

РОССИЙСКАЯ АКАДЕМИЯ НАУК  
ИНСТИТУТ ВОСТОКОВЕДЕНИЯ РАН

ПРОБЛЕМЫ ОБЩЕЙ  
И ВОСТОКОВЕДНОЙ  
ЛИНГВИСТИКИ

Сочетаемость языковых единиц  
и языковые модели.  
Памяти З. М. Шаляпиной (1946–2020)

Москва  
ИВ РАН  
2021

ББК 81  
П78

*Ответственный редактор*

А. И. Коган

*Редактор-составитель*

А. С. Панина

*Рецензенты:*

С. А. Бурлак, И. Б. Иткин

*Утверждено к печати Издательской комиссией  
Ученого совета Института востоковедения РАН*

П78 **Проблемы** общей и востоковедной лингвистики. Сочетаемость языковых единиц и языковые модели. Памяти З.М. Шаляпиной (1946-2020) : сб. ст. / отв. ред. А. И. Коган, ред.-сост. А. С. Панина; Ин-т востоковедения РАН – М.: ИВ РАН, 2021. – 321 с.

Studies in General and Oriental Linguistics. Language models, combinability and grammar. In memory of Z. Shalyapina (1946-2020) / Ed. by A. Kogan. – Moscow: IOS RAS, 2021. – 321 p.

ISBN 978-5-907543-08-9

DOI: 10.31696/978-5-907543-08-9

Сборник основан на материалах конференции «Востоковедные чтения 2021. Сочетаемость языковых единиц и языковые модели. Памяти З.М. Шаляпиной (1946-2020)» (Институт востоковедения РАН, 14-15.04.2021) и охватывает широкий круг тем, включая проблему слова, морфологические и синтаксические модели, фразеологию и лексическую семантику.

The volume is based on the proceedings of the conference on the languages of Asia and Africa held at the Institute of Oriental Studies on 14-15.04.2021, dedicated to the memory of Z. Shalyapina (1946-2020). Discussing various languages of Asia and Africa, it covers a broad range of topics from definition of word, morphological and syntactic word classes to phraseology lexical semantics.

## СОДЕРЖАНИЕ

### **З. М. Шаляпина: жизнь и труды**

Алпатов В. М. <i>Воспоминания о З.М. Шаляпиной</i> .....	5
Горелова Л. М. <i>Зоя Михайловна Шаляпина (другу и ученому)</i> .....	9
Демиц К. И., Костыркин А. В. <i>Разработка З. М. Шаляпиной концепции языковых ярусов в рамках трехмерной стратификационной модели языка</i> .....	15
Мальшева С. Д. <i>Библиография трудов Зои Михайловны Шаляпиной</i> .....	25

### **Грамматика, синтаксис**

Борисова Е. Г., Маннапова С. А. <i>Типология средств модификации содержания сообщения (на материале русских и китайских частиц)</i> .....	47
Гращенков П. В. <i>Контекстная валентность по [Шаляпина 2007] и несобственные участники прилагательных</i> .....	56
Лазуткина Е. М. <i>Процессы фразеологизации в синтагматике на разных уровнях языковой системы русского языка</i> .....	64
Погибенко Т. Г. <i>Возникновение новых языковых моделей как ответ на коммуникативные вызовы дискурса</i> .....	74
Рудницкая Е. Л. <i>Употребление и характеристики глагола bi- 'быть' в эвенкийском языке</i> .....	89
Beletskiy S. <i>Cigogo nominal demonstratives: morphology and semantics</i> .....	100
Lander Yu. <i>Fronted topics, clefted foci and the parameter of integration into the clause</i> .....	112
Yakubovich I. <i>Hittite agent form</i> .....	124

### **Морфология, проблемы слова**

Антонян К. В. <i>О грамматическом и словарном аспектах слова в современном китайском языке</i> .....	142
Белова А. Г. <i>Грамматическая модель и фонетическая система классического арабского языка</i> .....	151
Блинов А. А. <i>Акронимы в современном арабском языке</i> .....	162
Дербишева З. К. <i>Специфика дистрибуции агглютинативных аффиксов в глагольной словоформе турецкого языка</i> .....	181
Зверев А. С. <i>Графическое слово в зеркале японского языка: к уточнению понятия</i> .....	191
Столбова О. В. <i>Модель словоформы в чадских языках</i> .....	209
Эдельман Д. И. <i>Сочетаемость праязыковых элементов и структурная трансформация иранских языков</i> .....	218

### **Лексикология, лексикография**

Будкина Ю. В. <i>Структура лексико-семантического поля соматических единиц на материале фразеологического фонда персидского языка</i> .....	226
---	-----

Гасунс М. Ю. <i>Новое в санскритской компьютерной лингвистике</i> .....	236
Додыхудоева Л. Р. <i>Структурно-семантические особенности глаголов «тянуть, тащить», «курить» и их деривативов в таджикском и шугнанском языках</i> .....	246
Дудченко Г. Б. <i>Приграничный русско-китайский пиджин 90-х годов XX в.</i>	264
Емельченкова Е. Н. <i>Феномен японских заимствований в китайском языке: проблемы упорядочения терминологии</i> .....	275
Кретов А. А., Гасунс М. Ю., Леонченко В. В. <i>Параметрический анализ «Санскритско-русского словаря» В.А. Кочергиной</i> .....	287
Смирнитская А. А. <i>Моделирование семантического перехода: обзор направлений исследования (на материале словарной статьи <i>viṣu</i> 'падать' в словаре <i>Tamil Lexicon</i>)</i> .....	302
<b>Сведения об авторах</b> .....	319

Воспоминания о З. М. Шаляпиной

Владимир Михайлович Алпатов

Институт языкознания РАН, Москва, Россия,

v-alpatov@iling-ran.ru, ORCID 0000-0003-4323-2832

*Аннотация.* Автор вспоминает о своем многолетнем знакомстве с З.М. Шаляпиной, начавшемся еще в студенческие годы, о ее пути в науке от математической лингвистики к японско-русскому автоматическому переводу, о встрече в Японии и о научной и административной работе З.М. Шаляпиной в Отделе языков стран Азии и Африки в Институте востоковедения РАН.

*Ключевые слова:* З.М. Шаляпина, лингвистика в России, японистика в России, автоматический перевод

*Abstract.* The author reminiscences on his long acquaintance with Zoya Shalyapina, dating back to student days, on her academic interests leading her from computational linguistics to Japanese-Russian machine translation, her travels to Japan, and the many years of scientific and administrative work at the Department of Languages of Asia and Africa of the Institute of Oriental studies.

*Keywords:* Z. Shalyapina, Russian linguistics, Japan studies in Russia, machine translation

Я познакомился с Зоей Шаляпиной в конце апреля 1967 г. во время второй межвузовской конференции по структурной, прикладной лингвистике и машинному переводу в МГУ. Она тогда училась на третьем курсе соответствующего отделения в МГПИИЯ, я на четвёртом курсе ОСИПЛ МГУ.

Тогда в нашей стране вузовское образование в области, чаще всего неофициально именовавшейся математической лингвистикой, находилось на подъёме. Отделения, аналогичные нашим, появились в нескольких крупных городах. Это было частью общемирового процесса, но при этом в СССР, мало зная о том, что происходит за границей, до всего доходили сами, причём, как следовало из рассказов участников студенческих конференций, в разных вузах по-разному. В Москве появились вышеупомянутые отделения в МГУ и МГПИИЯ. Мы и они приглашали друг друга, встречались на конференциях и симпозиумах, но особой дружбы не было. Подготовка студентов была различной, причём в «ИнЯзе» она, как я теперь понимаю, была лучше продумана с точки зрения решения практических задач, что потом сказало на профессиональных возможностях Зои. Там студентов активно включали в развернувшуюся в институте прикладную деятельность. Работали и умелые прикладники, и тогдашние лидеры «передовой лингвистики» вплоть до «самого» И.А. Мельчука, который не допускался на филологический факультет

МГУ (хотя в Институте восточных языков при университете он иногда выступал). А в «главном» столичном вузе с преподаванием лингвистических и особенно прикладных дисциплин было немало затруднений, зато было значительно больше математики, что само по себе было бы и хорошо, но математическим дисциплинам не было противовеса (а если был, то скорее гуманитарный, как в моей японской группе). Тем не менее из обоих вузов вышло немало специалистов. К сожалению, отделение в МГПИИЯ было довольно скоро закрыто.

На конференцию приезжали студенты и из Ленинграда, и из Риги, и из Тбилиси, и из Новосибирска. Виден был энтузиазм, но, увы, вне Москвы и отчасти Ленинграда всё довольно быстро заглохло. На конференции я впервые столкнулся с тем, что советская студентка из Риги прочитала доклад на английском языке, что было абсолютно не принято. Она перед докладом пояснила, что английский язык знает лучше, чем русский. Кроме студентов, выступали мэтры (тогда я впервые услышал И. И. Ревзина).

В последний вечер устроили встречу участников в квартире уже создавшей семью Нади Браккер в Марьиной роще. Небольшая комната была забита народом. Были не только студенты, но и А.А. Зализняк, тогда уже доктор наук. Бывший кем-то вроде тамады С.В. Кодзасов сказал: «Всё-таки мы должны помянуть Комарова» (космонавт погиб накануне). Но потом было весело, особенно благодаря Зализняку, разъезжались уже ночью. Вот тогда я и познакомился с Зоей, кажется, единственной там инъязовкой. Мы разговорились, и сразу было видно: человек неординарный.

Но потом мы встречались редко, хотя о Зое я время от времени слышал; знал, что она весьма успешно работает в том же институте, который кончала. Но в 1977 г. мы с ней встретились уже на постоянной основе в академическом Институте востоковедения, где я давно работал, а она только что пришла уже кандидатом наук.

За три года до этого из ведомственного Института патентной информации в ИВ АН перешла в составе семи человек целая лаборатория автоматического перевода с японского языка на русский во главе с С.М. Шевенко. Тогда не очень было ясно, зачем эта лаборатория в институте, занимавшемся совсем другими проблемами, но мой учитель И.Ф. Вардуль сказал: «Семь ставок под ногами не валяются». Скоро выяснилось, что Шевенко, первой в Москве объявившая себя специалистом по машинному переводу с японского языка, не столько что-то делает, сколько делает вид. Если не считать двух лаборанток, оставшихся в лаборатории у Шаляпиной на многие годы, при Шевенко состояли случайные люди. Никто не знал, что делать.

И вдруг появилась Зоя. Я так и не знаю, почему она, уходя из МГПИИЯ, выбрала для дальнейшей работы именно эту лабораторию. Она была во многом замкнутым человеком (хотя при первом знакомстве она показалась совсем иной) и никогда не рассказывала мне об этом.

Надо сказать об истории вопроса. Ещё в начале 60-х гг. зав. кафедрой в МГУ В.А. Звегинцев пришёл к выводу, что надо готовить специалистов по разным наиболее значимым языкам, из которых на первое место поставил японский. Не знаю, может быть, подспудно у него была и другая цель: охватить новыми лингвисти-

ческими методами языка, далёкие от европейских. Последнее вольно или невольно удалось в большей степени, а заметных прикладников не получилось. А Зоя ни в студенческие годы, ни в МГПИИЯ не занималась японским языком. Однако затем она поставила себе цель выучить его и без отрыва от основного места работы поступила на курсы японского языка. К моменту прихода в Институт востоковедения она уже владела им в достаточной степени.

Формально она была подчинена Шевенко, но с самого начала добилась автономии и стала разрабатывать свою систему. Она постепенно избавилась от тех, кто не был занят делами лаборатории: одни вполне успешно нашли себе иное дело по душе, а другие ушли, в том числе в домохозяйки. Сначала она всё разрабатывала одна с помощью лаборанток, потом стали появляться специалисты.

Я уже давно был знаком с Зоей, но первые два года мы общались довольно мало и только по делам. Однако в августе 1979 г. ей удалось (каким образом, я опять-таки не знаю) получить почти месяц стажировки в Токио. Я в это время был там в научной командировке, и Зоя по приезде сразу со мной связалась. И именно тогда мы по-настоящему подружились, перейдя по её предложению на «ты». Она каждый день посещала курсы японского языка для иностранцев, жила в японской семье, много разговаривала на языке. Но времени оставалось много. Для меня это была вторая поездка в Японию, а до её приезда я там уже жил три месяца, поэтому надо было ей помогать. Сразу стало очевидно, насколько она любит путешествовать. Разумеется, надо было выбирать маршруты в пределах одного дня, но мы сумели выбраться и в Хаконэ, и в Камакуру, и в Эносиму и в другие токийские окрестности; в самом Токио были и в синтоистских и буддийских храмах, и на Гиндзе, и в зоопарке.

По дороге мы много разговаривали. Скоро я выяснил, что у нас большое сходство и характеров, и мнений по многим вопросам. В те годы я постоянно ощущал, что многие мои коллеги, особенно моего возраста и моложе, всё больше склонялись к идеям, доходившим через западные «голоса»; советский взгляд на мир был всё больше не в почёте. Однако с Зоей мы сходились во взглядах почти во всём. За месяц, когда мы были в Токио, главной сенсацией было бегство на Запад артиста балета Годунова, и помню, как она его осуждала. Вместе с Годуновым постоянно вспоминали Барышникова, и она сказала, что он оправдал свою фамилию. Много позже она возмущалась словами её коллеги, уезжавшего из СССР, о причинах отъезда: «Не хочу, чтобы мои дети маршировали вокруг Вечного огня». Зоя рассказывала, что в МГПИИЯ она не боялась яростно спорить по политическим вопросам со своими учителями (а я этого избегал). Тем не менее, её профессионализм они признавали.

Потом мы много общались долгие годы. Система Зои развивалась и совершенствовалась, постепенно всё больше её захватывая. В конце концов она принесла в жертву и свою страсть к путешествиям. Вскоре после вышеупомянутой поездки она ещё раз съездила в Японию на конференцию, всего на несколько дней, и побывала в Киото. А потом, если не ошибаюсь, она никогда уже в этой стране не была. В профессиональном отношении это ничего не давало, а на познание мира времени уже не было. И вообще она редко в последующие годы выезжала за пределы Москвы и Подмосковья. Она вела политические споры, когда их многие

избегали, но в годы «перестройки» и «возвращения в мировую цивилизацию» Зоя стала чуждаться политики. Я, наоборот, стал одно время заниматься этим, но привлечь Зою уже не удавалось, хотя взгляды она не меняла: Дело ей было важнее. И, вероятно, поэтому она так и не оформила свои многочисленные разработки в виде докторской диссертации. Зато книгу «Трехмерная стратификационная модель языка и его функционирования: К общей теории лингвистических моделей» с интересной и оригинальной лингвистической концепцией она всё же издала в 2007 году. Думаю, что её идеи ещё будут востребованы.

Но помимо разработки системы Зое пришлось заняться совсем другим делом, где она тоже нашла себя. В конце 1994 г. мне нужно было освободить должность заведующего Отделом языков в связи с переходом на должность заместителя директора. Отдел к тому времени перенёс ряд потерь и заметно уменьшился в составе. Когда встал вопрос о новом заведующем, уже не помню, кому первому пришла в голову кандидатура Зои. Она сперва казалась неожиданной: её лаборатория занимала несколько обособленное положение в отделе. Однако все сразу её поддержали, и она, сохраняя заведование лабораторией, стала возглавлять весь Отдел языков и руководила им больше, чем кто-либо другой: около четверти века. В отделе сформировалась спокойная, домашняя обстановка, почти исчезли склоки, которых когда-то было немало. Сотрудникам мало платили, но хотя бы не мешали работать. В отличие от некоторых других отделов института приходила молодёжь: играли роль традиции ОСИПЛ и созданного во многом по его образцу Института лингвистики РГГУ. А сотрудничество Зоинной лабораторией с практическими организациями сделало её чуть ли не единственным подразделением института, приносящим ему пусть небольшой, но доход. К этому времени уже стало ясно, что дают и что не дают системы автоматического перевода: какие-то ожидания оказались завышенными, но наиболее разработанные системы, пусть с пост-редактированием, дают результаты.

После перебазирования в дирекцию я стал реже бывать в Отделе языков, но постоянные контакты продолжались, в том числе и после моего ухода из Института востоковедения в 2012 году. Временами перезванивались. Последний раз мы долго общались во время конференции Отдела в апреле 2018 года. Потом я узнал, что она болеет...

Безусловно, Зоя была не только видным учёным и организатором науки, но и просто хорошим человеком. А это много.



**Зоя Михайловна Шаляпина (другу и ученому)**

**Лилия Михайловна Горелова**

Университет Окленда, Окленд, Новая Зеландия,  
lgorelova@gmail.com, ORCID 0000-0001-6911-0118

**Аннотация.** Это эссе посвящается моему многолетнему другу и известному ученому Зое Михайловне Шаляпиной, с которой мы вместе работали в Отделе языков Азии и Африки Института востоковедения РАН более сорока лет. Такие личностные качества, как выдающаяся работоспособность, преданность своему делу, перфекционизм во всем и строгий логический ум, позволили ей сформулировать важные теоретические принципы в области науки, которой она занималась, а именно машинного перевода, а благожелательность, постоянная забота о сотрудниках и умение видеть перспективы способствовали созданию современного и эффективного научного подразделения.

**Ключевые слова:** З.М. Шаляпина, лингвист, педагог, теоретические проблемы англо-русского и японско-русского машинного перевода, универсальные методы организации формализованного лингвистического описания

**Abstract.** This essay is dedicated to my long-term dear friend and a prominent scientist Zoya Mikhailovna Shalyapina, with whom we have worked together in the Department of Languages of Asia and Africa of the Institute of Oriental Studies of the Russian Academy of Sciences for over forty years. Her personal qualities such as dedication to work, perfectionism in everything and a strict logical mind allowed her to formulate important theoretical principles in computational linguistics and, in particular, machine translation. Her benevolence, constant concern for colleagues and the ability to foresee the upcoming trends have largely contributed to the creation of a modern and effective linguistic department.

**Keywords:** Z.M. Shalyapina, linguist, teacher, theoretical problems of English-Russian and Japanese-Russian machine translation, universal methods of organizing formalized linguistic description

Если вы откроете любой сайт, на котором помещены данные о Зое Михайловне Шаляпиной, включая [ru.wikipedia.org](http://ru.wikipedia.org), то прочтаете, что она – известный российский лингвист, с конца 1995 года была заведующей Отделом языков народов Азии и Африки Института востоковедения РАН (ИВ РАН), кандидат филологических наук, имеет более 200 публикаций (даты жизни: 14 апреля 1946 – 19 октября 2020)<sup>1</sup>.

Для меня же более сорока лет она была просто *Зюечка*, с которой я познакомилась сразу же, как только была принята на работу в Институт востоковедения АН СССР в 1980 году, а она уже работала там с 1977 года. Мы *увидели* друг друга мгновенно (*as quick as thought*), и все эти сорок лет нас связывали те искренние отношения, которые, видимо, и называются дружбой. Где бы меня ни носила

---

<sup>1</sup> [https://ru.wikipedia.org/wiki/Шаляпина\\_Зоя\\_Михайловна](https://ru.wikipedia.org/wiki/Шаляпина_Зоя_Михайловна)

жизнь с конца предыдущего века, мы никогда не упускали друг друга из вида и всегда знали, что с каждой из нас происходит.

Когда я уже не смогла летать два раза в год в Москву, сейчас в том числе и из-за ковида, я ей звонила, где-то раз в две недели, иногда – реже, и мы подолгу разговаривали. Конечно, я знала, что она тяжело больна, и была в курсе многих подробностей. Последний раз мы разговаривали недели за три до ее *ухода*; я ее спрашивала о самочувствии, но было очевидно, что она не хочет об этом говорить – только сказала, что уже, в основном, лежит. Я же просто не могла впустить в голову мысль о том, что она уже *уходила* и у нее хватало сил только на то, чтобы слушать. Это был наш последний разговор с ней. Где-то через неделю я стала названивать снова, как будто чувствовала, что мне *необходимо* с ней поговорить, но телефон не отвечал, – такое уже бывало, ремонт в их доме был нескончаемый, и время от времени что-то случалось с телефонным кабелем. Однако беспокойство меня не покидало, и я попросила дочь Марии Николаевны Орловской, нашей тоже недавно *ушедшей* коллеги по Отделу, Ирину, связаться либо с Зоей, либо с ее сестрой Галиной, и узнать, в чем же дело. Ответ от Ирины пришел на следующий день: Зочка *ушла* во сне 19-го, а 21-го октября ее кремировали. Для меня это был шок, – я никак не думала, что это случится так скоро, хотя могла бы догадываться. Видимо, она меня берегла, не посвящая в подробности, и *уходила* мужественно, наедине с собой и сестрой, принимая судьбу как должное.

Последний раз мы виделись года полтора тому назад, во время нашей встречи в одном из московских кафе, – эта была середина пути между ее домом и моим, тем, где я тогда обитала. Я ожидала Зою у нужного входа, в этом центре их много, и когда увидела, *как* она медленно идет, то у меня сжалось сердце, однако я и тогда не поняла, что это уже *совсем* плохой признак. Мы обнялись, зашли в кафе, нашли там укромный уголок, заказали какую-то еду и долго-долго разговаривали, пока обе не устали. Потом я ее пошла провожать в метро до поезда в ее сторону. У нее был такой взгляд, будто она со мной просталась. Может быть, это я сейчас додумываю, но взгляд был такой, будто бы в нем отражались все мировые беды. Печальный. Несколько растерянный. Особый.

До того я пыталась уговорить ее предпринять определенные усилия по лечению болезни, но прекрасно понимала, что если Зоя уже пришла к определенному решению, то отговаривать ее было бесполезно. *Зоя была человеком убеждений*.

Я не могу сказать, что мы во всем соглашались друг с другом; на кое-какие вещи у нас с ней были разные взгляды, особенно различались взгляды на политику. Но в наших отношениях было что-то большее, стоящее над политикой, – вероятно, какие-то общие мировоззренческие ценности.

Я знала, что Зоя – фантастически логична в любом своем дискурсе, вопрос только в предпосылках. Ну, а предпосылки – по большей части это тоже дело убеждений. *Втиснуться* в ее непрерываемую логику было просто невозможно, а вот усомниться в правильности предпосылок – можно. Однако редко кто на это отваживался. Обычно люди быстро уставали от Зоинной железной логики и от силы ее убежденности, не могли выдвинуть никаких контраргументов и отказывались от дальнейших споров. Однако понимание ее правоты часто приходило только по прошествии какого-то времени, иногда – продолжительного. Вспомни-

наю наши совместные походы на овощные базы<sup>2</sup>. Когда возникали перерывы в работе и начинались какие-то разговоры, а часто и споры, народ поначалу пытался возражать Зое, но минут через десять все как-то потихоньку рассредотачивались по своим местам, – никто не выдерживал такой напряженной работы ума и убежденности.

*Зоя Михайловна была фантастически работоспособна и фантастически тщательна во всем том, что она делала.* Как она умела работать и насколько она была профессиональна в своей работе, я впервые поняла, когда мы договорились поработать над моим текстом, написанным на английском языке для одного иностранного журнала, – было это еще в середине 80-х. Собственно, мы обе отдали свои статьи для перевода на английский язык одной девушке, окончившей МГУ и позиционировавшей себя как переводчика, – она подрабатывала переводами. Однако когда мы ознакомились с тем, что получилось, то поняли, что статьи нуждаются в серьезной правке. Я думаю, что Зоя сама бы перевела свою статью на английский намного лучше, она ведь закончила переводческий факультет Московского государственного педагогического института иностранных языков (МГПИИЯ) им. М. Горького, но, видимо, ей было жаль времени<sup>3</sup>. Мы просидели над переводом целый день, и я вернулась домой совершенно измученная. А для Зои эта интенсивность в работе была обычной, и тщательность анализа просто поражала. В ход шли словари, разнообразные справочники, и все непонятные моменты обсуждались до тех пор, пока текст не читался так, как был задуман.

Второй урок я получила – как нужно работать – немного погодя, когда в Институте проходила крупная международная конференция, и руководил ею тогдашний директор Института Е.М. Примаков (тоже где-то в это время)<sup>4</sup>. Наш Отдел языков должен был подготовить тезисы докладов по лингвистической тематике, и ответственной за эту работу была назначена З.М. Шаляпина (обращаясь к ее профессиональной деятельности, я думаю, уместно перейти к ее полному имени-отчеству-фамилии), которая взяла меня в помощники. Я честно отнеслась к работе; в общем, была тоже перфекционисткой в меру отпущенных мне возможностей,

---

<sup>2</sup> В условиях планового социалистического метода хозяйствования плодоовощные базы входили в систему государственной системы снабжения населения продовольствием. В периоды острой нехватки рабочей силы на них направлялись городские служащие, работники учебных и научных организаций. Это было принудительной работой, и мы часто оказывались в компании докторов наук в возрасте за 50 лет, которые должны были перебирать гнилые овощи на конвейере.

<sup>3</sup> В настоящее время МГПИИЯ преобразован в Московский государственный лингвистический университет (МГЛУ). Непосредственной специальностью З.М. были прикладная лингвистика, английский и итальянский языки, а ее дипломная работа была посвящена орфографии английского языка. В дополнение З.М. знала еще французский и немецкий языки. Японский она выучила уже после окончания института, на курсах.

<sup>4</sup> Е.М. Примаков (1929-2015) – ученый-арабист, директор Института востоковедения АН СССР (1977-1985), государственный деятель, занимавший в разное время высокие политические и правительственные посты.

однако после моей правки осталось еще много незамеченных оплошностей. Когда я спросила Зою, много ли ей пришлось править после меня, она с некоторым замешательством сказала, что – прилично. И мы снова прошлись по всем текстам. Вот после этого они уже стали совсем *perfect*.

Я знала, что З.М. Шаляпина пришла в Институт востоковедения после того, как проработала с 1968-го по 1977-й годы в Лаборатории машинного перевода ее родного Института, МГПИИЯ им. М. Горького, – в начале 70-х она возглавляла работу над системой англо-русского машинного перевода (АРАП), идея которой была в свое время предложена И.А. Мельчуком и О.С. Кулагиной. А в 1975 году она защитила кандидатскую диссертацию по теме «Принципы организации и записи семантической информации в автоматическом словаре». В общем, работа над этой системой была реализацией ее прикладных интересов в лингвистике, а именно – машинного перевода, идея которого была очень популярна в те годы. Я знаю и причины, по которым она ушла из лаборатории. Они были связаны с общей ситуацией в стране, когда всем стало не только не до машинного перевода, но и не до науки в целом.

В Институте востоковедения З.М. Шаляпина начала заниматься японско-русским машинным переводом, практически на пустом месте. Она пришла в группу Отдела, которую возглавляла С.М. Шевенко. Кроме того, в это подразделение входило несколько девочек-лаборанток. Не было никакого технического обеспечения, и вообще это была работа не для такой малочисленной группы, а для большого института с необходимым компьютерным арсеналом, который бы обслуживался штатом программистов, как, например, это уже происходило в Японии. З.М. Шаляпина *со своим логическим умом, работоспособностью и преданностью делу погрузилась в теоретические проблемы*. Именно тогда она занималась развитием тех принципов «словарно-ориентированной» организации лингвистического описания, которые были выработаны ею еще в Лаборатории МГПИИЯ. Именно эти принципы составляют теоретическую основу так называемого сущностного подхода к языку, который она развивала до своей кончины. Для того, чтобы представлять себе, как можно моделировать язык с помощью компьютеров, нужно было выработать свое собственное теоретическое и философское понимание устройства естественных языков, создать универсальные методы организации формализованного лингвистического описания и конкретные формализмы для записи морфологической, синтаксической и семантической информации о языковых единицах. Над всеми этими проблемами З.М. Шаляпина очень глубоко думала и воплощала их решение в своих статьях и книгах. Должна сказать, что ее работы настолько плотно насыщены сложным формальным аппаратом, что их трудно читать неспециалистам в этой области. Не думаю, что я имею даже и потенциальную возможность подробно говорить о научных результатах, полученных З. М. Шаляпиной во время работы в Институте востоковедения, – есть много специалистов и ее учеников, которые могут рассуждать на эту тему более профессионально; все-таки у меня совсем другая специальность<sup>5</sup>.

---

<sup>5</sup> Научная библиография З.М. Шаляпиной приведена на нескольких сайтах:  
[https://ru.wikipedia.org/wiki/Шаляпина\\_Зоя\\_Михайловна](https://ru.wikipedia.org/wiki/Шаляпина_Зоя_Михайловна)  
<https://www.ivran.ru/persons/ShalyapinaZoya>

А в конце 90-х прошлого столетия ситуация с научным финансированием несколько изменилась к лучшему, и появились какие-то надежды. У Зои Михайловны – тоже. К тому же в 1995 году она была назначена заведующей нашим тогда еще довольно большим Отделом языков. Я думаю, что она согласилась на эту работу, только имея в виду основную цель: сохранение своей группы, с помощью которой она бы могла продвигать свою главную научную задачу, и это было создание автоматизированной системы японско-русского машинного перевода. Появилась возможность получать гранты РГНФ и РФФИ, а также государственные гранты. В 1994-1996 годах она получила Государственную (президентскую) стипендию. Все эти средства тратились, во-первых, на приобретение компьютерной техники, которой могли пользоваться все сотрудники Отдела, и он в результате стал одним из самых компьютеризованных подразделений Института, а, во-вторых, появилась возможность приглашать на работу программистов и специалистов, которые могли бы работать по тематике «автоматизированный перевод». Тогда же стали появляться и аспиранты, которые были заинтересованы в этой специальности, – Зою Михайловну можно было часто видеть окруженной молодежью, которой она что-то рассказывала или объясняла, а студенты – внимали, иначе и не скажешь. Надо отметить, что кроме исследовательской деятельности в Институте, в 1988-2002 годах З.М. Шаляпина вела большую образовательную работу в разных вузах Москвы, прежде всего в Московском государственном университете (МГУ) и в Российском государственном гуманитарном университете (РГГУ).

И вся эта разнообразная деятельность была работой на износ, – не уверена, что кто-то еще об этом скажет или напишет. Но я-то знала об этом очень хорошо. На мой взгляд, особенно утомительной была работа в качестве заведующей Отделом. Контингент сотрудников был очень возрастной, к новым условиям многим из них было трудно приспособиться, – кое-кто так и не научился работать на компьютерах. В конце концов Зоя Михайловна приняла на работу сотрудницу, которая могла это делать за тех, кто не мог приводить свои монографии в должный вид. Это-то было еще полбеды. Поскольку возможности работать в Институте не было, Отдел лишился своих основных помещений, то в основном сотрудники работали дома. Однако в присутственные дни кое-кто считал возможным и необходимым приходить в Отдел, в ту комнату, где обычно сидела Зоя Михайловна за своим компьютером (она у нас называлась компьютерной), садиться около нее и рассказывать и о том, как продвигается их работа, и о том, как вообще им живется. Тут же заполнялись планкарты, которые должны были содержать отчеты за прошедший год или планы на будущий, и часто это делала сама заведующая, – если сотрудник не умел делать это самостоятельно. Хотя Зоя Михайловна была очень терпеливым и благожелательным человеком и вполне четко осознавала свои обязанности по отношению к Отделу, иногда и ее терпения не хватало. Я видела это только по выражению ее лица, когда заходила в компьютерную. И всегда удивлялась тому, как она может проявлять столько терпения. В результате очень часто она получала возможность заниматься своей непосредственной исследовательской работой, либо с программистами, либо с коллегами, только после того, как все сотрудники расходились по домам, и она добиралась домой уже совсем поздно вечером. Следующий день был потерян, даже и отдохнуть не удавалось, особенно если ей приходилось – вопреки своему физическому состоянию – делать какую-то срочную работу, а приходилось довольно часто. Мало того, некоторые со-

трудники даже в эти дни доставали ее по домашнему телефону, часто – со всякой ерундой. Я всегда удивлялась, почему же люди не могут подождать до присутственного дня, если уж так необходимо рассказать или спросить о чем-то. Потом догадалась: надо было у начальника создать впечатление, что ты просто горишь на работе – ни дня без строчки. В принципе, я могла бы позволить себе позвонить ей в такие дни, чтобы просто поболтать, но практически никогда этого не делала, понимая, что ей нужно побыть в тишине вдали от людей. Работала Зоя Михайловна и в праздничные дни, – у нее был особый пропуск в Институт, и весь обслуживающий персонал ее очень хорошо знал и уважал.

Далее, по правилам охраны труда (если бы они существовали), в компьютерной вообще нельзя было постоянно находиться: комнатка была маленькая, напичканная техникой, которая постоянно была включена, и в ней все время находились сотрудники, – некоторые там работали, а некоторые заходили пообщаться. Летом было душно и жарко, зимой – как когда. Иногда холодно, но дышать было нечем всегда. Кто-то приходил, кто-то уходил, велись громкие разговоры, – сосредоточиться во всем этом могла только Зоя Михайловна. Но эта необходимость быть сосредоточенной постоянно, даже когда другие могут расслабиться, назавтра оборачивалась глубокой усталостью и тяжелой слабостью. С возрастом – сильнее.

Нельзя сказать, что сотрудники, особенно молодежь, которая появилась в Отделе, не берегли своего начальника. Старались. Звали ее к чаю, который накрывали в соседней комнате, когда всем становилось уже очевидно, что нужно передохнуть. Но все равно многие часы, которые она проводила в этой несовместимой со здоровым образом жизни атмосфере, сидячий – в основном – образ жизни ни к чему хорошему привести не могли. Все эти обстоятельства плюс постоянное напряжение не могли не укоротить дни ее жизни.

Как-то ночью мне не спалось, я начала думать о том времени, которое я провела в Отделе языков в Институте востоковедения, и в голове сложилось общее количество сотрудников, которое ушло за это время. Много, очень много. Особенно много стало уходить в последнее время, поскольку и возраст такой подходил, и общая ситуация для пожилых людей в стране лучше не становилась. И обо всех Зоя Михайловна беспокоилась и всех провожала в последний путь. Устраивала юбилейные конференции, инициировала публикацию сборников статей в память ушедших сотрудников, – я знаю об этом, потому что время от времени она просила меня написать либо текст к юбилею сотрудника, либо рецензию на очередную книгу, которая издавалась в память об ушедшем коллеге.

Зоя Михайловна Шаляпина руководила Отделом в нелегкие для науки времена, превратила его в эффективное и современное подразделение, поддерживала в Отделе благожелательную рабочую атмосферу. Известно, что традиции легко разрушить, но трудно сохранить. Хочется, чтобы Отдел хранил традиции, для создания которых Зоя Михайловна Шаляпина приложила столько физических и духовных сил.

**Разработка З. М. Шаляпиной  
концепции языковых ярусов в рамках  
трехмерной стратификационной модели языка**

**Кирилл Игоревич Демич**

Институт востоковедения РАН, Москва, Россия

**Александр Вячеславович Костыркин**

Институт востоковедения РАН, Москва, Россия, kost@ivran.ru

***Аннотация.*** На фоне повышенного интереса к стратификационным описаниям языка в 1960 гг. наиболее продуктивными оказались схемы С. Лэма и И.Ф. Вардуля. Учитывающая тезис Ю. С. Маслова об ограничениях одномерной схемы «Трехмерная стратификационная модель языка» З. М. Шаляпиной явилась развитием многоярусной схемы Вардуля. С привлечением теоретико-множественных подходов ярусы модели строятся на основе выявления «концентроров» вокруг безусловных единиц с помощью спиралевидной индукции.

***Ключевые слова:*** стратификационные модели языка, полевые модели языка, нечеткие множества, философия языка, теория языка

***Abstract.*** 1960s showed increased interest in stratificational models of language. The schemes by S. Lamb and I. F. Vardul appeared to be the most productive. Limitations of single dimensional model were pointed by Yu. S. Maslov. The point was adopted by Z. M. Shalyapina in her development of Vardul's stratificational scheme. Given set-theory approaches, each level of her model is constructed from undisputable units' core towards periphery employing spiral-wise induction.

***Keywords:*** stratificational models of language, core-periphery linguistic models, fuzzy sets, philosophy of language, theory of language

Одним из первых подходов к идее языковых ярусов можно считать выделение Соссюром единиц «низшего порядка» – *sousunités* (т.е. морфем: *suffixes, préfixes, radicaux*) и связанных с ними «единиц шире слова» – *unités plus larges que les mots* [Saussure 1995: 148; Соссюр 1933: 107–108]. На фоне перехода от структуралистской парадигмы, в том числе и блумфилдианского типа, к генеративистской с ее повышенным интересом к синтаксическим структурам все более насущно стала ощущаться необходимость выработки общей уровневой модели языка, которая бы на единых универсальных основаниях согласовывала единицы разных уровней, прежде всего фонетического, морфологического и синтаксического.

Пристальный интерес к себе эта тема привлекла в 1960-е гг. Появившаяся в 1962 г., первоначально как рабочий отчет и пособие для студентов, и почти незамеченная в ученых кругах «Стратификационная грамматика» С. Лэма стала отрицанием лингвистической концепции Л. Блумфилда [Lamb 1966: iii; Лэм 1977: 5].

Другой важной вехой стал IX Международный конгресс лингвистов 1962 г., где широкой научной общественности Р. Якобсон представил Н. Хомского с его «Синтаксическими структурами» как восходящую звезду [Березин 2000: 19], а Э. Бенвенист сделал программный доклад, посвященный уровням лингвистического анализа [Бенвенист 1965]. В последовавших за этим работах отечественных лингвистов вопросам уровневой организации также было уделено специальное внимание: «Понятие давления системы и иерархия языковых единиц» (1962) Э. А. Макаева, «О понятии «уровня» в современном языкознании» (1962) и «О теории лингвистических уровней» (1964) С.Д. Кацнельсона, «Об основных и промежуточных ярусах в структуре языка» (1968) Ю.С. Маслова, «Проблема уровней как психолингвистическая проблема» (1969) А.А. Леонтьева, «О понятии уровня языковой системы» (1972) В.М. Солнцева.



Рис.1. Семь ярусов японского языка [Вардуль 1973:16]

В этом ряду особняком стоит многоярусная схема языка, разрабатывавшаяся И. Ф. Вардулем [Вардуль 1966, 1973]. По признанию З. М. Шаляпиной [Шаляпина 2005: 32; 1986: 77], она сильно повлияла на ее обобщающую концепцию, к рассмотрению которой мы в дальнейшем перейдем. Кроме того, учитывая, что и Вардуль, и Шаляпина – оба по профильной лингвистической специальности японисты, особый интерес представляет еще одна линия исследований в области стратификационных исследований языка, которую можно отрывочно проследить, – вклад собственно японистов. Прежде всего здесь стоит назвать Бернарда Блоха, сыгравшего немаловажную роль в продвижении и популяризации идей Хомского, а также Георга фон дер Габеленца с несколькими его положениями общелингвистического значения. Здесь особо важна созвучная Соссюру идея о трехчастном делении предмета лингвистических исследований на язык, речь и речевую деятельность (*Einzel-sprache, Rede, Sprachvermögen*) [Березин 2000: 11; Gabelentz 1901], притом, что *Sprachvermögen* Габеленца оказывается ближе к языковым компетенциям Хомского, чем речевая деятельность Соссюра (*language*). Так обобщенное Хомским принципиальное противопоставление *языковой компетенции* и *языковой деятельности* «стало общепризнанным в области формальных лингвистических моделей» [Шаляпина 2007: 10]. Кроме того, внимание Вардуля, на которого подход Габеленца оказал определенное влияние, привлекло различие трех видов субъектов и предикатов (грамматических, психологических и логических). По мнению Вардуля, структуры трех соответствующих им видов находят отражение как в различие-



нии «формального (грамматического) и актуального (смыслового) членений предложения», так и в хомскианском противопоставлении поверхностной (грамматической) и глубинной (семантической) структур предложения [Вардуль 2005: 38].

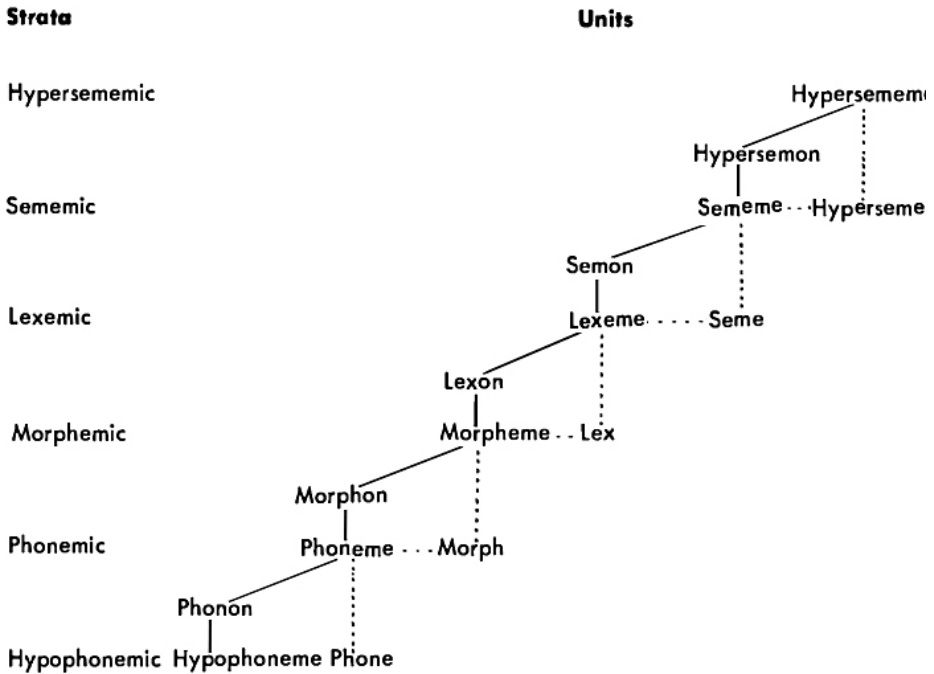


Рис.2. Некоторые единицы естественных языков [Lamb 1964: 20],  
ср. [Лэм 1977: 31]

Сначала остановимся на схеме Вардуля (см. рис. 1). Само ее построение началось с одной из его первых фундаментальных работ, а именно, с расщепления синтаксического уровня на актуальный и потенциальный, что в дальнейшем и потребовало выявления и упорядочения всех остальных уровней в рамках единой модели [Вардуль 1964]. Уже в следующих работах у него появляется понятие «яруса» как структурного уровня языковой системы [Вардуль 1965]. Сам иерархический принцип ярусной организации предполагает также и некоторую вертикальную ось. Хотя наличие такой оси автором прямо не было заявлено, но ввиду отсутствия предполагаемых измерений по другим осям, в этом смысле можно говорить о его схеме как о неявно линейной и вертикально упорядоченной (см. рис. 1). Прямая критика Вардулем схемы Лэма дает удачную возможность сопоставить оба подхода [Вардуль 1977: 18]. При несогласии Вардуля с внутренним устройством модели Лэма, обращает на себя внимание их внешнее сходство (см. рис. 2). Как и у Лэма, между выявляемыми объектами, требующими ярусного упорядочения на схеме, обнаруживаются связи со множеством ветвлений (см. рис. 3) и иерархически противоречивых отношений между конкурирующими ярусами

и объектами, схематическое представление которых фактически требовало бы методов, заимствованных из теории множеств, не всегда адекватных принципам стратификационного описания (см. рис. 4). Частичной попыткой разрешения этого противоречия у Вардуля и было привлечение таких методов [Солнцев 1972].

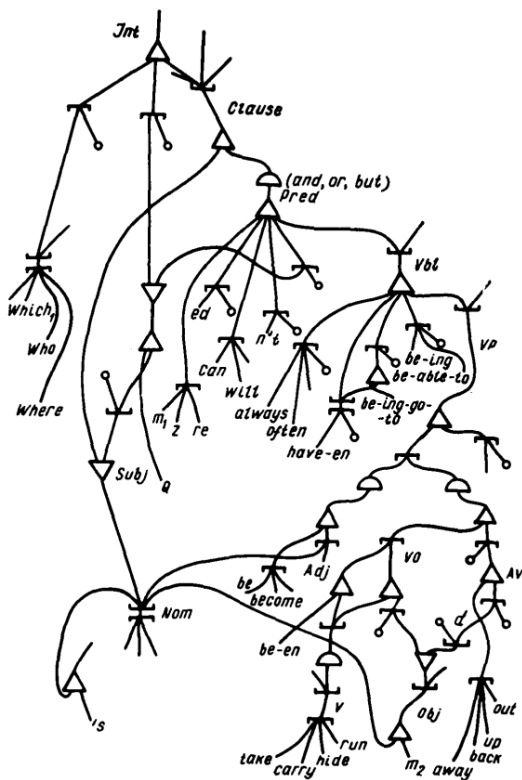


Рис. 3. Фрагмент английской лексотактики [Лэм 1977: 36]

В 1968 г. Ю. С. Маслов констатировал, что свойственная дескриптивизму сепарация ярусов при стратификации языковой структуры начинает преодолеваться, одним из удачных примеров чего он считает схему Лэма. По мнению Маслова, «установление «вертикальной» иерархии языковых единиц есть лишь одно из направлений в исследовании многомерной структуры языка», поскольку при установлении такой иерархии «нельзя игнорировать и переходные, смешанные и периферийные явления» в рассуждениях не только об уровнях или ярусах языковой структуры, но и о соответствующих уровнях лингвистического исследования. Маслов приходит к выводу, что некоторые явления, например, грамматические категории, очевидно, мыслятся как лежащие в другой плоскости, «перпендикулярно к основной иерархической оси».

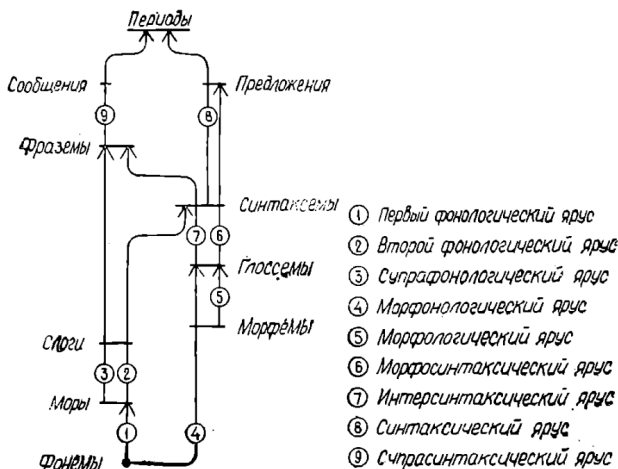


Рис. 4. Девять ярусов японского языка [Вардуть 1986:42]

Учитывая этот тезис Маслова, Шаляпина строит свои рассуждения на том, что наряду с иерархическими моделями языка, начиная с Лэма, построенными вокруг процедур кодирования и декодирования или анализа и синтеза речи, для которых мерезологические отношения между разноуровневыми единицами или отношения их вложенности имеют большое значение, существуют и принципиально иные, которые хотя и оперируют единицами разных языковых ярусов и тем самым апеллируют к ним, однако рассматривают преимущественно другие разноплановые процессы, как, например, «стандартная теория» Хомского. В «стандартной теории» построение строгой иерархической модели не является приоритетным, поскольку в центр внимания ставится генерация поверхностных структур из глубинных, с которыми вспомогательные механизмы их фонетической репрезентации и семантической интерпретации сопрягаются как периферийные [Шаляпина 2007: 12]. В первом случае, по наблюдению Шаляпиной, для иерархических моделей предполагается строго последовательный переход по цепочке уровней, когда изменение очередности перехода от уровня к уровню или «перескоков» между ними не предусматривается. Таким образом, делает она вывод, при попытке согласовать иерархический подход Лэма с генеративистским, как, например, это сделал Мельчук для теории «Смысл ↔ Текст», «требуется либо изменить состав или порядок расположения уровней внутри самой исходной модели, либо признать, что переход от уровня к уровню происходит не в плоскости этой модели, но за ее пределами» [Шаляпина 2007: 12]. В итоге в соответствии с идеей Маслова о том, что смешанные или промежуточные ярусы могут быть проекцией на основной оси явлений и категорий, являющихся содержанием другой плоскости, где они представлены более полноценно, Шаляпина предлагает различать три основные оси, «формируемые тремя базовыми функциями языкового знака – функциями обозначения, членения и обобщения» [Шаляпина 2007: 3].

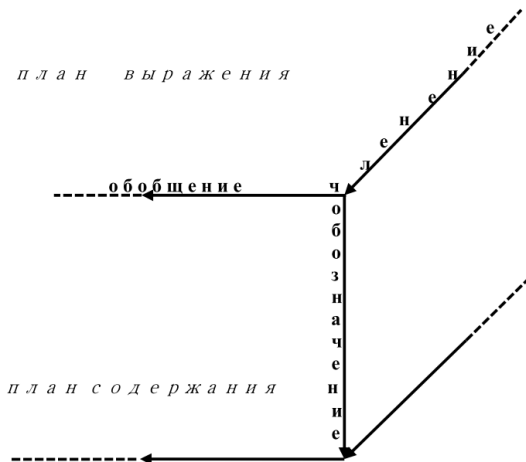


Рис. 5. Предварительная схема простейшего трехмерного варианта многомерной модели лингвистической компетенции [Шалапина 2007: 22]

Возвращаясь к схеме Вардуля, можно сказать, что отталкиваясь от нее и устраняя ее основное ограничение, связанное с тем, что как бы конкурирующие на скрытой основной оси ярусы и объекты не всегда отражают подлинные языковые явления и противоречия (см. рис. 4), но являются побочным эффектом ее одномерности, многомерная модель Шалапиной разгружает основную ось членения, характеризующую мереологическим отношением «часть–целое», от несвойственных ей противоречий, не связанных с вложенностью единиц разных ярусов, но в многомерной модели координируемых уже по другим осям (см. рис. 5), что позволяет вскрыть нереализованный потенциал исходной схемы – возможность на месте мнимой межъярусной конкуренции выявлять подлинную неопределенность, «размытость», межъярусных границ. Такая экспликация проблемы позволяет прямо привлекать для ее решения инструментарий теории нечетких множеств [Шалапина 2007: 92].

Шалапина строит свою многоуровневую модель в несколько шагов. На первом этапе при ее построении она опирается на постулат о гомогенности каждого яруса, согласно которому существует узкий класс «безусловных единиц», принадлежность которых к рассматриваемому ярусу не вызывает сомнений [Шалапина 2007: 59-60]. Таким образом принцип гомогенности ярусов позволяет строго и последовательно устанавливать границу между смежными ярусами и выявлять в пределах каждого яруса единицы, которые целесообразно определять как граничные. Благодаря этому Шалапина получает возможность ввести общий принцип минимальных и максимальных граничных единиц своих ярусов [Шалапина 2007: 46-47], который подразумевает корреляцию единиц смежных ярусов, соотносимых с одним и тем же текстовым сегментом.

Так, например, при установлении границы «между такими фундаментальными областями оси членения, как морфология и синтаксис» в качестве пограничных единиц выступают «словоформы как максимальные единицы морфологии

и одновременно минимальные единицы синтаксиса» [Шаляпина 2007: 73]. При этом свойства коррелятов могут «переходить» межъярусную границу, т.е. наследоваться [Шаляпина 2007: 49]. В соответствии с этим промежуточные ярусы, если они возникают при таком межъярусном наследовании, не разрушают основных мерологических отношений между объемлющими и объемлемыми единицами и не противоречат принципам координации по этой оси, а также при необходимости допускают возможность рекурсии.

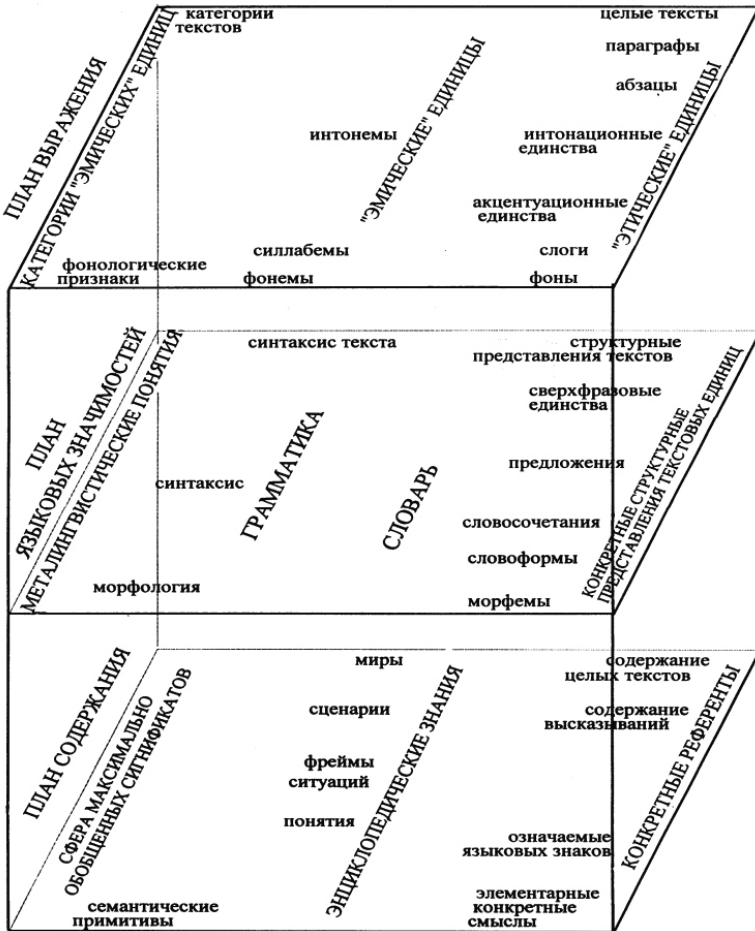


Рис. 6. Трехмерное пространство лингвистической компетенции [Шаляпина 2007: 252]

Важной особенностью лингво-философской концепции Шаляпиной, представленной такой многомерной моделью, является формализация противопоставления в языке центра и периферии, подход, восходящий прежде всего к исследованиям представителей Пражской лингвистической школы [Шаляпина 2007: 60].

Для его реализации она формулирует логическую схему, по которой для любого заданного яруса лингвистического описания могут строиться рекурсивные критерии, позволяющие устанавливать принадлежность к этому ярусу отдельных языковых единиц и тем самым определять *область распространения* яруса, что напоминает математический метод доказательства по индукции. *Область распространения* определяется таким образом, так сказать, по спирали: от первого, базового «концентра» безусловных единиц данного яруса к единицам второго, производного «концентра», выявляемым по их сочетаемости с безусловными единицами и т.д. [Шаляпина 2007: 61–62].

Другой принципиальной особенностью многомерной модели Шаляпиной является требование выделения ярусов не только в системе языка, но и в пределах конкретных текстов и их фрагментов [Шаляпина 2007: 49, 91]. Этот посыл содержательно и без побочных деформаций, развивая центральную для отечественного языкознания традицию Виноградова–Пешковского, наследует критерии «отдельности слова», предложенные А. И. Смирницким (1952) [Шаляпина 2007: 61]. Тогда как, например, Вардуль, апеллируя к глоссам Смирницкого, выводит свою «глоссемию» на основе «абстракции отождествления», предполагающей отвлечение «от пространственно-временных координат дискурсных сегментов» [Вардуль 1977: 49], отсекающее непосредственно данную конкретику эмпирического текстового слова [Демич, Костыркин 2018].

Кроме всего прочего, значение этих положений Шаляпиной и в том, что они имеют своим следствием универсальный вывод о существенной исчислимости *тривиальных коллокаций* [Ривелис 2007: 291–92], недооцененного иллюстративного материала, нередко рассматриваемого теоретиками как малоинформативный лексикографический «балласт» [Берков 1973: 120–21, 2004:64], но традиционно несущего большую нагрузку в японской национальной лексикографии и обладающего системной семантизирующей способностью при составлении двуязычных словарей, о чем свидетельствует словарная работа в рамках проекта ЭЛЕЯРС, долгое время возглавлявшегося Шаляпиной. Именно тривиальные коллокации, ярко отражая явления *межплановой лингвистической избыточности* [Шаляпина 2007: 264–65], наиболее полно задействуют «безусловный» ядерный языковой материал. Кроме того, этот вывод, как теперь ясно, особо значимый для теоретической и практической лексикографии, хорошо согласуется с тезисом о центральной роли словаря в лингвистическом описании, принятом и обобщенном Шаляпиной до принципиальной возможности описания всего языка словарем на основе общего принципа, согласно которому «словарь есть обобщение текста, грамматика есть обобщение словаря» [Шаляпина 1975: 135; 1991: 43]; [Костыркин, Шаляпина 2005: 213]; [Шаляпина 2007: 6, 144].

