook 2018, p. 25], в ближайшие десятилетия фонетических заимствований будет в китайском все больше.

Стоит, тем не менее, упомянуть небольшое количество выдающихся работ, посвященных проблематике фонетических заимствований в китайском языке. Одним из наиболее значимых китайских трудов стал Словарь заимствований в китайском языке, впервые вышедший в 1985 г. – [刘正琰 (liú zhèngtán) 1985]. Основная часть вхождений в нем представляет собой именно фонетические заимствования. Из публикаций на английском языке необходимо отметить две работы. В [Lin 2008] приводится описание основных исследований последних 30 лет в западной научной литературе. Из этого обзора видно, что основной интерес у лингвистов вызывал кантонский диалект китайского, значительно отличающийся от путунхуа. Вторая работа – [Miao 2005] – представляет подробное исследование фонетических заимствований в путунхуа из трех европейских языков. Методика этой работы будет подробнее описана ниже и ляжет в основу одного из разделов настоящей статьи. Что касается работ, посвященных фонетическим заимствованиям из русского языка, то единственным достаточно систематическим описанием этого явления стала статья русских и китайских лингвистов Амурского государственного университета, выпущенная в 2016 г. [Ин, Шипановская 2016]. К сожалению, данная статья отличается малым объемом проанализированных данных (около 50 слов) и отсутствием единообразия в классификации возможных фонетических трансформаций.

Очевидно также, что фонетические заимствования – большая проблема для автоматической обработки китайского языка. Усугубляется она тем, что в китайской системе письменности нет пробелов, а предлагаемое обычными алгоритмами NLP деление на слова достаточно европоцентричное [Huang, Chen 1996, p. 4]. В отличие от семантических калек, часто уже кодифицированных и не представляющих особенных проблем для китайского словоделения, подход к обработке фонетических заимствований машинными алгоритмами развит крайне слабо. В хрестоматийном труде *Introduction to Chinese Natural Language Processing* можно встретить лишь небольшой пассаж о распознавании транслитерированных заимствований, который является частным случаем распознавания слов, не включенных в словарь (проблема OOV) – [Wong, Xu 2010, pp. 39, 67]. Эвристики, которые применяются в автоматических переводчиках, например Google.Translate или Яндекс.Переводчик, пока достаточно далеки от идеала и в обнаружении фонетического заимствования, и в его правильном обратном переводе. Более досконально автоматический анализ европейских

фонетических заимствований был исследован на материале корейского и японского языков. Несмотря на типологические и генеалогические различия между каждым из трех языков, имеется и ряд принципиальных для нашего исследования сходств между ними: общие черты наблюдаются как в фонетике (в т. ч. жесткие ограничения на структуру слога), так и в графике (иероглифическая и/или слоговая письменность). В числе работ, посвященных японскому и корейскому, можно отметить [Кoo 2015] (заимствования из английского в корейский и японский) и [Fujii, Ishikawa 2001] (пара языков: английский и японский).

Настоящая работа предлагает начать систематичное заполнение лакуны в области исследований фонетических заимствований из русского языка в путунхуа. Мы попытались проанализировать основные аспекты в последовательности фонетических и графических трансформаций, которая происходит при транслитерации русских слов в стандартном китайском языке. Методы и результаты настоящего исследования были описаны в курсовой работе, защищенной в Высшей школе экономики в 2019 г. [Семенов 2019].

2. Материал, методика и ограничения исследования

Объектом нашей работы стали фонетические заимствования из русского в путунхуа. В первую очередь это имена собственные, однако в нескольких разделах были использованы заимствования имен нарицательных, зафиксированные в Словаре заимствований в китайском языке [刘正琰 (liú zhèngtán) 1985]. Отметим, что в этом словаре есть вхождения лексем, относящиеся к разным историческим периодам в России – например, к имперскому (слова «мазурка» или «сударь») или к советскому («ГПУ»). Это позволяет нам предположить, что они могли прийти в китайский язык в разные эпохи. Однако, во-первых, в самом словаре отсутствуют пометы, касающиеся времени (или хотя бы источников) заимствования, во-вторых, стоит принимать во внимание, что время использования лексики в русском языке не обязательно соответствует времени его заимствования. Например, если какое-то слово, касающееся имперских реалий, пришло в китайский через пиджин (об этом будет рассказано в разделе 4), можно ожидать, что оно сохранило черты китайского языка или диалектов XIX в.; но если оно пришло в китайский через переводы русской литературы XIX в., это, скорее всего, произошло не раньше середины XX в. В дальнейших исследованиях было бы полезно сгруппировать вхождения словаря по

предположительному времени их появления в китайском; однако на использованных для данного исследования материалах сделать это не представляется возможным. Исключение составляет лишь подборка слов, встречающихся одновременно в словарях пиджина и в Словаре заимствований в китайском языке.

Первый этап исследования (раздел 3) посвящен анализу трансформаций, которые происходят с согласными русского языка в процессе адаптации к китайской фонетической системе. Анализ выполнен в парадигме Теории оптимальности и повторяет методику, предложенную в диссертации Жуйцинь Мяо, защищенной в Университете Стоуни-Брук (Нью-Йорк) в 2005 г. [Miao 2005]. В этой работе рассматриваются фонетические заимствования из английского, немецкого и итальянского языков. Анализ для русского языка был проведен на материале имен нарицательных из Словаря заимствований в китайском языке (387 слов).

В разделе 4, посвященном проверке гипотезы о влиянии приграничных пиджинов на современную адаптацию фонетических заимствований из русского в путунхуа, за основу был взят материал словаря сибирского пиджина, собранный Е. Перехвальской в ее книге [Перехвальская 2008]. Необходимо отметить, что в рассмотрение не брался недавно опубликованный Словарь кяхтинского пиджина – [Попова, Таката 2017]. Мы приняли это решение по причине существенных различий в системе нотации в русских работах и в данном словаре.

Следующий этап работы (раздел 5) посвящен графической адаптации фонетических заимствований из русского в путунхуа. Раздел 5.1 посвящен сравнению нормы и узуса в транслитерациях русских имен собственных в китайском сегменте Интернета. В качестве образца нормы была использована таблица для транслитерации названий на русском языке, опубликованная Синьхуа [新华 (xīnhuá) 1993]. В качестве корпуса реальных употреблений имен собственных в китайском Интернете был автоматически собран список из более чем 13 тыс. имен собственных, о которых имеется статья в китайской Википедии. Это было сделано на материале Wikidata – масштабной многоязычной базы знаний, основанной на данных из Википедии (URL: <https://www.wikidata.org/wiki/>). Данный этап работы был осуществлен при помощи языка Python.

В последнем этапе исследования (раздел 5.2) анализируются частотность иероглифических N-грамм, использованных в транслитерациях заимствований, и их сравнение с N-граммами в контрольном корпусе, составленном из нескольких художественных произведений китайских авторов конца XX в. Данный анализ был также проведен при помощи языка Python.

Необходимо отдельно отметить ограничения последних двух разделов исследования. Во-первых, все статистические исследования были проведены на основе пары «русское слово – иероглифическое заимствование», минуя фонетическую транскрипцию китайского слова. Возможно, при дальнейшем анализе этот слой окажется очень нужен в качестве медиатора между исходным русским словом и иероглифической репрезентацией. Второе ограничение – использование материалов китайской Википедии для анализа узуса носителей путунхуа. Дело в том, что на территории КНР китайская Википедия заблокирована; соответственно основную часть пользователей этой энциклопедии предположительно составляют жители Китайской Республики, Гонконга, Макао, Сингапура и диаспоры. При этом в основной части вышеперечисленных государств, во-первых, не обязательно использование предписаний по транслитерации, выдвинутых Синьхуа, во-вторых, используются другие варианты языковой нормы (гоюй) и традиционная система китайских иероглифов. Следовательно, Wikidata, пользующаяся статьями Википедии, едва ли может претендовать на звание самой большой базы знаний именованных сущностей на китайском, так как в Китае существует своя интернет-энциклопедия Baidupedia (百度百科, URL: <https://baike.baidu.com/>), объем которой превышает общий объем англоязычной, немецкоязычной, русскоязычной и китайскоязычной Википедии. Это позволяет предположить, что в Baidupedia содержится больше статей о русских именованных сущностях, чем в рассматриваемой нами базе знаний. Избрание не самого большого китайскоязычного набора данных, аудитория которого сильно смещена в сторону Тайваня, Сингапура и Гонконга, могло породить сильный перекосяк в данных, на основе которых строилось наше исследование.

Почему же в таком случае была выбрана именно Wikidata? В первую очередь из-за интерфейса по автоматической выкачке данных: при помощи сравнительно простого поискового запроса она позволяет сформировать набор данных по всем именам собственным, в оригинале имеющим русское имя, а ныне переведенным на китайский язык. Такого объема вхождений невозможно добиться ни пользуясь бумажными версиями словарей заимствований, ни параллельными русско-китайскими корпусами, ни скачав списки имен собственных, находящиеся в онлайн-словарях (так как обычно их объем исчисляется сотнями). Baidupedia же не предлагает алгоритмов поиска по своей базе знаний или возможности выкачки всей энциклопедии, поэтому на сегодняшний день остается довольствоваться базой знаний Википедии.

3. Анализ фонетических заимствований в китайском языке в парадигме Теории оптимальности

Прежде чем начать сравнительный анализ фонетических заимствований из европейских языков и из русского в китайский, отметим основные параметры различия фонетики т. н. SAE-языков и путунхуа в области консонантизма. В фонетике стандартного китайского языка отсутствуют фонемные противопоставления по звонкости-глухости и по мягкости-твердости. Вместо этого большая часть согласных обладает оппозицией по придыхательности. Кроме этого, структуру слога в китайском можно задать формулой $(C_1)V(C_2)$, где C – согласный, V – гласный или дифтонг. C_1 , в соответствии с китаистикой, принято называть инициальной, сочетание $V(C_2)$ – финалью. При этом набор согласных, возможных в позиции C_2 , значительно ограничен по сравнению с C_1 .

Обратимся теперь к исследованиям фонетических заимствований в стандартном китайском. Как уже было отмечено в обзоре литературы, наиболее общим и последовательным трудом, изучающим адаптацию фонетических заимствований в путунхуа из европейских языков, стала диссертация Жуйцинь Мяо [Miao 2005]. Автор выполнила ее в парадигме Теории оптимальности (далее – ОТ), и основным объектом ее изучения стали трансформации отдельных согласных в китайских заимствованиях, а также стратегии адаптации консонантных кластеров в путунхуа. Мы применили методы ее работы и сравнили результаты, полученные на данных русских заимствований, с результатами, которые были сделаны Мяо для английского, немецкого и (в некоторых вопросах) итальянского языков. Был воспроизведен первый раздел ее работы, мы рассмотрели возможные трансформации согласных (сгруппировав их по способу артикуляции).

Учитывая, что большое количество иероглифов имеет вариативность в прочтении (как в тоне, так и в звуковом составе), сравнивать русские и иероглифические вхождения при отсутствии эксплицитной транскрипции было бы некорректно. Поэтому в рассмотрение был взят ресурс, где транскрипция была дана эксплицитно.

Таков Словарь заимствований в китайском языке [刘正燊 (Liú zhèngtán) 1985], где каждому словарному вхождению была приписана принятая в КНР транскрипция пиньинь. Из этого словаря было взято 387 слов, в большинстве своем имен нарицатель-

ных. Необходимо отметить, что не все из них являются исконно русскими, немалая часть из них пришла в русский из языков России и бывшего СССР. Анализ слов из этого ресурса был сделан вручную.

Последнее, о чем следует напомнить перед анализом заимствований из европейских языков, – между Международным фонетическим алфавитом (МФА, IPA) и официальной китайской транскрипцией пиньинь (拼音, pīnyīn) есть существенные различия, при этом обе системы взаимозаменяемы. В данной работе в целях удобства как для специалистов по китайскому языку, так и для лингвистов-типологов всегда будет приводиться и транскрипция пиньинь (в кавычках “ ”), и транскрипция МФА (в косях скобках //).

В работе Мяо были рассмотрены все согласные сегменты, встречающиеся в английских, немецких и итальянских заимствованиях, проанализированные по группам (в зависимости от способа образования). Такое деление было избрано вследствие работы в парадигме P-map (от Perceptual mapping), подразумевающей, среди прочего, что главным качеством согласного, улавливаемым перцептивно, является его способ артикуляции [Miao 2005, p. 90].

Для каждого согласного из языков-доноров была приведена сводная таблица всех соответствующих ему согласных в китайских фонетических адаптациях, где все варианты были упорядочены по частотности (в процентном соотношении). На основе таблиц была составлена иерархия ограничений (в данном случае – faithfulness constraints), выглядящая для большей части согласных (кроме губных и плавных) таким образом:

(1) IDENT (Manner) >> IDENT (Place) >>
>> IDENT (Voice/Asp),

где первое предписывает сохранение того же способа образования, второе – сохранение того же места образования, третье – переход звонкого согласного в непридыхательный, а глухого – в придыхательный [Miao 2005, p. 82].

Мы повторили данное исследование на русском языке и проанализировали частотность всех преобразований для каждого из согласных звуков. Ниже приведена таблица (табл. 1), где представлены обобщенные результаты, показывающие все возможные трансформации каждого из согласных русского слова. Для каждого согласного реализации упорядочены по частоте их встречаемости.

Таблица 1

Иерархии реализаций консонантных сегментов в фонетических заимствованиях

| Способ артикуляции | Сегмент | Твердый, onset | Твердый, coda | Мягкий, onset | Мягкий, coda | Пример наиболее частотной реализации |
|--------------------|------------|--|--|---|---------------------------------|---|
| взрывные | /б/, /б' / | “b” /p/ > “p” /p ^h / | “b” /p/ > “p” /p ^h / | “b” /p/ | н./д. | баян > 巴扬 “bāyáng” /pajaj/ |
| | /п/, /п' / | “p” /p ^h / > “b” /p/ > “f” /f/ | “p” /p ^h / > “b” /p/ > ø | “p” /p ^h / > “b” /p/ | н./д. | пуд > 普特 “pǔtè” /p ^h ut ^h ɕ/ |
| | /к/, /к' / | “k” /k ^h / > “g” /k/ > “j” /tɕ/ | “k” /k ^h / > “g” /k/ > “j” /tɕ/ | “j” /tɕ/ > “g” /k/ > “k” /k ^h / | н./д. | кагюша > 喀秋莎 “kāqiūshā” /k ^h ɑ: tɕ: ɕʊʂɑ/ |
| взрывные* | /г/, /г' / | “g” /k/ > “j” /tɕ/ | “g” /k/ > ø | “g” /k/ > “j” /tɕ/ > “q” /tɕ ^h /, /ɕ/ | н./д. | ГПУ > 格别乌 “gébíewū” /k ^h ɣ ^h ɣé:ū/ |
| | /д/, /д' / | “d” /t/ > “t” /t ^h / > “z” /ts/ | “d” /t/ > “t” /t ^h / > ø | “d” /t/ > “j” /tɕ/ | н./д. | дума > 杜马 “dǔmǎ” /t ^h ù.mǎ:/ |
| | /т/, /т' / | “t” /t ^h / > “d” /t/ > “q” /tɕ/ | “t” /t ^h / > “d” /t/ > ø | “q” /tɕ ^h / > “t” /t ^h / > “j” /tɕ/ > “d” /t/ | “j” /tɕ/, “t” /t ^h / | точка > 托其卡 “tōqíkā” /t ^h ɔ: tɕ ^h ɪ: k ^h ɑ:/ |

Окончание табл. 1

| Способ артикуляции | Сегмент | Твердый, onset | Твердый, coda | Мягкий, onset | Мягкий, coda | Пример наиболее частотной реализации |
|--------------------|------------|---|--|-----------------------------------|---|--|
| | | | | | | |
| Шелевые | /в/, /в' / | “w” /w/ > “f” /f/ > “b” /p/, “u” /-u-, /ø | “f” /f/ | “w” /w/ > “u” /-u/, “f” /f/ | н./д. | ведро > 维得罗 “wéidélúo” /wéitǐl ^w ó:/ |
| | /ф/, /ф' / | “f” /f/ > “h” /h/ | “f” /f/ | “f” /f/ | “f” /f/ | финка > 芬卡 [帽] “fēnkǎ” /fānkǎ:/ |
| | /з/, /з' / | “zh” /ʃ/, “s” /s/ > “z” /ts/ > “sh” /ʃ ^h / /ɕ/, “ch” /tʃ ^h / | “z” /ts/ > “s” /s/ | “j” /tɕ/ > “q” /tɕ ^h / | н./д. | азарин > 阿札林 “āzhálin” /ā.ʒá.lín/ |
| | /с/, /с' / | “s” /s/ > “sh” /s/ > “x” /ɕ/ | “s” /s/ > “z” /ts/ > “zh” /ʃ/ > “sh” /ɕ/ | “x” /ɕ/ > “s” /s/ | “x” /ɕ/ > “s” /s/ | сарафан > 萨腊范 “sàláfān” /sà.là.fān/ |
| | /ж/ | “zh” /ʃ/ > “r” /ʒ/, “sh” /ɕ/, “x” /ɕ/, “ch” /tʃ ^h / | н./д. | --- | --- | крыжачок > 克雷扎卓克 “kèléizhāzhuōkè” /k ^h ɕ:ləiʒá.ʒ ^h ó.k ^h ɕ:/ |
| | /ш/, /ш' / | “sh” /ɕ/ > “s” /s/, “x” /ɕ/ | “sh” /ɕ/ > “s” /s/, “x” /ɕ/ | н./д. | “sh” /ɕ/ > “x” /ɕ/ | шапка > 沙普克 “shāpǔkè” /šā.p ^h ú.k ^h ɕ:/ |
| | /х/, /х' / | “h” /h/ | “h” /h/ | “h” /h/, “x” /ɕ/ | н./д. | хорошо > 哈拉嗦 “hàlāsuo” /xà.là.s ^w ó:/ |
| | /ц/ | “ch” /tɕ ^h / > “c” /ts ^h / > “sh” /ɕ/ | “c” /ts ^h / | --- | --- | червонец > 切尔逢涅茨 “qīè'ěrféngniècí”, /tɕ ^h è.ə.fəŋniè.ts ^h z/ |
| | /ч/ | --- | --- | --- | “q” /tɕ ^h / > “ch” /tʃ ^h / > “j” /tɕ ^h / | ЧК > 契卡 “qīkǎ” /tɕ ^h ī.kǎ:/ |

| | | | | | | |
|---------|-----------|---|---------------------------------------|--|--|---|
| носовые | /м/, /м'/ | “m” /m/ | “m” /m/ > “ng.” /ŋ/ > “n.” /n./ | “m” /m/ | “n.” /n./, “ng.n” /ŋ.n/ | МиГ > 米格 “mǐgé” /mǐ.kk:/ |
| | /н/, /н'/ | “n.” /n./ > “n.n” /n.n/ > “ng.n” /ŋ.n/ > “ng.” /ŋ./ | “n.” /n./ > “ng.” /ŋ./ | “n.” /n./ > “n.n” /n.n/ > “ng.n” /ŋ.n/ | “n.” /n./ > “ng.” /ŋ./ > “n.” /n./ | НЭП > 纳普 “nàpǔ” /nà.p'ù:/ |
| плавные | /р/, /р'/ | “r” /r/ | “r.” /ʃ./ > ʃ > “i.” /i/ > “ng.” /ŋ./ | “r” /r/ > “r.” /ʃ./ | “r” /r/ > “r.” /ʃ./ | рало > 拉罗 “lāluó” /lā.l'wó:/ |
| | /л/, /л'/ | “l” /l/ > “n” /n/ | “r.” /ʃ./ > ʃ > “l” /l/ | “l” /l/ | “l” /l/ > “r.” /ʃ./ > ʃ | лезгинка > 列兹金卡 [舞] “l'iezjĩnká” /l'è.tsʒ.tẽmk'á:/ |
| глайды | /й/ | --- | --- | “y” /j/ > ʃ > “i.” /i/ > “i.” /j/ > ʃ | “i.” /i/ > “y” /ji/ > ʃ | ералаш > 叶拉拿西 “yè.lānaxī” /jè.lā.ná.xī:/ |

Примечания

Трансформация в ноль звука обозначается символом ʃ.

Несуществующий в русском языке сегмент обозначается прочерком (---); русские сегменты, примеров которых не нашлось в выборке, обозначены «н./д.».

Символ границы слога – «.»». Соответственно обозначение “n.” необходимо трактовать как «звук n в позиции coda», а «l» – как «звук l в позиции onset».

Сегменты /д/, /д'/, /т/, /т'/ выделены в отдельную группу взрывных, так как их мягкие варианты по способу образования близки к аффрикатам.

Можно увидеть, что основная часть реализаций согласных удовлетворяет иерархии ограничений, построенной для европейских языков. Наличие дополнительной (по сравнению с европейскими языками) оппозииции по мягкости в большинстве случаев также не добавляет исключений.

Так, русское /к/ в позиции onset переходит в /k^h/ – в 95% случаев, в /к/ – в 3% случаев, а /ф/ в onset реализуется в 93% как /f/, в 7% – как /h/. Это подтверждает гипотезу Мяо о том, что в первую очередь будут выполняться ограничения способа и места, а затем сохраняться переход «звонкий > непридыхательный; глухой > придыхательный».

Тем не менее существует два типа явлений, достойные более пристального внимания. В пункте 3.1 будут рассмотрены случаи адаптации согласных, удовлетворяющие основной иерархии, но обладающие сильной вариативностью. В пункте 3.2 будет рассмотрен ряд согласных, основные реализации которых в заимствованиях противоречат предложенной иерархии.

3.1. Широкая вариативность

3.1.1. Аффрикаты и мягкие взрывные

Первая такая группа – русские аффрикаты /ц/ и /ч/. Например, /ч/ может реализовываться четырьмя аффрикатами – “q” /tʃ^h/, “ch” /tʃ^h/, “zh” /tʃ/ или “j” /tʃ/.

Кроме того, примечательна реализация аффрикат в виде щелевых в наиболее редких случаях:

(2) царь > 沙尔 “shā’ěr” /ʃā:ʔ/

Этот процесс общий для заимствований из русского и из рассмотренных Мяо европейских языков, и он указывает на то, что аффрикаты – одна из наиболее неустойчивых групп согласных с точки зрения сохранения способа образования. В принципе, это не очень удивительный факт, так как типологически известен эффект спирантизации².

Здесь было бы логично упомянуть и о двух группах согласных, в которых часто происходят трансформации, на первый взгляд нарушающие иерархию. Речь идет, во-первых, о мягких сегментах

² Kerstens J., Ruys E., Zwarts J. Spirantisation [Электронный ресурс] // Lexicon of Linguistics. 1996. URL: <https://lexicon.hum.uu.nl/?lemma=Spirantisation> (дата обращения 25 марта 2020).

/г'/ и /к'/, а во-вторых, о /д'/, /т'/. Согласные из обеих этих групп часто превращаются в аффрикаты в китайских транслитерациях. Ниже можно увидеть примеры таких трансформаций.

Трансформация мягких заднеязычных в аффрикату:

(3) кисель > 吉协力 “jíxiéli” /tɕi:ɕé:lì:/

Трансформация мягких зубных в аффрикату:

(4) «Катюша» > 喀秋莎 “kāqiūshā” /kʰā:tɕʰə̃ʃā:/

Представляется достаточно простой интерпретация такого изменения сегментов в сторону аффрикат. Мягкие заднеязычные из-за смещения места образования и долгого шума при взрыве (в силу большой площади задней части языка) перцептивно могут напоминать аффрикативные согласные. Необходимо отметить, что Мяо обращала особое внимание на сочетание «заднеязычный согласный + передний гласный» в своей работе, где она указывала на аналогичное поведение (передачу заднеязычного переднеязычной аффрикатой) в заимствованиях из европейских языков (например, нем. “Kiel” /ki:l/ > кит. 基尔 “jī’ěr” /tɕiǝ/). Было отмечено, что это следствие палатализации [Miao 2005, p. 63]. Это кажется справедливым, а в русском языке с более выраженными процессами палатализации это становится еще более очевидным. Что касается мягких зубных /д'/ и /т'/, то в современном русском их можно вполне считать аффрикатами; отнесение их на письме к их твердым взрывным аналогам можно объяснить скорее традицией и влиянием орфографии, нежели отображением реальных фонетических особенностей этих сегментов. Китайские же заимствования очень явно отображают эту особенность русских согласных.

Таким образом, явления, кажущиеся на первый взгляд исключениями из предложенной Мяо теории в зоне мягких согласных, на самом деле являются лишь ее дополнительным подтверждением.

3.1.2. Плавные

Напомним, что в китайском языке распределение плавных согласных принципиально отличается от обычного для европейских языков. Наименее маркированным является боковой /л/, который может встречаться в позиции onset. Этот звук является «базовым» плавным в китайском языке, аналогично тому, каким в японском является одноударный /л/. Существует также ретрофлексный сегмент, который в позиции onset представляется как аппроксимант /ɭ/, а в позиции coda – как эризованный шва /ɻ/.

Это, видимо, является причиной некоторой асимметрии в преобразовании плавных (как русских, так и английских/немецких/итальянских), заключающейся в том, что в позиции onset оба согласных /р/ и /л/ в абсолютном большинстве случаев преобразуются в /l/, а в позиции coda – в /ʌ/. Примечательно, что в случае с мягкими сегментами /р'/, /л'/ происходит значительное расхождение: в позиции coda примерно половина вхождений /р'/ трансформируется в /ʌ/, в то время как остальные вхождения реализуются как onset /l/ (обычно в слоге /li/). Аналогичный процесс происходит при трансформации /л'/, где как onset /l/ реализуется больше половины вхождений (55%).

Реализация /р'/ в позиции coda в /l/:

(5) сударь > 苏达利 “sūdàlì” /sū:tá:lì/

Реализация /р'/ в позиции coda в /ʌ/:

(6) богатырь > 波加的尔 “bōjiādìěr” /pō:teā:tì: ʌ/

Для описания трансформации плавных Мяо вводит дополнительные ограничения маркированности. Оба из них предлагаются ставить в вершине иерархии ограничений. Первое из них – $*\sqrt{V} \gg *lV-$ – обозначает, что аппроксимант в onset более маркирован, чем /l/ [Miao 2005, с. 86]. Это оказалось верно. Второе ограничение, предложенное автором, – $*-Vl/-Vr \rightarrow /CV/$: запрет плавной coda становиться плавным onset в заимствованном слове [Miao 2005, с. 87]. Это ограничение оказывается опровергнуто в случае с русскими заимствованиями на материале мягких согласных.

3.1.3. Носовые

Что касается употребления носовых согласных в onset, то, аналогично заимствованиям из европейских языков, вариация согласных очень мала и в основном губно-губной /м/ и зубной /н/ совпадают со своими аналогами /m/ и /n/. Примечателен в данном случае единственный момент (не затронутый в работе Мяо): немалое количество вхождений в русском языке, где /н/ встречается в позиции между гласными, преобразуется в сочетание коды /n/ или /ŋ/ и следующего за ним в позиции onset /n/, как бы «усиливая» этот согласный (см. пример 7).

(7) Семинар > 习明纳尔 “xímíngnà'ěr” /éimíngnà: ʌ/

Что же касается позиции coda, то здесь (аналогично тому, как это происходит в европейских языках) примечателен переход /м/ в допустимую в китайском назализованную финаль (/п/ или /п̃/). Впрочем, этот процесс происходит далеко не во всех случаях (27%), уступая при этом эпентетической вставке гласного после /м/.

3.2. Согласные, трансформация которых противоречит иерархии ограничений

Случаи значимого несоответствия предложенной иерархии ограничений относятся к трансформации сегментов /ж/, /з/ и /з'/. Приведем два примера:

(8) жалейка > 扎列卡[管] “zhálièkǎ” /ʃá:l̃i:è:kʰǎ:/

(9) мазурка > 玛组卡[饼] “mǎzǔkǎ” /mǎ:t̃s̃ù:kʰǎ:/

В отличие от господствующей в западных заимствованиях трансформации звука в щелевые (в большинстве случаев в англ. – в /s/, /ʃ/ или /ç/), в заимствованиях из русского они преобразуются в аффрикаты /ʃ/ (30%) и /tʃ/ (20%). Ожидаемые же трансформации данных сегментов в щелевые представлены для каждого из звуков в меньшинстве. Мы попробуем проверить связь данного несоответствия с влиянием пиджина на стратегию фонетической адаптации русских слов.

Помимо этого, сегмент /ж/, особенно в позиции onset, имеет тенденцию к реализации в виде ретрофлексного аппроксиманта “r” /ɹ/ (17% на материале Словаря заимствований). Такая тенденция не наблюдается в европейских языках, где сегмент /з/ обычно реализуется как “y” /j/ (из-за близости по месту образования). При этом можно увидеть, что в обоих случаях (и в русском, и в европейском) фрикативные согласные могут меняться на аппроксиманты того же места образования. Это можно объяснить тем, что аппроксиманты и фрикативные различаются только шириной щели [Кодзасов, Кривнова 2001, с. 280]. Кроме того, возможно, на такой способ адаптации повлияла устоявшаяся норма китайско-русской практической транскрипции Палладия: буква «ж» обозначает в транслитерации китайских слов как раз данную ретрофлексную инициаль: например, имя великого китайского писателя 郭沫若 (“Guō Mòruò”, /kʷò:mò:ʷò:/) в системе Палладия передается как «Го Можо».

4. Проверка влияния пиджина и северных диалектов на адаптацию русских заимствований в путунхуа

Перед непосредственным обсуждением влияния пиджина на китайские идиомы необходимо локализовать объект нашего исследования. В семействе северных диалектов китайского ученые выделяют группу диалектов гуаньхуа [Завьялова 1996, с. 19]. Нас будут интересовать несколько диалектов из этой группы, носители которых живут в приграничных с Россией и Монголией регионах КНР.

Фонетика диалектов гуаньхуа, во-первых, отличается от фонетики путунхуа, во-вторых, неоднородна между говорами. Особенности консонантизма можно увидеть как в наборе согласных фонем (например, неразличение верхнезубных согласных и передне-твердотвёрдых), так и в позиционном распределении инициалей и финалей (например, взаимная заменяемость [-ŋ] и [-n] и более продуктивная эризация).

В зоне распространения приграничных северокайских говоров с конца XVIII в. стал распространяться вариативный русско-кайский пиджин, которым пользовались изначально русские и кайские крестьяне и купцы. Далее этот пиджин распространился на приграничные и даже на внутренние территории России, а также стал языком межнационального общения для малых народов этого региона (маньчжуров, удэ́гэ и т. д.) с русскими – [Перехвальская 2008, с. 73, 121]. Примечательно, что, несмотря на периоды изоляции между кайским и русским населением во второй половине XX в., а также на почти полное исчезновение русской общины в Маньчжурии в эпохи японской оккупации и Культурной революции в КНР, этот пиджин продолжает существовать (хоть и сильно видоизменившись), на что указывают недавние исследования [Цзе 2007].

Нас интересует в данном случае тот факт, что большое количество как российских, так и кайских ученых отмечают в своих работах, что «транзит» русских слов через пиджин и северные кайские диалекты мог оказаться важным этапом для адаптации русских заимствований в нормативном путунхуа. С вариациями этого предположения можно ознакомиться в [Ин, Шипановская 2016, с. 145]. Тем не менее нам не удалось встретить ни одной научной работы, где было бы доказано прямое влияние пиджина и/или северных диалектов кайского языка на адаптацию русских заимствований в путунхуа. Теперь мы обладаем трактовкой ОТ для трансформаций русских согласных в путунхуа, а также сравнением этих данных с аналогичными трансформациями в европейских

языках. Значит, мы можем попробовать проверить гипотезу о влиянии пиджина или диалектов на трансформацию русских согласных в путунхуа с лучшей верифицируемостью.

Для того чтобы обоснованно предполагать связь между ранними языковыми контактами русских и китайцев и современным процессом русских заимствований в путунхуа, мы предлагаем следующий набор требований, удовлетворение каждого из которых повышает вероятность того, что заимствование пришло в путунхуа при посредничестве пиджина или северо-китайских говоров:

1. Трансформация элементов в этих словах противоречит предсказаниям ОТ для русских и европейских заимствований.

2. Трансформация элементов в этих словах объясняется фонетикой/грамматикой пиджина или северных диалектов.

3. Существуют словарные вхождения в пиджине или диалектах, которые подтверждают возможность прямого заимствования той или иной трансформации из пиджина/диалекта в путунхуа.

В качестве главного источника по грамматике и фонетике пиджина была использована монография Е.В. Перехвальской, посвященная русским пиджинам и собственно русско-китайскому контактному языку, на котором автор специализировалась [Перехвальская 2008]. В данной монографии также приводится словарь всех найденных ею лексем и словоформ со всеми возможными транскрипциями. В качестве описания фонетики северных диалектов был взят труд О.И. Завьяловой «Диалекты китайского языка», упоминавшийся выше [Завьялова 1996]. Для анализа словарных вхождений северных диалектов были использованы русскоязычные статьи о русско-китайских языковых контактах в этом регионе [Ин, Шипановская 2016], [Ма 2015].

Что касается возможности трансфера диалектных фонетических черт в стратегию адаптации в путунхуа, то оказалось, что ни одна из наиболее знаковых особенностей диалектов не проявляется в словах пиджина, диалектов и путунхуа так, чтобы удовлетворить всем вышеперечисленным требованиям. Например, слово «хорошо» согласно Словарю заимствований в китайском языке выглядит в путунхуа как 哈拉嗦 “hālāsuo” /xā:lā:s^wó/. Можно было бы заподозрить трансформацию /ш/ > /s/ как проявление диалектной особенности (отсутствие оппозиции зубных и передненёбных щелевых), но такая трансформация вполне предсказывается иерархией ограничений ОТ, рассмотренной в предыдущем разделе. Более того, примеры из английских и немецких заимствований указывают на то, что процент аналогичных чередований в заимствованиях из европейских языков значительно

больше, чем в русских заимствованиях (25% в европейских заимствованиях против 8% в русских). Необходимо также отметить, что большое количество консонантных чередований, описанных в диалектной фонетике, не встречается ни в лексемах пиджина, ни в диалектных заимствованиях, ни в фонетических заимствованиях путунхуа. Это значит, что во всех вышеописанных случаях не следует множить сущности без необходимости и вводить пиджин и диалекты в качестве посредников в процессах фонетической адаптации русских заимствований в путунхуа.

Есть, однако, две особенности, удовлетворяющие всем требованиям к «пиджинному следу», обозначенным нами. Первую из них можно увидеть на закрытой группе лексем, и связана она с процессом первичного лексического заполнения пиджина. Некоторое количество существительных (34 из 193 зарегистрированных во всем корпусе) заканчиваются квази-суффиксом, записываемым обычно как «дза»: «яйдза», «монедза» (монеты, деньги), «купедза» и т. д. [Перехвальская 2008, с. 97]. Согласно наиболее распространенной гипотезе, этот суффикс появился из китайской морфемы 子 “zǐ” /tsi/, достаточно грамматикализовавшейся в современном китайском. Согласно [Моисеев 2013], в северных диалектах в принципе распространена практика частотного употребления подобных «классификаторов», в то время как для большинства современных заимствований (из русского или из европейских языков) такая стратегия не реализуется. Самое главное здесь то, что мы можем проследить цепочку вхождений слов с таким суффиксом в словаре пиджина, в диалектных словарях и затем в Словаре заимствований в путунхуа. Такова судьба слова «купец»: через пиджинное «купедза» это слово появляется в диалекте и в путунхуа в виде сочетания 谷瘠子 “gǔbīēzǐ” /kǔ:pj̥:tsʒ/, где мы можем увидеть реализацию суффикса, исходно прикрепленного к русской псевдооснове. Таких слов (включая позднюю замену родового слова на более подходящее по смыслу) в современном китайском языке, пришедших из пиджина и/или диалектов, сейчас можно встретить пять. Таким образом, мы можем с большой долей уверенности говорить о влиянии исходного морфемного парсинга русских слов носителями китайского на восприятие этих заимствований в нормативном китайском языке. Однако класс этих слов, как показывают остальные данные Словаря заимствований, непродуктивен, поэтому это не является главным объектом для рассмотрения в нашей работе.

Другая особенность трансформации русских заимствований лежит уже в области фонетики. Как уже упоминалось, большая часть современных заимствований, содержащих звуки /ж/, /з/ и /з'/, во-

преки предсказаниям ОТ трансформирует этот звук в аффрикаты /tʃ/ или /tʃʰ/. Если же посмотреть на вхождения в словаре пиджина, окажется, что и там абсолютное большинство вхождений русского звука /з/ реализуются так же (напр., “магазин” > [maatsinə], “паровоз” > [palawotʃə]). Такая же трансформация в пиджине актуальна и для другого фрикативного /ж/. Этот переход, уникальный тем, что противоречит выработанным ограничениям ОТ, но совпадает с рядом вхождений в корпусе данных пиджина, позволяет предположить, что это и есть одна из немногих продуктивных фонетических трансформаций, которая была заимствована из пиджина в путунхуа для адаптации русских заимствований.

5. Анализ графических преобразований в заимствованиях из русского языка на основе больших данных

Как правило, каждый слог в китайском языке обладает определенным лексическим значением и соответствует определенному иероглифу, составляя таким образом единство слога, морфемы и обозначающего ее графического символа [Хаматова 2003, с. 19]. При этом широко распространены как омофоны, так и омографы. Из-за этого часты случаи написания одного слога разными иероглифами и, наоборот, множественных вариантов прочтения одного и того же иероглифа.

Отсутствие прямого перевода слогов в иероглифы и обратно подразумевает, что для максимально точного анализа следовало бы рассматривать такие данные, где были бы эксплицитно указаны и произношение адаптированного русского заимствования, и его иероглифическая запись. Однако основная часть доступных на китайском языке ресурсов, которая могла бы нам для этого пригодиться, не содержит информации о транскрипции слов. Поэтому данные, которые были проанализированы в практической части работы, содержали только русский оригинал и его иероглифическую транслитерацию.

Как уже было упомянуто в пункте 2, в нижеследующих разделах был использован ресурс Wikidata, связанный со свободной интернет-энциклопедией Википедия. Основным его минусом является его малый объем (по сравнению с Baidupedia), основными плюсами – большое количество данных (по сравнению со словарями и справочниками) и удобный интерфейс для скачивания данных.

Проект Wikidata позволяет осуществить выкачку необходимых данных в разном формате при помощи языка SPARQL на странице

<https://query.wikidata.org/>. Мы последовательно выполнили несколько запросов на этом ресурсе и сохранили данные в формате .tsv, затем объединив их в один файл .csv. Поисковые запросы собирали следующую информацию об объектах: страну (из набора: Российская Федерация, СССР, Российская империя, Беларусь, Приднестровье), тип объекта (из набора: озера, реки, горы, острова, населенные пункты, персоны), названия объекта на русском и на китайском языках.

В результате поиска было получено 13 410 объектов. Из них 81% относились к РФ, 10% – к СССР. Выборка стран была сделана с учетом того, в каких государствах на русском языке говорит большинство населения и где он является официальным. Что касается распределения по типам объектов, то 43% всех объектов покрывали населенные пункты, далее (33%) шли имена людей, еще по 10–11% были заняты реками и озерами. Эти типы объектов были выбраны, так как ожидалось, что именно они будут содержать наибольшее количество фонетических заимствований (в отличие, например, от названий организаций и экономических компаний, которые, как показал предварительный анализ, в большинстве своем калькировались семантически).

Таблица 2

Описание набора данных для графического анализа транслитераций

| Столбец | Легенда и тип переменной | Значение |
|----------|--|--|
| id_ | Индекс объекта; порядковая | Натуральное число (с нуля) |
| wiki_id | Индекс объекта в системе Wikidata; порядковая | Натуральное число (с нуля) |
| label_ru | Название объекта на русском языке; номинальная | Любая строка кириллических символов |
| label_zh | Транслитерация объекта в системе Wikidata; номинальная | Любая строка из иероглифических символов |
| xinhua | Нормативная транслитерация по Синьхуа; номинальная | Любая строка из иероглифических символов |
| country | Название страны из выборки; номинальная | Значение из набора {Russia, USSR, Russian Empire, Belarus, Transnistria} |
| type | Название типа объекта из выборки; номинальная | Значение из набора {settlement, person, river, lake, island, mountain} |

Окончание табл. 2

| Столбец | Легенда и тип переменной | Значение |
|-----------------|---|-------------------------------------|
| levenstein_abs | Абсолютное расстояние Левенштейна между значениями label_zh и xinhua; количественная | Натуральное число (с нуля) |
| levenstein_norm | Нормированное расстояние Левенштейна между значениями label_zh и xinhua; количественная | Десятичная дробь в интервале [0, 1] |
| jaccard | Коэффициент Жаккара между значениями label_zh и xinhua; количественная | Десятичная дробь в интервале [0, 1] |

Предобработка данных состояла из нескольких шагов и выполнялась над обобщенным файлом .csv при помощи программы на языке Python с использованием библиотек HanziConv, re и pandas.

Названия, содержащиеся в объектах Wikidata, были записаны разными системами иероглифов – традиционной и упрощенной. Для чистоты анализа все иероглифические вхождения были переведены в упрощенную письменность при помощи библиотеки HanziConv.

Далее, при помощи регулярных выражений, были очищены от семантических элементов русские и китайские вхождения. Так, китайские названия были очищены от так называемых родовых слов – семантических морфем, употребляемых после имен собственных (как в примере 10). Русские названия были очищены от поясняющих слов («городское поселение», «река», «вулкан» и т. д.), которые могут стоять по обе стороны от собственно названия. Также были удалены все знаки препинания (включая кавычки) и латинские и числовые элементы в названиях на русском и китайском языках. Наконец, были унифицированы последовательности имени, отчества и фамилии в русских и китайских именах людей.

- (10) 伏尔加 -河
 “fúěrjīā -hé”
 /fú:ə ətɕā: -xɿ:/
 Волга (фон.) -река (род. слово)

После этого было добавлено четыре столбца, необходимые для одного из разделов статистического исследования: нормативная транслитерация, порожденная созданным нами алгоритмом, а также три метрики близости строк, сравнивающие китайское вхождение из Wikidata и нормативную транслитерацию на китайский язык. Использованные метрики – абсолютное расстояние Левенштейна, нормированное расстояние Левенштейна и индекс Жаккара. Три метрики были высчитаны для каждого объекта при помощи библиотеки `textdistances`. Подробнее об алгоритме и метриках будет рассказано ниже.

5.1. Сравнение нормы и узуса в китайских транслитерациях

С конца XX в. в Китайской Народной Республике существуют прескрипции к транслитерации иностранных слов. Занимается этими информационное агентство Синьхуа (新华社) – официальное правительственное издание и самое большое новостное агентство в КНР. В 1982 г. оно выпустило Русско-китайский словарь транскрипций (俄汉姓名译名手册). Этот словарь включал в себя все хотя бы единожды употребленные в изданиях имена собственные, взятые из русского языка, и состоял почти из 600 страниц. В 1993 г. Синьхуа выпустило более компактную и удобную в использовании таблицу для транслитерации русских слов, которая умещалась на одном листе и позволяла находить необходимые сочетания русских букв и выбрать подходящие для этих сочетаний иероглифы. Все примечания и исключения были компактно сформулированы в шести примечаниях, идущих под таблицей; большая их часть касалась фонетических поправок, связанных как с русской (например, для транслитерации сочетаний «чн» и «чт» следовало искать сочетания «шн» и «шт»), так и с китайской фонетикой (для транслитераций «мп» и «мб» следовало искать иероглифы, соответствующие сочетаниям «нп» и «нб»).

Мы решили проверить, насколько нормативные предписания, диктуемые агентством Синьхуа, соответствуют узусу, который в нашем исследовании представлен вхождениями из Wikidata. Для этого нами был создан код на языке Python, который бы получал на вход строку с русскими символами и выводил бы иероглифическую транслитерацию этой строки. Несмотря на то что в русскоязычном сегменте Интернета уже существует автоматический транслитератор русских слов (представленный на сайте БКРС: <https://bkrs>).

info/proper_convert.php), при проверке даже на небольшой выборке он выдает регулярные ошибки, поэтому мы решили сделать собственный транслитератор, создающий максимально «ортодоксальные» транслитерации. В качестве источника он принимает таблицу предписаний Синьхуа 1993 г., переработанную в удобный для считывания компьютером вид и сохраненную в формате .csv. В наборе данных из Википедии наш алгоритм принимал на вход слово из столбца с русским именем собственным и записывал его китайскую транслитерацию в новый столбец, предназначенный для нормативной транслитерации. Отметим, что в процессе транслитерации было найдено 30 объектов, правописание которых не могло быть предсказано созданным нами алгоритмом. Это произошло из-за отсутствия предписаний для конкретных комбинаций русских букв, в основном невозможных с точки зрения русской орфографии (имена собственные на языках России, например «Чыбыда»). Все эти случаи были исключены из рассмотрения.

Для сравнения предсказанной и реальной транслитераций было использовано три метрики близости строк. Первая – абсолютное расстояние Левенштейна. Эта метрика показывает, сколько символов в строке А необходимо вставить, убрать или заменить, чтобы получить строку В. Область значений этой метрики – натуральные числа от 0 (полное совпадение строк) до бесконечности. Вторая – нормированное расстояние Левенштейна – метрика, учитывающая не только количество замененных символов, но и длину исходной строки и показывающая отношение количества исходных символов, которые необходимо было заменить, к длине исходной строки. В данном случае область значений – десятичные дроби от 0 до 1 включительно, где 0 – полное совпадение строк, 1 – полная замена исходной строки. Третья метрика – индекс Жаккара. Эта метрика не учитывает порядка символов, зато рассматривает частотное распределение символов в строке А и в строке В. Чем более соответствуют друг другу символьные наборы и их частотность в двух строках, тем ближе значение индекса к 1, чем меньше – тем ближе значение индекса к 0. При обработке данных были сделаны попытки унифицировать последовательность элементов в русских вхождениях и их китайских переводах (в первую очередь это касается личных имен – они были приведены в порядок: (имя)?-(отчество)?-фамилия). Однако где-то могли остаться строки с несовпадающим порядком морфем или слов, и индекс Жаккара позволяет сравнить такие строки и не пометить как расхождение разный порядок слов в транслитерации Синьхуа и вхождении Wikidata.

Далее мы проанализировали распределение объектов нашего набора данных в зависимости от абсолютного расстояния Левенштейна. Расстояние, равное 0, оказалось у 24% объектов, равное 1 – у 20% объектов, равное 2 – у 14% объектов и равное 3 – у 15% объектов (см. график 1). Мы также рассмотрели среднее расстояние Левенштейна, сгруппировав данные по типам объектов (см. график 2). Оказалось, что наименьшее среднее расстояние свойственно именам собственным (чуть меньше 1), в то время как географические объекты имели среднее расстояние в промежутке между 1 и 2,5, а населенные пункты – более 3. Это позволяет предположить, что личные имена – тип объектов, при транслитерации которого наиболее последовательно соблюдаются прескрипции Синьхуа. Это подтверждает распределение нормированного расстояния Левенштейна на подвыборке личных имен в датасете, где около 80% объектов имеют метрику в интервале (0,0–0,1), а 15% объектов – в интервале (0,1–0,2). Кроме того, мы рассмотрели среднее расстояние Левенштейна, сгруппировав данные по странам (график 3), в результате чего получили наименьшее среднее расстояние по объектам в РФ (около 2), а наибольшее – в СССР (около 4).

Абсолютное расстояние Левенштейна между нормой и узусом

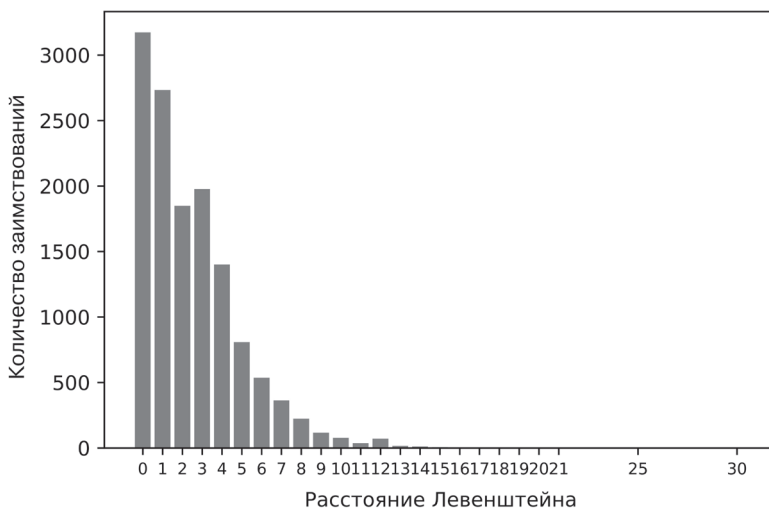


График 1. Частотное распределение абсолютного расстояния Левенштейна



График 2. Усредненное расстояние Левенштейна по типам объекта

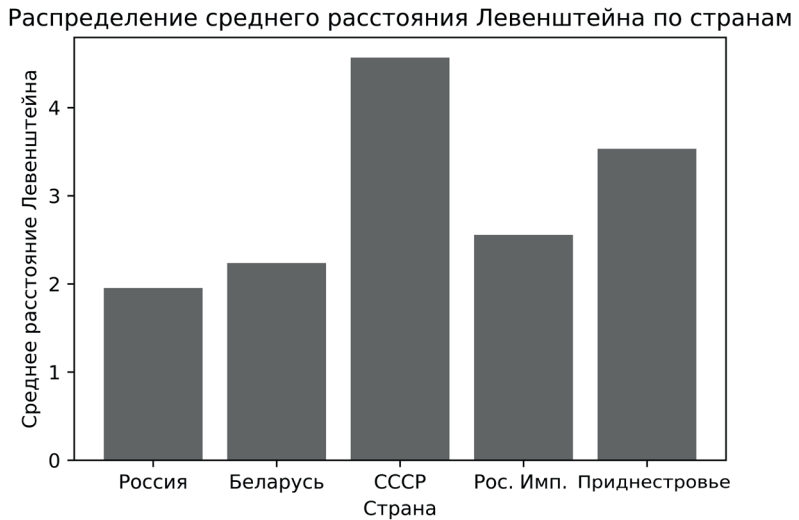


График 3. Усредненное расстояние Левенштейна по странам

Также были проанализированы нормированное расстояние Левенштейна (график 4) и индекс Жаккара (график 5) на всех данных. Гистограммы показывают, что получившиеся распределения «зеркально» отражают друг друга. При этом около 3 тыс. объектов (что составляет около 22%) показывают полное совпадение нормы и узуса по обоим метрикам, а также наблюдается два «всплеска» на уровнях 0,2 и 0,5 (нормированное расстояние Левенштейна) и 0,8 и 0,5 (индекс Жаккара).

Наконец, существует около 500 объектов, полностью не совпадающих по каждой из метрик. Ручной анализ этих вхождений показал, что основная их часть – полные семантические кальки (напр., Первомайское и «五一» “wǔyī” /ǔ:jī:/, досл. «первое мая») либо дальневосточные топонимы (напр., Владивосток и «海參崴» “hǎishēnwǎi” /xǎiṣhēnwǎi/ – «Излучина трепанга»).

Можно сделать вывод о том, что соответствие между предписаниями Синьхуа и реальными данными далеко не полное. Это не позволяет нам, например, автоматически сгенерировать еще несколько десятков тысяч китайских транслитераций на основе всего набора собственных имен, имеющих в русском разделе Википедии.

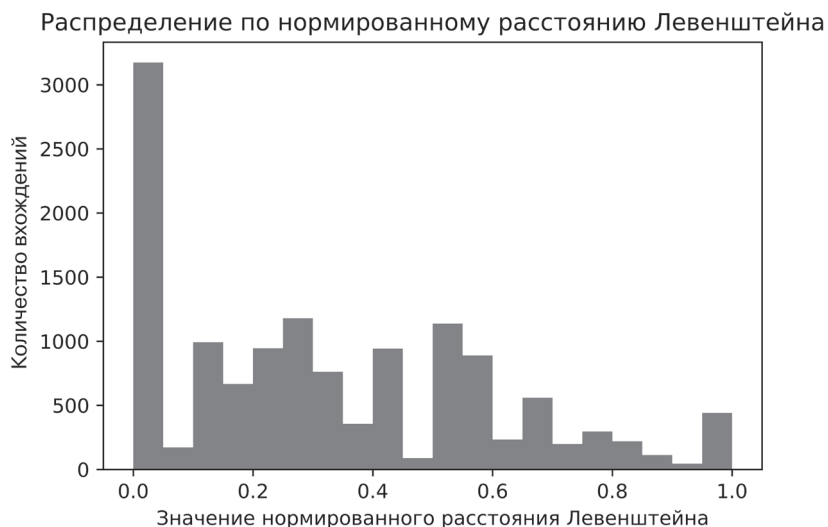


График 4. Частотное распределение по нормированному расстоянию Левенштейна

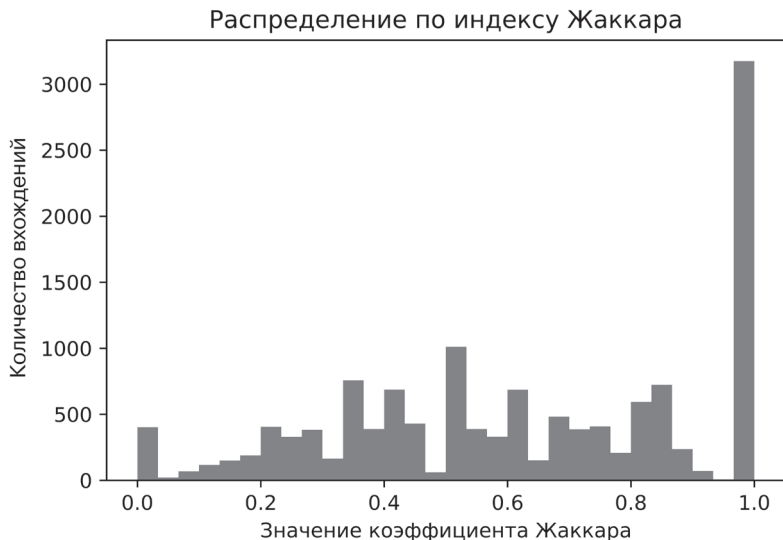


График 5. Частотное распределение по индексу Жаккара

Зависимость между расстоянием Левенштейна и дихотомией «личное имя VS географический объект» позволяет предположить, что в географических объектах, даже при использовании в целом фонетической транслитерации, части названий (прилагательные или морфемы в сложных словах) переводятся на китайский семантически мотивированными морфемами. Эта гипотеза выливается в достаточно обширное исследование, которое также было нами проведено, однако, чтобы не слишком отклоняться от проблематики данной статьи, промежуточные итоги этого исследования здесь приводиться не будут. Покажем лишь один пример такой частично семантической адаптации.

- | | |
|------------|------------------------|
| (11) 下- | 诺夫哥罗德 |
| “xià-“ | “nuòfūgēluódé” |
| /ɕà-/ | /nʷò:fʷū:kɣ̃:lʷó:tɣ̃:/ |
| низ (сем.) | Новгород (фон.) |

5.2. Анализ N-грамм в фонетических заимствованиях и в контрольных китайских текстах

В предыдущем разделе мы обнаружили, что нормативные предписания не описывают весь узус, представленный в Интернете. Это подтверждает и следующее наблюдение: было подсчитано символическое разнообразие иероглифов в списке всех переводов из Wikidata и в списке всех транслитераций по Синьхуа. Первых оказалось 877, вторых – 260. Можно ли в таком случае извлечь пользу из такого многообразия иероглифов, не ущемляющегося в рамки нормы?

Мы попробуем ответить на этот вопрос при помощи анализа частотности иероглифических (символьных) N-грамм. Для этого мы возьмем 1 тыс. самых частотных N-грамм, встречающихся в китайских заимствованиях из Wikidata, и проверим по контрольному корпусу современной китайской литературы, насколько часто эти N-граммы встречаются в аутентичных китайских текстах. Мы сделали сравнение при $N = 2$ (биграммы) и при $N = 3$ (триграммы). В качестве контрольного корпуса мы взяли художественные произведения трех современных китайских авторов (Мо Янь, Лю Чжэньюнь и Юй Хуа) и создали корпус размером 860 тыс. символов.

Анализ тысячи наиболее частотных биграмм в заимствованиях показал, что 936 из этих биграмм вообще не встречаются в контрольном корпусе, а большая часть оставшихся биграмм имеет ранг выше, чем 10 тыс. Что касается тысячи наиболее частотных триграмм, то уже 988 из них не встречаются в контрольном корпусе вовсе, а все остальные (за исключением одной) имеют ранг выше 400 тыс.

Это позволяет сделать вывод о том, что даже простые модели, построенные на символических N-граммах, должны будут показать хороший результат при дальнейшей реализации прикладных задач, таких как вычленение заимствованных именованных сущностей.

6. Перспективы исследования

Результаты, полученные в данном исследовании, преждевременно считать окончательными. Например, в разделах 3 и 4 был проигнорирован так называемый графический фактор, о котором не раз упоминала Мяо [Miao 2005, pp. 46, 56, 79], но который с трудом поддается исчислению и воспроизводимости, а следовательно, и верификации. Речь идет о том, что в пиньинь привычные для европейцев символы согласных обозначают звуки с совершенно другой артикуляционной характеристикой. Например, “z”, в боль-

шинстве европейских письменностей обозначающий щелевой, в пиньинь помечает аффрикату /ʈʂ/. При сегодняшнем культурном доминировании английского языка и латинской письменности нельзя отвести такое предположение: носители китайского, размышляя о возможной транскрипции того или иного русского имени собственного, обращают внимание в первую очередь на английскую транслитерацию русского слова, а затем на соответствие английских букв символам пиньинь. Тогда объяснение аффрикаты в транслитерации, например, названия города Рязань становится более прозаичным, чем через влияние пиджина (см. пример 12). Контраргументом здесь может послужить существование системы прямой китайско-русской транскрипции Палладия, которая более точно передает различия аффрикат и щелевых, обозначая рассмотренный выше звук как «цз»; однако эта система более известна именно в России и нацелена на адаптацию слов в обратном направлении — из китайского в русский.

(12) Рязань > (гипотеза) англ. Ryazan >

кит. 梁赞 “liángzàn” /lʰɑ̃ŋtsən/

Кроме этого, в первых разделах не было рассмотрено ни адаптации консонантных кластеров, ни адаптации гласных в заимствованиях в путунхуа. Особый интерес в этом разрезе представляет не раз высказанное предположение о том, что ударение в словах языка-донора влияет на тоновое маркирование в полученных словах (одну из вариаций этой гипотезы можно найти в [Ин, Шипановская 2016]). Дальнейшие исследования должны будут завершить описание всех сегментов и супraseгментных единиц русских заимствований в китайском.

Очевидно, найденные фонетические и (в особенности) графические закономерности вдохновляют и на прикладные исследования в сфере автоматической обработки китайского языка. Наиболее важной задачей в этой области видится создание алгоритма, позволяющего выделять в китайских текстах именованные сущности, являющиеся фонетическими заимствованиями из русского языка (возможно, с дальнейшим их обратным переводом на русский). Учитывая неспособность норм Синьхуа предсказать реально используемые фонетические заимствования, следует использовать статистические модели порождения и распознавания заимствований. Для этих целей могут быть использованы как традиционные статистические подходы — например, модель скрытых цепей Маркова, — так и нейросетевые модели. Важным аспектом при составлении работающих моделей для китайского языка станет предва-

рительное ознакомление с уже существующими подходами к выделению именованных сущностей в изолирующих языках, таких как индонезийский [Wibawa, Purwarianti 2016], [Gunawan, Suhartono, Purnomo, Ongk 2018].

В дальнейшем практическом исследовании необходимо будет учитывать и вхождения со вспомогательными (такие, как кавычки) и некитайскими (такие, как «буквенные слова» всех видов и вообще использование латиницы) символами, так как в Интернете именно такие обозначения нередко используются для передачи фонетических заимствований именованных сущностей из других языков. Один из показательных примеров – страница в китайской Википедии, посвященная российскому рэперу LOC-Dog:

- | | |
|--------------|---------------------|
| (13) LOC- | 狗 |
| /lok/ | “gǒu” /kǒu/ |
| LOC (граф.)- | собака (калька Dog) |

Алгоритмы, полученные в результате дальнейшего практического исследования, можно будет оценить на данных русско-китайского параллельного корпуса НКРЯ, в котором содержатся всевозможные виды русских заимствований (названия населенных пунктов, ручьев, уменьшительно-ласкательные русские имена и т. д.).

Благодарности

Автор искренне благодарит Б.В. Орехова (НИУ ВШЭ) за помощь в выработке методологии и О.В. Дерезу (НИУ ВШЭ) за помощь в формулировке работы на начальных этапах. Кроме того, автор выражает признательность Д.А. Скоринкину (НИУ ВШЭ) за консультации в области поисковых запросов на языке SPARQL и Л.С. Холкиной (РГГУ) за консультации по поводу прескриптивных норм Синьхуа.

Acknowledgements

The author is sincerely grateful to B.V. Orekhov (HSE University) for help in developing the methodology and to O.V. Dereza (HSE University) for help in formulating the work at the initial stages. In addition, the author expresses his gratitude to D.A. Skorinkin (HSE University) for consultations in the field of search queries in the SPARQL language and L.S. Kholkina (RSUH) for consultations on the Xinhua prescriptive norms.

Литература

- Завьялова 1996 – *Завьялова О.И.* Диалекты китайского языка. М.: Научная книга, 1996. 207 с.
- Ин, Шипановская 2016 – *Ин Ц., Шипановская Л.М.* Русские заимствования в китайском языке как результат языковых контактов // *Филологические науки. Вопросы теории и практики.* 2016. Т. № 7 (61): В 3 ч. Ч. 1. С. 144–152.
- Кодзасов, Кривнова 2001 – *Кодзасов С.В., Кривнова О.Ф.* Общая фонетика. М.: РГГУ, 2001. 590 с.
- Ма 2015 – *Ма Н.* Харбинский русско-китайский пиджин первой половины XX века и его влияние на русский и китайский языки и культуры // *Филологические науки. Вопросы теории и практики.* 2015. Т. 3. № 8 (50). С. 122–127.
- Перехвальская 2008 – *Перехвальская Е.В.* Русские пиджины. СПб.: Алетейя, 2008. 363 с.
- Семенас 2005 – *Семенас А.Л.* Лексика китайского языка. Базовый учебник. М.: АСТ: Восток–Запад, 2005. 310 с.
- Семенов 2019 – *Семенов К.И.* Стратегии преобразования иноязычных фонетических сочетаний в китайском языке (на материале заимствований из русского языка): Курсовая работа студента 2 курса бакалавриата. М.: НИУ ВШЭ, 2019. 55 с.
- Попова, Таката 2017 – *Словари кяхтинского пиджина / Под ред. И.Ф. Попова, Т. Таката.* М.: Наука, 2017. 603 с.
- Хаматова 2003 – *Хаматова А.А.* Словообразование современного китайского языка. М.: Муравей, 2003. 224 с.
- Цзе 2007 – *Цзе Я.* Забайкальско-маньчжурский препиджин: опыт социологического исследования // *Вопросы языкознания.* 2007. № 2. С. 67–74.
- Cook 2018 – *Cook A.* A typology of lexical borrowing in Modern Standard Chinese // *Lingua Sinica.* 2018. Т. 4, no. 1. P. 6.
- Fujii, Ishikawa 2001 – *Fujii A., Ishikawa T.* Japanese/English Cross-Language Information Retrieval: Exploration of Query Translation and Transliteration // *Computers and the Humanities.* 2001. Vol. 35, no. 4. P. 389–420.
- Gunawan, Suhartono, Purnomo, Ongk 2018 – *Gunawan W., Suhartono D., Purnomo F., Ongk A.* Named-Entity Recognition for Indonesian Language using Bidirectional LSTM-CNNs // *Procedia Computer Science.* 2018. Vol. 135. P. 425–432.
- Huang, Chen 1996 – *Huang C.-R., Chen K.-J.* Issues and topics in Chinese natural language processing. Taipei: Chinese University Press, 1996. 22 p.
- Koo 2015 – *Koo H.* An unsupervised method for identifying loanwords in Korean // *Language Resources and Evaluation.* 2015. Vol. 49, no. 2. P. 355–373.
- Lin 2008 – *Lin Y.-H.* Variable vowel adaptation in Standard Mandarin loanwords // *Journal of East Asian Linguistics.* 2008. Vol. 17, no. 4. P. 363–380.
- Miao 2005 – *Miao R.* Loanword Adaptation in Mandarin Chinese: Perceptual, Phonological and Sociolinguistic Factors. Ph.D. Thesis. Stony Brook University, Stony Brook, NY, 2005. 184 p.
- Nelson 2013 – *Nelson L.* English Loan Words in Mandarin Chinese: Phonology vs. Semantics // *Proc. of the National Conf. on Undergraduate Research (NCUR) 2013, University of Wisconsin La Crosse, WI April 11–13, 2013.* Madison, 2013. P. 497–505.

- Vervaeet 2017 – *Vervaeet R.* English Loanwords in the Chinese Lexicon. MA Thesis. Gent: Universiteit Gent, 2017. 128 c.
- Wibawa, Purwarianti 2016 – *Wibawa A., Purwarianti A.* Indonesian Named-entity Recognition for 15 Classes Using Ensemble Supervised Learning // 5th Workshop on Spoken Language Technologies for Under-resourced languages, SLTU, 9–12 May 2016. Yogyakarta, 2016. P. 221–228.
- Wong, Xu 2010 – *Wong K.-F., Xu R.* Introduction to Chinese Natural Language Processing. San Rafael, CA: Morgan & Claypool Publ., 2010. 148 p.
- 刘正琰 (liú zhèngtán) 1985 – 汉语外来词词典 (hànyǔ wàiláicí cídiǎn) / ed. 刘正琰 (liú zhèngtán). 上海 (shànghǎi): 上海辞书出版社 (shànghǎi císhū chūbǎnshè), 1985. 422 p.
- 新华 (xīnhuá) 1993 – 世界人名翻译大辞典 (shìjiè rénmíng fānyì dà cídiǎn), 新华社译名室编 (xīnhuá shè yì míng shì biān), 1993. 56 p.

References

- Cook, A. (2018), “A typology of lexical borrowing in Modern Standard Chinese”, *Lingua Sinica*, vol. 4, no. 1, p. 6.
- Fujii, A. and Ishikawa, T. (2001), “Japanese/English Cross-Language Information Retrieval: Exploration of Query Translation and Transliteration”, *Computers and the Humanities*, vol. 35, pp. 389–420.
- Gunawan, W., Suhartono, D., Purnomo, F. and Ongk, A. (2018), “Named-Entity Recognition for Indonesian Language using Bidirectional LSTM-CNNs”, *Procedia Computer Science*, vol. 135, pp. 425–432.
- Huang, C.-R. and Chen, K.-J. (1996), *Issues and Topics in Chinese Natural Language Processing*, Chinese University Press, Taipei, Taiwan.
- Jie, Y. (2007), “Baikal-Manchurian pre-pidgin: a sociological study”, *Voprosy yazykoznaniiya*, vol. 2, pp. 67–74.
- Khamatova, A.A. (2003), *Slovoobrazovaniye sovremennogo kitayskogo yazyka* [Word formation in Modern Chinese], Muravei, Moscow, Russia.
- Kodzazov, S.V. and Krivnova, O.F. (2001), *Obshchaya fonetika* [General phonetics], RGGU, Moscow, Russia.
- Koo, H. (2015), “An unsupervised method for identifying loanwords in Korean”, *Language Resources and Evaluation*, vol. 49, no. 2, pp. 355–373.
- Lin, Y.-H. (2008), “Variable vowel adaptation in Standard Mandarin loanwords”, *Journal of East Asian Linguistics*, vol. 17, no. 4, pp. 363–380.
- Liu, Zh.-T. (1985), *Hanyu wailaici cidian* [Chinese Loanword Dictionary], Shanghai cishu chubanshe, Shanghai, China.
- Ma, N. (2015), “Harbin Russian-Chinese pidgin in the first half of 20th century and its influence on Russian and Chinese languages and culture”, *Filologicheskiiye nauki. Voprosy teorii i praktiki*, vol. 3, no. 8 (50), pp. 122–127.
- Miao, R. (2005), *Loanword Adaptation in Mandarin Chinese: Perceptual, Phonological and Sociolinguistic Factors*, Ph.D. Thesis, Stony Brook University, Stony Brook, NY, USA.
- Nelson, L. (2013), “English Loanwords in Mandarin Chinese: Phonology vs. Semantics”, *Proc. of the National Conf. on Undergraduate Research (NCUR) 2013*, University of Wisconsin La Crosse, WI April 11–13, 2013, Madison, pp. 497–505.