## Литература

1. Бенвенист Э. Уровни лингвистического анализа. // *Новое в лингвистике*. Вып. IV. М.: Прогресс; 1965. С. 434–449.

2. Берков В.П. *Вопросы двуязычной лексикографии (словник)*. Л.: Изд-во ЛГУ; 1973.
3. Берков В.П. *Двуязычная лексикография*. 2-е изд. М.: Астрель, АСТ, Транзиткнига; 2004.
4. Вардуль И.Ф. *Очерки потенциального синтаксиса японского языка*, М.: Наука; 1964.
5. Вардуль И.Ф. Слово как часть речи // *Вопросы теории частей речи на материале языков различных типов. Тезисы докладов на открытом расширенном заседании Ученого совета Института языкознания АН СССР (Ноябрь 1965 г.)*. Л., 1965. С. 22–23.
6. Вардуль И.Ф. Ярусы языковой системы и универсалии. // *Конференция по проблемам изучения универсальных и ареальных свойств языков. Тезисы докладов*. М., 1966. стр. 18–21
7. Вардуль И.Ф. О ярусах японского языка // *Тезисы докладов IV научной конференции по японской филологии*. Март 1973. С. 15–20.
8. Вардуль И.Ф. *Основы описательной лингвистики*. М.: Наука. 1977.
9. Вардуль И.Ф. По поводу одной синтаксической идеи Г. Габеленца. // *Актуальные вопросы японского и общего языкознания* (Статья датирована сентябрем 1978 г.). М.: Восточная литература. 2005. С. 38–44.
10. Виноградов В.В. О формах слова. // *Исследования по русской грамматике*. М.: Наука. 1975. С. 33–50. Перв. публ. в: «Известия АН СССР, ОЛЯ». Т. III. Вып. I. 1944. С.31–44.
11. Демич К.И., Костыркин А.В. Японское слово в электронном словаре: Категории леммы, лексемы, словарного и заголовочного слова в структуре словарного описания // *Труды института востоковедения РАН. Вып.19: Проблемы общей и востоковедной лингвистики: Лексикология и лексикография*. Отв. ред. З.М. Шалапина. М.: ИВ РАН. 2018. С. 166–178.
12. Березин Ф.М. *О парадигмах в истории языкознания XX в. В сб.: Лингвистические исследования в конце XX в. Сборник обзоров*. М.: ИНИОН РАН. 2000. С. 9–25.
13. Леонтьев А.А. Проблема уровней как психолингвистическая проблема. // *Единицы разных уровней грамматического строя языка и их взаимодействие*. М.: Наука. 1969. С. 263–266.
14. Лэм, Сидней М. *Очерки стратификационной грамматики*. Минск: Вышэйшая школа. 1977.
15. Макаев Э.А. Понятие давления системы и иерархия языковых единиц. // *Вопросы языкознания*. 1962. № 5. С. 47–52.
16. Маслов Ю.С. Об основных и промежуточных ярусах в структуре языка // *Вопросы языкознания*. 1968. №4. С. 69–79.
17. Кацнельсон С.Д. О понятии «уровня» в современном языкознании // *Тезисы докладов на дискуссии о проблеме системности в языке (30 янв. – 2 февр. 1962)*. М., 1962. С. 5.
18. Кацнельсон С.Д. О теории лингвистических уровней. // *Вопросы общего языкознания*. М. 1964. С. 32–41.

19. Костыркин А. В., Шаляпина З.М.. Принципы описания японского синтаксиса в системе японско-русского автоматического перевода ЯРАП. // *Актуальные вопросы японского и общего языкознания*. М.: Восточная литература. 2005. С. 210–223.
20. Крылов С. А. Некоторые уточнения к определениям понятий словоформы и лексемы // *Семиотика и информатика*. Вып. 109. М., 1982. С. 118–136.
21. Ривелис Е. *Как возможен двуязычный словарь*. Stockholm. 2007.
22. Смирницкий А.И. К вопросу о слове (Проблема отдельности слова) // *Вопросы теории и истории языка в свете трудов И. В. Сталина по языкознанию*. М.: Изд. АН СССР. 1952. С. 182–203.
23. Солнцев В.М. О понятии уровня языковой системы // *Вопросы языкознания*. 1972. №3. С. 3–19.
24. Соссюр, Ф. де. *Курс общей лингвистики*. М.: СОЦЭКГИЗ. 1933. Пер. А. М. Сухотина.
25. Шаляпина З.М. Система АРАП: словарь, грамматика и их использование при автоматическом анализе // *Международный семинар по машинному переводу*. М., 1975. С. 135–138.
26. Шаляпина З.М. О морфолого-синтаксическом статусе некоторых японских грамматических единиц (в связи с общими проблемами разграничения морфологии и синтаксиса в естественном языке). // *Актуальные проблемы японского языкознания*. М.: Наука, 1986. С. 69–103.
27. Шаляпина З.М. Грамматика и ее соотношение со словарем при словоцентрическом подходе к языку (на опыте формализованного лингвистического описания) // *Вопросы языкознания*. 1991. № 5. С. 42–54.
28. Шаляпина З.М. И.Ф. Вардоль и компьютерная лингвистика. // *Актуальные вопросы японского и общего языкознания*. М.: Восточная литература. 2005. С. 31–34.
29. Amelung, Iwo. Naming Physics // Michael Lackner and Natascha Vittinghoff, eds., *Mapping Meanings: The Field of Learning in Late Qing China*. Leiden: Koninklijke Brill NV, 2004. P. 381–422.
30. Gabelentz, Georg von der. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig. Chr. Herm. Tauchnitz. 1901.
31. Eisler, Rudolf. *Wörterbuch der Philosophischen Begriffe und Ausdrücke*. Berlin, 1899.
32. Lamb, Sydney M. *Outline of stratificational grammar*. Washington. Georgetown University Press. 1966.
33. Saussure, Ferdinand de. *Cours de linguistique generale*. Payot. 1995.



## Библиография трудов Зои Михайловны Шаляпиной

**Светлана Дмитриевна Малышева**

Институт востоковедения РАН, Москва, Россия, malisheva.sveta2011@yandex.ru

**Ключевые слова:** З.М. Шаляпина, библиография, российское японоведение, японско-русский автоматический перевод, грамматические модели

## Bibliography of Zoya M. Shalyapina

**Keywords:** Z. Shalyapina, bibliography, Russian linguistics, Japanese studies in Russia, Japanese-Russian machine translation, grammatical models

### Монографии, брошюры

1. Исследования по японско-русской электронной лексикографии. Выпуск первый. – М.: Институт востоковедения РАН, 2019. – 186 с., ISBN 978-5-89282-860-4. (соавторы: Демич К.И., Костыркин А.В.)
2. Трехмерная стратификационная модель языка и его функционирования: К общей теории лингвистических моделей / Институт востоковедения РАН. Отв. ред. В.М. Алпатов. М.: Издательская фирма Восточная литература РАН, 2007. – 475 с.
3. Экспериментальный комплекс ЯРАП для лингвистических исследований в области японско-русского автоматического перевода: первая очередь / Институт востоковедения РАН. Москва, 2001. – 404 с. Библиогр. 77 наз. Рукопись деп. в ИНИОН РАН 5.11.01 № 56804. (соавторы: Модина Л.С., Канович М.И., Любченко В.И., Панина А.С., Сенина Н.И., Сивцева В.И., Тарасова Е.С., Хайлова И.М., Штернова О.А.)
4. Макет лингвистического обеспечения системы японско-русского автоматического перевода ЯРАП: общая структура и основные компоненты / АН СССР. Ин-т востоковедения. М., 1980. 235 с. Библиогр.: С. 228–232. Рукопись деп. в ИНИОН АН СССР № 5344 от 12.05.80 г.
5. Вопросы организации и записи семантической информации в автоматическом словаре. Диссертация на соискание учен. степени канд. филол. наук. – М.: МГПИИЯ им. М. Тореца, 1975. – 340 с.
6. Вопросы организации и записи семантической информации в автоматическом словаре. Автореферат диссертации на соискание учен. степени канд. филол. наук. – М.: МГПИИЯ им. М. Тореца, 1975. – 38 с.
7. Семантико-синтаксический анализ в системе англо-русского автоматического перевода (АРАП) / Публ. проблемной группы по экспериментальной и прикладной лингвистике. Вып. 47. – М.: ИРЯ АН СССР, 1974. – 58 с.
8. Экспериментальный фрагмент англо-русского толково-комбинаторного словаря в процессе перевода / Публ. проблемной группы по экспериментальной

- и прикладной лингвистике. Вып. 16. М.: ИРЯ АН СССР, 1971. – 37 с. (соавторы: Жолковский А.К., Мельчук И.А.)
9. Образцы словарных статей англо-русского автоматического семантического словаря / Публ. проблемной группы по экспериментальной и прикладной лингвистике. Вып. 9. М.: ИРЯ АН СССР, 1970. – 58 с. (соавторы: Мельчук И.А., Перцов Н.В., Поливанова А.К., Равич Р.Д., Убин И.И., Чехов А.С.)
10. Модель автоматического синтеза английской орфографии / Публ. проблемной группы по экспериментальной и прикладной лингвистике. Вып. 6. М.: ИРЯ АН СССР, 1970. – 46 с.

### Статьи в журналах и сборниках

1. Лексикография и лексикология в Отделе языков народов Азии и Африки ИВ РАН // Труды Института востоковедения РАН Вып.19. Проблемы общей и востоковедной лингвистики: Лексикология и лексикография Институт востоковедения РАН. 2018. М.: Институт востоковедения РАН, 2018. С. 10–20.
2. Основные работы в области лексикологии, выполненные в Отделе языков народов Азии и Африки ИВ РАН // Труды Института востоковедения РАН Вып. 19. Проблемы общей и востоковедной лингвистики: Лексикология и лексикография. 2018. Институт востоковедения РАН. 2018. М.: Институт востоковедения РАН, 2018. С. 21–31.
3. Лексикография и лексикология в отделе языков народов Азии и Африки ИВ РАН // Востоковедные чтения 2018. Лексикология и лексикография. К 200-летию ИВ РАН. Тезисы докладов межинститутской научной конференции. Институт востоковедения РАН. 2018. М.: Институт востоковедения РАН, 2018. С. 61–62.
4. Предисловие // Труды Института востоковедения РАН. Вып. 6. Востоковедные чтения 2018. Лексикология и лексикография. К 200-летию ИВ РАН. М.: Институт востоковедения РАН, 2018. С. 3–4.
5. Исследования по языкам южной и юго-восточной Азии в Отделе языков народов Азии и Африки Института востоковедения РАН // Труды Института востоковедения РАН. Вып. 6. Востоковедные чтения 2018. Лексикология и лексикография. К 200-летию ИВ РАН. М.: Институт востоковедения РАН, 2018. С. 5–39.
6. (Составление) Выборочная библиография по языкам Южной и Юго-Восточной Азии, исследуемым в Отделе языков народов Азии и Африки ИВ РАН // Труды Института востоковедения РАН. Вып. 6. Востоковедные чтения 2018. Лексикология и лексикография. К 200-летию ИВ РАН. М.: Институт востоковедения РАН, 2018. С. 40–71.
7. К научной биографии В.М. Алпатова // Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. Труды научной конференции, к 70-летию члена-корреспондента РАН В.М. Алпатова. В двух томах. Т. 1. М.: ИВ РАН, 2016. С. 5–7.

8. (Составление) В.М. Алпатов. Библиографическая справка // Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. К 70-летию члена-корреспондента РАН В.М. Алпатова. Труды научной конференции. Институт востоковедения РАН 22–23 апреля 2015 г. В двух томах / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина, А.С. Зверев. – Т. 1. М.: ИВ РАН, 2016. С. 8–38.
9. Предисловие ко второму тому // Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. К 70-летию члена-корреспондента РАН В.М. Алпатова. Труды научной конференции. Институт востоковедения РАН 22–23 апреля 2015 г. В двух томах / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина, А.С. Зверев. – Т. 1. М.: ИВ РАН, 2016. С. 3–4.
10. Еще раз о границах слов в японских глагольных образованиях // Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. К 70-летию члена-корреспондента РАН В.М. Алпатова. Труды научной конференции. Институт востоковедения РАН 22–23 апреля 2015 г. В двух томах / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина, А.С. Зверев. – Т. 2. М.: ИВ РАН, 2016. С. 46–61.
11. Предисловие // Проблемы общей и востоковедной лингвистики 2014. В мире арабского языка. К 90-летию со дня рождения Г.Ш. Шарбатова. Труды научной конференции. Институт востоковедения РАН. 2016. Т. 1. М.: ИВ РАН, 2016. С. 3–4.
12. Г.Ш. Шарбатов: 90 лет со дня рождения // Проблемы общей и востоковедной лингвистики 2014. В мире арабского языка. К 90-летию со дня рождения Г.Ш. Шарбатова. Труды научной конференции. Институт востоковедения РАН. 2016. Т.1. М.: ИВ РАН, 2016. С. 5–29.
13. Языковые уровни в их соотношении между собой // Проблемы общей и востоковедной лингвистики 2013. Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Труды научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина. М.: ИВ РАН, 2014. С. 49–75.
14. 90 лет со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика // Проблемы общей и востоковедной лингвистики 2013. Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Труды научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Отв. редактор З.М. Шаляпина. Сост., ред. А.С. Панина. М.: ИВ РАН, 2014. С. 5–10.
15. Предисловие // Проблемы общей и востоковедной лингвистики 2013. Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Труды научной конференции. Институт востоковедения РАН 20 ноября 2013 г. Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина. М.: ИВ РАН, 2014. С. 3–4.
16. Памяти Юрия Ароновича Рубинчика // Ирано-славика. 2014. Том 25. № 1–2. С. 40–44.
17. Система русского синтеза RussLan при словарном поиске по японско-русским словарям в русско-японском режиме // Русский язык: исторические судьбы

- и современность: V Международный конгресс исследователей русского языка (Москва, МГУ им. М.В. Ломоносова, филол. фак-т, 18–21 марта 2014 г.): Труды и материалы / Сост. М.Л. Ремнёва, А.А. Поликарпов, О.В. Кукушкина. М.: Изд-во Моск. ун-та, 2014. С. 582–583. (соавтор: Канович М.И.)
18. Нина Васильевна Солнцева: памяти ученого // Вестник РГГУ. № 8 (130). Научный журнал. Серия «Филологические науки. Языкознание». Московский лингвистический журнал. Том 8. М., 2014. С. 209–215.
  19. (Составление) Н.В. Солнцева. Библиографическая справка // Вестник РГГУ. № 8 (130). Научный журнал. Серия «Филологические науки. Языкознание». Московский лингвистический журнал. Том 8. М., 2014. С. 216–223.
  20. Understanding and Generation of Texts from the Standpoint of the Three-Dimensional Stratificational Linguistic Framework // Understanding by Communication / Elena Borisova and Olga Souleimanova, eds. Newcastle upon Tyne: Cambridge Scholars Publishing, 2013. P. 6–32.
  21. Предисловие. 90 лет со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика // Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Тезисы докладов научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Науч. ред. З.М. Шалапина. Сост., ред. А.С. Панина / Институт востоковедения РАН. М.: ИВ РАН, 2013. С. 1–3.
  22. Публикации о жизни и трудах В.Ф. Вардуля // Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Тезисы докладов научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Науч. ред. З.М. Шалапина. Сост., ред. А.С. Панина / Институт востоковедения РАН. М.: ИВ РАН, 2013. С. 3–5.
  23. Публикации о жизни и трудах Ю.А. Рубинчика // Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Тезисы докладов научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Науч. ред. З.М. Шалапина. Сост., ред. А.С. Панина / Институт востоковедения РАН. М.: ИВ РАН, 2013. С. 5–7.
  24. Языковые уровни в их соотношении между собой // Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Тезисы докладов научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Науч. ред. З.М. Шалапина. Сост., ред. А.С. Панина / Институт востоковедения РАН. М.: ИВ РАН, 2013. С. 63–66.
  25. (Составление) Лев Николаевич Морев (1928–2012) / Восток. Афро-Азиатские общества: история и современность. 2013, № 6. С. 202-204. (соавтор: Погибенко Т.Г.)
  26. Некоторые возможности автоматической оптимизации лингвистического обеспечения системы обработки текстов при сущностном подходе к его организации // Вестник РГГУ. 2013, № 8. Научный журнал. Серия «Филологические науки. Языкознание». Московский лингвистический журнал. Т. 15. С. 182–208.
  27. X Международная научная конференция «Языки стран Дальнего Востока, Юго-Восточной Азии и Западной Африки» // Вопросы языкознания. 2013, № 4. С. 155–159.

28. Что это значит – лингвист? // Караван. Памяти Ю.А. Рубинчика. Специальный выпуск / Ежемесячный электронный журнал. Культурное представительство при Посольстве Исламской Республики Иран в Российской Федерации Август 2013. С. 37–41. URL: [http://www.moscow.icro.ir/uploads/karavan\\_03.pdf](http://www.moscow.icro.ir/uploads/karavan_03.pdf)
29. Порядок слов при японско-русском автоматическом переводе // Языки Дальнего Востока, Юго-Восточной Азии и Западной Африки: материалы X Международной конференции (Москва, 21–22 ноября 2012 г.): [сб. статей] / МГУ, Ин-т стран Азии и Африки, СПбГУ. Москва: Ключ-С. 2012. С. 314–327.
30. Понимание и порождение текстов с точки зрения трехмерной стратификационной модели языка и его функционирования // Общение в аспекте понимания. Сборник статей. Кн. 1. М.: МГПУ, 2012. С. 8–20.
31. Система русского синтеза RussLan в двух разных коммуникативных задачах // Понимание в коммуникации: человек в информационном пространстве. Сборник научных трудов / Российское общество преподавателей русского языка и лит-ры; Московский городской педагогический ун-т, ЯГПУ им. К.Д. Ушинского. Т. 1. Ярославль: Изд-во ЯГПУ, 2012. С. 244–261. (соавтор: Канович М.И.)
32. Новая жизнь старых словарей: традиция и современность // 100 лет Николаю Александровичу Сыромятникову. Тезисы докладов межинститутской научной конференции 14 декабря 2011 г. М.: ИВ РАН, 2011. С. 15–21. (соавторы: Демич К.И., Костыркин А.В., Тарасова Е.С.)
33. Выбор переводных эквивалентов в системе японско-русского автоматического перевода ЯРАП // Языки Дальнего Востока, Юго-Восточной Азии и Западной Африки. Материалы IX Международной конференции (Москва, 27–28 октября 2011 г.). / МГУ им. М.В. Ломоносова. ИСАА СПбГУ: [под общ. ред. Б.В. Касевича, М.И. Каплуна, М.А. Диас-Гонсалес]. М.: ИД Ключ-С, 2011. С. 159–166.
34. Памяти Николая Александровича Сыромятникова (к 100-летию со дня рождения) // Вестник РГГУ, № 11(73)/11. Научный журнал. Серия Филологические науки. Языкознание. Московский лингвистический журнал. Т. 13. М.: РГГУ, 2011. С. 298–306.
35. Понимание и порождение текстов с точки зрения трехмерной стратификационной модели языка и его функционирования // Материалы конференции «Понимание в коммуникации – 5» / Сб. работ. Департамент образования города Москвы. Государственное образовательное учреждение высшего профессионального образования города Москвы «Московский городской педагогический университет» (ГОУ ВПО МГПУ). Институт гуманитарных наук. Институт иностранных языков. Институт языкознания РАН. ISBN 978-5-243-00280-6. УДК 316. ББК 60.524. Э94. М.: ГОУ ВПО МГПУ, 2011. С. 195–197.
36. Экспериментальный комплекс ЯРАП для исследований по японско-русскому автоматическому переводу: 2008 г. // Бюллетень Общества востоковедов РАН. Вып. 17: Труды межинститутской научной конференции Востоковедные чтения-2008. Москва, 8–10 октября 2008 г. М.: ИВ РАН, 2010. С. 359–397. (соавторы: Костыркин А.В., Канович М.И., Модина Л.С., Панина А.С., Тарасова Е.С.)

37. Отдел языков народов Азии и Африки ИВ РАН: пятое десятилетие научной деятельности // Бюллетень Общества востоковедов РАН. Вып. 17: Труды меж-институтской научной конференции Востоковедные чтения-2008. Москва, 8–10 октября 2008 г. М.: ИВ РАН, 2010. С. 11–26.
38. РАМЕЯ/и автоматизированное рабочее место лингвиста для иероглифических языков // Вестник РГГУ, № 9(52)/10. Научный журнал. Серия Филологические науки. Языкознание. Московский лингвистический журнал. Том 12. М.: РГГУ, 2010. С. 117–148. (соавторы: Демич К.И., Костыркин А.В.)
39. Средства семантико-синтаксической обработки в системе русского синтеза RussLan // Русский язык: исторические судьбы и современность. IV Международный конгресс исследователей русского языка. Москва, МГУ им. М.В. Ломоносова, филологический факультет. 20–23 марта 2010 года. Труды и материалы (Составители М.Л. Ремнева, А.А. Поликарпов). М.: Изд-во Московского университета, 2010. С. 518–519. (соавтор: Канович М.И.)
40. Анализ текста и его место в модели перевода // Понимание в коммуникации: Сб. статей: в 2 тт. Том 2 / Составитель и автор предисл. Е.Г. Борисова. Департамент образования города Москвы. Государственное образовательное учреждение высшего профессионального образования города Москвы «Московский городской педагогический университет» (ГОУ ВПО МГПУ). Исторический факультет. Кафедра массовых коммуникаций. М.: ГОУ ВПО МГПУ, 2009 (10). С. 194–218.
41. К вопросу о наследовании валентностей в японском языке // VIII международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки (Москва, 22–24 сентября 2009 г.) / Ин-т стран Азии и Африки при МГУ им. М.В. Ломоносова. СПбГУ. Ин-т научной информации по общественным наукам РАН. ИВ РАН. М.: ИД Ключ-С. 2009. С. 205–221.
42. Об одной пионерской статье, или Принципы формализации семантики в 1961 году // Вестник РГГУ, № 6/09. Ежемесячный журнал. Серия Языкознание. Московский лингвистический журнал. Том 11. М.: РГГУ, 2009. С. 9-27.
43. Анализ текста и его место в модели перевода // Материалы конференции Понимание в коммуникации – 4 / МГПУ. Департамент образования города Москвы. Государственное образовательное учреждение высшего профессионального образования города Москва «Московский городской педагогический университет (ГОУ ВПО МГПУ)». Кафедра Массовых коммуникаций. Институт языкознания РАН. М., 2009. С. 70.
44. Отдел языков народов Азии и Африки ИВ РАН на 50-м году своей научной деятельности // Востоковедные чтения 2008. Тезисы докладов научной конференции 8–10 октября 2008 г. М.: ИВ РАН, 2008, С. 33.
45. Экспериментальный комплекс ЯРАП для исследований по японско-русскому автоматическому переводу: 2008 г. // Востоковедные чтения 2008. Тезисы докладов научной конференции 8–10 октября 2008 г. М.: ИВ РАН, 2008. С. 15. (соавторы: Костыркин А.В., Канович М.И., Модина Л.С., Панина А.С., Тарасова Е.С.)

46. Аналитические показатели управления в синтаксической структуре // Вестник РГГУ, № 6/08. Ежемесячный научный журнал. Серия Языкознание. Московский лингвистический журнал, № 10. М.: РГГУ, 2008. С. 11–57.
47. Нихонго-росиаго кикай хоньяку пуродзекуто JARAP {Японско-русский проект машинного перевода ЯРАП} // 日本語ロシア語機械翻訳プロジェクトJARAP. Гэнго сёри гаккай дай 14 кай нэндзи тайкай. Хаппё ромбунсю. Proceedings of the Fourteenth Annual Meeting of the Association for Natural Language Processing NLP. 2008 нэн 3 гацу 18 нити 20 нити. Кайдзё. Токё дайгаку. The University of Tokyo. Токио: гэнго сёри гаккай. The Association for Natural Language Processing, 2008. P. 317–320. (соавторы: Костыркин А., Канович М., Модина Л., Панина А., Тарасова Е.)
48. Моделирование языка при автоматическом переводе (в рамках сущностного подхода к лингвистическому описанию). Программа спецкурса для студентов-лингвистов // Вестник РГГУ, № 8/07. Ежемесячный журнал. Серия Языкознание. Московский лингвистический журнал, 2007. Т. 9, № 2. С. 168–190.
49. The JARAP Experimental Environment for Japanese-Russian Automatic Translation // 38. ICANAS. 10-15/09/2007 Ankara/Türkiye. Bildiri özetleri kitabı. Abstracts. Ankara/Türkiye: Atatürk kültür, dil ve tarih yüksek kurumu, 2007. P. 85. (соавторы: Kostyrkin, A.V., Kanovich, M.I., Modina, L.S., Panina, A.S., Tarasova, Ye.S.)
50. Аналитические показатели «сильного» управления в сущностной модели естественно-языкового синтаксиса // Понимание в коммуникации. 2007. Язык. Человек. Концепция. Текст: Тезисы докладов Международной научной конференции (28 февраля 1 марта 2007 г.). М.: НИВЦ МГУ, 2007. С. 135–138.
51. Some Aspects of Valency Inheritance in Japanese // Kinship in the Altaic World. Proceedings of the 28th Permanent International Altaistic Conference, Moscow 10–15 July, 2005 // Elena V. Boikova and Rostislav B. Rybakov (eds.). Wiesbaden: Harrassowitz Verlag, 2006. P. 269–282.
52. Владимир Иванович Любченко // Московский лингвистический журнал. 2006. Т. 9, № 1. С. 182–184.
53. Аппарат R-отсылок как универсальное средство синтаксического синтеза (на опыте разработки системы русского синтеза RUSSLAN) / Компьютерная лингвистика и интеллектуальные технологии (см. в книгах). 2006. Т. 2006. С. 207. (соавтор: Канович М.И.)
54. Борис Дмитриевич Тихомиров // Московский лингвистический журнал. 2006. Т. 9, № 1. С. 102–106. (соавторы: Драмбян Л.В., Малевич О.Б., Семенова Е.К.)
55. Аппарат R-отсылок как универсальное средство синтаксического синтеза (на опыте разработки системы русского синтеза RussLan) // Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции Диалог'2006 (Бекасово, 31 мая 4 июня 2006 г.) / Под ред. Н.И. Лауфер, А.С. Нариньяни, В.П. Селегея. М.: РГГУ, 2006. С. 207–213. (соавтор: Канович М.И.)
56. Коммуникативная организация предложения и его включение в текст. Русский язык. Методическая газета учителей-словесников. М.: ИД Первое сентября. Апрель (1–15) 2006, № 7 (486). С. 44–45.

57. Принципы описания японского синтаксиса в системе японско-русского автоматического перевода ЯРАП // Актуальные вопросы японского и общего языкознания. Памяти И.Ф. Вардуля / Отв. ред. В.М. Алпатов, З.М. Шаляпина; сост. А.С. Панина. Российская академия наук. Институт востоковедения. М.: Издательская фирма «Восточная литература» РАН, 2005. С. 210-244. (соавтор: Костыркин А.В.)
58. И.Ф. Вардуль и компьютерная лингвистика // Актуальные вопросы японского и общего языкознания. Памяти И.Ф. Вардуля / Отв. ред. В.М. Алпатов, З.М. Шаляпина; сост. А.С. Панина; Российская академия наук. Институт востоковедения. М.: Издательская фирма «Восточная литература» РАН, 2005. С. 31-34.
59. Some aspects of valency inheritance in Japanese // Permanent International Altaistic Conference. 48th Meeting, Moscow. July 10-15, 2005. ABSTRACTS // Institute of Oriental Studies. Russian Academy of Sciences. М.: ИВ РАН, 2005. С. 77-79.
60. Многофункциональность аппарата лексических функций и некоторые возможности его применения и развития // East West Encounter: Second International Conference on Meaning ↔ Text Theory. Восток Запад. Вторая международная конференция по модели «Смысл ↔ Текст» / Отв. ред. Ю.Д. Апресян, Л.Л. Иомдин. М.: Языки славянской культуры, 2005. С. 447-457.
61. An Entity-based Linguistic Framework. The Case of Japanese Syntax // Paris Lectures in Japanese Linguistics / Andre Wlodarczyk, ed. Tokyo: Kurosio Publishers, 2005. P. 167-211.
62. Система японско-русского автоматического перевода ЯРАП/1: первые экспериментальные результаты // Бюллетень (Newsletter) Общества востоковедов РАН. 10. Москва: ИВ РАН, 2004/2005. С. 164-226.
63. The JaRAP system of Japanese-Russian automatic translation in 2004 // Международный конгресс востоковедов. Тезисы. International Congress of Asian and North African Studies. Abstracts. Москва: ICANAS XXXVII, 2004. P. 177-178. (соавторы: Kanovich M.I., Kostyrkin A.V., Lubchenko V.I., Modina L.S., Panina A.S., Senina N.I., Sivtseva V.I., Tarasova Ye.S., Shternova O.A.)
64. Система японско-русского автоматического перевода ЯРАП в 2004 г. // Международный конгресс востоковедов. Тезисы. International Congress of Asian and North African Studies. Abstracts. Москва: ICANAS XXXVII, 2004. С. 178-179. (соавторы: Канович М.И., Костыркин А.В., Любченко В.И., Модина Л.С., Панина А.С., Сенина Н.И., Сивцева В.И., Тарасова Е.С., Штернова О.А.)
65. RUSSLAN: A System of Russian Language Generation // Investigations into Formal Slavic Linguistics. Contributions of the Fourth European Conference on Formal Description of Slavic Languages FDSL IV held at Potsdam University, November 28-30, 2001. (соавторы: Borisova, E.G., Kanovich, M.I., Panina, A.S., Tarasova, Y.S., Shternova, O.A.)
66. RUSSLAN: A System of Russian Language Generation // Investigations into Formal Slavic Linguistics. Contributions of the Fourth European Conference on Formal Description of Slavic Languages FDSL IV held at Potsdam University, November 28-30, 2001. Linguistik International. Part I. Band 10 / Peter Kosta, Joanna Blaszczak, Jens Frasek, Ljudmila Geist, Marzena Zygis (eds.). Frankfurt am Main,



- Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang Verlag (Europaeischer Verlag der Wissenschaften), 2003. P. 385–403. (соавторы: Borisova, E.G., Kanovich, M.I., Panina, A.S., Tarasova, Y.S., Shternova, O.A.)
67. О нашем учителе // Московский лингвистический журнал. 2003. Т. 7, № 1. С. 6–8. (соавторы: Штернова О.А., Хайлова И.М., Видяпина В.В., Бурлова Г.М., Шагалова М.Г., Бакулов А.Д.)
  68. Программа курса «Проблемы и методы автоматического перевода» // Московский лингвистический журнал. 2003. Т. 7. № 1. С. 125–154.
  69. Межязыковой переход в системе японско-русского автоматического перевода ЯРАП // Языки Дальнего Востока, Юго-Восточной Азии и Западной Африки. Доклады VII-ой Международной конференции. Москва, 16–19 сентября 2003 г. Ч. II. М.: ИСАА при МГУ, 2003. С. 181–192.
  70. RUSSLAN: an experimental computational model of Russian generation (the pilot version released in June, 2002). URL: [http://isir.ras.ru/win/db/show\\_ref.asp?P=.id-31897.In-ru](http://isir.ras.ru/win/db/show_ref.asp?P=.id-31897.In-ru) (соавторы: Kanovich, M.I., Shternova, O.A., Borisova, E.G., Panina, A.S., Tarasova, Ye.S.)
  71. Основные направления исследований Отдела языков народов Азии Института востоковедения РАН // Бюллетень (Newsletter) Общества востоковедов Российской Академии наук. 6. Материалы II Всероссийского съезда востоковедов. Казань, 18–22 октября 1999 г. М.: ИВ РАН, 2002. С. 162–171.
  72. Japanese with relation to Russian in the JaRAP system of Japanese-Russian automatic translation // International Conference. Japan Phenomenon: Views from Europe / Russian Association of Japanologists. Japan Foundation. Institute of Oriental Studies (Russian Academy linguistik internationalof sciences). Japan Today. To World Oriental Congress Moscow, 2004. Moscow, 2001. P. 132–139.
  73. Структурные валентности как универсальный инструмент описания языковой сочетаемости (при сущностном подходе к ее моделированию) // Московский лингвистический журнал. 2001. Т. 5, № 2. С. 35–84.
  74. RUSSLAN: система русского лексико-синтаксического синтеза // Fourth European Conference on Formal Description of Slavic Languages FDSL-4. 28-30 November, 2001. Hosted by University of Potsdam. Department of Slavic Languages sponsored by German Research Foundation (DFG). ABSTRACTS. Potsdam, 2001. P. 144–145. (соавторы: Борисова Е.Г., Канович М.И., Панина А.С., Тарасова Е.С., Штернова О.А.)
  75. RUSSLAN: система русского лексико-синтаксического синтеза / Fourth European Conference on Formal Description of Slavic languages FDSL-4 Abstracts Potsdam University: 2001. P. 144–145. (соавторы: E.G. Borisova, M.I. Kanovich, A.S. Panina, Y.S. Tarasova, O.S. Shternova)
  76. Еще раз о границах японского синтаксиса // VI-я Международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки (25–28 сентября 2001 г.). Материалы и тезисы докладов. СПб.: СПбГУ, Восточный факультет, 2001. С. 160-174.
  77. Once more on the scope of Japanese syntax // VI-я Международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки (25–

- 28 сентября 2001 г.). Материалы и тезисы докладов. СПб.: СПбГУ, Восточный факультет, 2001. С. 174–176.
78. Проблемы русского лексико-синтаксического синтеза при сущностном подходе к языку // Русский язык: исторические судьбы и современность. Международный конгресс исследователей русского языка. Москва, филологический факультет МГУ им. М.В. Ломоносова. 13–16 марта 2001 года. Труды и материалы. (Под общей ред. М.Л. Ремневой и А.А. Поликарпова). М.: Изд-во Московского университета, 2001. С. 427–428. (соавторы: Борисова Е.Г., Канович М.И., Панина А.С., Тарасова Е.С., Штернова О.А.)
  79. Средства компьютерной обработки японских текстов, разрабатываемые в Институте востоковедения РАН // Российская Академия наук. Общество востоковедов. Бюллетень (Newsletter) 4. Материалы II Всероссийского съезда востоковедов. Казань, 13–22 октября 1999 г. М.: ИВ РАН, 2000. С. 104–118.
  80. Лексико-синтаксическая информация в словаре для автоматического перевода (В рамках сущностного подхода к лингвистическому описанию) // Словарь в современном мире. Материалы Третьей международной школы-семинара. Иваново 14–16 сентября 1999 г. Отв. ред. О.М. Карпова. Иваново, 2000. С. 117–121. (соавторы: Канович М.И., Костыркин А.В., Модина Л.С.)
  81. The JaRAP experimental environment as a tool for contrastive studies of Japanese and Russian // Вторая зимняя типологическая школа. Материалы международной школы-семинара молодых ученых по лингвистической типологии и антропологии. Москва: РГГУ, ИВ РАН, Международная типологическая ассоциация, Ассоциация «Профессионалы за сотрудничество», 2000. С. 72–77.
  82. Система ИРИС иероглифический редактор и справочник по японской графике и лексике // V-я международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки (8–10 сентября 1999 г.). Материалы и резюме докладов. СПб., 1999. С. 52–58. (соавторы: Любченко В.И., Костыркин А.В., Модина Л.С., Сенина Н.И., Сивцева В.М.)
  83. Организация лексико-синтаксических знаний в модели русского синтеза RU-MORS // Труды Международного семинара Диалог'99 по компьютерной лингвистике и ее приложениям. В двух томах. Под ред. А.С. Нариньяни. Т. 2. Приложения. Таруса, 1999. С. 326–333. (соавторы: Канович М.И., Штернова О.А.)
  84. Межъязыковые операции в модели японско-русского автоматического перевода ЯРАП // Электронные библиотеки и базы данных по истории Евразии в средние века. Вып. 7. М.: ИВ РАН, 1999. С. 172–231. (соавторы: Модина Л.С.)
  85. Оппозиция «часть–целое» и сущностный подход к моделированию языковой компетенции // Роман Якобсон: тексты, документы, исследования. М.: РГГУ, 1999. С. 541–551.
  86. Экспериментальный комплекс ЯРАП как инструмент сопоставительных лингвистических исследований // Общее и восточное языкознание. Сборник научных трудов, посвященных 70-летию члена-корреспондента РАН В.М. Солнцева. М.: Современный писатель, 1999. С. 64–78. (соавторы: Модина Л.С., Канович М.И.)

87. Towards a conception of linguistic competence underlying diverse modes of linguistic performance // Caron, B. (ed.). Proceedings of the XVIth International Congress of Linguists. CD-ROM. Pergamon, Oxford: Elsevier Science Ltd., 1998. Paper № 0430. 26 p.
88. Проблемы и перспективы автоматического перевода с японского языка на русский // Языки Азии и Африки: традиции, современное состояние и перспективы исследований. Материалы научной конференции 5–8 октября 1998 г. М.: ИВ РАН, 1998. С. 137–143. (соавторы: Модина Л.С., Канович М.И., Любченко В.И.)
89. Словари различных лингвистических уровней в модели японско-русского автоматического перевода ЯРАП // Актуальные проблемы теоретической и прикладной лексикографии. Межвузовский сборник научных трудов. Иваново: Юнона, 1997. С. 238–249. (соавторы: Модина Л.С., Канович М.И., Любченко В.И., Панина А.С., Сенина Н.И., Сивцева В.И., Тарасова Е.С., Хайлова И.М., Штернова О.А.)
90. Лексико-грамматическая классификация японских лингвистических существительных и ее функции в модели японско-русского автоматического перевода ЯРАП // IV Международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки. 17–20 сентября 1997 г. Тезисы докладов. Ч. I. А–П. М.: МГУ им. М.В. Ломоносова, Ин-т стран Азии и Африки; СПбГУ, Восточный факультет, 1997. С. 142–150. (соавтор: Модина Л.С.)
91. An Integrated Experimental Environment for Lexico-Morphological Research in Japanese-Russian Automatic Translation // Oriental Studies in the 20th Century: Achievements and Prospects. Abstract of the papers of CIS scholars for the 35th ICANAS (Budapest, Jul.7-12, 1997). Vol. 2. Moscow, 1997. P. 329–332. (соавторы: Modina, L.S., Kanovich, M.I., Lubchenko, V.I.)
92. Towards a conception of linguistic competence underlying diverse modes of linguistic performance // XVIth International Congress of Linguists. Abstracts. Paris, 1997. P. 271. (Франц. перевод: Sur la conception de compétence linguistique sousjacente des modes de performance linguistique divers, там же).
93. Об одном возможном подходе к задаче анализа японских текстов при их автоматической обработке // Проблемы компьютерной лингвистики. Минск: Минский Гос. лингвистический университет, 1997. С. 110–144.
94. Об одном формализме для записи толкований слов и словосочетаний // Обработка текста и когнитивные технологии. Вып. 1. М., Пушкино: ОНТИ ПНЦ РАН, 1997. С. 73–110.
95. Русский морфологический и морфолого-синтаксический синтез по неполной и/или неточной входной информации // Русский язык как иностранный: лингвистические проблемы (к 20-летию кафедры современного русского языка). М.: ИКАР, 1997. С. 137–147. (соавтор: Канович М.И.)
96. Опозиция «часть–целое» и сущностный подход к моделированию языковой компетенции // Материалы Международного конгресса «100 лет Р.О. Якобсону». Москва, 1996. С. 71–73.

97. Компьютерная среда ЯРАП для изучения и исследования особенностей японского и русского языков в процессе перевода. Материалы Международной выставки-ярмарки при II Международном конгрессе ЮНЕСКО «Образование и информатика» (1–5 июля 1996 г.). 20 с. (соавторы: Модина Л.С., Канович М.И., Любченко В.И., Панина А.С., Сенина Н.И., Сивцева В.И., Тарасова Е.С., Хайлова И.М., Штернова О.А.)
98. Словоизменение и простейшие виды словообразования в модели русской морфологии, ориентированной на задачи синтеза // Труды международного семинара ДИАЛОГ'96 по компьютерной лингвистике и ее приложениям. Пушкино 4–9 мая 1996 г. / Под ред. А.С. Нариньяни. М., 1996. С. 97–102. (соавтор: Канович М.И.)
99. Автоматический перевод: эволюция и современные тенденции // Вопросы языкознания. 1996, № 2. С. 105–117.
100. Машинный перевод // Программы и учебный план отделения теоретической и прикладной лингвистики. М.: МГУ, 1996. С. 78–82.
101. Принципы анализа японских текстов при моделировании японской лексико-морфологической системы // Труды международного семинара ДИАЛОГ'96 по компьютерной лингвистике и ее приложениям. Пушкино 4–9 мая 1996 г. / Под ред. А.С. Нариньяни. М., 1996. С. 169–174. (соавтор: Модина Л.С.)
102. Принципы организации лингвистических знаний в объектно-ориентированной модели лексико-морфологической системы японского языка // DIALOG'95. Труды Международного семинара по компьютерной лингвистике и ее приложениям. Казань, 31 мая – 4 июня 1995 г. Казань, 1995. С. 198–205. (соавтор: Модина Л.С.)
103. Разработка и экспериментальная верификация моделей лингвистических знаний и механизмов лексико-морфологического уровня для японского и русского языков в связи с задачами японско-русского автоматического перевода / Отчет о НИР № 95-01-01136-а. Российский фонд фундаментальных исследований // Информационный бюллетень РФФИ. 1995. Т. 3. № 1. С. 366. (соавторы: Канович М.И., Любченко В.И., Модина Л.С., Панина А.С., Сенина Н.И., Сивцева В.И., Тарасова Е.С., Хайлова И.М., Штернова О.А.)
104. Лингвистическая избыточность как фактор организации лексико-морфологической базы в модели русского словоизменительного синтеза // DIALOG'95. Труды Международного семинара по компьютерной лингвистике и ее приложениям. Казань, 31 мая – 4 июня 1995 г. Казань, 1995. С. 133–139. (соавтор: Канович М.И.)
105. Экспериментальная среда для обработки японских текстов в связи с задачами японско-русского автоматического перевода // Международная конференция. Памятники духовной, материальной и письменной культуры древнего и средневекового Востока (создание баз данных). 30 мая – 4 июня 1995 г. Тезисы докладов. М.: ИВ РАН, РФФИ, РГНФ. 1995. С. 48–50. (соавтор: Модина Л.С., Любченко В.И.)

106. Графико-морфологическая грамматика японского языка в действующей модели анализа японских письменных текстов // III Международная конференция «Языки Дальнего Востока, Юго-Восточной Азии и Западной Африки». Тезисы докладов. М.: МГУ, ИСАА. СПбГУ, 1995. С. 120–123. (соавтор: Модина Л.С.)
107. Автоматический перевод: эволюция и современные тенденции // Лингвистика на исходе XX века: итоги и перспективы. Тезисы международной конференции. Т. II. М.: МГУ, 1995. С. 562–563.
108. The RUMORS system of Russian synthesis // COLING 94. The 15th International Conference on Computational Linguistics. Proceedings. Vol. I. Kyoto, 1994. P. 177–179. (соавтор: Kanovich M.I.)
109. The JaRAP experimental system of Japanese-Russian automatic translation // COLING 94. The 15th International Conference on Computational Linguistics. Proceedings. Vol. I. Kyoto, 1994. P. 112–114. (соавтор: Modina L.S.)
110. Экспериментальный комплекс ЯРАП/л для лингвистических исследований по японско-русскому автоматическому переводу // Языки Юго-Восточной Азии, Дальнего Востока и Западной Африки: сходства и различия. М., 1993. С. 17–20. (соавтор: Модина Л.С.)
111. Грамматика и ее соотношение со словарем при словоцентрическом подходе к языку (На опыте формализованного лингвистического описания) // Вопросы языкознания. 1991. № 5. С. 42–54.
112. Система автоматического получения русского подстрочника японских научно-технических текстов // Восток. Прошлое и будущее народов. IV Всесоюзная конференция востоковедов (новые подходы в теории и методиках востоковедных исследований). Тезисы докладов и сообщений. Махачкала, 1–5 октября 1991. Т. 1. Языкознание. Культурология, философия, искусство, архитектура, музыка. Экономика. М., 1991. С. 81–85. (соавтор: Модина Л.С.)
113. Обработка сложных терминов в системе японско-русского автоматического перевода ЯРАП // Международная научная конференция «Вопросы стандартизации, интернационализации и автоматизации перевода терминологических единиц». Черновцы. 8–11 октября 1991 г. Тезисы докладов. Черновцы, 1991. С. 92–94. (соавтор: Модина Л.С.)
114. Словарь как единая форма организации формализованного лингвистического описания // Proceedings of the Fourteenth International Congress of Linguists. Berlin/ GDR, August 10 August 15, 1987. II // W. Baher, J. Schildt, D. Viehweger (eds.). Berlin: Akademie-Verlag, 1991. P. 1244–1247.
115. Автоматизированная система лексико-морфологического анализа японских текстов // III международная научно-техническая конференция «Программное обеспечение ЭВМ». Ноябрь 1990, Тверь. Секция 4. Опыт разработки и использования систем искусственного интеллекта. Тезисы докладов. Тверь, 1990. С. 65. (соавтор: Модина Л.С.)
116. Об одном подходе к выделению фразеологических единиц в связи с задачами автоматической обработки текстов // Фразеография в машинном фонде русского языка. М.: Наука, 1990. С. 188–195.

117. Отечественные системы машинного перевода // Справочник. Искусственный интеллект. Кн. 1. Системы общения и экспертные системы. Под ред. проф. Э.В. Попова. М.: Радио и связь, 1990. Глава 12. С. 248–261. (соавтор: Бакулов А.Д., Леонтьева Н.Н.)
118. Современное состояние машинного перевода // Справочник. Искусственный интеллект. Кн. 1. Системы общения и экспертные системы. Под ред. проф. Э.В. Попова. М.: Радио и связь, 1990. Глава 11. С. 216–248. (соавтор: Леонтьева Н.Н.)
119. Система СПРИНТ/Я: современное состояние // Международный семинар по машинному переводу «ЭВМ и перевод 89» (Тбилиси, 27 ноября – 2 декабря 1989 г.). Тезисы докладов. М.: ВЦП, ГРУЗНИИТИ, 1989. С. 101–104. (соавторы: Драмбян Л.В., Семенова Е.К., Тихомиров Б.Д.)
120. Принципы описания японской морфологии в системе ЯРАП // Международный семинар по машинному переводу «ЭВМ и перевод 89» (Тбилиси, 27 ноября – 2 декабря 1989 г.). Тезисы докладов. М.: ВЦП, ГРУЗНИИТИ, 1989. С. 321–324.
121. Текст как объект автоматического перевода // Текст и перевод. М.: Наука, 1988. С. 113–129.
122. Выступление на совещании по проблемам подготовки специалистов по прикладной лингвистике // Экспериментальная лингвистика в языковом вузе (Стенограмма научного совещания) / Публ. проблемной группы по экспериментальной и прикладной лингвистике, вып. 177. М.: ИРЯ АН СССР, 1987. С. 32–34.
123. О системе японско-русского автоматического перевода ЯРАП // Всесоюзная конференция «Совершенствование перевода научно-технической литературы и документов». Тезисы докладов и сообщений. М.: ВЦП, 1988. С. 223–225. (соавторы: Драмбян Л.В., Семенова Е.К., Тихомиров Б.Д.)
124. К проблеме моделирования языковой компетенции и языковой деятельности человека в рамках общей модели языковой действительности // Референция и проблемы текстообразования / Проблемная группа «Логический анализ языка». М.: Ин-т языкознания АН СССР, 1988. С. 214–234.
125. Three functions of language sign and the problem of their representation in linguistic description // *Linguistics: A Soviet Approach* / M.S. Andronov, Bhakti P. Malik (eds.), Swapan Banerjee (coed.) / *Indian Journal of Linguistics*. Calcutta, 1988. P. 110–142.
126. Принципы формального анализа японского текста в связи с задачей его перевода // Проблемы языков Азии и Африки. Материалы II Международного Симпозиума. Варшава-Краков, 10–15 ноября 1980 г. Варшава, 1987. *Problemy języków Azji i Afryki. Materiały II Międzynarodowego Symposium Warszawa-Kraków 10–15 listopada 1980*. Warszawa: Państwowe Wydawnictwo Naukowe, 1987. С. 111–126.
127. Словарь как единая форма организации формализованного лингвистического описания // *Abstracts of Section Papers and Round-Table Talks*. XIV International

Congress of Linguists. Berlin, August 10 to 15, 1987. Organized under the Auspices of CIPL. Berlin, May 1987. P. 188.

128. A model of human linguistic competence as a unified basis for a set of fundamentally dissimilar models of linguistic performance // ABSTRACTS. 8 Международный конгресс по логике, методологии и философии науки. LMPS'87. Moscow, USSR, 17–22 August 1987. Vol. 1. Sections 1–5, 7, 12. М.: ИФ АН СССР, 1987. С. 524–526.
129. К вопросу о валентности как средстве представления синтагматической информации в автоматическом словаре // Сочетание лингвистической и внелингвистической информации в автоматическом словаре (Материалы Всесоюзной школы-семинара). Ереван: Изд-во АН АрмССР, 1987. С. 187–189.
130. Общий порядок обработки японского текста при получении его нормализованного русского подстрочника // Отчет о научно-исследовательской работе «Разработать и ввести в эксплуатацию пакет прикладных программ, обеспечивающих перевод с японского языка на русский (ПОЯРАП) (результаты предпроектных исследований)» / УДК 681.326/681.39:801.3 № гос. регистрации Г.Р. № 01.86.0 039077. М.: ВЦП, 1986. С. 8–16.
131. Some problems of distinguishing between word and morpheme in Japanese // *Travaux de linguistique Japonaise*. Vol. VIII / Université de Paris 7. UER. Langues et civilisations de l'Asie Orientale. Actes du 4-e Colloque International d'Etudes Japonaises (section linguistique). Proceedings of the 4th International Studies Conference on Japan (Section of Linguistics). Paris, 1986. P. 131–124.
132. О морфолого-синтаксическом статусе некоторых японских грамматических единиц (в связи с общими проблемами разграничения морфологии и синтаксиса в естественном языке) // Актуальные проблемы японского языкознания. М.: Наука, 1986. С. 69–103.
133. О некоторых теоретических аспектах понятия «единица автоматического перевода» // Международная конференция «Теория и практика научно-технического перевода». Москва, 1985 г. Тезисы докладов. М.: ВЦП, 1985. С. 69–71.
134. Три функции языкового знака и проблемы их отражения в лингвистическом описании // Проблемы функциональной грамматики. М.: Наука, 1985. С. 103–120.
135. О двух типах условий кореферентности валентностей в предикатно-актантных конструкциях // Типология конструкций с предикатными актантами. Л.: Наука, 1985. С. 46–50.
136. Коммуникативная организация предложения: функциональное содержание и средства выражения // Восточное языкознание: грамматическое и актуальное членение предложения. М.: Наука, 1984. С. 146–157.
137. Организация и самоорганизация лингвистического обеспечения в системе автоматического перевода // Международный семинар по машинному переводу. Тезисы докладов. М.: ВЦП, 1983. С. 238–240.
138. О морфолого-синтаксическом статусе некоторых грамматических единиц японского языка (к определению слова в японском языке) // К III Междуна-

- родному симпозиуму по теоретическим проблемам языков Азии и Африки (Тезисы докладов советских языковедов). М.: Наука, 1983. С. 72–75.
139. Вариантность как свойство языковой системы // Изв. АН СССР. Серия литературы и языка, 1983. Т. 42. № 4. С. 383–391. (соавторы: Каменева М.С., Кобозева И.М.)
  140. Три функции языкового знака и их разграничение в лингвистическом описании. // Проблемы функциональной грамматики (Тезисы конференции). Звенигород, 1–2 апреля 1983 г. М.: Ин-т языкознания АН СССР, 1983. С. 45–47.
  141. Об одном способе формализации понятия лингвистической нормативности в целях автоматической обработки текстов // Формальное представление лингвистической информации. Н.: ВЦ СО АН СССР, 1982. С. 20–37.
  142. Лексико-синтаксическая структура предложения как показатель его коммуникативной организации // Теоретические проблемы восточного языкознания. Ч. 2. М.: Наука. Гл. ред. вост. литературы, 1982. С. 129–136.
  143. К построению формализованной модели языка, допускающей вариантность соотносящихся с ней моделей языковой деятельности человека // Вариантность как свойство языковой системы (тезисы докладов) Ч. 2. М.: Наука. Гл. ред. вост. литературы, 1982. С. 119–122.
  144. Об одном подходе к описанию межфразовой структуры текстов (в связи с разработкой формальной модели процесса перевода с японского языка на русский) // Вторая всесоюзная школа молодых востоковедов (Тбилиси, октябрь 1982 г.). Тезисы. Т. II. Языкознание, литературоведение. М.: ИВ АН СССР, 1982. С. 139–142.
  145. Межфразовые связи как способ реализации лексических валентностей // Аспекты изучения текста. Сборник научных трудов. М.: УДН, 1981. С. 80–87.
  146. Об одном типе японских глагольных конструкций // Всесоюзная конференция Совершенствование перевода научно-технической литературы и документов. Тезисы докладов и сообщений. М.: ВЦП, 1982. С. 207–208.
  147. Grammatic and semantic normativity of linguistic units and features as a factor of automatic text processing // COLING 82. Abstracts. E. Hajičova (ed.). Prague: Charles University, 1982. P. 255–258.
  148. Синтаксис и семантика в системе англо-русского автоматического перевода АРАП // Актуальные вопросы практической реализации систем автоматического перевода. Ч. 1. М.: МГУ, 1982. С. 184–204.
  149. О понятии лингвистической нормативности в свете задач автоматической обработки текста // Тезисы докладов. Семинар проекта ДИАЛОГ. Сангасте 1–6 марта 1982 г. Тарту: ТГУ, 1982. С. 93–95.
  150. Сегментация и морфологический анализ японского текста // Переработка текста методами инженерной лингвистики. Тезисы докладов. Минск: МГПИИЯ, 1982. С. 67–69. (соавтор: Хайлова И.М.)
  151. Система англо-русского автоматического перевода (АРАП) и японско-русского автоматического перевода (ЯРАП) // Синтаксический компонент в системах машинного перевода / Всесоюзный центр переводов. Серия 2. Машин-



- ный перевод и автоматизация информационных процессов. 5. М.: ВЦП, 1981. С. 62–69. (соавтор: Кудряшова И.М.)
152. Вопросы синтаксиса на Восьмой международной конференции по вычислительной лингвистике в Токио (COLING 80) // Синтаксический компонент в системах машинного перевода / Всесоюзный центр переводов. Сер. 2. Машинный перевод и автоматизация информационных процессов. 5. М.: ВЦП, 1981. С. 107–115.
153. Хроникальная заметка о II-м Международном симпозиуме ученых социалистических стран на тему «Теоретические проблемы языков Азии и Африки» // Вопросы языкознания. 1981, № 6. С. 150–153. (соавтор: Белова А.Г.)
154. Семантические элементы, семантические отношения и их взаимосвязь в системе элементарных семантических единиц // Представление знаний и моделирование процессов понимания. Новосибирск: ВЦ СО АН СССР, 1980. С. 31–46.
155. К вопросу об операционном критерии разграничения слова и морфемы в связи с задачей формального анализа японских текстов // Тезисы рабочего совещания по морфеме (ноябрь 1980 г.). М.: Наука. Гл. ред. вост. литературы, 1980. С. 108–115.
156. Принципы формального анализа японского текста в связи с задачей его перевода // Тезисы докладов советской делегации на 2-ом симпозиуме ученых социалистических стран на тему «Теоретические проблемы языков Азии и Африки» (Варшава-Краков, 9–16 ноября 1980 г.). М.: ИВ АН СССР, 1980. С. 60–63.
157. Problems of formal representation of text structure from the point of view of automatic translation // COLING 80. Proceedings of the 8th International Conference on Computational Linguistics. Sept. 30–Oct. 4, 1980, Tokyo. Tokyo, 1980. P. 174–182.
158. Автоматический перевод как моделирование переводческой деятельности человека // Международный форум по информации и документации. 1980. Т. 5, № 2. С. 16–20.
159. Automatic translation as a model of the human translation activity // International Forum on Information and documentation. 1980. Vol. 5, № 2. P. 20–23.
160. Формальный язык для записи толкований слов и словосочетаний // Проблемы кибернетики. Вып. 36. М., 1979. С. 247–278.
161. Язык для записи семантической информации о единицах естественного языка: общая структура и некоторые возможности использования // Синтаксический и семантический компонент лингвистического обеспечения. Новосибирск: ВЦ СО АН СССР, 1979. С. 33–46.
162. Вопросы формального представления структуры текста с точки зрения задач автоматического перевода // Международный семинар по машинному переводу. Москва, 1979 г. Тезисы докладов. М.: ВЦП, 1979. С. 130–132.
163. Об одном формальном свойстве системы элементарных семантических единиц // Семантика естественных и искусственных языков в специализированных системах. Тезисы. Л.: ЛГУ, 1979. С. 173–174.

164. К вопросу об организации связного текста и возможностях ее учета при типологическом сопоставлении языков // XIV Тихоокеанский научный конгресс. СССР, Хабаровск, август 1979. Комитет Л. Социальные и гуманитарные науки. Секция III. Этнокультурные проблемы изучения народов Тихоокеанского региона. Секция IV. Языки бассейна Тихого океана. Тезисы докладов / Тихоокеанская научная ассоциация. Т. II. М., 1979. С. 298–300.
165. Принципы выявления лексико-грамматического состава японского текста при его формальном анализе // Тезисы докладов IV научной конференции по японской филологии / МГУ им. М.В. Ломоносова. ИСАА. Кафедра японской филологии. М.: ИСАА, 1979. С. 56–58.
166. О понятии грамматики с точки зрения формализованного описания естественного языка // Семантические вопросы искусственного интеллекта. Киев, 1978. С. 28.
167. Лексико-синтаксическая структура предложения как показатель его коммуникативной организации // Тезисы докладов I международного симпозиума ученых социалистических стран на тему «Теоретические проблемы восточного языкознания». Ч. 2. М.: ИВ АН СССР, 1977. С. 91–94.
168. О записи и организации лингвистической информации в связи с задачей автоматического анализа текста // Автоматическая переработка текста методами прикладной лингвистики. Материалы II Всесоюзной конференции 6-7 октября 1973 г. Кишинев, 1977. С. 43–45.
169. Многоаспектный автоматический словарь с грамматическим обеспечением и его использование при автоматическом анализе текстов // «Вопросы кибернетики». Формализация языковых данных. Вып. 17. М.: Советское радио, 1977. С. 95–115.
170. К вопросу о средствах выражения коммуникативной организации предложения и их взаимосвязи // Проблемы синтаксической семантики. Материалы научной конференции. М., 1976. С. 261–263.
171. К проблеме коммуникативной организации текста и ее отражения в семантической записи // Вычислительная лингвистика. М.: Наука, 1976. С. 53–80.
172. Об одном подходе к формальному описанию семантики слова // О некоторых вопросах автоматической обработки текстовой информации / Институт кибернетики АН УССР. Киев, 1976. С. 35–50.
173. О формальном языке для описания лексической семантики // Семантические вопросы искусственного интеллекта. Киев, 1975. С. 20–21.
174. Система АРАП: словарь, грамматика и их использование при автоматическом анализе // Международный семинар по машинному переводу. Москва, 25–27 ноября 1975 г. Тезисы докладов и сообщений. М.: ВЦП. С. 135–138.
175. Об англо-русском многоаспектном словаре с грамматическим обеспечением // Вопросы информационной теории и практики. Сборник № 27. М.: ВИНТИ, 1975. С. 64–87. (соавторы: Афанасьева Л.А., Ельницкий Л.Л., Заворыгина Т.В., Павлова Н.Н., Перцов Н.В., Пивоварова Т.В., Потемкина Т.Г., Убин И.И., Чехов А.С., Шаталова М.Г., Эрастов К.О.)

176. Некоторые типы информации, задаваемые синтаксической структурой подчинительных сочетаний слов, и их отражение в семантической записи текста // *Материалы семинара по теоретическим проблемам синтаксиса (2-5 марта 1975 г.)*. Ч. II. Пермь: ПермГУ, 1975. С. 281–284.
177. К проблеме построения формальной модели процесса перевода // *Теория перевода и научные основы подготовки переводчиков. Материалы всесоюзной научной конференции*. Ч. II. М.: МГПИИЯ им. М. Тореца, 1975. С. 165–172.
178. О соотношении синтаксиса связного текста с синтаксисом предложения // «*Вопросы методики использования текста в обучении иностранным языкам*». Материалы конференции молодых научных работников. М.: МГПИИЯ им. М. Тореца, 1975. С. 217–225.
179. *Англо-русский многоаспектный автоматический словарь (АРМАС) // Машинный перевод и прикладная лингвистика*. Вып. 17. М.: МГПИИЯ им. М. Тореца, 1974. С. 7–67.
180. Семантические требования при автоматической обработке текстов // *Лингвистическое обеспечение автоматизированных систем управления и информационно-поисковых систем (Краткое содержание докладов научно-технической конференции 29–31 мая 1974 г.)*. Тюмень, 1974. С. 41–42.
181. Об одном подходе к задаче обнаружения синтактико-семантической структуры текста // *Лингвистика текста. Материалы научной конференции*. Ч. II. М.: МГПИИЯ им. М. Тореца, 1974. С. 171–177.
182. Семантические элементы как инструмент описания смысловой сочетаемости // *Тезисы научной конференции молодых научных работников «Проблемы компонентного анализа» (24–25 апреля 1973 г.)*. М.: МГПИИЯ им. М. Тореца, 1973. С. 132–133.
183. Семантические элементы как основа лексикографического описания: общие принципы и формальный аппарат // *Машинный перевод и прикладная лингвистика*. Вып. 15. М.: МГПИИЯ им. М. Тореца, 1972. С. 54–98.
184. Язык формальных толкований как средство описания смысла слов при автоматическом переводе // *Автоматическая обработка текстов на естественных языках. Тема 1026.1.2. Семинар стран членов СЭВ. Ереван 13–18 ноября 1972 г.* Ереван: АрмНИИНТИ, 1972. – 16 с.
185. Язык для записи лингвистической информации в автоматическом словаре // *Машинный перевод и прикладная лингвистика*. Вып. 16. М.: МГПИИЯ им. М. Тореца, 1972. С. 113–152. (соавтор: Арсентьева Н.Г.)
186. Всесоюзная конференция по структурно-математическим методам моделирования языка // *Машинный перевод и прикладная лингвистика*. Вып. 14. М.: МГПИИЯ им. М. Тореца, 1971. С. 221–229. (соавторы: Белецкая И.П., Белецкий М.И., Перебейнос В.И., Скороходько Э.Ф.)
187. О структуре словаря для автоматического перевода с использованием семантики // *Автоматическая переработка текста методами прикладной лингвистики. Материалы всесоюзной конференции 6–8 октября 1971 г.* Кишинев, 1971. С. 180–182.

188. Словарные статьи СКОРОСТЬ. КОРОБКА СКОРОСТЕЙ // Апресян Ю.Д., Мельчук И.А., Перцов Н.П., Шаляпина З.М. Материалы к толково-комбинаторному словарю русского языка – 6 / Публ. проблемной группы по экспериментальной и прикладной лингвистике. Вып. 38. М.: ИРЯ АН СССР, 1971. С. 3–18. (соавторы: Апресян Ю.Д., Мельчук И.А.)
189. Зимняя школа по лингвистическим проблемам автоматической обработки текстов // Машинный перевод и прикладная лингвистика. Вып. 13. М.: МГПИИЯ им. М. Тореза, 1970. С. 218–221. (соавторы: Перлина Н., Саввина Е., Чепиго Т., Эрастов К.)
190. Об основных проблемах построения модели английской орфографии // Машинный перевод и прикладная лингвистика. Вып. 13. М.: МГПИИЯ им. М. Тореза, 1970. С. 151–176.
191. К проблеме семантической конверсии (на материале английского языка) // Структурно-математические методы моделирования языка. Тезисы докладов и сообщений всесоюзной научной конференции. Ч. II. Киев: КГУ, 1970. С. 137–138.
192. Конференция по проблеме прикладной лингвистики // Русский язык в национальной школе, 1970, № 2. С. 82–85.
193. Некоторые вопросы семантики в связи с построением модели английского словообразования // Материалы XXV научной студенческой конференции. Литературоведение, лингвистика. Тарту: ТГУ, 1970. С. 70–72.
194. Об орфографическом компоненте в модели морфологии английского языка (переход от фонемной транскрипции к орфографической записи) // Проблемы прикладной лингвистики. Тезисы межвузовской конференции 16–19 декабря 1969 г. М.: МГПИИЯ им. М. Тореза, 1969. С. 351–355.
195. О способах выражения семантических параметров в английском языке // Машинный перевод и прикладная лингвистика. Вып. 11. М.: МГПИИЯ им. М. Тореза, 1969. С. 106–121.
196. О некоторых вопросах, связанных с составлением семантического словаря // Третья межвузовская студенческая конференция по проблемам структурной и прикладной лингвистики. Тезисы докладов. 20–25 апреля 1967 г. М.: МГУ, 1967. С. 41–43.

### Рецензии

1. Рец. на кн.: Холодович А.А. Проблемы грамматической теории // Вопросы языкознания. 1982, № 1. С. 121–125.
2. Review of: Y. Brandt Corstius. Exercises in Computational Linguistics / Mathematisch Centrum, Amsterdam. Mathematical Centre Tracts 30. Amsterdam, 1970 // Linguistics. An International Review. 129. June 1, 1974. P. 87–96. (соавтор: Mel'chuk I.)
3. Review of: R.L. Venezky. The structure of English Orthography / Janua Linguarum, series minor, 82. The Hague, Mouton, 1970 // Linguistics. An international review. 84. June 1972. P. 89–100. (соавтор: Shevoroshkin V.)

4. Review of: Arne Zettersten. A statistical study of the graphic system of the present-day American English. Lund: Studentlitteratur, 1969 // *Linguistics. An International Review*. 91. October 15, 1972. P. 115–117.
5. Рец. на кн.: И.К. Бельская. Язык человека и машина // *Научно-техническая информация*. 1970, № 10. С. 45–47.
6. Рец. на кн.: Robert P. Stockwell, J. Donald Bowen, John W. Martin. *The Grammatical Structures of English and Spanish*. Chicago and London, 1965 // *Филологические науки*. 1968, № 3. С. 117–118.
7. Рец. на кн.: Robert P. Stockwell and J. Donald Bowen. *The Sounds of English and Spanish*. Chicago and London, 1965 // *Филологические науки*. 1967, № 1. С. 105–107.

### **Научное и ответственное редактирование**

1. Труды востоковедения РАН. Вып. 19. Проблемы общей и востоковедной лингвистики: Лексикология и лексикография / Отв. ред. З.М. Шаляпина. Ред.-сост. А.С. Панина. – М.: ИВ РАН, 2018. – 516 с.
2. Труды востоковедения РАН. Вып. 6. Проблемы общей и востоковедной лингвистики: Языки Южной и Юго-Восточной Азии (материалы научной конференции, ИВ РАН, 26 октября 2016г.) / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина, А.И. Коган. – М.: ИВ РАН, 2018. – 240 с.
3. Востоковедные чтения 2016. Языки Южной и Юго-Восточной Азии. Тезисы докладов Межинститутской научной конференции ФГБУН Институт востоковедения РАН, 26 октября 2016г. / Науч. ред. З.М. Шаляпина, сост. А.С. Панина. М.: ИВ РАН, 2016. – 44 с.
4. Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. К 70-летию члена-корреспондента РАН В.М. Алпатова. Труды научной конференции. Институт востоковедения РАН 22–23 апреля 2015г. В двух томах / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина, А.С. Зверев. – Т. 1. – М.: ИВ РАН, 2016. – 232 с.
5. Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. К 70-летию члена-корреспондента РАН В.М. Алпатова. Труды научной конференции. Институт востоковедения РАН 22–23 апреля 2015г. В двух томах / Отв. ред. З.М. Шаляпина. Сост., ред. А.С. Панина, А.С. Зверев. – Т. 2. М.: ИВ РАН, 2016. – 216 с.
6. Проблемы общей и востоковедной лингвистики 2013. Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Труды научной конференции. Институт востоковедения РАН 20 ноября 2013 г. Отв. редактор З.М. Шаляпина. Сост., ред. А.С. Панина / Институт востоковедения РАН. М.: ИВ РАН, 2014. – 312 с.
7. Два выдающихся востоковеда. К 90-летию со дня рождения И.Ф. Вардуля и Ю.А. Рубинчика. Тезисы докладов научной конференции. Институт востоковедения РАН 20 ноября 2013 г. / Науч. ред. З.М. Шаляпина. Сост., ред. А.С. Панина / Институт востоковедения РАН. М.: ИВ РАН, 2013. – 70 с.

8. Бюллетень Общества востоковедов РАН. Вып. 17: Труды межинститутской научной конференции «Востоковедные чтения 2008»: Москва, 8–10 октября 2008 г. / Ред. коллегия: Л.Б. Алаев, Д.Д. Васильев, М.С. Мейре, В.В. Наумкин (председатель), А.А. Столяров (секретарь), З.М. Шаляпина (научн. ред.). М.: Учреждение Российской академии наук «Институт востоковедения РАН», 2010. – 702 с.
9. Актуальные вопросы японского и общего языкознания. Памяти И.Ф. Вардуля // Отв. ред. В.М. Алпатов, З.М. Шаляпина; сост. А.С. Панина; Российская академия наук. Институт востоковедения. М.: Издательская фирма «Восточная литература» РАН, 2005. – 416с.
10. Японско-русские термины по обработке металла давлением. Сост. А.Л. Семенов. Под ред. к.ф.н. З.М. Шаляпиной / ГКНТ и АН СССР. ВЦП научно-техн. литературы и документации. Тетради новых терминов. № 106. – М.: ВЦП, 1987. – 43 с.
11. Актуальные вопросы практической реализации систем автоматического перевода. Материалы первого совместного советско-французского семинара, состоявшегося в Москве в 1977 г. Ч. 1. Отв. ред. В.М. Андрющенко, науч. ред. З.М. Шаляпина. М.: МГУ, 1982. – 240 с.
12. Актуальные вопросы практической реализации систем автоматического перевода. Материалы первого совместного советско-французского семинара, состоявшегося в Москве в 1977 г. Ч. 2. Отв. ред. В.М. Андрющенко, науч. ред. З.М. Шаляпина. М.: МГУ, 1982. – 300 с.

### **Библиографическая справка**

1. Шаляпина З.М. // Милибанд С.Д. Востоковеды России: XX – начало XXI в. Библиографический словарь: в 2 кн. Институт востоковедения РАН. М.: Восточная литература, 2008. Кн. II: Н–Я. С. 639.
2. С.И. Гиндин. Язык, текст, перевод, увиденные по-своему // Русский язык. Методическая газета учителей-словесников. М.: Издательский дом "Первое сентября", апрель (1–15) 2006, №7 (486). С. 44.
3. Shalyapina Zoya M. // Dictionary of International Biography. A biographical record of contemporary achievement / International Biography Center. Twenty-fifth edition. 1997. Cambridge (England): Melrose Press Ltd., 1997. P. 355.

**Типология средств модификации содержания сообщения  
(на материале русских и китайских частиц)**

**Елена Георгиевна Борисова**

Московский городской педагогический университет, Москва, Россия,  
ORCID 0000-0003-3878-5344

**Софья Андреевна Маннапова**

Московский городской педагогический университет, Москва, Россия,  
sonyachita14@mail.ru, ORCID 0000-0002-5314-6009

***Аннотация.*** В статье рассматриваются принципы типологии служебных единиц, используемых для модификации значений высказываний в процессе коммуникации. На основании данных русского и китайского языков определяются свойства единиц, используемых говорящим для управления пониманием сообщения со стороны слушающего: привлечение внимания к части сообщения («усиление», «выделение»), маркировка степени известности и связанных с ней импликатур (значения причины, следствия, уступки) и ряд других. Разрабатывается набор признаков, на основании которых можно было бы проводить анализ соответствующих служебных единиц (усилительных или модальных частиц, дискурсивных слов и других классов) с целью их сравнения в масштабе знаний о всех языках и создания типологической классификации.

***Ключевые слова:*** управление пониманием, место в высказывании, топикализация, коннотативное значение, частица

***Abstract.*** The paper concerns the principles of typological classification of functional lexemes modifying the content of utterances according to the needs of communication. Data from Russian and Chinese is used to reveal the qualities of the entities used by the Speaker to govern the Hearer's understanding: to attract attention of the Hearer to certain parts of the utterance (underlining, emphasizing), to mark the degree of the Hearer's awareness about something, possibly accompanied by implicatures (cause, consequence, concession) etc. The authors propose a set of features that can be used for analyzing and comparing the functional words (modal particles, discourse markers, adverbials etc.) in all languages with the prospects of developing typological classification.

***Keywords:*** governing understanding, syntactic position, topicalization, connotation, implicature, particle

## **1. Типология и место в ней служебных слов**

Изучение языка – это не только описание тех или иных явлений в многочисленных языках народов мира, но и создание обобщающих моделей и теорий, какой, к примеру, является пока еще не вполне оцененная Трехмерная модель языка З. М. Шаляпиной. Между конкретными описаниями и обобщающими моделями

можно разместить типологические исследования, позволяющие систематизировать конкретные факты в языках мира и выявить имеющиеся в них закономерности и тенденции, которые могут быть положены в основу какой-либо гипотезы или модели устройства языка.

Типология развивается уже около полутора столетий. Можно отметить значительный прогресс в этой сфере знаний, достигнутый исследователями разных стран (А. Е. Кибрик, Дж. Гринберг), в том числе и питомцами Института востоковедения М. Копчевской-Тамм, А. Айхенвальд, В. И. Подлесской и многими другими.

Объектом типологии оказываются явления фонетики и фонологии, грамматики (морфологии в плане содержания категорий и поведения морфем, а также синтаксиса). В последнее время активизируются усилия по созданию разных видов семантической типологии, начало которой положено ранними работами Б. Ю. Городецкого: типология семантической деривации (Е. В. Рахилина, Г. И. Кустова, Анна А. Зализняк), особенно метафоризации (А. Н. Баранов). В то же время различные языковые явления, имеющие прагматические корни, мало представлены в типологических исследованиях.

## 2. Как модифицируется понимание сообщения

Целью нашего исследования является рассмотрение аспектов, которые можно было бы положить в основу семантической типологии служебных единиц, в первую очередь – слов-модификаторов содержания сообщения [Борисова 2018]. К таковым можно отнести разряд усилительных частиц в русском языке, модальных частиц в немецком, фразовых частиц в китайском. В то же время, к примеру, в английском языке сюда входят не только довольно редкие в этом языке модальные частицы, но и более близкие к значащим частям речи адвербиалы типа *in fact*, *actually* т.п. Наконец, модификация понимания может осуществляться и с помощью междометий, союзов, что приводит к появлению у них значений, близких к значениям частиц. В русском языке это союзы *и*, *а*, *но*, междометие *ну*.

В каком-то смысле эта же модификация может осуществляться и полнозначными словами, и целыми высказываниями. Пример из русского языка:

(1) а. *Заявляю с полной ответственностью: больше я это прощать не намерена.*

vs.

б. *Я-таки не намерена больше прощать.*

*Да я тебе этого больше прощать не стану!*

Примеры из китайского языка:

(2) а. 我请你帮助他。 vs. б. 你帮助他吧。  
*Wǒ qǐng nǐ bāngzhū tā.* *Nǐ bāngzhū tā ba.*  
«Я прошу тебя ему помочь». «Помоги ему [частица]».

(3) а. 我禁止你去。 vs. б. 你别去了。  
*Wǒ jìnzhǐ nǐ qù.* *Nǐ bié qù le.*  
«Я запрещаю тебе идти». «Не ходи [частица]».



Во многих языках средства модификации высказываний еще не получили достаточно полного описания. Даже для европейских языков чаще всего речь идет о простом перечислении этих единиц, возможно, еще о дополнительных (оценочных, эмоциональных) оттенках. В этом случае типологическое сравнение будет сведено к маркировке таких оттенков. Более тщательные описания частиц-модификаторов, массово появляющиеся в наши дни, позволяют отметить ряд параметров, которые, будучи выявленными для пары языков (в настоящем исследовании для русского и китайского), в дальнейшем могут быть применены и к единицам других языков.

### 3. К чему относятся модификаторы

Одной из важных характеристик модифицирующих единиц является их сфера действия: к чему могут относиться частицы в синтаксическом и семантическом плане [Богуславский 1996]. В пределах одного языка можно найти частицы, которые способны модифицировать сообщения одного типа, но не способны участвовать в модификации сообщений другого типа. Например, в русском языке утверждения модифицируются почти всеми усилительными частицами, а обращение – очень небольшим числом частиц: *же, ну, редко да, -таки*:

(4) *Вольха! Ну Вольха же!* (настоятельное обращение)

Схожая ситуация наблюдается и в китайском языке, где обращение маркируется преимущественно только частицей *啊 a* и ее фонетическими разновидностями:

(5) 儿子啊, 有些事情我得跟你说清楚啊。

*Erzi a, yǒu xiē shìqíng wǒ děi gēn nǐ shuōqīngchū a.*

«Сын [частица], есть вещи, которые мне нужно тебе объяснить».

Что касается позиции частиц в предложениях, то здесь тоже бывают различные варианты. Так, в русском языке характерной оказывается позиция в начале предложения (*a, da*), а также после слова, к которому относится частица (*же, -то, -таки*). В китайском языке распространенной позицией оказывается конечная: это конец предложения (*啊 a, 吧 ba, 呢 ne, 吗 ma, 的 de, 了 le*) или конец части предложения (*啊 a, 吧 ba, 呢 ne*):

(6) 现在还早呢。

*Xiànzài hái zǎo ne.*

«Сейчас еще рано [частица]».

(7) 这个人总爱说漂亮话, 实际上呢, 不做一点扎扎实实的工作。

*Zhègè rén zǒng ài shuō piàoliang huà, shíjìshàng ne, bù zuò yídiǎnr zhāzhā shíshí de gōngzuò.*

«Этот человек всегда любит разглагольствовать, а на самом деле [частица] ничего тщательно не делает».

Для русского языка конечная позиция не характерна, разве что она вызвана соответствующим расположением слова, к которому частица относится:

(8) *Ну как? – Пришел-таки.*

Сфера действия, равно как и позиция в предложении, может быть связана не с синтаксической организацией, а с актуальным (тема-рематическим) членением высказывания. Некоторые частицы русского языка маркируют только ремю (-же, и), некоторые особенно выделяют тему (-то). Китайские частицы, способные занимать позицию в середине предложения (啊 *a*, 吧 *ba* и 呢 *ne*), маркируют только предцизируемую часть сообщения, иногда совпадающую с темой, выполняя при этом также функцию привлечения внимания собеседника:

- (9) 这个事儿啊, 不是你想得那么简单。  
*Zhègè shìr a, bù shì nǐ xiǎngde nàme jiǎndān.*  
«Это дело [частица] не такое простое, как ты думаешь».
- (10) 我爸爸吧, 特别喜欢足球。  
*Wǒ bàba ba, tèbié xǐhuan zúqiú.*  
«Мой отец [частица] очень любит футбол».

Заметим, что одной из первых работ по частицам в неевропейском языке (а именно языке сомали) была работа А. К. Жолковского [Жолковский 1971], в которой описывалась частица, маркирующая ремю.

#### 4. Что могут модифицировать частицы

Частицы в русском языке и, по-видимому, в других языках вносят в содержание высказывания различные модальные оттенки, заставляющие слушателя несколько иначе понимать сообщение с частицей, по сравнению с сообщением без нее. Так, русские частицы способны сделать побуждение более настоятельным или даже категоричным:

- (11) *Отвечай же!*
- (12) *Да не подходи к нему.*

Аналогично, в китайском языке частица 了 часто используется совместно с лексической единицей 别 *bié* (*не*) для выражения категоричного запрета:

- (13) 千万别看了。  
*Qiānwàn bié kàn le.*  
«Ни в коем случае не смотри [частица]».

Интенсификация побуждения может осуществляться рядом немецких модальных частиц, которые в целом демонстрируют схожие функции:

- (14) *Komm doch.* «Иди же».

В то же время некоторые частицы русского языка могут превращать побуждение в снисходительное разрешение:

- (15) *Ладно, бери уж, раз просишь.*

В китайском языке тоже распространены варианты смягчения побуждения с помощью частиц 吧 *ba* и 啊 *a*:

(16)а. 你先坐一会儿吧。  
*Nǐ xiān zuò yíhuìr ba.*  
«Посидите пока немного [частица]»;

б. 快请客人来呀!  
*Kuài qǐng kèrén lái ya!*  
«Поскорей же пригласи гостя войти!»

Частицы осуществляют модификацию и в рамках вопросительных предложений. Некоторые частицы способны сделать вопрос более настоятельным, например:

(17) *Зачем же ты туда полез?*

В китайском языке также есть частицы, которые некоторым образом модифицируют вопрос, однако, как правило, не усиливают, а напротив – смягчают его интонацию:

(18)а. 这到底是龙还是猫啊?  
*Zhè dàodǐ shì long hái shì māo a?*  
«Это, в конце концов, дракон или кошка [частица]»?

б. 回家，你去哪儿呢?  
*Wǒ huí jiā, nǐ qù nǎr ne?*  
«Я возвращаюсь домой, а ты куда пойдешь [частица]»?

Иногда частица вносит существенную модификацию в вопросительное предложение, например, превращает вопрос в риторический:

(19) *И зачем я его не послушал?*

В китайском языке есть частицы, которые участвуют непосредственно в формировании определенного типа вопроса, т.е. без таких частиц данный тип вопроса в принципе не может быть образован. Например, основная функция частицы 吗 *ma* – образование общего вопроса:

(20) 你是中国人吗?  
*Nǐ shì Zhōngguó rén ma?*  
«Ты китаец [частица]?»

Частица 吧 *ba* служит для создания уточняющего вопроса, когда говорящий почти полностью уверен в положительном ответе слушающего, например (ср. с 吗 *ma*):

(21) 你是中国人吧?  
*Nǐ shì Zhōngguó rén ba?*  
«Ты, должно быть, китаец [частица]»?

В повествовательных предложениях русские частицы выполняют различные функции, чаще всего подводимые в словарях под понятие усиления, которое может интерпретироваться как «привлечение внимания»:

(22) *А нам зарплату повысили.*

В некоторых случаях эта же идея интерпретируется как категоричность утверждения:

(23) *Да он же неграмотный.*

Аналогично, в китайском языке уровень категоричности утверждения повышается с помощью частиц 的 *de* и 了 *le*:

(24) а. 你的病一定会好的。

*Nǐde bìng yīdìng huì hǎo de.*

«Тебе обязательно станет лучше [частица]».

б. 我不去了。

*Wǒ bù qù le.*

«Я не пойду [частица]».

В то же время в некоторых случаях добавление частиц, напротив, делает утверждение менее настоятельным и в этом смысле более приемлемым для кооперативного общения:

(25) *Я тут кое-что набросал.*

Добавление частицы показывает, что говорящий не настаивает на вовлечении собеседника в беседу [Борисова, Овчинникова 2005].

Многие китайские фразовые частицы также выполняют функцию смягчения утверждения, например, частица 啊 *a* (в фонетической и графической модификации 哇 *wa*):

(26) 今天天气多好哇。

*Jīntiān tiānqì duōhǎo wa.*

«Сегодня такая хорошая погода [частица]».

Очень ярко разница в функциях между обозначенными двумя группами частиц (повышающими и снижающими уровень категоричности утверждения) прослеживается в китайской реплике согласия: 好啊/吧/了的。 *Hǎo a/ba/le/de*. «Хорошо [частица]». Согласие с частицами 了 *le* и 的 *de* расценивается как твердое и уверенное, а согласие с частицами 吧 *ba* и 啊 *a* звучит значительно мягче. При этом частица 吧 *ba* делает согласие неуверенным (как будто говорящий соглашается нехотя), а частица 啊 *a* помимо прочего еще подчеркивает положительный модус (говорящий соглашается с радостью).

В целом функции частиц более или менее совпадают в разных языках, хотя существует немало особенностей, нюансов в рамках выполнения частицами своих основных функций в каждом языке.

## 5. Семантические аспекты частиц

Частицы и другие служебные слова иногда описываются как единицы без собственного лексического значения, способные только вносить изменения в понимание высказываний. Однако невозможность в некоторых случаях замены одной частицы на другую, выполняющую ту же функцию, показывает, что они не

синонимичны. Кроме того, в пользу признания за частицами лексического значения говорит и тот факт, что можно проследить семантическую деривацию и внутри семантики служебной лексики, и в смысле обращения к этимологическим (микродиахроническим) связям между частицами и другими словами (союзами, междометиями, местоимениями), от которых они произошли.

Представляется, что в семантике каждой частицы выделяется общая часть, с которой связаны другие, частные значения. Так, для частицы *же*, имеющей индоевропейские корни уже в статусе частицы, отмечается [Храковский, Володин 1982] общее значение известности слушающему, из чего можно вывести другие ее значения, в том числе и отмеченные выше. Другие русские частицы имеют семантические связи со словами современного языка: частица *и* – с союзом *и*, общим значением которого является добавление информации; частица *-то* – с местоимением *то* и т.п. Для некоторых китайских частиц (啊 *a* и 吧 *ba*) можно отметить интересубъективную функцию<sup>1</sup>, т.е. функцию ориентации на слушающего, которая тесно связана с категорией вежливости.

В аспекте типологического описания семантики частиц следует обратить внимание на способности частиц маркировать синтаксические отношения на основании своих основных значений. Так, в русском языке частицы *-то*, *же*, *ведь* могут обозначать причину, частица *-то* – уступку:

(27) *Надень шарф – на улице-то минус.* – причина;

(28) *Съешь-то он съест, да кто ж ему даст!* – уступка.

В китайском языке частицы 啊 *a* и 吧 *ba* маркируют первую часть условных предложений, однако эта функция, вероятно, будет связана не столько с основной интересубъективной функцией этих частиц, сколько прежде всего с их способностью маркировать преддицируемый компонент высказывания (см. выше):

(29) 要是我呀, 就给他提意见。

*Yàoshi wǒ ya, jiù gěi tā tí yìjiàn.*

«Если бы это был я [частица], то я бы высказал свою точку зрения».

(30) 去吧, 得花很多时间; 不去吧, 又不太礼貌。

*Qù ba, děi huā hěn duō shíjiān; bù qù ba, yòu bù tài lǐmào.*

«Если пойду [частица], потрачу много времени; если не пойду [частица], это тоже будет не очень вежливо».

Следует отметить, что позиция китайской частицы в середине предложения характеризуется практически полной утратой частицей своего модального значения и ее использованием исключительно в служебной функции (для разметки пауз, оформления определенного типа предложения и т.п.) [Курдюмов 2014, с.39], то есть можно говорить о процессе грамматикализации китайских модальных частиц.

Таким образом, семантика частиц – модификаторов смысла тоже может использоваться при создании типологического описания подобных единиц в языках мира.

<sup>1</sup> Более подробно об интересубъективности и интересубъективной функции см. [吴祥福 2004; 王敏, 杨坤 2010; 饶安芳 2013; 张博宇 2015].

## Выводы

Мы рассмотрели модификаторы значений высказываний (частицы) в двух языках: русском и китайском. Данное сопоставление (с привлечением в необходимых случаях данных немецкого языка) позволяет сделать вывод о возможности широкого сравнения схожих по функциям единиц в языках мира с перспективой создания типологического описания этих единиц. Основаниями для сравнения могут быть формальные параметры: употребление в высказываниях разных типов (повествовательных, вопросительных, повелительных), расположение в предложении (в начале, в конце, в середине), включение в сферу действия синтаксических единиц или единиц тема-рематического членения. Кроме того, важным оказывается такой параметр, как модификация частицами значений высказываний и оттенки этой модификации (категоричность или смягченность утверждения, настоятельность просьбы и т.п.). Также при разработке типологии частиц представляется возможным учитывать параметр семантической деривации: установление этимологических связей, признание факта о развитии различных значений частицы из одного общего, выведение маркировки синтаксических отношений.

В статье не были затронуты вопросы фонетических характеристик рассматриваемых единиц, однако и здесь тоже намечаются возможности сравнения (безударность и т.п.).

Представляется, что создание типологического описания модификаторов значения может играть важную роль для типологии языков, поскольку в этом случае описываются единицы, участвующие в некоторых речевых действиях, отличающихся от действий с полными единицами. Это управление вниманием, пониманием сообщения [Борисова 2018], модификация аффективного состояния собеседника и ряд других характеристик. Внимание к подобным действиям и средствам, используемым при их осуществлении, в языкознании то возникает, то затухает. Однако развитие когнитивной науки показывает важность подобных явлений для более полного описания языка и речи.

## Источники

Су Ж., Ван Л., Старостина С. П. *Практический курс китайского языка. Разговорный практикум*: учеб. изд. М.: АСТ: Восток-Запад, 2007. 219 с.

裸婚时代 (Naked Wedding). Телесериал. / 总导演. 滕华涛, 导演. 曹盾; 总. 龚立波, 李也平等; 领衔主演: 文章, 姚笛; 联合主演: 凯丽, 丁嘉丽等. 北京: 北京光彩世纪文化艺术有限公司. 2011. DVD-ROM.

实用对外汉语教学语法 (Practical Chinese Grammar for Foreigners). / 陆庆和著. 北京: 北京大学出版社, 2006. 501 页.

实用汉语语法 (Practical Chinese Grammar) / 房玉清著. 3 版. 北京: 北京语言大学出版社, 2009. 513 页.

实用现代汉语语法 (Modern Chinese Grammar Practice) / 刘月华等著. 增订本. 北京: 商务印书馆, 2001. 1005 页.

他的国 (His Country) / 韩寒著. 沈阳: 万卷出版公司, 2009. 213 页.

## Литература

1. Богуславский И. М. *Сфера действия лексических единиц*. М.: Языки русской культуры, 1996. 468 с.
2. Борисова Е. Г. Управление пониманием. Языковые единицы, регулирующие понимание сообщения // *Вестник Московского университета*. Сер. 9. Филология. 2018. № 6. С. 34–50.
3. Борисова Е. Г., Овчинникова Т. Е. Пространства усиления // *Диалог. Компьютерная лингвистика и интеллектуальные технологии. Труды международной конференции «Диалог-2005»*. М.: Наука, 2005. С. 559–564.
4. Жолковский А. К. *Синтаксис сомали. Глубинные и поверхностные структуры*. М.: Наука, 1971. 266 с.
5. Курдюмов В. А. Морфологический уровень китайского языка как изолирующего топикового. Сущность и содержание позиции частицы // *Вестник Московского городского педагогического университета. Серия «Филология. Теория языка. Языковое образование»*. 2014. № 1 (13). С. 37–46.
6. Храковский В. С., Володин А. И. Опыт анализа семантико-синтаксических свойств усилительной частицы же (ж) в императивных конструкциях // *Семантика служебных слов* / Левицкий Ю. А. (ред). Пермь: ПермГУ, 1982. С. 23–33.
7. 饶安芳. 汉语人称指示语中的话语交互主观性 (Discourse Intersubjectivity of Chinese Personal Deixis) // *安顺学院学报*. 2013. Vol. 15. № 2. Pp. 55–57.
8. 王敏, 杨坤. 交互主观性及其在话语中的体现 (Intersubjectivity and its Realization in Discourse) // *外语学刊*. 2010. № 1. Pp. 48–51.
9. 吴祥福. 近年来语法化研究的进展 (Progress in Recent Studies on Grammaticalization) // *外语教学与研究 (外国语文双月刊)*. 2004. Vol. 36. № 1. Pp. 18–24.
10. 张博宇. 话语标记语的主观性与交互主观性探析 (Study on Subjectivity and Intersubjectivity of Discourse Markers) // *外语学刊*. 2015. № 3. Pp. 79–83.

## Контекстная валентность по [Шаляпина 2007] и несобственные участники прилагательных

**Павел Валерьевич Гращенко**

Московский государственный университет, Москва, Россия,

Институт востоковедения РАН, Москва, Россия,

pavelvalgra@gmail.com, ORCID 0000-0001-9754-2452

**Аннотация.** В статье исследовано явление контекстной валентности, состоящее в том, что у ряда лексем могут образовываться синтаксические зависимые, отсутствующие в лексической структуре. Данному явлению уделено пристальное внимание в монографии З. М. Шаляпиной, [Шаляпина 2007]. Одна из ситуаций, типичных для появления контекстной валентности в русском языке, – возникновение предложных групп с *для*. Как правило подобные случаи связаны с передачей дополнительного значения ‘достижения определенного стандарта’. Предлагается подход, при котором участники фактически проецируются в синтаксис элементами, *de facto* задающими ситуацию сравнения.

**Ключевые слова:** прилагательные, русский язык, контекстная валентность, оценочная структура, степень проявления признака

**Abstract.** The paper studies the phenomenon of contextual valency. Context valency takes place in case when a lexeme can join a syntactic item that is absent in its lexical structure. This phenomenon has received close attention in the monograph of Z.M.Shalyapina, (Shalyapina 2007). One of cases when a contextual valency appears in Russian is the occurrence of prepositional phrases with ‘for’ (*dlya*). As a rule, such cases are associated with an additional value ‘a certain standard reached’. It is proposed that heads relevant for the comparison projects additional participants into syntactic structure.

**Keywords:** adjectives, Russian, context valency, comparison structure, degree

### 1. Влияние интенсификаторов на свойства зависимых прилагательных

Семантическая (гипер)роль бенефактива у прилагательных обладает одной интересной особенностью. Она может проявляться у семантических классов прилагательных, не обладающих способностью к управлению. Так, у прилагательных классов РАЗМЕР, ФИЗИЧЕСКИЕ СВОЙСТВА, ЧЕРТЫ ХАРАКТЕРА И СОСТОЯНИЯ ЧЕЛОВЕКА, которые обычно не имеют синтаксических участников, такие участники появляются при употреблении с ними интенсификатора (*слишком, очень, весьма* и т.д.):

- (1) ??*вход, узкий для коляски*  
??*высокая для ребенка температура*  
??*вода в ванной горячая для малыша*  
??*умный для старшеклассника юноша*



(2) интернет:

*Вход в заведение слишком узкий для коляски.*

*Вызвать скорую, это очень высокая для ребенка температура..*

*Слышала (от врачей), что если водичка в ванной чересчур горячая для малыша, ...*

*Очень умный для старшеклассника Кэндзи Койсо пытается устранить вызванную его просвещенной глупостью проблему в глобальной сети.*

Еще один контекст, при котором роль, отсутствующая или маргинальная при других – например, атрибутивных – употреблениях, становится (более) грамматичной – предикативное употребление в краткой форме [Арутюнова 1988: 240]:

(3) интернет:

*Дверь осторожно приоткрылась, и лихорадочно-бледный маг еле протиснулся в открывшийся проем, словно тот был узок для него.*

*Эта температура высока для стабильного соединения оцинкованных сталей без повреждения покрытия.*

*Так же возмутилась и свекровь, что вода горяча для малыша...*

*Четырехъядерный процессор M8s Kodi Xbmc Amlogic S812 умен для Android.*

Влияние интенсификаторов на свойства зависимых прилагательных было хорошо исследовано для английского языка, см. [Nissenbaum, Schwarz 2010] и приведенные там ссылки. Первым, кто сделал подобное наблюдение для английского языка, был Р. Джекендофф, см. [Jackendoff 1972], однако он, как и последующие исследователи английских прилагательных, больше говорил о зависимости (свойств) целевых участников (инфинитивов) от присутствия интенсификаторов:

(4) английский

a. *The mattress is thin.*

‘Матрац тонкий’.

b. *\*The mattress is thin to sleep on.*

‘\*Матрац тонкий, чтобы спать на нем’.

c. *The mattress is too thin to sleep on.*

‘Матрац слишком тонкий, чтобы спать на нем’ [Jackendoff 1972: 227].

(5) английский

a. *The football is soft.*

‘Мяч мягкий’.

b. *\*The football is soft to kick.*

‘\*Мяч мягкий, чтобы ударить.’

c. *The football is soft enough to kick.*

‘Мяч достаточно мягкий, чтобы ударить по нему’ [там же].

Как можно видеть по русским переводам, такая же зависимость целевых инфинитивов от интенсификаторов наблюдается и в русском языке.

## 2. Контекстная валентность по [Шаляпина 2007]

Для объяснения причин возникновения подобных синтаксических участников, не входящих явным образом в число участников семантических, мы прибегаем к понятию «контекстной валентности» по [Шаляпина 2007].

В [Шаляпина 2007: 177] вводится и получает детальное описание феномен так называемой «косвенной реализации» или «делегирования» валентностей. Данное явление определяется следующим образом: «При косвенной реализации валентности исходный носитель валентности **передает (делегировует)** ее одной из сущностей, с которыми он в данном контексте связан (непосредственно или опосредованно) отношениями зависимостей, а та **займствует** эту его валентность для реализации в качестве одной из своих собственных валентностей».

По наблюдению З.М. Шаляпиной, одним из случаев, где регулярно встречаются контекстные валентности, является группа прилагательного: «в той же роли часто выступают единицы, обозначающие оценочную характеристику определяемого субъекта по тому или иному параметру: *он смел / решителен / осторожен / <в поступках> / удачлив <в любви> / высок <ростом> / красив <лицом> / мягок <по характеру>*» [Шаляпина 2007: 182].

Очевидно, что оценочная структура, присутствующая в неуправляющих прилагательных на семантическом уровне, получает синтаксическое выражение благодаря введению интенсификаторов. В терминах конфигурационного подхода наличие интенсификатора может свидетельствовать о существовании специальной функциональной проекции, адъюнктом которой является интенсификатор. Та же самая проекция ответственна и за введение комплемента с предлогом *для*.

Похожий эффект наблюдается в случае именной проекции – в ряде контекстов у некоторых существительных посессор получает дативное оформление, см. [Grashchenkov, Markman 2010]:

- (6) *Дима был мне другом.  
Она была мне сестрой.*

Дативное управление или, альтернативно, предложная группа с *для* (*Он был для меня другом*) появляется лишь в определенном (предикативном) контексте, ср.: *\*Он обидел друга Диме* и т.д. Похожие участники – как целевой стандарт, так и целевой инфинитив – «наводятся» также и модификаторами в глагольной группе:

- (4) интернет:  
а. *Хорошо стреляешь для художника, прямо в сердце.*  
б. *Сегодня я слишком устал, чтобы продолжать путь...*
- (8) а. *\*Стреляешь для художника, прямо в сердце.*  
б. *\*Я устал, чтобы продолжать путь...*

## 3. Интенсификация и структура группы аппликатива

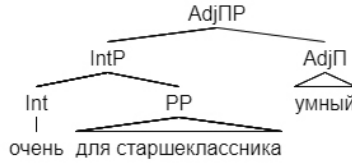
Начиная с диссертации [Pylkkänen 2002] принято считать, что подобные участники вводятся в структуру специальной аппликативной проекцией, см. также применения данного подхода к русскому материалу в [Пшехотская 2012]. Как нам представляется, в случае стативных предикатов из-за отсутствия фазы, соответст-

вующей каузативной и процессной стадиям, бенефактив переосмысливается как экспериенцер / целевой стандарт.

Можно предположить, что за целевыми стандартами стоят два типа структур.

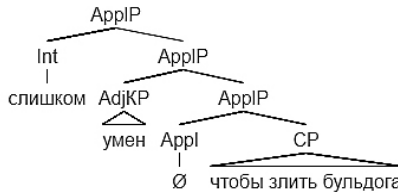
В случае атрибутивного употребления групп прилагательного типа *слишком умный для старшеклассника юноша* можно применить анализ, подходящий также и для аналитического компаратива. Модификаторы типа *очень*, *слишком* и т.д., как и сравнительные модификаторы типа *более*, обладают способностью привносить собственные аргументы, ср. *\*(очень) умный для старшеклассника юноша*, *\*(более) трудный, чем математика, предмет*. Мы будем анализировать вершины *очень*, *слишком* и проч. как адьюнкты к ПФ прилагательного. В позиции компонента возглавляемых ими проекций находятся предложные группы с *для*:

(9)



В случае предикативного употребления с сентенциальным компонентом *чтобы* можно предположить, что над группой прилагательного надстраивается аппликативная проекция:

(10)



Такая конфигурация, т.е. расположение экспериенцера / целевого стандарта / целевого инфинитива в позиции компонента, а группы прилагательного – в спецификаторе, подтверждается двумя фактами. Во-первых, имеющиеся у прилагательных компоненты находятся ближе к нему, чем целевой участник (также занимающий позицию компонента, но более высокой проекции):

(11) интернет:

*Не обращай внимания на моего брата, он слишком [зол на мир], чтобы говорить с другими нормально!*

*Старый добрый Клуб Писателей оказался слишком [стар и слишком добр к себе], чтобы принять такие нововведения.*

*... в «Реале» слишком [довольны Навасом и Касильей], чтобы давать 50 миллионов за человека, который через год достанется бесплатно.*

*... этот проект слишком [дорог нам], чтобы бросать все на полпути.*

В примерах (11) проекции группы прилагательного с имеющимися у них лексическими зависимыми находятся выше в структуре, чем *чтобы*-инфинитивы, как мы предполагаем – в позиции спецификатора проекции AppIP.

Во-вторых, лексический комплемент прилагательного, даже если он выражен косвеннопадежной (а не предложной) группой, не си-командует *чтобы*-клаузой. Если си-командование имеет место, именная группа не может быть кореферентна подлежащему *чтобы*-клаузы, выраженному местоимением:

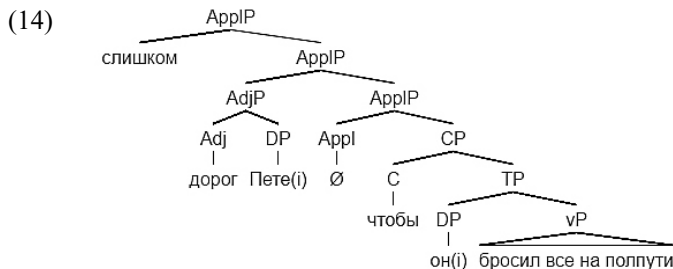
(12) *Други Пети, приехал в гости, чтобы он\*<sub>i,j</sub> немного отдохнул.*

В данном примере местоименное подлежащее может относиться лишь к именной группе *Пети*, т.к. только эта именная группа не си-командует подлежащим *он*. В предложениях ниже подлежащее-прономинал финитной *чтобы*-клаузы может относиться к лексическому комплементу прилагательного:

(13) *В «Реале» слишком довольны Навасом, чтобы он ушел в другой клуб.*

*Этот проект слишком дорог Пете, чтобы он бросил все на полпути.*

Данные факты говорят о том, что ни на каком отрезке деривации между дативным участником и зависимой *чтобы*-клаузой не имеется отношения си-командования. Это соответствует предложенной нами выше структуре<sup>1</sup>, где группа прилагательного находится в спецификаторе аппликативной проекции, а экспериенцер / целевой стандарт – в комплементе:



В [Гращенков 2018] сделано следующее наблюдение: большинство валентных прилагательных обязательно имеют краткую форму, ср.: *безразличен / беспристрастен / благожелателен к чему-либо; идентичен / одинаков с чем-либо; жеснат на ком-либо; интересен / омерзителен чем-либо* и т.д. Это обобщение регулярно нарушается участником *для*:

(15) интернет:

*В этом историческом для нас сезоне, когда вновь откроет двери Основная сцена Большого театра...*

*Это не был классический для нашего футбола случай подсиживания ассистентом...*

<sup>1</sup> Хотя это, естественно, не единственная структура, удовлетворяющая данному условию.

*В столь поздний для Питера час желающих совершать променад не наблюдается.*

*В минувшем сезоне, ставшем дебютным для красно-белых, они набрали 50 очков в 30 матчах.*

У прилагательных с суффиксом *-ск-*, как известно, краткой формы не бывает, то же самое можно сказать и про *поздний*, *дебютный* и некоторые другие лексемы, вполне допускающие управление предложной группой с *для*. Это также свидетельствует о том, что источник управления в таких случаях – не лексическая структура прилагательных. Обе предложенные выше структуры вводят участника с *для* независимо от прилагательного. Одна структура содержит *для*-группу в комплементе интенсификатора, другая – в комплементе проекции Appl.

Подытоживая: экспериенциальный участник, соответствующий субъекту оценки в семантической структуре прилагательных, имеет универсальную оценочную природу.

#### 4. Оценка и участник с *для-Gen*

Исследованиям места прилагательных среди частей речи посвящено большое количество теоретической и типологической литературы, см., например, [Алпатов 1986, 1990], [Dixon 1977, 2004] и др.

Семантически прилагательные примечательны в первую очередь тем, что многие из них (хотя и не все, ср.: *равный*, *мертвый*...) образуют градуальные значения. В отличие от именных предикатов типа ‘быть домом’, признак, задаваемый адъективными предикатами, устроен иначе. В случае ‘быть домом’ мы можем разделить все объекты на дома и не-дома. Значение ‘быть высоким’ не содержит точного указания на то, для каких случаев оно истинно, а для каких – ложно. Упрощая, можно сказать, что часть домов будет подпадать под описание *высокий дом*, часть – точно не будет, и наряду с этим будет часть объектов, относительно которой нельзя точно утверждать, являются они высоким домом или нет.

Для прилагательного, таким образом, характерна ситуация оценки, включающая субъект, параметр и обстоятельства оценки. Один и тот же дом будет высоким для одного и не будет таковым для другого; высокий дом в деревне будет отличаться от высокого дома в городе и т. д. Принято считать, что значение прилагательных содержит не только отсылку к некоторому состоянию, но и переменную, характеризующую степень проявления этого состояния (*degree variable*). Положительная степень прилагательного (*высокий*) при этом оказывается семантически сложнее сравнительной степени (*выше*), т. к. содержит указание на то, что степень проявления задаваемого прилагательным свойства достигла в данной ситуации (и по мнению говорящего) определенного уровня или «стандарта».

Можно заметить, что в примерах выше (*хорошо стреляешь для художника*, *поздний для Питера час*, *в этом историческом для нас сезоне*) исходные лексемы (*стрелять*, *поздний*, *исторический*) не являются оценочными. Предложная группа с *для* задает как раз степень проявления признака (*degree*), релевантную для того, чтобы стало возможным наступление определенной ситуации (факультативно представленной *чтобы*-комплементом).

Можно сказать, что интенсификаторы в подобных конструкциях похожи на показатели компаратива, также способствующие у прилагательных «открытию»

новых валентностей. Семантически компаратив сопоставляет степень проявления некоторого качества (*John is tall*) со стандартом (*than 6 feet*). Логическая форма для *John is taller than 6 feet*, ‘Джон выше 6 футов’, выглядит так [Bhatt, Pancheva 2004: 4]:

(16) [<sub>DegP</sub> -er than 6 feet]<sub>i</sub> John is [<sub>AP t<sub>i</sub></sub> tall]

В ряде языков компаративная конструкция образуется не маркированием прилагательного, а просто добавлением стандарта сравнения в косвеннопадежной форме. Так, например, в дравидийском языке каннада, имеющем самостоятельную категорию прилагательных, сравнительная конструкция образуется присоединением к объекту сравнения показателя датива:

(17) каннада  
*idu adakk-inta dodda mane.*  
этот тот(ДАТ)-СОМП большой дом  
‘Этот дом больше, чем тот’ [Bhat 1994: 26].

Можно предположить, что прилагательные отличает от других частей речи именно оценочная семантика и связанная с ней синтаксическая структура, см. [Beck 2002], [Bhat 1994], [Шрамм 1979], [Вольф 1978]. Попытка образовать аналогичную конструкцию на основе имен или глаголов приводит в каннада к неграмматичности:

(18) каннада  
*\*allig-inta jana illi se:ridda:re.*  
там(ДАТ)-СОМП люди здесь собраться  
‘Больше народу здесь, чем там’ [Bhat 1994: 26].

(19) каннада  
*\*avanu nanag-inta o:d-idda:ne*  
он я(ДАТ)-СОМП прочел  
‘Он прочел больше, чем я’ [там же].

Таким образом, именно введение оценочной семантики, задающей степень проявления признака, позволяет интенсификаторам с одной стороны и компаративам с другой «открывать» дополнительную валентность стандарта сравнения (и факультативно – актуализировать ситуацию проявления стандарта клаузой со *чтобы*).

## 5. Заключение

Понятие контекстной валентности по [Шаляпина 2007] помогает описать частую в случае прилагательных ситуацию, когда в синтаксис вводится участник, не проецируемый синтаксической вершиной непосредственно. Контекстные валентности универсальны. В случае прилагательных (и иногда – других частей речи) они связаны с универсальной для естественного языка оценочной структурой.

Синтаксические явления, связанные с контекстными валентностями, могут быть представлены как в формате зависимостей (собственно валентности), так и в формате функциональных проекций и составляющих.

## Литература

1. Алпатов В.М. О разных подходах к выделению частей речи // *Вопросы языкознания*. 1986. № 4. С. 37–46.
2. Алпатов В.М. Из истории изучения частей речи // Алпатов, В.М. (ред.). *Части речи: теория и типология*. М.: Наука, 1990. С. 6–24.
3. Арутюнова Н.Д. *Типы языковых значений: Оценка. Событие. Факт*. М.: Наука, 1988. 341 с.
4. Вольф Е.М. *Грамматика и семантика прилагательного: На материале иберо-романских языков*. М.: Наука, 1978. 199 с.
5. Гращенков П.В. «Грамматика прилагательного. Типология адъективности и атрибутивности». М.: Языки славянской культуры, 2018. 432 с.
6. Пшехотская Е.А. *Косвенное дополнение как субкатегоризованный и несубкатегоризованный актант (на материале русского языка)*. Дис. ... канд. филол. наук. М.: МГУ, ОТиПЛ, 2012. 151 с.
7. Шаляпина З.М. *Трёхмерная стратификационная модель языка и его функционирования: к общей теории лингвистических моделей*. М.: Восточная литература РАН, 2007. 485 с.
8. Шрамм А.Н. *Очерки по семантике качественных прилагательных: на материале современного русского языка*. Л.: Изд-во Ленингр. гос. ун-та, 1979. 135 с.
9. Beck, D. 2002. *The typology of parts of speech systems: the markedness of adjectives*. New York; London: Routledge. 219 p.
10. Bhat D. N. S. *The adjectival category: Criteria for differentiation and identification*. Amsterdam; Philadelphia: John Benjamins, 1994. 309 p.
11. Dixon R. M. W. Where have all the adjectives gone? // *Studies in language*. 1977. № 1. P. 19–80.
12. Dixon R. M. W. Adjective classes in typological perspective // *Adjective classes. A cross-linguistic typology* / ed. by R.M.W. Dixon and A.Y. Aikhenvald. Oxford University Press, 2004. P. 1–49.
13. Grashchenkov P., Markman V.G. Non-core arguments in verbal and nominal predication: High and low applicatives and possessor raising // *The Proceedings of WCCFL 27*. Los Angeles, CA: UCLA, 2010. P. 185–193.
14. Jackendoff R. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press. 1972. 400 p.
15. Nissenbaum J., Schwarz B. Parasitic degree phrases // *Natural language semantics*. March 2011. Vol. 19, issue 1. P. 1–38.
16. Pykkänen L. *Introducing arguments*. Doctoral dissertation. Cambridge, MA: MIT. 2002. 137 p.
17. Bhatt R., Pancheva R. Late merger of degree clauses // *Linguistic Inquiry*. 2004. Vol. 35. P. 1–46.