- Perekhval'skaya, E.V. (2008), *Russkiye pidzhiny* [Russian pidgins], Aleteiya, Saint-Petersburg, Russia.
- Wibawa, A. and Purwarianti, A. (2016), "Indonesian Named-entity Recognition for 15 Classes Using Ensemble Supervised Learning", *5th Workshop on Spoken Language Technologies for Under-resourced languages, SLTU*, 9–12 May, 2016, Yogyakarta, pp. 221–228.
- Semenas, A.L. (2005), *Leksika kitayskogo yazyka. Bazovyy uchebnyk* [The Chinese lexicon. Basic Textbook], AST: Vostok-Zapad, Moscow, Russia.
- Semenov, K.I. (2019), *Strategii preobrazovaniya inoyazychnykh foneticheskikh sochetaniy v kitayskom yazyke (na materiale zaimstvovaniy iz russkogo yazyka)* [Adaptation Strategies for Foreign Phonetic Borrowings in Chinese (Based on Russian Loanwords)], 2nd year BA term paper, NRU "Higher School of Economics", Moscow, Russia.
- Popova, I.F. and Takata, T. (2017), *Slovari kyakhtinskogo pidzhina* [Dictionaries of Kyakhta pidgin], Nauka, Moscow, Russia.
- Vervaet, R. (2017), *English Loanwords in the Chinese Lexicon*, MA Thesis, Universiteit Gent, Gent, Belgium.
- Wong, K.-F. and Xu, R. (2010), *Introduction to Chinese Natural Language Processing*, Morgan & Claypool Publ., San Rafael, CA, USA.
- Xinhua (1993), *Shijie renming fanyi da cidian* [Translation dictionary for world personal names], Xinhua she yiming shi bian, Beijing, China.
- Ying, J. and Shipanovskaya, L.M. (2016), "Russian loanwords in Chinese as a result of language contacts", *Filologicheskiye nauki. Voprosy teorii i praktiki*, no. 7 (61), pt. 1, pp. 144–152.
- Zav'yalova, O.I. (1996), *Dialekty kitayskogo yazyka* [Dialects of Chinese], Nauchnaya kniga, Moscow, Russia.

Информация об авторе

Кирилл И. Семенов, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия; 101000, Россия, Москва, ул. Мясницкая, 20; kir.semenow@yandex.ru

Information about the author

Kirill I. Semenov, HSE University, Moscow, Russia; bld. 20, Myasnit-skaya Str., Moscow, Russia, 101000; kir.semenow@yandex.ru

«Латинизация» китайского языка:
на примере китаеязычной
интернет-коммуникации

Александр Н. Сбоев

*Дальневосточный федеральный университет
Владивосток, Россия, sboevalexander@mail.ru*

Аннотация. Процесс латинизации китайского языка можно проследить с момента появления алфавита пиньинь, который был утвержден 61 год назад и функционирует до сих пор, однако область его применения до недавнего времени ограничивалась лишь сферой образования, где он выступал как вспомогательный инструмент при изучении иероглифов. Тем не менее сейчас слова, записанные буквами, появляются все чаще в бытовой коммуникации, а основным источником таких слов является интернет-коммуникация. Стоит отметить, что часть буквенных слов, функционирующих в китаеязычной интернет-коммуникации, заимствованы, но в данной статье в фокусе исследования находятся именно китайские слова, которые стали записываться полностью с использованием алфавита пиньинь. При такой записи не используются диакритические знаки для обозначения тона. Возможно ли, что буквенная запись с использованием алфавита пиньинь будет вытеснять традиционную иероглифическую запись? В настоящий момент такое утверждение кажется абсурдным, но происходящие в современном китайском языке процессы все же заставляют задуматься над этим.

Ключевые слова: китайский язык, интернет-коммуникация, буквенные слова, латинизация, алфавит

Для цитирования: Сбоев А.Н. «Латинизация» китайского языка: на примере китаеязычной интернет-коммуникации // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 64–83. DOI: 10.28995/2686-7249-2020-7-64-83

Chinese language “latinization”.
(On the basis of Chinese Internet communication)

Aleksandr N. Sboev

Far Eastern Federal University, Vladivostok, Russia, sboevalexander@mail.ru

Abstract. The process of latinization of the Chinese language goes back to 1958, when Pinyin alphabet was created. The usage of Pinyin alphabet for a long time was limited only to the sphere of education, where the alphabet acted as an “extra tool” for studying the hieroglyphic script. Nowadays the usage of Pinyin alphabet has become wider: Chinese words in dictionaries are listed in alphabetical order; the alphabet is used for typing hieroglyphs in word processors; words written in latin letters are widely used in Chinese Internet communication, etc. While some of such “literal” words currently used in Chinese Internet communication are borrowings from other languages (mostly from English), this article is focused on wholly Chinese words that are written only in Pinyin alphabet without tone diacritics. Is it possible that writing in Pinyin alphabet will replace the traditional hieroglyphic writing? Today it seems absurd, but time will show.

Keywords: Chinese language, internet communication, letter words, pinyin, latinization, alphabet

For citation: Sboev, A.N. (2020), “Chinese language “latinization”. (On the basis of Chinese Internet communication)”, *RSUH/RGGU Bulletin. RSUH/RGGU Bulletin. “Literary Theory. Linguistics. Cultural Studies” Series*, no. 7, pp. 64-83, DOI: 10.28995/2686-7249-2020-7-64-83

Введение

Проекты создания алфавита китайского языка начались еще до образования Китайской Республики. Многочисленные алфавитные системы для китайского языка разного назначения стали появляться со второй половины XIX в. [Завьялова 2014, с. 250], созданы они были на латинской графике, арабских цифрах, произвольных знаках, китайской национальной графике [Софронов 2007, с. 611]. Лу Чжуанчжан является автором азбуки 切音新字 (цеинь синьцзы) и учебника «Наглядная азбука» (имуляожань чуцзе 一目了然初阶 1892 г.), где для записи китайских слов он использовал латиницу [Саськова 2016], в первоначальном виде учебник был предназначен для диалектов Амоя (Сямэня), Чжанчжоу, Цюаньчжоу [Софронов 2007, с. 611].

С наступлением XX в. иероглифы все чаще воспринимались как препятствие на пути модернизации Китая в области науки, культуры и политики [Zhou 2014, p. 82]. В 1918 г. была опубликована

официальная алфавитная система 注音字母 чжунинь цзыму. Графика этого алфавита основана на японской азбуке катакана, в нем также были использованы простые иероглифы или части сложных иероглифических знаков [Завьялова 2014, с. 259]. В 1928 г. был опубликован «романизированный алфавит для государственного языка» 国语罗马字 гоюй ломацзы, ставший первой официальной китайской алфавитной системой на латинском языке. Однако, как отмечает О.И. Завьялова: «Он никогда широко не употреблялся, его можно обнаружить главным образом в словарях, основная причина – сложное обозначение тонов через разное написание одного и того же слова без использования диакритических знаков или цифр» [Завьялова 2014, с. 261]. Этот алфавит использовался в методике преподавания китайского языка как иностранного вне Китая [Софронов 2007, с. 620]. В конце 1920-х гг. в СССР приступили к созданию нового алфавита для китайского языка на латинской графике, он был рекомендован проживавшим в России китайцам: «...алфавит получил название латинизированного алфавита китайского языка – латин хуа цзыму, его официальным назначением было обучение письму неграмотных взрослых» [Софронов 2007, с. 621].

В 1930 г. был разработан новый вариант чжунинь цзыму, который был ориентирован на пекинское произношение и получил название 注音符号 [Завьялова 2014, с. 40–41]. Однако, как отмечает Г.П. Сердюченко: «Ни одна попытка распространения в старом Китае фонетической письменности не получила государственной поддержки» [Сердюченко 1959, с. 28].

Председатель комитета по реформе китайской письменности У Юйчжан, видевший, как советскими филологами создавалась транскрипционная система китайского языка на раннем этапе, говорил, что «традиционная китайская письменность является продуктом феодализма, она стала инструментом угнетения пролетариата и камнем преткновения для массового образования. Традиционная китайская письменность не подходит для современной эпохи. Китай должен заменить свою устаревшую систему символов на пиньинь» [The Progress of Pinyin 2008].

Академик В.М. Алексеев в предисловии работы «Китайская иероглифическая письменность и ее латинизация» пишет, что «доводы за латиницу и против иероглифики имеют главным образом политическую установку» [Алексеев 1932, с. 20–21]. Далее В.М. Алексеев указывает, что существовали пять позиций в отношении латинизации китайского языка: 1. Радикальная реформа китайской письменности. Сторонники данной позиции исходят из положения, что в XX в. нет места идеографическому (иероглифическому) письму. 2. Защита китайского языка, но не иероглифики,

которая также может подлежать уничтожению и замене фонетическим (алфавитным) письмом. 3. Против замены китайского языка и письменности. 4. Применение к иероглифике алфавитной нотации для того, чтобы обучить школьников единому для Китая чтению иероглифических знаков, а в дальнейшем полностью перейти на алфавитную запись. 5. Консервативная позиция, не выдвигающая никаких новых предложений, но стоящая в оппозиции ко всем другим [Алексеев 1932, с. 20–21].

В 1958 г. V сессией ВСНП первого созыва одобрен проект китайского алфавита, который был назван 拼音字母 пиньинь цзыму – «фонетический алфавит». Как отмечает В.М. Софронов: «Алфавит пиньинь цзыму является прямым наследником латинизированного алфавита, разработанного в начале 30-х годов в Советском Союзе» [Софронов 2007, с. 627]. Тем не менее «самой распространенной вплоть до 1970-х гг. системой записи китайских слов в словарях и текстах на английском языке была система Уэйда–Джайлса» [Завьялова 2014, с. 251].

В период разработки алфавита пиньинь были те, кто поддерживали его, а также те, кто выступали против. «Сторонники латиницы исходили, прежде всего, из того, что иероглифы являются пережитком прошлого, они отрывают Китай от мирового развития, к тому же иероглифическая письменность сложна для изучения. Противники же, наоборот, были против разрушений тысячелетней истории» [Алпатов 2013, с. 305]. Сторонники иероглифической письменности отмечали, что латинизация китайского языка приведет к распаду языка, выделению диалектов в отдельные языки, и указывали на существенную роль иероглифов при снятии вопросов омонимии, однако, что касается последнего пункта, то В.М. Алексеев считал, что этот вопрос в большинстве случаев снимается контекстом [Алексеев 1932, с. 49].

Мао Цзэдун в своих публичных заявлениях говорил, что пиньинь будет играть лишь второстепенную роль и что основной упор будет сделан на упрощение символов. Однако в более поздних замечаниях, сделанных для ограниченной аудитории, он подчеркнул, что это всего лишь временная уступка реальности, и в 1957 г. на заседании ЦК КПК Мао вновь заявил о своей поддержке алфавитного письма [DeFrancis 2006, p. 5].

С 1970-х гг. для записи китайских слов повсеместно используется система пиньинь. С 1977 г. пиньинь применяется в качестве международной стандартной системы записи китайских имен собственных и географических названий, а с 1982 г. пиньинь утвержден в качестве нормы записи латинскими буквами любых китайских слов. С 1 января 2009 г. пиньинь официально введена на Тай-

ване. Во введенном 1 января 2001 г. Законе о языке и письменности пиньинь определяется как единый стандарт алфавитной записи и транскрипции слов путунхуа (введенный в 1955 г. официальный китайский язык, основанный на северных диалектах, в котором пекинское произношение взято за нормативное), который применяется в тех случаях, когда использование иероглифов является неудобным или невозможным [Завьялова 2014, с. 266].

В.М. Софронов полагает, что «сфера практического применения алфавита пиньинь цзыму остается крайне узкой» [Софронов 2007, с. 634], тем не менее, как считает М.В. Гурин, «переход на алфавит остается (теоретически) дальней стратегической целью языковой политики КНР» [Гурин 2011, с. 88].

Как считает Чжоу Югуан: «Фонетический алфавит пиньинь цзыму достиг только статуса “транскрипции”, но не официальной “алфавитной письменности”» [Чжоу Югуан 1989, с. 388]. Здесь следует отметить, что термин «транскрипция» в отношении пиньинь употребляется не в значении «фонетическая транскрипция», и на наш взгляд, правильнее было бы назвать пиньинь фонетическим алфавитом, так как отдельная буква данного алфавита может обозначать не только одну фонему. Как отмечал Н.А. Спешнев: «Китайский алфавит не может служить эквивалентом фонетической транскрипции, что особенно наглядно выявляется в тех случаях, когда речь идет об описании звуков их комбинаторных вариантов, ассимиляции, эризации» [Спешнев 1980, с. 6]. Об этом также писала Хилари Чапелл: «Недостатком пиньинь является то, что он не раскрывает правил произнесения той или иной буквы, в связи с чем у англоговорящих могут возникнуть трудности с правильным произнесением звуков, выраженных буквами x, sh, q» [Chappell 1980, p. 111]. Таким образом, пиньинь не является фонетической транскрипцией в полной мере для иностранцев» [Chappell 1980, p. 115].

А.Н. Алексахин называет пиньинь «звуко-буквенным стандартом слова» и, соглашаясь с китайскими лингвистами, полагает, что преимущество данного вида записи слов в экономии: вместо 10 тыс. иероглифов применяются всего 26 букв латинского алфавита [Алексахин 2010, с. 20]. Причем только четыре звука обозначаются двумя буквами – zh, ch, sh, ng [Chappell 1980, p. 112].

Как считают китайские лингвисты Чэн Таоань и Чэн Чжиюн:

С 80-х гг. прошлого века в Китае отказались от плана латинизации китайского языка, а «Проект фонетического алфавита китайского языка» является лишь системой знаков, обозначающих чтение иероглифов, эту систему не стоит рассматривать как самостоятельное письмо [Чэн Таоань 2012].

Мао Гун отмечал, что пиньинь является лишь вспомогательным средством и используется как некий «инструмент»; таким образом, он утверждал, что в Китае нет идеи полной замены иероглифов латинскими буквами [Чэн Таоань 2012].

Чжан Фэн и Ма Сюэдань вообще говорят о том, что «пиньинизация китайского языка непременно приведет к утрате интереса к китайскому языку» [Чжан Фэн 2010, с. 728].

Су Пэйчэн, выделяя положительные черты записи пиньинь, такие как простая структура, упорядоченность, небольшое количество букв, известность чтения слова при первом же взгляде на него, в то же время отмечает, что невозможно изменить китайские иероглифы на текст пиньинь, основываясь только на этих пунктах. Потому что у китайских иероглифов есть еще и много других преимуществ [Су Пэйчэн 2018, с. 100].

Отдельно стоит пояснить ситуацию на Тайване, где широко употреблялся алфавит чжуинь цзыму, а «официальным алфавитом на латинской основе на Тайване вплоть до последнего времени считался также гоюй ломацзы, хотя на практике обычно использовалась старая английская транскрипция Уэйда–Джайлса. Эта транскрипция до сих пор сохраняется в латинском написании личных имен не только многих тайваньцев, но также других китайцев, живущих за пределами континентального Китая» [Завьялова 2014, с. 251]. В 2002 г. была официально введена новая латинизированная система тунъюн пиньинь, с 1 января 2009 г. Тайвань перешел на пиньинь [Завьялова 2014, с. 53], которая считается единственной официальной системой [Завьялова 2014, с. 265].

Для набора иероглифов на телефонах, компьютерах и других электронных устройствах с помощью обычной клавиатуры на Тайване был разработан способ ввода иероглифов, основанный на алфавите чжуинь цзыму [Завьялова 2014, с. 197].

Современный этап латинизации – буквенные слова

Что же касается современного этапа латинизации китайского языка, то в настоящее время в китайском языке появляется все больше слов, в записи которых присутствуют буквы. Изучением буквенных слов в китайском языке занимаются китайские, отечественные и западные лингвисты.

Ю.С. Сухина полагает, что «многообразие типов буквенных слов, многообразие способов и механизмов интеграции буквенных элементов в систему китайского языка, активное образование бук-

венных аббревиатур в самом китайском языке позволяют выделить буквенные слова в отдельный вид современной китайской лексики» [Сухина 2010, с. 104].

Автор-составитель Словаря буквенных слов китайского языка, изданного в 2008 году, Лю Юнцюань указывает, что расширение международных контактов, появление новых предметов и явлений, которые трудно отразить в иероглифической записи, способствуют увеличению количества буквенных слов в китайском языке [Лю Юнцюань 2008, с. 1].

Лю Юнцюань отмечает четыре преимущества записи слов в буквенном виде: быстрота, облегченность написания (для написания буквенного слова требуется меньше черт, чем при написании слова в иероглифическом виде), ясность (наглядность), широта использования (многие буквенные слова могут быть понятны не только в Китае, но и за его пределами) [Лю Юнцюань 2008, с. 1–2].

Лю Юнцюань выделяет несколько классификаций буквенных слов. 1. В зависимости от принадлежности букв к тому или иному алфавиту:

a. Буквенные слова, которые записываются с помощью букв латинского алфавита (таких большинство). Например, 电子化 *huà* ‘электронизация’.

b. Буквенные слова, которые записываются с помощью букв греческого алфавита (таких намного меньше). Например, 测试 *cèshì* ‘бета-тестирование’.

2. В зависимости от однородности написания:

a. Буквенные слова, запись которых состоит только из букв. Например, FOB ‘цена FOB’.

b. Буквенные слова, которые записываются буквами в сочетании с символами, цифрами или иероглифами. Например, 3D, X光 *X guāng* ‘рентген’.

3. В зависимости от формы написания:

a. Буквенные слова, которые записываются полностью прописными буквами. Например, ISO ‘Международная организация по стандартизации’.

b. Буквенные слова, которые записываются полностью строчными буквами. Например, ml ‘миллилитр’.

c. Буквенные слова, которые начинаются с прописной буквы, а затем пишутся строчными. Например, Dbase ‘система управления базами данных Dbase’.

d. Буквенные слова, которые могут записываться и полностью прописными буквами, и полностью строчными, и в сочетании прописных и строчных букв. Например, EMAIL, Email, email ‘электронная почта’.

4. В зависимости от количества букв:

- a. Буквенные слова, в составе которых только одна буква. Обычно записывается в сочетании с иероглифами. Например, В超 *B chāo* ‘УЗИ’.
- b. Буквенные слова, в составе которых две буквы. Могут употребляться отдельно, а также в сочетании с иероглифами. Например, РС ‘персональный компьютер’.
- c. Буквенные слова, в составе которых более двух букв. Такие слова чаще употребляются самостоятельно, без сочетания с иероглифами или числами. Например, BASIC语言 *BASIC yǔyán* ‘язык BASIC’.

5. В зависимости от местоположения букв в слове:

- a. Буквенные слова, запись которых начинается с букв. Например, T恤 *T xiù* ‘футболка’.
- b. Буквенные слова, запись которых оканчивается буквой. Например, 卡拉 OK *kālā OK* ‘караоке’.
- c. Буквенные слова, в которых буквы стоят в середине слова. Например, 三C革命 *sān C gé mìng* ‘Революция трех C: компьютер (computer), контроль (control), коммуникация (communication)’ [Лу Юнцюань 2008, с. 322–323].

К.А. Маркина дает следующее определение буквенным словам: «Буквенные слова – это слова современного китайского языка, записываемые либо полностью буквами иностранных (преимущественно латинского) алфавитов, либо с их участием» [Маркина 2008, с. 3].

К.А. Маркина выделяет три типа буквенных слов в современном китайском языке. Как показали наши исследования, в интернет-лексике присутствуют все три типа:

- a) заимствованные буквенные слова, заимствованные буквенные аббревиатуры, а также некоторое количество несокращенных иностранных слов, которые употребляются в китайском языке в их полной буквенной записи, как и в языке-источнике;
- b) собственно китайские буквенные образования, созданные на основе фонетического алфавита пиньинь;
- в) буквенные слова (как аббревиатурные, так и несокращенные) в сочетании с китайскими иероглифическими компонентами [Маркина 2008, с. 3].

Как пишет Е. Дж. Юань, на форумах в целях обхода цензуры во время обсуждения внутриполитических острых тем пользователи Сети могут применять такие буквенные слова, как **MZ** вместо 民主 *mínzhǔ* ‘демократия’ или **TG** вместо 土共 *tǔ gòng* ‘Коммунистическая партия Китая (КПК)’, например:

- (1) 兄弟我告诉你, **MZ** 可是万金油, 永远没有错的。

Xiōngdì wǒ gàosù nǐ, MZ kěshì wànjīnyóu, yǒngyuǎn méiyǒu cuò de.

‘Брат, я скажу тебе, что **демократия** это панацея, она никогда не подведет.’

- (2) **TG** 愿意与 **DL** 谈判, 但是美国却见 **DL**,

这就破坏了中央与 **DL** 的谈判的基础。

*TG yuànyuì yǔ DL tánpàn, dànshì měiguó què jiàn DL,
zhè jiù pòhuài le zhōngyāng yǔ DL de tánpàn de jīchǔ.*

‘**КПК** хочет поговорить с **Далай-ламой**. Но встреча США с ним разрушила основу переговоров центра с **Далай Ламой**.’
[Yuan 2012, pp. 273–274]

Чэнь Шэри Юн соглашается с утверждением Е.Дж. Юань, отмечая, что применение данной языковой стратегии пользователями Интернета обусловлено стремлением обойти интернет-цензуру [Chen 2014].

Помимо обхода цензуры, как отмечает Е.Дж. Юань, буквенные слова могут использоваться для классификации участников дискуссии в соответствии с их политическими убеждениями, отраженными в комментариях, которые они делают. Например, **FQ** (сокращение от **愤怒青年 fènnù qīngnián** досл. ‘сердитый юноша’) обозначает участников дискуссии, которые публикуют эмоциональные комментарии с крайними политическими взглядами на социальные вопросы. Такие участники дискуссии скорее всего являются молодыми студентами, находящимися под сильным влиянием коммунистического политического мировоззрения, которое коренится в классовой борьбе и является антагонистическим. Слово **JY** (аббревиатура от **精英 jīngyīng** ‘элита’) – это ярлык для тех людей, которые часто выражают либеральные, правые политические взгляды. Считается, что эти участники дискуссии принадлежат к социальной элите. Эти люди обычно делают комментарии в пользу принятия демократической политической системы в западном стиле и нелиберальной экономики в Китае. Например:

- (3) 我算半个新人了, 一直潜水乱逛, 最近看国关的帖子, 真的感觉要感谢无数为了 **MZ** 奋斗的 **JY** 们, 本来说实话一直对 **TG** 有很多不满.....

Wǒ suàn bàn gè xīnrén le, yīzhí qiǎnshuǐ luàn guàng, zuìjìn kàn guóguān de tiēzi, zhēn de gǎnjué yào gǎnxiè wúshù wèile MZ fèndòu de JYmen, běnlái shuō shíhuà yīzhí duì TG yǒu hěnduō bùmǎn...

‘Можно считать меня новичком на форуме, я обычно не оставляю никаких сообщений, а просто бездумно серфю по Интернету. Недавно я наткнулся на один пост о государстве. Я действительно чувствую, что должно отблагодарить **социальные элиты**, которые усердно борются за **демократию** в Китае. Честно говоря, у меня много претензий к **КПК**...’ [Yuan 2012, p. 275].

В.В. Бачурин выделяет две причины использования слов, записанных латинскими буквами: «1) стремление к оптимизации языкового кода, так как лексическая единица, построенная из китайских слогов, может быть очень громоздкой, неудобной для произношения и восприятия: 高速缓冲存储器 *gāosù huǎnchōn cúnchúqì* (Cache), 小型计算机系统接口 *xiǎoxíng jìsuànjī xìtǒng jiēkǒu* (SCSI, Small Computer System Interface), 单指令流多数据流 *dān zhīlǐngliú duō shùjùliú* (SIMD, Single Instruction Multiple Data). 2) для программистов и специалистов ИТ, геймеров и различных групп и пользователей сети употребление алфавитных слов, аббревиатур и гибридов также играет роль маркера принадлежности к группе, является частью подязыка профессионального или интернет-общества» [Бачурин 2018, с. 113].

Стоит заметить, что на Тайване, в отличие от материкового Китая, не так часто используют буквенные слова, вероятно, потому, что популярным способом записи является не пиньинь, а чжуйнь, который основан на иероглифической графике [Yang 2007, pp. 10–11].

Латинизация в китаеязычной интернет-коммуникации

В данной работе мы решили проследить проявление процесса латинизации китайского языка в интернет-коммуникации.

Латинизация китайского языка в интернет-коммуникации проявляется в двух аспектах: буквенных заимствованиях (это слова в полной записи как в языке-источнике и всевозможные виды буквенных сокращений) и буквенных словах на основе пиньинь.

В 2012 г. издательством «Шанъу иньшугуань» выпущен Словарь интернет-языка Синьхуа, составителем которого является Ван Лэй [Ван Лэй 2012]. Он содержит 2946 лексических единиц. Словарь основан на алфавите пиньинь, хотя и включает неассимилированные заимствования из иностранных языков [Сбоев 2018, с. 118].

В 2014 г. был опубликован Новейший словарь терминов интернет-коммуникации [Новейший словарь терминов интернет-коммуникации 2014]. В словаре насчитывается 2011 лексических единиц и 327 эмограмм. Новейший словарь терминов интернет-коммуникации состоит из нескольких разделов: слова в иероглифическом виде, буквенные слова, слова, записанные цифрами, эмограммы. В первых двух разделах лексика расположена на основе алфавита пиньинь, цифровые слова расположены в порядке следования чисел от 0 до 9 [Сбоев 2018, с. 118].

Буквенные слова, представленные в Словаре интернет-языка Синьхуа и в Новейшем словаре терминов интернет-коммуникации, можно разделить на несколько видов.

1. Буквенные аббревиатуры от английских слов: **AAA** (образовано от “anytime, anyplace, anywhere”) – в любое время, в любом месте, где угодно; **IDNY** (образовано от “I don't know you” – я тебя не знаю); **DIYer** (образовано от “do it yourself” плюс суффикс *-er*) – тот, кто любит все делать сам, и др.

В словаре Синьхуа всего 467 таких слов, что составляет 15,8% от всего объема словаря, а Новейший словарь содержит 290 слов этой группы (14,4% от всего объема словаря).

2. Буквенные слова, записываемые первыми буквами алфавита пиньинь китайских морфем: **DX** (大虾 *dàxiā*) – продвинутый интернет-пользователь, **BC** (白痴 *báichī*) – *идуот*, **SG** (帅哥 *shuàigē*) – *красавчик* и др.

В словаре Синьхуа всего 161 слово данной группы, что составляет 5,5% от всего объема словаря, в Новейшем словаре – 130 слов (6,5% от всего объема словаря).

3. Буквенные слова, заимствованные из японского языка и записываемые латинскими буквами: **otaku** (от японского “おたく”) – любители анимэ.

Словарь Синьхуа содержит всего 4 таких слова (0,13% от всего объема словаря), Новейший словарь – 2 слова (0,09% от всего объема словаря).

4. Английское слово, записанное в том же виде, как оно используется в языке-оригинале, например **kick** – бить, пинать.

В словаре Синьхуа 45 слов данной группы, что составляет 1,5% от всего объема словаря, в Новейшем словаре – 26 слов (1,3% от всего объема словаря).

5. Буквенные слова, состоящие из одной буквы. Это может быть английское слово, начинающееся с этой буквы, или инициаль китайского слова, например **L** (от китайского 乐 *lè* или английского “laugh”) – смешной, радостный; **D** (от китайского 顶 *dǐng* ‘поддерживать’ или 弟 *dì* ‘братец (обращение к парню)’, 的 *de* ‘служебное слово’, 呆 *dāi* ‘дурак’), **g** (от английского “grin”) – ухмыляться.

Всего 21 слово в словаре Синьхуа (0,7% от всего объема словаря) и 25 слов в Новейшем словаре (1,2% от всего объема словаря).

6. Китайское слово, написанное полностью при помощи алфавита пиньинь, но без указания тона. Например, **ai** (爱 *ài*) – любить, **jiong** (囧 *jiǒng*) – грустный и др.

Словарь Синьхуа содержит 20 таких слов (0,67% от всего объема словаря), Новейший словарь – 10 (0,5% от всего объема словаря).

7. Буквенные слова, внешне отражающие то, что они обозначают: **orz** (вид человека, согнувшего колени и опустившего голову) может обозначать: грусть, тоска; сногсшибательный; большое уважение.

В словаре Синьхуа всего 9 слов данной группы (0,3% от всего объема словаря), в Новейшем словаре – 1 слово (0,05% от всего объема словаря).

Некоторые буквенные слова относятся как к первой, так и ко второй группе. Их значение определяется контекстом, например **GG** – старший брат (от китайского 哥哥 *gēge*), Google (от китайского 谷歌 *gǔgē*), должен идти (от английского “**gotta go**”), отличная игра (от английского “**good game**”) и др. [Словарь интернет-языка Синьхуа 2012].

Нами также были замечены некоторые особенности ассимиляции двух языков – китайского и английского. Интерес привлекло буквенное слово **JMS** – 姐妹们 *jiěmèimen* ‘сестры’ в силу того, что множественное число в аббревиатуре передано английским суффиксом множественного числа *-s*, в результате чего аббревиатура приняла букву ‘S’ в качестве последнего элемента. Так в китайском языке появилось новое буквенное слово на основе фонетического алфавита пиньинь с заимствованным грамматическим элементом. Другим примером смешения языков выступило слово **Ungeliable** – 不给力 *bù gěilì* ‘никуда не годится’, в котором китайский корень вставляется в готовую английскую модель, состоящую из английской отрицательной приставки *un-* и суффикса прилагательных *-able*.

Наиболее интересной на наш взгляд представляется тенденция к записыванию слов китайского языка полностью буквами алфавита пиньинь без соответствующего диакритического знака, обозначающего тон.

Причины записи китайских слов с использованием алфавита пиньинь

Китайские слова в полной записи с использованием алфавита пиньинь в интернет-коммуникации используются в следующих случаях.

1. Для замены:

1.1. Замена табуированного слова: то есть слово, записанное алфавитом пиньинь, используется как эвфемизм. Например:

- (4) 出了社会你就知道学校尼玛就是 **shi!**

Chūle shèhuì nǐ jiù zhīdào xuéxiào ní mǎ jiùshì shi

‘Выйдя в общество, ты сразу поймешь, что университет, твою мать, это **дерьмо**’.

1.2. Для замены слова, которое редко используется, и если нет времени искать его в выпадающей строке при наборе на клавиатуре. Например:

- (5) 真的皮草是把动物连毛带皮的扒下。而平时的羊毛衫，只是 **hao** 下来的毛。

Zhēn de pícǎo shì bǎ dòngwù lián máo dài pí de bā xià. Ér píngshí de yáng máo shān, zhǐshì hao xiàlái de máo

‘В настоящих шубах используется шкура и шерсть животных, а в обычных шерстяных свитерах лишь **выдернутые** куски шерсти’.

В данном предложении слово **hao** заменяет слово 薅 *hāo* ‘выдергивать’, которое используется в повседневном общении довольно редко.

1.3. Для замены иероглифа, который читается иначе на диалекте. Например:

- (6) 如果嫁人前S号变M号我就去**shi!**

Rúguǒ jià rénqián S hào biàn M hào wǒ jiù qù shi

‘Если до того, как я выйду замуж, мой размер одежды изменится с S на M, то я сразу умру’.

В данном предложении слово **shi** заменяет 死 *sǐ* ‘умереть’.

1.4. При замене сразу нескольких слов, имеющих одинаковое звучание. Например, слово **ta** заменяет 他 *tā* ‘он’ и 她 *tā* ‘она’. В китайском языке при записи местоимения третьего лица с использованием пиньинь не будет указан пол, в отличие от иероглифической записи, где по ключу слева это можно увидеть сразу [Турр 2018, р. 20].

2. Для привлечения внимания. Естественно, что слово, записанное буквами, в окружении слов, записанных иероглифами, привлечет внимание. Обычно в этих целях используются какие-либо эмоционально окрашенные слова. Например: 给! 我! *gun! Gěi! Wǒ! Gun!* ‘Катись отсюда!’

3. В качестве звукоподражания. Преимущества: 1) может подражать большему количеству звуков и схожесть звучания больше, так как буквенное слово может записываться не по правилам записи слогов в путунхуа, например, слово **puei** ‘апчхи’ является звукоподражанием чиханию; 2) буквенная запись нивелирует путаницу в употреблении того или иного иероглифа для обозначения звука,

например, для обозначения стука в дверь можно использовать как слово 碰碰 *pèngpèng*, так и слово 砰砰 *pēngpēng*, буквенная запись данного слова будет идентична – *pengpeng*.

Помимо этого, буквенная запись звукоподражаний может обозначать несколько разных звуков. Например, слово **biu** может обозначать звук выстрела – ‘пиф-паф’.

- (7) 鸟被 **biubiu** 两枪打死

Niǎo bèi biubiu liǎng qiāng dǎ sǐ

‘в птицу выстрелили два раза, **пиф-паф**, и она была убита’.

Также данное слово может обозначать быстрые изменения:

- (8) 孙悟空大闹了天宫之后, 被如来佛祖压在了五行山下,

“**biu**” 的一下, 五百年过去了...

Sūn Wùkōng dà nàole tiāngōng zhīhòu, bèi rúlái fózǔ yā zài le wǔháng shānxia, “biu” de yīxià, wǔbǎi nián guòquē...

‘После того как Сун Укун устроил переполох в небесном дворце, он был заточен Буддой под горой Пяти стихий, ‘**вжух**’ – и пятьсот лет прошло...’

Чжан И полагает, что иногда причина использования слова в полной записи пиньинь бывает неясна и это, возможно, происходит не целенаправленно, но часто слова, записанные инициалами слогов или полностью в записи пиньинь, используются, чтобы избежать онлайн-цензуры. Например:

- (9) 5分钟高清无码 **MAO** 片。

5 fēnzhōng gāoqīng wú mǎ MAO piàn.

‘Пятиминутный **порнофильм** без мозаики’ [Zhang 2017, p. 157]

Не Сяоли считает, что существует две основные проблемы с использованием слов в полной записи с помощью пиньинь:

1. Неунифицированное написание. Буквенные слова могут начинаться то со строчной, то с прописной буквы и даже быть полностью записаны прописными буквами. Чжоу Цзяньминь, однако, отмечает, что употребление буквенного слова, полностью записанного прописными буквами, позволяет выразить восклицание, выкрик, что невозможно было бы передать, используя иероглифическую запись. Например:

- (10) **GUN!** 不要让我再见到你!

***GUN!** Bùyào ràng wǒ zàijiàn dào nǐ*

‘**Убирайся!** Чтоб я тебя больше не видел!’ [Не Сяоли 2017, с. 73].

2. Раздельное или слитное написание между словами. Буквенная запись слов должна соответствовать правилам раздельного через пробел написания отдельных слов и слитного написания одного слова, даже если оно состоит из нескольких слогов. Однако в интернет-коммуникации пользователи могут нарушать оба этих правила. Не Сяоли приводит в пример слово **chou buyaolian** (в иероглифической записи 臭不要脸 *chòu bù yào liǎn*) 'бесстыжий'. Китайский лингвист отмечает, что фактически это является одним словом и должно записываться слитно, но в таком случае его прочтение будет затруднительно, а значение непонятно, так как слог *ya* может по сути делиться на *ya* и *o*, слог *lian* – на *li* и *an*, при слитной письменной записи буквами возникает сложность в разграничении слогов [Не Сяоли 2017, с. 40].