## Процессы фразеологизации в синтагматике на разных уровнях языковой системы русского языка

Елена Михайловна Лазуткина

Институт русского языка РАН, Москва, Россия,

lazutkelenaa@yandex.ru, ORCID 0000-0003-0135-8277

**Аннотация.** В статье представлен семантический анализ значимых единиц разных уровней языковой системы с точки зрения функциональной лингвистики. Автор исследует процессы в синтагматике в аспекте действия языковой универсалии «фразеологизация». Морфологические форманты, лексемы, словосочетания и модели предложений рассматриваются как элементы сложившихся или складывающихся парадигматических отношений в русском языке.

**Ключевые слова:** теория языка, семантический анализ, синтагматика, фразеологизация, парадигматика

**Abstract.** The article presents a semantic analysis of significant units of different levels of the language system from the point of view of functional linguistics. The author examines the processes in syntagmatics in the aspect of the action of the linguistic universal "phraseologization". Morphological formants, lexemes, word combinations and sentence models are considered as elements of established or emerging paradigmatic relations in the Russian language.

**Keywords:** language theory, semantic analysis, syntagmatics, phraseologization, paradigmatics

Среди новаторских идей и разработок З.М. Шаляпиной в области синтагматики есть понятие реляции. При реляции роль слова в линейной последовательности языковых элементов выполняет отношение между словами: оно имеет возможность вступать в некусное или юнктивное отношение с другими словами. Зоя Михайловна считала реляцию феноменом психолингвистической природы [Шаляпина 1976: 53-80]. Е.Г. Борисова характеризует введение этого понятия как «выявление дополнительного способа выдвигания в позицию «хозяина» самого отношения между двумя членами словосочетания; например: *производить впечатление* (Oper<sub>1</sub>), *стоит задача* (Func<sub>0</sub>)» [Борисова 2018: 157].

В статье рассматриваются вопросы синтагматики, близкие теоретической концепции З.М. Шаляпиной: граница слова, лексическая законченность слова, грамматическая оформленность слова, членимость слова, состав и членимость словосочетания, композиционная семантика предложения, преобразование информационной структуры модели предложения.

Наша работа строится на положениях современной функциональной лингвистики: «принимается постулат о том, что языковая форма всегда, в конечном итоге, мотивируется смыслом». Основной принцип исследования – «вместо задачи описания языка задача построения объяснительной науки о языке, вместо про-



цедур открытия речевых регулярностей моделирование деятельности носителя языка» [Алпатов 2018: 148-150].

Фразеологизация также имеет психолингвистическую основу. В.В. Виноградов, объясняя механизм образования фразеологизмов, подчеркивает ведущую роль ключевых элементов («опорных частей») в несвободных сочетаниях: они выступают как семантическая доминанта, которая формирует «целостное единство», «синтаксически составные слова», «сложные лексические единицы». Семантика одного из компонентов словосочетания оказывается в тени семантически главенствующего компонента, в то время как этот компонент выполняет роль средства грамматического оформления словосочетания как фрагмента высказывания. «Фразеологические единства как бы поглощают индивидуальность слова, хотя и не лишают его смысла» [Виноградов 1977: 133]. При фразеологизации происходит формирование «смысловых сгустков», которые выполняют функцию слова в процессе линейного развертывания речи.

Такое понимание механизма фразеологизации разделяет М.В. Панов: «Целое не столько определяется частями, сколько определяет их <...> целое сохраняет свою определенность; часть же – именно потому, что она часть фразеологизма – семантически опустошается; целое становится более всех своих частей» [Панов 2004: 67-68].

Фразеологизация проявляет свое действие на разных уровнях языковой системы, вследствие чего считается универсальным свойством языка. См.: «Специфика связного значения неразрывно сопряжена не только с лексическим составом языка, но и грамматическим его строем [Телия 1981: 3]. В.В. Виноградов считал необходимым выделение фразеологии в отдельную лингвистическую дисциплину.

## **I. Границы слова и структурно-семантические отношения частей производного слова**

### **1. Границы слова. «Внешнее» грамматическое оформление асертивного смысла с помощью отрицательной частицы *не*.**

Пробел в письменной фиксации речи не всегда может быть значим для определения границы слова. Выделяя признаки слова, Ф.Ф. Фортунатов подчеркивает содержательную целостность слова: «Отдельным словом является такой комплекс звуков речи, который имеет в языке значение отдельно от других звуков и звуковых комплексов, являющихся словами, и который при этом не разлагается на два или несколько отдельных слов без изменения или без утраты значения хотя бы той или другой части звукового комплекса. <...> Например, звуковой комплекс *неправда* (ложь) представляет одно слово, ...так как будучи разложен на отдельные слова *не* и *правда* теряет данное значение» [Фортунатов 2010: 112].

М.В. Панов также акцентирует такое качество слова, как смысловая неразложимость его состава. Понятие лексической законченности, как замечает М.В. Панов, «вытекает из общенародного, проверенного многолетней практикой взгляда на слова, взгляда эмпирического, который как раз и требует научного обоснования» [Панов 2004: 58].

У некоторых слов отрицательная частица пишется раздельно с глаголом, но является элементом *внутренней формы слова* – в том смысле, который вкладывал

в этот термин А. А. Потебня, т.е. образа значения. Уместно привести слова Э. Бенвениста о том, что смыслообразующим элементом любого высказывания является ассерция: «К грамматической связи, объединяющей члены высказывания, имплицитно добавляется «это есть!», которое устанавливает связь между языковым рядом и системой действительности» [Бенвенист 1974: 170]. Это мнение согласуется с положением когнитивной лингвистики, в соответствии с которым концепт «отрицание» в онтогенезе существует вкупе и во взаимодействии с концептом «утверждение».

Таким образом, графический облик слова и его состав могут быть явлениями разного порядка – традиции правописания и грамматической оформленности лексем; ср.: *Коту недостает ласки; Коту не хватает ласки.*

**а) Лексический комплекс «отрицание + глагол» в конструкциях с ассертивным смыслом.**

Рассмотрим предложения с глаголами, имеющими значение «отрицательно оцениваемый говорящим признак». Эти комплексы показывают высокую степень спаянности глагола с отрицанием *не*.

Следующий пример служит иллюстрацией двух типов сочетания глагола с отрицанием «не» – связанного и свободного: *Старуха его работала санитаркой в больнице, поэтому спирт у них никогда не переводился и никогда не переводился зря* (С. Соколов). Первая часть данного предложения показывает очевидность ассертивного смысла и высокую спаянность комплекса «отрицание + глагол» (*спирт не переводился* – ‘был в наличии’). Вторая часть предложения содержит тот же глагол в свободном сочетании с отрицанием *не*. Частица *не* слабо связана с глаголом, и она играет свою основную роль: отрицает негативный смысл глагола *переводить* – ‘тратить’, формируя положительное значение сочетания *не переводился* – ‘сохранялся’.

Как правило, фразеологизированный комплекс «глагол с отрицанием» обнаруживает морфологическую ригидность: это формы несовершенного вида настоящего времени (реже – прошедшего); например: *Пускай говорит. От его слов тебя не убудет!* (А. Островский); *Улыбка не сходила с его лица; Вчерашний слушай не идет из головы.*

**б) Устойчивые выражения с подчеркнутым положительным модальным компонентом в глаголе;** например: *Он не преминул сказать об ошибках прежнего председателя правления; Они не упускают случая напомнить об обещаниях комиссии.*

**в) Предложения характеристики, имеющие общее значение «утверждение отрицательного предикативного признака как свойства, качества субъекта»;** глаголы-предикаты с низкой степенью членимости и высокой степенью фразеологизации; например: *Время не ждет; Работа не стоит (Работа продолжается); Парень не подкачает; За мной не заржавеет; На небе звезд не счесть* (много).

**г) Предложения, обозначающие вневременное, постоянное свойство субъекта;** например: *Эта ткань не мнется; Сосед не пьет, не курит; Бумага не салилась.*

д) Предложения с отрицательным модальным компонентом в глаголе, характеризующим лицо; например: *Он не водил машину* ('не умел водить'); *Девушка не готовит* ('не умеет готовить еду').

е) Предложения, обозначающие нарушение функции чего-нибудь – какого-нибудь устройства, прибора, его поломку; например: *Машинка не шила*; *Почта не ходит*; *Молния не застегивается*.

Таким образом, отрицательная частица *не*, которая по правилам орфографии пишется раздельно с глаголами, может входить в число средств лексической законченности глагола и его грамматической оформленности, т.е. выполнять роль «внешнего оформления» ассертивного смысла.

Дискурсная норма употребления этих глаголов должна учитываться в лексикографической практике: комплексы с ассертивным смыслом должны считаться самостоятельными значениями некоторых групп полисемантических глаголов, а их словарная статья в толковом словаре должна указывать на морфологические характеристики глагольных форм и допустимые дейктические параметры высказывания.

**2. «Внутреннее» оформление ассертивного смысла. Прилагательные, предикативы и наречия с префиксом *не-*, имеющие качественное значение.** Префикс *не-* теряет отрицательное значение и формирует конкретное качественное значение слова, обуславливает толкование лексемы как словарной единицы; например: *Нет ни в чем вам благодати, С счастьем у вас разлад: И прекрасны вы некстати, И умны вы невпопад* (А.С. Пушкин).

См. также: *невидимый соперник*; *непромокаемый плащ*; *незнакомый посетитель*; *посмотрела неласково*; *мне это неприятно*; *приходил нечасто*; *им этого не доставало*. Ср. слова М.В. Панова об ассертивном смысле словосочетания *неудовлетворительная отметка*: «*Неудовлетворительная* – одно целостное слово <...> оно фразеологично, оно не только отрицает, но и утверждает» [Панов 2004: 81].

**3. Слова, образовавшиеся в результате лексико-синтаксического способа словообразования.**

Этот продуктивный в славянских языках способ словообразования показывает нам давно закрепившиеся в системе слова – образцы фразеологизации лексических комплексов низкой степени членности и высокой степени идиоматичности. При этом одна из частей слова «опустошается» и служит средством грамматического оформления всего образования; например: *тотчас, сейчас, сегодня, впередсмотрящий, сумасшедший*; некоторые фамилии типа *Перебейнос, Небаба, Немадруг*.

**4. Изменение морфемного состава производного слова при фразеологизации.**

В основе анализа семантики производного слова, определения его номинативной сущности лежит постулат о «завершенности» лексического значения и его грамматической оформленности. Соотношение частей производного слова является способом существования его значения, но морфемный состав слова лишь намекает его членение, и изменение его лексического значения ведет к изменению его структуры. «Слова – это смысловые единства, части которых не составляют сво-

бодного сочетания. <...> Следовательно, всякое сложное по своему составу слово есть фразеологизм» [Панов 2004: 68].

Сравним топонимы с суффиксами *-щин(а)/-чин(а)*. Они образуются непосредственно от прилагательных с суф. *-ск-*, а опосредствованно мотивированы существительными, представляющими собой географические или этнические наименования. ср. например: *Брянщина, Владимирщина, Вологодщина, Полтавщина, Смоленщина* и т.д. В слове низкая степень морфемной членимости, суффиксы обуславливают флексию *-а* и вхождение слова в разряд существительных женского рода. Тем не менее, появились административные названия *Сабуровщино, Детчино*, что свидетельствует о потере внутренней формы топонима, о фразеологизации комплекса «корень слова и суффикс», о десемантизации суффикса *-щин(а)/-чин(а)*. Эти номинации потеряли ассоциативную связь с понятием ‘крупный надел земли’, ‘земля отцов’.

Слова переходят в разряд топонимов среднего рода с суффиксом *-ин(о)* с обобщающим словом *село*: *Пущино, Гущино, Галкино*. Сохранение в топонимах флексии *-а*, предсказанной суффиксом *-щин(а) / -чин(а)*, ввело бы их в разряд наименований женского рода с родовым словом *деревня*; например: *Бухловка, Лобня, Сабуровщина, Детчина*.

## II. Фразеологизация в модели словосочетания

Существуют устойчивые схемы словосочетаний, обозначающих эмоциональное состояние человека, в строении которых находит отражение внеязыковая ситуация; например: *Вот и сейчас, на мгновение подняв печальные глаза, она предложила посидеть возле* (В. Личутин). Ср. также: *счастливая улыбка, блаженное лицо, унылые плечи, свирепый взгляд*. Возникновение этих словосочетаний может быть объяснено в традиции функционально-когнитивной лингвистики – как смысловое расширение схемы подчинительной связи слов в результате когерентности (связанности), выраженной в дискурсе информации и невыраженной, подразумеваемой – той, которой владеет каждый носитель языка.

В поговорке *Повинную голову и меч не сечет* словосочетанию *повинная голова*, возможно, предшествовало устойчивое выражение *поклонная голова*. Концепт «вина» получил устойчивую ассоциацию с образом – позой человека со склоненной головой, что способствовало идиоматизации данного словосочетания и возникновению у него символического значения. А.А. Потебня упоминает вариант поговорки, существовавший до фразеологизации: *Поклонную голову и меч не сечет* [Потебня 1990: 195].

Подобные выражения представляют собой синтагматические последовательности, которые соотносятся как минимум с двумя грамматическими основами. В некоторых терминсистемах на основе компрессии выражений различной линейной протяженности формируются аналогичные устойчивые номинативные комплексы; например: *Управление распределенной обработкой сигналов*.

## III. Фразеологизация на уровне модели предложения

Чем выше системный уровень языковой единицы, тем сложнее метод изучения ее семантики. В исследовании синтаксической семантики применяется комп-

лексный метод, включающий дедуктивные построения, синтагматический и парадигматический анализ. Инструментом анализа семантики предложения является понятие синтаксической позиции, которое ввел Т.П. Ломтев. Термин *синтаксическая позиция* обозначает значимое положение определенной формы слова в модели предложения, ее «реляционные свойства» [Ломтев 1960: 15 – 28].

Синтаксическую позицию следует назвать главной единицей смысловой организации предложения, поскольку она делает очевидным сходство семантических признаков форм слов, имеющих разные лексические значения и входящих в разные семантические классы, но которым предопределена системой языка одна и та же функция. Это понятие показывает, что модель предложения связана с миром знания через содержательные структуры языка более высокой степени абстракции, нежели значение слова, – через макросемантические разряды лексики, а значит когнитивные категории, базовые концепты.

Параметры компонентов модели предложения находятся в различном соотношении в аспекте триады «форма – значение – функция». Н.Ю. Шведова справедливо подчеркивает важность семантики формы языковой единицы. Ср.: «Описывая разные способы выражения той или иной семантической категории, нельзя забывать о том, что синонимия никогда не равняется полной идентичности, так как форма обязательно накладывает на общее значение заключенной в ней категории свое собственное грамматическое значение» [Шведова 2005: 263].

### **1. Фразеологизация актантной структуры предложения при деривации декаузации.**

Рассмотрим изменения в следующих диатезах: *Мороз покрывает лужи тонким ледком – Лужи начало покрывать тонким ледком – Тонкий ледок начал покрывать лужи – Лужи начали покрываться тонким ледком.*

Инвариантное значение моделей – их сигнификативное ядро – «локально направленное действие на объект, имеющее своим следствием полноту его охвата (или заполнения)».

При декаузации происходит фразеологизация актантной структуры предложения: семантика словоформы в субъектной позиции соединяет в себе значение субъекта действия и значение средства воздействия: *Мороз покрывает лужи тонким ледком – Тонкий ледок начал покрывать лужи.* Семантика схемы: ‘направленное на объект действие субстанции, способной к самостоятельному перемещению, имеющее своим следствием пространственную смежность с объектом’.

Позиция субъекта в именит. пад. в декаузативной конструкции имеет особые содержательные характеристики – «определенность», «множественность». Она имеет пропозитивное значение и предназначена для макросемантического разряда имен существительных, которые представляют концепт «множество» (множество как континуум). Этой спецификацией компенсируется отсутствие в предложении словоформы со значением средства воздействия.

### **2. Фразеологизация предикатного ядра.**

Процесс фразеологизации рассмотрим на примере модели предложения ‘стихийное изменение состояния субъекта (человека) под воздействием чувств, ощущений и внешних сил’, которая является дериватом декаузативной модели ‘на-

правленное на объект действие движущейся субстанции...’ типа *Тонкий ледок начал покрывать лужи*.

Модель предложения является интерпретативной структурой и отражает особенности мировосприятия носителей русского языка. Чувства, эмоции, ощущения трактуются языковой системой как концепт «субстанция, способная к самостоятельному перемещению». Действие чувств, эмоций, состояний и мыслей на человека подводится под базовый концепт «локально направленное действие, имеющее своим следствием полноту охвата»; например: *Берлиоза охватил необоснованный, но столь сильный страх, что ему захотелось тотчас же бежать с Патриарших без оглядки* (М. Булгаков).

Обратный порядок слов является стилистически нейтральным, свойственным данной модели. В позиции начала предложения одушевленное существительное со значением лица. Это говорит о том, что для говорящих на русском языке важно сделать именно человека темой высказывания, хотя он является объектом воздействия. Причина изменения состояния человека (*страх*), хотя является преддицируемым компонентом, подлежащим, занимает в предложении периферийную позицию. Глагол-предикат семантически опустошается, сохраняет только сему «полнота охвата». Таким образом, подлежащее и сказуемое выступают как единый фразеологизированный комплекс с характеризующим значением в функции преддицирующего члена конструкции. См. также: *От радости ходить по земле охватывает его душу ветерок счастья* (Ю. Бондарев); *Однажды, вернувшись с тяги, Я лег подремать на диван. Разносчик болотной влаги, Меня прознобил туман* (С. Есенин); *Её лицо заливают краска смущения*.

### 3. Фразеологизация состава преддицирующей части в структуре предложения.

Исходная модель деривационной сети предложений с событийными глаголами локально направленного действия имеет семантику ‘целенаправленное действие субъекта, изменяющее состояние объекта и имеющее своим следствием полноту его охвата’; например: *Томатной пастой заправляя борщ, лицо от звуков Моцарта не морщь!* (Е. Евтушенко); *Мне дали пульмановский вагон, загрузили его повидлом, джемом и вареньем...* (П. Проскурин).

В модификации исходной модели и в ее дериватах может происходить перераспределение смысловой нагрузки между компонентами модели как приспособление актантной структуры предложения при сохранении его сигнификативного ядра.

Варианты фразеологизации: нулевой актант или инкорпорированный актант в глагольную лексему; при этом глагол-предикат семантически опустошается, сохраняет только сему «полнота охвата». Семантически главенствующим компонентом при фразеологизации становится значение каузатора. См. разные способы его представления.

а) В качестве нулевого члена (подразумеваемого или обобщенного средства воздействия). Принято считать нуль анафорическим элементом; см. например: *Надо перегородить проход (чем-то). Губы обметало; Луга залило; Дети обрызгали скатерть*. Естественно, важную роль в формировании в глагольном окружении нулевой позиции играет семантика словоформы в вин. пад., которая сужает круг возможных распространителей глагола. Префиксальные глаголы совершен-

ного вида с собой непроизвольного действия, имеющего своим следствием отрицательный результат – пачкание поверхности чего-л., – как правило, употребляются с нулевой словоформой в твор. пад.; например: *вывозить (прост.) платье, закапать письмо, закрапать пол, залить скатерть, замазать руки, замарать, замусолить книжку, заплескать пол.*

б) В качестве инкорпорированного члена (названного глагольной лексемой). *Бегемот отрезал кусок ананаса, посолил его, поперчил* (М. Булгаков); *Дорогу опять асфальтируют; Девушка надушилась; Ветки заиндевели; Дорогу заснежило.*

Подобные способы перераспределения содержательной нагрузки между синтаксическими позициями и видоизменение семантической структуры предложения легко считываются носителями языка как известные формулы линейной организации высказываний.

#### IV. Фразеологизация как усложнение информационной структуры предложения

Когнитивной категории экзистенции, «сущностной» категории, которой Аристотель придавал особый статус, присущ высокий уровень парадигматических отношений. Она находит разнообразные способы своего представления, в том числе в предложениях с акциональными глаголами. Строение бытийных моделей предложений с данными глаголами определяет синтаксическая категория лица: информация о процессах и событиях преподносится в трактовке говорящего как факт, положение дел; например: *В избу набились ребятишки.*

Если предложение с событийным предикатом дает представление об этапах протекания действия-события и о его участниках, то в синтаксическом деривате (в бытийном предложении) нет детализации нюансов события. Таким образом, денотативный план, обозначенный акциональным глаголом, включается в структуру пропозиции, поглощается моделью бытийного предложения. «Когнитивный цикл развивается от восприятия к знанию» [Арутюнова 1988: 112].

«Материальный носитель» бытийного смысла – системно закрепленный обратный порядок слов модели предложения; формы 1-го и 2-го лица глагола не употребляются.

Превращение события в факт – это результат «работы» сознания говорящего, отказ от характеристик реального события («бегство» от времени и пространства) во имя последующего «оперирования» этой данностью, уже получившей его осмысление в определенном стилистическом ключе, в намеченной оценочной палитре, – для создания континуума текста, смыслового расширения текста.

Н.Д. Арутюнова выделяет «семантические составляющие» бытийных предложений, которые она называет «заранее заданным комплексом признаков»: «(1) область бытования (локализатор) – пространственный или временной параметр, (2) показатель существования (глагол), (3) бытующий предмет» [Арутюнова 1999: 742].

Рассмотрим следующие разновидности бытийных предложений с акциональными глаголами.

а) **Конструкции с количественной характеристикой, с накопительным компонентом в глагольном значении.** Модели предложений, содержащие акциональный глагол СВ в форме прош. вр. с результативным значением, существуют в языке в двух ипостасях: с препозицией обстоятельственной словоформы со значением места или без этой словоформы; например: *В дом набежало народу / В дом набежал народ / Набежало народу; Под куртку напозли муравьи / Напозли муравьи.*

б) **Предложения с глаголами явления, возникновения.** В таких предложениях нередко присутствует темпоральный указатель, или он подразумевается; например: *В клубе даётся сеанс одновременной игры в шахматы; На коже делаются пятна; Приближалась довольно скучная пора – Стоял ноябрь уж у двора* (А.С. Пушкин); *В народе удержались (не исчезли) старые обычаи; У него завелись деньги; Куда девалась моя книжка?*

в) **Предложения наличия, положения в пространстве.** В группу входят стативные и процессуальные глаголы. *Над ним довлеет страх; В пейзаже доминируют золотистые цвета.* Глаголы самопроизвольного движения редуцируют сему акциональности и актуализируют сему бытийности; например: ‘быть, течь’ (о текучей массе) *откуда, куда, по чему: С гор движутся грязевые потоки, сель; Течет лавина, лава.*

г) **Соматические речения с очевидным или с подразумеваемым субъектом восприятия.** Бытийные предложения с соматическими речениями обычно обозначают неконтролируемые процессы, происходящие в человеке и с человеком как с физическим телом и вместилищем интеллектуальных и психических процессов. Они сообщают о положении дел, подлежащее и сказуемое выступают как единый комплекс с характеризующим значением. Модус говорящего может сливаться с модусом называемого в предложении субъекта состояния. Ср.: *И только душу захлестнуло Сребристой мглой из-под подков* (А. Блок).

д) **Предложения, организованные глаголами восприятия.** Н.Д. Арутюнова относит их к «пограничным» случаям бытийности: «На границе между бытийными и субъектно-предикатными структурами расположена обширная область предложений с глаголами восприятия и воспринимаемых признаков (*синеть, блестеть*), явления, положения в пространстве и именем неопределенной референции» [Арутюнова 1999: 790]. Ср.: *До деревни докатились раскаты грома; Под елью виднелись лисички; Из-за реки донеслась песня; Тень от меловой горы сюда не достигала, и весь берег заливала луна* (М. Булгаков).

## Вывод

В языке постоянно происходит формализация смыслов. Большую роль в этом процессе играет фразеологизация.

Исследование семантики системных единиц разной линейной протяженности с помощью методологического конструкта «фразеологизация» дает возможность объяснить синтагматику значимых единиц разных языковых уровней, строение моделей предложений, обнаружить сложившиеся и складывающиеся парадигматические системные отношения и выявить способы репрезентации концептов.



Предпочтения «коллективного говорящего», которые обуславливают семантические преобразования, формирование новых номинаций, преобразование моделей предложения отражают особенности интерпретативных процессов в современном русском языке: семантическое развитие системных единиц, в том числе моделей предложения, в сторону появления модальных и оценочных компонентов значения, показывают усложнение информационного плана моделей предложений как прием расширения смыслового пространства текста.

### Литература

1. Алпатов В.М. Структурализм и современная лингвистика // *Вестник Института востоковедения РАН*. 2018. № 5. С.146–150.
2. Арутюнова Н.Д. *Типы языковых значений: Оценка. Событие. Факт*. М.: Наука, 1988. 341 с.
3. Арутюнова Н.Д. Бытийные предложения в современном русском языке // *Арутюнова Н.Д. Язык и мир человека*. М.: Языки русской культуры, 1999. I-XV. С. 737–792.
4. Бенвенист Э. *Общая лингвистика*. М.: Прогресс, 1974. 448 с.
5. Борисова Е.Г. Нексус, юнкция, реляция и частица *и* // *Вестник Института востоковедения РАН*. 2018. № 5. С. 157–159.
6. Виноградов В.В. Основные понятия русской фразеологии как лингвистической дисциплины // Виноградов В.В. *Избранные работы. Лексикология и лексикография*. М.: Наука, 1977. С.118–139.
7. Ломтев Т.П. Об абсолютных и реляционных свойствах синтаксических единиц (О понятии позиции в теории синтаксиса) // *Филологические науки*. М.: 1960. № 4. С. 15–28.
8. Панов М.В. О слове как единице языка // Панов М.В. *Труды по общему языкознанию и русскому языку*. Т.1. М.: Языки славянской культуры, 2004. С. 51-87.
9. Потемкина А.А. *Из записок по русской грамматике*. Т. 1-2. М.: 1958. 536 с.
10. Потемкина А. А. *Теоретическая поэтика*. М.: Высшая школа, 1990. 344 с.
11. Степанов Ю.С. Система языка в связи с семиотикой // Степанов Ю.С. *Язык и Метод. К современной философии языка*. М.: Языки русской культуры, 1998. С.57–60.
12. Шаляпина З.М. К проблеме коммуникативной организации текста и ее отражения в семантической записи // *Вычислительная лингвистика*. М.: 1976. С.53–80.
13. Шведова Н.Ю. Место семантики в описательной грамматике (синтаксис) // Шведова Н.Ю. *Русский язык. Избранные работы* – М.: Языки славянской культуры, 2005. С. 243–263.

**Возникновение новых языковых моделей  
как ответ на коммуникативные вызовы дискурса**

**Тамара Григорьевна Погибенко**

Институт востоковедения РАН, Москва, Россия,  
t.pogibenko@mail.ru, ORCID 0000-0002-6667-1929

***Аннотация.*** Цель данной статьи – показать справедливость тезиса П. Хопера о том, что грамматика является эмерджентным свойством текстов и что структура является эпифеноменальным побочным продуктом дискурса. В статье на материале австроазиатских языков описываются ситуации создания новых языковых моделей, которые возникают как ответ на коммуникативные вызовы дискурса. Приводятся примеры из практики полевой работы с носителями мон-кхмерских языков, показывающие, как с использованием языкового ресурса родного языка они могут создавать языковые модели для решения сложных коммуникативных задач, возникающих в дискурсе. Описывается возникновение моделей для решения коммуникативных задач при переходе к письменной форме языка: параллельное возникновение предложений со сдвигом ремы в древне-кхмерском и никобарском языке кар и создание расщепленных предложений в древне-кхмерском. На примере результативной и счетной конструкций утверждается, что вариативность как в синхронии, так и в диахронии может быть следствием исправления нежелательных побочных эффектов синтаксической компрессии.

***Ключевые слова:*** мон-кхмерские языки, эмерджентная грамматика, возникновение моделей, вариативность, происхождение счетной конструкции

***Abstract.*** The purpose of this paper is to illustrate the idea put forward by P. Hopper who proposed that grammar is an emergent property of texts and that structure is an epiphenomenal by-product of discourse. Situations of model creation are analyzed using Austroasiatic data. The models under consideration are believed to come up as an answer to challenges of discourse. The examples taken from fieldwork practice with native Mon-Khmer speakers show that new models can be created as a solution to demanding tasks of communication making use of language resources. The rise of models at periods of transition to writing is described: that is the parallel rise of sentences with rheme shift in Old Khmer and Kar and the rise of split sentences in Old Khmer. It is maintained that variability of result constructions and numeral constructions both in synchrony and diachrony may be driven by the need to correct the undesirable effects of syntactic compression.

***Keywords:*** Mon-Khmer languages, emergent grammar, model rise, variability, origin of numeral constructions

### **Введение**

Эволюционирование языковых моделей – широко обсуждаемая тема в современной лингвистике последних десятилетий, в рамках которой рассматриваются

сценарии возникновения грамматики и синтаксической сложности. Грамматикализация – еще одно широко представленное в научной литературе направление, в рамках которого сформировалось представление об эмерджентной грамматике – постоянном возникновении новых структур в результате разного рода повторения в дискурсе. Идею о том, что грамматика является эмерджентным свойством дискурса, т.е. создается для обслуживания возникающих потребностей коммуникации, высказал П. Хоппер:

«Хоппер высказал идею, что исследования по грамматикализации ставят под сомнение предположение о существовании некоего априорного грамматического компонента, являющегося предпосылкой для дискурса и условием для коммуникации. Взамен он выдвинул тезис о том, что грамматика является эмерджентным свойством текстов. «Структура» в этом случае является эпифеноменальным побочным продуктом дискурса. Идея «эмерджентной грамматики» нашла аналогии в работах нескольких лингвистов, таких как Дюбуа, Томпсон и Химельман» [цит. по Norper 1996: 231].

В данной статье на материале австроазиатских языков будут рассмотрены ситуации создания и использования новых языковых моделей, которые возникают как ответ на коммуникативные вызовы дискурса. Мы рассмотрим следующие случаи.

1) Речевое поведение информанта, когда перед ним возникает задача описания положения дел с участием референтов, не входящих в фонд знаний исследователя.

2) Ситуации, возникающие в ходе полевой работы, когда информант сталкивается с тем, что в его языке отсутствует модель, при помощи которой можно адекватным образом перевести анкетное предложение, отразив его смысл в полном объеме. Как показывает практика полевой работы, в таких случаях носитель языка может прибегнуть к созданию новой модели с использованием языкового ресурса родного языка.

3) Возникновение новых моделей при переходе к письменной форме языка.

4) Создание моделей, связанных со спецификой семантической структуры предикатов оценки.

5) Возникновение вариантов модели как результат исправления нежелательных побочных эффектов синтаксической компрессии на примере результативной и счетной конструкций.

### **1. Модели для описания перемещенных референтов и положений дел (пахынг и кату)**

Как пишет Т. Гивон, на этапе перехода к описанию перемещенных референтов возникает острая необходимость в пополнении аппарата языковых средств, отслеживающих референты (referent tracking). Создание общего фонда средств обозначения референтов (shared reference) становится насущной задачей [Givon 2009: 319]. Рассмотрим два примера из практики полевой работы. Это примеры речевого поведения информантов, когда перед ними возникает задача описания некоего положения дел с участием референтов, не входящих в фонд знаний исследователя.

дователя. В обоих случаях информанты создают речевые модели, апеллирующие к вымышленному общему фонду знаний коммуникантов<sup>1</sup>.

Так, при переводе предложения «*В этом лесу много зверей*» на язык пахынг<sup>2</sup> информант построил предложение: «*Мы с тобой идем по лесу, я показываю, ты видишь много зверей*». Свое речевое поведение информант объяснил тем, что иначе исследователь не узнает, что в его лесу много зверей. В данном случае информант выбрал режим межличностной коммуникации типа «здесь и сейчас», который, как отмечает Т. Гивон, свойственен коммуникации детей раннего возраста и ранним этапам коммуникации человека «when all referents are equally accessible to all participants on the scene» [Givon 2009: 318].

Похожая ситуация возникла у Н. Костелло при сборе материала по языку кату в Лаосе<sup>3</sup>. При описании ситуаций с участием референтов, заведомо не входящих в фонд знаний лингвиста-исследователя, рассказчица использовала необычную модель речевого поведения: она употребляла клише «*ты знаешь, ты видела*» [Costello 1993].

## 2. Модель с расщепленными группами для выражения нескольких рем в языке ма

Следующий пример является образцом языкового творчества информанта, который попытался создать модель предложения для решения сложной коммуникативной задачи при переводе на язык ма<sup>4</sup> диагностического предложения «*Наконец, на той стороне реки мой сын увидел живого слона*».

- (1) ма    *pràj*                    *sìn*                    *dà*                    *ne*    *sv*                    *dù*                    *nəm*  
           в конце концов    смотреть    сторона    тот    видеть    один    CLF  
           *kəp*                    *klaw*                    *ʔaŋ*    *sv*                    *rəwəs*    *kis*  
           ребенок +            мужчина = сын    1SG    видеть    слон    живой  
           ‘Наконец, на той стороне реки мой сын увидел живого слона’  
           (букв. ‘Наконец увидел на той стороне, увидел его, мой сын увидел живого слона’) (Погибенко, Ефимов, 2020, СКД 181).

Как известно, деление на тему и рему в предложении может включать несколько уровней. В работе [Тестелец 2001:445] рассматривается предложение «*Тонкие стволы берез белели редко и отчетливо*», в котором выделяется первая рема «*белели редко и отчетливо*», в ней далее выделяются две части с разной

<sup>1</sup> Ср. употребление русских вводных слов и выражений *знаешь, знаете (ли), видишь (ли), понимаешь*. В словарях отмечается, что они служат для привлечения внимания, призыва к собеседнику. Однако внутренняя форма этих слов говорит о том, что они так или иначе апеллируют к фонду знаний собеседника, а их употребление связано с модусом введения новой информации в фонд знаний адресата сообщения.

<sup>2</sup> Язык пахынг – один из языков малых народов Вьетнама, входит в семью хмонг-миен (мяо-яо).

<sup>3</sup> Язык кату – один из языков малых народов Вьетнама и Лаоса, входит в катуическую группу мон-кхмерской ветви австроазиатской семьи языков.

<sup>4</sup> Язык ма – один из языков малых народов Вьетнама, входит в бахнарическую группу мон-кхмерской ветви австроазиатской семьи языков.

коммуникативной функцией: «белели» – вторичная тема и «редко и отчетливо» – вторичная рема.

В предложении (1) при переводе на язык ма информант счел необходимым выделить три ремы разных уровней.

тема	рема
<i>мой сын</i>	<i>на той стороне реки увидел живого слона</i>
<i>увидел</i>	<i>1. на той стороне реки</i>
<i>увидел</i>	<i>2. слона</i>
<i>слона</i>	<i>3. живого</i>

Для симметричного размещения трех рем в синтаксической структуре предложения он проделал следующее. Расщепил на три состава результативный комплекс *sin sv* ‘увидеть’, повторив второй компонент: *sin ... sv ... sv*. Расщепил именную группу с классификатором и с модификатором *dù nət rəwəs kɪs* ‘живой (т.е. реальный, настоящий) слон’ на два состава: *dù nət ... rəwəs kɪs*. После каждого компонента расщепленного результативного комплекса разместил именные группы, репрезентирующие ремы 1, 2, 3. В результате получилась следующая структура:

$G_1 + \text{Рема}_1 + G_2 + \text{Рема}_2 + G_2 + \text{Рема}_3$

Очевидным представляется функциональное совершенство этого сложного синтаксического построения. В нем нет никаких нарушений, используется только грамматический ресурс, которым обладает результативная конструкция в языке ма: возможность инкорпорации объектной ИГ, возможность дистантного расположения частей конструкции, лексическая самостоятельность ее компонентов. Это, в свою очередь, возможно благодаря тому, что изолирующие языки слоговые. В таких языках компоненты многосложных единиц легко отделяются, подвергаются синонимической замене и уstraняются.

Полевые материалы не позволяют выяснить, является ли нормой расщепление интродуктивной ИГ в языке ма или это языковое творчество информанта, виртуозно владеющего грамматическим ресурсом родного языка – языка изолирующего типа, где тактические возможности компонентов предложения строго ограничены грамматической значимостью порядка слов.

Такое языковое творчество информанта тем более значимо, что он перевел с диагностического предложения на вьетнамском языке, в котором результативный комплекс не разрывается, и компоненты, репрезентирующие коммуникативные блоки 1, 2, 3, размещены после него.

(2) вьет. *Suối cùng con trai tôi đã nhìn thấy*  
наконец ребенок мужчина 1SG PRF смотреть видеть  
*bên kia sông một con voi sống*  
сторона тот река один CLF слон живой

Можно только предполагать, что если бы язык ма был литературным или если бы коммуникативные задачи подобного рода достаточно часто возникали в дискурсе, то была бы вероятность того, что конструкция, представленная в (1), со временем утвердилась бы в языке. Это предположение вполне обосновано, поскольку модель с расщепленной ИГ отмечается в родственных языках, например, в кхмерском.

В кхмерском языке модель с расщепленной ИГ также может использоваться для коммуникативного выделения нескольких рем. Например, в (3) задействован с некоторыми изменениями тот же ресурс, что и в (1): в расщепленный результативный комплекс *ba:n mɔ:k* ‘получить, раздобыть’ инкорпорируется вершина *chat* ‘зонт’ расщепленной ИГ *chat haek nuj* ‘этот разорванный зонт’. В этом предложении нет повтора  $\Gamma_2$ , как это имеет место в (1), поэтому создаются только две позиции для размещения трех рем – вершинного имени, локативной ИГ и модификаторов вершинного имени<sup>5</sup>.

$\Gamma_1 + \text{Рема}_1 + \Gamma_2 + \text{Рема}_2 + (\text{F}_2) + \text{Рема}_3$

- (3) кхм. យីអើ! ម៉ែបាននីត្រមកពីណាវែហែកហ្នឹង?  
*ji: ʔaə mae ba:n chat mɔ:k pi: na:*  
 INTERJ NTERJ матушка получать зонт приходит; REZ откуда  
**haek nuj**  
 разорванный этот  
 ‘Матушка, откуда у тебя этот разорванный зонт?’ (Тосанаводей: 5)

Для сравнения, в (4) коммуникативный статус ремы имеет только «разорванный», а «зонт» имеет коммуникативный статус темы:

- (4) кхм. ម៉ែចង់ត្រឃើងវែហែកអញ្ចឹងហើយម៉ែ!  
*tes chat jə:ŋ haek ʔaŋsəŋ haəj mae*  
 почему зонт 1PL рванный такой REZ матушка  
 ‘Матушка, почему наш зонт такой рванный?’ (Тосанаводей: 5).

В (5) также используется стратегия расщепления интродуктивной ИГ *ru:p*: *ciəm tuəj*.

- (5) кхм. លោកពូ សូមមេត្តា ... គូររូបចៀមឲ្យគ្នាមួយមក!  
*lò:k pu: soum mətta: ku: ru:p ciəm*  
 дядюшка пожалуйста будьте добры рисовать изображение баран  
*ʔaoj khniə tuəj mɔ:k*  
 DAT 1SG один IMP  
 ‘Пожалуйста... нарисуйте мне барашка’ (МП: 13).

### 3. Создание новых синтаксических моделей при переходе к письменной форме языка

#### 3.1. Модель предложения со сдвигом ремы в древне-кхмерском и языке кар

В древне-кхмерском языке и никобарском языке кар на этапе перехода от устной речи к письменной возникает одна и та же модель предложения со сдви-

<sup>5</sup> Пример взят из текста, в котором рассказывается, как мудрому феодалу удалось разрешить спор двух женщин, каждая из которых утверждала, что зонт принадлежит ей. Женщины порвали зонт, пытаясь вырвать его друг у друга. Каждой из них было велено пойти домой с разорванным зонтом, и реакция домочадцев показала, кому он принадлежит. Предложение (3) – это слова детей женщины, которая пришла домой с чужим зонтом, предложение (4) – слова детей женщины, которой принадлежит зонт.

гом ремы с использованием одного и того же языкового ресурса – маркера зависимой предикации *ta*.

Данный факт тем более примечателен, что между этими языками мало общего. Их родство отдаленное: их относят к разным ветвям австроазиатской семьи. Они принадлежат к разным языковым типам: древне-кхмерский – изолирующий язык, никобарские языки агглютинирующие с элементами синтетизма и изоляции. У них разный статус: древне-кхмерский – государственный язык, кар – бесписьменный<sup>6</sup> разговорный язык автохтонного населения острова Кар-Никобар, входящего в группу Никобарских островов. Наконец, между этими языками расстояние в несколько веков: древне-кхмерский – язык эпиграфических надписей Камбоджи начала 7 – середины 15 в., письменный материал по языку кар – перевод евангелия от Марка – относится к языку конца 19 – начала 20 вв.

Появление новой модели в обоих языках связано с тем, что в устной речи ритмико-интонационные средства в значительной степени компенсируют отсутствие сегментных средств, присущих языку письменному, а опора на контекст в разговорном языке значительно больше. Поэтому решение задач коммуникации при переходе от устной речи к письменному языку требует создания новых синтаксических моделей.

Это в особенности актуально для древне-кхмерской эпиграфики. Эти тексты выгравированы на камне в храмах и посвящены указам короля, описанию имущественных отношений, родословных с указанием заслуг и королевских милостей, описанию даров храмам со скрупулезным перечислением даров и дарителей, описанию организации функционирования храмов с перечислением прав и налагаемых обязанностей. Коммуникативная структура таких текстов должна была обеспечивать точность указания объектов и субъектов описываемых положений дел и правильную расстановку акцентов.

В русском языке предложения со сдвинутой ремой получаются в результате линейно-акцентного преобразования: *Иван↑ ee любит↓* > *Любит ee↑ Иван↓*. При этом, как отмечает Е. В. Падучева, смысл предложения со сдвинутой ремой – это предложение тождества: *Иван есть тот, кто ee любит* [Падучева 1985].

В древне-кхмерском языке предложения со сдвигом ремы образуются при помощи маркировки сказуемого маркером зависимой предикации *ta*<sup>7</sup>, связкой тождества *gi* или их комбинацией [Погибенко 2012: 257-59]. Такой морфосинтаксис отражает логико-семантическую структуру предложений со сдвигом ремы, о которой говорит Е. В. Падучева. Рассмотрим примеры.

(6) д.кхм. *vvaṃ āc ti srac ta khloñ a<sup>a</sup>nak nau<sup>8</sup>*  
NEG мочь DP зависеть DN начальник люди PTCL

<sup>6</sup> Хотя язык кар использовался как язык административной документации колониальных властей.

<sup>7</sup> В древне-кхмерском маркер *ta* имеет несколько функций, см. [Погибенко 2012, 2020].

<sup>8</sup> В ДАК *nai* дается как начальная частица, топикализирующая подлежащее, в частности смещенное. Однако из примеров, там приведенных, следует, что *nai* маркирует сдвиг ремы на подлежащее, как и в данном примере.

*<sup>a</sup>nak ta pradhāna ta noḥ deça noḥ*  
 люди DP верховный DN(GEN) этот страна этот  
*noḥ gi ta paripālana*  
 этот (= вот) COP ID DP сохранять

‘(Эта постройка) не может подчиняться начальнику над персоналом, но знатные люди этой страны – вот кто хранит ее’ (K351, 11-13)<sup>9</sup>.

В следующем предложении сдвиг ремы происходит на имя *tāñ*. Как считает Ж. Седес [IC V: 69], это имя человека, который был организатором церемонии дарения даров господина *Mratañ Kumārasvāmi*. В этом предложении было важно точно указать, кто даритель, а кто организатор церемонии дарения<sup>10</sup>.

(7) д.кхм. *amnoy mratañ kumārasvāmi ai ta vraḥ tilakeçvara*  
 дары господин PN LOC DAT бог PN  
*tāñ gui ta yajamāna ta vraḥ*  
 PN COP ID DP даритель DAT бог  
 ‘Дары господина *Kumārasvāmi* богу *Tilakeçvara*, жертвователю – *Tāñ*’  
 (K664, 69)

В языке кар сказуемое предложения со сдвигом ремы также выступает в форме ЗП с маркером *tə* (*ta*). В (8) сдвиг ремы происходит на темпоральную составляющую, в (9) – на субъектную:

(8) кар *ət heṅ tak pəri tə-akahalo:n tə ɽe, ɽaŋuə*  
 NEG один CLF но *tə*-знать ACC 3SG.OBL.INAN 3SG.NOM.INAN  
*saka:mə ɽe:tak ɽinre:*  
 день время тоже

‘О дне же том или часе, никто не знает’ (букв. ‘Никто не знает того, тот день и то время’) Марк, 13-32 (Whitehead, 1918: 10).

(9) кар *tə kale:k takə ən ɽəh ɽi:ə*  
 DP сильный очень 3SG.NOM этот ребенок  
 ‘Этот ребенок – очень сильный’ (Whitehead, 1993: 110).

### 3.2. Модель расщепленного предложения в древне-кхмерском языке

В древне-кхмерских текстах отмечаются предложения, построенные по модели расщепленного предложения. Они используются для выделения той части, на которую падает смысловое ударение. В них могут использоваться разные маркеры релятивизации и их комбинации: *ta, tel, man, ta tel, ta man, man tel*. Наиболее типичной является структура, в которой в начале главного предложения находится связка тождества *gi* (*gui*), а в начале придаточного – маркер релятивизации:

(10) *gi* + Главное предложение DP + Придаточное предложение

<sup>9</sup> K351, 11-13 – номер и строки надписи по корпусу IC.

<sup>10</sup> Слово *tāñ* имеет несколько значений, в ДРАК в этом предложении оно переводится как ‘госпожа’, т.е. подразумевается, что заказчиком церемонии дарения была супруга господина *Kumārasvāmi*.



В современном языке расщепленные предложения имеют такую же структуру, за одним исключением: связка тождества может опускаться.

- (11) кхм. គឺគាត់ហើយដែលបានសរសេរសំបុត្រនេះ  
*ku:* kōət haəj dæɛl ba:n sɔːse: sɔmbot nih  
 COP ID 3SG PTCL DP PRF писать письмо этот  
 ‘Это он написал письмо’ (КРС).

В (12) представлен один из вариантов предложения утращения, обращенного к потенциальным вандалам, которое помешалось в конце надписи на древне-кхмерском языке. В данном предложении решается задача коммуникативного выделения названий адов<sup>11</sup>.

- (12) д.кхм. *ge ta sak gi ge ta camlāk gi*  
 3SG DP грабить здесь 3SG DP высекать здесь  
*lān vraḥ avicīnaraka mahāraurava <...>*  
 уничтожить (надписи) храм Авичи-нарака Махараурава  
*gi nā ge tel kaṃvoṅ doṅ ge pitāmātā ge*  
 COP ID место 3SG DP пребывать вместе PL предки 3SG  
 ‘Те, кто грабят здесь, те, кто делают надписи здесь, высекают здесь, стирают надписи храма, Авичи-нарака, Махараурава <перечислены другие ады> – это там, где они (будут) пребывать со своими предками’ (K728, 5).

Рассмотрим предложение, которое свидетельствует о том, что техника клефтинга в древне-кхмерском языке еще находится в процессе развития. В (13) при помощи клефтинга выделяется ИГ *noḥ sre* ‘это поле’. Отклонения от структуры, представленной в (10), следующие. Во-первых, здесь используется грамматический ресурс, сходный с тем, который был использован в предложении (1) на языке ма. Комплексный маркер зависимой предикации *man tel* расщеплен на два состава: *man* размещен в главном предложении, *tel* – в придаточном. Во вторых, в придаточном предложении порядок слов изменен на обратный.

- (13) д.кхм. *gui man gui noḥ sre tel oṃ ta vraḥ*  
 COP.ID DP COP.ID это поле DP давать DAT бог  
*poṅ rudrabhava*  
 TTL PN  
 ‘Это то поле, которое Понь Рудрабхавы дал богу (букв. ‘Это которое, это поле, которое дал богу Понь Рудрабхавы’) (K79, 20-21).

#### 4. Предикаты оценки как коммуникативный вызов дискурса

Хорошо известно, что предикаты оценки часто требуют особых моделей для своего употребления. Специфический морфосинтаксис предикатов оценки связан со спецификой их семантической структуры, в которой важное место отводится прагматическому компоненту. Приведем два примера.

<sup>11</sup> В списке адов первым идет Авичи-нарака – самый глубокий, восьмой круг горячего ада, где преступники беспрестанно умирают и вновь рождаются.

Пример первый. В ходе полевой работы по изучению языков малых народов Вьетнама возникла ситуация, когда информант отказался переводить предложение «*Эта женщина красивая*», сказав, что оно «неправильное», и заменил его на «правильное» предложение «*Эта женщина красивая, как девушка*». В данном случае языковое поведение информанта, создавшего особую модель для перевода предложения с предикатом «красивый», объясняется, используя выражение Н. Д. Арутюновой, «семантической опустошенностью предикатов оценки». По этому поводу Н. Д. Арутюнова пишет:

«Оценочные предикаты информативно недостаточны. В тексте, так или иначе, компенсируется их смысловая неполнота и неоднозначность, проистекающая из нестабильности их смыслового объема и тех нормативов, на которых основана оценка. Самым простым приемом является дескриптивное развертывание оценки, состоящее в том, что вслед за оценочным предикатом эксплицируется эталон или фактическое положение вещей» [Арутюнова 1988: 92].

Как видим, наш информант попытался ответить на коммуникативный вызов дискурса тем, что включил в структуру предложения синтаксический компонент, выражающий понятие компаративного класса.

Пример второй. С проблемой опустошенности предиката оценки «огромный» столкнулся составитель надписи (K393) на древне-кхмерском языке. Он также ответил на коммуникативный вызов дискурса созданием языковой модели, при помощи которой постарался преодолеть «семантическую опустошенность» данного предиката.

В семантическую структуру предикатов оценки входят прагматические компоненты, которые являются переменными сущностями. Выражения *большое озеро, красивая женщина, вкусный чай, высокий каблук* и т.п. не дают представления о характеризуемом объекте в полном объеме. Референт выражения *большое озеро*, обозначенный на карте страны и на плане дачного кооператива, будет иметь разные размеры. Чтобы донести смысл выражения «огромное озеро» в том объеме, в котором требовалось в данном конкретном дискурсе, и дать правильное представление о размерах озера, составитель надписи использовал стратегию создания лексического комплекса *mahimā pi chloñ* с глаголом «пересекать», в котором инкорпорированным участником<sup>12</sup> является субъект оценки – адресат данного сообщения.

(14) д.кхм. *vraḥ travāñ ti mahimā pi chloñ*  
ТТЛ озеро место огромный чтобы пересечь  
'озеро, огромное чтобы переплыть' (K393).

Используя данную стратегию, составитель надписи переводит предикат «огромный» в разряд предикатов утилитарной оценки<sup>13</sup>, что делает смысл сообщения понятным адресату. Примечательно то, что в санскритской части надписи, которая обращена к адресату другого рода, указываются конкретные топографические параметры подаренных земель в восьми направлениях.

<sup>12</sup> Об инкорпорированном участнике см. [Падучева 2004: 57-58].

<sup>13</sup> Об утилитарной оценке см. [Арутюнова 1988: 75].

В случае данной надписи коммуникативный вызов дискурса предельно ясен. Как можно понять из самой надписи, ее автор – лицо высокого ранга, пользующееся милостями короля. Его акт дарения имеет целью возродить храм, который опустел и более не функционирует. Надпись содержит описание большого количества его даров храму. Важность и ценность этих даров подчеркивается на протяжении всего текста: золотые и серебряные предметы, украшения – это «достойные дары», драгоценные камни (бриллианты, сапфиры, изумруды и др.) – все «высокого качества», «лучшие из камней», «прекрасные», «отличные». Он подарил и заселил 23 деревни, «вновь заселенные», с рекой, прорыл каналы, сделал насыпи, «чтобы всем переходить (по ним)» – все это он сделал с заботой о «страждущих», «как отец для детей». Естественно, что в таком дискурсе составитель надписи счел выражение «огромное озеро» недостаточно информативным, тем более что слово *mahimā* – санскритское заимствование.

Ж. Сёдес, редактор и переводчик опубликованных надписей Камбоджи, не уловил суть такого языкового поведения автора надписи и перевел слово *chloñ* как ‘инаугурация’<sup>14</sup>, синтаксически связав его с последующей частью текста: «Mention <...> d’une grande pièce d’eau (*vrah travāñ ti mahimā*). Description de ce qui semble avoir été une inauguration (*chloñ*)» [IC, VII: 69]. Сегментация текста на языке изолирующей типологии бывает не простой задачей, но не в данном случае.

## 5. Вариативность паттернов как результат «исправления ошибок» при возникновении новых языковых моделей

*A speaker is like a lousy auto mechanic: every time he fixes something in the language, he screws up something else.*

*Joseph H. Greenberg*

Говорящий как нерадивый автомеханик: всякий раз, когда он что-то исправляет в языке, он что-то другое портит.

Джозеф Гринберг

В данном разделе мы рассмотрим примеры изменения языковых моделей в диахронии как урегулирование конфликтов стратегий выражения, что является ответом на возникающие коммуникативные вызовы, связанные с основной функцией языка выражать информацию «точно» и «эффективно».

В синхронии вариативность в языке одной и той же модели выглядит либо как избыточность, либо как неправильное воспроизведение. В примерах, которые мы рассмотрим, вариативность паттернов является побочным продуктом в процессе создания новой модели.

### 5.1 Результативная конструкция в языке ма

В языке ма в результативной конструкции возможен разный порядок компонентов:

---

<sup>14</sup> В древне-кхмерском языке *chloñ* ~ *chloññ* ~ *chlañ* имеет значения: ‘1. to pass from one side to the other, cross, traverse, 2. to carry across, transmit, communicate; to bring into effect (through religious or magical means), 3. to mark the beginning of; to celebrate, inaugurate; to consecrate, dedicate’ (DAK).

- а. TA – V<sub>1</sub> – NP<sub>OBJ</sub> – V<sub>REZ</sub>  
 б. TA – V<sub>1</sub> – V<sub>REZ</sub> – NP<sub>OBJ</sub>  
 в. V<sub>1</sub> – NP<sub>OBJ</sub> – TA – V<sub>REZ</sub>

Примеры:

- (15) ма *b̀n̄n* *l̄n̄* *l̄e* *po* *sre* *ḡəs*  
 деревня+деревня=жители PRF мотыжить поле завершать(ся); REZ  
 ‘Крестьяне обработали поле’ (Погибенко, Ефимов 2020, ПРП 65, Вар.1).
- (16) ма *b̀n̄n* *l̄n̄* *l̄e* *po* *ḡəs* *sre*  
 деревня+деревня=жители PRF мотыжить завершать(ся); REZ поле  
 ‘Крестьяне обработали поле’ (Погибенко, Ефимов 2020, ПРП 65).
- (17) ма *b̀n̄n* *l̄n̄* *po* *sre* *l̄e* *ḡəs*  
 деревня+деревня=жители мотыжить поле PRF завершать(ся); REZ  
 ‘Крестьяне обработали поле’ (Погибенко, Ефимов 2020, ПРП 65, Вар.2).

Эти три варианта результативной конструкции суть свидетельства многоэтапных преобразований в диахронии, которые привели к возникновению серийной конструкции данного типа. Варианты «а» и «б» являются отражением разных этапов синтаксической компрессии. Вариант «а» возникает в результате корферентного сокращения субъектной ИГ во второй предикации. Вариант «б» – следующий этап компрессии, он приводит к универбализации глаголов.

В варианте «в» темпорально-аспектуальный показатель находится при результативном глаголе. В языке ма это не распространенная позиция для аспектуального показателя в серийной конструкции. Кроме того, она нарушает принцип универбальности, в соответствии с которым серийный комплекс должен иметь объединенную аспектуально-темпоральную периферию. Возникновение такого варианта, несомненно, связано с прагматическим фактором. Порядок следования V<sub>1</sub>V<sub>REZ</sub> унаследован от исходной двусентенциальной структуры, в ней он был мотивирован принципом иконичности. В серийной конструкции с универбальным комплексом иконический принцип теряет силу. Ему на смену приходит фактор прагматический, в соответствии с которым фазисное значение занимает более высокое место в семантической и коммуникативной структуре предложения. Вот почему аспектуальный показатель переносится к главному (фазисному) глаголу.

Этот прагматический фактор настолько силен, что он становится драйвером следующего этапа преобразований: в результативной конструкции с фазисным глаголом происходит изменение порядка следования компонентов:

$$V_1 V_{PHAS} \rightarrow V_{PHAS} V_1$$

В результате возникает конструкция с предикатным актантом, в которой принцип прагматической выделенности фазисного глагола отражается в порядке слов.

- (18) ма *khaj* *ʔv̄?* *dà* *ḡəs* *ci* *thə*  
 3SG NEG еще завершать(ся) писать письмо  
 ‘Она не написала письма’ (Погибенко, Ефимов 2020, ВИД 59).

В языке кату находим два варианта результативной конструкции, в которых задача обеспечения коммуникативной выделенности фазисного и результативного

компонента решается другими способами. В (19) и (20) они выносятся в начало предложения.

(19) кату *song*            *ku*        *têêng*    *buông*    *thôi*  
 заканчивать 1SG    делать    лодка    достаточно  
 ‘Я закончил делать лодку’ (Costello, 1993: 66)

(20) кату *tak*            *nhi*        *bəər*      *đò*        *viêng*    *tapaai*  
 достигать    DU        два        3 SG/PL    гнаться    заяц  
*mót*        *kalóóng*    *tóng*  
 входить    внутри    дыра  
 ‘Когда они догнали зайца, заяц спрятался в нору’ (Costello 1993: 258).

В (21) конструкция того же типа, что в (18), но не с кореферентным сокращением ИГ<sub>СУБ</sub> при втором глаголе, а с анафорической заменой.

(21) кату *sasau*    *tamoog*    *ku*        *luch*        *đò*        *đống*  
 внук        сестра    1SG    переставать 3SG    давать  
 ‘Внук моей сестры перестал давать (рис)’ (Costello 1993: 640).

В лингвистической литературе часто задается вопрос, откуда в естественных языках возникает многообразие синтаксических конструкций. Разные тактические варианты результативных конструкций в языках ма и кату могут служить ответом. Эти варианты возникли под действием разноуровневых факторов: синтаксических, семантических, прагматических, которые могут привести не только к созданию разных вариантов данной конструкции, но и к возникновению конструкций другого типа.

«Существующие сегодня грамматические структуры и их умпомрачительное разнообразие (*mind-blowing diversity*) лучше понимаются и объясняются, если проследить пути развития, которые привели к их возникновению» [Givon 2009: xvii].

## 5.2 Счетная конструкция в мон-кхмерских языках

В данном разделе мы рассмотрим порядок слов в счетной конструкции. Мы попытаемся доказать, что модель с продвижением группы «числительное – классификатор» в начало счетной конструкции может возникать в диахронии в результате урегулирования конфликтов стратегий выражения, что является ответом на возникающие коммуникативные вызовы. В примерах, которые мы рассмотрим, вариативность паттернов является побочным продуктом процесса создания счетной конструкции.

Дж. Гринберг выделял четыре типа счетной конструкции в языках мира [Greenberg 1972]: два с постпозицией счетных групп «Классификатор – Числительное» / «Числительное – Классификатор» по отношению к имени и два с препозицией этих групп по отношению к имени.

Гринберг полагал, что выбор одной из четырех моделей напрямую связан с основным порядком слов в данном языке. Для мон-кхмерских языков это утверждение является ошибочным, в них порядок слов одинаковый, но в кон-

струкциях с классификатором разных. Более того, есть языки, в которых используются конструкции и с препозицией, и с постпозицией имени (кату).

Ниже приводятся типы порядка слов в счетной конструкции, которые выделяет Дж. Гринберг (1, 2, 3, 4), два типа порядка слов, которые не отмечаются в языках мира (5, 6), счетная конструкция, в которой вместо классификатора используется повтор имени (7).

1. Числительное – Классификатор – Имя	ксингмул, ма, вьетнамский, китайский, кату
2. Имя – Числительное – Классификатор	вьетнамский (устар.), кхмерский, древне-кхмерский, кату
3. Имя – Классификатор – Числительное	древне-кхмерский
4. Классификатор – Числительное – Имя	(Гринберг, 1972) не отмечен в МК
5. * Классификатор – Имя – Числительное	не отмечен
6. * Числительное – Имя – Классификатор	не отмечен
7. Имя – Числительное – Имя	тайский, бирманский

Гринберг отмечает, что в языках мира не встречается порядок следования (5) и (6) [Aikhenvald 2000: 105]. Это утверждение верно и для МК языков, в них классификатор всегда примыкает к числительному. В [Погибенко 2012: 378-397] делается попытка объяснить отсутствие такого словоупорядка особенностями возникновения счетной конструкции. Там говорится, что счетная конструкция возникает в результате синтаксической компрессии исходной последовательности двух предложений, связанных по смыслу и описывающих одну ситуацию, в которой числительное является сказуемым второго предложения. Синтаксическое сжатие шло по следующим путям.

а) Изменение порядка слов на VO во второй, семантически зависимой предикации<sup>15</sup>: *Охотник убил обезьян, обезьян три.* → *Охотник убил обезьян, три обезьяны.*

б) Корреферентная замена повторяющейся ИГ на классификатор: *Охотник убил обезьян три CLF.*

В результате синтаксической компрессии числительное уходит с позиции вершинного предиката второго предложения и становится зависимым компонентом в именной группе в первом предложении. Тем самым оно утрачивает коммуникативную выделенность. Трансформация (в) – продвижение числительного с классификатором в позицию вершины ИГ – позволяет восстановить коммуникативную выделенность числительного.

в) *Охотник убил три CLF обезьяны.*

Последнее предположение подтверждается данными языка кату. В кату словоупорядок (1) используется в интродуктивных ИГ, в которых числительное является ремой, а словоупорядок (2) используется в идентифицирующих ИГ (второй и последующих номинациях объекта, уже введенного в дискурс), где числительное не является ремой [Погибенко 2012: 388-392].

<sup>15</sup> В МК языках отмечается часто, например, в придаточных предложениях.

Итак, по нашему предположению, словопорядок (1) возникал при прохождении всех трех этапов грамматикализации: (а), (б) и (в), словопорядок (2) – при прохождении этапов (а) и (б), словопорядок (3) возникал на этапе (б), минуя этап (а), словопорядок (4) возникал на этапе (в), минуя этап (а).

Предложенная схема грамматикализации счетной конструкции не только объясняет все отмеченные в языках мира типы порядка слов в счетной конструкции, но и делает понятным отсутствие порядка слов (5) и (6). Более того, предложенная схема объясняет происхождение структуры (7), когда в позиции классификатора используется повтор считаемого имени или его части. Например, бирм. *eij tə-eij* ‘один дом’ (букв. ‘дом один дом’) [ЛЭС 1990: 75]. Такая конструкция возникает, когда на шкале грамматикализации отсутствует этап (б) – замена повторяющегося имени на классификатор.

### Сокращения и условные обозначения

ДКМ	древне-кхмерский	LOC	локатив
д.кхм.	древне-кхмерский	NEG	отрицание
вьет.	вьетнамский	NOM	номинатив
кхм.	кхмерский	OBJ	объект
СКД	Анкета «Сказуемое – дополнение»	OBL	косвенный падеж
ACC	аккузатив	PHAS	фазисный глагол
CLF	классификатор	PL	множ. число
COP ID	связка тождества	PN	имя собственное
DAT	датель	PRF	перфектив
DN	маркер зависимой именной группы	PTCL	частица
DP	маркер зависимой предикации	REZ	показатель результата
GEN	генитив	SG	ед. число
IMP	императив	TA	время-вид
INAN	неодушевленный	TTL	название титула
INTERJ	междометие		

### Источники

- ЛЭС – Лингвистический энциклопедический словарь. М., 1990.
- МП – អង្គទុន ជីសាំងតិចស៊ុយប៊ើរី ព្រះអង្គម្ចាស់តូច ភ្នំពេញ, ២០១៥ (Антуан де Сент-Экзюпери. *Маленький принц*. Пномпень, 2015).
- Тосанаводей – ទស្សនាវដ្តីរឿងព្រេងខ្មែរថ្មី. ភ្នំពេញ, ២០០៦ (*Новый сборник кхмерского фольклора*. Пномпень, 2006.)
- DAK – Jenner Ph. N. *A Dictionary of Angkorian Khmer*. Canberra, 2009.
- DPAK – Jenner Ph. N. *A Dictionary of Pre-Angkorian Khmer*. Canberra, 2009.

### Литература

1. Аругюнова Н. Д. *Типы языковых значений. Оценка. Событие. Факт.* М., 1988.
2. КРС – Горгониев Ю. А. *Кхмерско-русский словарь.* М., 1984.
3. Падучева Е. В. *Высказывание и его соотнесенность с действительностью.* М., 1985.
4. Падучева Е. В. *Динамические модели в семантике лексики.* М., 2004.
5. Погибенко Т. Г. Связка тождества *кш:* в кхмерском языке // *Китайское языкознание. Изолирующие языки. XII Международная конференция. Материалы.* М. 2004. С. 242–250.
6. Погибенко Т. Г. *Австроазиатские языки: проблемы грамматической реконструкции.* М.: ИВ РАН, 2012.
7. Погибенко Т. Г. Никобарские языки и древне-кхмерский: формы зависимой предикации // *Вестник Института востоковедения РАН.* 2020. № 3 (13). М.: ИВ РАН. С. 317–332.
8. Погибенко Т.Г., Ефимов А.Ю. *Материалы российско-вьетнамской лингвистической экспедиции. Язык ма.* М., 2020.
9. Тестелец Я. Г. *Введение в общий синтаксис.* М., 2001.
10. Costello N. *Katu Folktales and Society.* Vientiane, 1993.
11. Givon T. *The Genesis of Syntactic Complexity: Diachrony, Ontogeny, Neuro-Cognition, Evolution.* John Benjamins Publishing Company, 2009.
12. Greenberg J. Numeral classifiers and substantival number; problems in the genesis of a linguistic type // *Proceedings of the XIth International Congress of Linguistics.* Bologna, 1974. Vol.1.
13. Hopper P. Some recent trends in grammaticalization // *Annual Review of Anthropology.* 1996. № 25. P. 217–236.
14. Pogibenko T. Copula of identity in Old Khmer // *Mon-Khmer Studies.* Salaya, 2010. Vol.39. P. 61–74.
15. Whitehead H. *St. Mark. In Car-Nicobarese.* Rangoon 1918.
16. Whitehead H. *Dictionary of the Car-Nicobar Language.* New Delhi, 1993. (Reprint from 1925 edition).



## Употребление и характеристики глагола *bi-* ‘быть’ в эвенкийском языке

Елена Леонидовна Рудницкая

Институт востоковедения РАН, Москва, Россия,  
erudnitskaya@gmail.com, ORCID 0000-0002-4455-262X

**Аннотация.** В статье рассматриваются отличительные черты и употребление глагола *bi-* ‘быть’ в экзистенциальном употреблении в устной речи. В разных говорах этот глагол может употребляться с нерегулярным показателем небудущего времени *-si/-hi* или с регулярным *-rA*. В литературном языке используется только форма с *-si/-hi*. Предлагается, что форма с *-rA* проникла в устный язык из непского говора, на котором был основан первый вариант эвенкийского литературного языка. Также предлагается объяснение, почему форма с *-rA* имеет в большинстве случаев только значение местонахождения, но не обладания (эти два значения характерны для экзистенциального глагола).

**Ключевые слова:** вспомогательный глагол, связочный глагол, настоящее время, экзистенциальное употребление, устный язык, литературный язык, модель управления, обладание, местонахождение, эвенкийский язык

**Abstract.** The paper examines special features and use of the verb *bi-* in its existential use in the Evenki oral speech. In different dialects, this verb can be used with the irregular marker of non-future tense *-si/-hi* or with the regular one, *-rA*. In the literary language, only the irregular form is used. We suppose that the *-rA* form has penetrated into the oral language from the Nepa dialect, on which the first version of the Evenki literary language was based. We also suggest the explanation as to why the *-rA* form usually occurs with the location meaning, but not in the possession meaning (these two meanings are equally frequent interpretations of an existential verb).

**Keywords:** auxiliary, copula, presence, existential use, oral language, literary language, subcategorization frame, possession, location, Evenki

### 1. Введение

Глагол ‘быть’ – сложный объект исследования. В зависимости от языка у этого глагола может различаться сфера употребления. Исследователи на материале европейских языков рассматривают вопрос о так называемом «экзистенциальном» значении глагола ‘быть’: [Арутюнова 1976; Babby 1980; Borschev, Partee 2002; Hazout 2004; etc.]. В английском языке *be* может быть и вспомогательным (*The boy is crying*), и связочным (*John is young/ John is in the garden*), и экзистенциальным (*There is an apple in the fridge*). Различия между тремя данными употреблениями *be* определяются по лексико-синтаксическому контексту, а морфологически *be* во всех трех случаях оформляется в целом одинаково. Во многих же не индоевропейских языках, в том числе в алтайских, экзистенциальное значение глагола ‘быть’ выражается другой основой – например, *iss-*, в отличие от (клити-

ческой) связки *-i-*, в корейском языке, ср. (1а-б), или экзистенциальный и отрицательный экзистенциальный глаголы *var-* и *yok-* в турецком [Wagner-Nagy 2011: 177].

- (1) а. *ce-nun yengkwuk salam-i-eyyo*  
 я-FAM-TOP английский человек-be.COP-POLITE  
 ‘Я англичанин’ [Yeon, Brown 2011: 155].
- б. *neyngcangko-ey oleynci cwusu-ka iss-ta*  
 холодильник-LOC апельсин сок-NOM быть.EX-DECL  
 ‘В холодильнике есть апельсиновый сок’ [Yeon, Brown 2011: 152].

В эвенкийском языке есть отрицательное экзистенциальное предикативное слово *ācin* ‘NEG.EX’ [Рудницкая, в печати], отличное от вспомогательного отрицательного глагола *ə-* ‘NEG’, а просто экзистенциальное значение выражается глаголом *bi-*, также являющимся вспомогательным глаголом и связкой. Рассмотрим (2а-в) как примеры на эти три значения<sup>1,2</sup>.

- (2) а. *koj-no bi-hi-m gun-də-rə-n*  
 заблудиться-PTCP.PF быть-NFUT-1SG говорить-IPFV-NFUT-3SG  
 ‘‘[Я] заблудился’’, – [он] говорит’ (Малые языки 2012–2021, Эконда)  
 [Yeon, Brown 2011: 152].
- б. *Tari-l oro-r burgu-l (bi-si-Ø)*  
 тот-PL олень-PL толстый-PL (быть-PRS-3PL)  
 ‘Те олени толстые’ [Nedjalkov 1997: 277].
- в. *tər turu-dū žə d’u-n bi-hi-n mon-ŋi-n*  
 тот Тура-DAT FOC дом-PS.3SG быть-NFUT-3SG RFL-PROPR-3SG  
 ‘[У нее] есть собственный дом в Туре’ (Малые языки 2012–2021, Тутончаны).

Эвенкийский язык похож на русский тем, что оба языка относятся к языкам типа «В-languages» («бытийные языки») по [Harves, Kaune 2012]. В эвенкийском у экзистенциального *bi-* есть существенные морфосинтаксические отличия от вспомогательного и связочного *bi-*, связанные с факультативностью только первых

<sup>1</sup> По работе [Василевич 1948: 64], перед гласной и в интервокальном положении звуку [s] в южном диалекте соответствует [h] в северном диалекте. Однако Г.М. Василевич фиксирует состояние эвенкийского языка на начало / середину XIX в. В работе [Клячко, в печати] по материалам корпуса устных рассказов (сноска 2) в языке XXI в. эта закономерность нарушена. *Xu-* вместо *cu-* может встречаться в устной речи носителей, говоры которых относятся и к южным, и к северным (перечень поселков по диалектам дается далее в основном тексте). В целом этот критерий Г.М. Василевич не соблюдается. Мы этот вопрос не рассматриваем, считая обе формы эквивалентными вариантами.

<sup>2</sup> Большинство примеров взято из корпуса устных рассказов Малые языки 2012–2021: «Малые языки Сибири: наше культурное наследие»; URL: <http://minlang.srcr.msu.ru/> (дата обращения 27.05.2021). Поиск: <http://gisly.net/corpus/> (дата обращения 23.05.2021).

двух в настоящем времени, с морфологическим оформлением в настоящем/«небудущем» (NFUT) времени (только *-si-/-hi-* в связочной/вспомогательной функции и *-si-/-hi-/-rA-* в экзистенциальной функции) и др. Эти вопросы будут рассмотрены в п. 2. Чтобы выдвинуть гипотезу об особенностях значения *bi-rA-* в экзистенциальном употреблении (в основном местонахождение), экзистенциальное употребление *bi-rA-* сравнивается с отрицательным экзистенциальным предикативным словом *āčīn* ‘NEG.EX’ (п. 2.1).

## 2. Экзистенциальное употребление *bi-* в настоящем и «небудущем» времени: морфологическое оформление и особенности значения

Форма *bi-hi-/ bi-si-* является исключением. Эта форма считается нерегулярной формой глагола *bi-*, в то время как регулярный аффикс настоящего времени (PRS) – *-d'ArA* [Nedjalkov 1997: 236-237; Bulatova, Grenoble 1999; Булатова 2002; Болдырев 2007: 664], ср. (3):

- (3) ...*akni-li-n*                      *baka-d'ara-Ø*                      *tar*    *aŋi-wa-n...*  
старший.брат-PL-PS.3SG    находить-PRS-3PL    тот    это-ACC-PS.3SG  
{Старшие братья,} ‘старшие братья находят тот это самое’, {это самое}  
[мешочек]. (Малые языки 2012–2021, Полигус)

Показатель настоящего времени PRS *-d'ArA* можно разделить на *-d'A* ‘IPFV’ и *-rA* ‘NFUT’ [Nedjalkov 1997: 236-237]. И.В. Недялков предполагает грамматикализацию *-d'A + -rA* → *-d'ArA* в значении PRS и как доказательство этого отмечает (без примеров), что последовательность *-d'A-d'ArA* ‘IPFV’ + ‘PRS’ возможна. Однако при таком анализе неполным является набор значений настоящего времени: глагол на *-d'ArA* может обозначать прогрессив, потенциалис, настоящее историческое (привязано к точке отсчета в прошлом). У *-d'ArA* нет значения постоянно-/повторяющегося действия или состояния, характерного для PRS: это значение передается *-rA* ‘NFUT’ или сочетанием аспектуального показателя с *-rA*. Чтобы восполнить данный семантический пробел, исследователи вводят аналитическое время «настоящее хабитуальное»: *-pkī/-vkī + bi-* ‘PTCP.HAB + bi-[PRS]’ [Nedjalkov 1997: 235-236]. Другой вариант «настоящего хабитуального»: *-ŋnA-rA* ‘HAB-NFUT’ [Болдырев 2007: 670-671]. В целом существование «грамматикализованного показателя *-d'ArA* ‘PRS’» не кажется полностью очевидным, так что некоторые исследователи вообще не считают *-d'ArA* ‘PRS’ отдельным временем, ограничиваясь «небудущим» ‘NFUT’ временем на *-rA* [Константинова 1964: 180-181].

Мы будем придерживаться последней точки зрения. В этом случае *bi-si-n* в (2б), как и в (2а) и (2в), нужно глоссировать как ‘быть-IPFV-NFUT-3SG’, а *baka-d'ara-Ø* в (3) – как ‘находить-IPFV-NFUT-3PL’; глосса PRS должна быть исключена. Такого глоссирования мы придерживаемся в дальнейшем; форма *bi-hi-/si-* относится к времени ‘NFUT’ (*-rA*), являясь исключением.

В работе [Константинова 1964: 174-180] обращается внимание на возможность регулярной формы *bi-rA-* со значением настоящего времени. В (4) мы видим форму *bi-d'ə-rə-*, этот пример корректен, даже если считать, что *-d'ə-rə* в *bi-d'ə-rə* это форма PRS. О.А. Константинова отмечает, что *bi-rA-* – это диалектная форма,

не относящаяся к литературному языку<sup>3</sup>. В (4) показано употребление *bi-rə-* со значением местонахождения, на которое указывает О.А. Константинова и которое в устных текстах наиболее частотно.

- (4) *Dulin-dū-tin oro-r bi-d'ə-rə-n ... {dulin-dū-tin}*  
 середина-ДАТ-PS.3PL олень-PL БЫТЬ-IPFV-NFUT-PS.3SG  
 'Среди оленей находится, ...' {среди них} (про потерянного оленя)  
 (Малые языки 2012–2021, Стрелка-Чуня).

[Болдырев 2000: 67-68], не уточняя, относится ли *bi-rA-* к литературному языку, указывает на более широкий круг значений *bi-rA-* (Б.В. Болдырев также не приводит источники). По эвенкийско-русскому словарю Б.В. Болдырева [Болдырев 2000], *bi-rA-* имеет круг употреблений, традиционно ассоциирующийся именно с экзистенциальными, ср. (5а-г)<sup>4</sup>, причем значение настоящего времени может иметься не только в присутствии показателя *-d'ə*, как мы видели в (4), но и без него (5б-г). В этом случае форму *bi-rA-* можно считать приблизительным положительным коррелятом отрицательного экзистенциального *āchin*: *āchin* обозначает отсутствие объекта в каком-либо месте, отрицает обладание объектом или существование объекта, ср. (6).

- (5) а. *ar moha-l-dulā irəktə-l bi-rə-Ø*  
 тот лесок-PL-ALL лиственница-PL БЫТЬ-NFUT-3PL  
 'Там к лесу лиственницы были.' (Малые языки 2012–2021, Тура). [наличие]
- б. *Jakutskaj-dulā-rikta umūkōn mesto bi-rə-n.*  
 Якутск-ALL-LIM один место БЫТЬ-NFUT-3SG  
 'Только в Якутске одно место есть.' (Малые языки 2012–2021, Эконда).  
 [наличие]
- в. *Tug-da əhī-kōn wnuki-l-wi bi-rə-Ø*  
 так сейчас-ATTEN внук-PL-PS.1SG БЫТЬ-NFUT-3PL  
 'Так-то теперь внуки есть' (Малые языки 2012–2021, Тура). [обладание]
- г. *huwul-in oldo aja bi-rə-n*  
 весь-PS.3SG рыба хороший БЫТЬ-NFUT-3SG  
 {В озере} 'Вся рыба хорошая есть' (Малые языки 2012–2021, Чиринда).  
 [наличие]
- (6) *əhī-kokōn zweroferma ašin o-sa*  
 сейчас-INTS звероферма.NOM NEG.EX статья-PTCP.ANT  
 'Теперь зверофермы не стало' (Малые языки 2012–2021, Тура).  
 [отрицание существования объекта]

<sup>3</sup> О.А. Константинова приводит примеры на *bi-rA-* без указания на источники. В наших материалах по письменной речи XXI вв. (КИЭА РАН 2011–2017: Корпусы Института этнографии и антропологии РАН, в которые входят в основном нормализованные газетные и переводные тексты; URL: <http://corpora.iea.ras.ru/corpora/texts.php> (дата обращения: 10.05.2021)), эта форма не встретилась.

<sup>4</sup> Мы приводим не примеры Б.В. Болдырева, а аналогичные по смыслу и форме примеры из корпуса рассказов, см. сноску 2.

Форма *bi-riA-* в значении местонахождения употребляются с самыми разными аспектуальными аффиксами, помимо *-d'A* 'IPFV', ср. (7а-в). А форма *bi-si/-hi-* не допускает аспектуальных аффиксов, в отличие, например, от русского вспомогательного или связочного глагола *быть*: *Он бывает занят/ скупым*.

В значении местонахождения *bi-riA-* ведет себя как динамичный (стадиальный), а не статичный (индивидуальный) предикат, и подобные грамматические особенности есть не только в эвенкийском, но и в других алтайских языках. Ср. *iss-nun* (быть.EX-ATR.DYN) 'находящийся' в отличие от *i-n* (быть-ATR) 'являющийся' в корейском [Yeon, Brown 2011: 331-332].

(7) а. *tadū-wər bi-l-də-n*  
там-RFL.PL быть-INCH-NFUT-3SG  
{Назад вернулся и} '...там стал жить' (Малые языки 2012–2021, Эконда).

б. *d'u-dū bi-hi-n m'əsto idū bi-d'ə-rə-Ø <...>*  
дом.чум-DAT быть-NFUT-3SG место где быть  
*amtil-wi a-d'i-ŋkī-tin*  
родители-PS.1SG спать-IPFV-PSTITER-3PL

'В чуме есть место, где мои родители находятся, спят' (Малые языки 2012–2021, И. Увачан, Тутончаны).

в. ... *tadū i bi-li-ŋnə-rə-Ø*  
там и быть-INCH-NAV-NFUT-3PL  
{Раз они привыкли, где находятся,} 'там и стоят' (Малые языки 2012–2021, Белый Яр).

Особенно интересен пример (7б). В нем есть обе формы: *bi-hi-n* употребляется в экзистенциальном значении 'есть' (место), а *bi-d'ə-rə-Ø* – в значении местонахождения: (родители) 'находятся'.

Как показывают приведенные примеры, в целом рассказчики разграничивают вспомогательное/связочное употребление *bi-*, с одной стороны, и экзистенциальное, с другой. Однако в одних районах у формы *bi-riA-* экзистенциальное значение сводится к значению местонахождения, например, в Тутончанах, ср. примеры (4), (7б-в), а в других (примеры (5а-г)) – к более широкому спектру экзистенциальных значений.

Откуда появилась форма *bi-riA*<sup>5</sup>, и почему большинство исследователей не включают ее в литературный язык, что подтверждается и нашими данными по устным vs. письменным источникам (см. сноску 3)?

<sup>5</sup> Настоящее время на *-ra* распространено в тунгусских языках. Например, согласно [Суник 1962: 49], *-ra* в близкородственном северо-тунгусском эвенском языке это показатель PRS для глаголов регулярного спряжения, а *-da* и *-si* – для неправильных глаголов (ср. эвенкийское *bi-si/-hi-*). Согласно [Суник 1962: 54], *-ra* в нанайском (южно-тунгусском) языке это показатель PRS для I спряжения. Однако, согласно работе [Певнов 2018: 444], у эвенкийского языка было больше языковых контактов с предком ороцкого языка (вымерший язык, южная группа). Эти контакты отражаются в основном в лексике; данные по глагольным категориям у нас отсутствуют. По работе [Ваек 2020], у ороцкого языка (вымерший язык, южная группа) также



В непском говоре, как видно из примеров (8а-б), формы на *-pā (-rA)* имеют не только значение местоположения, как в (8а), но и более широкое экзистенциальное значение, как в (8б), где речь идет о наличии времени. Мы предполагаем, что формы на *-pā (-rA)* получили более ограниченную трактовку в южных говорах (только значение местоположения) потому, что южные говоры ближе к литературному языку, в котором форма *bi-rA*, очевидно, отсутствует. В южных говорах нелокативные экзистенциальные значения выражаются формой *bi-si/-hi-*.

Есть еще одна гипотеза для объяснения того, что экзистенциальные употребления у многих носителей сводятся к значению местоположения.

Раз у носителей говоров эвенкийского языка разное представление о значении *bi-rA-*, включающее в себя не обязательно только местонахождение, мнение, что *bi-rA-* – отдельный лексический глагол местонахождения [Болдырев 2007: 665], представляется необоснованным.

Мы показали вариативность употребления регулярного (*-rA*) и нерегулярного (*-si/-hi*) показателя небудущего ‘NFUT’ времени *bi-* в экзистенциальном/локативном значении по устным говорам. Эта вариативность происходит из распространения непской экзистенциальной формы на *-pā* по другим говорам в конце 1920-х гг. Регулярная форма, возможно, относится к литературному языку, но недостаточно обоснована трактовка ее как омонима вспомогательного/ связочного *bi-*<sup>7</sup>.

### 2.1. Сравнение *bi-rA-* и отрицательного экзистенциального предикатного слова *aʕin*: почему *bi-rA-* интерпретируется многими рассказчиками как глагол местонахождения?

В отличие от вспомогательного *bi-* [Шаляпина 2007: 181], у «экзистенциального» *bi-* валентности свои, а не полученные путем делегирования, или наследования<sup>8</sup>. Например, в русском сочетании *Иван будет писать письмо* агентивный актант *писать (он)* синтаксически не связан непосредственно с глаголом *писать*, а является зависимым вспомогательного глагола *будет*. Таким образом, *писать* делегирует свою валентность ‘агенса’ вспомогательному глаголу *будет*, а *будет* заимствует ее и реализует как собственную валентность, заполняемую местоимением *он*. В экзистенциальном употреблении у *будет* есть собственные валентности: [Borschev, Partee 2002], как и другие исследователи, предлагают семантический анализ экзистенциального глагола (на примере русского *быть*) как имеющего два аргумента: тема (“inner”) и местонахождение/локация или обладатель, посессор (“outer”): *В комнате будет кондиционер, У него будут дети*. Так, валентность

<sup>7</sup> Ср. русский глагол *быть* – это один глагол; он в экзистенциальном значении не обязательно опускается в настоящем времени: *Он (\*есть) болен* vs. *У него (есть) время*.

<sup>8</sup> Наследование (*делегирувание*) валентностей это один из случаев контекстных валентностей, куда включаются случаи несоответствия синтаксической валентности семантической [Мельчук 1974: 134–135; Апресян 1974: 125, 148]. З.М. Шаляпина дает определение делегирования валентностей (с. 177): «носитель валентности *передает (делегирует)* ее одной из сущностей, с которой он в данном контексте связан (непосредственно или опосредованно) отношениями зависимости, а та *заимствует* эту его валентность для реализации в качестве одной из собственных валентностей».

‘тема’ экзистенциального *будет* заполняется именем в именительном падеже (*кондиционер, дети*), а валентность ‘местонахождение/ обладатель’ – сочетаниями *в комнате и у него*.

Интересно, что валентности экзистенциального *bi-* в эвенкийском языке могут подвергаться делегированию. Такая ситуация встречается, например, в конструкциях, обозначающих наличие у кого-то родственников, ср. (5б) со стандартным выражением валентностей *bi-* и (9), где валентность ‘обладатель’ делегируется второму актанту по валентности ‘тема’, так что первый актант – зависимое второго (*ənin-ŋəhə* ‘мать-DEADREL’ – зависимое *akin-in* ‘старший.брат-PS.3SG’) в посессивной конструкции. В русском языке невозможна конструкция типа (9) с «внутренним посессором».

(5) б. *Jakutskaj-dulā-rikta umūkōn mesto bi-rə-n.*  
 Якутск-ALL-LIM один место БЫТЬ-NFUT-3SG  
 ‘Только в Якутске одно место есть.’ (Малые языки 2012–2021, Эконда)  
 [наличие]

(9) *utōlā-pti-dū [ənin-ŋəhə akin-in] bi-s’ō-n*  
 прежде-ATR.TEMP мать-DEADREL старший.брат-PS.3SG БЫТЬ-PST-3SG  
*andrej jefimowič*  
 Андрей Ефимович  
 ‘Раньше у моей покойной матери был старший брат, Андрей Ефимович.’  
 (Малые языки 2012–2021, Кислокан) [*bi-s’ō-n* → *akin-in* → *ənin-ŋəhə*]

Аналогичная валентностная структура и у отрицательного экзистенциального предиката [Matushansky 2020]: ср. (6) выше (п. 2) и (10а-б).

(10) а. *ədi-n ač̄in bi-č̄o-n ənin-ŋi-w*  
 муж-PS.3SG NEG.EX БЫТЬ-PST-3SG мать-PROPR-PS.1SG  
 ‘Мужа не было у моей матери’ (Малые языки 2012–2021, Ербогачён, северный говор).

б. *ač̄i-r o-ra-Ø bagdakə-l=tī tālā*  
 EG.EX-PL статья-NFUT-3PL дикий.олень-PL=FOC туда/там  
*učasko-l-dū-wun*  
 участок-PL-DAT-PS.1PL(EXCL)  
 {Уже сейчас} ‘не стало диких оленей-то на наших участках’ (Малые языки 2012–2021, Стрелка-Чуныя).

Согласно работе [Рудницкая, в печати], отрицательные экзистенциальные предложения с *ač̄in* выражают отношения обладания (10а) не реже, а скорее даже чаще, чем местонахождения (10б). Почему же «положительные» экзистенциальные предложения выражают в основном отношение местоположения, настолько, что Б. В. Болдырев предлагает выделить отдельный глагол «*bi-* местоположения» (см. выше, п. 2)? В наших данных конструкции с *bi-*, выражающие значение обладания, встречаются относительно редко: (5в), (9).

Мы предлагаем такую гипотезу: обладание и совместность в эвенкийском часто передаются именами обладания / совместности (см. [Константинова 1966: 154-155; Певнов 2019]): на *-č̄i* ‘ATR’; *-taj* ‘ATR.PROPR, *-lĀn* ‘ATR.PROPR’, *-nĀn*



‘COM.REL’. Поэтому конструкция с *bi-* используется для указания на обладание реже, чем для указания на местоположение. Отрицательная экзистенциальная конструкция с *ač̣in* используется и для указания на необладание, и на отсутствие где-либо, без альтернативных конструкций. Подобная асимметрия приводит к «перекосу» значения экзистенциального *bi-* в сторону местонахождения. Ср. примеры (11)-(13) на конструкции с *ač̣in* (необладание, примеры (а)), и обладание (примеры (б-в) с глаголом *bi-* или именами обладания).

- (11) а. *Nuŋa-r-tin* [*hutə-ja* ***aṣ̌i-r***] *bi-šō-tin*  
 3PL-PL-PS.3PL ребенок-PART NEG.EX-PL быть-PST-3PL  
 ‘У них не было детей’ (букв. ‘Они были **бездетные**’) (Малые языки 2012–2021, Хантайское озеро, северный говор).
- б. *ahī-ši* ***bi-sō-n***  
 женщина-ATR быть-PST-3SG  
 ‘Женатый был’ (Малые языки 2012–2021, Тутончаны).
- (12) а. *ər=kə* [*oron-ə* *aṣ̌in*] *əwəŋki ə-to-n*  
 этот=FOC олень-PART NEG.EX эвенк NEG-FUTCNT-3SG  
*utōlə-pti* *əmtu* *in-ə*  
 прежде-ATR.TEMP <неясная форма> жить-PTCP.NEG  
 ‘Ведь безоленный эвенк не мог раньше жить’ (Малые языки 2012–2021, Кислокан).
- б. *huru-w-rə-Ø* *tigə-wər* *mū-taj.*  
 пойти-TR-NFUT-3PL посуда-RFL.PL вода-ATR.PROPR  
 ‘Унесли посуду с водой.’ (Малые языки 2012–2021, Юкта).
- в. *gē* *gē* *kərgən əpet* ***bišo*** *atirkān*  
*dur* *otolgī-ši*  
 два сын-ATR  
 ‘Вторая семья тоже была, старуха с двумя сыновьями’ (Малые языки 2012–2021, Чиринда).
- (13) а. [*Эвэнки-л ачи-р-ду-тын*] [*ачи-р-ду-тын* *оро-р*],  
 эвенк-PL NEG.EX-PL-DAT-PS.3PL NEG.EX-PL-DAT-PS.3PL олень-PL  
 ‘Без эвенков, без оленей...’ {экспедиция не могла бы сделать столько работы} (КИЭА РАН 2011–2017, газета).
- б. *oro-r-lon* *nulgi-solā-tin*  
 олень-PL-ATR.PROPR кочевать-CVB.ANT-3PL  
 ‘С оленями аргишили’... {сюда в каком году приехал?} (Малые языки 2012–2021, Эконда)
- в. *ollo-l-o* *kətə-kōkūn-mə* *wā-rə-w* *amī-nən*  
 рыба-PL-PART много-INTS-ACC добыть-NFUT-1PL(EXCL) отец-COM.FAM  
 ‘Рыбы много добываем с отцом’ (Малые языки 2012–2021, Полигус).

В этой статье была коротко рассмотрена модель управления экзистенциального глагола *bi-* и ее особенности, а также значение глагола *bi-*: было проведено сравнение экзистенциального *bi-* и его «отрицательного аналога» предикативного слова *ač̣in* 'NEG.EX'. Обладая одинаковой моделью управления (тема + местоположение/ обладание), *bi-* и *ač̣in* различаются по сфере употребления. В подавляющем большинстве случаев *bi-* выражает отношение местонахождения, в то время как *ač̣in* выражает одинаково часто оба отношения. Это связано с наличием в эвенкийском языке множественных показателей обладания, так что имена с этими показателями преобладают, по крайней мере в устной речи, и употребляются вместо конструкции с *bi-*.

### Глоссы

1 – 1-е лицо, 3 – 3-е лицо, ACC – аккузатив, ANT – (причастие) предшествования, ALL – аллатив, ATR – атрибутив, ATTEN – уменьшительность, DAT – датив, DEAD-REL – умерший родственник, DYN – динамический, EX – экзистенциальный, EXCL – эксклюзивное множественное число, FOC – энклитика (частица), маркирующая фокус или коммуникативное выделение, FUT – будущее время, FUTCNT – ближайшее будущее время, HAB – хабитуалис; INCH – инхоатив, INTS – интенсификатор, IPFV – имперфектив, LIM – лимитатив, NEG – отрицание, NFUT – небудущее время, PART – партитив, PF – перфектный конверб, PL – множественное число, PRS – настоящее время, PROPR – обладатель отчуждаемой принадлежности, PS – посессивный аффикс, PST – прошедшее время, PSTITER – прошедшее итеративное время, PTCP – причастие, RFL – рефлексив, SG – единственное число, TEMP – темпоральный атрибутив, TR – показатель транзитивности.

### Источники

КИЭА РАН 2011–2017 – *Корпусы Института этнографии и антропологии РАН.*

URL: <http://corpora.ica.ras.ru/corpora/texts.php>. Дата обращения 10.05.2021.

Малые языки 2012–2021 – *Малые языки Сибири: наше культурное наследие.*

URL: <http://minlang.srcs.msu.ru/>. Дата обращения 15.05.2021.

Поиск – URL: <http://gisly.net/corpus/>. Дата обращения 20.05.2021.

### Литература

1. Апреян Ю. Д. *Лексическая семантика (синонимические средства языка)*. М., 1974.
2. Арутюнова Н. Д. *Предложение и его смысл: логико-семантические проблемы*. М., 1976.
3. Болдырев Б. В. *Эвенкийско-русский словарь*. Новосибирск, 2000. Ч. 1, Ч. 2.
4. Болдырев Б. В. *Морфология эвенкийского языка* / Ред. Роббек В.А. Новосибирск, Наука, 2007.
5. Булатова Н. Я. *Эвенкийский язык в таблицах*. Спб., 2002.

6. Василевич Г. М. *Очерки диалектов эвенкийского (тунгусского) языка*. Л., 1948.
7. Клячко Е.Л. Диалектные особенности публикуемых текстов. Предисловие к сборнику: *История страны в рассказах о жизни кетов, селькупов и эвенков* (в печати) / О.А. Казакевич, Е.М. Будянская, Ю.Е. Галямина, Е.Л. Клячко, Н.К. Митрофанова.
8. Константинова О. А. *Эвенкийский язык. Фонетика. Морфология*. М., 1964.
9. Мельчук И.А. *Опыт теории лингвистических моделей «Смысл ↔ Текст»*. М., 1974.
10. Певнов А. М. Лингвистические свидетельства исторических контактов ороков и орочей на Сахалине // *Acta Linguistica Petropolitana. Труды Института лингвистических исследований*. 1918. Т.14. №1. С. 441–463.
11. Певнов А. М. Формы со значением объектного обладания/ необладания в эвенкийском языке // Е.В. Головкин, А.М. Певнов, М.Ю. Пупынина, А.А. Сюрюн, А.Ю. Урманчиева, ред. *Лингвистика и другие важные вещи: памяти Александра Павловича Володина (1935–2017)*. Санкт-Петербург, 2019. С. 302-339.
12. Рудницкая Е. Л. Отрицательная экзистенциальная конструкция с *āchin* в эвенкийском языке: внутренняя синтаксическая структура // *Урало-алтайские исследования* (в печати).
13. Суник О. П. *Глагол в тунгусо-маньчжурских языках: Морфологическая структура и система форм глагольного слова*. Л., 1962.
14. Шаляпина З.М. *Трехмерная стратификационная модель языка и его функционирования (К общей теории лингвистических моделей)*. М., 2007.
15. Babby, Leonard. *Existential Sentences and Negation in Russian*. Ann Arbor (MI), 1980.
16. Baek, Sangyub. *Grammatical peculiarities of Uilta in the third group of Tungusic*. Доклад на II Международной конференции по уральским, алтайским и палеоазиатским языкам. СПб (онлайн), 08–10.2020.
17. Bulatova, Nadežda, Grenoble, Lenore. *Evenki // Languages of the world/ Material 141*, München, 1999.
18. Hazout, Ilan. The syntax of existential constructions // *Linguistic Inquiry*. 2004. № 35 (3). P. 393–430.
19. Matushansky, Ora. *Missing types*. Доклад на конференции Caritive constructions in the languages of the world. ИЛИ РАН, СПб (онлайн). 31.11.2020 – 02.12.2020.
20. Nedjalkov, Igor. *Evenki*. London / New York, Routledge, 1997.
21. Partee, Barbara, Borschev, Vladimir. Genitive of negation and scope of negation in Russian existential sentences // Toman J. ed. *Annual Workshop on Formal Approaches to Slavic Linguistics: the Second Ann Arbor Meeting 2001 (FASL 10)*. Michigan, 2002. P. 181–200.
22. Wagner-Nagy, Beata. *On the Typology of Negation in Ob-Ugric and Samoyedic Languages*. Helsinki, 2011.
23. Yeon, Jaehoon, Brown, Lucien. *Korean. A comprehensive grammar*. London – New York, 2011.

## Cigogo Nominal Demonstratives: Morphology and Semantics

**Stanislav B. Beletskiy**

National Research University Higher School of Economics, Moscow, Russia;

University of Dodoma, Dodoma

stas.beletskiy@gmail.com, ORCID 0000-0003-3372-1480

**Abstract.** This paper presents the system of nominal demonstratives in Cigogo language spoken in Tanzania. The system includes 4 sets of forms build by reduplication of the pronominal prefix that is further suffixed with formants *no*, *o*, *lya*. Besides the special deictic (proximal, medial, distant to/from speaker/hearer) and anaphoric functions, they fulfill a wide range of discourse functions: anaphoric, cataphoric, recognitional, transposed.

**Keywords:** Cigogo, demonstrative, deixis, anaphora, word-formation

**Аннотация.** В статье представлена система указательных местоимений языка гого (Танзания). Система включает в себя 4 серии форм, образованных путем дубликации прономинального префикса, с последующей суффиксацией формантами *no*, *o*, *lya*. Помимо пространственно-дейктических (близость к говорящему/слушающему, удаленность от говорящего/слушающего) и анафорических функций указательные местоимения выполняют ряд дискурсивных функций: анафора и катафора в нарративе, введение новой информации, метареферентация и др.

**Ключевые слова:** язык гого, указательные местоимения, дейксис, анафора, словообразование

### 1. Introduction

Gogo (sw. Kigogo, autonyme: Cigogo) is a Bantu language classified as G11 by both Guthrie [Guthrie 1971] and Maho [Maho 2009] and as ‘gog’ by ISO 639-3 [Ethnologue]. The language is spoken by 1,080,000 speakers who live in Central Tanzania in Dodoma Region and the neighboring districts of Iringa, Manyara and Singida regions [Atlasi ya lugha za Tanzania 2009]. The language consists of six dialects spoken in different parts of *Ugogo* (Gogo land): Cinyamaseya (north), Cinyambwa (west), Citumba (east), Cinyatakama, Kozisamba (south), Cinyaugogo/Ngogo (core), the latter is being considered the superior dialect (‘pure’ Cigogo) as the Bible was translated into it. Speakers of all dialects have mutual understanding as the varieties differ mainly in tonal representations and vocabulary [Balisidya 1977].

Cigogo posses 5 contrastive vowels: [i, e, a, o, u] and two tones (high and low).

The Cigogo consonant inventory includes voiced and voiceless plosives: [p/b, t/d, c/j, k/g], voiced and voiceless fricatives: [f/v, s/z, h], voiced and voiceless nasals: [m/ṃ, n/ṅ, ɲ/ṅ̃, ŋ/ṅ̃], and approximants [l, j, w]. In writing, the voiceless nasals are represented with di- and trigraphs <mh, nh, nyh, ng’h>.

The preferred syllable type is CV. The collision of vowels on morphemic boundaries is resolved by a number of assimilatory strategies (gliding, merger, deletion, lengthening) which are summarized in the Table 1 below.

<i>V1</i> \ <i>V2</i>	<i>i</i>	<i>e</i>	<i>a</i>	<i>o</i>	<i>u</i>
<i>i</i>	iyi	ye	ya	yo	yu
<i>a</i>	e	e	a(a)	ao	au
<i>o</i>	wi	we	wa	o(w)	o(w)
<i>u</i>	wi	we	wa	wo	wu

Table 1. Vowel Assimilation in Cigogo  
(modified from [Großerhode 1997: 21])

The assimilatory processes observed among consonants are:

- palatalization of the velars:  $g + i = ji$ ;  $k + i = ci$ , except of the reflexive infix -ki-;
- destopping of the voiceless stops after a nasal that results into voiceless (aspirated) nasals:  $N + t = nh$ ,  $N + p = mh$ ,  $N + c = nyh$ ,  $N + k = ng'h$ ;
- full spirantization after the super high \*i and \*u.

Cigogo shows a typical for Bantu languages of the zone G system of nominal classes that is summarized in the Table 2 below.

Nominal prefix in both nouns and adjectives can be augmented by preprefix *u-*, if the prefix contains this vowel; otherwise it is augmented by preprefix *i-*.

Concord Class \	Nominal prefix	Subject marker	Object marker	Pronominal prefix	Possessive -a
1	mu	a	mu	yu	wa
2	wa	wa	wa	wa	wa
3	mu	u	u	u	wa
4	mi	i	i	i	ya
5	(l)i	li	li	li	lya
6	ma	ga	ga	ga	ga
7	ci	ci	ci	ci	ca
8	vi	vi	vi	vi	vya
9	N	i	i	i	ya
10	N/zi	zi	zi	zi	za
11	(l)u	lu	lu	lu	lwa
12	ka	ka	ka	ka	ka
14	(w)u	u	u	wu	wa
15	ku	ku	ku	ku	kwa
16	ha	ha	ha	ha	ha
17	ku	ku	ku	ku	kwa
18	mu	mu	mu	mu	mwa

Table 2. Cigogo Nominal Concorde  
(modified from [Rugemalira 2013:29] and [Großerhode 1997: 26])

## 2. Research Background

According to the “Electronic Bibliography for African Languages and Linguistics” [Maho 2009], Cigogo has gained considerable attention from scientists. The bibliography includes 45 publications dedicated to both linguistic (grammar descriptions, vocabulary lists and dictionaries, an exercise book, a collection of fairy tales, the Bible), and non-linguistic topics (lists of animals, birds and plants; publications about culture and history). However, there are only two grammars of Cigogo – a prescriptive [Cordell 1941] and a descriptive [Rugemalira 2013]. The most considerable difference between them consists in the representation of the system of demonstratives. Let us compare the models suggested by the two authors.

Cordell distinguishes 4 sets of demonstrative adjectives and pronouns. The set 1 consists of the forms resulting from the reduplication of the pronominal prefix in all nominal classes. The general semantics of the set is proximity to the speaker. The set 2 is derived from the set 1 by suffixing *-hya* in all classes with the general meaning of remoteness from the speaker. The set 3 is made from set 1 by changing the final vowel to *-o* expressing the idea of referring to something before mentioned. Set 4 is made from set 1 by suffixing *-no*. “These are used e.g. to announce the reappearance of something that has been lost” [Cordell 1941: 57]. All demonstratives can be shortened by replacing the first syllable with *a-*. “The short form tends to be used when the Demonstrative is used as a pronoun (by itself) <...>” [Cordell 1941: 25].

Rugemalira goes further in classifying Cigogo demonstratives into 5 types based on their ability to show different levels of nearness / remoteness [Rugemalira 2013: 35-36]. Type 1 is morphologically equal to set 1 by Cordell and is used by the speaker to refer to something that s/he can touch. Type 2 is morphologically equal to set 4 and is used by the speaker to refer to something that is close to him/her but out of reach. Type 3 is morphologically equal to full forms of set 3 and is used by the speaker to refer to something that is far from him/her but close to the hearer or to refer to something before mentioned. Type 4 is morphologically equal to short forms of set 3 and is used by the speaker to refer to something that is far from both him/her and the hearer but is in view to both of them. Type 5 is morphologically equal to set 2 and is used by the speaker to refer to something that is far from both him/her and the hearer and is not in view to neither of them. Further he notes that it is common for native speakers to merge semantically types 1 and 2, and 3 and 4.

The overview shows that both systems describe the basic uses of Cigogo demonstratives (special-deictic) and do not cover extended uses. Moreover, splitting the demonstratives with *-o* into two semantically different types by Rugemalira suggests the need for thorough investigation of the semantic differences among full and short forms. Furthermore, both systems are not situated into the broader typological contexts – the general theory of demonstratives and demonstratives in Bantu languages. The present paper aims at filling in these gaps.

## 3. Theoretical Framework

### 3.1. Cross-linguistic overview of Demonstratives

Cross-linguistically a demonstrative is defined as “any item, other than 1st and 2nd person pronouns, which can have pointing (or deictic) reference” (Dixon 2003: 61-

62). Demonstratives build a closed grammatical system and overlap in their functions with other systems (articles, personal pronouns, interrogatives, etc.). There are four frequent types of demonstratives: nominal, local adverbial, verbal, manner adverbial.

A nominal demonstrative refers to an object and can form a NP with a noun or pronoun (*this stone is hot*) or a complete NP (*this is hot*). Nominal demonstratives may interact with the system of personal pronouns substituting or co-occurring with 3rd person. In some languages nominal demonstratives may be connected with definite articles synchronically and diachronically like it is the case with German forms *die* (F and PL), *der* (M), *das* (N) which express definiteness when unstressed and reference when stressed. In languages with noun classes nominal demonstratives are often marked with class concords.

Local adverbial demonstratives point to a place and occur alone (*put it there*) or with a noun taking a local marking (sw. *nyumbani kule* ‘there at home’). Universally languages possess at least two forms of local adverbials contrasting special relations (here/there) or visibility.

Verbal demonstratives are subclass of verbs deictically referring to an action. They can occur as a self-standing predicates or in combination with a lexical verb (‘do it like this’).

Manner adverbial demonstratives are derived from nominal demonstratives and function as modifiers to verbs (‘do in this way’).

The most frequent functions of demonstratives include deictic reference, syntactic function, anaphora and cataphora, identification, new information and discourse organization.

Deictic reference can be speaker-anchored or addressee-anchored and encode radial spatial categories (proximal (here), medial (there), distal (over there)), height and stance (close/far to/from speaker and on the same level, higher or lower than the speaker), visibility (visible, near me; visible, near him; invisible, near me, etc.).

Syntactic functions of demonstratives may vary from language to language allowing them to appear in any core or peripheral function in a clause.

Anaphora is a reference to something earlier in the discourse. Cataphora is a reference to something later in the discourse.

Demonstratives used for identification or recognition are usually accompanied with a relative clause (*It was that sort of gluggy rice which the Japanese go in for*).

Demonstratives also can mark new information in discourse (*There’s this new girl at school today*) and organize it.

Using slightly different terminology, Levinson suggest the following taxonomy of different uses of demonstratives (see Figure 1).

Exophoric uses can be accompanied with gestures (gestural use) or used without gestural component (symbolic use). Gestural use can be contrastive (comparing two items) and non-contrastive. A special type of exophoric use of demonstratives is their transposed use when the deictic ground (origo) is transposed from the moment of speech to other time or place (in a narrative context) or another perspective – to the addressee.

Discourse deictic use is the use of demonstrative to refer to a chunk of discourse itself (*Bloop! It sounded like that*).

Non-deictic uses include anaphora and cataphora as references to something mentioned or projected in discourse, empathetic (*That goddamn son of a bitch*) and recognitional (*Do you remember that wonderful day in Morocco?*).

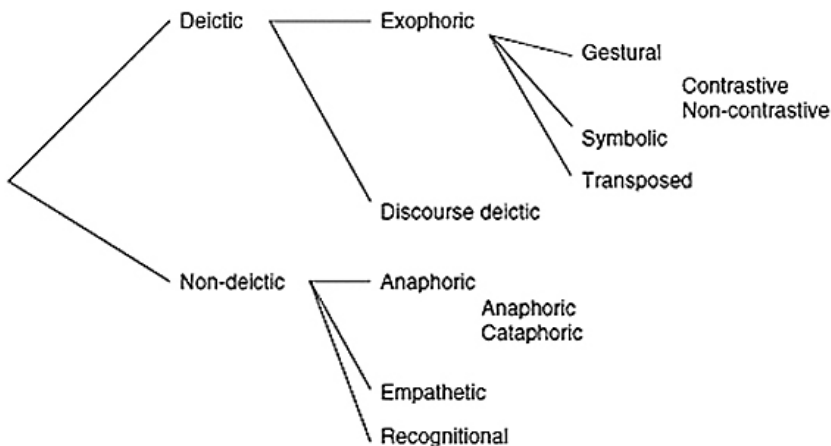


Figure 1. The Distinct Uses of Demonstratives  
[Levinson 2018: 10]

### 3.2. Demonstratives in Bantu Languages

Bantu languages show nominal demonstrative systems with up to 5 morphologically and semantically differing sets. Accordingly, the typology of these forms is based on the assumption that the Proto Bantu possessed up to 5 sets of nominal demonstratives which were preserved or blurred in modern-day Bantu languages. Weier (1995: 22) refers to Meeussen (1967), among other authors, to exemplify the hypothetical sets of Bantu nominal demonstratives:

- [1] “PP ‘that, the’ (a weak anaphoric)
- [2] V-PP (LL?) ‘this’ (as opposed to PP-día)
- [3] PP-nóo ‘this (near me)’ (as opposed to V-PP-o)
- [4] V-PP-o (LLL?) ‘that (near you)’, ‘that (mentioned)’
- [5] PP-día ‘that (over there)’.”

Further, Weier agrees with de Blois (1970) about the nature and functions of the augment (V). It must have developed from the weak anaphoric demonstrative; synchronically it stays in complementary distribution with nominal demonstratives: if a demonstrative is used before the noun the augment is omitted, but if it used after the noun the augment reappears.

Based on this assumption, Weier models the development of the demonstrative forms in the following way: the initial weak form [1] was reduplicated to generate speaker-anchored proximal forms [2], reduplicated and suffixed with -nóo, -o, -día to generate speaker-anchored medial forms [3], addressee-anchored proximal/medial and discourse anaphoric forms [4], as well as distal forms [5].



The augment in its development underwent several stages and formed systems of 5 types:

- 1) closed symmetrical 3-item system, in which the augment is equal to PP, thus possesses the CV-structure;
- 2) open asymmetrical 3-item system, in which the augment loses the first consonant and lowers one or more of its high vowels (e-a-u, e-a-o, i-a-o);
- 3) 2-item system, where the system of type (2) erodes further while substituting a lower vowel with a higher one (i-i-u, e-o-o, i-a-i, e-a-e);
- 4) ‘stark’ 1-item system, where the system of type (3) is reduced to the predominant vowel;
- 5) latent (absent) augment.

#### 4. Data Collection and Analyses

This research is based on the collection of mixed types of data:

- 1) 12 questionnaires about morpho-syntactic properties of Bantu languages [Batibo 1990];
- 2) 50 traditional narratives (fairy tales) collected throughout 2015-2020 in Dodoma region;
- 3) 9 focus-group discussions about the use of medical plants recorded and transcribed in 2019;
- 4) published data [Cordell 1941; Rugemalira 2013].