Заключение

В настоящее время запись слов латиницей без обозначения тонов широко используется на вывесках магазинов, в заголовках газет, названиях компаний, дорожных указателях, где эти надписи зачастую идут параллельно с иероглифическими. Помимо этого, латинская запись применяется для передачи китайских географических названий, имен собственных.

В эпоху интернационализации и компьютеризации латинизация китайского языка переживает новый этап. В китаеязычной интернет-коммуникации нередко встречаются слова, записанные латинскими буквами. Это заимствованные слова в исходном написании, сокращенные иностранные слова и слова, записанные с помощью алфавита пиньинь. Слова, полностью записанные с помощью пиньинь в интернет-коммуникации, не имеют знака тона, и их значение определяется в первую очередь контекстом. Написание слов с помощью пиньинь дает некоторые преимущества по сравнению с иероглифической записью слова, например возможность выразить восклицание, записав слово полностью прописными буквами. Мы отмечаем, что говорить о полной замене иероглифического письма на буквенное преждевременно, ведь использование буквенного слова в интернет-коммуникации всегда обусловлено какими-то дополнительными мотивами, однако происходящие сейчас процессы в китайском языке требуют внимания и дальнейшего исследования.

Благодарности

Автор статьи выражает благодарность коллективу кафедры китаеведения Восточного института – Школы региональных и международных исследований Дальневосточного федерального университета и лично профессору кафедры Анне Александровне Хаматовой за помощь в выполнении данного исследования. Автор признателен рецензентам за выдвинутые критические замечания к статье, устранения которых способствовали улучшению качества работы.

Acknowledgments

The author of the article expresses gratitude to the staff of the Department of Sinology of the Eastern Institute – School of Regional and International Studies of the Far Eastern Federal University and personally to Professor of the Department Anna Aleksandrovna Khamatova for her help in carrying out this study. The author is grateful to the reviewers for the critical comments on the article shortcomings, the elimination of which contributed to the improvement of the quality of the work.

Литература

- Алексахин 2010 – *Алексахин А.Н.* Алфавит китайского языка путунхуа. Буква – фонема – звук речи – слог – слово. 2-е изд., испр. и доп. М.: Восточная книга, 2010. 184 с.
- Алексеев 1932 – *Алексеев В.М.* Китайская иероглифическая письменность и ее латинизация. Л.: Академия наук СССР, 1932. 178 с.
- Алпатов 2013 – *Алпатов В.М.* Размышления над книгой В.М. Алексеева «Китайская иероглифическая письменность и ее латинизация» // Общество и государство в Китае. 2013. № 43 (2). С. 300–314.
- Бачурин 2018 – *Бачурин В.В.* Буквенные слова и гибридные образования в китайской ит-терминологии и интернет-лексике [Электронный ресурс] // Международный научно-исследовательский журнал. 2018. № 9 (75). Ч. 2. С. 112–115. URL: <https://research-journal.org/languages/bukvennyye-slova-i-gibridnye-obrazovaniya-v-kitajskoj-it-terminologii-i-internet-leksike/> (дата обращения 14 февраля 2020).
- Гурин 2011 – *Гурин М.В.* Китайская иероглифика в современном мире // Историческая и социально-образовательная мысль. 2011. № 3 (8). С. 87–91.
- Завьялова 2014 – *Завьялова О.И.* Большой мир китайского языка. 2-е изд. М.: Восточная книга, 2014. 320 с.
- Маркина 2008 – *Маркина К.А.* Новые тенденции в развитии китайской лексики: буквенные слова китайского языка как проявление его интеграции с элементами иной типологии: Автореф. дис. ... канд. филол. наук. М., 2008. 23 с.

- Саськова 2016 – *Саськова Е.В.* Алфавит китайского языка как элемент языковой политики КНР [Электронный ресурс] // Современные научные исследования и инновации. 2016. № 1. URL: <http://web.snauka.ru/issues/2016/01/62296> (дата обращения 03 февраля 2020).
- Сбоев 2018 – *Сбоев А.Н.* Словари и сборники интернет-лексики современного китайского языка // Труды Института востоковедения РАН. 2018. № 19. С. 115–124.
- Сердюченко 1959 – *Сердюченко Г.П.* Китайская письменность и ее реформа. М.: Восточная литература, 1959. 54 с.
- Софронов 2007 – *Софронов В.М.* Китайский язык и китайская письменность: Курс лекций. М.: АСТ: Восток–Запад, 2007. 628 с.
- Спешнев 1980 – *Спешнев Н.А.* Фонетика китайского языка. Л.: Изд-во Ленинградского университета, 1980. 141 с.
- Сухина 2010 – *Сухина Ю.С.* Особенности интеграции буквенных слов в китайскую языковую систему // Слово: Фольклорно-диалектологический альманах. 2010. № 8. С. 104–108.
- Чжоу 1989 – *Чжоу Югуан.* Модернизация китайского языка и письменности // Новое в зарубежной лингвистике. Языкознание в Китае. 1989. Вып. XXII. С. 376–398.
- Ван Лэй 2012 – *Ван Лэй (汪磊).* Xinhua wangluo yuyan cidian (新华网络语言词典) [Словарь интернет-языка Синьхуа]. Пекин: Шаньту иньшугуань, 2012. 233 с.
- Лю Юнцюань 2008 – *Nanyu zimuci cidian (汉语字母词典)* [Словарь буквенных слов китайского языка] / Сост.: Лю Юнцюань (刘涌泉). Пекин: Вайюй цзяосюэ чубаньшэ, 2008. 332 с.
- Не Сяоли 2017 – *Не Сяоли (聂小丽).* Wangluo yuyan zhong hanyu pinyin hanzi de shiyong ji qi guifanhua wenti (网络语言中汉语拼音文字的使用及其规范化问题) [Использование пиньинь в китайском интернет-языке и его стандартизация] // Вестник Лэшаньского педагогического института. 2017. Т. 32. no. 2. P. 36–57.
- Новейший словарь терминов интернет-коммуникации 2014 – *Zuixin wangluo jiaoji yongyu cidian (最新网络交际用语词典)* [Новейший словарь терминов интернет-коммуникации] / Сост.: Чжоу Цзяньминь, Сюн Иминь (周建民, 熊一民主编), испр. и доп. изд. Пекин: Чжунго шэхуэй кэсюэ чубаньшэ, 2014. 522 с.
- Су Пэйчэн 2018 – *Су Пэйчэн (苏培成).* Nanyu pinyinhua de fansi (汉语拼音化的反思) [Размышления о пиньинизации китайского языка] // The Study of Chinese Characters and Language. 2018. Vol. 3. P. 97–107.
- Чжан Фэн, Ма Сюэдань 2010 – *Чжан Фэн, Ма Сюэдань (张丰, 麻雪丹).* Ye tan “hanyu pinyinhua” (也谈“汉语拼音化”) [Рассуждения на тему «Пиньинизация китайского языка»] // Сборник докладов 6-й научно-технической конференции пр. Цзилинь «Низкоуглеродистая экономика и научное развитие», г. Чанчунь, 6 октября. Чанчунь, 2010. С. 725–727.
- Чэн Таоань, Чэн Чжиюн 2012 – *Чэн Таоань, Чэн Чжиюн (程陶庵, 程志泳).* Hanzi ladinghua yiyi – cong 56 nian qian Mao Dun zhi Ye Laishi de fuxin tanqi (汉字拉丁化疑义——从 56 年前茅盾致叶籁士的复信谈起) [Сомнение в пользу латинизации китайского языка – рассуждения с ответа Мао Дуня Е. Лайши 56 лет назад] // Газета «Гуанмин жибао». 2012. 11-й вып.
- Chappell 1980 – *Chappell H.* The Romanization Debate // The Australian Journal of Chinese Affairs. 1980. Vol. 4. P. 105–118.

- Chen 2014 – *Chen S.Y.* From OMG to TMD – Internet and Pinyin acronyms in Mandarin Chinese [Электронный ресурс] // Language@Internet. 2014. № 11. article 3. URL: <http://www.languageatinternet.org/articles/2014/chen> (дата обращения 12 мая 2019).
- DeFrancis 2006 – *DeFrancis J.* The Prospects for Chinese Writing Reform. Philadelphia, PA: University of Pennsylvania, 2006. 29 p. (Sino-Platonic Papers, № 171).
- The Progress of Pinyin 2008 – The Progress of Pinyin [Электронный ресурс] // Center for the Study of Intelligence. 2008. Volume 5. Issue 1. URL: https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/vol5no1/html/v05i1a06p_0001.htm (дата обращения 12 мая 2019).
- Trupp 2018 – *Trupp G.* Blogs are the Mirrors to Ourselves: Examining Emerging Semantic Conventions in an Online World [Электронный ресурс] // Sociology. 2018. 35 p. URL: <https://www.semanticscholar.org/paper/Blogs-are-the-Mirrors-to-Ourselves%3A-Examining-in-an-Trupp/2704355625181bf474421df6e228fbe84d-3b121e> (дата обращения 2 марта 2020).
- Yang 2007 – *Yang C.* Chinese Internet language: A sociolinguistic analysis of adaptations of the Chinese writing system [Электронный ресурс] // Language@Internet. 2007. Vol. 4. article 2. URL: <http://www.languageatinternet.org/articles/2007/1142> (дата обращения 2 марта 2020).
- Yuan 2012 – *Yuan E.J.* Language use as social practice on the Chinese Internet // Computer-mediated communication across cultures: International interactions in online environment / K.S. Amant, S. Kelsey (eds.). Pennsylvania, PA: IGI Global, 2012. P. 266–281.
- Zhang 2017 – *Zhang Yi.* Heteroglossic Chinese Online Literacy Practices On Micro-Blogging and Video-Sharing Sites [Электронный ресурс] // Scholar Commons. 2017. URL: <http://scholarcommons.usf.edu/etd/6788> (дата обращения 2 августа 2020).
- Zhou 2014 – *Zhou Yirou.* Script Crisis and Literary Modernity in China, 1916–1958: Ph.D. Thesis. New York: Columbia University, 2014. 256 p.

References

- Aleksakhin, A.N. (2010), *Alfavit kitayskogo yazyka putunkhua. Bukva – fonema – zvuk rechi – slog – slovo* [The alphabet of Chinese language (Putonghua). Letter – phoneme – speech sound – syllable – word], Vostochnaya kniga, Moscow, Russia.
- Alekseev, V.M. (1932), *Kitayskaya ieroglificheskaya pis'mennost' i ee latinizatsiya* [Chinese hieroglyphic writing and its Latinization], Akademiya nauk SSSR, Leningrad, USSR.
- Alpatov, V.M. (2013), “Reflection on the book “Chinese hieroglyphic writing and its Latinization” by V.M. Alekseyev, *Obshchestvo i gosudarstvo v Kitae*, no. 43 (2), pp. 300–314.
- Bachurin, V.V. (2018), “Literal words and hybrid formations in Chinese IT-terminology and online vocabulary”, *Mezhdunarodnyi nauchno-issledovatel'skii zhurnal*, no. 9, pp. 112–115, [Online], available at: <https://research-journal.org/languages/bukvennye-slova-i-gibridnye-obrazovaniya-v-kitajskoj-it-terminologii-i-inter-net-leksike/> (Accessed 14 February 2020).

- Center for the Study of Intelligence (2008), “The Progress of Pinyin”, vol. 5, no. 1 [Online], available at: https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/vol5no1/html/v05i1a06p_0001.htm (Accessed 3 February 2020).
- Chappell, H. (1980), “The Romanization Debate”, *The Australian Journal of Chinese Affairs*, no. 4, pp. 105–118.
- Chen, S.Y. (2014), “From OMG to TMD – Internet and Pinyin acronyms in Mandarin Chinese”, *Language@Internet*, no. 11, [Online], available at: <http://www.languageatinternet.org/articles/2014/chen> (Accessed 12 May 2019).
- Cheng, Taoan and Cheng, Zhiyong (2012), “Hanzi ladinghua yiyi – cong 56 nian qian Mao Dun zhi Ye Laishi de fuxin tanqi” [Doubt about the benefits of latinizing Chinese – reasoning from the answer of Mao Dun to Ye Laishi 56 years ago], *Guangming ribao*, no 11.
- DeFrancis, J. (2006), *The Prospects for Chinese Writing Reform*. Philadelphia, PA: University of Pennsylvania, 2006. 29 p. (Sino-Platonic Papers, № 171).
- Gurin, M.V. (2011), “Chinese hieroglyphic writing in the modern world”, *Istoricheskaya i sotsial'no-obrazovatel'naya mysl'*, no. 3 (8), pp. 87–91.
- Liu, Yongquan (2008), *Hanyu zimuci cidian* [Chinese letter word dictionary], Foreign Language Teaching and Research Press, Beijing, China.
- Markina, K.A. (2008), *Novye tendentsii v razvitiï kitaiskoi leksiki: bukvennye slova kitaiskogo yazyka kak proyavlenie ego integratsii s elementami inoi tipologii* [New trends in the development of Chinese vocabulary. Letter words of the Chinese language as the demonstration of its integration with the elements of the other typology], Ph.D. Thesis, Moscow, Russia.
- Nie, Xiaoli (2017), “Wangluo yuyan zhong hanyu pinyin hanzi de shiyong ji qi guifanhua wenti” [The Use of Chinese Pinyin in Internet Language and Its Standardization], *Journal of Leshan teachers college*, vol. 32, no. 2, pp. 36–57.
- Sas'kova, E.V. (2016), “The alphabet of the Chinese language as an element of the PRC's language policy”, *Sovremennyye nauchnyye issledovaniya i innovatsii*, no. 1, [Online], available at: <http://web.snauka.ru/issues/2016/01/62296> (Accessed 3 February 2020).
- Sboev, A.N. (2018), “Dictionaries and compilations of modern Chinese online vocabulary”, *Trudy Instituta vostokovedeniya RAN*, no. 19, pp. 115–124.
- Serdyuchenko, G.P. (1959), *Kitayskaya pis'mennost' i eye reforma* [Chinese writing system and its reform], Vostochnaya literatura, Moscow, Russia.
- Sofronov, V.M. (2007), *Kitaiskii yazyk i kitaiskaya pis'mennost'*. Kurs lektsii [Chinese language and Chinese writing system. Lecture course], AST: Vostok-Zapad, Moscow, Russia.
- Speshnev, N.A. (1980), *Fonetika kitaiskogo yazyka* [Phonetics of the Chinese language], Izdatel'stvo Leningradskogo universiteta, Leningrad, USSR.
- Sukhina, YU.S. (2010), “Features of the letter words' integration into the Chinese language system”, *Slovo: Fol'klorno-dialektologicheskii al'manakh*, no. 8, pp. 104–108.
- Su, Peicheng (2018), “Hanyu pinyinhua de fansi” [Reflections on the Chinese Pinyinization], *The Study of Chinese Characters and Language*, no. 3, pp. 97–107.
- Trupp, G. (2018), “Blogs are the Mirrors to Ourselves: Examining Emerging Semantic Conventions in an Online World”, *Sociology*, [Online], available at: <https://www.semanticscholar.org/paper/Blogs-are-the-Mirrors-to-Ourselves%3A-Examining-in-an-Trupp/2704355625181bf474421df6e228f8e84d3b121e> (Accessed 3 February 2020).

- Wang, Lei (2012), *Xinhua wangluo yuyan cidian* [Xinhua Internet Language Dictionary], The Commercial Press, Beijing, China.
- Yang, C. (2007), “Chinese Internet language: A sociolinguistic analysis of adaptations of the Chinese writing system”, *Language@Internet*, no. 4, [Online], available at: <http://www.languageatinternet.org/articles/2007/1142> (Accessed 3 February 2020).
- Yuan, E.J. (2012), “Language use as social practice on the Chinese Internet”, in Amant, K.S. and Kelsey, S. (eds.), *Computer-mediated Communication Across Cultures: International Interactions in Online Environment*, IGI Global, Pennsylvania, PA, pp. 266–281.
- Zav'yalova, O.I. (2014), *Bol'shoi mir Kitaiskogo yazyka* [The great world of Chinese language], Vostochnaya kniga, Moscow, Russia.
- Zhang, Feng and Ma, Xuedan (2010), “Ye tan ‘hanyu pinyinhua’” [About “Chinese Pinyinization”], *Low-carbon economy and scientific development: Proceedings of the 6th Annual Scientific and Technical Conference of Jilin Province*, Changchun, 6 October, Changchun, China, pp. 725–727.
- Zhou, Youguang (1989), “Modernization of Chinese language and writing”, *Novoye v zarubezhnoi lingvistike, Linguistics in China*, vol. 22, pp. 376–398.
- Zhou, Jianmin and Xiong, Yimin (2014), *Zuixin wangluo jiaoji yongyu cidian* [New Dictionary of online communication], China Social Sciences Press, Beijing, China.
- Zhang, Yi (2017), “Heteroglossic Chinese Online Literacy Practices On Micro-Blogging and Video-Sharing Sites”, *Scholar Commons*, [Online], available at: <http://scholarcommons.usf.edu/etd/6788> (Accessed 2 August 2020).
- Zhou Yurou (2014), *Script Crisis and Literary Modernity in China, 1916–1958*, Ph.D. Thesis, Columbia University, New York, USA, 256 p.

Информация об авторе

Александр Н. Сбоев, кандидат филологических наук, Дальневосточный федеральный университет, Владивосток, Россия; 690090, Владивосток, Россия, ул. Суханова, д. 8; sboevalexander@mail.ru

Information about the author

Aleksandr N. Sboev, Cand. of Sci. (Philology), Far Eastern Federal University, Vladivostok, Russia; bld. 8, Sukhanova Str., Vladivostok, Russia, 690090; sboevalexander@mail.ru

УДК 811.352.3

DOI: 10.28995/2686-7249-2020-7-84-106

Материалы к типологии ассоциативной множественности: кубанский диалект кабардино-черкесского языка

Ирина Г. Багирокова

*Институт языкознания РАН, Москва, Россия
Национальный исследовательский университет «Высшая школа экономики»
Москва, Россия, ibagirokova@yandex.ru*

Алина И. Баранова

*Кельн, Германия
baranovaalinka@gmail.com*

Юрий А. Ландер

*Национальный исследовательский университет «Высшая школа экономики»
Москва, Россия, yulander@yandex.ru*

Аннотация. В статье описываются две конструкции ассоциативной множественности в кубанском диалекте кабардино-черкесского языка. Первая конструкция представляет собой классический пример употребления суффикса, также используемого для маркирования аддитивного множественного числа, с именами собственными – преимущественно для обозначения семьи фокального референта. Во второй конструкции появляется аналитический показатель, который может присоединяться к широкому кругу именных групп (в том числе неодушевленных, нереферентных и сочиненных), причем управляя их падежом. Для того чтобы описать особенности аналитической конструкции, предлагается задействовать в типологии ассоциативности не только «каноническую» иерархию одушевленности, но и ряд иерархий, включающий, в частности, иерархии определенности, числовой индивидуализованности и морфосинтаксической автономности.

Ключевые слова: ассоциативная множественность, кабардино-черкесский язык, определенность, индивидуализованность, морфосинтаксическая автономность

Для цитирования: Багирокова И.Г., Баранова А.И., Ландер Ю.А. Материалы к типологии ассоциативной множественности: кубанский диалект кабардино-черкесского языка // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 84–106. DOI: 10.28995/2686-7249-2020-7-84-106

© Багирокова И.Г., Баранова А.И., Ландер Ю.А., 2020

Data for the typology of associative plurals: Kuban Kabardian

Irina G. Bagirokova

*Institute of Linguistics RAS, Moscow, Russia
HSE University, Moscow, Russia, ibagirokova@yandex.ru*

Alina I. Baranova

*Cologne, Germany
baranovaalinka@gmail.com*

Yury A. Lander

HSE University, Moscow, Russia, yulander@yandex.ru

Abstract. Kabardian (Northwest Caucasian) displays two associative plural constructions. The first pattern exploits the suffix which is also used for the expression of additive plural: it is added to proper names and normally provides a reference to the family of the focal referent. Within the second pattern, a specific associative plural marker follows a syntactically autonomous nominal. The latter pattern possesses several specific properties: the associative plural marker governs the case of the focal nominal, which can be represented even by inanimate, non-specific and coordinate NPs. To describe the Kabardian associative plural system, we would suggest using not only a simplified version of Animacy Hierarchy (as is often done in typological literature) but involving several other hierarchies including those of definiteness/referentiality, number individuation, and morphosyntactic autonomy.

Keywords: associative plural, Kabardian, referentiality, individuation, morphosyntactic autonomy

For citation: Bagirokova, I.G., Baranova, A.I. and Lander, Yu.A. (2020), "Data for the typology of associative plurals: Kuban Kabardian", *RSUH/RGGU Bulletin. "Literary Theory. Linguistics. Cultural Studies" Series*, no. 7, pp. 84-106, DOI: 10.28995/2686-7249-2020-7-84-106

С шумным ветром улетело
Настроение плохое.
Сердце сразу захотело
Жить, любить и все такое.

Л. Леушина

1. Введение

Настоящая статья посвящена конструкциям, выражающим ассоциативную множественность в кабардино-черкесском языке (абхазо-адыгская семья): мы рассматриваем тип множественной референции, при котором референция к одним элементам множества

носит эксплицитный характер, а к другим – неэксплицитный [Даниэль 2000]. Так, в примерах (1) и (2) выделяется один *фокальный* участник (Руслан и Замира соответственно)¹, но соответствующая именная группа указывает на множество, элементом которого является фокальный участник².

- (1) *wəna.ʃhe-r* *rʷəslan-xe* *ja-tje-x^w-a*
 крыша-abs Руслан-PL 3PL.IO-LOC-падать-PST
 ‘Крыша упала на семью Руслана.’
- (2) *zamjəre* *səme* *nəpawə-r* *ja-p-a*
 Замира APL ребёнок-ABS 3PL.ERG-воспитывать-PST
 ‘Замира и те, кто с ней, воспитали ребёнка.’

Ассоциативная множественность противопоставлена в первую очередь аддитивной множественности, при которой каждый элемент описываемого множества подходит под описание через основу формы множественного числа [Барулин 1980], как в (3). Однако существуют и другие типы конструкций, выражающих множественность, – сочинительные и комитативные, в которых элементы множества перечисляются [Даниэль 2000; Архипов 2009], ср. (4).

- (3) *lə-xe-m* *pχe-xe-r* *ze-p-jə-x-a*
 мужчина-PL-OBL дрова-PL-ABS REC.IO-LOC-3SG.ERG-пилить-PST
 ‘Мужчины напилили дрова.’
- (4) *fjəzjəke-m-re* *matjematjəke-m-re* *č'etərje*
 физика-OBL-COORD математика-OBL-COORD четыре
qə-s-h-rjə *sə-šə -h-a*
 DIR-1SG.ERG-уносить-ADD 1SG.ABS-LOC-выходить-PST
 ‘По физике и математике я получила четверку и поступила.’

¹ Термин «фокальный участник» представляет собой соответствие “focal element” в [Daniel, Moravcsik 2005]. В [Даниэль 1999; 2000] используется термин «фокусный участник» – на наш взгляд, неудобный, поскольку термин «фокус» используется преимущественно для описания коммуникативной структуры. В [Даниэль 2000] участник, относительно которого определяется референция группы, именуется также эксплицитным референтом.

² Обсуждение семантики ассоциативных форм см. [Даниэль 1999; 2000; Moravcsik 2003; Daniel, Moravcsik 2005].

Как видно из (1)–(2), в кабардино-черкесском языке есть две конструкции ассоциативной множественности, морфологическая и аналитическая. Ниже они описываются для кубанского диалекта (говор аула Ходзь Республики Адыгея)³. Кабардино-черкесский язык представляет определенный интерес для типологии ассоциативности по двум причинам. Во-первых, две конструкции здесь различаются сферой употребления, причем одна из этих конструкций, аналитическая, используется чрезвычайно широко⁴. Во-вторых, кабардино-черкесские конструкции нетривиальным образом взаимодействуют с другими именными категориями – категорией числа и категорией падежа.

Кабардино-черкесский язык характеризуется как близкий к эталону агглютинации и полисинтетический – обладающий сложной морфологией, которая передает значительное количество информации о ситуации внутри слов. Помимо развитого вершинного маркирования (выражения актантов в сказуемом) в кабардино-черкесском языке имеется и зависимостное маркирование – падежная система. Традиционно (см., например, [Грамматика 1957; Грамматика 1970; Кабардино-черкесский язык 2006]) выделяется четыре падежа: два ядерных – абсолютный и косвенный (в литературе обычно именуемые именительным и эргативным), а также инструменталис (творительный падеж) и адвербиалис (обстоятельный падеж). Показатели падежей, как правило, появляются в именной группе один раз – на вершине (которая в норме замыкает группу). Суффикс инструменталиса в некоторых условиях – прежде всего при определенности именной группы – сочетается в одной форме с суффиксом косвенного падежа⁵. Маркеры ядерных падежей могут отсутствовать на личных местоимениях, именах собственных, в нереперентных именных группах и в посессивных

³ Кубанский диалект располагается в отдалении от основных кабардино-черкесских диалектов и отличается от них во многих отношениях [Кумахов 1969]. Соответственно, хотя основные факты, излагаемые в нашей статье, по-видимому, верны и для литературного кабардино-черкесского языка, нельзя исключать, что мы фиксируем и некоторые инновации кубанского диалекта.

⁴ Конструкции ассоциативной множественности присутствуют и в других живых абхазо-адыгских языках; см., в частности, обсуждение адыгейской ассоциативной множественности в [Bagirokova et al., to appear]. Конструкция с полноценным аналитическим ассоциативом с широкой сферой использования, насколько нам известно, зафиксирована только для кабардино-черкесского языка.

⁵ См. подробное обсуждение аналогичного явления в близкородственном адыгейском языке в [Сердобольская, Кузнецова 2009].

группах. Противопоставляются две формы аддитивного числа – немаркированная форма (обычно выражающая единственное число) и форма множественного числа, маркируемая суффиксом *-xe*. Для нас существенно, что в аддитивных формах множественного числа консультанты обычно настаивают на присутствии падежного маркера независимо от других свойств именной группы.

Во втором разделе статьи мы касаемся основных параметров типологии ассоциативной множественности. В третьем и четвертом разделах подробно рассмотрены конструкции с *-xe* и *səte*. Пятый раздел содержит некоторые предложения, касающиеся описания типологии ассоциативной множественности. В последнем разделе представлены выводы и возможные направления дальнейших исследований.

2. Параметры типологии ассоциативных конструкций

В литературе по ассоциативной множественности (см., например, [Daniel, Moravcsik 2005]) обычно обсуждаются три параметра типологии этих конструкций: (i) тип группы, (ii) тип фокального элемента, (iii) формальный тип маркирования.

Тип группы. Согласно [Daniel, Moravcsik 2005], конструкции ассоциативной множественности могут обозначать следующие группы (*X* – фокальный элемент): (i) ‘*X* и его семья’; (ii) ‘*X* и его друзья или знакомые’; (iii) ‘*X* и другие члены временно существующей группы, членом которой является *X*’. Языки могут разрешать либо исключительно первую интерпретацию ассоциативных конструкций, либо первую и вторую, либо все три возможные интерпретации (см. также [Даниэль 2000]).

Допустим и другой подход к семантике ассоциативности: отношение между участниками группы может извлекаться либо из лексической семантики имени-основы, либо из контекста и только в отдельных случаях приписываться самой конструкцией. К первому случаю относятся, в частности, примеры, когда форма ассоциативной множественности от слова ‘отец’ означает ‘родители’, ко второму – обозначения ситуативно заданных групп, к третьему – прежде всего конструкции, ограничивающие группы описанием семьи независимо от фокального элемента. Тогда интерпретация ассоциативных конструкций оказывается подобна, например, интерпретации посессивных конструкций, в которых посессивное отношение также извлекается либо из лексической семантики, либо из контекста (см., например, [Partee 1997]).

Говоря о типе группы, следует заметить, что ассоциативная множественность, выражая замкнутую совокупность объектов, противопоставляется в литературе некоторым сходным значениям – прежде всего симилиативной множественности, выражению открытого, «размытого» класса, включающего эксплицированный референт и другие объекты, обладающие чертами сходства с ним [Даниэль 2000]; ср. выражения вроде *танцы-шманцы, танцы и всё такое*. Это противопоставление очевидно в языках, использующих специфицированные симилиативные конструкции. Вместе с тем, если ассоциативная и симилиативная множественность охватываются одной конструкцией, провести четкую границу между ними оказывается чрезвычайно сложно: симилиативная множественности может изображаться как та же конструкция, но с ослаблением отношения, меньшей привязкой его к описываемой ситуации и, как следствие, с меньшей определенностью выражаемой группы.

Тип фокального элемента. В [Daniel, Moravcsik 2005] фактически представлена иерархия, позволяющая описать лексические ограничения на существительные, которые образуют конструкции ассоциативной множественности:

- (5) Имена собственные > Термины родства > Прочие нарицательные имена, обозначающие людей > Имена, не обозначающие людей

Чем ниже позиция имени на этой шкале, тем менее вероятно образование ассоциативной конструкции. Если в языке образуются ассоциативные формы от существительных, находящихся ниже на шкале, в нем ожидаются и ассоциативные формы от существительных, которые располагаются выше.

Легко заметить, что приведенная иерархия представляет собой разновидность шкалы одушевленности [Corbett 2001, pp. 83–84]. Правда, в этой версии шкалы одушевленности отсутствуют местоимения, присутствующие в других версиях, но множественное число местоимений 1-го и 2-го лица часто несет ассоциативное значение (ср. *мы* ‘я и те, кто со мной’, *вы* ‘ты и те, кто с тобой’; см., например, [Даниэль 2000; Vassilieva 2005]). Соответственно добавить их в иерархию (5) не представляет труда.

Формальный тип маркирования. Значение ассоциативности выражается разными способами. М.А. Даниэль [Даниэль 2000] противопоставляет морфологические, аналитические и синтаксические ассоциативы. Морфологические ассоциативы подразумевают присутствие морфологически связанного показателя, который может совпадать с показателем аддитивной множественности или представлять собой специфицированный маркер ассоциатив-

Таким образом, с точки зрения функции для кабардино-черкесского языка можно говорить об ассоциативных формах. Ниже мы увидим, что эти формы не ограничиваются употреблением, описанными в грамматиках.

Формальные особенности. Несмотря на сходство с аддитивными формами, у морфологических ассоциативов обнаруживаются весьма интересные особенности – в первую очередь во взаимодействии с категорией падежа. В то время как в аддитивных формах множественного числа, маркируемых *-xe*, показатели падежей обычно оказываются обязательными, в ассоциативных формах с тем же суффиксом показатели ядерных падежей опускаются [Кумахов 1971, с. 5–6] – как в позиции, которая обычно требует косвенного падежа (9), так и в позиции, которая обычно требует абсолютного падежа (10).

- (10) *amjəne-xe* *twəʃ'anə-m* *ʃə-t-λak^v-a*
 Амина-PL магазин-OBL LOC-1PL.ERG-видеть-PST
 'Мы видели семью Амины в магазине.'

Отсутствие падежного показателя не является для ассоциативов с *-xe* обязательным:

- (11) *ashad-xe(-m)* *ja-ʒ'a-p-ʔ-a?*
 Асхад-PL-OBL 3PL.IO-PRF-2SG.ERG-говорить-PST
 'Ты сказал Асхаду и его семье?'

Формы с выраженным показателем косвенного падежа употребляются с большей легкостью, чем формы с показателем абсолютного падежа. Это коррелирует и с тем, насколько легко падежные показатели опускаются в других конструкциях, – согласно нашим предварительным наблюдениям, по крайней мере при нереперентных именных группах отсутствие абсолютного показателя допускается чаще, чем отсутствие косвенного.

Показатели неядерных падежей сочетаются с ассоциативными формами, но в отличие от маркеров ядерных падежей не опускаются. Ср. присоединение адвербиалиса (12) и инструменталиса (13) (появление косвенного падежа в сочетании с последним ожидаемо факультативно):

- (12) *tə* *wəne-r* *z-je-r* *mwəse-x-wə=ra?*
 этот дом-ABS REL.IO-POSS+иметься-ABS Муса-PL-ADV=COP+Q
 'Это дом Мусы и его семьи?' (Букв.: 'Тот, чей этот дом, – Муса и его семья?')

- (13) *rome-xe(-m)-č'e* *də-r-je-kʷe-č'ə-rjə*,
 Рома-PL-OBL-INS 1PL.ABS-TRANS-DAT-идти-выходить-ADD
mwəse-xe(-m)-č'e *də-qə-de-č'ə-ž'-a*
 Муса-PL-OBL-INS 1PL.ABS-DIR-LOC-выходить-RE-PST
 'Мы прошли в сторону дома Ромы и его семьи и вышли через то место, где живет Муса с семьей.'

Функционирование. При элицитации в отсутствие контекста носители практически всегда настаивают на том, что суффиксальные ассоциативные формы выражают отношения родства. Тем не менее иногда допускается и другой перевод:

- (14) *ashad-xe(-r)* *wəramə-m* *tje-t(-xe)*
 Асхад-PL-ABS улица-OBL LOC-стоять-PL
 'Асхад и его друзья на улице.'

Показатель *-xe* в ассоциативной функции присоединяется исключительно к именам собственным, указывающим на людей,⁶ и не сочетается с терминами родства (ср. примеры 15а и 15b). Разумеется, имена собственные, образующие ассоциативные формы, всегда представляют собой определенные именные группы, указывающие на единичный референт.

- (15) а. *marjəne-xe* *ja-pšante-m* *praznikə-r*
 Марина-PL 3PL.IO+POSS-двор-OBL праздник-ABS
šə-xe-d-ve.wənexʷə-č'-a
 LOC-LOC-1PL.ERG-отмечать-выходить-PST
 'Праздник мы отмечали во дворе семьи Марины.'
- б. **praznikə-r* *šə-xe-ve.wənexʷə-č'-a*
 праздник-ABS LOC-LOC-отмечать-выходить-PST
s-j-ade-xe-m *ja-pšante-m*
 1SG.IO-POSS-отец-PL-OBL 3PL.IO+POSS-двор-OBL
 (Ожид. 'Праздник отмечался во дворе семьи моего отца.')

⁶ Употребление с именами собственными, указывающими на животных, иногда допускается, но на наш взгляд оно свидетельствует о явлении вроде персонификации – поднятии референта/именной группы по шкале одушевленности.

Ассоциативный *-xe* присоединяется и к сложным выражениям – сочетаниям фамилии и имени (16). В адыгских языках сочетания такого рода могут образовывать «именной комплекс» – сложные образования, обнаруживающие свойства единого слова (см. о них [Lander 2017]). Однако в кабардино-черкесском языке, по-видимому, происходит переосмысление именного комплекса как словосочетания. В результате сказать с уверенностью, что в (16) суффикс присоединяется к сочетанию слов или образованию, более «спяянному» с морфологической точки зрения, по-видимому, невозможно.