All pieces of data were migrated to SIL FieldWorks Explorer for annotation and creation of corpus. The corpus consists of 7040 nominal tokens. The annotated corpus was used to generate concordances of all types of demonstratives which were analyzed in terms of both morphology and semantics.

#### 5. Results

The data shows that Cigogo possesses 4 morphologically and semantically contrastive types of nominal demonstratives, including local adverbials and manner adverbials that fulfill a wide range of deictic and non-deictic functions. However the data does not attest verbal demonstratives which are rare cross-linguistically.

Cigogo possess a complex system of noun classes, and nominal demonstratives obligatorily mark classes of the nouns they refer to through the use of pronominal prefixes. The overview of all attested forms is given in the Table 3 below. It should be noted that the table includes surface realizations of deep structures which are governed by the phonetic processes described in the Introduction and by pragmatic context.

Class	PP	<i>DEMI</i>	<i>DEM2</i>	<i>DEM3</i>	<i>DEM4</i>
1	<i>yu</i>	yuyu ayu	yuyuno ayuno	yuyo ayo	yuyulya (a)yulya
2	<i>wa</i>	wawa awa	wawano awano	wao ao	wawalya (a)walya
3	<i>u</i>	wuwu au	wuwuno auno	(w)uwo ao	wuwulya (a)wulya
4	<i>i</i>	iyi ai	iyino aino	(y)iyo ayo	iyilya (a)yilya

5	<i>li</i>	lili ali	lilino alino	lilyo alyo	lililya (a)lilya
6	<i>ga</i>	gaga aga	gagano agano	gago ago	gagalya (a)galya
7	<i>ci</i>	cici aci	cicino acino	cico aco	cicilya (a)cilya
8	<i>vi</i>	vivi avi	vivino avino	vivyo avyo	vivilya (a)vilya
9	<i>i</i>	iyiyi ai	iyiyino aino	iyoyo ayo	iyiyilya (a)ilya
10	<i>zi</i>	zizi azi	zizino azino	zizo azo	zizilya (a)zilya
11	<i>lu</i>	lulu a(l)u	luluno a(l)uno	lu(l)o a(l)o	lululya (a)lulya
14	<i>u</i>	wuwu a(w)u	wuwuno awuno	wuwo ao	wuwulya (a)ulya
15	<i>ku</i>	kuku aku	kukuno akuno	kuko ako	kukulya (a)kulya
16	<i>ha</i>	baha aha	hahano (a)hano	baho aho	bahalya (a)halya
17	<i>ku</i>	kuku aku	kukuno (a)kuno	kuko ako	kukulya (a)kulya
18	<i>mu</i>	mumu amu	mumuno (a)muno/munu	mumo amo	mumulya (a)mulya

Table 3. Overview of Cigogo Nominal Demonstrative Forms

The demonstratives of locative classes 16, 17, 18 are locative adverbials (here, there, over there, inside, etc.). The demonstratives of classes 7 and 8 can be used as manner adverbs (this/that way). In this paper, we do not analyze adverbials because these subtypes require a separate piece of research.

### 5.1. Morphology of Cigogo Nominal Demonstratives

According to their morphology and deictic uses, nominal demonstratives in Cigogo can be divided into 4 distinct types which reflect the hypothetical Proto Bantu system. The Meeussen's *weak anaphora* type (bare PP) is not included in the Table 4 because in our data it functions as existential (copula).

<i>Type</i>	<i>Deixis</i>	<i>Proximal</i>	<i>Medial</i>	<i>Distal</i>
	<i>Structure</i>			
<i>DEM1</i>	<i>PP+PP</i>	speaker-anchored	-	-
<i>DEM2</i>	<i>PP+PP+no</i>	-	speaker-anchored	-
<i>DEM3</i>	<i>PP+PP+o</i>	addressee-anchored		visible
<i>DEM4</i>	<i>PP+PP+lya</i>	-	-	invisible

Table 4. Types of Nominal Demonstratives in Cigogo



- (4b) *A-ku-no*                      *ku-nyuma*    *Ndevu*    *y-a-lil-a*  
 AUG-CL17-DEM2    CL17-back    beard    3SG-PST-cry-FV  
 ‘There back [at Hombolo] Beard cried [when he realized that his friend went to Dar es Salaam]’.

#### 5.2.4. Symbolic use

The examples (5) and (6) show that in some contexts (symbolic use) full and short forms (DEM1) are interchangeable.

- (6) *M-simo*    *a-u*                                      *u-ku-ci-lajiz-a*  
 CL3-story    {AUG-CL3}DEM1    CL3-PRST-OM\_1PL-teach-FV  
 ‘This story teaches us’.

- (7) *I-simo*    *yi-yi*                                      *i-ku-ci-lanjiz-a*  
 CL9-story    {CL9-CL9}DEM1    CL9-PRST-OM\_1PL-teach-FV  
 ‘This story teaches us.’

#### 5.2.5. Discourse deictic use

The difference in discourse deictic use of DEM2 and DEM4 in (7) and (8) is explained by the parameter ‘visibility’ and ‘proximity’. In (7) the referent is visible and near but is treated as a before-mentioned piece of information; in (8) it is imaginable thus it is far (in the forest) and not visible.

- (7) *Z: A-u*                                      *mkwata*  
 {AUG-CL3}DEM1    *mkwata*  
 ‘Z: This *Strychnos innocua*’.  
*Y: Mkwata*    *a-u-no*                                      *ci-ka-u-tumi-ag-a*  
 Mkwata    AUG-CL3-DEM2    2PL-PST-OM\_ CL3-use-HAb-FV  
 ‘Y: This *Strychnos innocua* we use [as food]’.

- (8) *Zi-kw-injil-a*                      *mu-m-hango*                      *y-a*                      *m-pela*  
 CL10-PRST-enter-FV    CL18-CL9-hollow    CL9-of                      CL9-baobab  
*u-laza*                      *bit-a*                      *ba-ha-dya*  
 2SG-need    go-FV    CL16-CL16-DEM4  
 ‘When they [bees] come into the hole in a baobab tree go there [to collect honey]’.

#### 5.2.6. Anaphoric use

Anaphoric function is fulfilled by DEM1, DEM3, and DEM4 although only DEM3 (-o) is believed to be a basic device for this function. DEM1 seems to mark transposed use (as if the speaker and the hearer were involved in the story), while DEM4 marks the story as a narrative distant from both the speaker and the hearer. The difference between DEM1 and DEM2 may be explained by the iconicity principle: in (10), DEM1 is used in the immediate context (next sentence), while in (9), DEM3 indicates distance between the referent and the anaphora (the referent is situated in the first sentence and the anaphoric expression is found in the last sentence of the story). The difference in DEM4 structures (latent augment in (11) and closed symmetric in (12)) may be again explained by the iconicity principle – but in terms of the size of the referent: a small chick vs. a large group of 25 people.

- (9) *Si-y-a-sum-a*                      *langa ku-lawilila Ø-siku*      *yi-y-o*  
 NEG-3SG-PST-can-FV see      CL17-start CL9-day      CL9-CL9-DEM3  
 ‘It [Snake] couldn’t see starting from that day [when it was invited to a party]’.
- (10) *Wa-ka-w-a*                      *lel-a*                      *wa-datu*      *w-a*                      *ci-c-ekulu.*  
 3PL-PST-be-FV      give\_birth-FV      CL2-three      CL2-of      CL7-CL7-female  
*Wa-na*                      *wa-wa*                      *wa-ka-wiluk-a*  
 CL2-child      {CL2-CL2}DEM1      CL2-PST-gow-FV  
 ‘... and they gave birth to three daughters. These children grew up’.
- (11) *Ny-ina*                      *I-m-wewe*                      *y-a-z-a mala*  
 CL9-mother      AUG-CL9-hawk      3SG-PST-come-FV      finish  
*ku-mu-dy-a*                      *mw-ana*                      *yu-dya*  
 INF-OM\_CL1-eat-FV      CL1-child      CL3-DEM4  
 ‘[After] Mother Hawk finished eating that child [who was left alone]’.
- (12) *W-a-wil-il-a*                      *ba-ha*                      *i-ka-w-a*  
 3PL-PST-come-APL-FV      {CL16-CL16}DEM1      CL9-PST-be-FV  
*ku-z-a*                      *wa-funik-a*                      *wa-nhu*  
 INF-come-FV      OM\_3PL-cover-FV      CL2-person  
*w-ose*      *wa-wa-lya*                      *ishirini*      *ne [na=i]*      *Ø-hano*  
 CL2-all      CL2-CL2-DEM4      twenty      and-AUG      CL9-five  
 ‘When they came down [water] covered all of those 25 people [who were in the cave]’.

#### 5.2.7. Cataphoric use

DEM1 has the cataphoric function while marking the projected piece of discourse.

- (13) *A-ku*                      *y-a-ki-tow-a*                      *ma-gulu*      *g-akwe*  
 {AUG-CL17}DEM1      3SG-PST-REFL-beat-FV      CL6-leg      CL6-his  
 ‘Here [when] he stamped his feet down’.

#### 5.2.8. Recognitional use

Under recognitional use we examine two similar functions: introducing new information (14) and refreshing the common knowledge shared by both the speaker and the hearer, also called metarepresentational use (15). New information is introduced with DEM1. Reference to common knowledge is marked by DEM2 and DEM4. In (15) the speaker is starting a story with the well known reference to the remote (magical) time in the past. The full form of DEM4 *vivilya* might be used to show the great distance between the present moment and the narrative past.

- (14) *A-di*                      *i-pela.*                      *I-pela*                      *a-di-no*  
 {AUG-CL5}DEM1      CL7-guava      CL7-guava      AUG-CL7-DEM2  
*yani*      *di-ku-tumik-a*                      *kila*                      *Ø-sehemu*  
 that\_is      CL7-PRST-use-FV      every      CL9-part  
 ‘This is guava. This guava is used everywhere’.
- (15) *Ba-h-o*                      *vi-lo*                      *vi-vi-lya*  
 CL16-CL16-DEM2      CL8-night      CL8-CL8-DEM4  
 ‘There many nights back [Once upon a time].’

## 6. Conclusion

Cigogo possess a closed system of nominal demonstratives that include four sets of forms distributed across 16 nominal classes. The four types reflect the reconstructed Proto Bantu system and are built with reduplicated pronominal prefixes (DEM1), suffixed by -no (DEM2), -o (DEM3), and -lya (DEM4).

They fulfill a wide range of deictic (primary) and non-deictic (extended) functions. The summary of forms and functions is given in the Table 5.

Type	Deixis Structure	Proximal	Medial	Distal	Extended use
DEM1	<i>PP+PP</i>	speaker-anchored	-	-	symbolic, anaphoric, transposed, recognitional
DEM2	<i>PP+PP+no</i>	-	speaker-anchored	-	discourse deictic, anaphoric, recognitional
DEM3	<i>PP+PP+o</i>	addressee-anchored		visible	discourse deictic
DEM4	<i>PP+PP+lya</i>	-	-	invisible	discourse deictic, anaphoric, recognitional

Table 5. Forms and Functions of Cigogo Nominal Demonstratives

## Resources

Ethnologue – ISO 639-3. URL: <https://iso639-3.sil.org/>. Retrieved 08.07.2021.

## References

1. *Atlasi ya lugha za Tanzania* [Atlas of Tanzanian languages]: Chuo Kikuu cha Dar es Salaam [University of Dar es Salaam], 2009.
2. Balisidya, M. L. *Mabadiliko katika fasihi simulizi ya Wagogo*. Dar es Salaam: The University of Dar es Salaam printing press, 1977.
3. Batibo H.M. *Kidadisi cha maumbo-miundo katika lugha za Kibantu* [Questionnaire on morphological characteristics in Bantu languages]. University of Dar es Salaam, 1990.
4. Blois de, K.F. The Augment in the Bantu Languages // *Africana Linguistica*. 1970. № 4. P. 85–166.
5. Cordell, O. T. *Cigogo grammar* [Manuscript], 1941.
6. Dixon, R. M. W. Demonstratives: A cross-linguistic typology. // *Studies in Language*. 2003. Vol. 27. Issue 1. P. 61–112.
7. Großerhode R. *Ton und Druck im Gogo*. Inaugural-Dissertation zur Erlangung des Doktordgrades. Universität Bayreuth, 1997.

8. Guthrie, Malcolm. *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages, 2: Bantu prehistory, inventory and indexes*. London: Gregg International, 1971.
9. Levinson, S. Introduction: Demonstratives: Patterns in Diversity // S. Levinson, S. Cutfield, M. Dunn, N. Enfield, & S. Meira (Eds.), *Demonstratives in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press, 2018. P. 1–42.
10. Maho Jouni Filip. *Electronic Bibliography for African Languages and Linguistics*. 2009. URL: <http://goto.glocalnet.net/jfmaho/bob.pdf> (retrieved on 16.04.2016)
11. Meeussen, A.E. Bantu Grammatical Reconstruction. I// *Africana Linguistica* 1967. № 3. P. 79–122.
12. Rugemalira, J. *Srufi ya Kigogo*. Dar es Salaam: Dar es Salaam University press, 2019.

## Fronted topics, clefted foci and the parameter of integration into the clause

Yury A. Lander

National research university Higher school of economics, Moscow, Russia  
yulander@yandex.ru, ORCID 0000-0003-1798-7174

**Abstract.** The paper examines the similarities between two families of constructions: fronted nominal topics and clefted nominal foci. These may originate from the fact that both fronted nominal topics and clefted foci dissociate the expression of the information structure and the expression of the semantics associated with the relevant state-of-affair (SOA). These topics and foci can be said to vary both cross-linguistically and diachronically in the extent to which they are integrated into the clause describing the SOA. This approach provides an analysis for some typologically unusual constructions.

**Keywords:** nominal topic, nominal focus, clefting construction, topic dislocation, state-of-affair, information structure

**Acknowledgments:** I am grateful to the audience of the conference “Vostokovednye čtenija – 2021” for discussion. Special thanks to my consultants in Abaza, Tanti Dargwa and West Circassian. All errors are mine. This work is partly based on the research supported by Russian Science Foundation (Grant No. 17-18-01184), based at the Russian State University for Humanities. I dedicate this paper to the memory of Zoya Shalyapina, who was always eager to discuss whatever issue I was interested in.

**Аннотация.** В работе обсуждается сходство между конструкциями с выдвинутыми топиками, с одной стороны, и клефтами и псевдоклефтами, с другой стороны. Это сходство, как кажется, проистекает из того, что все эти явления связаны с диссоциацией описания релевантной ситуации и информационной структуры. Топикализованные и фокализованные группы в описываемых конструкциях могут быть в разной степени интегрированы в клаузу, описывающую релевантную ситуацию, причем эта интеграция имеет в том числе и диахроническое измерение. Такой подход позволяет, в частности, предложить анализ для некоторых типологически необычных конструкций.

**Ключевые слова:** выдвижение топика, клефт, псевдоклефт, релевантная ситуация, информационная структура предложения

**Благодарности:** автор благодарен участникам конференции «Востоковедные чтения 2021» (ИВ РАН) и консультантам по абазинскому, даргинскому и западно-черкесскому языкам. Работа частично основана на исследованиях, проводимых в РГГУ при поддержке Российского научного фонда (грант № 17-18-01184). Статья посвящается памяти З.М. Шаляпиной, которая всегда готова была обсуждать интересующие меня научные проблемы.

### 1. Introduction

In this paper I consider similarities between two families of constructions, namely fronted nominal topics (1) and clefted nominal foci (2).



- (1) *Lee, I've been following his progress very much over the last month.*  
[Radford 2018:41]
- (2) *What he brought was a donkey.* [Higgins 1979:2]

Cross-linguistically these constructions display much variation in morphosyntactic properties. I argue that this variation originates from the fact that both fronted nominal topics and clefted foci manifest dissociation of the expression of the information structure and the expression of the semantics associated with the relevant state-of-affair (SOA) to a different extent.

The structure of this paper is as follows. Sections 2 and 3 briefly present the phenomena we are interested here. The facts described there are widely discussed in literature and should not be taken as something new. Section 4 contains an attempt to interpret the commonalities between fronted topics and nominal foci. In Section 5 I illustrate how the perspective provided here can be useful in analyzing certain curious patterns in Malay/Indonesian, Tanti Dargwa and West Caucasian languages. The last section contains concluding remarks.

## 2. Fronted nominal topics

As is well-known, topics tend to occupy the initial position in the utterance. Thus, Givón [2001:278] mentions “the use of referent fronting in ‘marked’ topicalizing or contrastive constructions that seem to be attested in all languages, regardless of type and degree of word-order rigidity”. In the following example from Japanese, the topic phrase (which takes a dedicated marker) precedes all other constituents:

- (3) Japanese (isolate)  
*[kono uta wa] [jon ga sakkyoku shi, pooru ga*  
 this song TOP John NOM compose do:INF, Paul NOM  
*sakushi shita]*  
 write lyrics:PST  
 ‘This song, John composed, and Paul wrote lyrics.’ [Iwasaki 2013:238]

One can distinguish between three kinds of constructions with fronted topics [Surányi 2016]:

- Hanging Topic Left Dislocation constructions, where the topic shows no obvious syntactic connections (except for linear adjoining) to the syntactically complete sentence that follows it, as in (4), where the topic ‘that day’ is marked by the default absolutive case rather than the oblique case which is expected for a temporal adjunct and hence is not likely to constitute a part of the clause ‘he didn’t do anything’;
- (4) West Circassian (West Caucasian)  
*a mafe-r, a-š' z-jə ə-šá-Ɂ-ep!*  
 that day-ABS that-OBL one-ADD 3SG.ERG-do-PST-NEG  
 ‘That day, he didn’t do anything.’ (elicited)

- Contrastive Topic Left Dislocation constructions, where the fronted topic necessarily coexists with a resumptive element in the subsequent sentence and can show other properties connecting it to this sentence (e.g., case marking), as in (5);

(5) Hungarian (Uralic)

*az édes süteményeket, azokat szereti*  
the sweet cakes.ACC those.ACC likes  
'The sweet cakes, he likes (lit., them).' [Surányi 2016:428]

• Plain Topicalization, where the topic retains its grammatical function within the sentence but appears at the very beginning of the sentence, as in (6).

(6) Russian (Indo-European, Slavic)

*Slona=to ja i ne primetil.*  
elephant.ACC=TOP I.NOM ADD NEG noticed  
'I have not noticed the elephant.' (Ivan Krylov's fable "The inquisitive man", 1814).

Thus, fronted topics can be integrated into the sentence to a different degree. Hanging topics like in (4) are often described as not being parts of the sentence at all (cf. [Lambrecht 2001] for an extensive discussion) and normally get the "most basic" form (e.g., the absolutive form in West Circassian above). Case-matching effects like that shown in (5), where the dislocated topic takes the same case as the resumptive pronoun, suggest that such topics depend on the following sentence syntactically to a greater extent. Nonetheless according to the most typical syntactic approaches, they do not get any role-based grammatical function because their place in the sentence is already occupied by the corresponding pronoun. Finally, in plain topicalization constructions, the topicalized element retains its grammatical function within the sentence.

Several notes are in order here.

First, in the descriptive practice it is very hard or probably even impossible to make clear-cut distinctions proposed above. For instance, not all languages have flagging markers (cases or adpositions) that indicate the relations of a given element to the clause, hence it is sometimes impossible to distinguish between Hanging Topic Left Dislocation and Contrastive Topic Left Dislocation. Further, since in many languages activated discourse referents tend to get a null expression [Ariel 1990, Kibrik 2011] and topic phrases necessarily activate discourse referents, it is not always possible to understand whether a topic has a correlate expression in the subsequent clause. Not surprisingly, some authors divide the range of fronted topics in a different manner. For example, Lambrecht [2001] only distinguishes between dislocation and (plain) topicalization, while López [2016] prefers to contrast between H-dislocation which only includes hanging topics and D-dislocation which is associated with some integration of the topic into the clause reflected in the presence of grammatical constraints.

Second, the presence of a dedicated topic marker does not imply integration of the topic into the clause. For example, the Japanese topic marker *wa* illustrated in (3) is sometimes thought to belong to the same class of "particles" as case markers [Kuno 1973] (even though it can combine with some of them; cf. (7) below), but this alone is not a sufficient argument for the *wa*-marked topic to be considered a part of the following sentence.

Third, it is not the case that any marker that *usually* implies syntactic dependency should be considered as a marker of the syntactic dependency *per se*. In fact, the primary function of many markers which specify role-related semantics may be semantic rather than syntactic (cf. [Lander and Vyunova 2018]). Hence the examples like (7),

where the topic marker combines with a marker specifying the semantic role, should not be taken as evidence for the real integration of the topic into the sentence.

- (7) Japanese (isolate)
- |               |           |            |             |                  |           |             |          |               |
|---------------|-----------|------------|-------------|------------------|-----------|-------------|----------|---------------|
| [ <i>sono</i> | <i>ko</i> | <i>ni</i>  | <i>wa</i> ] | [ <i>jon</i>     | <i>ga</i> | <i>hana</i> | <i>o</i> | <i>okuri,</i> |
| this          | child     | DAT        | TOP         | John             | NOM       | flower      | ACC      | send:INF      |
| <i>pooru</i>  | <i>ga</i> | <i>uta</i> | <i>o</i>    | <i>uttatta</i> ] |           |             |          |               |
| Paul          | NOM       | song       | ACC         | sing:PST         |           |             |          |               |
- ‘To this child, John sent flowers and Paul sang a song.’ [Iwasaki 2013:238]

Fourth, there can be many more kinds of evidence pointing to greater or lesser integration of topics into the clause. For example, according to Rudnitskaya [2010], in Korean topics can bind subject-oriented reflexive pronouns even in configurations where there is a subject in the clause; cf. the following example (still marginal, as indicated by the question mark):

- (8) Korean (isolate)
- ? *Mia-nun nwuna-ka caki hakkyo-eyse ce-yil yeypu-ta*  
 Mia-TOP sister-NOM RFL school-LOC most beautiful-DCL  
 ‘As for Mia, her sister is the most beautiful in her (Mia’s, \*sister’s) class.’  
 [Rudnitskaya 2010:125]

Such syntactic properties suggest greater integration of topics into sentences (see also [Her 1991; Alsagoff 1992a; 1992b] *inter alia* who suggested that they topics be interpreted as grammatical functions on a par with subjects, objects, etc.). In fact, the integration of the topic into the clause may have many more dimensions, so the above-mentioned trichotomy described by Surányi [2016] may be just a simplification.<sup>1</sup> This has been recently emphasized by Andrason et al. [2016], who argued that the very concept of left dislocation has a radial structure involving a prototype and manifold deviations from it.

### 3. Clefted nominal foci

Lambrecht [2001:1066] states that “[s]ince focal denotata are by definition communicatively indispensable elements of propositions, and since propositions are expressed in clauses, focus constituents by necessity occur clause-internally”. This wording is not accurate enough, though, since many languages allow the focused element to appear outside of the clause which describes the relevant SOA. This is observed in clefting constructions including clefts (9) and pseudoclefts (10):

- (9) Cleft
- It was [this]<sub>FOCUS</sub> that caused it.*

<sup>1</sup> Lambrecht [2001:1050] mentions four criteria important for recognizing prototypical dislocation constructions, namely (i) extra-clausal position of a constituent, (ii) possible alternative intra-clausal position, (iii) pronominal coindexation, (iv) special prosody. At the same time, he notes that only the first criterion is a necessary condition for qualifying a construction as an instance of dislocation, hence other criteria constitute independent parameters which probably reflect the integration of a topic into the clause.

(10) Pseudocleft

*What caused it was [this]<sub>FOCUS</sub>.*

Pseudocleft constructions, which seem very widespread at least in some linguistic areas, exploit the natural association of the predicate and the focus: the relevant SOA is expressed by a relative clause that constitutes the subject and the focus formally constitutes the predicate. Cleft constructions, which presumably originate from the same association, probably result from more complex processes, so I will not discuss them here, although my general ideas hold for them as well.

As first noticed by Higgins [1973] for English and later discussed for a number of different languages, foci in such constructions often show connectivity effects, acting in some respects as if they are parts of the relative clause. For example, in (11) and (12) the focus takes an accusative marker which is expected if it is governed by the verb in the relative clause. This phenomenon is akin to case-matching effects observed with some dislocated topics.

(11) Hebrew (Afroasiatic, Semitic)

*ma še-kaninu ba-šuk [ze et ha-sveder ha-kaxol]<sub>FOCUS</sub>*  
what that.we.bought in.the.market COP ACC the-sweater the-blue  
'What we bought in the market was the blue sweater.' [Heller 2002:244]

(12) Russian (Indo-European, Slavic)

*kogo bespolezno iskat' v Sevilje, tak éto*  
who.ACC useless look.for.INF in Seville.LOC so this.N  
*cirjulnika*  
barber.ACC  
'It is a barber for whom it is useless to look in Seville.' (RNC)

Connectivity effects are not restricted to case marking. For instance, foci in clefts and pseudoclefts can be constituted by reflexive pronouns that are normally bound by subjects. Importantly, in these structures reflexive pronouns take antecedents from inside relative clauses rather than are bound by their formal subjects (i.e. headless relatives); cf. (13). I consider this unexpected binding to be parallel to the unexpected binding with fronted topics discussed above.

(13) Abaza (West Caucasian)

*ʕaləj awəj a-tʒə z-zə-j-č'pa-z*  
Ali that DEF-house REL.IO-BEN-3M.ERG-make-PST.NFIN  
*[j-qa akʷa-p]<sub>FOCUS</sub>*  
3M.IO-head 3N.COP-DCL  
Lit., 'For whom Ali built that house is himself.' (elicited)

Effects of this kind do not necessarily come together (cf. [Heller 2002]), hence different clefts/pseudoclefts may differ in this respect. If we interpret the connectivity effects as signs of integration of the focus into the clause describing the relevant SOA, this picture turns out to be of the same nature as topic integration into the clause.

## 4. Information structure vs SOA-related semantics

It is convenient to contrast between the SOA-related semantics and the information structure. Quite often, the choice of a construction is motivated by considerations related to them both. A canonical example is presented by voice contrasts, which cannot be described without appeal to role semantics but at the same time are widely used for topicalization of some argument. Still, the expression of the SOA-related semantics and the information structure can be dissociated, even if such patterns look more marked (i.e. they are structurally more complex and are more restricted in use, at least in some languages). That is what happens in constructions with dislocated topics and clefted foci: there is a clause that provides a basic description of the relevant SOA and there is an external position which is responsible for the information structure formation.

The constructions showing dissociation between the SOA-related semantics and the information structure are also more complex for processing, because they require an additional step to create the semantic links between a separated element and the description of a SOA. Hence we expect the development of separating constructions into more integrated ones, which are more optimal for communication because they recover role relations in a more direct manner. This is indeed what we observe both for topics and for foci. Indeed, this idea is fully consistent with the well-known ideas of grammaticalization of topics into subjects [Shibatani 1990] on the one hand and on the development of monoclausal constructions from clefts and pseudoclefts [Heine and Reh 1984:181–182; Harris and Campbell 1995:151–168; Aannestad 2021:20–21] on the other hand. Given that the effects of integration are numerous and presumably need not strongly correlate with each other, it is likely to be gradual rather than abrupt. As a result, we expect to see constructions that display only partial integration of foci or topics into the clause describing the relevant SOA.

In fact, we can view the occurrence of connectivity effects on a par with formal separation as manifestation of this partial dissociation discussed here. Below we will look at more constructions which present challenges for a synchronic analysis but can be regarded as the results of partial integration of topics and foci into the clause.

## 5. More examples of partial dissociation

### 5.1. West Caucasian internally-headed clefts

When expressing the narrow focus, West Caucasian languages actively use pseudocleft constructions like (13) above and (14), where the backgrounded SOA is described by a non-finite relative clause and the focused part appears as the predicate (sometimes accompanied by a copula); see, e.g., [Sumbatova 2009] for West Circassian, [Kindlein 2016] for Kabardian. Interestingly, even then the focused constituent can appear linearly embedded into the relative clause (15) (see also [Kindlein 2016] for an extensive discussion of this construction in Kabardian). For this pattern I will use the term “internally-headed pseudocleft” (by analogy with internally-headed relative clauses) and contrast it with externally-headed pseudoclefts like (14).

(14) West Circassian (West Caucasian)

<i>tə-š'ə-zə-ka-ʔe-re-r-jə</i>	<i>[gʷəke-r a.rə]</i> <sub>FOCUS</sub>
1PL.ABS-LOC-REL.ERG-CAUS-be-DYN-ABS-ADD	hope-ABS 3.COP
‘What makes us live is (our) hope.’ (WCC)	

## (15) West Circassian (West Caucasian)

*adəge-q<sup>w</sup>aje-r* [*adəge-rjespwəbljəke-r a.rə*]<sub>FOCUS</sub>  
 Circassian-cheese Adyghe-republic-ABS 3.COP  
*qə-zə-š<sup>w</sup>-a-šə-š<sup>w</sup>tə-r*  
 C1SL-REL.IO-LOC-3PL.ERG-do-FUT-ABS

‘It is in Republic of Adyghe where the Circassian cheese will be made.’ (WCC)

Both constructions show connectivity effects. Yet, internally-headed pseudoclefts seemingly display more connectivity effects than externally-headed pseudoclefts. In Abaza, for example, only the former allow the focus to take an its instrumental marker defined by its role in the relative clause (which is by the way also marked by a prefix within the predicate of the relative clause), as demonstrated in (16):

## (16) Abaza (West Caucasian)

a. *acajnka z-la-sə-r-gələ-wa* [*arəj a-ruč<sup>w</sup>’ka(\*-la)*  
 mark REL.IO-INS-1SG.ERG-CAUS-stand-IPF this DEF-pen-INS  
*ak<sup>w</sup>-p*]<sub>FOCUS</sub>  
 3N.COP-DCL

b. *acajnka* [*arəj a-ruč<sup>w</sup>’ka(-la) ak<sup>w</sup>-p*]<sub>FOCUS</sub>  
 mark this DEF-pen-INS 3N.COP-DCL  
*j-z-la-sə-r-gələ-wa*  
 3N.ABS-REL.IO-INS-1SG.ERG-CAUS-stand-IPF

‘What I give grades with is this pen.’ (elicited)

None of these constructions can be considered a simple clause because both of them include two syntactically unrelated predicates – a finite copula and a non-finite predicate of the relative clause. Nonetheless, foci in internally-headed pseudoclefts are (quite expectedly) more integrated into the clause describing the SOA than external foci and the word order is just one of the manifestations of this.

### 5.2. Agreement with highlighted arguments in Tanti Dargwa

Tanti Dargwa belongs to the East Caucasian family and like almost all of its relatives displays ergative strategies of encoding both in case marking and in the agreement of the predicate. Still, as was first described by Nina Sumbatova, in some periphrastic constructions involving a lexical predicate and a copula, the latter may show agreement with either the actor (17a) or the undergoer (17b), even though the lexical predicate should agree with the absolutive argument:

## (17) Tanti Dargwa (East Caucasian)

a. *murad-li t’ant’i-b qali b-irq’-u-le=sa-b*  
 Murad-ERG Tanti-N(ESS) house N-do.IPF-PRS-CVB=COP-N

b. *murad-li t’ant’i-b qali b-irq’-u-le=sa-j*  
 Murad-ERG Tanti-N(ESS) house N-do.IPF-PRS-CVB=COP-M

‘Murad is building a house in Tanti.’ [Sumbatova and Lander 2014: 451]

Sumbatova conjectured that the copula agrees with its own null subject, which is functionally the theme of the episode and may be coreferent to either the actor or the undergoer. This representation later received support from examples like (18), unknown

by Sumbatova at that time, where the possessor serves as the topic (as indicated by its unmarked rather than genitive form) and in some sense “fulfill” the position postulated by Sumbatova on the basis of exclusively structural considerations. One problem acknowledged by Sumbatova is, however, that the copula can also agree with the element which is explicitly marked as focused by serving as the host of the copula clitic, as in (19).

(18) Tanti Dargwa (East Caucasian)

*ʃaʰmad qali b-ik:-ub-le=sa-j*

Ahmad house N-burn.PF-PRET-CVB=COP-M

‘Ahmad, (his) house burnt.’ (elicited)

(19) Tanti Dargwa (East Caucasian)

*dali=sa-j=da rurs:i quli-r r-alt-un-se*

I.ERG=COP-M=1 girl house.LOC-F(ESS) F-leave.IPF-PRS-ATTR

‘It is me who leaves (my) daughter at home.’

[Sumbatova and Lander 2014: 458]

A simple empirical generalization is that the copula agrees with some prominent element which is highlighted as the topic or the focus – and it is a matter of interpretation and assumptions on the organization of grammar whether an obligatory external position for it should be postulated. Of course, in either case the highlighted element is contrasted with the rest of the clause by controlling such “trigger-happy agreement” (in terms of [Comrie 2003]) on the external copula and hence retains behavior of a less-integrated element. But in many other respects (including case marking and controlling the agreement on the “internal” predicate”), it is normally integrated into the clause. If this is indeed a case of partial integration, we do not need to postulate an external position for all examples, even though it can exist in examples like (18).

### 5.3. “Second passive” in Malay/Indonesian

In Malay/Indonesian there is a specific construction which is often described as a kind of passive [Alieva et al. 1972: 332–337; Chung 1976; Sneddon 1996: 248–249]. Here the transitive verb does not take any voice morphology (unlike in most other constructions) and is immediately preceded by its actor (normally pronominal), while the undergoer usually precedes the whole complex. Hence we find the structure like [Undergoer [Actor Verb]], illustrated in (20).

(20) Standard Indonesian (Austronesian)

*rumah ini bisa kau tempat-i*

house this can you occupy-TR

‘...this house can be occupied by you.’ (SEALang)

As with passive constructions in many other languages, we can think that at least originally this configuration was intended to make the undergoer more topical; in fact, there are approaches that explicitly treat the undergoer in this pattern as the topic [Alsagoff 1991; 1992]. The undergoer is fully integrated into the clause: for example, it is the only argument that can be relativized, which is one of the defining subject properties in languages of West Indonesia [Ogloblin 1996]. The actor, on the other hand, also displays some subject properties – for example, it can antecede the reflexive undergoer, even in lects that clearly prohibit semantic binding [Arka, Manning 1998]:

(21) Indonesian (Austronesian)

*Diri.saya saya serah-kan ke polisi*  
RFL.1 I surrender-TR to police  
'I surrendered myself to police.' [Arka, Manning 1998: 7]

Interestingly, the actor is obligatory here, so this construction clearly declines from the prototypical passive (as described, for example, by Siewierska [1984] and Shibatani [1985]) and is akin to ergative constructions in other languages. But if we consider this an ergative construction, it would violate the well-established generalization that locutor actors prefer accusative encoding as compared with non-locutor actors. Indeed, in Malay/Indonesian the situation is quite the opposite, since the "second passive" construction appears primarily with locutor pronominal actors.

The specific features of the construction under discussion can be easily explained, if we assume that it originates from a left-dislocation construction but the former topic in it has been already integrated into the clause, acquired some properties of the subject but left some other subject properties with the actor.

## 6. Conclusion

To sum up,

- the SOA-related semantics and the information structure can be expressed independently, in which case the topic or the focus are expressed outside of the clause expressing the semantic relations,
- nonetheless there is a pressure to express them in a single syntactic construction, and this pressure may lead to integration of initially separated topics and foci into the clause describing the relevant SOA,
- since this integration is a gradual diachronic process, we observe constructions that display elements only partly integrated, i.e. showing some properties of elements of the clause (described as connectivity effects) and other properties suggesting their independent status.

Presumably, when the dissociation between the SOA-related semantics and the information structure disappears, configurations simultaneously motivated by both components can arise. But this is certainly not the only (and possibly even not the default) way for the appearance of such configurations. Alternatively, the combination of features related to the SOA-related semantics and the information structure may be due to the obligatorification of tendencies (such as the tendency for the actor to be topical).

Note that both ways presuppose that the resulted structures are highly grammaticalized. If grammaticalization leads to the disappearance of transparency [Lander 2015], then the need for dissociation between the SOA-related semantics and the information structure may show up again at this point, so new separating constructions can appear. This scenario, however, is still a matter of speculation, so many more traces of such a circle should be found to make it a valuable hypothesis.



## Abbreviations used in glosses

ABS absolutive, ACC accusative, ADD additive, BEN benefactive, ATTR attributive form, CAUS causative, COP copula, CVB converb, DAT dative, DCL declarative, DEF definite, CISL cislocative, DYN dynamic, ERG ergative, ESS essive, F feminine, FUT future, INF infinitive, INS instrumental, IO indirect object, IPF imperfective, LOC locative, M masculine, N neuter, NEG negation, NFIN non-finite, NOM nominative, OBL oblique, PF perfective, PL plural, PRET preterit, PRS present tense, PST past, REL relative, RFL reflexive, SG singular, TOP topic, TR transitive.

## Sources

RNC – *Russian National Corpus*.

URL: <https://ruscorpora.ru/new/>. Accessed on 27.09.2021.

SEALang – *SEALang Library Indonesian Text Corpus*.

URL: <http://sealang.net/indonesia/corpus.htm>. Accessed on 27.09.2021.

WCC – Arkhangelskiy, Timofey, Irina Bagirokova, and Yuri Lander. 2018. *West Circassian (Adyghe) Corpus*.

URL: <http://adyghe.web-corpora.net/>. Accessed on 27.09.2021.

## References

1. Aannestad, Aidan. *A Typology of Morphological Argument Focus Marking*. M.A. thesis, University of North Dakota.
2. Alieva, Natalia F., Vladimir D. Arakin, Alexander K. Ogloblin, Ülo Sirk. *Grammatika Indonezijskogo Jazyka*. Moscow: Nauka; 1972. (In Russ.).
3. Alsagoff, Lubna. 1992a. The notions of Subject and *Topic* in Malay. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Grammar of Event Structure (1991)*. Berkeley: BLS; 1992:3–14.
4. Alsagoff, Lubna. 1992b. *Topic in Malay: the Other Subject*. Ph.D. diss., Stanford University; 1992
5. Andrason, Alexander, Josh Westbury, and Christo H. J. van der Merwe. 2016. Left dislocation: An exploration in linguistic typology. *Stellenbosch Papers in Linguistics Plus*. 2016;50:1–20
6. Ariel, Mira. *Accessing Noun-Phrase Antecedents*. London: Routledge; 1990.
7. Arka, I Wayan and Christopher D. Manning, 1998. Voice and grammatical relations in Indonesian: A new perspective // Mirriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG98 Conference*. Stanford: CSLI Publ.  
URL: <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG3-1998/>. Accessed on 27.09.2021.
8. Chung, Sandra. On the subject of two passives in Indonesian. Charles N. Li, ed., *Subject and Topic*. New York: Academic Press; 1976:57–98.
9. Comrie, Bernard. When agreement gets trigger-happy // *Transactions of the Philological Society*. 2003;101(2):313–337.

10. Givón, T. *Syntax: An Introduction*. Vol. 1. Amsterdam, Philadelphia: John Benjamins; 2001.
11. Harris, Alice C. and Lyle Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge: Cambridge University Press; 1995.
12. Heine, Bernd and Mechthild Reh. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske; 1984.
13. Heller, Daphna. 2002. On the relation of connectivity and specificational pseudoclefts // *Natural Language Semantics*. 2002; 10(4):243–284.
14. Her, One-Soon. 1991. Topic as a grammatical function in Chinese. *Lingua*. 1991;84(1):1–23.
15. Higgins, F. Roger. *The Pseudo-cleft Construction in English*. New York: Garland; 1979.
16. Iwasaki, Shoichi. *Japanese*. Amsterdam/Philadelphia: John Benjamins; 2013.
17. Kibrik, Andrej A. *Reference in Discourse*. Oxford: Oxford University Press; 2011.
18. Kindlein, David. *Aspects of Information Structure in Written Kabardian Narrative*. M.A. thesis, University of Gloucestershire; 2016
19. Kuno, Susumu. *The Structure of the Japanese Language*. Cambridge (Mass.), London: The MIT Press; 1973
20. Lambrecht, Knud. Dislocation. Martin Haspelmath et al., eds., *Language Typology and Language Universals: An International Handbook*. Vol. 2. Berlin, New York: Walter de Gruyter; 2001:1050–1078.
21. Lander, Yury. On how compositionality relates to syntactic prototypes and grammaticalization. Peter Arkadiev et al., eds., *Donum semanticum: Opera linguistica et logica in honorem Barbarae Partee a discipulis amicisque Rossicis oblata*. Moscow: LRC Publishers; 2015:148–157.
22. Lander, Yury and Anastasia Vyunova. How independent can dependent marking be? *Vestnik Instituta vostokovedenija RAN*. 2018;5:185–192.
23. López, Luis. Dislocations and information structure. Caroline Féry and Shinichiro Ishihara, eds., *The Oxford Handbook of Information Structure*. Oxford: Oxford University Press; 2016:402–421.
24. Ogloblin, Alexander K. *Očerk diaxroničeskoj tipologii malajsko-javanskix jazykov*. Moscow: Novoe tysjačeletie; 1996. (In Russ.)
25. Radford, Andrew. *Colloquial English: Structure and Variation*. Cambridge: Cambridge University Press; 2018.
26. Rudnitskaya Elena L. *Spornye voprosy korejskoj grammatiki. Teoretičeskie problemy i sposoby ix rešenija*. Moscow: Vostočnaja literature RAN; 2010. (In Russ.)
27. Shibatani, Masayoshi. Passives and related constructions: a prototype approach. *Language*. 1985; 61(4): 821–848.
28. Shibatani, Masayoshi. Grammaticalization of Topic into Subject. Elizabeth Closs Traugott and Bernd Heine, eds., *Approaches to Grammaticalization*. Vol. 2. Amsterdam/Philadelphia: John Benjamins. 1990: 93–133.
29. Siewierska, Anna. *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm; 1984.
30. Sneddon, John N. *Indonesian: A Comprehensive Grammar*. London, New York: Routledge; 1996.

31. Sumbatova, Nina R. Kommunikativnaja struktura adygejskogo predloženija: perspektiva i fokus // Yakov G. Testeleťs, ed., *Aspekty polisintetizma: Očerki po grammatike adygejskogo jazyka*. Moscow: RGGU. 2009:559–611. (In Russ.)
32. Sumbatova, Nina R. and Yury A. Lander (with Magomed Kh. Mamaev). *Darginskij govor selenija Tanty: grammatičeskij očerk, voprosy sintaksisa*. Moscow: Jazyki slavjanskoj kul'tury; 2014. (In Russ.)
33. Surányi, Balázs. Discourse-configurationality. // Caroline Féry and Shinichiro Ishihara, eds., *The Oxford Handbook of Information Structure*. Oxford: Oxford University Press; 2016:422–440.

**Hittite agent form****Ilya S. Yakubovich**Institute of Oriental studies of RAS, Moscow, Russia,  
sogdiana783@gmail.com, ORCID 0000-0002-7576-7930

**Abstract.** The Hittite agent form in *-anza* is used in lieu of the base forms of neuter nouns when the respective nouns appear in subject position in the transitive clauses. It is argued that the *anza*-extension is to be analyzed as /-ant-s/, where the morpheme /-ant-/ transfers the noun to the common gender, while /-s/ is the nominative ending. The agent form can be regarded as part of the paradigm of the Hittite neuter nouns, but it coexists with the derivational transgender suffixes and has come into being through the grammaticization of one of such suffixes. The creation of the agent form was facilitated by the paradigmatic gap in the declension of Anatolian neuter nouns and a broader phenomenon of privative ergativity, which can be reconstructed at the Proto-Anatolian level.

**Keywords:** Hittite, Anatolian, agent form, paradigmatic gap, privative ergativity

**Acknowledgments:** I am grateful to all the participants of the conference *Historical-Comparative Linguistics in the 21st century* (Pavia, 2011), who participated in the discussion of my talk, and particularly to Silvia Luraghi, who organized this event. The editors of this volume kindly accepted this contribution, even though my presentation at the Z. M. Shal'japina memorial conference was on a different topic. Stephen Durnford and Vera Tsukanova did much to improve the style of this paper. None of the above scholars is to be blamed for its shortcomings.

**Аннотация.** Агентивная форма с суффиксом *-anza* используется вместо основной словарной формы хеттских существительных среднего рода в тех случаях, когда они выступают в позиции подлежащего в переходных клаузах. В статье предлагается анализировать формант *-anza* как /-ant-s/, где морфема /-ant-/ отвечает за перевод существительных из среднего рода в общий, а /-s/ является окончанием именительного падежа. Агентивная форма является частью парадигмы хеттских существительных среднего рода, сосуществуя при этом с деривационными механизмами изменения рода существительных и представляя собой результат грамматикализации одного из таких механизмов. Появлению агентивной формы способствовала неполнота парадигмы анатолийских существительных среднего рода, что в свою очередь является одним из проявлений более общего феномена привативной эргативности, реконструируемого для анатолийского праязыка.

**Ключевые слова:** хеттский язык, анатолийские языки, агентивная форма, неполная парадигма, привативная эргативность

**Благодарности:** Автор признателен участникам конференции *Historical-Comparative Linguistics in the 21st century* (Павия, 2011) за обсуждение его доклада и в особенности Сильвии Лураги, организатору конференции. Редакторы данного сборника приняли данную статью, хотя доклад на конференции памяти З. М. Шал'япиной был на другую тему, а С. Дёрнфорд и В. Л. Цуканова помогли стилистически улучшить текст.

## 1. Introduction

One of the features of the Hittite, which sharply sets it apart from most ancient (non-Anatolian) Indo-European languages, is the phenomenon of split-ergativity. The languages belonging to the Anatolian branch of the Indo-Anatolian (Early Indo-European) language family typically distinguish between the common and neuter gender within the nominal system. Prototypically, the Hittite neuter nouns are inanimate, while the nouns of the common gender can be either animate or inanimate. The Hittite nouns and adjectives of the common gender take the nominative case if they function as subjects, in either transitive or intransitive clauses, and the accusative case if they function as direct objects. The behaviour of the neuter nouns is quite different: as in the other ancient Indo-European languages, they show the identical case marking in the capacity of intransitive subjects (S) and direct objects (O), as in (1-2), but the transitive subjects, semantically agents (A), feature the same lexemes provided with a special suffix, as in (3). More specifically, the form *išhan-anza* is derived from *išhan-*, the oblique stem of *ēšhar* ‘blood’ (n.), extended with the element *-anza*.

- (1) KUB 14.3 iv 52, CTH 181

nu	apāt	ēšhar	kuwapi	paizz[i]
CONN	that.NOM.SG.N	blood.NOM.SG	where	go.3SG.PRES

‘When that blood flows...’.

- (2) KBo 22.1 obv. 24–25, CTH 272

nu	ŠA	<sup>LU</sup> MÁŠDA	ēšhar=šet	natta	šanhiškatteni
CONN	GEN	poor-man	blood=his.ACC.SG.N	NEG	seek.2PL.PRES

‘You do not seek the blood of a poor man’.

- (3) KUB 14.14 rev. 11, CTH 378

nu	<sup>KUR URU</sup> HATTI=ya	apāš	išhan-anza	arḫa	namma
CONN	Hattusa=and	that.NOM.SG.C	blood.NOM.SG.C!	away	also

zinne[šta]  
finish.3SG.PRET  
‘That blood(shed) further finished off (the land of) Hattusa as well’.

By definition, the ergative alignment consists in the distinct marking of the transitive subjects while the marking of the intransitive subjects and direct objects is identical. In this sense, the difference between *išhananza* in (3), on the one hand, and *ēšhar* in (1) and (2), on the other hand, clearly represents an argument for the ergative alignment in a particular segment of Hittite grammar. The significance for morphology of this syntactic observation is less clear: should one add the ergative case to the paradigms of the Hittite grammar, should one treat the agent form as a derived stem, or can one find a middle ground? Another controversial question concerns the implications of Hittite split-ergativity for Anatolian and Indo-European comparative linguistics. Is it to be treated as a Hittite or Anatolian innovation, an archaism of Early Indo-European, which is lost from Core Indo-European, or something in between?

These are the issues with which I was engaging when I made a presentation entitled ‘Ergativity in Hittite’ at the conference *Historical-Comparative Linguistics in the 21<sup>st</sup> century*, organized by Silvia Luraghi within the framework of the *Humboldt Kolleg*

and held in Pavia in September 2011. No proceedings of this conference were ever published, while my participation in several other projects prevented me at the time from preparing a journal article on the basis of my talk or otherwise pursuing research on this topic. Recently, however, a number of papers (notably [Goedegebuure 2013], [Jasanoff 2017], [Pooth et al. 2019], [Pooth and Orqueda 2021]) developed and carried forward certain claims based on my presentation, citing them with reference to its accompanying handout. While I was naturally gratified to see that my ideas are gaining acceptance, I was simultaneously embarrassed by the fact that my colleagues had to resort to quoting from an unpublished source, sometimes with the note “available upon request”.

The present paper serves the modest goal of filling in this gap in the printed discourse. I deliberately abstain from discussing the secondary sources that postdate my Pavia presentation, since doing otherwise would extend the paper far beyond the limits of this volume, while there is no guarantee that my intrusion into the area where I currently feel just a guest would take the discussion further ahead. For those who wish to cite my paper in the future, I recommend the format [Yakubovich 2011/2021], explicitly acknowledging its status as the elaboration of my earlier work. The only substantial innovation of this version vis-à-vis the earlier handout is the consistent use of the term “agent form”, which is also reflected in the title of the present paper.

A more personal reason for publishing this research is the desire to contribute to the journal issue devoted to the memory of Zoja Mikhajlovna Shaljapina, my former senior colleague at the Institute of Oriental Studies, Russian Academy of Sciences. A year before the Pavia conference she welcomed me to its Department of Languages, of which she was then the head, and I continued to work under her benevolent leadership until her retirement in 2018. Given her general interest in what was known as structural linguistics in the days of her youth, as well as her particular engagement with the combinatorial properties of linguistic units, it seems appropriate to dedicate my belated publication on this topic to her memory.

The rest of the paper will be organized as follows. After a brief survey of previous research in Section Two, I will proceed to the synchronic morphological analysis of the Hittite agent form in Section Three. Section Four will be devoted to the question whether it should be treated as derivational or inflectional. The last two sections will address the historical problems: in Section Five I will address the formal precursors of the agent form, while in Section Six I will tackle the functional motivation for its development.

## 2. Research History

The first coherent description of the phenomenon under discussion was made in Laroche 1962. Laroche discovered that neuter nouns cannot function in Hittite as the subject of a transitive verb, but must be replaced by a special form in *-anz(a)* (/ants/) in the singular and *-anteš* (/antes/) in the plural. He acknowledged that the distribution of these two markers is functionally similar to that of ergative case endings in other languages but nevertheless suggested that their primary function is gender-switching (“*transfert d’un inanimé dans la classe animée*”).

The immediate reaction to Laroche’s hypothesis was rather sceptical. [Benveniste 1962] pleaded for the analysis of *-ant-* as a derivational morpheme with a semantic function. The derivatives in *-ant-*, according to him, “*désigneront des notions matérielles*”

*es transférées au rang de puissances actives*”. He objected to the existence of the ergative construction in Hittite on the grounds that the ergative nouns prototypically denote animate agents, which is now an outdated assumption (cf. below in this section). He also pointed out that the same suffix was occasionally attached to the common gender nouns and indicated in that it is likely to be ultimately related to the other Indo-European derivatives endowed with the suffix *\*-ont-*. The same line of thinking was later developed by [Neu 1989] and [Carruba 1992], who opted for a “personifying” and “individualizing” function of the *-ant-* suffix, respectively.

Nonetheless, the hypothesis of the ergative alignment in Hittite underwent a revival in later scholarship. A crucial role in this development was played by [Garrett 1990a], a paper published in the leading linguistic journal, which argued that the Hittite language possessed a genuine ergative case. Garrett rejected Laroche’s analysis of *-ant-* as a gender-switching suffix on the grounds that “a derivational suffix whose function is to permit nouns of a particular gender to stand in a particular syntactic function would ... have no parallel in Indo-European or elsewhere”. He advanced the derivation of Hitt. *-anz(a)* from the reconstructed instrumental marker *\*-an-ti* and suggested that its plural counterpart *-ant-eš* is due to the secondary accretion of the nominative plural ending */-es/* of the common gender.

To be sure, the new stance on Hittite ergativity was not universally accepted. Thus, [Oettinger 2001] generally followed the synchronic analysis of [Laroche 1962], but argued that the gender-switching *-ant-* could represent the grammaticization of the derivational suffix without such a function (e.g. *ḫamešḫa-* (c.) vs. *ḫamešḫ-ant-* (c.) ‘spring’), to which he attributed a “personifying” value. According to a different approach defended in [Josephson 2004], Hitt. *-ant-* was originally a morpheme deriving countable entities from the collectives, although “the singulative and decollectivizing *-ant-* permits an agentive function, which allows for subjecthood and nominative case form”.

In contrast, [Patri 2007] took Garret’s historical analysis as the starting point but suggested that Hittite neuter subjects take the synchronic ablative endings in transitive constructions in lieu of the nominative markers. He objected on alleged typological grounds to the status of *-anz(a)* as an ergative ending restricted to the class of neuter nouns, stating that that “*ce sont pratiquement toujours les noms animés qui font preuve d’un alignement ergatif, très rarement l’inverse*”. This assertion, however, contradicts a widely accepted typological universal known as Silverstein’s hierarchy, according to which nominative-accusative case marking correlates with higher animacy, while ergative case marking correlates with lower animacy [Silverstein 1976: 113].

The standard reference grammar of the Hittite language [Hoffner and Melchert 2008] listed *-anza* as the ergative singular ending and *-anteš* as its plural counterpart, essentially following Garrett’s views. Nevertheless, voices of dissent continued to be heard. Thus, [Dardano 2010] dwelled on the instances where forms in *-anza* derived from neuter nouns occur in intransitive constructions, which speaks against their ergative function. She supported the analysis of *-ant-* as an animating suffix along the line of [Benveniste 1962]. On his part, [Rizza 2010] subscribed to the analysis of [Laroche 1962]: “*si potrebbe trovare, infine, spazio, per l’ergatività in eteo senza caso ergativo per la classe dei neutri*”. The same claim that the structure of the Hittite language features the ergative construction but not the ergative case is found in [Shatskov 2011], with an emphasis on the coexistence of different strategies of gender-switching.

Melchert’s views on Hittite ergativity are laid out in the paper that was based on his 2007 presentation, but appeared in print four years later [Melchert 2011a]. This paper took an agnostic stance on the origin of the ergative ending *-anza* but defended its synchronic existence against both the derivational hypothesis of [Benveniste 1962] and the instrumental hypothesis of [Patri 2007]. Melchert’s main empirical argument for the proposed analysis was based on the resumption of some ergative subjects by neuter clitics, which speaks against the view that the nouns in *-anza* underlyingly belong to the common gender.

The typology of the previous approaches to the ergative alignment in Hittite can be summarized in the following table, where only the most relevant papers have been cited:

Table 1: Diverging views on Hittite ergativity

	Morpheme <i>-ant-</i>	Morpheme <i>-anz(a)</i>
Ergative construction	Oettinger 2001, Rizza 2010, Shatskov 2011	Laroche 1962, Garrett 1990a, Melchert 2011a
No ergative construction	Benveniste 1962, Josephson 2004, Dardano 2010	Patri 2007

### 3. Morphological segmentation

The embarrassment of choice reflected in the table above clearly calls for a new review of the Hittite empirical data. It seems appropriate to begin with establishing the morphemic boundaries in the agent form. Hittite is an inflectional language, which implies that one morpheme normally combines information about the number, case, and gender (where appropriate) of the inflected nouns and adjectives. Therefore, the proponents of the ergative ending *-anza* (right column in Table 1) assume by definition that this is a unitary morpheme. In contrast, their opponents, seemingly without exception, operate with the morphemic combination */-ant-/+/-s/*, where */-s/* is the nominative ending of the common gender. The question “one morpheme vs. two morphemes” belongs to the domain of morphology and morphosyntax and must be resolved with the reliance on the standard combinatorial methods, which have been commonly used in descriptive linguistics since the early twentieth century.