- (16) *qʷəʃhe zewərbja-xe(-m) ja-ʒʹegʷə-m*
 Кушхов Заурбий-PL-OBL 3PL.IO+POSS-свадьба-OBL
sə-qə-r-a-ke-blek-a
 1SG.ABS-DIR-DAT-3PL.ERG-CAUS-близкий-PST
 ‘Семья Заурбия Кушхова пригласила меня на свадьбу.’

Иногда суффиксальные ассоциативные формы допускают референцию к группе, в данный момент не включающей фокального референта (17).

- (17) *mwəse-xe ʋʷe-lə-žʹ-a-xe, awe jezə mwəse še-s-qəm,*
 Муса-PL LOC-лежать-RE-PST-PL но сам Муса LOC-сидеть-NEG
qale-m kʷ-a
 город-OBL идти-PST
 ‘Семья Мусы уже легла спать, но самого Мусы (дома) нет, он уехал в город.’

Вероятно, с ассоциативной формой в кабардино-черкесском языке связано регулярное обозначение фамилий формами, образованными суффиксом *-xe* от собственных имен. В литературном языке и в этом случае регулярно отсутствуют падежные маркеры [Кумахов 1971, с. 5–6], но в кубанском диалекте – как и в случае ассоциативной множественности – появление падежных маркеров факультативно:

- (18) *mjerjem.xe(-m) ze-xe-s ja-ʃə-n-wə*
 Меремовы-OBL REC.IO-LOC-сидеть 3PL.ERG-делать-MOD-ADV
r-a-χʷə-h-a
 LOC-3PL.ERG-случаться-уносить-PST
 ‘Меремовы решили устроить вечеринку.’

- (19) *anzwer ja-ze.xe.la.č'e-č'e e qʷəšhe.xe(-m)-č'e kʷ-a*
 Анзор POSS-строение-INS Кушховы-OBL-INS идти-PST
 'Анзор фигурой пошел в Кушховых.'

4. Конструкция с показателем *səme*

Кабардино-черкесский язык принадлежит к языкам со специализированным аналитическим показателем ассоциативной множественности. Таким маркером является служебное слово *səme* (ср. пример 2). В литературе обычно сообщается, что этот показатель присоединяется к (личным) именам собственным [Турчанинов, Цагов 1940, с. 53; Кумахов 1971, с. 25; Кабардино-черкесский язык 2006, с. 103; Жилетежев 2010]⁷, хотя, как мы увидим ниже, в кубанском диалекте это ограничение не работает. Более того, у *səme* обнаруживаются и другие особенности, указывающие, что перед нами совсем другое явление, нежели простая суффиксальная ассоциативная конструкция.

Формальные особенности. Показатель *səme* присоединяется к полной именной группе с фиксированным косвенным падежом – можно сказать, что падеж этот управляется показателем ассоциативной множественности. Так как падежные показатели на именах собственных в кабардино-черкесском языке регулярно отсутствуют, чтобы проиллюстрировать это, приходится обратиться к примерам с нарицательными существительными:

- (20) *s-j-ane-s-j-ade-xe-r ja-χʷen-a*
 1SG.IO-POSS-мать-1SG.IO-POSS-отец-PL-ABS 3PL.IO+DAT-ругать-PST
s-jə-padrugə-m səme
 1SG.IO-POSS-подруга-OBL APL
 'Мои родители ругали мою подругу и тех, кто с ней.'
- (21) *mə je.fa.kʷe-m səme d-wəž'eχʷ-a*
 этот пьяница-OBL APL 1PL.ERG-скучать-PST
 'Компания этого пьяницы уже нам надоела.'

⁷ В [Кабардино-черкесский язык 2006, с. 364] сообщается, что аналитический показатель ассоциативности «[с]очетается только с нарицательным именем человека или с вопросительным местоимением *xəm?* “кто”». Поскольку все примеры на сочетание показателя с именами далее включают имена собственные, в данном случае можно думать об опечатке.

Как показывают приведенные примеры, конструкция с *səme* употребляется и в позиции, в которой существительные обычно маркируются косвенным падежом, как в примере (20), где *sjəpadrugem səme* выступает в качестве непрямого объекта при глаголе 'ругать', и в позиции, предполагающей маркирование существительного показателем абсолютива (ср. пример (21) – в качестве пациенса переходного глагола). В отличие от обычных именных групп в этих позициях, конструкция с *səme* не присоединяет показатели ядерных падежей.

Показатели неядерных падежей к *səme* присоединяются:

- (22) *sjəme* *səme* *qə-šə-r-je-ka-ž'-wə* *zare* *səme-č' e*
 Сима APL DIR-LOC-LOC-DAT-CAUS- Зара APL-INS
 отправляться-ADV
jə-wəxə-ž'-wə *psew-m-jə* *q-jə-ž'e*
 LOC-заканчивать-RE-ADV все-OBL-ADD DIR-3PL.IO+DAT-звать
 'Зови всех, начиная с Симы и ее семьи и заканчивая Зарой с ее компанией.'
- (23) *mew-b-dje* *šə-t-xe-r* *zawər* *səm-wə* *pe.re?*
 тот-OBL-APUD LOC-стоять-PL-ABS Заур APL-ADV Q
 'Те, что там стоят, – это Заур со своей компанией?'

Другая особенность *səme* состоит в том, что этот показатель может следовать за сочинительной конструкцией, обозначающей несколько элементов множества. Так, в (24) эксплицитно представлена референция к двум элементам множества⁸. Ассоциативная конструкция, образованная от формы множественного числа вроде (25), признается недопустимой.

- (24) *jəjəne-re* *rwəslan-re* *səme* *škola-r*
 Ирина-COORD Руслан-COORD APL школа-ABS
rjəmont *ja-š-a*
 ремонт 3PL.ERG-делать-PST
 'Ирина, Руслан и те, кто с ними, отремонтировали школу.'

⁸ Аналогичный пример сочинительной конструкции в сочетании с *səme*, по-видимому взятый из фольклорных текстов, приводит Р. Матасович в своем обзоре кабардино-черкесской грамматики [Matasović 2010, p. 18]. См. также [Кабардино-черкесский язык 2006, с. 103].

- (25) *a χəʒ'ebz-xe-t səme wə-rja-b'əs-wə
 тот девушка-PL-OBL APL 2SG.ABS-3PL.IO+POSS-спутник-ADV
 wə-s-ʃ'əpsə-ne-qəm
 2SG.ABS-1SG.ERG-пускать-FUT-NEG
 (Ожид.: 'Я тебя с компанией этих девушек не пущу.')

Несмотря на то что в [Даниэль 2000] не отрицается возможность эксплицитного выражения более одного элемента множества в ассоциативных конструкциях, автор говорит, что убедительный пример такой структуры встречается только в конструкциях с личными местоимениями множественного числа. Кабардино-черкесские конструкции типа (24), однако, демонстрируют, что существует по крайней мере один язык, где обычные сочинительные группы могут образовывать ассоциативы.

Функционирование. В отличие от ассоциативного показателя *-xe*, *səme* не обнаруживает узких ограничений на тип фокального элемента. Именные группы, с которыми сочетается *səme*, включают группы, образованные указательными местоимениями (26)⁹, именами собственными (27), терминами родства (28), другими нарицательными одушевленными группами, обозначающие людей (29), и даже (хотя и не у всех наших консультантов) именные группы, не обозначающие людей, – одушевленные (30) и неодушевленные (31):

- (26) a-bə.m səme s-ja-psele-b-a-t
 тот-OBL APL 1SG.ABS-3PL.IO+DAT-говорить-PST-PST-RS
 'Я поговорил с ней и ее компанией.'
- (27) sase səme neg'əfe-xe-t
 Саса APL веселый-PL-IPF
 'Саса и те, кто с ней, были очень веселыми.'

⁹ Зафиксировано употребление *səme* с вопросительным местоимением:

- (i) xet səme a w-jə-b'əs-a-te-r?
 кто səme тот 2SG.IO-POSS-спутник-PST-RS-ABS
 'Кто (мн. ч.) это с тобой был?'

Не исключено, что здесь представлена не ассоциативная конструкция, а конструкция выражения группы (как в примере (33) далее).

- (28) *s-jə-tjotje-m* *səme* *nəpawə-r*
 1SG.IO-POSS-тётя-OBL APL ребёнок-ABS
ja-p-a
 3PL.ERG-воспитывать-PST
 ‘Моя тётя и те, кто с ней, воспитали ребёнка.’
- (29) *ʃale-m* *səme* *qale-m* *qʷ-a*
 парень-OBL APL город-OBL идти-PST
 ‘Парень и те, кто с ним, уехали в город.’
- (30) *s-jə-melə-m* *səme* *qʷed-a*
 1SG.IO-POSS-баран-OBL APL исчезать-PST
 ‘Мой баран и другие (животные) пропали.’
- (31) *a* *skripke-m* *səme* *qʷed-a*
 тот скрипка-OBL APL исчезать-PST
 ‘Та скрипка и другие (предметы) пропали.’

При одушевленном фокальном элементе, называющем человека, конструкция может обозначать множество, объединяющее фокальный референт и его родственников, друзей или знакомых, а также случайную группу, членом которой оказался фокальный референт. При одушевленном фокальном элементе, не обозначающем человека, и неодушевленном фокальном элементе также подразумевается не всегда жестко определенная группа. Таким образом, конструкция с *səme* допускает все три интерпретации, описанные в [Daniel, Moravcsik 2005].

Впрочем, как показывает следующий пример, использование *səme* не ограничивается конструкциями ассоциативной множественности:

- (32) *txələm* *səme* *ten-č'-jə* *ʃ-a-ʃe*
 книга-OBL *səme* где-INS-ADD LOC-3PL.ERG-продавать
 ‘Учебники (букв. книги) и тому подобное продают везде.’

В (32) *txələm* ‘книга’ не имеет уникального референта, «вокруг» которого строится множество. Группа, образованная *səme*, не указывает на какое-либо ограниченное множество и может описывать неоднородный класс предметов, включающий книги,

тетради и т. п., выделяемый лишь по функциональному сходству. Фактически конструкция с *səme* здесь выражает симилятивную множественность, а не ассоциативную множественность в строгом смысле. С учетом этого *səme* нельзя описывать как исключительно ассоциативный показатель, поскольку набор его функций шире. В то же время примеры вроде (32) принимаются лишь частью наших консультантов и, вероятно, представляют собой инновацию.

Заметим, что *səme* используется и в конструкциях, в которых эксплицировано все множество, – после перечисления всех участников группы (ср. [Кабардино-черкесский язык 2006, с. 103]):

- (33) *wəne.ze.ʔ^wə.xə-m* *qə-z-de-ʔepəq^w-a-xe-r* *anjəsjet*,
 уборка.дома-OBL DIR-1SG.IO-COM-помогать-PST-PL-ABS Анисет
- nwərijəjet*, *zwərijəjet* *səme*
 Нуриет Зуриет *səme*
- ‘В уборке дома мне помогли (букв. те, кто помогал мне в уборке дома) Анисет, Нуриет и Зуриет.’

5. От кабардино-черкесского языка к типологии ассоциативности

Итак, в кабардино-черкесском языке имеется две конструкции, претендующие на статус ассоциативных. При этом они разительно отличаются друг от друга – и особенности этих конструкций не всегда можно описать в терминах, принятых в типологии ассоциативной множественности. В связи с этим ниже предлагаются дополнительные параметры, которые следует учитывать при обсуждении типологии ассоциативности.

1. Прежде всего отметим недостаточность иерархии (5) для описания специфики конструкции с *səme* – ее исключительно широкой сферы употребления, отличающей ее от аналогичных конструкций (как от конструкции с *-xe*, так и от конструкций ассоциативной множественности во многих других языках). Эта проблема решается, если мы допускаем, что явления, описываемые через шкалу одушевленности, отражают не одну иерархию, а несколько. В [Croft 2002] противопоставляются четыре независимых, хотя и связанных, измерения:

- (34) а. Лицо: Первое, второе > третье
- б. Референциальный тип вершины именной группы:
 Местоимение > имя собственное > [термин родства] >
 > нарицательное имя
- в. Одушевленность: Человек > Не-человек (одушевленный) >
 > Неодушевленный
- г. Определенность: Определенная именная группа >
 > Неопределенная референтная именная группа >
 > Нереферентная именная группа

Конструкции ассоциативной множественности могут изучаться с точки зрения каждой из этих иерархий. Очевидно, что они всегда тяготеют к верху иерархий (34). В то же время языки ограничивают ассоциативные конструкции точками на разных иерархиях. Например, мы можем говорить о языках, в которых ассоциативное значение возникает только у местоимений первого и второго лица в соответствии с иерархией (34а), о языках, ограничивающих ассоциативные конструкции местоимениями и именами собственными (34б), о языках, требующих, чтобы фокальный элемент описывал человека или по крайней мере имел одушевленного референта (34с). Для кабардино-черкесской аналитической ассоциативной конструкции важно, что фокальный элемент не просто бывает неодушевленным, но и не обязательно должен быть определенным, то есть способен занимать достаточно низкие позиции в иерархии (34д). При появлении неопределенного (вероятно, нереферентного или родового) фокального референта как раз и наблюдается использование конструкции с *səte* для выражения симилятивной множественности, как в (32). В этом смысле симилятивная множественность является своего рода продолжением ассоциативной множественности, так что не очевидно, что эти два значения должны быть строго разграничены.

2. Однако и иерархий (34) для описания кабардино-черкесской ассоциативной множественности недостаточно: помимо прочего мы должны отразить, что *səte* присоединяется к сочинительным конструкциям. Мы предполагаем, что это свойство отражает степень индивидуализованности фокального элемента в соответствии с иерархией (35):

- (35) Единственное число > Перечисление элементов >
 > Аддитивное множественное число

Установление референции группы при выражении ассоциативной множественности требует, чтобы точка отсчета для этой референции (фокальный элемент) была четко установлена, противопоставлена другим объектам, другими словами — индивидуализована. Единственный, уникальный элемент обладает высокой степенью индивидуализации по сравнению с любым множеством. Но четко заданное множество (с перечисленными элементами) лучше противопоставлено прочему, чем нечетко заданное множество, определяемое исключительно через характеристики его элементов (как в случае аддитивной множественности). Следовательно, сочинительное выражение множественности действительно должно легче выступать в качестве фокального элемента, чем формы обычного аддитивного множественного числа.

3. Наконец, отдельный параметр, существенный для типологии выражения множественности, составляет морфосинтаксическая автономность элемента, от которого образуется конструкция множественного числа. Мы представляем ее следующим образом¹⁰:

- (36) Неавтономная основа/часть слова > Полное слово >
 > Сложная группа с внутренней подчинительной структурой >
 > Сложная группа с внутренней сочинительной структурой

Постулируя эти четыре категории, мы исходим из допущения, что в языках в норме противопоставляются морфологические и синтаксические правила. Неавтономная основа — единица, неподвластная синтаксическим правилам, полное слово — минимальная единица, с которыми они могут работать, сложная группа — группа, состоящая из нескольких слов¹¹.

Показатели ассоциативной множественности регулярно присоединяются к словам и сложным именным группам, даже если эти показатели удобно описывать как аффиксы. Ср. (37), где посесивное определение относится скорее не ко всему обозначаемому множеству, а лишь к фокальному референту. Здесь кажется естественным признать, что маркер ассоциативной множественности присоединяется к целой именной группе:

¹⁰ Хотя (36) приводится в виде иерархии, для наших целей направленность этой иерархии не принципиальна.

¹¹ Мы намеренно не затрагиваем здесь вопрос о семантической композиции. Типы (36) во многих случаях могут иметь корреляты в семантических типах (например, в наличии/отсутствии спецификации по тем или иным категориям, включая референцию), но мы предполагаем, что эта корреляция не абсолютна.

(37) Тантынский даргинский [Сумбатова, Ландер 2014, с. 51]

[*ʒeʃla* *t:at:i-qali* *b-ač'-ib*

ты:GEN отец-APL Н.PL-приходить:PFV-PRET

‘Пришёл твой отец со своими друзьями.’

Конструкции ассоциативной множественности тяготеют скорее к середине иерархии (36). Тем не менее в кабардино-черкесском языке показатель *səme* способен присоединяться как к подчинительной, так и к сочинительной сложной группе. Более того, в результате деморфологизации именного комплекса суффикс *-xe* появляется и на сложной группе, в том числе и в ассоциативном значении (см. пример 16).

6. Заключение

Итак, для описания конструкций ассоциативной множественности в кабардино-черкесском языке недостаточно стандартных параметров, используемых в типологии ассоциативности. В связи с этим мы предложили, во-первых, использовать вместо так называемой шкалы одушевленности несколько шкал индивидуализации – лицо, тип именной группы, одушевленность, референциальный статус, а кроме того, дополнительные параметры, включающие числовую индивидуализацию и формальный тип элемента, от которого образуется конструкция.

Эти параметры важны не только для ассоциативной множественности, но и для других типов множественной референции. В частности, выражение аддитивной множественности может быть привязано не только к одушевленности, но и, например, к определенности: существуют языки, в которых противопоставление по числу реализуется только для определенных или только для референтных именных групп [Corbett 2001, pp. 278–279]. Не менее релевантна для аддитивной множественности и шкала (36). Формы аддитивной множественности, наблюдаемые в языках средневропейского стандарта, обычно образуются от неавтономных основ. В то же время в других языках показатели числа (аналитические или синтетические) присоединяются к единицам, свободно функционирующим на синтаксическом уровне, – словам или сложным группам.

Сочинительное и комитативное выражение множественности также может быть привязано к разным параметрам. Так, нереферентные группы иногда сочиняются хуже, чем референтные, а оду-

шевленность порой влияет на выбор сочинительной конструкции [Haspelmath 2004, p. 12] и является типичной (хотя и не обязательной) для комитативной конструкции [Архипов 2009, с. 96]. Сочинительное и комитативное выражение множественности не типично для основ и более типично для слов и сложных групп, но оно наблюдается и при неавтономных элементах – например, в так называемых компаундах *dvandva*.

Для установления четких зависимостей между семантически и формальными типами лингвистических единиц и возможностью/вероятностью тех или иных конструкций выражения множественности требуются дополнительные исследования на большом типологическом материале. Тем не менее отрадно думать, что ассоциативная множественность вписывается в схемы, не ограниченные ею самой.

Благодарности

Авторы признательны за обсуждение кабардино-черкесского материала участникам Адыгейских экспедиций НИУ ВШЭ и РГГУ 2015 и 2016 гг. (а. Ходзь и Блечепсин) и в особенности П.М. Аркадьеву, за обсуждение общих вопросов типологии ассоциативности – М.А. Даниэлю, за комментарии по ранней версии статьи – анонимному рецензенту. Ответственность за возможные ошибки лежит на авторах.

Acknowledgments

The authors are grateful to the participants of the Circassian field trips to the villages of Hodz and Blechepsin (organized by HSE University and RSUH in 2015 and 2016) and in particular to Peter Arkadiev for a discussion of Kabardian data. We also thank Michael Daniel for discussions of the typology of associative plurals and an anonymous reviewer for comments on an early version of this paper. All errors are ours.

Сокращения

ABS – абсолютив, ADD – аддитив, ADV – адвербиалис, APL – ассоциативная множественность, APUD – ‘возле’, CAUS – каузатив, COM – комитатив, COORD – сочинение, COP – копула, DAT – дативный преверб, DCL – декларатив, DIR – директив, EMPH – эмфаза, ERG – префикс эргативной серии, FUT – будущее время, GEN – генитив, H – ‘человек’, INS – инструменталис, IO – непрямой объект, IPF – имперфект, LOC – локативный преверб, LOCANAPH – локативное местоимение, MOD – модальный показатель, NEG – отрицание, OBL – косвенный падеж, PFV – перфектив, PL – множественное число,

POSS – посессив, PRET – претерит, PRF – этимологический префикс при глаголе ‘говорить’, PST – прошедшее время, Q – показатель вопроса, RE – рефактив, REC – реципрок, REL – относительный префикс, RS – ретроспективный сдвиг, SG – единственное число, TRANS – транслатив.

Литература

- Архипов 2009 – *Архипов А.В.* Типология комитативных конструкций. М.: Знак, 2009.
- Барулин 1980 – *Барулин А.Н.* Категория числа в местоимениях // Исследования в области грамматики и типологии языков / Ред. В.М. Андрищенко. М.: МГУ, 1980. С. 145–164.
- Богданов 1968 – *Богданов В.Н.* Особый случай диалектного согласования сказуемого с подлежащим по смыслу и категория представительности // Научные доклады высшей школы. Филологические науки. 1968. № 4. С. 68–75.
- Грамматика 1957 – Грамматика кабардино-черкесского литературного языка. М.: Изд. АН СССР, 1957.
- Грамматика 1970 – Грамматика кабардино-черкесского литературного языка. Ч. 1: Фонетика и морфология. М.: Наука, 1970.
- Даниэль 1999 – *Даниэль М.А.* Семантические типы ассоциативности (на материале багвалинского языка) // Типология и теория языка: от описания к объяснению: К 60-летию А.Е. Кибрика / Ред. Е.В. Рахилина, Я.Г. Тестелец. М.: Языки русской культуры, 1999. С. 362–370.
- Даниэль 2000 – *Даниэль М.А.* Типология ассоциативной множественности: Дис. ... канд. филол. наук. М.: РГГУ, 1999.
- Жилетежев 2010 – *Жилетежев Х.Ч.* Особенности выражения числа собственных имен существительных в кабардино-черкесском языке // Известия СОИГСИ. Школа молодых ученых. 2010. Вып. 4. С. 148–155.
- Кабардино-черкесский язык 2006 – Кабардино-черкесский язык: монография: в 2 т. Нальчик: Эль-Фа, 2006. Т. 1.
- Кумахов 1969 – *Кумахов М.А.* Кубанский диалект // Очерки кабардино-черкесской диалектологии / Ред. М.А. Кумахов. Нальчик, 1969.
- Кумахов 1971 – *Кумахов М.А.* Словоизменение адыгских языков. М.: Наука, 1971.
- Сердобольская, Кузнецова 2009 – *Сердобольская Н.В., Кузнецова Ю.Л.* Двойное падежное маркирование: уникальный случай адыгейского языка // Аспекты полисинтетизма: очерки по грамматике адыгейского языка / Ред. Я.Г. Тестелец. М.: РГГУ, 2009. С. 166–200.
- Сумбатова, Ландер 2014 – *Сумбатова Н.Р., Ландер Ю.А.* (при уч. М.Х. Мамаева). Даргинский говор селения Танты: грамматический очерк, вопросы синтаксиса. М.: Языки славянской культуры, 2014.
- Турчанинов, Цагов 1940 – *Турчанинов Г., Цагов М.* Грамматика кабардинского языка. I. М.; Л.: Изд. АН СССР, 1940.
- Bagirokova et al., to appear – Bagirokova, I., Lander, Yu. and Phelan, P. Number in West Circassian // Number in the world's languages / P. Acquaviva, M. Daniel (eds.). Berlin: Mouton de Gruyter.

- Corbett 2001 – *Corbett G.G.* Number. Cambridge: Cambridge University Press, 2001.
- Croft 2002 – *Croft W.* Typology and universals. 2nd ed. Cambridge: Cambridge University Press, 2002.
- Daniel, Moravcsik 2005 – *Daniel M., Moravcsik E.* The associative plural // The world atlas of language structures / M. Haspelmath, M.S. Dryer, D. Gil, B. Comrie (eds.). Oxford: Oxford University Press, 2005. P. 150–153.
- Haspelmath 2004 – *Haspelmath M.* Coordinating constructions: An overview // *Coordinating constructions* / M. Haspelmath (ed.). Amsterdam: John Benjamins, 2004. P. 3–39.
- Hellwig 2011 – *Hellwig B.* A grammar of Goemai. Berlin: Mouton de Gruyter, 2011.
- Lander 2017 – *Lander Yu.A.* Nominal complex in West Circassian: between morphology and syntax // *Studies in Language*. 2017. Vol. 41. no. 1. P. 76–98.
- Matasović 2010 – *Matasović R.* A short grammar of East Circassian (Kabardian). 9th version. Zagreb, 2010.
- Moravcsik 2003 – *Moravcsik E.* A semantic analysis of associative plurals // *Studies in Language*. 2003. Vol. 27. P. 469–503.
- Partee 1997 – *Partee B.H.* Genitives – A case study // *The handbook of logic and language* / J. Van Benthem, A. ter Meulen (eds.). Amsterdam: Elsevier, 1997. P. 464–470.
- Vassilieva 2005 – *Vassilieva M.B.* Associative and pronominal plurality: Ph.D. Thesis. New York: Stony Brook University, 2005.

References

- Arkhipov, A.V. (2009), *Tipologiya komitativnykh konstruktivnykh* [A typology of comitative constructions], Znack, Moscow, Russia.
- Bagirokova, I., Lander, Yu. and Phelan, P. (To appear), “Number in West Circassian”, in Acquaviva, P. and Daniel, M. (eds.), *Number in the world's languages*, Mouton de Gruyter, Berlin, Germany.
- Barulin, A.N. (1980), “The category of number in pronouns”, in Andryushchenko, V.M. (ed.), *Issledovaniya v oblasti grammatiki i tipologii yazykov* [Studies in grammar and linguistic typology], Moscow State University, Moscow, Russia, pp. 145–164.
- Bogdanov, V.N. (1968), “A peculiar case of dialectal subject-predicate semantic concord and the category of representative plural”, *Research papers of higher school, Filologicheskoye nauki*, vol. 4, pp. 68–75.
- Corbett, G.G. (2001), *Number*, Cambridge University Press, Cambridge, UK.
- Croft, W. (2002), *Typology and universals*, 2nd ed., Cambridge University Press, Cambridge, UK.
- Daniel, M.A. (1999), “Semantic types of associativity (Bagvalal data)”, in Rachilina, Ye.V. and Testelets, Ya.G. (eds.), *Tipologiya i teoriya yazyka: ot opisaniya k ob'yasneniyu* [Typology and theory of language: from description to explanation. Toward the sixtieth anniversary of A.E. Kibrik], *Yazyki russkoi kul'tury*, Moscow, Russia, pp. 362–370.

- Daniel, M.A. (2000), *Tipologiya assotsiativnoy mnozhestvennosti* [The Typology of associative plurality], Ph.D. Thesis, RSUH, Moscow, Russia.
- Daniel, M. and Moravcsik, E. (2005), "The associative plural", in Haspelmath, M., Dryer, M.S., Gil, D. and Comrie, B. (eds.), *The world atlas of language structures*, Oxford University Press, Oxford, UK, pp. 150–153.
- Grammatika kabardino-cherkesskogo literaturnogo yazyka* [A Grammar of Standard Kabardian] (1957), Academy of Sciences of the USSR, Moscow, USSR.
- Grammatika kabardino-cherkesskogo literaturnogo yazyka* [A Grammar of Standard Kabardian] (1970), part 1: Phonetics and Morphology, Nauka, Moscow, USSR.
- Haspelmath, M. (2004), "Coordinating constructions: An overview", in Haspelmath, M. (ed.), *Coordinating constructions*, John Benjamins, Amsterdam, Netherlands, pp. 3–39.
- Hellwig, B. (2011), *A grammar of Goemai*, Mouton de Gruyter, Berlin, Germany.
- Kabardino-cherkesskii yazyk* [Kabardian language] (2006), in 2 vols., vol. 1, El'-Fa, Nalchik, Russia.
- Kumakhov, M.A. (1969), "The Kuban dialect", in Kumakhov, M.A. (ed.) *Ocherki kabardino-cherkesskoy dialektologii* [Essays in Kabardian dialectology], Nalchik, USSR.
- Kumakhov, M.A. (1971), *Slovoizmeneniye adygsikh yazykov* [Inflection in Circassian languages], Nauka, Moscow, USSR.
- Lander, Yu.A. (2017), "Nominal complex in West Circassian: between morphology and syntax", *Studies in Language*, vol. 41, no. 1, pp. 76–98.
- Matasović, R. (2010), *A short grammar of East Circassian (Kabardian)*, 9th version, Zagreb, Croatia.
- Moravcsik, E. (2003), "A semantic analysis of associative plurals", *Studies in Language*, vol. 27, pp. 469–503.
- Partee, B.H. (1997), "Genitives – a case study", in van Benthem, J. and ter Meulen, A. (eds.), *The handbook of logic and language*, Elsevier, Amsterdam, Netherlands, pp. 464–470.
- Serdobol'skaya, N.V. and Kuznetsova, Yu.L. (2009), "Double case marking. A peculiar phenomenon in Adyghe", in Testelefs, Ya.G. (ed.), *Aspekty polisintetizma: ocherki po grammatike adygeiskogo yazyka* [Aspects of polysynthesis. Essays on the grammar of the Adyghe language], RSUH, Moscow, Russia, pp. 166–200.
- Sumbatova, N.R., Lander, Yu.A., with Mamaev, M.H. (2014), *Darginskii govor seleniya Tanty: grammaticheskii ocherk, voprosy sintaksisa* [The Dargwa variety of the Tanti village. A grammatical sketch, aspects of syntax], Yazyki slavyanskoi kul'tury, Moscow, Russia.
- Turchaninov, G., Tsagov, M. (1940), *Grammatika kabardinskogo yazyka* [A Grammar of Kabardian], vol. 1, Academy of Sciences of the USSR, Moscow-Leningrad, USSR.
- Vassilieva, M.B. (2005), *Associative and pronominal plurality*, Ph.D. Thesis, Stony Brook University, New York, USA.
- Zhiletezhov, Kh.Ch. (2010), "On expressing number in proper nouns in the Kabardian", *Izvestiya SOIGSI. Shkola molodykh uchenykh*, vol. 4, pp. 148–155.

Информация об авторах

Ирина Г. Багирокова, кандидат филологических наук, доцент, Институт языкознания РАН, Москва, Россия; 125009, Россия, Москва, Большой Кисловский пер., д. 1 стр. 1;

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия; 105066, Россия, Москва, Старая Басманная ул., д. 21/4; ibagirokova@yandex.ru

Алина И. Баранова, независимый исследователь, Кельн, Германия; baranovaalinka@gmail.com

Юрий А. Ландер, кандидат филологических наук, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия; 105066, Россия, Москва, Старая Басманная ул., д. 21/4; yulander@yandex.ru

Information about the authors

Irina G. Bagirokova, Cand. of Sci. (Philology), Institute of Linguistics RAS, Moscow, Russia; bld. 1/1, Bolshoi Kislovskii Lane, Moscow, Russia, 125009;

HSE University, Moscow, Russia; bld. 21/4, Staraya Basmannaya Str., Moscow, Russia, 105066; ibagirokova@yandex.ru

Alina I. Baranova, independent researcher, Cologne, Germany; baranovaalinka@gmail.com

Yury A. Lander, Cand. of Sci. (Philology), HSE University, Moscow, Russia; bld. 21/4, Staraya Basmannaya Str., Moscow, Russia, 105066; yulander@yandex.ru

Синтаксическая сложность академического текста:
корпусное исследование работ русскоязычных
студентов, изучающих английский язык,
в сравнении с текстами профессиональных авторов

Ольга И. Виноградова

*Национальный исследовательский университет «Высшая школа экономики»
Москва, Россия, olgavinogr@gmail.com*

Елизавета А. Смирнова

*Национальный исследовательский университет «Высшая школа экономики»
Пермь, Россия, smelizabetha@yandex.ru*

Анна В. Виклова

*Национальный исследовательский университет «Высшая школа экономики»
Москва, Россия, annakisseleva@gmail.com*

Ирина М. Пантелеева

Москва, Россия, impanteleyeva@gmail.com

Аннотация. В статье анализируются данные получения автоматизированной оценки синтаксической сложности академического текста на базе размеченных частеречной разметкой удовлетворяющих требованиям сравнимости корпусов двух экспертных и двух учебных – в двух тематических областях. Исследование проводилось с использованием двух различных компьютерных инструментов – парсера UDPipe и комплекса программ AntConc. Были сделаны следующие выводы: экспертные академические тексты показывают высокую клаузальную и фразовую сложность, уровень которой выше в текстах по менеджменту, нежели по экономике. По сравнению с текстами студентов клаузальная сложность текстов профессионалов существенно выше, в то время как параметры фразовой сложности были почти одного уровня или даже незначительно ниже, чем у студентов. Наиболее значимые различия наблюдаются в употреблении генитивов с предлогом *of*, нормализованная частотность которых в учебных корпусах превышает оную в профессиональных статьях примерно в 1,7 раз в обеих областях. Результаты противоречат выводам предыдущих исследований о том, что с повышением уровня владения языком клаузальная сложность снижается, а фразовая растет.

Ключевые слова: синтаксическая сложность, фразовая сложность, клау-
зальная сложность, академический английский, компьютерные инстру-
менты, учебные корпуса

Для цитирования: Виноградова О.И., Смирнова Е.А., Виклова А.В., Пан-
телеева И.М. Синтаксическая сложность академического текста: корпус-
ное исследование работ русскоязычных студентов, изучающих англий-
ский язык, в сравнении с текстами профессиональных авторов // Вестник
РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020.
№ 7. С. 107–129. DOI: 10.28995/2686-7249-2020-7-107-129

Syntactic Complexity of Academic Text: A Corpus Study of Written Production by Learners of English with Russian L1 in Comparison with Expert Texts of English Authors

Olga I. Vinogradova

HSE University, Moscow, Russia, olgavinogr@gmail.com

Elizaveta A. Smirnova

HSE University, Perm, Russia, cmelizaveta@yandex.r

Anna V. Viklova

HSE University, Moscow, Russia, annakisseleva@gmail.com

Irina M. Panteleeva

Moscow, Russia, impanteleyeva@gmail.com

Abstract. The paper analyses data of automated evaluation of syntactical complexity of scholarly texts from four comparable corpora – two of expert texts and two of learner texts – in two subject areas. The research applied two different computer tools – UDPipe parser and AntConc set of programs. The following conclusions were made:

– Expert scholarly texts demonstrated high clausal and phrasal complexity, and its level was higher in Business Studies than in Economics.

– Clausal complexity in expert texts was much higher than in student texts, while phrasal complexity was at about the same level as, or even slightly lower than that of students.

The largest difference was attested for genitive construction with preposition *of*, which was 1.7 times as frequent in student texts in both subject areas as in expert texts. The results contradict the conclusion of the prior research that with the growth in proficiency clausal complexity decreases, while phrasal complexity grows.