First, it is easy to confirm that the combination of the morphemes */-ant-/* and */-s/* indeed yields the graphic sequence *-an-za* in Hittite. The nominal and adjectival *-nt-* stems, such as *appant-* ‘captive’ or *hūmant-* ‘all’, show nom.sg forms *ap-pa-an-za* and *hu-u-ma-an-za* respectively. But this is not their only similarity with the forms in *-anza* attested in the ergative constructions. Just as *ap-pa-an-za* ‘captive’ can be contrasted with *ap-pa-a-an-te-eš* ‘captives’, so *ud-da-na-an-za* ‘word’ can be contrasted with *ud-da-na-an-te-eš* ‘words’ in the position of the transitive subject. This parallelism can be more extensively illustrated by the table below. All the forms transcribed there are based on the paradigms listed in [Hoffner and Melchert 2008: 123–124].

Table 2 “Ergative” derivatives and common gender *ant-*stems

	water (“erg.”)	word (“erg.”)	captive	all
nom.sg.	<i>/widen-ant-s/</i>	<i>/uttan-ant-s/</i>	<i>/app-ant-s/</i>	<i>/xomant-s/</i>
nom.pl.	<i>/widen-ant-es/</i>	<i>/uttan-ant-es/</i>	<i>/app-ant-es/</i>	<i>/xomant-es/</i>



This proportion alone would probably represent a sufficient reason to segment *ud-da-na-an-za* as /uttan-ant-s/ rather than *\*\*/uttan-ants(a)/*. Yet, there are also additional arguments for such a segmentation. As was already mentioned above, the Hittite language features the opposition between common and neuter genders, which is (mostly) lexical for the nouns and inflectional for the adjectives. While *uttar* ‘word’ and *wātar* ‘water’ are neuter nouns in Hittite, /uttan-ant-es/ and /widen-ant-es/ must be analyzed as forms of the common gender, because the nominative plural ending of the neuter gender is /-a/ rather than /-es/. The segmentation /uttan-ant-s/ yields another form of the common gender, with its characteristic nominative singular ending /-s/, and is conducive to identifying /-ant-/ as the gender-switching marker. In contrast, if one prefers the analysis *\*\*/uttan-ants(a)/*, one is left in the dark about the mechanism responsible for the transfer of /uttan-ant-es/ to the common gender. Pace [Garrett 1990: 290], a proportional analogy of the type *appanza / appanteš = uddananza / X → X = uddananteš* could hardly work if the Hittite native speakers processed *appanza* and *uddananza* as forms with different morphological segmentation and different gender assignment.

Fortunately, there are independent ways to show that *uddananza* and similar forms were assigned the common gender in Hittite. [Patri 2007: 52–53] lists noun phrases such as *apā-š išhan-anza* ‘that blood(shed)’, *kā-š tuppiy-anza* ‘this tablet’ and *kui-š witenanza* ‘which water’, which clearly suggest that the “ergative” nouns can agree with nominative modifiers of the common gender. There are no opposite examples, such as *\*\*apā-t išhan-anza*, displaying the neuter ending *-t* instead of the common gender ending *-š*, despite the fact that *wātar* ‘water’, *tuppi-* ‘tablet’, and *ēšhar* ‘blood’ are all neuter nouns. If the agent form is modified by the adjective instead of the pronoun, two competing strategies are attested: the adjectives can take either the nominative ending *-š* (*šuppi-š A-anza* ‘pure water’, *ištarniya-š annašn-anza* ‘internal \*annaššar’) or the same element *-anza* as their syntactic heads *idalaw-anza* GIG-*anza* ‘evil illness’, *ḪUL-uw-anza uddan-anza* ‘evil word’ (for further examples, cf. [Shatskov 2011: 144–145]). Under the two-morpheme analysis, we obtain the expected pattern of gender agreement in both cases: /suppi-s widen-ant-s/ ‘pure water’ vs. /idalaw-ant-s uttan-ant-s/ ‘evil word’. Once again, there are no examples that plead for the interpretation of the derived forms in *-anza* as neuter nouns.

The data pertaining to clause agreement are consistent with those pertaining to phrase agreement. To be sure, the singular forms in *-anza* are unsuitable as a test: predictably for an ancient Indo-European language, Hittite shows no difference between the predicates agreeing with singular subjects depending on their gender. In contrast, the clauses featuring plural subjects display the gender-based discrepancy: the plural subjects of the common gender trigger number agreement on the predicate, while the neuter plural subjects agree with the singular predicates (Hoffner and Melchert 2008, §15.16 with example (4)). Yet, the plural forms in *-anteš* derived from neuter nouns show the agreement pattern that is typical of the common gender, as can be illustrated by example (5), taken from [Garrett 1990a: 266–267].

- (4) KBo 3.34 ii 18, CTH 8  
<sup>m</sup>Āškali=ma                      uddār                      arāiš  
 Askali.DAT.SG=but              word.NOM.PL              rise.3SG.PRET.SG  
 ‘But accusations arose against Askali’.

- (5) KBo 55.52 obv. 11, CTH 449  
 [SA]G.DU-ann=a ishīšn-anteš appanzi  
 head<sup>2</sup>.ACC.SG=and binding.NOM.PL.C! seize.3PL.PRES  
 ‘And the bindings will seize the head?’

The force of this last argument need not be exaggerated: we have seen that the morphology of *-anteš* transparently points in the direction of the common gender, so this agreement pattern is only natural. The point here is that if the formation of *-anteš* were a result of a purely mechanical proportional analogy, as per [Garrett 1990a: 290], then we could expect them to preserve the lexical gender of the Hittite base noun *išhieššar* (n.) ‘binding’. On the whole, the patterns of plural formation, phrase agreement, and clause agreement supply three independent arguments in favour of the assumption that the agent form of the (underlying) neuter nouns belongs to the common gender. After the segmentation of the endings */-s/* and */-es/*, we are left with the suffix */-ant-/*. In the instance of nouns, it emerges as a marker of gender-switch, while in the instance of adjectives (*idalaw-ant-* etc.) it appears to carry no grammatical function and to represent a redundant agreement marker.

Against this background, we can consider the main combinatorial argument that was advanced against the analysis in terms of gender-switching. It has to do with discourse cohesion, namely the resumption of the agent form by the neuter anaphoric pronoun *=at* ‘it’ in two Hittite examples [Melchert 2011a: 162–163]. Their translations should be sufficient to illustrate this point: (a) ‘Water (A-*az*) washes the roof and it (=at) flows down the drain’; (b) ‘[They have give]n compensation for the blood(shed) (*ēšhar*). That blood(shed) (*išhananza*) further finished off the Land of Hattusa as well, so that the Land of Hattusa has already made compensation for it (=at)’. Significant for interpreting these both examples is the syntactic function of the neuter clitic *=at* as intransitive subject and direct object respectively. In fact, we shall see in Section Six that this clitic pronoun simply cannot appear in subject position in the Hittite transitive clauses, which constitutes yet another ergative feature of the Hittite grammar. But as long as the resumptive pronoun does not appear in the same syntactic position as the agent form, one can make a legitimate argument that it simply refers back to the underlying neuter gender of its antecedent. Example (b) is particularly revealing: here the neuter dictionary form *ēšhar* coexists with the secondary common gender form *išhananza*, but as long as the syntactic structure allows it, the anaphoric clitic preserves the gender of the base form.

The additional evidence offered by some of the followers of the one-morpheme analysis as confirmation of their views consists in the derivation of *-anza* from the ablative (later ablative-instrumental) ending, which in most cases is recorded with the *-az* sign in Hittite. In some cases, as in A-*az* ‘water’ cited in the previous paragraph, the agent form can also be endowed with the ending *-az*, probably reflecting phonetic [-ãts]. The nasal vowel formation was a change in progress in Late Hittite, possibly reflecting the influence of the Luwian substrate [Yakubovich 2010: 324–325]. Conversely, some ablative-instrumental forms can exceptionally show the ending *-anza*, for example *witenanza* ‘with water’ [Garrett 1990a: 277]. The fact that this change is virtually restricted to oblique stems ending in *-n* [Patri 2007: 39–40] is compatible with the interpretation of this phenomenon as optional post-nasalization. The precise phonetic explanations are, however, besides the point as long as both the ablative ending *-anza* and the

agent form in *-az* are rare and exceptional, which seems to be recognized by all the participants of the debate. Perhaps their connection could be explored as a historical hypothesis if it were fully supported by synchronic combinatorial evidence, but given that this evidence overwhelmingly contradicts it, such a solution has nothing to recommend itself.

#### 4. Inflection or derivation?

So far, we have seen that the element *-anza* is to be decomposed into the nominative ending */-s/* and the gender-switching morpheme */-ant-/*. For some scholars this is tantamount to saying that */-ant-/* is a derivational morpheme, as suggested by the following formulation: “*Ist -anza als -ant-s zu erklären (d.h. -(a)nt- ist als derivationelles Suffix zu analysieren plus die Endung des Nominativs Singular) oder ist -anza als flexionale Endung zu erklären?*” [Dardano 2010: 175]. The same dichotomy between the ergative case ending and the derivational suffix *-ant-* is implicit in the earlier work with my participation [Kassian and Yakubovich 2007: 436–437]. A more nuanced set of alternatives is offered in [Rizza 2010: 150]: “*Credo si debba scegliere: o è un fenomeno flessivo y allora il sistema è ergativo, o è un fenomeno derivazionale (di mozione, sostanzialmente), o al limite entrambi (flessivo y derivazionale: metaplasmo di genere per ragioni sintattiche) e allora è comunque riducibile a derivazionale, per quanto concerne la classe nominale (metaplasmo) e le conseguenze per l'accordo con nominali retti*”.

The morphological status of */-ant-/* must, however, be studied before one explores its consequences for evaluating the status of Hittite ergativity. The principle “one inflectional morpheme per word form”, which seemed to guide some of the preceding research, is by no means a universal: the agglutinative languages routinely attach combinations of inflectional morphemes to inflected forms, for example, those expressing number and case. One, therefore, cannot use the inflectional character of the ending */-s/* in order to deny the same status to */-ant-/*. There is, however, a set of universal tests for determining whether a particular morpheme is inflectional or derivational, which is summarized, for example, in [Haspelmath 2002: 70]. Not all these tests are equally relevant to our situation, and some of them might be considered redundant for our purposes, but the three of them listed below appear to be sufficient to settle the case of the Hittite morpheme under discussion.

The relevance of *-ant-* to the syntax need not detain us for long, since this topic has already been implicitly treated in the previous section. As a gender-switching suffix, *-ant-* always affects the phrase agreement features and sometimes also affects the clause agreement features of the relevant nouns. This criterion cannot, however, be decisive, since the derivational morphemes can also influence the syntactic features of the base noun.

Table 3: Distinction between inflection and derivation

Derivation	Inflection
Not relevant to the syntax	Relevant to the syntax
New concept	Same concept as base
Limited applicability	Unlimited applicability

The preservation of a lexical meaning provides a more direct evidence, although it is more difficult to ascertain it in the instance of an extinct language. It seems, however, reasonable to assume that if two cognate forms are used with the same reference, then, in the absence of the contrary evidence, they stand for the same concept. In this connection, we can return to the example from the Edict of Telipinu, which has already been cited in the previous section: “[They have give]n compensation for the blood(shed) (*ēšhar*). That blood(shed) (*išhananza*) further finished off the Land of Hattusa as well” (KUB 14.14 rev. 10-11). Here *ēšhar* and *išhananza* ‘blood(shed)’ obviously refer to the same entity, so the difference between the two forms must be grammatical, not lexical. Furthermore, the resumption of the forms in /-ant-s/ by neuter pronouns, again addressed in the previous section, can only be interpreted as an indication that these pronouns stand for the neuter forms of the paradigm. For example, if the neuter noun =*at* (nom.sg.n.) functions as the anaphor of A-*az* = /widenants/ ‘water’ (c.), the default assumption is that it stands for *wātar* ‘water’ (n.), which in turn supports the hypothesis that the two forms belong to the same paradigm.

The last test concerns the obligatory character of inflectional morphemes within the domains of their application, cf. [Haspelmath 2002: 75]. In the instance of markers of the Hittite agent form, this domain is limited to underlyingly neuter nouns, since the Hittite nouns of the common gender consistently show nominative alignment, as is also the case in other ancient Indo-European languages. For example, the Hittite noun *aruna-* (c.) ‘sea’ features the nominative form *arunaš*, which can function as subject in both transitive and intransitive clauses, and the accusative form *arunan*, which is used in the function of the direct object. Therefore, our enquiry is limited to the question whether the nominative-accusative forms of neuter nouns ever occur as transitive subjects in Hittite. The answer turns out to be negative: the only alleged counterexample, namely *ḫandaiš* in (7), is not probative, as per [Dardano 2010: 177], since assigning the common gender to this noun represents a viable alternative analysis.

- (7) KBo 3.23 i 5–6, CTH 24  
 mān[=an]            ḫandaiš            walaḫzi  
 when=she.ACC.C heat.NOM.SG.C!    strike.3SG.PRES  
 ‘When the heat strikes [him] ...’.

To be sure, there are some complications to the proposed test. On the one hand, we find instances of apparent over-application of gender-switching: the common gender forms in *-ant-* may appear in intransitive contexts, cf. [Neu 1989: 2]. Thus, in (8) below, the secondary common gender stem *ḫuwallišsan-ant-* belongs to the paradigm of the Hittite noun *ḫuwalliš(n)-* (n.) ‘cone’. The likely reason for the use of the agent form in this case is that the subject semantically behaves as agent: the heated cone<sup>7</sup> produces a hissing sound, as opposed to passively floating in water, and is therefore perceived to be in control of the situation. In terms of generative linguistics, this is an unergative rather than unaccusative clause, while typologists can use it as an argument for the elements of split-S system rather than pure ergativity in Hittite. What primarily matters for our purposes is that the occasional over-application of a grammatical marker need not cancel its inflectional character. For example, George Orwell’s aphoristic sentence ‘All animals are equal, but some animals are more equal than others’ features a playful use of the comparative degree with what normally functions as a relational adjective,

but this is not a reason to doubt the status of the comparative degree as an inflectional feature of the gradable adjectives. A typological parallel from Hittite involves the medio-passive verbs: while in the majority of cases the active and medio-passive paradigms can be attached to the same stems but are associated with syntactic differences, there are also verbs compatible solely with medio-passive endings, while in the instance of other verbs such endings may occur in constructions that are otherwise typical of active forms [Hoffner and Melchert 2008: 230–231].

- (8) KUB 7.58 i 3–4, CTH 779 (restored based on KUB 45.20 i 20)  
 (...) *ḫuwalliššananzašša* [(G)IM-an] *weteni* *anda taškupāizzi*  
 cone<sup>?</sup>.NOM.SG.C!=and as water.DAT.SG in hiss.3SG.PRES  
 ‘As (heated stone) or cone<sup>?</sup> hisses in water...’.

Another problem is that the morpheme *-ant-* is just one of several markers of secondary common gender forms. Thus in (7) above one has to assume that the neuter *s*-stem *ḫandaiš-* ‘heat’ was re-assigned to the common gender without adding any special marker, i.e. */xandais-/* → *ḫandaiš*, or even underwent reanalysis in the position of a transitive subject, i.e. */xandai-/* → *ḫandaiš*. In the instance of Hitt. *nēpiš-* (n.) ‘sky, heaven’, its (singular) agent form is *nepišanza*, but *nepiša-* (c.) is also attested (nom.sg. [*ne-p*]ī-ša-aš, acc.sg. *ne-pi-ša-an*), and there are more examples of common gender derivatives formed by the thematization of the stem [Shatskov 2011: 146–147]. Furthermore, the gender-switch in Hittite can also be accomplished by the addition of the suffix (originally compound element) *-šepa-* / *-zepa-*: for example, Hitt. *tēkan-* (n.) ‘earth’ can be contrasted with its common gender counterpart *taganzepa-* [Kloekhorst 2008: 812]. It is, however, crucial that the gender-switching markers other than *-ant-* normally trigger the formation of new lexemes endowed with their own paradigms, while in the instance of the transgender suffix *-ant-* such cases constitute rare exceptions. The competition and coexistence of agent form and derived nouns of the common gender is schematically represented in Table Four below. While the lexical derivatives of the common gender are certainly of interest for tracing the origins of the Hittite agent form, as will be elaborated in the next section, they need not undermine the status of the agent form in *-ant-* as member of inflectional paradigms. The lexical derivative *nepiša-* could coexist with the agent form *nepišanza* just as the collective noun *Gebirge* can coexist with the plural form *Berge* ‘mountains’ in German.

Table 4: Competing strategies of gender-switching in Hittite

	‘water’	‘earth’	‘sky’
Base form	<i>wātar</i>	<i>tēkan</i>	<i>nēpiš</i>
Agent form	<i>witananza</i>	∅	<i>nepišanza</i>
Derived noun	∅	<i>taganzepa-</i>	<i>nepiša-</i>

Finally, one must acknowledge that the paradigms of most Hittite neuter nouns plainly lack the attested agent form, as does *tēkan* (n.) ‘earth’ in the above table. This situation can hardly be called exceptional: the limited character of the Hittite corpus inevitably leads to the presence of accidental gaps in the majority of the available paradigms, both nominal and verbal. Yet, there is also an intrinsic reason why the agent

form is rare: the Hittite neuter nouns generally denote non-living entities, which prototypically do not control the acts where they are involved. The examples cited above, in particular (3) and (5) show that the use of neuter nouns as subjects of transitive verbs is typical of metaphoric constructions. This observation might again be relevant for tracing the historical origins of the agent form, but it need not affect the conclusions about its synchronic morphological status.

Thus, there are no formal obstacles to treating the Hittite agent form in *-ant-* as part of a nominal paradigm. In practical terms, this means that they need not be listed as separate dictionary lemmata, although, given their relatively rare attestations, they can be listed as separate stems next to the base forms, e.g. *ēšhar* (n.) ‘water’, cf. *išhan-ant-* (c.) ‘id.’ This suggestion broadly conforms to the practice adopted, for example, in [Tischler 2001]. The question of how one should refer to such formations is more complicated. The term adopted in this paper essentially pays homage to (8) and similar examples, suggesting that the occurrence of the agent form is not limited to the ergative constructions per se. Yet, given the frequent use of the label “ergative” in connection with the *-ant-* forms in the previous literature as well as the observed close correlation between agent form and ergative construction in Hittite, the term “ergative form” (still abbreviated as *erg.*) emerges as an acceptable compromise.

For formal reasons, given in Section Two above, I must strongly object to the term “ergative case”: *išhananza* is structurally a nominative case form provided with a gender-switching operator. Yet, nothing in the previous discussion is to be taken as an argument against the existence of the ergative construction in Hittite. Structural mismatches between different levels of linguistic description need not undermine its validity but rather bear witness to the inherent complexity of language organization. In this sense, the mismatch between the split-ergative surface syntax and the absence of the morphological ergative case in Hittite can be considered together with the mismatches between morphosyntactic and (deep) syntactic ergativity that have been discussed in connection with certain other world languages [Comrie 1989: 72–74, 114]

## 5. Prehistory of the agent form

In order to trace the origins of the transgender marker *-ant-*, it seems appropriate to take into consideration the function of its likely cognates in the related Anatolian languages, such as Luwian and Lycian. In Luwian we find a number of derived common-gender forms endowed with the cognate morpheme *-ant-*, some of them functionally very similar to the Hittite agent form. Yet the best-preserved and most-cited Luwian example belonging to this group displays a feature that distinguishes it from the Hittite examples treated thus far: in one of the two cases, the suffix is added to the underlying common-gender noun.

- (9) KUB 9.6+ ii 14–15, CTH 759  
 āšša=ti                      ēlḫādu                      tappašantiš                      tiyammantiš  
 mouth.ACC.SG=3SG.REFL    wash.3SG.IMPV    sky.NOM.SG.C!    earth.NOM.SG  
 ‘Let Heaven and Earth wash their mouth(s)’.

The forms *tappašantiš* and *tiyammantiš* can be compared with *tappaš-* (n.) ‘sky, heaven’ and *tiyamm(i)-* (c.) ‘earth’ respectively. The Luwian language, unlike Hittite,

nearly eliminated the consonantal stems of the common gender, so the secondary vowel *-i-* is inserted between the morpheme *-ant-* and the nominative ending *-š*. This is, of course, yet another confirmation of the analysis of Hitt. *-anza* as */-ant-/ + /-s/*, advocated in Section Three: in the instance of Luwian, the comparable forms cannot be analysed as synchronic ablatives by any stretch of the imagination. Nevertheless, the synchronic derivation of *tiyammantiš* ‘earth’ from a common gender noun requires explanation. Is this just an overshot caused by its juxtaposition with the agent form *tappašantiš* ‘heaven’, a reflection of the historical neuter gender of Proto-Anatolian *\*deḡom* ‘earth’, or a derivational strategy to be distinguished from the formation of the agent form? And if one chooses the latter option, should *tappašantiš* be conversely grouped together with *tiyammantiš* as yet another manifestation of the same strategy?

An additional piece of evidence of the same kind comes from Lycian, an Anatolian language that is more closely related to Luwian than to Hittite. [Garrett 1990a: 275] mentions Lyc. *tesēti* (secondary variant *teseti*) ‘oaths’, which he contrasts with the collective noun *tasa* (n.) with the same meaning and analyses as formal counterparts of Hittite “ergative plural” forms. While the comparison is indeed formally plausible (Lyc. *-ēt-* ~ Luw *-ant-*), the singular forms *tesī*, occurring more than ten times in the Lycian corpus, can formally be analysed as nominative or accusative forms of the common gender. Since assigning *tes-ēt-i* to the paradigm of a common gender form would not be in line with the distribution of the “ergative case” proposed for Hittite, [Melchert 2004: 63] assigns all the available forms of *tesī* to dat.sg. This procedure, however, is entirely circular: as long as there are other instances of the forms in *-ant-* derived from common gender nouns in the Anatolian languages, one can alternatively assign *tes-ēt-i* to the same class and retain the conventional interpretation of *tesī* as a direct case form, as per [Neumann 2007: 350]. Decisive support for this alternative hypothesis comes from contextual evidence: as is clear from Garrett’s translation ‘And the Lycian oaths will strike/seize him’, Lyc. *tesēti/teseti* are expected to punish the evildoers. This is also the function of Hitt. *linkiyanteš*, the deities that pursue oath-breakers, as in (10), which formally represent the *-ant-*derivatives from Hitt. *lingai-* (c.) ‘oath’ [Melchert 2011a: 162], cf. [Kloekhorst 2008: 527]. Interpreting *tesēti* and *linkiyanteš* in a uniform fashion is clearly preferable to advancing two separate explanations for these formations.

- (10) KBo 6.34 iii 16–17, CTH 427  
 n=an                      linkiyanteš      eppir  
 CONN=(s)he.ACC.C    oath.NOM.PL    seize.3PL.PRET  
 ‘The Oaths seized him’.

While both *tesēti* and *linkiyanteš* function as agents of punishment, the application of the tests outlined in the previous section shows that they cannot be taken as inflectional forms. They do not modify the grammatical properties of *tesī/lingai-* and do not contain a suffix that could be freely attached to all the nouns of the common gender. As for the semantic difference between *tesī* and *tesēti*, or between *lingai-* and *linkiyanteš*, one can argue that it has to do with deification: i.e. ‘oath’ vs. ‘Oath’. Incidentally, this was not the only way to derive divine names from inanimate objects: Hitt. *taganzepa-* (c.), already mentioned in the previous section, can be translated as ‘(personified) Earth, goddess of the earth’. In this instance, too, the analysis in terms of the agent form is out

of the question: the attestations of Hitt. *taganzepa-* are not limited to the nominative, but we also find dat.sg. *da-ga-an-zi-pí*, abl. *da-ga-an-zi-pa-az* etc. [Kloekhorst 2008: 812–813]. This observation, in turn, undermines the identification of Luw. *tiyammant(i)-* ‘earth’ as the agent form, even though it is not attested in cases other than the nominative. It is easy to see that the context of (9) is fully compatible with the deification of *tiyammantiš*.

As was already mentioned in the previous section, the Hittite noun *nēpiš-* (n.) ‘sky, heaven’ has the common gender counterpart *nepiša-*. In this case, too, we are clearly dealing with a lexical derivative: besides the nominative from *[ne-p]í-ša-aš* in (11), acc.sg. *[ne-p]í-ša-an* is also attested [Kloekhorst 2008: 603]. In the light of these data, Luw. *tappašantiš* in (9) can be regarded as a borderline case. On the one hand, its syntactic behaviour represents a perfect match to that of the Hittite ergative forms, as it marks the gender-switch of the underlying neuter noun in agent position. On the other hand, as long as one assumes that *tiyammantiš* stands for ‘(personified) Earth’, it is hard to avoid the conclusion that *tappašantiš* functions in the same passage as ‘(personified) Heaven’, an equivalent of *nepiša-* rather than *nepišanza*.

(11) KUB 44.4 ii 2, CTH 520

[nep]išaš=za            GE<sub>6</sub>-i            wa(š)šiyat  
 sky.NOM.SG.C=REFL   black.ACC.SG.N   wear.3SG.PRET.MED  
 ‘[He]aven dressed in black’.

So far, we have considered only those cases where the derived *-ant-* stems are only attested in the nominative, thus sharing a property with the agent form. Yet this is not always the case. Hitt. *udnē* (n.) ‘country’ can be contrasted with *udniyant-* (c.) ‘people, population’; the meaning of the latter form was described as ‘*das belebt gedachte und als handelndes Subjekt fungierende Land*’ [Tischler 2010: 150]. There is no doubt that the suffix *-ant-* in *udniyant-* is identical to that in *linkiyanteš* ‘Oaths’, while the difference between its personifying and deifying functions is lexically conditioned (some scholars use “animatizing” as the blanket term). The morphosyntactic differences between the attested forms of the two nouns are more substantial: in addition to the nominative, *udniyant-* (c.) also features the genitive and dative forms, similarly to the Hittite nouns endowed with other gender-switching suffixes, such as *-a-* and particularly *-šepal/-zepa-*.

Table 5: Common gender derivatives of the Hittite words for ‘country’ and ‘earth’

	country (n.)	people of the country (c.)	earth (n.)	earth goddess (c.)
Acc. sg.	<i>udnē</i>	∅	<i>tēkan</i>	<i>taganzepan</i>
Gen. sg.	<i>udniyaš</i>	<i>udniyandaš</i>	<i>taknaš</i>	<i>taganzepaš</i>
Dat. sg.	<i>udniya</i>	<i>udniyanti</i>	<i>taknī</i>	<i>taganzepi</i>

The discussion of the forms above may seem hair-splitting and of little consequence if one merely takes into consideration its synchronic dimension, cf. [Garrett 1990: 288]. The rationale for addressing these fine distinctions in earnest is fleshing out the grammaticization cline involving the Anatolian *-ant-* forms. Hitt. *udniyant-* (c.) clearly patterns with lexical derivatives, Lyc. *tesēti* and Hitt. *linkiyanteš* do it less transparently, Luw. *tappašantiš* and *tiyammantiš* are amenable to different interpretations,



while a large group of the Hittite nouns in *-ant-* can unquestionably be subsumed under the category of the agent form. Since the grammaticization of morphemes is commonly acknowledged to occur far more frequently than their de-grammaticization, those Anatolian derivatives in *-ant-* that deviate in their function from the prototypical agent form must be taken as archaisms pending evidence to the contrary, cf. [Oettinger 2001: 312]. Conversely, the agent form typical of Hittite emerges as an innovation, which probably postdates the separation of the Anatolian group and was implemented to various extent in individual Anatolian languages.

## 6. Anatolian privative ergativity

One of the reasons why the grammaticization of the *-ant-* suffix for marking the agent form in Hittite is not universally accepted concerns the lack of clarity regarding the functional reasons for such a development. This can be illustrated with the citation from [Garrett 1990a: 278–279]: “The view that the ‘ergatives’ somehow originated as derived *ant-*stems requires ... the proposal that a constraint against A-function neuters somehow developed in Anatolian, and that somehow a derivational suffix with a specific semantic value was pressed into service to “animatize” neuters just in A function”. Indeed, if *ēšhar* could once function as subject of transitive clauses, what was the point in coining the new form *išhananza*?

In my view, the hypothesis that the Anatolian neuter nouns could originally be used in agent function need not be taken for granted. As argued in the previous section, the derivational animatizing suffix *-ant-* can be reconstructed for Proto-Anatolian, while existence of other animatizing derivatives can be demonstrated at least for Hittite (cf. Table Four and Table Five above). It stands to reason that the agent form was unlikely to undergo grammaticization in response to a constraint on neuter agents, which must have existed beforehand. It is more economical to assume that it came into being as a more predictable and regular alternative to the derivational animatizing forms of the common gender.

Furthermore, as duly noted in previous scholarship, the constraint on neuter nouns used as agents in Hittite does not constitute the only support for the split pattern of morphological alignment in this language. The Hittite clitic pronouns of the third person are likewise prevented from appearing in the agent position by the language structure (12), even though they can function as both direct objects (13) and intransitive subjects (14), and thus feature both the nominative and accusative case forms. In fact, [Garrett 1997] has demonstrated that the pro-drop syntax is also characteristic of those intransitive clauses where the underlying subjects fully control the situation and thus represent agents in the clause argument structure. This is a significant parallel to the constraint on neuter nouns, which also appears to be extended to some intransitive agents (cf. example (8) above). The main difference between the two constraints lies in the extension of the pro-drop to the clitics of the common gender, such as =*aš* ‘he’ in (13) and =*an* ‘him’ in (14). As shown in [Melchert 2011b], the similar constraint on anaphoric clitics functioning as agents is also present in Luwian and thus can be reconstructed for Proto-Anatolian.

(12) KUB 33.96 i 7, CTH 345

nu	<sup>d</sup> 10-ni	IGI-anda	idalawatar	šan[ <sup>h</sup> iskizzi]
CONN	Storm-god.DAT.SG	against	evil.ACC.SG	seek.3SG.PRES
	‘(He) s[EEKS] evil against the Storm-god’.			

(13) KBo 32.14 ii 7, CTH 789

mān=an                      pahhuenanza              arḫa              warnuzzi  
 OPT=(s)he.ACC.C    fire.NOM.SG.C!    away              burn.3SG.PRES  
 ‘May the fire burn him up!’

(14) KBo 3.4 i 11, CTH 61

nu=war=aš=za                      DINGIR-LIM-iš              DÙ-at  
 CONN=QUOT=(s)he.NOM.C=REFL    god.NOM.SG              become.3SG.PRET.MED  
 ‘He became a god’.

To be sure, this is not the only paradigmatic gap in the system of the Hittite clitic pronouns. As illustrated below, the “ergative” type of pro-drop involving the third person pronouns of the neuter gender contrasts with its perfectly “nominative” counterpart in the instance of the first and second person pronouns, for which only the object forms are attested. The third person clitics of the common gender reflect, strictly speaking, an intermediate tripartite system, even though one can argue that the presence vs. absence of a pronominal clitic reveals a more basic distinction than the variation in its form. Thus, the clitic pronouns offer another example of split-ergativity embedded into the Hittite grammar. It is comforting to observe that the splits in nominal and pronominal splits both conform to Silverstein’s hierarchy, where the higher vs. lower degree of animacy correlates with nominative vs. ergative case marking respectively, while the first and second person pronouns are treated as the most active among all the possible arguments.

Table 6: Direct Case Forms of Hittite Pronominal Clitics

	1sg.	2sg.	3sg.c.	3sg.n.	1pl.	2pl.	3pl.c.	3pl.n.
A	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
S	Ø	Ø	=aš	=at	Ø	Ø	=e, =at	=at
O	=mu	=ta	=an	=at	=at	=naš	=šmaš	=at

A less trivial similarity between the two types of Anatolian ergative constructions involves the presence of zero slots in the paradigm. This is self-evident for pronominal arguments, while in the instance of their nominal counterparts, the same conclusion follows from the reconstruction undertaken in the preceding section. Indeed, if one assumes the grammaticization of *-ant-* as the marker of agent form is a relatively recent phenomenon, this is tantamount to reconstructing an empty slot in the paradigms of the neuter nouns. This reconstruction is captured in the table below with reference to the Anatolian proto-language. A technical term that can be suggested for such a state of affairs is “privative ergativity”. There was no ergative case in Proto-Anatolian, but the forms *\*değom* ‘earth’ and *\*nebos* ‘sky’ can be analyzed as absolutive case forms.

Table 7: Direct Case Forms of Anatolian Nouns

	Anat. ‘cow’ (c.)	Anat. ‘sheep’ (c.)	Anat. ‘earth’ (n.)	Anat. ‘sky’ (n.)
A	*g <sup>w</sup> ou-s	*xowi-s	Ø	
S			*değom	*nebos
O	*g <sup>w</sup> ō-m	*xowi-m		

[Garrett 1990b: 228] describes the Hittite pronominal system as somewhat redundant: “One problem raised by the enclitic subjects is thus essentially functional: given their relative inutility, why were they created in Proto-Anatolian at all?” The reconstruction of the parallel patterns of private ergativity in both nominal and pronominal declension is conducive to answering Garrett’s question. If the third person subject clitic pronouns are indeed secondary in the Anatolian languages, the first to spread to the subject position was arguably the neuter clitic *\*=od > =at*, which came to be associated with the same core thematic roles as the neuter nouns in the absolutive case it stood for. Since the animate third person pronouns occupy the intermediate position between the first/second person pronouns and the inanimate third person pronouns in Silverstein’s hierarchy, creating the intermediate tripartite system for this class was one of the possible options within the system of pronominal clitics.

There is, however, no guarantee that the pronominal system in Table Six reflects an Anatolian innovation. It is true that the subject clitic pronouns are absent in the Ancient (non-Anatolian) Indo-European languages and thus cannot be reconstructed for Core Indo-European. Yet, as long as one accepts that the Anatolian group was the first one to separate itself from the Proto-Indo-European language community, as most people do nowadays, the elimination of the subject clitic pronoun can easily be regarded as an innovation of Core Indo-European. The functional motivation for this process would be the elimination of split-ergativity in favour of the nominative alignment in Core Indo-European.

The pronominal system does not yield us any real cues for determining whether the Anatolian alignment pattern is more archaic than the Core Indo-European one or vice versa. More information, however, can be gleaned from the nominal system. Building up on the discussion of [Ivanov and Gamkrelidze 1995], [Schmidt 2010: 247–250] stresses the similarity between the Hittite pairs of neuter nouns and their common gender animatizing derivatives on the one hand and the near-synonymous pairs of reconstructed Proto-Indo-European neuter (“inactive”) and masculine/feminine (“active”) nouns on the other hand. In the latter case, there is no formal relationship between the members of the relevant pairs, which renders them fully unpredictable, as one can infer from the table below.

Table 8: Opposition between neuter and masculine/feminine quasi-synonyms in PIE

<i>*nebhos</i> ‘(cloudy) ‘sky’ (n.)	<i>*paxur</i> (peh <sub>2</sub> ur) ‘fire’ (n.)	<i>*wodor</i> ‘water’ (n.)
<i>*dyēu-</i> ‘sky’ (m.)	<i>*ṅgni-</i> ‘fire’ (m.)	<i>*xap-</i> (h <sub>2</sub> ep-) ‘(flowing) water’ (f.)

Although the detailed discussion of the Early Indo-European (Indo-Anatolian) case marking must remain beyond the scope of this paper, one can conclude it by advancing a hypothesis about the inability of “inactive” nouns to act as agents at that stage. If so, they must have been replaced by their “active” counterparts in subject position in the transitive clauses (as well as in some intransitive clauses requiring agents as subjects). In the history of the Anatolian group, lexical suppletion was gradually replaced by the morphological derivation of the “active” nouns, such as Hitt. *taganzepa-* (c.) or *udniyant-* (c.) ‘population’. Under the proposed scenario, the grammaticized forms in *-anza* merely represent the final step in transforming a linguistic system characterized by pri-

vative ergativity into a full-fledged split-ergative system. Regardless of the validity of my tentative hypothesis projecting privative ergativity back into Early Indo-European, its status as a common Anatolian phenomenon appears to be reasonably secure, and so is the functional rationale for creating the Hittite agent form.

## References

1. Benveniste, Émile. Les substantifs en *-ant-* du Hittite // *Bulletin de la Societe de Linguistique de Paris* 57 (1962). P. 44–51.
2. Carruba, Onofrio. Le notazioni dell’agente animato in anatolico (e l’ergativo). O. Carruba, ed., *Per una grammatica ittita*. Pavia: Iuculano; 1992:61–98.
3. Comrie, Bernard. *Language Universals and Linguistic Typology*. Second Edition. Chicago: University of Chicago; 1989.
4. Dardano, Paola. Zur anatolischen Morphosyntax: das Suffix *-ant-* und seine Bildungen. A. Süel, ed., *Acts of the VIIth International Congress of Hittitology*. Ankara, 2010:173–88.
5. Gamkrelidze, Tamaz V. and Vyacheslav Vs. Ivanov. *Indo-European and Indo-Europeans*, Vol. I. Berlin: Mouton de Gruyter; 1995.
6. Garrett, Andrew. The Origin of NP Split Ergativity. *Language*. 1990(66):261–296. Cited as Garrett 1990a.
7. Garrett, Andrew. Hittite Enclitic Subjects and Transitive Verbs. *Journal of Cuneiform Studies* 1990(42):227–242. Cited as Garrett 1990b.
8. Garrett, Andrew. Wackernagel's Law and unaccusativity in Hittite. A. Halpern and A. Zwicky, eds, *Approaching second: Second position clitics and related phenomena*. Stanford, CA: Center for the Study of Language and Information, 1997:85–103.
9. Goedegebuure, Petra. Split-ergativity in Hittite. *Zeitschrift für Assyriologie und vorderasiatische Archäologie* 2013;(102/2):270–303.
10. Haspelmath, Martin. *Understanding Morphology*. London: Arnold; 2002.
11. Hoffner, Harry A. and H. Craig Melchert. *A grammar of the Hittite language*. Winona Lake, IN: Eisenbrauns; 2008.
12. Jasanoff, Jay. The impact of Hittite and Tocharian: Rethinking Indo-European in the 20th century and beyond. J. Klein et al., eds, *Handbook of Comparative and Historical Indo-European Linguistics*. Berlin: de Gruyter; 2017:220–238.
13. Josephson, Folke. Semantics and typology of Hittite *-ant-*. Clackson J. and B. Olsen, eds, *Indo-European Word Formation*. Copenhagen, Museum Tusulanum Press; 2004:91–118.
14. Kassian, Alexei and Ilya Yakubovich. Muršili II’s Prayer to Telibinu. D. Groddek and M. Zorman, eds, *Tabularia Hethaeorum: hethitologische Beiträge Silvin Košak zum 65. Geburtstag*. Wiesbaden: Harrassowitz; 2007:423–54.
15. Kloekhorst, Alwin. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden: Brill; 2008.
16. Laroche, Emmanuel. Un “ergatif” en indo-europeen d’Asie Mineure // *Bulletin de la Société de Linguistique de Paris*. 1962;(57):23–43.
17. Melchert, H. Craig. *A Dictionary of the Lycian Language*. Ann Arbor: Beach Stave; 2004.
18. Melchert, H. Craig. The Problem of the Ergative Case in Hittite. M. Mazoyer et al., eds, *Grammatical Case in the Languages of Middle East and Europe*. Chicago: Oriental Institute; 2011. Cited as Melchert 2011a.

19. Melchert, H. Craig. Enclitic Subject Pronouns in Hieroglyphic Luvian // *Aramazd (Armenian Journal of Near Eastern Studies)*. 2011;(6/2):73–86. Cited as Melchert 2011b.
20. Neu, Erich. Zum Alter der personifizierenden *-ant-* Bildung des Hethitischen. Ein Beitrag zur Geschichte der indogermanischen Genuskategorie. *Historische Sprachforschung*. 1989;(102):1–15.
21. Neumann, Günter. *Glossar des Lykischen* (prepared for publication by Johann Tischler). Wiesbaden: Harrassowitz; 2007.
22. Oettinger, Norbert. Neue Gedanken über das *nt-*Suffix // O. Carruba and W. Meid, (eds.), *Anatolisch und Indogermanisch*. Innsbruck: Innsitut der Sprachwissenschaft; 2001: 301–316.
23. Patri, Sylvain. *L’alignement syntaxique dans les langues indo-européennes d’Anatolie*. Wiesbaden: Harrassowitz; 2007.
24. Pooth, Roland, Peter Alexander Kerkhof, Leonid Kulikov, and Jóhanna Barðdal. The origin of non-canonical case marking of subjects in Proto-Indo-European: Accusative, ergative, or semantic alignment. *Indogermanische Forschungen*. 2019;(124/1):245–263.
25. Pooth, Roland and Veronica Orqueda. Alignment Change and the Emergence of the Thematic Conjugation from Proto-Indo-European to Indo-European: A Wedding of Hypotheses. *Transactions of the Philological Society*, 2021. To appear.
26. Rizza, Alfredo. Contributi allo studio dell’ergatività in anatolico: basi teoreticotipologiche. Sopra alcune recenti pubblicazioni. *Atti del Sodalizio Glottologico Milanese* (n.s.). 2010;(3): 144–162, 2010.
27. Schmidt, Karl Horst. 2010. *Miscellanea Anatolica*. J. Klinger et al., eds, *Investigationes Anatolicae: Gedenkschrift für Erich Neu*. Wiesbaden: Harrassowitz; 2010: 241–52.
28. Shatskov, Andrej. The ergative case in Hittite. *Acta Linguistica Petropolitana* VII/1 (2011): 143–52. (In Russian with an English abstract).
29. Silverstein, Michael. Hierarchy of features and ergativity. R.M.W. Dixon, ed., *Grammatical Categories in Australian Languages*. New Jersey: Humanities Press; 1976: 112–171.
30. Tischler, Johann. *Hethitisches Handwörterbuch, mit dem Wortschatz der Nachbarsprachen*. Innsbruck: Institut der Sprachen und Literaturen der Universität Innsbruck; 2001.
31. Tischler, Johann. *Hethitisches Etymologisches Glossar*, Lieferung 15: U. Innsbruck: Institut der Sprachen und Literaturen der Universität Innsbruck; 2010.
32. Yakubovich, Ilya. *Sociolinguistics of the Luwian Language*. Leiden: Brill; 2010.

## О грамматическом и словарном аспектах слова в современном китайском языке

**Ксения Владиславовна Антонян**

Институт языкознания РАН, Москва, Россия,  
kvantonian@yandex.ru, ORCID 0000-0002-8531-3641

**Аннотация.** Текст может быть разбит на единицы, обладающие грамматическими свойствами слова. Основными критериями выделения слова являются следующие: (1) синтаксическая самостоятельность; 2) неразрывность; 3) цельно-оформленность. Однако в китайском языке выделенные по этим критериям единицы («грамматические слова») далеко не всегда являются единицами словаря («инвентарными словами»). В статье этот феномен рассматривается на примере результативных конструкций.

**Ключевые слова:** китайский язык, слово, единицы текста, единицы словаря, результативные конструкции

**Abstract.** Word as a grammatical unit is characterized by syntactic freedom, inseparability of its components; grammatical markers are related to the word as a whole. On the other side, a word is a unit of a certain inventory (lexicon); it is reproduced rather than produced in speech. In Mandarin Chinese and typologically similar languages grammatical words are not necessarily units of lexicon. In this paper this phenomenon is studied on the material of resultative verb compounds.

**Keywords:** Mandarin Chinese, word, grammatical unit, unit of lexicon, resultative verb compound

Вопрос об основных единицах языка и, в частности, о слове принадлежит к числу основополагающих и, возможно, «вечных» вопросов описания китайского языка.

Вопросу о том, что такое слово и каковы критерии его выделения, посвящено огромное количество работ как в рамках общего языкознания, так и в рамках изучения отдельных языков. В настоящее время этот вопрос сохраняет свою актуальность, а его рассмотрение выходит на новый уровень в свете изучения национальных лингвистических традиций, с одной стороны, и привлечения данных смежных наук – психолингвистики и нейролингвистики – с другой [Алпатов 2016; Алпатов 2020].

Зоя Михайловна Шаляпина писала: «Проблема соотношения словаря и грамматики представляет собой одну из наиболее кардинальных лингвистических проблем, и от того, как она решается в рамках конкретных лингвистических описаний, существенно зависят их состав и строение» [Шаляпина 1991: 42].

В китайском языке (КЯ) слово как единица языка не является столь очевидной и интуитивно понятной единицей, как в европейских языках, грамматики которых являются, безусловно, словоцентрическими.

Для КЯ наиболее очевидной и интуитивно понятной является другая единица – слогоморфема, или значимый однослог (на письме ей соответствует иероглиф).

Единицами, аналогичными слову, являются в КЯ, во-первых, синтаксически самостоятельные слогоморфемы, а во-вторых, их устойчивые сочетания (в основном двусложные). Выделение уровня слов ставит ряд вопросов – о критериях ограничения этого уровня от уровня морфем, с одной стороны, и уровня словосочетаний, с другой стороны.

Помимо задачи выделения единиц основных языковых уровней на основании тех или иных критериев, перед лингвистом стоит не менее важная задача – описание свойств этих единиц, которые оказываются неоднородными (даже в рамках одного уровня) и делятся на сильно различающиеся подклассы. Мы полагаем, что только исследование и подробное описание свойств единиц китайского языка, функционально подобных слову, может дать нам подлинное понимание их природы.

В данной работе мы рассмотрим соотношение двух аспектов слова: грамматического (морфологического) и словарного (лексикологического). Это значит, что мы рассмотрим соотношение слова как грамматической единицы (единицы текста) и слова как единицы инвентаря воспроизводимых единиц (словаря).

Текст может быть разбит на единицы, обладающие грамматическими свойствами слова. Основными критериями выделения слова являются, как правило, следующие: 1) синтаксическая самостоятельность; 2) цельнооформленность: морфологический показатель оформляет слово в целом; 3) невозможность вставки единицы того же уровня, то есть другого слова.

Пример разбиения на такие единицы текстов, написанных на современном китайском языке и других языках Азии и Африки, приведен, в частности, в работе «Квантитативная типология языков Азии и Африки». Авторы теоретической части данной монографии – С.Е. Яхонтов и В.Б. Касевич – исходят из того, что «слово характеризуется, во-первых, относительной цельностью и, во-вторых, относительной свободой, способностью к самостоятельному употреблению» [КТ 1982: 18]. Как основной авторы используют критерий, предложенный П.С. Кузнецовым: слово определяется как звуковая последовательность, в которую не может быть включена другая последовательность того же уровня. Используется также критерий цельнооформленности, введенный А.И. Смирницким (морфологический показатель оформляет слово в целом, а не его часть или словосочетание).

Однако выделенные из текста в соответствии с грамматическими критериями единицы (их можно назвать «грамматическими словами»; в текстах на СКЯ значительную часть таких единиц составляют двуслоги) далеко не всегда являются «инвентарными словами», то есть не являются устойчивыми и воспроизводимыми. Более того, в ряде случаев «грамматическое слово» может оказываться непонятым (даже носителю языка) вне контекста.

Рассмотрим этот феномен на примере результативных конструкций (РК) в современном китайском языке (СКЯ).

РК состоит из двух компонентов, первый из которых выражен глаголом и обозначает действие, а второй (так называемый дополнительный элемент результата) – глаголом или прилагательным, обозначающим результат этого действия, ср. *kan-wan* ‘читать’-‘закончить’ – ‘дочитать до конца’, *xi-bai* ‘стирать’-‘белый’ – ‘отстирать добела (дочиста)’. Характерная черта РК – способность образовывать так называемые формы потенциального наклонения (ФПН), обозначающие возможность или невозможность достижения действием результата. ФПН образуются путем инфлексии морфем *de* (исходное значение этой служебной морфемы – ‘получить’, ‘достигнуть’) или *bu* ‘не’, ср. *kan-de-wan* ‘можно дочитать’, *kan-bu-wan* ‘невозможно дочитать’.

Мы употребляем термин «конструкция» в том значении, в котором его вводит С. Е. Яхонтов: «Две единицы, связанные грамматическими отношениями, составляют конструкцию (являясь ее компонентами). Вместе с тем эти две единицы образуют сложную единицу того же или более высокого уровня, которая в свою очередь может быть компонентом другой конструкции» [Яхонтов 1982: 157; Яхонтов 2016: 314]. Этот термин позволяет уйти от дихотомии «слово – словосочетание», строгое следование которой нередко приводит к вынужденному разделению однотипно устроенных сложных комплексов.

Т.П. Задоевко, которая была первым в отечественном языкознании исследователем, подробно описавшим РК, первоначально использовала для их обозначения термин «результативное глагольное образование» [Задоевко 1958: 76-77], который также позволяет уйти от дихотомии «слово – словосочетание». В дальнейшем она стала использовать для их обозначения термин «результативный глагол», получивший широкое распространение как в научной, так и в педагогической практике [Задоевко, Хуан 1986: 266].

В западном китаеведении для обозначения РК, как правило, используется термин *resultative verb compound*. Чжао Юаньжэнь называет их *verb-complement compounds* [Chao 1968: 435]. Понятие *compound* шире, чем понятие слова, и также позволяет уйти от ограничений, накладываемых дихотомией «слово – словосочетание». Некоторые лингвисты пользуются термином *les verbes résultatifs* ‘результативные глаголы’ [Cartier 1972].

В китайской научно-педагогической традиции, представленной, в частности, в различных изданиях авторитетной «Практической грамматики современного китайского языка» [Liu et al. 1983; Liu et al. 2012], рассматриваемые конструкции называются *dongbu jiegou* ‘конструкция глагол – дополнительный элемент’ и рассматриваются в разделе «Синтаксис». Как синтаксические конструкции рассматривали РК видные китайские лингвисты Гао Минкай и Ван Ли [Wang 1989: 255-261]. Впрочем, некоторые китайские лингвисты рассматривают РК как сложные слова: в работах Чжоу Чимина они называются *shidongxing fushi dongci* ‘сложные глаголы каузативного характера’ (Zhou 1957: 187).

Эти терминологические расхождения неслучайны.

РК удовлетворяют критериям «грамматического слова», поскольку они синтаксически самостоятельны, цельнооформленны (отрицание и грамматические по-



казатели относятся к РК в целом) и не допускают никаких вставок, кроме инфиксов *de/bu* (так образуются ФПН). Однако РК образуют открытый класс, и лишь часть их бывает представлена в словарях. Как правило, это устойчиво воспроизводимые РК. Все РК, по-видимому, и не могут быть учтены словарем. То, что они образуются свободно, их класс открыт, а значение, как правило, выводится из значений их компонентов, сближает их со словосочетаниями.

Интересно отметить, что некоторые типы РК заведомо «несловарны». Такие – заведомо «несловарные» – РК и будут предметом рассмотрения в данной работе. Мы рассмотрим три типа «несловарных» РК.

Во-первых, это окказионализмы и авторские неологизмы. Ср. *kan-bao* ‘смотреть’-‘сытый’ – ‘насытиться путем смотрения’ по аналогии с *chi-bao* ‘наестся’ и *he-bao* ‘напиться’.

В «Словаре сочетаемости глаголов» [DCYF 1987] морфема *bao* ‘сытый’ приведена в разделах *dongjieshi* ‘результативная форма’ словарных статей только двух глаголов: *chi* ‘есть’ и *he* ‘пить’ (то есть возможны РК *chi-bao* ‘наестся’ и *he-bao* ‘напиться’). Однако в романе Цянь Чжуншу «Осажденная крепость» нам встретился следующий пример:

*Xiao Fang, ni dao mei sheng bing? O, wo mingbai le! Bao xiaojie xiuse ke can, ni kan-bao le bu yong chi fan le.*

«Сяо Фан, ты не заболел? О, я понял! Барышня Бао уж такая аппетитная, вот ты и насмотрелся так, что теперь можно и не обедать» [Qian 1991: 19].

В предложении:

*ni kan- bao le bu yong chi fan le*  
ты смотреть сытый ПСВ/МЧ не нужно есть пища МЧ

морфема *bao* ‘сытый’ присоединяется к глаголу *kan* ‘смотреть’; образуемая РК обозначает буквально ‘насытиться с помощью смотрения’. Ясно, однако, что выражаемый в данном примере смысл является очень нестандартным, а сочетание *kan-bao* – очень нетипичным. Исчислить и предусмотреть появление сочетаний такого типа, по-видимому, невозможно. Не вполне ясно также, имеет ли смысл отражать сочетание *kan-bao* в словаре: оно тесно связано с ситуацией, может быть понято только в определенном контексте и, будучи весьма остроумным, играет всеми красками только в этом контексте.

Рассмотрим второй тип РК, которые оказываются непонятными, будучи вычлененными из предложения, в котором они выполняют роль сказуемого.

Это синтаксические структуры, которые в работах китайских грамматистов описываются как предложения, в которых глагол присоединяет дополнение не напрямую, а только через дополнительный элемент (ДЭ) результата [Liu et al. 1983: 334-335; DCYF 1987: 6; Liu et al. 2012: 534-535]. Глагол сам по себе, не оформленный ДЭ результата, не мог бы присоединить такое дополнение.

Вот примеры таких предложений.

*Haizi ba keren dou ku- rao le*  
ребенок ПВД гости ч плакать убежать ПСВ/МЧ

«Ребенок плакал так, что гости убежали» (букв. «Ребенок своим плачем вынул гостей убежать»)

*Ta ba toufa dou ao- bai le.*  
он ПВД волосы Ч страдать белый ПСВ/МЧ

«Он поседел от страданий» (букв. «Своими страданиями он воздействовал на волосы так, что они побелели»).

Глаголы *ku* ‘плакать’ и *ao* ‘страдать’ сами по себе не могут присоединять дополнения *keren* ‘гости’ и *toufa* ‘волосы’ (*\*ku keren*, *\*ao toufa*). Дополнения в этих предложениях семантически связаны не с глаголом, а с дополнительным элементом (вторым компонентом РК). В СКЯ дополнительные элементы *pao* ‘бежать’ и *bai* ‘белый’ также не могут присоединять такие дополнения (*\*pao keren*, *\*bai toufa*), поэтому можно говорить только о семантической связи. Но в древнекитайском языке это было возможно: глаголы различных типов, в том числе и непереходные, а также и прилагательные могли использоваться в каузативном значении и присоединять при этом дополнение, обозначающее объект воздействия [Яхонтов 2016 (1969): 172-178]. В древнекитайском языке форма *pao keren* имела бы значение ‘заставить гостей убежать’, а форма *bai toufa* – ‘сделать волосы белыми’. Возможность каузативного употребления непереходных глаголов, в том числе глаголов движения, а также прилагательных на более ранних стадиях развития КЯ и обеспечила возможность формирования таких не вполне обычных РК.

Как отмечает А.Картье, по смыслу подобные предложения очень похожи на предложения с дополнительным элементом степени (ДЭС): эти конструкции легко трансформируются одна в другую. Ср. следующие примеры, где под номером (1) приведены предложения с ДЭС, а под номером (2) – предложения со сказуемым-РК.

(1) а. *Kan de yanjing dou hua le.*  
читать Ч-ДЭС глаза Ч рябить МЧ

б. *Ba yanjing dou kan- hua le.*  
ПВД глаза Ч читать рябить ПСВ/МЧ  
«В глазах зарябило от чтения».

(2) а. *Shuo de shetou diao le.*  
говорить Ч-ДЭС язык упасть МЧ

б. *Ba shetou shuo- diao le.*  
ПВД язык говорить упасть ПСВ/МЧ  
«От болтовни язык уже отваливается» [Cartier 1972, p. 36-37].

Итак, конструкции *ku-pao* и *ao-bai* обладают признаками грамматических слов. Но тем не менее они не являются и не могут являться единицами инвентаря, так как вне контекста они непонятны. Эксперименты с информантами показали, что на слух при изолированном произнесении данные конструкции не опознаются; если же они предъявлялись в иероглифической записи, информанту далеко не сразу удавалось придумать ситуацию, в которой данная конструкция могла бы быть употреблена.

Подобные конструкции существуют и функционируют как слова только в составе предложения, будучи синтаксически и семантически связанными с конкрет-

ными актантами. Они не являются, по терминологии В.А. Плуногяна, автономными [Плуногян 2000: 18].