Keywords: syntactic complexity, phrasal complexity, clausal complexity, academic English, computer tools, learner corpora

For citation: Vinogradova, O.I., Smirnova, E.A., Viklova, A.V. and Panteleeva, I.M. (2020), “Syntactic complexity of academic text: A corpus study of written works by Russian-speaking students studying English in comparison with expert texts of English authors”, *RSUH/RGGU Bulletin. “Literary Theory. Linguistics. Cultural Studies” Series*, no. 7, pp. 107-129, DOI: 10.28995/2686-7249-2020-7-107-129

Введение

Исследование сложности академического текста входит в число приоритетных направлений современной корпусной лингвистики, о чем свидетельствует большое количество опубликованных работ, посвященных данному вопросу (см., например, [Biber et al. 2011]; [Biber, Gray 2016]; [Crossley, McNamara 2014]; [Lu 2011]). Интерес к данной теме обусловлен несколькими факторами. Во-первых, глобализация всех социальных сфер не обошла стороной и образование, в результате чего повышение качества преподавания академического английского языка, *lingua franca* научного сообщества, стало важной задачей, которую ставят перед собой многие вузы мира. Во-вторых, определение сложности текста и объективация данного параметра дают исследователям возможность обрабатывать и анализировать академический текст с помощью компьютерных инструментов. Полученные данные могут в дальнейшем применяться в научной деятельности и педагогической практике.

В данной статье мы попытаемся ответить на следующие исследовательские вопросы:

1. Какими показателями сложности обладает академический текст, написанный профессиональным автором? Как эти показатели различаются в текстах по разным дисциплинам?
2. Есть ли значимые различия между текстами, написанными экспертами, и текстами студентов, изучающих эту же дисциплину? Какие параметры сложности демонстрируют наибольшие различия?
3. Какие инструменты автоматической оценки сложности оказываются более эффективными?

Обзор предшествующих исследований

По определениям в [Biber, Gray 2016, p. 141], грамматическая сложность научного текста включает два аспекта: фразовую и клаузальную сложность. Фразовая сложность представляет собой степень концентрации существительных и прилагательных, функционирующих в качестве определений существительного, тогда как клаузальная сложность понимается как степень концентрации придаточных предложений в функции либо именной части сказуемого, либо обстоятельственных придаточных. Как показано в [Staples et al. 2016], по мере развития у студентов навыков письменной речи фразовая сложность повышается, а клаузальная снижается.

Ряд работ посвящен изучению отдельных синтаксических конструкций, ассоциируемых с клаузальной сложностью. Например,

клефтовые конструкции с местоимением *it* исследуются в работах [Suárez, Álvarez 2010]; [Hinkel 2003]. Как указано в работе [Biber et al. 1999, p. 963], конструкции с *it* особенно характерны для научной прозы, поскольку они позволяют максимально точно сформулировать мысль, расставив нужные акценты. Особое внимание уделяется условным придаточным (см., например, [Swales, Feak 2004]; [Warchal 2010]; [Mazgutova, Kormoz 2015]). К. Уорчал [Warchal 2010] отмечает, что этот тип придаточных, выполняющий широкий спектр функций, играет важную роль в научных текстах, где используется для построения логической аргументации. Как сказано в пособии [Swales, Feak 2004, p. 260], в научных текстах ирреальные условные придаточные используются для критики работ предшественников. По мнению Байбера и соавторов [Biber et al. 1999], условные придаточные предваряют и развивают аргументацию. Относительные придаточные также характерны для академического письма (см., например, [Byrnes, Sinicrope 2008]; [Biber et al. 1999]), и их частотность нередко служит показателем синтаксической сложности текста (например, [Jucker 1992]). Как отмечается в работе [Siermann et al. 2011], нефинитные клаузы, а именно причастные обороты, часто встречаются в научных текстах, где могут функционировать в качестве обстоятельственных придаточных в сопровождении союзов. Нефинитные клаузы отличаются компактностью, и их использование помогает уменьшить длину предложения, одновременно повышая его сложность (ср. *When considered from this angle, the difficulties appear insurmountable* ‘Если посмотреть с этой точки зрения, трудности кажутся непреодолимыми’¹ по сравнению с *When the difficulties are considered from this angle, they appear insurmountable* ‘Если посмотреть на трудности с этой точки зрения, они кажутся непреодолимыми’) [Siermann et al. 2011, p. 152]. О частотности причастных оборотов в письменном научном дискурсе см. также [Biber et al. 1999, p. 606]. Использование нефинитных клауз и парентетических конструкций для выражения отношения автора в научных работах, написанных испанскими и нидерландскими студентами продвинутого уровня, исследовалось в работе [Hannay, Caro 2008]. По наблюдениям авторов, эти конструкции чаще встречаются в текстах носителей испанского языка.

Некоторых исследователей интересовали не отдельные признаки грамматической сложности, а их группы. Так, Байбер и Грей [Biber, Gray 2016] рассматривали финитные и нефинитные обстоятельственные, изъяснительные и определительные клаузы как маркеры клаузальной сложности. По их данным, большинство

¹ Здесь и далее – перевод авторов.

этих грамматических конструкций более частотны в устной речи, нежели в письменном научном дискурсе. С. Стэйплз и соавторы [Staples et al. 2016] изучали в академических текстах различных дисциплин признаки клаузуальной сложности, такие как финитные обстоятельственные клаузы, придаточные с союзными словами на *wh-*, выступающие как именная часть сказуемого, изъяснительные придаточные с союзом *that* и союзы между однородными клаузами. Они пришли к выводу, что клаузуальная сложность более характерна для текстов из области искусств, гуманитарных и социальных наук и реже проявляется в естественнонаучных текстах.

Б. Грей [Gray 2013] составила список маркеров позиции автора и рассмотрела вариативность их использования в научных статьях, посвященных теоретическим, количественным и качественным исследованиям. Как оказалось, вариативность возникает под влиянием нескольких параметров, помимо различий между научными областями. К этим параметрам относятся используемые в статье подтверждающие данные, наличие или отсутствие данных, принадлежность работы к качественной, количественной или теоретической исследовательской парадигме. В том же русле Гарднер и соавторы [Gardner et al. 2018] проследили, как грамматические признаки объединяются в кластеры по четырем параметрам студенческих письменных текстов, а именно (i) сжатая информация о процедуре исследования или выражение позиции авторов по отношению к работам других исследователей; (ii) личная позиция авторов; (iii) описание возможных событий или произошедших событий и (iv) информационная плотность исследования. При этом авторы выделяли параметры, присутствовавшие в текстах различных дисциплин, жанров и уровней владения языком. Основываясь на результатах этих исследований, в работе [Nesi and Gardner 2019] утверждается, что клаузуальная сложность текста более характерна для гуманитарных дисциплин и более «разговорных» жанров. В первом случае она достигается за счет использования эпистемических обстоятельственных оборотов, существительных, выражающих отношение автора, и придаточных предложений с союзом *that*; во втором случае чаще используются глаголы, выражающие авторскую позицию.

В контексте усвоения иностранного языка синтаксическая сложность текста является средством определения уровня владения языком, как показано в работах [Foster, Skehan 1996], [Wolfe-Quintero et al. 1998], [Ortega 2003], [Lu 2011]. Синтаксическая сложность письменного текста на иностранном языке, как правило, рассматривается в связи с развитием языковой компетенции учащегося: например, какие синтаксические конструкции осваивают-

ся раньше, а какие позже [Hawkins 2001]; как синтаксическая сложность текста влияет на оценку его качества [Crossley, McNamara 2014]; как меняется репертуар синтаксических конструкций, которыми владеет учащийся, после интенсивного курса обучения [Mazgutova, Kormos 2015].

Несмотря на большое количество работ, посвященных различным параметрам сложности текста, насколько нам известно, ранее авторы не пытались сравнить эти параметры сразу в двух плоскостях: в различных дисциплинах и в текстах профессиональных авторов и учащихся. Кроме того, в нашем исследовании применяются различные подходы к изучению сложности текста, что позволит получить более полную картину данного явления, а также оценить эффективность выбранных методов.

Материалы и методы

Корпусы

Исследование проведено на материале четырех размеченных корпусов: двух экспертных и двух учебных.

Первый корпус текстов профессиональных авторов (примерно 694 тыс. слов) состоит из статей, опубликованных в рецензируемых журналах по менеджменту: *Journal of Management* (JOM) – 25 статей; *Journal of Management Studies* (JMS) – 20 статей; *Academy of Management Journal* (AMJ) – 17 статей (всего 62 статьи) в период с 2013 по 2016 г. Все источники – ведущие научные журналы с индексом Хирша 145, 119 и 227 соответственно (по данным за 2017 г.), что свидетельствует об их авторитетности в международном сообществе исследователей в области бизнеса и менеджмента.

Второй экспертный корпус из 679 тыс. слов включает статьи, опубликованные в рецензируемых журналах по экономике: *Quarterly Journal of Economics* (QJE) – 19 статей; *Journal of Financial Economics* (JFE) – 18 статей; *International Journal of Production Economics* (IJPE) – 19 статей (всего 56 статей) в период с 2013 по 2016 г. Индекс Хирша упомянутых журналов составляет соответственно 219, 206 и 141 (по данным за 2017 г.) – таким образом, это ведущие издания в данной области. Полный список статей, вошедших в экспертные корпуса по менеджменту и экономике, доступен по ссылке <https://docs.google.com/document/d/1GBkzdJ00hHqkT XSZv4HYVjAoQ5foQtY3pc6POqiaG58/edit?usp=sharing>.

Первый из учебных корпусов включает 58 текстов по менеджменту (130 тыс. слов), написанных студентами четвертого курса,

проходившими обучение по программе «Менеджмент» в Национальном исследовательском университете «Высшая школа экономики» в Перми в 2015 г. Это предварительное описание дипломной работы, написанное в рамках курса «Английский для академических целей». Каждый текст представляет собой краткое описание дипломного проекта на английском языке (примерно 2500 слов), включающее несколько разделов: введение, обзор литературы, методы, полученные или предполагаемые результаты, заключение. Все студенты — носители русского языка с уровнем владения английским примерно B2-C1.

В состав второго учебного корпуса (ок. 112 тыс. слов) входит 53 текста, написанных студентами-экономистами четвертого курса того же университета. Уровень владения английским у студентов-экономистов примерно такой же, что и у студентов программы «Менеджмент», и их родным языком также является русский. Исследовательские проекты по менеджменту и экономике имеют одинаковую структуру. Доступ к учебным корпусам возможен по следующей ссылке: <https://perm.hse.ru/bi/sfcr/corpora>

Методы исследования

Из отобранных текстов были удалены все таблицы, формулы, графики и ссылки.

На первом этапе исследования распределение показателей сложности в учебных и экспертных текстах анализировалось при помощи парсера UDPipe [Straka 2017]. В исследовании [Lyashevskaya, Panteleeva 2018], также посвященном работе с этим парсером, были выявлены критерии синтаксической сложности, которые основаны на синтаксических отношениях, устанавливаемых этим парсером. Показателями синтаксической сложности являются, например, среднее число обстоятельственных клауз, атрибутивных конструкций и относительных клауз в тексте. В случае несопоставимости текстов по длине в качестве метрик используются отношение числа тех же клауз к количеству предложений в тексте и отношение числа тех же клауз к количеству токенов.

На следующем этапе мы проверили частотность реализации параметров синтаксической сложности текста в исследуемых четырех корпусах. Подсчеты параметров на этом этапе были произведены с помощью программы AntConc [Anthony 2014], которая часто используется как инструмент автоматического анализа текстов при исследовании корпусных данных. В качестве объекта исследования были выбраны признаки фразовой и клаузуальной сложно-

сти научного текста, предложенные в работе [Staples et al. 2016]. К чертам фразовой сложности относится частотное распределение существительных; прилагательных, функционирующих в качестве определений (например, *empirical results*); существительных, функционирующих в качестве определений в позиции перед существительными (например, *stimulation results*); конструкции с предлогом *of* между двумя существительными (например, *the results of the analysis*). К клаузуальной сложности относится частотное распределение обстоятельственных, изъяснительных и относительных придаточных предложений. Мы изучали обстоятельственные придаточные следующих типов:

- условные клаузы с союзами *if, unless, in the event that, provided that*, например:
 - (1) *Unless the case can be made, there is little chance that the regulatory authorities will seek to change current practice.*
‘Если не удастся вынести этот вопрос на решение в суде, то маловероятно, что органы надзора попытаются изменить существующее положение вещей.’ (IJPE 2016-6);
- уступительные клаузы с союзами *although, even though, despite the fact that, in spite of the fact that*, например:
 - (2) *Even though the paper assumes a simple tracking function class, the proposed method is easily, readily extendible to a more complicated function with a much higher number of parameters.*
‘Хотя в работе рассматривается класс функций простого отслеживания, предложенный метод можно легко расширить на более сложные функции с гораздо большим числом параметров’ (IJPE 2014-1);
- временные клаузы с союзами *after, before, when, until, as soon as*, например:
 - (3) *After one component was completely varnished, the photocell system stopped paint-spraying until the subsequent piece along the chain conveyor reached the right position in the paint booth.*
‘После того как один компонент был полностью покрыт лаком, система фотоэлементов прекратила напыление краски до тех пор, пока следующий предмет на цепном конвейере не достиг нужного положения в покрасочной камере’ (IJPE 2015-1);
- клаузы причины с союзами *because, since*, например:
 - (4) *First, since the metal content of each mining block is not known, the associated financial value of a block is stochastic.*

'Во-первых, поскольку содержание металла в каждом шахтном блоке неизвестно, соответствующая финансовая ценность блока является стохастической величиной' (IJPE 2014-1);

- клаузы результата с союзом *so that*, например:
- (5) *This will require a broader more systemic approach to encouraging sustainable practices including investment and financing practices, **so that** all members of the dairy supply chain can co-operate and contribute to energy reduction.*

'Для этого потребуется более широкий и системный подход к внедрению устойчивых практик, в том числе инвестирования и финансирования, чтобы все участники цепочки поставок молочной продукции могли объединить усилия и внести свой вклад в снижение энергопотребления' (IJPE 2014-3);

- клаузы контраста с союзами *while, whereas*, например:
- (6) ***Whereas** previous work in the field has shown that customers value split lines when evaluating product quality... this study describes a means to quantify that preference in a useful way for manufacturers.*

'Хотя предшествующие исследования в данной области показали, что при оценке качества продукта покупатели ценят линии разьема... в этом исследовании описывается способ количественного измерения предпочтения, который будет полезен производителям' (IJPE 2015-2).

Выбор данных видов придаточных предложений обусловлен их высокой частотностью в академическом дискурсе, что убедительно показано в работе [Biber et al. 1999] и отмечено несколькими авторами в качестве типичной черты письменной речи людей, профессионально владеющих английским языком [Hinkel 2003]; [Siermann et al. 2011].

Результаты

При проведении первого этапа исследований авторы разработали программу на языке Python, которая для извлечения синтаксической и лексической информации использует модель english-ud-2.3 парсера UDPipe, а для извлечения морфологической информации — комплекс nltk пакета PorterStemmer. Используя эти средства, программа выполняет следующие задачи: 1) лемматизация; 2) стемминг; 3) выделение конструкций; 4) классификация

по оценкам. На основании анализа критериев сложности текста, о которых шла речь в разделе 3.2, были выбраны основные признаки для работы программы, и по ним производилась оценка сложности текстов. Программа прежде всего вывела подсчет результатов парсинга по пяти параметрам, приведенным в таблице 1. Заметим, что обстоятельственные придаточные клаузы включают в себя клаузы как с финитными, так и с нефинитными формами глагола. Число обстоятельственных, определительных и относительных клауз (первые три строки таблицы 1) является отношением суммарного числа соответствующих клауз в выборке корпуса к общему количеству текстов в выборке – тем самым это среднее количество соответствующих клауз на текст.

Используемые обозначения:

1. Тексты статей – это тексты статей из экспертных корпусов, а тексты студентов – это тексты предварительного описания на английском языке дипломной работы их учебных корпусов.

2. LL – логарифмическое правдоподобие (log likelihood). Данная функция используется для оценки статистической значимости разницы между показателями двух корпусов [Dunning 1993]. Вслед за Смитом [Smith 2009], значения логарифмической правдоподобности интерпретировались следующим образом: $LL \geq 6,63$ означает, что результат значим на уровне 1% ($p < 0,01$), т. е. вероятность того, что разница между показателями случайна, составляет 1%; $LL \geq 3,84$ означает, что результат значим на уровне 5% ($p < 0,05$).

3. %DIFF – размер эффекта. Данный показатель используется в работе для оценки размера разницы между нормализованной частотностью параметров сложности в двух корпусах и показывает пропорцию этой разницы. Отрицательные значения показывают, что параметр обладает более высокой нормализованной частотностью в справочном корпусе [Gabrielatos, Marchi 2012]. В нашем случае это означает, что частотность выше в экспертном корпусе, чем в учебном.

Как видно из таблицы, наибольшая разница между экспертными и учебными корпусами наблюдается в использовании определительных клауз, которые преобладают в текстах профессиональных авторов по экономике, и обстоятельственных клауз, доминирующих в экспертных текстах по менеджменту. Все значения параметров сложности оказались ниже в учебных корпусах, чем в экспертных. Однако следует отметить, что все различия между студенческими и экспертными корпусами оказались статистически незначимы.

Таблица 1

Средние на текст параметры синтаксической сложности экспертных и учебных текстов в двух предметных областях, полученные парсером UDPipe

| | Экономика (тексты) | | LL | %DIFF | Менеджмент (тексты) | | LL | %DIFF |
|--------------------------------|--------------------|--------|------|--------|---------------------|--------|------|--------|
| | студентов | статей | | | студентов | статей | | |
| Число обстоятельственных клауз | 39 | 240 | 0,17 | -7,24 | 45 | 276 | 1,71 | -18,57 |
| Число определительных клауз | 39 | 279 | 1,76 | -24,67 | 49 | 255 | 0,07 | -4,02 |
| Число относительных клауз | 21 | 98 | 0,71 | 18,67 | 26 | 122 | 0,08 | 6,44 |
| Число предложений | 119 | 728 | 0,43 | -6,61 | 131 | 622 | 0,27 | 5,19 |

По этой причине мы применили дополнительное взвешивание для достижения сопоставимости текстов студенческих работ и экспертных статей путем дальнейшего усреднения. Для этого мы число обстоятельственных клауз, атрибутивных конструкций и относительных клауз в строках 1–3 таблицы 1 поделили на среднее число предложений и на среднее число токенов в тексте. Результаты подсчета приведены в таблице 2.

Таблица 2

Усредненные параметры синтаксической сложности экспертных и учебных текстов в двух предметных областях, полученные парсером UDPipe

| | Экономика (тексты) | | Менеджмент (тексты) | |
|--|--------------------|--------|---------------------|--------|
| | студентов | статей | студентов | статей |
| Число обстоятельственных клауз на одно предложение | 0,331 | 0,329 | 0,341 | 0,444 |
| Число определительных клауз на одно предложение | 0,328 | 0,384 | 0,372 | 0,409 |
| Число относительных клауз на одно предложение | 0,177 | 0,135 | 0,197 | 0,196 |

Окончание табл. 2

| | Экономика (тексты) | | Менеджмент (тексты) | |
|--|--------------------|--------|---------------------|--------|
| | студентов | статей | студентов | статей |
| Число обстоятельственных клауз на один токен | 0,016 | 0,017 | 0,017 | 0,02 |
| Число определительных клауз на один токен | 0,016 | 0,019 | 0,018 | 0,018 |
| Число относительных клауз на один токен | 0,009 | 0,007 | 0,01 | 0,009 |

Данные таблицы 2 демонстрируют приблизительно одинаковый уровень синтаксической сложности экспертных и студенческих текстов в усредненных параметрах на предложение и на токен, из чего напрашивается вывод, что для более тонкого анализа синтаксической сложности нужны более дробные типы конструкций, а не только все относительные или все определительные конструкции. Чтобы убедиться в этом, мы рассмотрели тот параметр в текстах по экономике, в котором показатель в студенческих текстах превышает соответствующий показатель в текстах экспертных статей, – число относительных клауз на число предложений. Мы рассмотрели несколько предложений, где таких клауз больше одной, и в них видно, что использование относительных клауз в студенческих работах из вышеупомянутых корпусов является избыточным (предложения приведены в авторской орфографии и пунктуации, перевод максимально сохраняет авторский вариант, иногда в ущерб правильности предложения на русском языке):

- (7) *However, in other case the next model would be made, including the Chief characteristics, **which** are relevant only in family firms and find out **which** ones are important.*

‘Однако в ином случае была бы создана следующая модель, включающая основные характеристики, которые являются значимыми лишь в семейных компаниях, чтобы определить, какие являются важными’ (E-2015-9).

- (8) *The other approach to regional competitiveness **which** is associated with the most known theory for **which** Paul Krugman became known and for **which** he had received the Nobel Prize is the New Economic Geography.*

‘Другой подход к региональной конкурентоспособности, которая ассоциируется с самой известной теорией, благодаря которой прославился Пол Кругман и за которую он получил Нобелевскую премию, – это «новая экономическая география»’ (E-2015-46).

Иными словами, эти примеры свидетельствуют о том, что студенты, используя относительные клаузы, усложняют предложения настолько, что их понимание затрудняется, в то время как в статьях профессиональных изданий авторы используют те же конструкции для облегчения восприятия смысла текста, например:

- (9) *The supply chain is driven by the demand signal, **which** is composed of a replacement component for forklifts being retired after completing their useful life and a growth component (ЦРЭ-2013-1).*

‘Цепочка поставок управляется спросом, который представляет собой необходимость замены вилочных погрузчиков, списанных по истечении срока эксплуатации, в сочетании с ростом потребности в них’.

Автор статьи в первой клаузе использовал нерестриктивное относительное придаточное предложение для добавления информации к термину *demand signal*, а во второй клаузе в качестве рестриктивного определения слова *forklifts* использована определительная причастная конструкция, а не второе относительное придаточное предложение.

Результаты анализа клаузуальной и фразовой сложности научного текста представлены в таблице 3. В скобках указана нормализованная частотность на 1000 слов текста, что позволяет сопоставлять полученные результаты. Прочие условные обозначения те же, что были приведены выше перед таблицей 1.

Количественный анализ использования синтаксических конструкций, свидетельствующих о клаузуальной сложности текста, в целом подтвердил естественное предположение, что синтаксические конструкции преобладают в текстах, написанных профессиональными авторами. Наибольшая статистически значимая разница по сравнению со студенческими текстами (более чем в два раза) наблюдается в употреблении следующих обстоятельственных придаточных предложений: условия, уступки (в текстах по менеджменту), времени (в текстах по менеджменту) и результата (в текстах по экономике). Однако в функционировании обстоятельственных придаточных контраста прослеживаются различные тенденции при сравнении учебных текстов с профессиональными в зависимости от предметной области: в корпусе экономических текстов они статистически значимо чаще встречаются в работах студентов, а в корпусе по менеджменту – в текстах профессиональных авторов. Аналогичная ситуация наблюдается с изъяснительными придаточными: их слишком часто, по сравнению с профессионалами, используют студенты-экономисты и слишком редко – студенты-менеджеры. Исследователям еще предстоит понять,

Таблица 3

Параметры клаузуальной и фразовой сложности экспертных и учебных текстов в двух предметных областях, полученные инструментом AntConc

| Параметр | Экономика | | | | Менеджмент | | | |
|---|------------------|---------------|-------|---------|------------------|---------------|--------|---------|
| | тексты студентов | тексты статей | LL | %DIFF | тексты студентов | тексты статей | LL | %DIFF |
| Клаузуальная сложность | | | | | | | | |
| Финитные условные клаузы | 100 (0,89) | 1372 (2,02) | 79,02 | -126,31 | 68 (0,52) | 763 (1,10) | 42,45 | -110,18 |
| Уступительные клаузы | 47 (0,42) | 326 (0,48) | 0,77 | -14,41 | 34 (0,26) | 464 (0,69) | 36,76 | -155,64 |
| Временные клаузы | 95 (0,85) | 991 (1,46) | 29,67 | -72,07 | 94 (0,72) | 1252 (1,80) | 95,44 | -149,49 |
| Клаузы места | 79 (0,71) | 810 (1,19) | 22,97 | -69,12 | 47 (0,36) | 416 (0,60) | 12,35 | -65,80 |
| Причинные клаузы | 198 (1,77) | 1046 (1,54) | 3,06 | 12,86 | 158 (1,22) | 1126 (1,62) | 12,43 | -33,50 |
| Клаузы результата | 5 (0,04) | 177 (0,26) | 27,79 | -483,92 | 19 (0,15) | 60 (0,09) | 3,61 | 40,85 |
| Клаузы прогнотивопоставления | 65 (0,58) | 190 (0,28) | 22,64 | 51,78 | 59 (0,45) | 732 (1,05) | 49,48 | -132,40 |
| Всего финитных обстоятельственных клауз | 589 (5,26) | 4912 (7,23) | 58,14 | -37,56 | 498 (3,83) | 4813 (6,94) | 186,78 | -81,04 |

| | | | | | | | | |
|--|----------------|-----------------|--------|--------|----------------|-----------------|--------|--------|
| Клаузы с союзным словом на <i>wh</i> - | 40 (0,36) | 113 (0,17) | 15,07 | 53,40 | 38 (0,29) | 366 (0,53) | 14,07 | -80,42 |
| Глагол+клауза с союзным словом <i>that</i> | 486 (4,34) | 2937 (4,33) | 0 | 0,32 | 433 (3,33) | 4039 (5,82) | 141,53 | -74,73 |
| Фразовая сложность | | | | | | | | |
| Существительные | 32923 (293,96) | 197539 (290,93) | 3,01 | 1,03 | 39301 (302,32) | 206294 (297,25) | 9,37 | 1,67 |
| Адъективные определения | 7288 (65,07) | 39677 (58,43) | 69,51 | 10,20 | 8763 (67,41) | 48703 (70,18) | 4,11 | -1,47 |
| Существительные как определения | 4757 (42,47) | 39969 (58,86) | 493,26 | -38,59 | 5609 (43,15) | 36183 (52,14) | 182,16 | -20,84 |
| Генитивные конструкции с предлогом <i>of</i> | 1983 (17,71) | 7169 (10,56) | 374,94 | 40,37 | 2277 (17,52) | 7633 (11,00) | 347,83 | 37,21 |

в чем кроется причина таких особенностей, но полученные результаты следует учитывать преподавателям английского языка при обучении студентов разных специальностей.

Интересно, что анализ языковых единиц, свидетельствующих о фразовой сложности текста, выявил противоположную тенденцию. Единственной конструкцией, преобладающей в текстах экспертов по обеим дисциплинам, оказались существительные в функции определения.

Сравнение экспертных корпусов между собой показало, что статьи по менеджменту обладают более высокой клаузуальной и фразовой сложностью по рассматриваемым параметрам, чем статьи по экономике. Параметрами, преобладающими в текстах экспертов-экономистов, являются обстоятельственные придаточные условия и места, а также конструкции с существительными, выступающими в роли определения.

Разница в полученных показателях сложности экспертных текстов между экономикой и менеджментом не такая явная, как в более тонких показателях, полученных с использованием AntConc; если среднее число обстоятельственных клауз в текстах экспертных статей по экономике несколько меньше, чем в текстах по менеджменту (240 и 276 соответственно), то среднее число определительных клауз у экономистов примерно настолько же выше, чем у экспертов в области менеджмента (279 и 255 соответственно). Однако все показатели синтаксической сложности по рассматриваемым параметрам в экспертных текстах по менеджменту выше, чем в статьях по экономике. Тексты как студентов-экономистов, так и студентов-менеджеров имеют более низкие показатели сложности текста по сравнению с экспертными текстами по всем параметрам клаузуальной синтаксической сложности в двух рассматриваемых областях, но ситуация с параметрами сложности на одно предложение совершенно иная – показатели учебных текстов почти такие же, как в экспертных текстах той же области, – немного ниже или даже чуть выше, чем в экспертных, а число относительных клауз на одно предложение в экономических текстах, как говорилось выше, даже вполне ощутимо превышает число таких клауз в статьях экспертов.

Обсуждение результатов и выводы

Результаты количественного анализа эмпирического материала позволяют сделать следующие выводы:

1. Академический текст, написанный профессиональным автором, обладает высокой степенью как клаузуальной, так и фразовой

сложности, и оба этих вида синтаксической сложности больше распространены в текстах по менеджменту, чем в текстах по экономике.

2. По сравнению с текстами обучающихся клаузуальная сложность текстов профессионалов существенно выше, в то время как фразовая сложность сопоставима или незначительно ниже. Наиболее значимые различия (примерно в 1,7 раз больше в учебных текстах) наблюдаются в употреблении модификаторов с предлогом *of*, что, вероятно всего, объясняется тем, что изучающие язык предпочитают эту хорошо известную конструкцию всем остальным способам выражения приименных модификаторов, представленным в английском языке, а именно: притяжательные формы существительного перед определяемой именной группой; предшествующее определяемой именной группе существительное в функции определения; конструкции с другими предлогами после определяемой именной группы.

Для русскоязычных студентов конструкция с предлогом *of* очевидно является английским эквивалентом русского родительного падежа – формы с высокой частотностью. Порядок слов в этой конструкции совпадает с порядком слов в аналогичной русской конструкции, где главное слово предшествует зависимому (ср. *assets of the company* ‘активы компании’), что делает ее более понятной и простой для учащихся. В английской конструкции с существительным в качестве определения, напротив, главное слово следует за зависимым, что может вызывать определенные трудности у русскоязычных студентов (ср. гораздо более частотное *company assets* ‘активы компании’). В результате такого предпочтения в устной и письменной продукции русскоязычных изучающих нередко появляются фразы с чрезмерным присутствием предлога *of* – как, например, в следующем примере:

(10) *According to this research the biggest mistake of the participants of the negotiation is the absence of understanding of the different models of business and culture (E-2015-7).*

‘В соответствии с этим исследованием, самая большая ошибка участников переговоров – это отсутствие понимания различных моделей ведения бизнеса и культурных отличий’.

Другая распространенная в студенческом корпусе по менеджменту черта фразовой сложности – прилагательные в функции определения – также, по свидетельству корпусных данных, очевидно не вызывает сложности у русских изучающих английский язык вследствие наличия такой же конструкции и с таким же порядком слов в русском языке.

3. Для автоматизированных подсчетов параметров синтаксической сложности текста следует использовать методы, включающие

рассмотрение синтаксических конструкций с максимальной категориальной конкретизацией, так как в процессе освоения нового языка изучающими наблюдается интерференция с родным языком и возникают тенденции, приводящие к изменению частотности одного конкретного типа среди всех прочих разновидностей синтаксической конструкции – ср. результаты, полученные в [Ellis 1997], [Bhela 1999, pp. 26–27 и 29] и во многих других исследованиях процессов овладения иностранным языком. Используя два различных метода оценки синтаксической сложности текста: подсчет показателей применения UDPipe парсера и использование инструментов подсчета в AntConc, – мы получили результаты, которые противоречат выводам работы [Staples et al. 2016] о том, что с повышением уровня владения языком клаузальная сложность снижается, а фразовая растет. По нашим данным, параметры обоих видов синтаксической сложности в экспертных текстах в целом выше, чем в текстах студентов в тех же областях. Результаты, полученные на первом этапе, показывают бóльшую согласованность средних показателей сложности в тексте относительно его длины, нежели средних показателей на предложение или на число токенов.

В контексте преподавания английского языка для академических целей полученные результаты говорят о необходимости тренировать обучающихся в использовании сложноподчиненных предложений всех видов с целью приближения клаузальной сложности их текстов к уровню этого вида в текстах профессиональных авторов. В отношении фразовой сложности при обучении академическому английскому представляется необходимым делать акцент на употреблении существительных в качестве определений, особо привлекая внимание к тем случаям, когда учащиеся злоупотребляют выбором конструкций под влиянием интерференции со своим родным языком. Различие в значениях показателей сложности между дисциплинами является достаточно предсказуемым и свидетельствует о необходимости дисциплинарно-ориентированного подхода (*discipline-specific approach*) в преподавании английского языка для академических целей.

Таким образом, в результате проведенного исследования были получены ответы на три вопроса, поставленные авторами во введении:

- Академический текст, написанный профессиональным автором, демонстрирует черты как клаузальной, так и фразовой сложности. Были выявлены дисциплинарные различия в показателях сложности, например преобладание обстоятельственных придаточных предложений условия и места в экономических текстах, что можно объяснить их использованием в комментариях к различным формулам и моделям.

- Тексты профессиональных авторов в целом демонстрируют большую клаузальную и фразовую сложность. Наибольшие различия наблюдаются в употреблении различных видов придаточных предложений.
- Для оценки сложности текста представляется целесообразным рассматривать параметры сложности по отдельности, а не целыми группами, чтобы получить более подробную картину использования авторами различных синтаксических конструкций. Для этих целей можно применять инструменты для корпусного анализа текста, например AntConc.

Благодарности

Авторы выражают признательность П.М. Аркадьеву и анонимному рецензенту за ценные замечания к тексту статьи. Никто из указанных лиц не несет ответственности за возможные ошибки.

Публикация подготовлена по результатам проекта «2019 – Автоматизированная проверка текста, написанного на английском языке русскоязычными авторами (ADWISER)».

Acknowledgements

The authors are grateful to Peter Arkadiev and the anonymous reviewer for the important advice on the text of the paper. Any faults and shortcomings are the responsibility of the authors only.

The research was carried out within the HSE project 2019 – Automated detection of writing inaccuracies for students of English in Russia (ADWISER).

Литература

- Anthony 2014 – *Anthony L.* AntConc 3.4.4. Tokyo: Waseda University, 2014.
- Bhela 1999 – *Bhela B.* Native language interference in learning a second language: Exploratory case studies of native language interference with target language usage // *International Education Journal*. 1999. Vol. 1, no. 1. P. 22–31.
- Biber et al. 1999 – *Biber D., Johansson S., Leech G., Conrad S., Finegan E., Quirk R.* Longman grammar of spoken and written English. London: Longman, 1999.
- Biber et al. 2011 – *Biber D., Gray B., Poonpon K.* Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? // *TESOL Quarterly*. 2011. Vol. 45, no. 1. P. 5–35.
- Biber, Gray 2016 – *Biber D., Gray B.* Grammatical complexity in academic English: Linguistic change in writing. Cambridge: Cambridge University Press, 2016.
- Byrnes, Sinicrope 2008 – *Byrnes H., Sinicrope C.* Advancedness and the development of relativization in L2 German: a curriculum-based longitudinal study // *The longitudinal study of advanced L2 capacities* / L. Ortega, H. Byrnes, (ed.). New York: Routledge, 2008. P. 109–138.

- Crossley, McNamara 2014 – *Crossley S.A., McNamara D.S.* Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners // *Journal of Second Language Writing*. 2014. Vol. 26. P. 66–79.
- Dunning 1993 – *Dunning T.* Accurate methods for the statistics of surprise and coincidence // *Computational Linguistics*. 1993. Vol. 19, no. 1. P. 61–74.
- Ellis 1997 – *Ellis R.* *Second Language Acquisition*. Oxford: Oxford University Press, 1997.
- Gabrielatos, Marchi 2012 – *Gabrielatos C., Marchi A.* Keyness: Appropriate Metrics and Practical Issues // *CADS International Conference 2012. Corpus-assisted Discourse Studies: More than the sum of Discourse Analysis and Computing?* Bologna: University of Bologna, 2012.
- Gray 2013 – *Gray B.* More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles // *Corpora*. 2013. Vol. 8, no. 2. P. 153–181.
- Foster, Skehan 1996 – *Foster P., Skehan P.* The influence of planning and task type on second language performance, *Studies in Second Language Acquisition*. 1996. Vol. 18. P. 299–323.
- Gardner et al. 2018 – *Gardner S., Nesi H., Biber D.* Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing // *Applied Linguistics*. 2018. Vol. 10. P. 646–674.
- Hannay, Caro 2008 – *Hannay M., Caro E.M.* Thematic choice in the written English of advanced Spanish and Dutch learners // *Linking up contrastive and learner corpus research* / G. Gilquin, S. Papp, M.B. Díez-Bedmar (ed.). Amsterdam: Rodopi, 2008. P. 227–253.
- Hawkins 2001 – *Hawkins R.* *Second language syntax: A generative introduction*. Oxford: Wiley-Blackwell, 2001.
- Hinkel 2003 – *Hinkel E.* Simplicity without elegance: Features of sentences in L1 and L2 academic texts // *TESOL Quarterly*. 2003. Vol. 37, no. 2. P. 275–301.
- Jucker 1992 – *Jucker A.H.* *Social stylistics: Syntactic variation in British newspapers*. New York: Mouton de Gruyter, 1992.
- Lu 2011 – Lu X. A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development // *TESOL Quarterly*. 2011. Vol. 45, no. 1. P. 36–62.
- Lyashevskaya, Panteleeva 2018 – *Lyashevskaya O., Panteleeva I.* REALEC learner treebank: annotation principles and evaluation of automatic parsing [Электронный ресурс] // *Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories 2018 (TLT16)*. Prague, 2018. P. 80–87. URL: <https://www.aclweb.org/anthology/W17-7612.pdf> (дата обращения 26 нояб. 2019).
- Mazgutov, Kormos 2015 – *Mazgutova D., Kormos J.* Syntactic and lexical development in an intensive English for Academic Purposes programme // *Journal of Second Language Writing*. 2015. Vol. 29. P. 3–15.
- Nesi, Gardner 2019 – *Nesi H., Gardner S.* Complex, but in what way? A step towards greater understanding of academic writing proficiency [Электронный ресурс] // *Online Proceedings of the 51st Annual Meeting of the British Association for Applied Linguistics: Taking Risks in Applied Linguistics*, 6-8 September, 2018, York St John University / C. Danjo, I. Meddegama, D. O'Brien, Prudhoe J., L. Walz, R. Wicaksono (ed.). York: York St. John University, 2019. URL: <https://custom.cvent.com/01664CE00C344F7BA62E39C4CFE91FA8/files/0f77de05e-b81461a8037170680562243.pdf> (дата обращения 22 июня 2019).
- Ortega 2003 – *Ortega L.* Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing // *Applied Linguistics*. 2003. Vol. 24. P. 492–518.

- Siepmann et. al 2001 – *Siepmann D., Gallagher J.D., Hannay M., Mackenzie J.L.* Writing in English: A guide for advanced learners. Tübingen: A. Francke Verlag, 2011.
- Smith 2009 – *Smith N.* Corpus Linguistics: A Practical Web-Based Course. 2009 [Электронный ресурс]. URL: https://www.lancaster.ac.uk/fss/courses/ling/corpus/blue/108_4.htm (дата обращения 22 февр. 2020).
- Staples et al. 2016 – *Staples S., Egbert J., Biber D., Gray B.* Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre // *Written Communication*. 2016. Vol. 33, no. 2. P. 149–183.
- Straka 2017 – *Straka M.* CoNLL 2017 shared task – UDPipe baseline models and supplementary materials // LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics. Prague: Charles University, 2017.
- Suárez, Álvarez 2010 – *Suárez S.D., Álvarez E.G.* The use of it-clefts in the written production of Spanish advanced learners of English // *Linguistics & the Human Sciences*. 2010. Vol. 6. P. 151–171.
- Swales, Feak 2004 – *Swales J.M., Feak C.B.* Academic writing for graduate students: Essential tasks and skills. Ann Arbor, MI: University of Michigan Press, 2004.
- Warchal 2010 – *Warchal K.* Moulding interpersonal relations through conditional clauses, consensus-building strategies in written academic discourse // *Journal of English for Academic Purposes*. 2010. Vol. 9. P. 140–150.
- Wolfe-Quintero et al. 1998 – *Wolfe-Quintero K., Inagaki S., Kim H.Y.* *Second language development in writing: Measures of fluency, accuracy, and complexity*, Honolulu, HI: University of Hawaii Press, 1998.

References

- Anthony, L. (2014), *AntConc 3.4.4*, Waseda University, Tokyo, Japan.
- Bhela, B. (1999), “Native language interference in learning a second language: Exploratory case studies of native language interference with target language usage”, *International Education Journal*, vol. 1, no. 1, pp. 22–31.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. and Quirk, R. (1999), *Longman grammar of spoken and written English*, Longman, London, UK.
- Biber, D., Gray, B. and Poonpon, K. (2011), “Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?”, *TESOL Quarterly*, vol. 45, no. 1, pp. 5–35.
- Biber, D. and Gray, B. (2016), *Grammatical complexity in academic English: Linguistic change in writing*, Cambridge University Press, Cambridge, UK.
- Byrnes, H. and Sinicrope, C. (2008), “Advancedness and the development of relativization in L2 German: a curriculum-based longitudinal study”, in Ortega, L. and Byrnes, H. (ed.), *The longitudinal study of advanced L2 capacities*, Routledge, New York, USA, pp. 109–138.
- Crossley, S.A. and McNamara, D.S. (2014), “Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners”, *Journal of Second Language Writing*, vol. 26, pp. 66–79.
- Dunning, T. (1993), “Accurate methods for the statistics of surprise and coincidence”, *Computational Linguistics*, vol. 19, no. 1, pp. 61–74.
- Ellis, R. (1997), *Second Language Acquisition*, Oxford University Press, Oxford, UK.
- Gabrielatos, C. and Marchi, A. (2012), “Keyness: Appropriate Metrics and Practical Issues”, in *CADS International Conference 2012. Corpus-assisted Discourse Studies*:

- More than the sum of Discourse Analysis and Computing?* University of Bologna, Bologna, Italy.
- Gray, B. (2013), "More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles", *Corpora*, vol. 8, no. 2, pp. 153–181.
- Foster, P. and Skehan, P. (1996), "The influence of planning and task type on second language performance", *Studies in Second Language Acquisition*, vol. 18, pp. 299–323.
- Gardner, S., Nesi, H. and Biber, D. (2018), "Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing", *Applied Linguistics*, vol. 10, pp. 646–674.
- Hannay, M. and Caro, E.M. (2008), "Thematic choice in the written English of advanced Spanish and Dutch learners", in Gilquin, G., Papp, S. and Díez-Bedmar, M.B. (eds.), *Linking up contrastive and learner corpus research*, Rodopi, Amsterdam, Netherlands, pp. 227–253.
- Hawkins, R. (2001), *Second language syntax: A generative introduction*, Wiley-Blackwell, Oxford, UK.
- Hinkel, E. (2003), "Simplicity without elegance: Features of sentences in L1 and L2 academic texts", *TESOL Quarterly*, vol. 37, no. 2, pp. 275–301.
- Jucker, A.H. (1992), *Social stylistics: Syntactic variation in British newspapers*, Mouton de Gruyter, New York, USA.
- Lu, X. (2011), "A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development", *TESOL Quarterly*, vol. 45, no. 1, pp. 36–62.
- Lyashevskaya, O. and Panteleeva, I. (2018), "REALEC learner treebank: annotation principles and evaluation of automatic parsing", *Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories 2018 (TLT16)*, Prague, pp. 80–87, [Online], available at: <https://www.aclweb.org/anthology/W17-7612.pdf> (Accessed 26 November 2019).
- Mazgutova, D. and Kormos, J. (2015), "Syntactic and lexical development in an intensive English for Academic Purposes programme", *Journal of Second Language Writing*, vol. 29, pp. 3–15.
- Nesi, H. and Gardner, S. (2019), "Complex, but in what way? A step towards greater understanding of academic writing proficiency", *Online Proceedings of the 5^{4st} Annual Meeting of the British Association for Applied Linguistics: Taking Risks in Applied Linguistics, 6–8 September, 2018*, York St. John University, York [Online], available at: <https://custom.cvent.com/01664CE00C344F7BA62E-39C4CFE91FA8/files/0f77de05eb81461a80371706880562243.pdf> (Accessed 22 June 2019)
- Ortega, L. (2003), "Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing", *Applied Linguistics*, vol. 24, pp. 492–518.
- Siepmann, D., Gallagher, J.D., Hannay, M. and Mackenzie, J.L. (2011), *Writing in English: A guide for advanced learners*, A. Francke Verlag, Tübingen, Germany.
- Smith, N. (2009), *Corpus Linguistics: A Practical Web-Based Course* [Online], available at: https://www.lancaster.ac.uk/fss/courses/ling/corpus/blue/108_4.htm (Accessed 22 February 2020).
- Staples, S., Egbert, J., Biber, D. and Gray, B. (2016), "Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre", *Written Communication*, vol. 33, no. 2, pp. 149–183.
- Straka, M. (2017), "CoNLL 2017 shared task – UDPipe baseline models and supplementary materials", *LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics*, Charles University, Prague.

- Suárez, S.D. and Álvarez, E.G. (2010), "The use of it-clefts in the written production of Spanish advanced learners of English", *Linguistics & the Human Sciences*, vol. 6, pp. 151–171.
- Swales, J.M. and Feak, C.B. (2004). *Academic writing for graduate students: Essential tasks and skills*, University of Michigan Press, Ann Arbor, MI, USA.
- Warchal, K. (2010), "Moulding interpersonal relations through conditional clauses, consensus-building strategies in written academic discourse", *Journal of English for Academic Purposes*, vol. 9, pp. 140–150.
- Wolfe-Quintero, K., Inagaki, S. and Kim, H.Y. (1998), *Second language development in writing: Measures of fluency, accuracy, and complexity*, University of Hawaii Press, Honolulu, HI, USA.

Информация об авторах

Ольга И. Виноградова, кандидат филологических наук, доцент, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия; 115054, Россия, Москва, ул. Старая Басманная, д. 21/4, стр. 1; olgavinogr@gmail.com

Елизавета А. Смирнова, кандидат филологических наук, доцент, Национальный исследовательский университет «Высшая школа экономики», Пермь, Россия; 614070, Россия, Пермь, ул. Студенческая, д. 38; cmelizaveta@yandex.ru

Анна В. Виклова, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия; 115054, Россия, Москва, ул. Старая Басманная, д. 21/4, стр. 1; annakisseleva@gmail.com

Ирина М. Пантелеева, независимый исследователь. Москва, Россия; impanteleyeva@gmail.com

Information about the authors

Olga I. Vinogradova, Cand. of Sci. (Philology), associate professor, HSE University, Moscow, Russia; 21/4 Staraya Basmannaya Str., bld. 1, Moscow 115054, Russia; olgavinogr@gmail.com

Elizaveta A. Smirnova, Cand. of Sci. (Philology), associate professor, HSE University Perm, Russia; 38 Studencheskaya Str., Perm 614070, Russia; cmelizaveta@yandex.ru

Anna V. Viklova, HSE University Moscow, Russia; 21/4 Staraya Basmannaya Str., bld. 1, Moscow 115054, Russia; annakisseleva@gmail.com

Irina M. Panteleeva, independent researcher, Moscow, Russia; impanteleyeva@gmail.com

Russian onomatopoeic verbal interjections. Why use 'non-words' instead of ordinary ones?

Oksana A. Kanerva

Helsinki University, Helsinki, Finland, oksana.kanerva@helsinki.fi

Abstract. This paper considers such expressive words in Russian as onomatopoeic verbal interjections (*bukh* 'bang', *stuk* 'knock', *bul* 'plop'). First of all, it focuses on the reasons for using those words. This study suggests that native speakers of Russian may deliberately prefer those linguistic units over words from other classes (nouns, verbs or adverbs). It happens when such factors as Zipf's law and the principles of iconicity and economy in grammar come into play. Secondly, the article claims that while morphologically simple, syntactically mobile and with transparent onomatopoeia-based meaning, those words are not that primitive. Onomatopoeic verbal interjections are capable of conveying multiple shades of meaning due to such diachronic processes as lexical shifts and extensions they have undergone. That kind of qualities enable speakers to imply more than they want to explicitly utter under certain conditions. And finally, using corpus data the research distinguishes such pragmatic functions of those linguistic units (*bukh* 'bang', *stuk* 'knock', *bul* 'plop') as quasi-referential, emotive, phatic, poetic and metalinguistic function.

Keywords: Russian, onomatopoeic verbal interjections, pragmatic functions, Zipf's law, iconicity and economy in grammar

For citation: Kanerva, O.A. (2020), "Russian onomatopoeic verbal interjections. Why use 'non-words' instead of ordinary ones?", *RSUH/RGGU Bulletin. "Literary Theory. Linguistics. Cultural Studies" Series*, no. 7, pp. 130-146, DOI: 10.28995/2686-7249-2020-7-130-146

Звукоподражательные глагольные междометия русского языка: зачем использовать «неполноценные» слова вместо обычных?

Оксана А. Канерва

Хельсинкский университет, Хельсинки, Финляндия, oksana.kanerva@helsinki.fi

Аннотация. В статье рассматриваются такие экспрессивные языковые единицы, как звукоподражательные глагольные междометия (например, *бух*, *стук*, *буль*). Прежде всего, данное исследование фокусируется на причинах использования этих слов. В нем утверждается, что носители русского языка преднамеренно употребляют их вместо слов из других грам-

матических классов (например, существительных, глаголов или наречий). На это влияют такие факторы, как закон Ципфа, принципы иконичности и экономии в грамматике. Кроме того, звукоподражательные глагольные междометия никак не являются примитивными, несмотря на их простую морфологическую форму, синтаксическую мобильность и прозрачное лексическое значение, основывающееся на звукоподражании. Эти слова способны передавать разнообразные оттенки значений благодаря расширению и смещению их семантики в процессе исторического развития. Подобные качества позволяют говорящему имплицировать больше, чем он хочет открыто сказать. В заключение на основе корпусных данных в исследовании выделяются такие функции слов типа *бух*, *стук* и *буль*, как квазиреферентная, эмотивная, фатическая, поэтическая и метаязыковая.

Ключевые слова: русский язык, звукоподражательные глагольные междометия, прагматические функции, закон Ципфа, принципы иконичности и экономии в грамматике

Для цитирования: Канерва О.А. Звукоподражательные глагольные междометия русского языка: зачем использовать «неполноценные» слова вместо обычных? // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 130–146. DOI: 10.28995/2686-7249-2020-7-130-146

Introduction

The object of this study is a class of words in Russian referred to as onomatopoeic verbal interjections (e.g. *bukh* ‘bang’, *stuk* ‘knock’ and *bul’* ‘plop’). They are often assigned to the class of interjections, [see Shvedova 1980; Kartsevskii 1984; Kryk 1992; Wierzbicka 2003; Cuenca 2006]. Moreover, they are viewed as sound imitations [Wierzbicka 2003, p. 290; Oswalt 1994, p. 293]. Additionally, a number of scholars [Kartsevskii 1984, p. 132; Nikitina 2012, p. 165; Kanerva and Viimaranta 2018, p. 94] claim that these words have verbal features, i.e. they can function syntactically as predicates. Although interjections have been referred to as “non-words” [Ameka 1992, p. 101], this study insists on grammatical and contextual complexity of words like *bukh*, *stuk* and *bul’*.

This research focuses on identifying factors which drive speakers of Russian to use onomatopoeic verbal interjections in the first place. It aims at exploring pragmatic discourse functions of these words. The research questions can be formulated as follows:

- 1) In what contexts are onomatopoeic verbal interjections used?
- 2) What considerations drive native speakers of Russian to use onomatopoeic verbal interjections instead of ordinary words?

3) What pragmatic functions do onomatopoeic verbal interjections carry in discourse?

The material for this study was collected with the help of the New Explanatory Dictionary of the Russian Language by Efremova [Efremova 2000]. The search was made for all linguistic units marked both as predicates and interjections. This article is mainly going to concentrate on onomatopoeic verbal interjections. In order to distinguish them from verbal interjections, the entries for each word had to be checked separately for the description of the corresponding sound it depicts. This helped retrieve 76 onomatopoeic verbal interjections from Efremova [Efremova 2000]: *apchkhi, akh, akh-akh, bakh, babakh, bats, bom, bryak, bul'tykh, bul', bul'-bul', bum, bukh, vzhik, grokh, din'-din', don, zhzhik, zvyak, ik, kap, paf, paff, pif-paf, plyukh, pykh, skrip, stuk, trakh, trakh-tararakh, tararakh, tik, tik-tak, top, tren'-bren', tuk, tuk-tuk, tyuk, tyu-tyu, tyap, ukh, fu, fuk, f'yu, f'yu-f'yu, khlest, khlest', khlobyst', khlop, khlyst', khlyup, khlyask, khlyast', khryast', khrup, khrust', khryast', chok, chik, chirik, chirk, chikh, chmok, chok, chukh, sharakh, shark, shvark, shlep, shi, shmyg, shmyak, shork, shu-shu, shchelk*.

Two sub-corpora (the Main Corpus and the Corpus of Spoken Russian) of the Russian National Corpus (<http://ruscorpora.ru>) were analysed in this study. The Main Corpus of the Russian National Corpus (RNC) is represented by written texts of different genres and styles. In 2018 its total volume was 283,431,966 words and the search in it offered 4514 tokens featuring onomatopoeic verbal interjections from the list given above (approx. 16 tokens per million). The Corpus of Spoken Russian of the Russian National Corpus contains not only private conversations, but also public speeches and transcripts of films/cartoons, which belong to the prepared speech. In 2018 its total volume was 12,113,491 words and the search in it gave 203 tokens (approx. 17 tokens per million).

Features of onomatopoeic verbal interjections

Words like *bukh, stuk* and *bul'* have been characterized as “invariable linguistic units, which do not contain inflectional or derivational morphemes; aim at depicting acoustic events or indicating kinetic imagery; in some contexts express suddenness in connection with an indicated action; and function syntactically as independent utterances/clauses or become grammatically integrated by taking a predicate” [Kanerva 2018, p. 18]. From this definition it follows that onomatopoeic verbal interjections display interjectional, onomatopoeic and predicative features.

Clark and Gerrig [Clark and Gerrig 1990, pp. 788–789] claim that all sound imitations depict natural sounds and concentrate on the object of representation. There is a consensus among scholars that interjections indicate mental state of the speaker. Meinard [Meinard 2015, p. 167] distinguishes onomatopoeias from interjections and concludes that “isolated onomatopoeias focus on an object of conceptualization and fulfill the referential function of language whereas interjections focus on the modus, the subject of conceptualization, and fulfill the conative, phatic or expressive functions of language”. It means that interjections are capable of fulfilling pragmatic functions in discourse. In order to distinguish onomatopoeias from interjections, the main attention has to be paid to the mode of representation for each of these categories [Meinard, p. 156]. According to Nikitina [Nikitina 2012, p. 165], all verbal interjections or “verboids” (such as *morg* ‘blink’, *chlest* ‘plunge’, *nyrk* ‘lash’), in their turn, indicate kinetic movements. An important conclusion can now be drawn: none of the three word classes mentioned is used for description. They either indicate or depict certain events or states. There is strong typological evidence that onomatopoeias and interjections belong to expressive vocabulary [see Hinton et al. 1994; Foolen 2012; Dingemanse and Akita 2017]. This testifies in favor of regarding onomatopoeic verbal interjections as expressive linguistic units.

Words similar to Russian *bukh*, *stuk* and *bul* are believed to have given rise to a set of inflected nouns and verbs [Oswalt 1994, p. 302; Kor Chahine 2008, p. 157]. Thus, in the course of lexicalization a specific group of noun and verb derivatives has developed from regular onomatopoeias depicting certain natural sounds. As a result of this process, almost every onomatopoeic verbal interjection in Russian has a homonymic form of a masculine noun in the nominative (1) and accusative case (2) with zero ending:

(1) *poslyshalsya stuk v dver'*

‘knocking on the door was heard’,

(2) *on uslyshal stuk v dver'*

‘he heard (somebody) knocking on the door’.

In all other cases there are overt markers for all these nouns in accordance with general rules of Russian inflection: *stuk-a* ‘knock-GEN. SG’, *stuk-om* ‘knock-INTR.SG’, etc.

It is important to mention that derived words are conventionalized lexical items. The morpheme, as prototypically defined, is a syn-

tagmatic element which combines with other such elements to form words [Waugh 1995, p. 202]. Absence of inflectional and derivational morphemes is one more argument in favor of regarding onomatopoeic verbal interjections as non-conventionalized linguistic units. Simple morphological structure, however, does not deprive Russian words like *bukh*, *stuk* and *bul'* of a predicative function in certain contexts. What is more, it is completely possible to use them as corresponding inflected verbs. Additionally, a number of works have shown that they even possess such verbal features as aspect, tense, sometimes transitivity and voice [Kartsevskii 1984, p. 132; Viimaranta 2015]. As a result, it is possible to conclude that they behave syntactically similarly to verbs: they are not set off punctuationally from the rest of the sentence, can govern nouns denoting participants of the situation. For example:

(3) *chto-to bukhh v steklo*

'something banged into the (window)glass'

(4) *bukhh kulakom po stolu*

'(he/she) banged the table with (his/her) fist'

There are several restrictions to this: words like *bukh*, *stuk* and *bul'* are almost never used in negative clauses or questions [Nikitina 2012, p. 165]. Contrary to homonymic nouns and inflected verbs, which are fully conventionalized and integrated into the language system, onomatopoeic verbal interjections in a predicative function preserve their expressiveness to some extent [Kanerva 2018, p. 22]. Semantic shifts have made it possible for these words to display interjectional, onomatopoeic and predicative features in different contexts.

According to Sharonov [Sharonov 2008, §2.4], there is a conventionality scale within this very class of words. On the one end of this scale there are words conventionally linked to the source of sound. They have highly conventional, language specific spelling and low context dependence (*tik-tak* for a clock, *chukh-chukh* for a train). On the other end of it there are words which iconically depict the sound itself. The connection with the source of this sound is revealed through the context (*bum*, *tram*). In the middle of this scale there is a less conventionalized group of words, which show connection with different types of source objects (*bul'*, *khrust'*). Their morphological expression is still very familiar for native speakers.

*Iconic and economic motivation
for using onomatopoeic verbal interjections*

First of all, the factor which works in favor of using onomatopoeic verbal interjections instead of ordinary words is the “iconicity meta-principle”. It has been formulated by Givón [Givón 1985, p. 189] in the following way: “all other things being equal, a coded experience is easier to store, retrieve, and communicate if the code is maximally isomorphic to the experience”. The basic idea behind it is that iconic instances are easier to process than non-iconic ones. Without a doubt, there is a strong iconic correspondence between form and meaning of words like *bukh*, *stuk* and *bul'* owing to their sound-symbolic nature. Subsequently, speakers can access the meaning of these linguistic units with less effort due to their ability to iconically depict sound or indicate action in relation with this sound. A number of research papers [see Kor Chahine 2008, Nikitina 2012, Kanerva and Viimaranta 2018] as well as the corpus data show that Russian onomatopoeic verbal interjections are commonly used in predicative function instead of ordinary verbs:

- (5) *Dyatel nosom tuk da tuk*. [N.K. Rerikh. Listy dnevnika (1945)]
‘The woodpecker was drumming with its beak.’
- (6) [...] *on i bultykh v vodu*. [I.S. Shmelev. Leto Gospodne (1927–1944)]
‘He splashed into the water.’
- (7) *ya dver' na zamok shchelk*. [B.V. Savinkov (V. Ropshin).
To, chego ne bylo (1918)]
‘I locked the door.’

Haiman’s concept [Haiman 1983, p. 801] of “economic motivation” stipulates that “the conceptual simplicity of a notion corresponds to the simplicity of its expression”. Consequently, “reduction of form is economically motivated index of familiarity” [Haiman, p. 802]. This concept explains laconic morphological expression of words like *bukh*, *stuk* and *bul'*, which, by definition, do not host inflectional or derivational morphemes. This principle implies that familiar entities are usually morphologically shorter. Such familiarity is guaranteed by the semantic predictability of these words on the level of sound imitation.

Economic motivation for the usage of onomatopoeic verbal interjections can be seen in Grice’s [Grice 1975, p. 45] “maxim of quantity”, which demands from the speaker to make the utterance as informative as it is required by the speech situation. At the same time, this utterance should not convey more information than needed. Concise mor-

phological form in combination with iconic nature of words like *bukh*, *stuk* and *bul'* meets both demands set by the “maxim of quantity”. These two features make it possible to keep the utterances short and maximally informative at the same time. The Russian National Corpus offers conversation strings where speakers use rather short phrases containing words similar to *bukh*, *stuk* and *bul'*. It seems to be sufficient to convey meaning in full.

- (8) *Zamerzla! Tuk / tuk / tuk / ona zhiva!* [Dyuimovochka (1964), cartoon]
 ‘It (the swallow) has frozen to death! [Thumbelina is checking the heartbeat] Thump/thump/thump/ it is alive!’
- (9) – *I salatik iz kapusty takoi / nu / takoi...*
 – ***Chik-chik-chik-chik!***
 – *Da / i eshche tam mozno vsykie pasty...* [From the collection of the RNC (2006), private conversation]
 ‘– And a cabbage salad, well, like that...
 – Slash-slash-slash!
 – Yes, and also different types of pasta are possible (to make)...’

Zipf’s law [1935], which insists on communicative effectiveness with least effort, also motivates speakers to frequently use onomatopoeic verbal interjections. As far as there is a direct correlation between predictability of meaning and low amount of coding material [Givón 1991, pp. 87–89], less effort is required from the speaker to depict acoustic events or indicate kinetic imagery by resorting to these words.

The research data suggests that onomatopoeic verbal interjections are not rare. The New Explanatory Dictionary of the Russian Language by Efremova [Efremova 2000] offers 76 conventionalized linguistic units, i.e. the ones which are well established in the language. It is logical to assume that non-conventionalized items (*om* for chewing, *bryn'* for strumming a guitar), which have not yet been added to dictionaries and are not that familiar to all native speakers, may number in hundreds. The dictionary and the Corpus give more evidence in support of the role of Zipf’s law. As it has been mentioned, onomatopoeic verbal interjections tend to have corresponding inflected verbs or even aspectual pairs. For example, *bats*-INTJ has both *batsnut'*-PF and *batsat'*-IPF, but *chlobyst'*-INTJ has only *chlobystnut'*-PF and *shu-shu*-INTJ has only *shushukat'*-IPF. As a matter of fact, many of the onomatopoeic verbal interjections given in the list do not have any verb equivalents (*apch-khi*, *bom*, *don*, *fu*, *zhzhik*, etc.). Moreover, the search in the Main Corpus for some of these linguistic units demonstrates that they are more fre-

quently used than verbs. For example, there are 550 tokens with *bats*, but only 31 tokens with *batsat*' and 23 with *batsnut*'; 165 with *chik*, but 34 with *chiknut*' and 47 with *chikat*'.

In conclusion, onomatopoeic verbal interjections constitute a rare case when Zipf's law and the principles of iconicity and economy in grammar work well together. Sound-symbolic nature of these words makes them easier to process, morphologically simpler, communicatively efficient, effort saving and, as a result, promotes their frequent usage.

Pragmatic motivation for using onomatopoeic verbal interjections

There are other factors that determine the usage of onomatopoeic verbal interjections. These linguistic units are semantically complicated in a way that they can convey various shades of meaning and involve multiple implications. This quality has developed owing to such diachronic processes in language as semantic extensions and lexical shifts [Oswalt 1994, p. 302; Kryk 1992, pp. 199–200]. As a result, Russian onomatopoeic verbal interjections do not only depict natural sounds, they can also indicate actions or their characteristics to such an extent that the suddenness of something happening can express the speaker's attitude towards it. When it comes to using these linguistic units to render occasional semantic implications, pragmatic motivation works in favor of it. Discursively, they serve certain pragmatic functions. In such utterances they are grammatically processed in the same manner as interjections. According to Wilkins [Wilkins 1992, p. 129], "the size of an utterance (i.e. the number of morphemes an utterance contains) is inversely proportional to the amount of information that is recoverable from context". Interjections are sensitive to context and can only be interpreted in relation to it [Cuenca 2006, p. 35]. In order to guess what is implied by the speaker, the pragmatic context has to be resorted to.

In those cases when the speaker wishes to insinuate the contents of an utterance without naming the event or the state of affairs directly, using an onomatopoeic verbal interjection can be quite handy. This effect is achieved with the help of the differences between an object, its meaning, and a mental image of this object as "a part or mode of the single person's mind" [Frege 1949 [1892], p. 88]. While the meaning of a sign is "common property of many", the mental image is subjective. "The image of one person is not that of another. Hence, the various differences between the images connected with one and the same meaning" [Frege 1949 [1892], pp. 87–88].

Even more possible implications can be traced when an onomatopoeic verbal interjection is put in quotation marks in written text. This punctuation mark is often applied to indicate that the word/phrase is used in an unusual way [Valgina 1979]. As it has been experimentally shown by Kanerva and Viimaranta [Kanerva and Viimaranta 2018, p. 89], such expressions allow elicitation of a vast amount of possible scenarios of the speech event.

- (10) *Stop-khlop, i vot – klient uzhe daet pokazaniya!*
[A. Rubanov. Sazhaite, i vyrastet (2005)]
‘Stop-slam-INTJ, and here you are – the client is already testifying!’
- (11) *Kak ona, pomnish, pela: “Ta-ta-ta, yyyyyyyyyy... Da-da... yyyyyyyyyy, khop-khir’ep, din’-din’ i tam eshche blin-blin-blin”.*
[E. Pishchikova. Pyatietazhnaya Rossiya (2007)]
‘Do you remember how it was singing: “Ta-ta-ta, uuuuuuuuu... Da-da... uuuuuuuuu, d...mn, ding-ding-INTJ, and then also drat-drat-drat”
- (12) – *S kosoi, v prostyne, kak polozheno... Chrust-chrust, zvyak-zvyak... – Bred!* [S. Lungin. Vidennoe nayavu (1989–1996)]
‘With a scythe, in a sheet, the way it should be... Crack-crack, clink-clink-INTJ... – Delirium!’

These utterances taken from the Russian National Corpus illustrate the variety of semantic implications these units can convey. For instance, they imply arrest and imprisonment (10), point at repair works in a block of flats (11), mimic the sounds produced by Death when he is walking with his bones crunching and chains clinking (12).

Pragmatic functions of onomatopoeic verbal interjections

Boronnikova and Verižnikova [Boronnikova and Verižnikova 2014] on the basis of Jakobson’s [Jakobson 1960, pp. 350–377] classification of language functions put onomatopoeic interjections and verbal interjections into the quasi-referential functional type. This is the first function which can be singled out. These words are primarily used to depict a speech event in connection with sound or action related to this sound. That is why they are quite often accompanied by deictic elements as mentioned before.

Secondly, onomatopoeic verbal interjections are not void of the affective function, either. Like other interjections, they belong to expres-

sive vocabulary and can be used to indicate mental states and emotions. Indeed, in certain contexts they can clearly express shock (13), surprise (14) or excitement (15):

- (13) *Zhili mirno, / sosushchestvovali, / i vdrug – / **bakh!** / – v tyur'mu posadili.* [Obshchestvennoe mnenie (2004), interview]
 ‘They lived peacefully, co-existed, but suddenly – bang!-INTJ – were sent to prison.’
- (14) *Ne uspeli my zamochit' v koryte portyanki i bel'e – **bats!** – gosti k nam.* [V. Astaf'ev. Oberton (1995–1996)]
 ‘We have not even had time to soak our footcloths and underwear in the washtub – bang!-INTJ – guests are coming.’
- (15) *On i razbogatel' hotel, kak mechtali geroi Fedora Mikhailovicha: **trakh** – i razbogatel.* [A. Genis. Dovlatov i okrestnosti (1998)]
 ‘He also wanted to become rich, the way Fedor Mikhailovich’s characters dreamt: bang-INTJ – he has become rich.’

Expressive linguistic units are characterized as “the vocal gestures which are symptoms of the speaker’s mental state” [Ameka 1992, p. 113]. Such words can be identified as performing the affective function if the speaker’s attitude is traceable through the contents of the utterance.

The third function onomatopoeic verbal interjections can perform is to draw listener’s attention. In this article Jakobson’s [Jakobson 1960] term “phatic” refers to attracting attention and involving the listener. As a result, the basic criterion for identifying this function is to discover evidence from the context that these words help establish or maintain communicative contact.

- (16) *Vysokoe prividenie v nee, ya za nim, uzh tam, i – **bryak**, peredo mnoyu na polu uzhasnyi mertvoets.* [I.I. Lazhechnikov. Poslednii Novik (1833)]
 ‘A tall ghost (went) through (the door), I followed it, naturally, and – plop-INTJ, a horrible dead body on the floor in front of me.’
 Further in the text it is mentioned that the narrator, encouraged by the attention of the terrified listeners, continued his story.
- (17) *Ofitser, ne slushaya, skazal: – **Paf! Paf!*** [A. Kuznetsov. Babii zhar (1965–1970)]
 ‘The officer said without listening: – Pow! Pow!-INTJ’

The characters, to whom *paf* was addressed, were not planned to be executed. The intention was to utilize their house for the needs of the army. This objective could be met by intimidating the owners so that they do not show any resistance. And the context explicitly states that this aim was achieved: their attention was captured by the horrendous implications of these words.

- (18) *Potom podoshla voina, i za voinoi – pykh-trakh: vzdybil narod – myatezh, ogon' i burya.* [V.Ya. Shishkov. Vataga (1923)]

‘And then the war came, and after the war – puff-bang-INTJ:
the people rose up – rebellion, fire and storm.’

Although this utterance can be characterized as an internal monologue of the character, he imagined nonetheless how his words would encourage the others to join the mutiny.

- (19) *Voobshche, naskol'ko ya ponimayu, fiziki tol'ko i delayut, chto sledyat za rozhdeniem i raspadom chastits. Stuknuli proton v proton – khlop! – energiya prevratilas' v kvark.* [A. Torgashev and V. Rubakov. Kakie bozony nuzhny narodu // Russkii reporter, 2012]

‘In general, as far as I understand, physicists are mainly busy with watching the creation and decay of particles. They hit a proton against a proton – bang!-INTJ – the energy has been transformed into a quark.’

Example (19) is from an interview taken from an acknowledged physicist. It differs from the previous examples in a sense that this utterance is a side comment made by the journalist. The main purpose of this utterance is to comfort the reader who at times might be unfamiliar with the intricacies of nuclear synthesis. Such simplifications promote a greater involvement by showing that the author and the reader possess similar levels of knowledge.

The usage of verbal interjections in examples (16)–(19) is optional. The phatic function differs from the affective one in the sense that, when aiming at attracting attention, these words do not convey any additional information. Nevertheless, they accentuate the events and make the listener or the reader more involved.

In some cases, even poetic function can be observed, like in examples (20)–(22). In these utterances onomatopoeic verbal interjections are used to make the text more aesthetically appealing.

- (20) *Rachinskii, kogda-to slovesno “porovshii” menya za dekadentskie obrazy, fyrkaet dymom na yurkosti Shpetta: – “Paf, paf!” – vyletayut iz ust ego kluby.* [A. Belyi. Mezhdru dvukh revolyutsii (1934)]

‘Rachinsky, who used to verbally whip me for decadent imagery, is now snorting with smoke to Shpett’s witty remarks: – “Puff, puff!” – Puffs are flying from his mouth.’

- (21) *Komar saditsya na cheloveka, / chelovek udaryayet po nemu ladon'yu, / i oba, / Vashe velichestvo... bakh! trákh! tararákh! vzryváyutsya. A tepér' predstávo'te sebé, / Vashe velichestvo, / chto u nas ne odín takóí komár, / a millión [...]*

– **Tararákh!**

– **Tararákh**, / *Vashe velichestvo!* [Kain XVIII (1963), feature film]

‘A mosquito is landing on a man, the man is slapping it, and both of them, Your Majesty... bang! bang! bang! are exploding. And now imagine it to yourself, Your Majesty, that there is not only one mosquito, but a million [...]

– Bang!

– Bang, Your Majesty!’

- (22) *Chudno: smotrit Matrena na svoego na milogo druga: u Petra telo eshche skvoznoe, vidno, kak purpurnaya v nem perelivaetsya krov', a s levoi storony grudi, gde serdtse, lapchatyi plyashet ogon' – i tuda, i syuda: tuk-tuk-tuk, tuk-tuk-tuk.* [A. Belyi. Serebryanyi golub' (1909)]

‘How odd: Matrena is looking at her dear friend: Petr has such a reach-through body, one can see how purple blood is flowing in it, and on the left side of the chest, where the heart is, palmate flames are dancing – here and there: thump-thump-thump, thump-thump-thump.’

It does not mean, however, that two or even more functions could not be combined in one and the same utterance. One can certainly claim that in examples (21) and (22) the speaker or author does not only aim at making the utterance aesthetically appealing (i.e. poetic), but also wishes to attract attention.

The metalinguistic function is not frequent but can be traced when verbal interjections are used to imitate the original code of language, i.e. when they mimic the onomatopoeia initially employed to depict a certain acoustic event. In this case, instead of poeticizing or being phatic, they convey information on another level (it is a meta-imitation of the original reference to the corresponding sound). As a matter of fact, this information does not deal with signaling emotions, either. For example, the expression *sdelat' tyap-lyap* can be translated into English as ‘to do (something) pell-mell (sloppily, in a slipshod way)’. Both *tyap* and *lyap* exist as onomatopoeias. *Tyap*, for example, can refer to a sound produced by a pickaxe, while *lyap* can stand for plopping of something liquid.

- (23) – *Vzyl i propil: i net tebe – “fuk”; i – voz’mu; i opyat’ zhe – prop’yu.*
[A. Belyi. Moskva, part 1: Moskovskii chudak (1926)]

‘I have just blown it on a booze: and it’s gone – “puff -INTJ”; and then I will do it again.’

Originally, *fuk* was used to imitate the sound of something flying past very fast or burning down. In this context it refers to items or goods disappearing in an instant, like money or valuables being swapped for a drink.

Finally, some of these words can be “ideophonic”, i.e. they may refer to an idea portrayed by a sound. Oswalt [Oswalt 1994, p. 302] offers the following example in English, in which the sound of Cupid’s arrow symbolizes love at first sight:

- (24) *Our eyes met. Zing! Cupid!* [Radio commercial (1985)]

The basic idea behind it is that such words are deemed to evoke vivid sensory imagery. Similar utterances can be found in the Russian National Corpus:

- (25) *Predstavlyaetsya takoi priyatnyi belokuryi obraz: pukhlen’kie shchechki, chulki ne sinie, a s kruzhavchikami, vidnymi v razreze yubki. Cok-cok, dobroe utro, mal’chiki!* [A.A. Matveeva. *Obstoyatel’stvo vremeni* (2012)]

‘Such a pleasant fair-haired image is drawn in mind: plump cheeks, stockings, and not deep blue but with embroidery, visible through the slit. Clank-clank, good morning, boys!’

- (26) *Shchelk – i menya net, mne teper’ èto ne nuzhno videt’, peremalyvat’, toskovat’.* [A. Ilichevskii. *Pers* (2009)]

‘Click-INTJ – and I do not exist anymore, now I do not need to see it, consider, and miss.’

In example (25) the sound of high heels clicking on the floor stands for an attractive young woman. The sound of a switch being turned off symbolizes death, as shown in example (26). When used in this function, these words offer a vivid image of the whole concept. They do not merely depict, but rather give an instantaneous stereographic projection of an idea or situation.

Conclusions

Onomatopoeic verbal interjections are used instead of ordinary words due to a number of factors. Simple morphological structure, transparent onomatopoeia-based meaning and syntactic mobility make them more economical, easy to process and rather frequent. At the same time, words like *bukh*, *stuk* and *bul'* are not that primitive. They tend to be contextually very complex. This paper demonstrates that onomatopoeic verbal interjections perform various pragmatic functions in discourse. With their help it is possible to depict natural sounds or indicate action connected to this sound; express mental or emotional states of the speaker; attract listener's attention; poetically stylize the utterance; imitate the original onomatopoeia in order to generate new meanings; depict an idea symbolized by a certain sound imitation.

References

- Ameka, F. (1992), "Interjections: The universal yet neglected part of speech", *Journal of Pragmatics*, vol. 18, no. 2, pp. 101–118.
- Boronnikova N.V. and Verizhnikova E.V. (2014), "Interjections in the Macedonian language (developing the dictionary concept)", in Karanfilovski, M. ed., *Rusko-makedonski jazični, literaturni i kulturnivski*, vol. 5, Faculty of Philology "Blazhe Koneschi", Skopje, Macedonia, pp. 59–82.
- Clark, H.H. and Gerrig, R.J. (1990), "Quotations as Demonstrations", *Language*, vol. 66, no. 4, pp. 764–805.
- Cuenca, M. J. (2000), "Defining the indefinable? Interjections", *Syntaxis*, vol. 3, pp. 29–44.
- Dingemans, M. and Akita, K. (2017), "An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese", *Journal of Linguistics*, vol. 53, no. 3, pp. 501–532.
- Efremova, T.F. (2000). *Novyi slovar' russkogo yazyka: V 2 t.* [The new dictionary of the Russian language. In 2 vol.], *Russkii yazyk*, Moscow, Russia.
- Foolen, A. (2012), "The relevance of emotion for language and linguistics", in Foolen, A., Lüdke, U.M., Racine, T.P. and Zlatev, J. (eds.), *Moving ourselves, moving others: motion and emotion in intersubjectivity, consciousness and language*, John Benjamins, Amsterdam, Netherlands, pp. 349–368.
- Frege, G. (1949 [1892]), "On sense and nominatum", in Feigl, H. and Sellars, W. (eds.), *Readings in philosophical analysis*, Appleton-Century-Crofts, New York, USA, pp. 85–102.
- Givón, T. (1985), "Iconicity, isomorphism and non-arbitrary coding in syntax", in Haiman, J. (ed.), *Iconicity in syntax*, John Benjamins, Amsterdam, Netherlands, pp. 187–219.
- Givón, T. (1991), "Isomorphism in the grammatical code: Cognitive and biological considerations", *Studies in Language*, vol. 15, no. 1, pp. 85–114.

- Grice, H. P. (1975), "Logic and conversation", in Cole, P. and Morgan, J.C. (eds.), *Syntax and semantics, vol. 3: Speech acts*, Academic Press, New York, USA, pp. 41–58.
- Haiman, J. (1983), "Iconic and economic motivation", *Language*, vol. 59, pp. 781–819.
- Hinton, L., J.J. Ohala and Nichols J. (1994), *Sound Symbolism*, Cambridge University Press, Cambridge, UK.
- Jakobson, R. (1960), "Linguistics and poetics", in Sebeok, T.A. (ed.), *Style in language*, MIT Press, Cambridge, UK, pp. 350–377.
- Kanerva, O. (2018), "Correlation between expressiveness and syntactic independence of Russian onomatopoeic verbal interjections", *Poljarnyj Vestnik*, vol. 21, pp. 15–30.
- Kanerva, O.A. and Viimaranta, J. (2018), "Funktional'naya motivatsiya znakov preprianiya so zvukopodrazhatel'nymi glagol'nymi mezhdometiyami" [Functional motivation of punctuation with onomatopoeic verbal interjections], *Vestnik Tomskogo Gosudarstvennogo universiteta. Filologiya*, vol. 52, pp. 83–97.
- Kartsevskii, S. (1984), "Introduction to the study of interjections", *Voprosy yazykoznaviya*, vol. 6, pp. 127–137.
- Kor Chahine, I. (2008), "Splash! – splash – to splash. On the evolution of narrative predicates in the light of corpus data", in Mustajoki, A., Kopotev, M., Birjulin, L. and Protassova, E. (eds.), *Instrumentarij rusistiki: korpusnye podhody*, Department of Slavonic and Baltic Languages and Literatures, Helsinki, Finland, pp. 152–162. (Slavica Helsingiensia 34).
- Kryk, B. (1992), "The pragmatics of interjections: The case of Polish no", *Journal of Pragmatics* 18, pp. 193–207.
- Meinard, M. (2015), "Distinguishing onomatopoeias from interjections", *Journal of Pragmatics*, vol. 76, pp. 150–168.
- Nikitina, T. (2012), "Russian verboids: A case study in expressive vocabulary", *Linguistics*, vol. 50, no. 2, pp. 165–189.
- Oswalt, R.L. (1994), "Inanimate imitatives in English", in Hinton, L., Nichols, J., and Ohala, J.J. (eds.), *Sound Symbolism*, Cambridge University Press, Cambridge, UK, pp. 293–308.
- Sharonov, I.A. (2008), *Mezhdometiya v rechi, tekste i slovare* [Interjections in speech, text and dictionary], RGGU, Moscow, Russia.
- Shvedova, N.Ju. (ed.) (1980), *Russkaya grammatika* [Russian grammar], Nauka, Moscow, Russia.
- Valgina, N.S. (1979), *Russkaya punktuatsiya: printsipy i naznachenie: Posobie dlya uchitelei* [Russian punctuation. Principles and functions. Teacher's Guide], Prosveshchenie, Moscow, Russia.
- Viimaranta, J. (2015), *Verbal aspect in onomatopoeic interjections in Russian*, Paper presented on 13.12.2015 at Slavic Cognitive Linguistics Conference, Oxford, UK.
- Waugh, L.R. and Newfield, M. (1995), "Iconicity in the lexicon and its relevance for a theory of Morphology", in Landsberg, M.E. (ed.), *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, Mouton de Gruyter, Berlin, Germany, pp. 189–222.
- Wierzbicka, A. (ed.) (2003), "Interjections across cultures", in *Cross-Cultural Pragmatics: The Semantics of Human Interaction*, Mouton de Gruyter, Berlin, Germany, pp. 285–340.
- Wilkins, D. (1992), "Interjections as deictics", *Journal of Pragmatics*, vol. 18, no. 2, pp. 119–158.
- Zipf, G. (1935), *The psychobiology of language*, Houghton Mifflin, Boston, USA.

Литература

- Боронникова, Верижникова 2014 – *Боронникова Н.В., Верижникова Е.В.* Междометия в македонском языке (разработка концепции словаря) // Руско-македонски јазични, литературни и културни врски 5 / Во редакција на Максим Каранфиловски. Скопје: Филолошки факултет «Блаже Конески», 2014. С. 59–82.
- Валгина 1979 – *Валгина Н.С.* Русская пунктуация: принципы и назначение: Пособие для учителей. М.: Просвещение, 1979.
- Ефремова 2000 – *Ефремова Т.Ф.* Новый словарь русского языка: В 2 т. М.: Русский язык, 2000.
- Каневра, Виймаранта 2018 – *Каневра О.А., Виймаранта Й.* Функциональная мотивация знаков препинания со звукоподражательными глагольными междометиями // Вестн. Том. гос. ун-та. Филология. 2018. № 52. С. 83–97.
- Карцевский 1984 – *Карцевский С.* Введение в изучение междометий // Вопросы языкознания. 1984. № 6. С. 127–137.
- Кор Шаин 2008 – *Кор Шаин И.* Плех! – плех – плехнуть(ся). К вопросу об эволюции нарративных предикатов в свете корпусных данных // Инструментарий русистики: корпусные подходы / Под ред. А. Мустайоки, М.В. Копотева, Л.А. Бирюлина, Е.Ю. Протасовой. Хельсинки: Department of Slavonic and Baltic Languages and Literatures, 2008. С. 152–162. (Slavica Helsingiensia 34)
- Русская грамматика 1980 – *Русская грамматика: В 2 т. / Гл. ред. Н.Ю. Шведова.* М.: Наука, 1980.
- Шаронов 2008 – *Шаронов И.А.* Междометия в речи, тексте и словаре. М.: РГГУ, 2008. 296 с.
- Ameka 1992 – *Ameka F.* Interjections: The universal yet neglected part of speech // *Journal of Pragmatics*. 1992. Vol. 18, no. 2. P. 101–118.
- Clark, Gerrig 1990 – *Clark H.H., Gerrig R.J.* Quotations as Demonstrations // *Language*. 1990. Vol. 66, no. 4. P. 764–805.
- Cuenca 2000 – *Cuenca M.J.* Defining the indefinable? Interjections // *Syntaxis*. 2000. Vol. 3. P. 29–44.
- Dingemanse, Akita 2017 – *Dingemanse M., Akita K.* An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese // *Journal of Linguistics*. 2017. Vol. 53, no. 3. P. 501–532.
- Foolen 2012 – *Foolen A.* The relevance of emotion for language and linguistics // *Moving ourselves, moving others: motion and emotion in intersubjectivity, consciousness and language* / A. Foolen, U.M. Lüdke, T.P. Racine, J. Zlatev (eds.). Amsterdam: John Benjamins, 2012. P. 349–368.
- Frege 1949 [1892] – *Frege G.* On sense and nominatum // *Readings in philosophical analysis* / H. Feigl, W. Sellars, (eds.). New York: Appleton-Century-Crofts, 1949 [1892]. P. 85–102.
- Givón 1985 – *Givón T.* Iconicity, isomorphism and non-arbitrary coding in syntax // *Iconicity in syntax* / J. Haiman (ed.). Amsterdam: John Benjamins, 1985. P. 187–219.
- Givón 1991 – *Givón T.* Isomorphism in the grammatical code: Cognitive and biological considerations // *Studies in Language*. 1991. Vol. 15, no. 1. P. 85–114.
- Grice 1975 – *Grice H.P.* Logic and conversation // *Syntax and semantics*, vol. 3: *Speech acts* / P. Cole, J.C. Morgan (eds.). New York: Academic Press, 1975. P. 41–58.

- Haiman 1983 – *Haiman J.* Iconic and economic motivation // *Language*. 1983. Vol. 59. P. 781–819.
- Hinton, Ohala, Nichols 1994 – *Hinton L., Ohala J.J., Nichols J.* *Sound Symbolism*. Cambridge: Cambridge University Press, 1994.
- Jakobson 1960 – *Jakobson R.* *Linguistics and poetics* // *Style in language* / T.A. Sebeok (ed.). Cambridge: MIT Press, 1960. P. 350–377.
- Kanerva 2018 – *Kanerva O.* Correlation between expressiveness and syntactic independence of Russian onomatopoeic verbal interjections // *Poljarnyj Vestnik*. 2018. Vol. 21. P. 15–30.
- Kryk 1992 – *Kryk B.* The pragmatics of interjections: The case of Polish no // *Journal of Pragmatics*. 1992. Vol. 18. P. 193–207.
- Mainer 2015 – *Mainer M.* Distinguishing onomatopoeias from interjections // *Journal of Pragmatics*. 2015. Vol. 76. P. 150–168.
- Nikitina 2012 – *Nikitina T.* Russian verboids: A case study in expressive vocabulary // *Linguistics*. 2012. Vol. 50, no. 2. P. 165–189.
- Oswalt 1994 – *Oswalt R.L.* Inanimate imitatives in English // *Sound Symbolism* / L. Hinton, J. Nichols, J.J. Ohala (eds.). Cambridge: Cambridge University Press, 1994. P. 293–308.
- Viimaranta 2015 – *Viimaranta J.* Verbal aspect in onomatopoeic interjections in Russian. Paper presented on 13.12.2015 at Slavic Cognitive Linguistics Conference. Oxford, 2015.
- Waugh, Newfield 1995 – *Waugh L.R., Newfield M.* Iconicity in the lexicon and its relevance for a theory of Morphology // *Syntactic Iconicity and Linguistic Freezes: The Human Dimension* / M.E. Landsberg (ed.). Berlin: Mouton de Gruyter, 1995. P. 189–222.
- Wierzbicka 2003 – *Interjections across cultures* // *Cross-Cultural Pragmatics: The Semantics of Human Interaction* / A. Wierzbicka (ed.). Berlin: Mouton de Gruyter, 2003. P. 285–340.
- Wilkins 1992 – *Wilkins D.* Interjections as deictics // *Journal of Pragmatics*. 1992. Vol. 18, no. 2. P. 119–158.
- Zipf 1935 – *Zipf G.* *The psychobiology of language*. Boston: Houghton Mifflin, 1935.

Информация об авторе

Оксана А. Канерва, MA, Ph.D. Хельсинкский университет, Хельсинки, Финляндия; 00100, Финляндия, Хельсинки, ул. Йлиопистонкату, д. 4; oksana.kanerva@helsinki.fi

Information about the author

Oksana A. Kanerva, MA, Ph.D., Helsinki University, Helsinki, Finland; bld. 4, Yliopistonkatu Str., Helsinki, Finland, 00100; oksana.kanerva@helsinki.fi

In memoriam

УДК 81(092)

DOI: 10.28995/2686-7249-2020-7-147-154

Владимир Зиновьевич Санников (1931–2019)

Леонид П. Крысин

*Институт русского языка им. В.В. Виноградова РАН
Москва, Россия, leonid-krysin@mail.ru*

Леонид Л. Иомдин

*Институт проблем передачи информации им. А.А. Харкевича РАН
Москва, Россия, iomdin@iitp.ru*

18 декабря 2019 г. скончался Владимир Зиновьевич Санников – известный русский языковед, автор ряда работ в таких разных областях лингвистики, как история древнерусского языка, русская морфология и русское словообразование, лингвистическая статистика, автоматическая обработка текстов, теория речевых актов.

Особенно значителен вклад В.З. Санникова в изучение синтаксиса современного русского языка. В своих книгах, посвященных синтаксической проблематике, он учитывал и подвергал анализу не только собственно «строєвые» свойства русского предложения, но и многообразные семантические и прагматические факторы, влияющие на характер поведения слова в составе высказывания. Эта исследовательская концепция представлена в монографиях В.З. Санникова «Русские сочинительные конструкции» (1989) и, в особенности, «Русский синтаксис в семантико-прагматическом пространстве» (2008).

Некоторые работы В.З. Санникова свидетельствуют о желании автора выйти за пределы собственно лингвистического анализа и объяснить такие стороны существования человека и общества, которые хотя и связаны в значительной степени с языком, но содержат черты, относящиеся к психологии человека, к особенностям его интеллектуальной и эмоциональной жизни. Таковы работы

© Крысин Л.П., Иомдин Л.Л., 2020

В.З. Санникова, посвященные природе комического: «Русский каламбур» (1995), «Русский язык в зеркале языковой игры» (2002, 2-е издание), «Русская языковая шутка: от Пушкина до наших дней» (2003), «Краткий словарь русских острот» (2012).

В.З. Санников известен и как автор двух томов талантливо и одновременно бесхитростно написанных воспоминаний: «Записки простодушного» (2003) и «Записки простодушного. Жизнь в Москве» (2018). Далек не всякий человек способен найти в собственном характере какую-то доминирующую черту. В.З. Санникову это удалось: его мемуары называются «Записки простодушного», и именно это свойство – простодушие – было главным в его личности. И еще он был по-детски доверчив и добр.

Он родился 3 августа 1931 г. в городе Воткинске. Предки его – и по отцовской, и по материнской линии – были крестьяне Пермской губернии. Володя был старшим ребенком, а кроме него в семье были два его брата и сестра. Филологическое образование В.З. Санников получил в Пермском государственном университете. По окончании университета он поехал в Москву и в сентябре 1955 г. поступил в аспирантуру Института языкознания АН СССР. Закончив аспирантуру, он сначала работал в том же институте в должности младшего научного сотрудника, а в 1958 г., когда был образован Институт русского языка АН СССР, перешел на работу в этот институт.

Сначала профессиональные интересы молодого ученого были сосредоточены в области истории русского языка: он и начал свою научную деятельность по окончании аспирантуры в секторе сравнительно-исторического изучения восточнославянских языков и в этом секторе занимался исследованием синтаксиса древнерусского языка. Результаты этой исследовательской работы отражены в его кандидатской диссертации «Сложноподчиненное предложение в древнерусских памятниках XI–XIV вв.», которую он защитил в 1965 г. Спустя некоторое время В.З. Санников заинтересовался более современной научной тематикой. В частности, в составе группы молодых ученых по заданию правительства СССР участвовал в создании системы автоматической обработки деловой документации. Интерес к этой проблематике сохранился у В.З. Санникова и впоследствии, когда по политическим причинам (руководство Института русского языка преследовало его и еще нескольких сотрудников института за участие в поддержке несправедливо осужденных правозащитников) он вынужден был уволиться из академического института и поступить на работу в ведомственный институт Министерства электротехнической промышленности СССР – «Информэлектро». В этом институте под

руководством будущего академика (а в ту пору кандидата филологических наук) Ю.Д. Апресяна и по его инициативе небольшая группа лингвистов и математиков начала разработку системы французско-русского автоматического перевода технических текстов.

В созданной группе В.З. Санников работал над системой автоматического морфологического анализа и синтеза русского текста. Значительная часть этой деятельности состояла в создании компьютерного морфологического словаря русского языка. Эта работа была продолжена и в дальнейшем, когда в середине 80-х гг. XX в. этой исследовательской группе в почти полном ее составе удалось перейти на работу в Институт проблем передачи информации АН СССР, где была образована Лаборатория компьютерной лингвистики под руководством Ю.Д. Апресяна. Здесь В.З. Санников работал практически до конца своей служебной деятельности. Автоматическая система морфологического анализа и синтеза русского языка, в создании которой В.З. Санников сыграл решающую роль, легла несколько позднее в основу поисковой системы компании Яндекс: первый вариант этой поисковой системы практически целиком опирался на компьютерный словарь В.З. Санникова.

Одновременно с решением этих технических задач, где В.З. Санников был абсолютно точен и предлагал весьма нетривиальные нововведения (касающиеся, например, автоматического анализа сложных слов), он продолжал заниматься и такими более близкими его душе направлениями исследования, как синтаксис русского языка и лингвистическая природа комического. В 1987 г. он успешно защитил докторскую диссертацию на тему «Русские сочинительные конструкции».

Почти до самых последних дней своей жизни В.З. Санников занимался наукой. В частности, в журнале «Известия РАН. Серия литературы и языка» он опубликовал большие статьи «Об оценочных сложноподчиненных предложениях в русском языке» (2015), «Об оценочных суффиксах русского языка» (2018), «О парных речевых актах русского языка» (2019). С публикацией последней статьи ее автор успел познакомиться за несколько дней до своей кончины...

Владимир Зиновьевич Санников был прекрасным человеком и замечательным ученым. Хочется, чтобы память о нем сохранилась в сердцах тех, кто знал его, кому он был душевно близок, и тех, кому были интересны его оригинальные научные исследования.

Основные труды В.З. Санникова

Книги

- Санников В.З.* Русские сочинительные конструкции. Семантика, прагматика, синтаксис. М.: Наука, 1989.
- Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л., Лазурский А.В., Перцов Н.В., Санников В.З., Цинман Л.Л.* Лингвистическое обеспечение системы ЭТАП-2. М.: Наука, 1989.
- Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л., Лазурский А.В., Санников В.З., Цинман Л.Л.* Лингвистический процессор для сложных информационных систем. М.: Наука, 1992.
- Санников В.З.* Русский каламбур. 1200 каламбуров старых и современных. New Jersey: Hermitage Publishers, 1995.
- Санников В.З.* Русский язык в зеркале языковой игры. М.: Языки русской культуры, 1999; 2-е изд. М., 2002.
- Санников В.З.* Русская языковая шутка: от Пушкина до наших дней. М.: Аграф, 2003.
- Санников В.З.* Русский синтаксис в семантико-прагматическом пространстве. М.: Языки славянских культур, 2008.
- Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л., Санников В.З.* Теоретические проблемы русского синтаксиса. Взаимодействие грамматики и словаря. М.: Языки славянских культур, 2010.
- Санников В.З.* Краткий словарь русских острот. М.: Языки славянских культур, 2012.

Диссертации

- Санников В.З.* Сложноподчиненное предложение в древнерусских памятниках XI–XIV вв.: Дис. ... канд. филол. наук. М., 1965.
- Санников В.З.* Русские сочинительные конструкции: Дис. ... д-ра филол. наук. М., 1987.

Мемуары

- Санников В.З.* Записки простодушного. М.: Аграф, 2003.
- Санников В.З.* Записки простодушного. Жизнь в Москве. М.: Языки славянских культур, 2018.

Leonid P. Krysin

*V.V. Vinogradov Institute of Russian Language of the Russian Academy of Sciences,
Moscow, Russia, leonid-krysin@mail.ru*

Leonid L. Iomdin

*Institute for Information Transmission Problems of the Russian Academy of Sciences
(Kharkevich Institute), Moscow, Russia, iomdin@iitp.ru*

Vladimir Zinovyevich Sannikov (1931–2019)

Vladimir Zinovyevich Sannikov, a well-known Russian linguist, passed away on December 18, 2019. He was the author of a number of works in the history of the Old Russian language, Russian morphology and Russian word formation, linguistic statistics, natural language processing, and the theory of speech acts.

Yet Vladimir Sannikov's most important scientific contribution was in the domain of modern Russian syntax. In his books and papers focused on syntactic issues, he analyzed both fundamental syntactic phenomena of the Russian sentence and multiple semantic and pragmatic factors affecting the word's behavior in an utterance. This approach is presented in Sannikov's book on Russian coordinative constructions (1989) and especially in an innovative work on Russian syntax in the semantic and pragmatic space (2008).

Some of Vladimir Sannikov's works go far beyond linguistics proper and explain such aspects of the existence of a person and society, which, although they are largely associated with language, contain features related to human psychology, intellectual and emotional life of man. Such are the works of V.Z. Sannikov on the nature of the comic: "Russian pun" (1995), Russian language in the mirror of a language game" (2002, 2nd edition), "A concise dictionary of Russian witticisms" (2012).

V.Z. Sannikov is also known as the author of two volumes of talented and at the same time ingeniously written memoirs: "Notes of a simple-minded" (2003) and "Notes of a simple-minded. Life in Moscow" (2018). Not every person is able to find some dominant trait in his own character. Indeed, simplemindedness was one of the main traits of his personality. He was childishly trustful and very kind.

He was born on the 3rd of August, 1931, in the Ural town of Votkinsk. His paternal and maternal ancestors were peasants of the Perm province. Volodya was the eldest child, and besides him there were two of his brothers and a sister in the family. He got his philological education in the Perm State University. Upon graduation, he moved to Moscow and in September of 1955 became a post-graduate student at the Institute of Linguistics, Russian Academy of Sciences. Upon com-

pleting the program, he worked as a researcher at that very Institute till 1958 when at the Academy of Sciences of the USSR the Institute of the Russian Language was formed and he went to work there.

First he focused on the history of Russian language. The results of this research work are reflected in his Ph.D. thesis "A complex sentence in Old Russian monuments of the 11th–14th centuries", which he defended in 1965. Later, his scientific interests shifted to computational linguistics and natural language processing. When he was dismissed from the Russian Language institute for political reasons (because he had signed several letters in support of the dissidents who were unjustly convicted by Soviet authorities), he was able to join a research group of an applied science institute, *Informelektro*.

The group was headed by Juri Apresjan, the future Academician and a prominent linguist, and developed one of the first systems of machine translation in the Soviet Union. In this project, Vladimir Sannikov was responsible for the development of the morphological component of Russian text analysis and generation. He continued with this work in the Academic Institute of Information Transmission Problems, where the bulk of Apresjan's research group was invited in mid-1980s. In the Laboratory of Computational Linguistics of this Institute, formed by this group, Vladimir Sannikov worked until very recently, almost until his death. The automatic system of morphological analysis and synthesis of the Russian language, in the creation of which V.Z. Sannikov played a decisive role, formed the basis of the Yandex search engine somewhat later: the first version of that search engine relied almost entirely on V.Z. Sannikov's computer dictionary.

In parallel with the solution of those technical issues, where V.Z. Sannikov was absolutely accurate and proposed very non-trivial innovations (concerning, for example, automatic analysis of composite words), he continued to study the areas of research that were closer to his soul, such as the syntax of the Russian language and the linguistic nature of the comic. In 1987 he successfully defended his doctoral dissertation on the topic "Russian coordinative constructions".

Almost until the very last days of his life V.Z. Sannikov was pursuing his linguistic research. In particular, in the journal "Izvestia RAN (RAS). Series of Literature and Language", he published great articles: "On evaluative complex sentences in the Russian language" (2015), "On evaluative suffixes of the Russian language" (2018), "On paired speech acts of the Russian language" (2019). The author was able to see his last article published several days before his death ...

Vladimir Zinovievich Sannikov was a wonderful person and a first-class scientist. The fond memory of him will stay in the hearts of those who knew him, to whom he was spiritually close, and who were interested in his original scientific views and findings.

The main works of V.Z. Sannikov

Books

- Sannikov, V.Z. (1989), *Russkie sochinitel'nye konstruksii. Semantika, pragmatika, sintaksis* [Russian coordinative constructions. Semantics, pragmatics, syntax], Nauka, Moscow, Russia.
- Apresyan, Yu.D., Boguslavskii, I.M., Iomdin, L.L., Lazurskii, A.V., Pertsov, N.V., Sannikov, V.Z. and Tsinman, L.L. (1989), *Lingvisticheskoe obespechenie sistemy ETAP-2* [Linguistic support of the ETAP-2 system], Nauka, Moscow, Russia.
- Apresyan, Yu.D., Boguslavskii, I.M., Iomdin, L.L., Lazurskii, A.V., Sannikov, V.Z. and Tsinman, L.L. (1992), *Lingvisticheskii protsessor dlya slozhnykh informatsionnykh sistem* [Linguistic processor for complex information systems], Nauka, Moscow, Russia.
- Sannikov, V.Z. (1995), *Russkii kalambur. 1200 kalamburov starykh i sovremennykh* [Russian pun. 1200 puns old and modern], Hermitage Publishers, New Jersey, USA.
- Sannikov, V.Z. (1999), *Russkii yazyk v zerkale yazykovoi igry* [Russian language in the mirror of a language game], Yazyki russkoi kul'tury, Moscow, Russia.
- Sannikov, V.Z. (2003), *Russkaya yazykovaya shutka: ot Pushkina do nashikh dnei* [The Russian language joke. From Pushkin to the present day], Agraf, Moscow, Russia.
- Sannikov, V.Z. (2008), *Russkii sintaksis v semantiko-pragmaticheskom prostranstve* [Russian syntax in the semantic-pragmatic space], Yazyki slavyanskikh kul'tur, Moscow, Russia.
- Apresyan, Yu.D., Boguslavskii, I.M., Iomdin, L.L. and Sannikov, V.Z. (2010), *Teoreticheskie problemy russkogo sintaksisa. Vzaimodeistvie grammatiki i slovary* [Theoretical issues of Russian syntax. Interaction of grammar and vocabulary], Yazyki slavyanskikh kul'tur, Moscow, Russia.
- Sannikov, V.Z. (2012), *Kratkii slovar' russkikh ostrot* [A concise dictionary of Russian witticisms], Yazyki slavyanskikh kul'tur, Moscow, Russia.

Dissertations

- Sannikov, V.Z. (1965), *Slozhnopodchinennoe predlozhenie v drevnerusskikh pamyatnikakh XI–XIV vv.* [A complex sentence in Old Russian monuments of the 11th – 14th centuries], Cand. of Sci. Thesis, Moscow, Russia.
- Sannikov, V.Z. (1987), *Russkie sochinitel'nye konstruksii* [Russian coordinative constructions], Dr. of Sci. Thesis, Moscow, Russia.

Memoirs

- Sannikov, V.Z. (2003), *Zapiski prostodushnogo* [Notes of a simple-minded] Agraf, Moscow, Russia.
- Sannikov, V.Z. (2018), *Zapiski prostodushnogo. Zhizn' v Moskve* [Notes of a simple-minded. Life in Moscow], Yazyki slavyanskikh kul'tur, Moscow, Russia.

Информация об авторах

Леонид П. Крысин, доктор филологических наук, профессор, Институт русского языка им. В.В. Виноградова РАН, Москва, Россия; 119019, Россия, Москва, ул. Волхонка, д. 18/2; leonid-krysin@mail.ru

Леонид Л. Иомдин, кандидат филологических наук, профессор, Институт проблем передачи информации им. А.А. Харкевича РАН, Москва, Россия; 127994, Россия, Москва, Большой Каретный пер., д. 19, стр. 1; iomdin@iitp.ru

Information about the authors

Leonid P. Krysin, Dr. of Sci. (Philology), professor, V.V. Vinogradov Institute of Russian Language of the Russian Academy of Sciences, Moscow, Russia; bldg. 18/2, Volkhonka Str., Moscow, Russia, 119019; leonid-krysin@mail.ru

Leonid L. Iomdin, Cand. of Sci. (Philology), professor, Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Moscow, Russia; bldg. 1, bld. 19, Bol'shoi Karenyi Lane, Moscow, Russia, 127994, iomdin@iitp.ru

Дизайн обложки

Е.В. Амосова

Корректор

Н.К. Егорова

Компьютерная верстка

Е.Б. Рагузина

Подписано в печать 26.08.2020.
Формат 60×90¹/₁₆
Уч.-изд. л. 7,8. Усл. печ. л. 9,8.
Тираж 1050 экз. Заказ № 994

Издательский центр
Российского государственного
гуманитарного университета
125993, Москва, Миусская пл., 6
www.rggi.ru
www.knigirggi.ru