Класс подобных конструкций является, судя по всему, открытым. Ср. следующие примеры:

*Han- ya le sangzi*  
кричать сильный ПСВ горло  
«Осип от крика» (букв. «криком сделал горло осипшим»).

*Chi- huai le duzi*  
есть- плохой ПСВ живот  
«Получил расстройство желудка» (букв. «едой испортил желудок»).

*Va cao ba- teng le shou.*  
вытягивать траву вытягивать- болеть ПСВ руки  
«Руки заболели от прополки».

*qi- beng le duzi*  
сердиться лопнуть ПСВ живот  
«страшно рассердиться» (букв. «живот лопнул от злости»).

*qi- zha le fei*  
сердиться взорваться ПСВ легкие  
«страшно рассердиться» (букв. «легкие взорвались от злости»).

*pao- duan le tui*  
бегать ломать ПСВ ноги  
«оттоптать все ноги» (букв. «бегая, переломать ноги»), ср.

*Wei le gei ni mai zhe shuang xie, wo pao-duan le tui, keshi ni hai bu lingqing.*  
«Чтобы купить тебе эти туфли, я все ноги оттоптал, а ты всё недовольна».

*xiao- wan le yao*  
смеяться согнутый ПСВ поясница  
«согнуться от смеха» [Liu et al. 1983: 334-335; DCYF 1987: 6; Liu et al. 2012: 536-537].»

Конструкции *han-ya*, *chi-huai*, *ba-teng*, *qi-beng*, *qi-zha*, *pao-duan*, *xiao-wan* по своим грамматическим свойствам являются словами. Тем не менее они не являются «инвентарными словами», поскольку вне контекста они просто непонятны.

Рассмотрим третий тип «неинвентарных» РК. Это РК, второй компонент которых представляет собой глагол, управляющий именной группой со значением места. С.Е. Яхонтов и Н.Н. Коротков называют их глагольно-предложными сочетаниями [Яхонтов 1957: 96, Коротков 1968: 65-66]. Рассмотрим, например, предложения, содержащие конструкцию «глагол – *dao/shang* – именная группа со значением места», ср.

*Wo ba ta song- dao le menkou.*  
я ПВД он провожать достигнуть ПСВ дверь  
«Я проводил его до двери».

*Ba luò shuǐ de haizi jiù-shàng àn.*  
ПВД упасть вода АЧ ребенок спасти на (ДЭ) берег

«Вытащить упавшего в воду ребенка» (букв. «спасти упавшего в воду ребенка на берег»).

Примерно ли считать глагольные комплексы *song-dao* ‘проводить’-‘достигнуть’ и *jiù-shàng* ‘спасти’-‘поднять(ся) на’ результативными конструкциями? Мы полагаем, что да, поскольку они обладают характерным для РК свойством – возможностью инфиксировать *de/bu* с образованием ФПН. ФПН, образованные от глагольно-предложных конструкций, встречаются сравнительно редко, но они возможны. Ср. следующие примеры:

*Tianhei qián zánmen gān-de dào xián- lǐ ma?*  
темнота перед мы спешить- ИНФ ДЭ уезд- в ВЧ  
«Успеем ли мы до темноты вернуться в уезд?» [ХНВС 2007: 127]

*Kānlái, qī diǎn zhōng shí gān- bǔ- dào*  
по-видимому семь часов времени Св спешить ИНФ ДЭ  
*chǎng- lǐ lè.*  
завод- в МЧ  
«По-видимому, мы действительно не успеем приехать на завод к семи часам» [DCYF 1987: 285].

Итак, функционирующие в роли сказуемого морфемные комплексы *song-dao* ‘проводить’-‘достигнуть’ и *jiù-shàng* ‘спасти’-‘поднять(ся) на’ грамматически характеризуются теми же свойствами, что и обычные РК, а именно: (1) произносятся беспаузно, (2) неспособны к разъединению (за исключением возможности инфиксировать *de/bu*), (3) обладают свойством цельнооформленности (показатель совершенного вида *le* оформляет конструкцию в целом). Таким образом, глагольно-предложные комплексы представляют собой один из подтипов РК и обладают грамматическими свойствами слов. Тем не менее они не являются единицами инвентаря. Они семантически и синтаксически неполны вне определенных элементов контекста (неавтономны) и функционируют как слова только в контексте.

Рассмотренные примеры показывают, что в СКЯ не все единицы текста, обладающие грамматическими свойствами слова, являются в то же время единицами инвентаря. Более того, в целом ряде случаев грамматические слова оказываются синтаксически и семантически неполными вне контекста и поэтому вообще не имеют шансов превратиться в единицы инвентаря. Как представляется, приведенные наблюдения еще раз подтверждают тезис о несовпадении морфологического (грамматического) и лексикологического (словарного) аспектов слова в современном китайском языке и в типологически сходных с ним языках [Касевич 1988: 173-176]. Кроме того, они иллюстрируют неоднородность класса РК по параметру «словарности»: некоторые типы РК устойчивы и воспроизводимы, в то время как другие – заведомо «несловарны».

## Список сокращений

- АЧ – атрибутивная частица  
ВЧ – вопросительная частица  
ДЭ – дополнительный элемент  
ИНФ – инфикс  
КТ – «Квантитативная типология языков Азии и Африки»  
КЯ – китайский язык  
МЧ – модальная частица  
ПВД – предлог, вводящий дополнение  
ПСВ – показатель совершенного вида  
ПСВ/МЧ – показатель совершенного вида или модальная частица (в некоторых контекстах ПСВ *le* и МЧ *le* неразличимы)  
РК – результативная конструкция  
Св – связка  
СКЯ – современный китайский язык  
ФПН – формы потенциального наклонения  
Ч – частица  
Ч-ДЭС – частица, вводящая дополнительный элемент степени

## Источники

- DCYF – *Dongci yongfa cidian* (Словарь сочетаемости глаголов). Shanghai, 1987.  
ХНВС – *Xiandai Hanyu babai ci* (Восемьсот слов современного китайского языка). Beijing, [1981] 2007.  
Qian, Zhongshu. *Wei cheng* (Осажденная крепость). Beijing, 1991.

## Литература

1. Алпатов В. М. Проблема слова и психолингвистика // *Вопросы психолингвистики*. 2016. № 2 (28). С. 26–35.
2. Алпатов В. М. Японская лингвистическая традиция XIX-XX вв. как синтез национальной традиции и влияния западной науки // *Научный результат. Вопросы теоретической и прикладной лингвистики*. 2020. Т. 6, № 1. С. 12–19.
3. Задоенко Т. П. Результативное глагольное образование в современном китайском языке. // *Советское китаеведение*. 1958. № 2. С. 76–84.
4. Задоенко Т. П., Хуан Шуин. *Основы китайского языка. Основной курс*. М., 1986.
5. Касевич В. Б. *Семантика. Синтаксис. Морфология*. М., 1988.
6. КТ – *Квантитативная типология языков Азии и Африки*. Под ред. В. Б. Касевича и С. Е. Яхонтова. Л., 1982.
7. Коротков Н. Н. *Основные особенности морфологического строя китайского языка*. М, 1968.
8. Плунгян В. А. *Общая морфология. Введение в проблематику*. М., 2000.

9. Шаляпина З. М. Грамматика и ее соотношение со словарем при словоцентрическом подходе к языку (на опыте формализованного лингвистического описания) // *Вопросы языкознания*. 1991. № 5. С. 42–54.
10. Яхонтов С. Е. *Категория глагола в китайском языке*. Л., 1957.
11. Яхонтов С. Е. [1969] Выражение каузативности в древнекитайском языке // *Проблемы китайского и общего языкознания. К 90-летию С. Е. Яхонтова*. Отв. ред. Е. Н. Колпачкова. СПб., 2016. С. 171–181.
12. Яхонтов С. Е. Основные понятия грамматики изолирующих языков // *Теоретические проблемы восточного языкознания*. Отв. ред. В. М. Солнцев, И. Ф. Вардуть, Ю. Я. Плам. Ч. II. М., 1982. С. 156–163.
13. Яхонтов С. Е. [1982] Основные понятия грамматики изолирующих языков // *Проблемы китайского и общего языкознания. К 90-летию С. Е. Яхонтова*. Отв. ред. Е. Н. Колпачкова. СПб., 2016. С. 313–318.
14. Cartier, Alice. *Les verbes résultatifs en chinois moderne*. Paris, 1972.
15. Chao, Yuen-ren. *A Grammar of Spoken Chinese*. Berkeley, 1968.
16. Liu Yuehua deng zhu. *Shiyong xiandai Hanyu yufa*. (Лю Юэхуа и др. *Практическая грамматика СКЯ*). Beijing, 1983.
17. Liu Yuehua deng zhu. *Shiyong xiandai Hanyu yufa*. (Лю Юэхуа и др. *Практическая грамматика СКЯ*). Beijing, 2012.
18. Wang Li. *Hanyu yufa shi (История грамматической системы китайского языка)*. Beijing, 1989.
19. Zhou Chiming. Hanyu de shidongxing fushi dongci (Сложные глаголы каузативного характера в китайском языке) // *Shandong daxue xuebao (Ученые записки Шаньдунского университета)*. 1957. № 1. С. 187–210.

## Грамматическая модель и фонетическая система классического арабского языка

**Анна Григорьевна Белова**

Институт востоковедения РАН, Москва, Россия,  
himyarit@yandex.ru, ORCID 0000-0001-6748-0342

**Аннотация.** Грамматическая модель классического арабского языка, сложившаяся на основе трехсогласного корня и способа внутренней флексии, проявляет тенденцию к единообразию и регулярности своих форм. Фонетическая система классического арабского языка, сложившаяся на основе строгих правил слога и позиционных чередований, также проявляет тенденцию к устойчивости и регулярности своих законов. Столкновение грамматической модели и фонетического правила приводит к нарушению грамматической модели как по синхронным, так и по историческим причинам. Грамматическая модель находит соответствующие способы обеспечить максимальную регулярность форм.

**Ключевые слова:** модель, парадигма, фонетическое правило, внешняя/внутренняя флексия, глагол, корень, основа, классический арабский язык

**Abstract.** The grammatical model of Classical Arabic was formed on the basis of the three-consonantal root and internal inflection. Such a model tends to be uniform and regular in its forms. The phonetic system of Classical Arabic was formed on the basis of strict syllable rules and positional sound alternation. This system shows a tendency to the stability and regularity of its laws. The collision of the grammatical model and phonetic rule leads to a violation of the grammatical model. The reasons for the collision can be both historical and synchronous. The grammatical model finds appropriate ways to provide maximum regularity of their forms.

**Keywords:** model, paradigm, phonetic rule, external/internal inflexion, verb, root, stem, Classical Arabic

Грамматическая модель, отражающая особенности классического арабского языка на синхронном уровне, не может быть полностью понята без учета истории формирования основных единиц, с которыми она работает. Также грамматическая модель не может быть понята без учета особенностей фонетической системы арабского языка. Анализ модели и ее свойств и особенностей опирается на предшествующие исследования автора в области формирования корня и его структуры на сравнительном материале арабского и других семитских языков [Белова 1991; 1993; 2002; 2004; Belova 2012; 2015].

\*\*\*

Синхронная морфологическая система классического арабского языка (далее сокр. КЛАЯ) представлена двумя парадигмами способов формообразования и деривации:

парадигмой способов, внешних по отношению к корневой основе слова (различные виды аффиксов);

парадигмой внутренних способов, традиционно называемых «внутренней флексией», которая затрагивает структуру и фонологический состав корневой основы<sup>1</sup>.

В КЛАЯ наиболее представительной в этой области является глагольная система.

Категориальные значения глагольных форм, как правило, получают выражение в сложном сочетании этих двух способов. Ср., например, финитные формы префиксального и суффиксального спряжения глагола \**drs* «учить что-либо, изучать что-либо» первичной (I stem) и производной (II stem) основ:

(а) оппозиции основ префиксального и суффиксального спряжения

	imperf.	perf.
I stem	<i>-drus-</i>	<i>daras-</i>
II stem	<i>-darris-</i>	<i>darras-</i>

I: «учить»; II: «обучать, преподавать», оппозиция формообразовательная и оппозиция деривационная (II stem – интенсив, транзитив) выражены внутренней флексией (изменением вокализма и структуры основы);

(б) категориальные оппозиции лица, числа, рода, наклонения выражены внешним способом, префиксально и окончаниями: *ya-drus-ū-na* «они учат (что-либо)», где *ya-* префикс – показатель 3 л., *-ū#* показатель мн. числа м.рода, *-na* – показатель наклонения; также в отглагольной деривации: *dāris-u-n* (part. act. I stem, m. sg. nom., indefinit.) «учащий(-ся), изучающий» и *ma-drūs-u-n* (part. pass. I stem), m., sg., nom., indefinit.) «изучаемый, изученный», как видно из сопоставления двух производных форм, сочетаются внутренний и внешний способы.

Поскольку в системе КЛАЯ как формообразование (т.е. образование форм слова, не обусловленное синтаксически), так и деривация (образование новых слов) могут использовать одновременно и внешний способ (аффиксация) и внутренний способ (внутреннюю флексию, затрагивающую корневой состав основы слова); понятие модели, т.е. готовой образцовой формы, оказывается удобным рабочим приемом в изучении и рассмотрении арабской грамматики и лексики.

Понятие и термин «модель» традиционно используется в арабском и, шире, семитском языкознании в соответствующих терминах: *pattern/model* [SL 1997: 191-193; EALL, I, 2006: 573-577]; *Typ* [Grundriss 1982: 63 и сл.]; «*Grundformen*» как «*Basen*» [Brocklemann 1966(1908), S.285-288]; *schème* [Chartouni-Grand'Henry 2000, с.26]. В арабском языке понятие и термин *wazn-* «модель, по которой образуется слово» [Гранде 1963: 101-102; 587] < *wazn-* «вес, весовая гиря; размер

<sup>1</sup> В другой системе понятий внутренний способ формообразования и деривации описывается в терминах: прерывистая морфема, диффикс, трансфикс, операция (способ геминации корневого согласного) [Старинин 1963, с.3-5, 25-39; Плунгян 2000, с.67-70]. Понятие прерывистой морфемы могло бы унифицировать систему грамматического описания и свести все способы, как однородные, к одной парадигме. Однако на практике выделение и идентификация прерывистых аффиксов оказываются слишком громоздкими, так как требуют дополнительных пояснений и уточнений.



стиха в поэзии, соответствующий правилам просодии» [ВК II: 1530-1531]<sup>2</sup>. Термин входит в систему понятий словообразования и словоизменения арабской лингвистической традиции [Chartouni-Grand'Henry 2000: 26 и сл.; EALL, II 2007: 448].

В системе арабской грамматики этот термин связан с понятием *qiyās-* («аналогия»), которое обозначает образование новой формы слова в соответствии с уже известным и установленным образцом (моделью) [Гранде 1963: 575; Chartouni-Grand Henry 2000: 14; EALL, I, 2006: 191-194].

Модели глагольного формообразования и деривации входят в парадигму внутренних способов, включающих в ряде форм и внешний способ аффиксации.

Образцовой парадигмой служит система моделей образования финитных форм и отглагольных имен (причастия, имени действия) «правильного» глагола. Правильным глаголом считается глагол, образующийся от консонантного корня, состоящего из трех фонетически устойчивых согласных. Устойчивые корневые согласные сохраняются и не изменяются во всех финитных формах и в отглагольных именах.

Примером может послужить формообразовательная и деривационная парадигма моделей от упомянутой выше трехсогласной глагольной основы – *drus*<sup>3</sup> «учить, изучать что-либо».

Однако не все глаголы КЛЯЯ укладываются в регулярную парадигму. Термин и понятие «правильный» глагол появился в арабском языкознании в противопоставлении нескольким видам глаголов, которые, естественно, относятся к «неправильным», нарушающим регулярные модели парадигмы.

Нарушения регулярных моделей могут быть связаны как с особенностями фонетической системы КЛЯЯ, так и с историческими причинами. Сложившаяся глагольная парадигма КЛЯЯ, как показывает анализ материала, по-разному реагирует на историческое наследие КЛЯЯ и на правила сложившейся фонетической системы КЛЯЯ.

## 1. Идеальная модель и старые фонетические законы

В отличие от реконструкции корневой гласной прасемитских именных основ<sup>4</sup>, реконструкция корневого гласного глагольных основ прасемитского состояния остается открытым вопросом. С исторической точки зрения функциональное различие глагольных основ по вокализму *\*u: \*i: \*a* должно было быть связано

<sup>2</sup> Коннотация исходной лексической единицы – «вес, соответствующий норме, правильный вес».

<sup>3</sup> В арабском корневом фонде следует отметить два корневых гнезда консонантного состава *DRS*: 1) *drs* «топтать; молотить (зерно); утаптывать – стирать с лица земли»; 2) «читать текст > познавать, изучать, учить» [ВК I; 688]. Оба гнезда имеют соответствия в большинстве семитских языков [DRS, IV: 316]. Этимологическая связь значений не установлена. В арабском языке в значении (2) может быть древним культурным заимствованием из сирийского арамейского. Таким образом в арабском возникают омонимы: *dars-* (1) «молотьба» и *dars-* (2) «урок; учение».

<sup>4</sup> См. опыт реконструкции анатомической и фаунистической лексики в [SED I, II 2000, 2005]; о корневом глагольном вокализме [Коган 2009: 43-46].

с различием древней характеристики действия и состояния по направленности, по времени и способу свершения. Эти формы восходили не только к прасемитскому периоду, но и глубже – к афразийской языковой общности [Diakonoff 1988: 85-87].

В известной нам грамматической системе КЛЯЯ гласный (V) первичной основы имперфекта структуры  $-C_1C_2VC_3$  (так наз. «типовой гласный») функционально не определяется и входит в сферу словарной, т.е. лексической, информации. Однако сопоставительное исследование имперфектных глагольных основ КЛЯЯ и соответствующих основ семитских языков, сохранивших передачу глагольного вокализма, показывают: 60% от глагольных основ  $-*u-$  характеризуются значением транзитива или активного движения; 60% от глагольных основ  $-*i-$  характеризуются значением интранзитива или состояния; основы  $-*a-$  в равной степени могут быть как транзитивными, так и интранзитивными [Белова 1993: 48].

Однако значимым признаком глагольных основ  $-a-$  является их консонантный состав, в котором второй ( $C_2$ ) или третий ( $C_3$ ) корневой согласный является ларингалом или фарингалом (обобщено обозначаемый как Н):  $ʔ/ʕ/h/h$ . Ср. КЛЯЯ *ya-ftuq-u* «он протыкает/пробивает» и *ya-ftah-u* «он пробивает (канал)» > «открывает». Переход  $u/i-$  глагольной основы в  $-a-$  под влиянием указанных корневых согласных отмечается для большинства семитских языков [Brockelmann 1908: 194 и сл.].

Описанные явления могут свидетельствовать о том, что лексико-грамматическая функция глагольного вокализма «уступила» более сильному действию фонетического закона регрессивной ассимиляции. Морфологическая модель уступает фонетике. Слабость этой модели, конечно, была обусловлена и тем фактом, что и оппозиция  $-*u-$  :  $*i-$  в первичных глагольных основах не была достаточно регулярной и грамматически значимой.

Пережитком старой системы глагольного вокализма первичной основы в КЛЯЯ остается неполная регулярная связь между вокализмом имперфектной основы  $-C_1C_2VC_3$  и вокализмом перфектной основы  $C_1VC_2VC_3$ . Ср. вокализм имперфектных и перфектных основ:

$-C_1C_2uC_3$ . (imperf.):  $C_1aC_2aC_3$ - (perf.)

(2 m.sg.)	(2 m. sg.)
<i>ta-drus-u</i> <sup>5</sup>	<i>daras-ta</i>
<i>ta-sma</i> ‘-u	<i>sami</i> ‘-ta
<i>ta-mlk-u</i>	<i>malak-ta</i>
<i>ta-kbur-u</i>	<i>kabur-ta</i>

Колебания в вокализме обеих основ в системе КЛЯЯ и слабая грамматическая функциональность вокализма свидетельствуют о разрушении старой модели.

В парадигме финитных форм первичной глагольной основы КЛЯЯ жестко действует только модель пассива. Модель образуется внутренним способом – изменением как гласного основы, так и гласного в префиксе. Первичный вокализм

<sup>5</sup> *drs-* «учить, изучать»; *sm* ‘ «слышать, слушать»; *mlk-* «владеть; властвовать»; *kbr-* «быть, стать большим, взрослым».

глагольной основы (*a/i/u*) безразличен для модели пассива: в основе перфекта изменяется только основа.

Act <imperf. 2 m.sg.>	Pass
<i>ta-drus-u</i> <sup>6</sup>	<i>tu-dras-u</i> <sup>7</sup>
<i>ta-fil-u</i>	<i>tu-fil-u</i>
<i>ta-ḍbat-u</i>	<i>tu-ḍbat-u</i>
<i>ta-ḥmad-u</i>	<i>tu-ḥmad-u</i>
<perf.2 m.sg.>	
<i>daras-ta</i>	<i>duris-ta</i>
<i>fatal-ta</i>	<i>futil-ta</i>
<i>ḍabat-ta</i>	<i>ḍubiṭ-ta</i>
<i>ḥamad-ta</i>	<i>ḥumid-ta</i>

Модель пассива имперфектной основы расширяет свой показатель на префикс; префикс, таким образом, тоже «участвует» в показании пассива. Напротив, модель перфектной основы «учитывает» более сложную функцию суффиксального показателя *-ta*, обозначающего 2-ое лицо, муж. род. и ед. число. Вокализм этого показателя в данной форме подчинить пассиву невозможно: *-tu* – показатель 1-го лица ед. числа; *-ti* – показатель 2-го лица жен. р. ед. числа. Изменение вокализма суффиксов, связанных с системой личных местоимений, могло бы нарушить общую грамматическую парадигму категории лица.

Производная модель каузативной основы (IV stem) *tu-C<sub>1</sub>C<sub>2</sub>iC<sub>3</sub>*- уже не может быть нарушена старым фонетическим законом. Ср.:

*ta-ḍhak-* (I stem) : *tu-ḍḥik-* (IV stem)<sup>8</sup>

*ta-smah-* (I stem) : *tu-smiḥ-* (IV stem).

Фарингальный *ḥ* в той же синтагматической позиции уже не может изменить гласный *-i-*, являющийся составной частью модели IV stem.

Аналогично действуют деривационные модели отглагольных имен. Ср. *ta-s'al-* (imperf. 2 m. sg.) «ты спрашиваешь; просишь»: имя действия (V stem) *tasa''ul-* «попрошайничество»; модель пассивного причастия (I stem) *mas'ul-* «спрашиваемый > ответственное лицо».

Как показывает материал, в глагольной системе КЛАЯ первичные глагольные основы еще не полностью «освоены» регулярной парадигмой категориальных моделей и сохраняют следы воздействия не грамматического, а фонетического правила. Нерегулярные чередования гласных первичной глагольной основы в грамматических формах и практически нулевая лексическая функция могут свидетельствовать о том, что на определенном этапе исторического развития гласный корневой глагольной основы выполнял главным образом просодическую функ-

<sup>6</sup> *drs*(I stem) «учить, изучать»; *fil* (I stem) «скручивать»; *ḍbt* (I stem) «хватать обеими руками»; *ḥmd* «хвалить».

<sup>7</sup> *drs*(I stem) «учить, изучать»; *fil* (I stem) «скручивать»; *ḍbt* (I stem) «хватать обеими руками»; *ḥmd* «хвалить».

<sup>8</sup> *ḍhk* (I stem) «смеяться», (IV stem) «смешить»; *smḥ* «даровать; позволять», (IV stem) «снисходить к чему-л.».

цию – функцию сохранения слоговой структуры консонантной основы. В такой просодической функции фонологическая функция гласного теряла свое значение, и корневой гласный мог изменяться под влиянием более сильного корневого согласного.

В то же время фонологическая функция гласного остается стабильной в его грамматической функции. В системе новой модели КЛЯЯ гласный уже не может подаваться позиционным фонетическим влияниям. Таким образом можно уточнить частое упрощенное утверждение о том, что в арабском языке, как и вообще в семитских языках, «корень состоит только из согласных, а гласные выражают грамматические отношения» [Юшманов 1985: 21; Гранде 1963: 10-11; Diakonoff 1988: 44; SL 1997: 191-192; 203-207; EALL, IV, 2009: 93]<sup>9</sup>

## 2. Идеальная модель и «неправильные» глаголы. Модель и структура глагольного корня

В первой части статьи были рассмотрены отношения между морфологической моделью и вокализмом глагольных форм. Во второй части рассматриваются отношения между моделью, консонантным составом и структурой глагольного корня.

Панхронический подсчет консонантных корней в лексическом фонде КЛЯЯ выявляет 5590 трехгласных глагольных корней. Ни одного глагольного корня, состоящего из меньшего количества согласных, не обнаружено [EALL, IV: 94-95]<sup>10</sup>.

Трехгласный «неправильный» глагол отличается от правильного тем, что один, два (и даже все три) из его согласных являются «слабыми», т.е. подвергаются фонетическим изменениям вплоть до полного падения. К таким «слабым» согласным в КЛЯЯ относятся смычный ларингальный ' и сонорные глайды w, y. Наиболее «слабыми» являются w и y. Изменения и падения корневых согласных нарушают регулярную модель трехгласной парадигмы.

Нарушения «правильной» модели глагола «учить, изучать» в формах первичной основы «слабых глаголов»:

	финитные формы 2 л. м.р. ед.ч.			отглагольные имена
1) * <i>drs</i>	(v) <i>drus</i>	<i>ta-drus-u</i>	<i>daras-ta</i>	<i>dāris-</i>
2) * <i>qwl</i>	<i>qul</i>	<i>ta-qūl-u</i>	<i>qul-ta</i>	<i>qā'il-</i>

<sup>9</sup> Карл Брокельман, один из патриархов классической семитологии, в своем фундаментальном труде (1908 г.), не затрагивая вопроса о возможной реконструкции корневого вокализма в семитских языках, вместо понятия абстрактного консонантного корня вводит понятие «первичная основа» (Grundform/Base) [Brokelmann 1908: 286-287]. Учитывая данные более широких сравнительно-исторических исследований и реконструкций на материале семитских и афразийских языков, И. М. Дьяконов предполагает возможность реконструкции корневого вокализма первичных имен, но сомневается в возможности реконструировать корневой вокализм первичных глагольных основ и отглагольных дериваций [Diakonoff 1988: 44-45].

<sup>10</sup> Четырехгласные глаголы в КЛЯЯ функционируют, не нарушая парадигму моделей, так как «укладываются» в производную модель II stem.

3) <i>wfy</i>	<i>fī</i>	<i>ta-fī</i>	<i>wafay-ta</i>	<i>wāfī</i>
4) <i>w'y</i>	<i>'i</i>	<i>ta-'ī</i>	<i>wa'ay-ta</i>	<i>wā'ī</i>

Примечание: *qwl* «сказать» – глагол с одним «слабым» *w*; *wfy* – «выполнять» – глагол с двумя «слабыми»; *w'y* «обещать» – глагол с тремя «слабыми».

Причины отклонения от модели обусловлены разнородными факторами. Формы (2) частично обусловлены фонетически: долгий гласный в закрытом слоге сокращается, но появление долгого слога в основе, т.е. *-ū-* < *\*taqwul-u* и переход *\*qāwil-* > *qā'il-* не могут быть объяснены фонетическими законами системы КЛЯЯ. Таким же образом не могут быть объяснены падения «слабого» *w* в имперфектной основе (и соответственно – в императиве) в формах (3) и (4)<sup>11</sup>.

В парадигме производной модели II stem (интенсив, транзитив) «слабые» согласные не нарушают регулярные формы:

	финитные формы 2 л. м.р. ед.ч.			отглагольные имена
	imperat.	imperf.	perf.	part.
«обучать»	<i>darris</i>	<i>tu-darris-u</i>	<i>darras-ta</i>	<i>mu-darris-</i>
«заставить сказать»	<i>qawwil</i>	<i>tu-qawwil-u</i>	<i>qawwal-ta</i>	<i>mu-qawwil-</i>
«выполнять полностью»	<i>waffī</i>	<i>tu-waffī</i>	<i>waffay-ta</i>	<i>mu-waffī</i>

Примечание. Реальная форма II stem для *\*w'y* в арабском словарном фонде не отмечена.

Как показывают сравнительные примеры, производная парадигма легко справляется со «слабыми» согласными в более сложных моделях. То же явление характерно и для модели имен действия и первичной основы.

Отдельно следует обратить внимание на модель имени действия II stem *ta-wfiy-at-*, в которой «слабый» *w* сохраняется в поствокальной позиции после префикса *ta-* в полном соответствии с фонетическими правилами КЛЯЯ. Сопоставление с этой формой модели I stem *ta-fī* (?) < *ta-wfī* свидетельствует о том, что нарушение модели КЛЯЯ обусловлено иной причиной<sup>12</sup>.

### 3. Идеальная глагольная модель и двухсогласная корневая основа

В глагольной системе КЛЯЯ активно функционируют деноминативы. Образование глагольных основ от именных корневых основ трехсогласной или более

<sup>11</sup> Гипотезы о возможных исторических изменениях в слоговых сочетаниях глайдов *w/y* с гласными см. [Юшманов 1985: 66-74; СЯ, I, 2009: 49]; о том, что часть глагольных основ типа (2, 3, 4) может восходить к двухсогласным глагольным основам [Djakonoff 1970: 453-480; id. 1988: 42-51; Zaborski 1971: 51-98; EALL, I, 2006: 313-316]; проблемы структуры корня на арабском материале [Белова 2002: 29-36; id. 2004].

<sup>12</sup> Карл Брокельман объясняет образование императивно-имперфектной модели I stem от глаголов *\*wCC* с утратой первого *\*w* в данной позиции как результат диссимиляции с последующим слогом *-Ci-* и относит это явление к прасемитскому состоянию [Brockelman 1908: 256-257; 596-597]. Действительно, в словарном фонде КЛЯЯ можно отметить модели с сохранением первого «слабого» *w* перед слогом на *-Ca-*: *ta-wǧal-u* (*wǧl* «бояться») и на *-Cu-*: *ta-wṭur-u* (*wṭr* «быть мягким»). Список таких основ достаточно велик. Не исключено, что эти формы – новообразования.

структуры осуществляется в соответствии с моделью I stem для трехгласной или по производным моделям (II stem, V stem) от более сложных. Отыменная глагольная модель от двухгласных именных основ образуется за счет расширения корня до трехгласного (триконсонтизации корня). При помощи какого согласного? При помощи согласного, который не изменит лексическое значение именного корня, но «уложит» его в соответствующую парадигму моделей. В этой структурно-грамматической функции могут выступать только глайды w/y, фонетически близкие соответствующим гласным фонемам u/i, а также щелевой ларингальный h, который также функционирует как «расширитель» консонантного корня, но чаще – в деривации именных форм.

Ср. ниже примеры деноминативных глаголов и расширения двухгласных основ:

	финитные формы I stem 2 л. м.р. ед.ч.			отглагольные имена
	imperat.	imperf.	perf.	part.
yad- «рука»	'V-ydi > 'īdi	ta-ydī	yaday-ta	yādī- < yādiy-
'ah- «брат»	'V-'ḥu/i > 'ūḥu/i	ta-'ḥū/ī	'ahaw-ta	'āḥī- < 'āḥiy-
(1) san-at- «год»	(i)-snah	ta-snah-u	sanih-ta	sānih-
(2) sn- (V stem):	imper. ta-sanna; imperf. ta-ta-sannā perf. ta-sannay-ta			

Примечание. Arab yad- < Sem \*yad-/\*'id «hand» [SED, I, 2000: 262-263] > v. \*ydy «задеть, ранить кого-либо в руку»; «протянуть руку» > «оказать помощь»; III stem (взаимно-направленное действие) «передавать из рук в руки» [BK II: 1623-1625].

Arab 'ah- < Sem \*'ah(w) [DRS: 15] > v. 'ḥw «быть, стать братом»; III stem «поступать по-братски»; V stem «держаться по-братски» [BK I: 17-18].

Arab san-at- (n.f.) «год», ср. в Sem: Akk šatt- Ug šnt pl. šn-m, šn-t; Hbr šanā Aram šana «год». В КЛАЗЯ v. snh/snw/y (I stem, V stem) «стать многолетним, постареть»; III stem – «заключить договор сроком на год» [BK I: 1153-1155; Bustāni: 978-979; KB: 997; Riemschneider 1973: 272].

Деноминативы как новообразования укладываются в регулярную модель за счет добавления недостающего третьего корневого согласного. Правило расширения консонантного корня устойчиво и определено:

1) добавочный согласный должен быть максимально нейтральным, не изменяющим лексическое значение первичного корня;

2) добавочные согласные составляют группу определенного фонетического ряда (w/y/h). Арабская традиция относит их к «слабым», так как они более других подвержены фонетическим изменениям;

3) выбор добавочного согласного определяется фонетическим составом первичной корневой основы (здесь, как в примере с корнем \*san- «год», могут быть варианты);

4) добавочный согласный занимает конечную позицию в синтагматической последовательности корневых согласных.

Группа фонем-расширителей артикуляционно и акустически граничит с соответствующими гласными \*u/\*i/\*a фонологической системы КЛАЗЯ. В историческом плане фонемы-расширители восходят к сонантам, которые могли выпол-

нять и слогообразующие функции [Diakonoff 1988: 34, 40, 42-43]. В фонологической системе КЛАЯ группа *w/y/h* утрачивает слогообразующую функцию, однако сохраняет двойственный характер. Эти фонемы могут быть корневыми, т.е. выполнять лексическую и, соответственно, фонематическую функцию. Ср., например: *'ah-* (н.) «брат»; *wdd* (в.) «любить»; *yad-* (н.) «рука»; *whb* (в.) «дарить»; *ra 's-* (н.) «голова». С другой стороны, как уже было показано, они могут расширять консонантный корень и, кроме того, выполнять техническую функцию в слогаделении: прикрывать неприкрытый слог *'u-drus* (imperat.) в изолированной позиции; устранять стечение гласных: *kubrā* (adj.) «самая большая» – *kubrā+āt-* (pl.) > *kubrayāt-* и т. п.

Полная глагольная парадигма КЛАЯ сформировалась на основе трехсогласного корня в плане собственно языкового материала, в плане организации этого материала – преимущественно на основе способов внутренней флексии.

В отличие от внешних способов формообразования и деривации, которые могут присоединяться к основе слова любой структуры, способы внутренней флексии (чередование гласных, геминация согласного) обусловлены консонантно-вокалической структурой корневой основы глагола.

«Трехсогласная» модель, базирующаяся на внутренней флексии, не является идеальной для реального лексического состава КЛАЯ и его фонетической системы.

Историческое наследие КЛАЯ – пережитки древнего вокализма гласного корня, правила позиционных фонетических изменений первичных глагольных основ еще сопротивляются регулярной модели.

С другой стороны, трехсогласная модель успешно перестраивает первичные двухсогласные корневые основы. Расширение двухсогласных корней происходит за счет определенной группы фонем, которые не могут изменить лексическое значение первичного корня. В этой структурной (или «технической») функции эти фонемы-расширители сближаются с особым типом аффиксации.

Рассмотренные отношения между глагольной моделью КЛАЯ и его фонетической системой, относящимися к двум разным уровням языка, показывают, что давление системы/модели структурного яруса ограничивается фонетическими факторами.

### Сокращения и условные обозначения

act. – active	m. – masculine gender	pl. – plural
adj. – adjective	n. – noun	sg. – singular
f. – feminine gender	n. act. – noun of action	sub. – subjunctive
indef. – indefinite	nom. – nominative	v. – verb
indicat. – indicative	part. – participle	C – (корневой)
imper. – imperative	pass. – passive	согласный
imperf. – imperfect	perf. – perfect	V – гласный

### Названия языков

Akk – аккадский, Arab – арабский, Aram – арамейский, арамейские, Hbr – древнееврейский, Sem – семитский, семитские, Ug – угаритский

## Источники

- БК – Albert de Biberstein-Kazimirski. *Dictionnaire arabe-français*. I–II vv. Paris: Maisonneuve, 1860 (repr.).
- Bustānī – al-Bustānī, Butros. *Qaṭru l-Muḥīṭ* (A concise Arabic Dictionary). Librairie du Liban, Beirut, 1969 (ed. repr. 1869).
- DRS – D. Cohen. *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*; éd. Mouton, La Haye – Paris, 1970–1999.
- K/B – Ludwig Koehler, Walter Baumgartner. *Lexicon in Veteris Testamenti Libros*. Leiden: Brill, 1958.
- Leslau Wolf. *Comparative dictionary of Ge'ez*. Wiesbaden: Otto Harrassowitz, 1987.
- Riemschneider Kaspar K. *Lehrbuch des Akkadischen*. Veb Verlag Enzyklopädie, Leipzig, 1973.

## Литература

1. Белова А.Г. Структура семитского корня и семитская морфологическая система // *Вопросы языкознания*, 1991. № 1. С. 79–90.
2. Белова А.Г. К вопросу о реконструкции семитского корневого вокализма // *Вопросы языкознания*, 1993. № 6. С. 28–56.
3. Белова А.Г. К вопросу о структуре семитского корня (комплементы и фонетические расширители) // *История и языки Древнего Востока: Памяти И.М. Дьяконова*. СПб.: Orientalia, 2002. С. 29–36.
4. Белова А.Г. *Сравнительно-исторический аспект многозначности в корнеслове арабского языка*. М.: ИВ РАН, 2004.
5. Гранде Б.М. *Курс арабской грамматики в сравнительно-историческом освещении*. М.: Изд.-во восточной литературы, 1963.
6. Дьяконов И.М., Порхомовский В.Я. О принципах афразийской реконструкции // *Valcanica. Лингвистические исследования*. М., 1979.
7. Коган Л.Е. – Семитские языки // *СЯ*, I, 2009: 43–46.
8. Плунгян В.А. *Общая морфология, Введение в проблематику*. М.: МГУ, Эдиториал УРСС. М., 2000.
9. Старинин В.П. *Структура семитского слова*. М.: ИНА АН СССР, Изд.-во восточной литературы. 1963.
10. *СЯ*, I – *Языки мира. Семитские языки. Аккадский язык. Северозападно-семитские языки*. М.: ИЯ РАН. ACADEMIA, 2009.
11. Юшманов Н.В. *Грамматика литературного арабского языка*. 3-е изд. М.: Наука, 1985 (1928).
12. Юшманов Н.В. Структура семитского корня // *Избранные труды по общей фонетике, семитологии и арабской классической морфологии*. М.: Вост. лит. РАН, 1998 (1938). С. 126–199.



13. Belova Anna. Vestiges du vocalisme radical en sémitique // *Folia Orientalia*. vol. V. XLIX. Krakow, 2012. P. 69–76.
14. Bélova Anna. Racine à deux consonnes et morphologie à trois consonnes (l'arabe classique) // *Folia Orientalia*. vol. LII, Krakow, 2015. P. 53–61.
15. Brockelmann Carl. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Bd. I-II, Berlin: G.O. Hildesheim, 1908–1913.
16. Grand'Henry Jacques. *Grammaire arabe à l'usage des Arabes*. Traduction française et commentaires des Eléments d'arabe, morphologie et syntaxe, II de Rachid Chartouni (Beyrouth). Louvain-la-Neuve: Peeters, 2000.
17. Diakonoff I. M. Problems of root-structure in Proto-Semitic // *Archiv Orientalni*, 1970. Vol. 38, № 4. P.453–480.
18. Diakonoff I. M. *Afrasian languages*. M.: Nauka, 1988.
19. EALL – *Encyclopedia of Arabic Language and Linguistics*. Gen. Ed. Kees Versteegh. Vol. I–V, Leiden – Boston: Brill, 2006-2009.
20. *Grundriss der Arabischen Philologie*. Bd. I: Sprachwissenschaft. Hrs. von Wolfdietrich Fischer. Wiesbaden: Dr. Ludwig Reichert Verlag, 1982.
21. SED, I – Alexander Milirarev, Leonid Kogan et al., *Semitic Etymological Dictionary*. Vol. I. Anatomy of Man and Animals. Ugarit-Verlag, Münster, 2000.
22. SL – *The Semitic Languages*. Ed. by Robert Hetzron. London: Routledge, 1997.
23. Rainer M. Voigt. *Die infirmen Verbaltypen des Arabischen und das Biradikalismus-Problem*. Stuttgart, F. Steiner, 1988.
24. Zaborski Andrzej. Biconsonantal Verbal Roots in Semitic // *Zeszyty Naukowe Uniwersitet Jagellonskiego*. CCLXIX. Prace jezykoznawcze 35. Krakov. P. 51–98.
25. Zaborski Andrzej. Biconsonantal roots and triconsonantal root variation in Semitic: Solutions and prospects // *Semitic studies in honor of Wolf Leslau*. Ed. Alan S. Kaye. II. Wiesbaden, O. Harrassowitz, 1991. P. 1675–1703.
26. Zaborski Andrzej. Biradikalism // *Encyclopedia of Arabic Language and Linguistics*. Vol. I. Gen. ed. Kees Versteegh. Leinen – Boston, Brill, 2006. P. 313–316.

## Акронимы в современном арабском языке

Андрей Андреевич Блинов

Институт востоковедения РАН, Москва, Россия,  
progrock1@yandex.ru, ORCID 0000-0002-9694-8706

**Аннотация.** Статья посвящена процессу акронимизации, который возник в арабском языке сравнительно недавно. Настоящая работа дает представление о развитии этого явления с течением времени. Ее основная цель заключается в исследовании лингвистических условий функционирования акронимов в арабском языке. В данной статье рассматривается история этого словообразовательного процесса до наших дней, изложены взгляды ведущих европейских и арабских ученых на данную проблематику, перечислены основные акронимы в арабском языке с их подробным описанием, представлены модели, по которым они образуются, и выделены общие механизмы их формирования.

**Ключевые слова:** арабский язык, современный арабский язык, акронимы, акронимизация, арабский мир

**Abstract.** This article explores the acronymization process, which has occurred in Arabic quite recently. The paper provides insight into the development of this phenomenon over the course of time. Its main purpose lies in the study of linguistic conditions of operation of acronyms in Arabic. The present article reviews the history of this word-formation process until now, outlines the viewpoints of leading European and Arab scientists on this topic, lists major acronyms in Arabic with their detailed description, presents the models on which they are based and highlights the general mechanisms of their formation.

**Keywords:** Arabic, modern Arabic, acronyms, acronymization, Arab World

### 1. Введение

Новые лексемы создаются в языке с помощью различных способов словообразования. Например, в английском языке наибольшее распространение получили аффиксация, словосложение и конверсия [Kiparsky 1982: 7]. Кроме того, существует целый ряд реже встречающихся способов словообразования, таких как контаминация, усечение, формирование неологизмов, регрессивная деривация, аббревиация и акронимизация (см. более подробно [Booij 2007, Haspelmath & Sims 2010, Lieber 2010, Plag 2003]).

Невзирая на их ограниченную продуктивность, инициальные аббревиатуры и, в частности, акронимы обнаруживаются во многих языках. Поскольку процессы аббревиации и акронимизации требуют от реципиента минимальных усилий для обработки полученной информации, можно заключить, что распространение таких словообразовательных процессов более вероятно, чем в случае других моделей. В действительности, как указывается в настоящей статье, ввиду того что акронимы произносятся как отдельные слова, например *AIDS* (*Acquired Immune*

Deficiency Syndrome) «СПИД (синдром приобретенного иммунодефицита)», их использование более предпочтительно по сравнению с обычными аббревиатурами.

Говоря об акронимах в арабском языке, нельзя не отметить их недостаточную изученность в целом, поэтому данный феномен, несомненно, заслуживает самого пристального внимания. Акронимы появились в арабском языке совсем недавно, начиная с 60-х гг. прошлого века. Это свидетельствует о том, что знакомство с акронимами, произошедшее после начала вещания радиостанций и телеканалов, может повлиять на другой язык в определенных случаях, особенно в ситуациях языкового контактирования. В частности, в рамках данного исследования рассматривается влияние на акронимизацию как словообразовательный процесс в арабском языке со стороны других европейских языков – главным образом английского. Данная работа также дает представление о развитии этого явления в арабском языке с течением времени. Ее основная цель заключается в исследовании лингвистических условий функционирования акронимов в арабском языке.

## 2. Общая информация

### 2.1. Инициальные аббревиатуры и акронимы

Инициальные аббревиатуры по праву считаются одним из самых популярных видов сокращений. Сопоставляя их с другими схожими моделями, немецкий лингвист Инго Плаг отмечает, что инициальные аббревиатуры (например, *USA* «США», *FBI* «ФБР» и т.д.) и слова-гибриды (например, *smog* «смог», *motel* «мотель» и т.д.) имеют аналогичную структуру, потому что все они являются совокупностью частей различных слов. Аббревиацию объединяет с усечением и контаминацией то, что она характеризуется потерей языкового материала (а не его добавлением, как в случае аффиксации), однако, в отличие от усечения и контаминации, просодические категории не играют значительной роли в процессе аббревиации. Вместо этого первостепенное значение здесь имеет орфография» [Plag 2003: 160-161].

Определения инициальных аббревиатур, представленные в работах ведущих зарубежных исследователей, зачастую совпадают между собой. Например, как пишет американский языковед Рошель Либер в своей книге «Введение к морфологии», «они являются результатом использования начальных букв слов, составляющих название или фразу, для создания новых лексем» [Lieber 2010: 53]. Тем не менее данное определение не учитывает некоторые инициальные аббревиатуры, образованные от сложных слов, например в немецком языке. Поэтому более подходящим в данном случае может быть следующее определение: «инициальная аббревиатура – это новая лексема, созданная на основе начальных букв последовательности нескольких слов в ходе процесса словообразования».

Если говорить о происхождении слова «акроним», то оно образовалось от древнегреческих слов *akros* «высший, крайний» (здесь в знач. «начальный») и *онута* «имя, название». Хотя считается, что само слово «акроним» окончательно оформилось только в 1943 г. [Thomas, Ethridge & Ruffner 1965: 7], британский лингвист Брайан Фостер отмечает, что акронимы «широко использовались еще во время Первой мировой войны» [Foster 1972: 196].

Трактовка понятия «акроним», приведенная в данной работе, отражает современный подход к определению этого термина исходя из слитности/раздельнос-

ти произношения составляющих его компонентов. Тем не менее изначально акронимы приравнивались к инициальным аббревиатурам и употреблялись в качестве одного из их синонимов. Так, американский исследователь Дуайт Болингер использовал термин «акроним» для обозначения «произносимых вместе начальных букв сложных названий» социальных и политических организаций [Bolinger 1968: 116]. Такие названия, по его мнению, могли произноситься по буквам, как в слове *FBI*, или слитно, как в слове *NATO*. Тем не менее автор не объяснял причины, которые могли повлиять на каждый тип произношения. В «Оксфордском словаре современного английского языка для продвинутых учащихся» 1974 г. его составитель Альберт Хорнби определял акроним как «слово, образованное от начальных букв имени». Американский ученый Эдриан Акмаджян рассматривал акронимизацию как «лишь один из процессов аббревиации, или сокращения» [Акмаджян et al 1984: 69]. Британский лингвист Рэндольф Кверк разделил акронимы на два типа: (1) акронимы, которые произносятся как отдельные буквы, например *VIP* «ВИП» вместо *Very Important Person* «очень важное лицо»; (2) акронимы, которые произносятся как отдельное слово, например *NATO* «НАТО» вместо *North Atlantic Treaty Organization* «Организация Североатлантического договора» [Quirk et al 1985: 1580]. Английский переводчик Питер Ньюмарк считал акронимы начальными буквами фразы, используемыми «для указания на соответствующий объект, учреждение или процедуру». Он употреблял термин «интернационализмы» для описания акронимов, обозначающих названия международных учреждений, например *FAO* «ФАО» вместо *Food and Agriculture Organization* «Продовольственная и сельскохозяйственная организация» [Newmark 1988: 200].

Другую точку зрения на акронимы выражали, например, Эдвард Финеган и Нико Беснир, причисляя акронимы к особым формам сокращений, увеличивающих словарный запас языка. Они ограничивали область употребления акронимов только теми единицами, которые произносились как отдельное слово [Finegan & Besnier 1989: 110]. В дальнейшем данные взгляды стали доминировать в ученой среде, хотя многие словари до сих пор отождествляют акронимы с инициальными аббревиатурами, интерпретируя это понятие в более широком плане. Таким образом, долгое время акронимы рассматривались различными учеными в рамках более крупной категории инициальных аббревиатур без учета принципа слитности/раздельности их произношения и только недавно выделились в отдельный класс.

Инициальные аббревиатуры в настоящее время подразделяются на два вида: акронимы (т.е. звуковые аббревиатуры) и алфавитизмы (т.е. буквенные аббревиатуры) (Rua 2002: 31). Акронимы – это комбинация начальных букв последовательности слов, произносимая как отдельная лексема. Примеры акронимов можно найти в различных языках, таких как английский, например *NATO* (*North Atlantic Treaty Organization*), или французский, например *OTAN* (*Organisation du Traité de l'Atlantique Nord*) «НАТО (Организация Североатлантического договора)». Напротив, алфавитизмы – это комбинация начальных букв слов, произносимая как отдельные звуки. Примеры таких аббревиатур также встречаются во многих языках – например, английском (*SMS* (*Short Message Service*) «СМС (служба коротких сообщений)») или французском (*SVP* (*S'il vous plaît*) «пожалуйста»). Таким образом, в отличие от акронимов алфавитизмы произносятся как серия букв, на-

пример *CIA* (Central Intelligence Agency) /ci: ai ei/ «ЦРУ (Центральное разведывательное управление)», в то время как акронимы, например *OPEC* /oupek/ (Organization of Petroleum Exporting Countries) «ОПЕК (Организация стран-экспортеров нефти)», произносятся как отдельное слово.

С точки зрения орфографии акронимы могут записываться как прописными буквами, например *NASA* (National Aeronautics and Space Administration) «НАСА (Национальное управление по аэронавтике и исследованию космического пространства)», так и строчными буквами, например *laser* (light amplification by stimulated emission of radiation) «лазер (усиление света посредством вынужденного излучения)». Наоборот, алфавитизмы всегда записываются только прописными буквами, например *GMO* (Genetically Modified Organism) «ГМО (генетически модифицированный организм)» или *NBA* (National Basketball Association) «НБА (Национальная баскетбольная ассоциация)». Даже несмотря на то что алфавитизмы представлены в современном арабском языке в достаточной степени, например *mīm tā' fā'* (munazzamatu t-tahrīri l-filasṭīniyyah) «ООП (Организация освобождения Палестины)» либо *ḡīm 'ayn mīm* (al-ḡumhūriyyatu l-'arabiyyatu l-muttaḥidah) «Объединенная Арабская Республика», в рамках настоящего исследования подробно освещается только процесс акронимизации ввиду его низкой исследованности в соответствующей литературе.

Анализируя роль акронимизации в современном мире, Р. Хартманн и Ф. Сторк указывают на то, что акронимы создаются для ускорения и облегчения коммуникации из-за их краткости и легкой запоминаемости как в письменной, так и в устной речи [Hartmann & Stork 1976: 1]. Таким образом, при формировании акронимов чрезвычайно важны именно краткость и точность. В этом отношении, как говорится в работе Мод Эрманн, которая совместно с другими учеными исследовала несколько пар акронимов в 22 различных языках, «подобно другим языковым единицам, акронимы имеют высокое референтное значение в том смысле, что большую часть времени они служат референтным ориентиром текстового содержания» [Ehrmann et al 2013: 237]. Иными словами, акрономизация – это словообразовательный процесс, направленный на создание новых слов для обозначения понятий, которые часто используются в СМИ либо с трудом поддаются описанию. В таком случае произнесение полной формы акронима требует дополнительных временных затрат, особенно при употреблении таких распространенных акронимов, как *NASA*, *UNESCO*, *NATO*, *AIDS* и т.д. Египетский ученый Махмуд Хиджази объясняет причины популярности акронимов следующим образом: а) акронимы экономят время ввиду их предельной лаконичности; б) акронимы занимают меньше места при печати в тех случаях, когда для выражения полноценных конструкций и/или сложных слов употребляются отдельные буквы; в) использование акронимов упрощает понимание исходной конструкции или сложного слова [Ḥiǧāzī 2001: 95]. Поэтому выбор акронимов в качестве предпочтительного способа аббревиации является отличительным признаком многих языков.

Обзор существующей научной литературы показывает, что на сегодняшний день отсутствуют крупные исследования, посвященные процессу акронимизации в арабском языке, – как в нашей стране, так и за рубежом. Для устранения этого пробела была подготовлена данная статья, в которой рассматривается история

данного словообразовательного процесса до наших дней, изложены взгляды ведущих европейских и арабских ученых на данную проблематику, перечислены основные акронимы в арабском языке с их подробным описанием, представлены модели, по которым они образуются, и выделены общие механизмы их формирования.

## 2.2. История инициальных аббревиатур и акронимов

Потребность в экономии языковых средств возникла несколько тысяч лет назад. Американский ученый Гарланд Кэннон утверждает, что инициальные аббревиатуры существовали на протяжении многих веков, поэтому несколько примеров этого явления можно найти в иврите, например *MILH* (*Mi Iolh Lnv Hshmilh*) «Кто вознесется на небо за нас?», или латыни, например *SPQR* (*Senatus Populusque Romanus*) «Сенат и народ Рима» [Cannon 1989: 99]. Другим известным в древности акронимом, упоминаемым в христианской Библии, является слово *INRI* (*Iesus Nazareus Rex Iudaeorum*) «Иисус Назарянин, Царь Иудейский». Кроме того, появившееся в XVII в. английское слово *cabal* (Clifford, Arlington, Buckingham, Ashley, Lauderdale) образовалось от начальных букв фамилий участников так называемого «кабального» совета при Карле II.

Впоследствии, начиная с XIX в., возникли новые аббревиатуры, которые начали использоваться достаточно часто. Так, в Великобритании прочно вошли в употребление такие аббревиатуры на основе латыни, как *A.B.* (*Artium Baccalaureus*) «бакалавр искусств» и *D.C.L.* (*Doctor Civilis Legis*) «доктор гражданского права», а в США имели хождение следующие сокращения: *C.O.D.* (*Cash on Delivery*) «наложенный платеж», *N.G.* (*No Go*) «негодный» и *O.K.* (*Oil Korrect*) «все в порядке». Известный публицист и сторонник развития американского варианта английского языка Хенри Луис Менкен писал, описывая этот процесс, что в нем проявляется «характерная черта американцев превращать сложные понятия в простейшие аббревиатуры» [Mencken 1919: 22]. Во время Первой мировой войны было зафиксировано систематическое увеличение количества аббревиатур, поскольку жители США постоянно использовали сокращенные формы в законотворчестве, бизнесе и повседневной деятельности.

В 1945 г. компания «Мерриам-Вебстер» добавила 72 инициальные аббревиатуры в приложение ко второму изданию своего знаменитого словаря, вышедшего в 1934 г., сделав их частью американского лексикона [Russell 1946: 220]. В 1951 г. британский исследователь Роберт Хенри Роббинс собрал 126 устных и письменных аббревиатур только в сфере авиации, тем самым подчеркивая их стремительный рост, который продолжился в последующие годы (см. более подробно [Robbins 1951: 67-70]).

В 1960 г. в США был опубликован первый «Словарь акронимов» (Р. Томас и др., 1960). Через пять лет увидело свет второе издание данной книги под названием «Словарь акронимов и инициальных аббревиатур» (Р. Томас и др., 1965). Вскоре были выпущены другие словари по данной тематике, например «Акронимы и аббревиатуры в библиотечной и информационной работе» (А. Монтгомери, 1975), «Словарь акронимов и аббревиатур Пью в области менеджмента, технологий и информатики» (Э. Пью, 1981) и «Словарь акронимов, инициальных аббре-

виатур и сокращений», 18-е издание (Д. Моссман, 1994). В частности, авторы «Словаря акронимов и инициальных аббревиатур» (Р. Томас и др., 1965) связывали возникновение новых акронимов с беспрецедентным прогрессом электроники и технологий. Использование акронимов, по их мнению, было призвано «экономить драгоценные дюймы газетной бумаги и драгоценные секунды эфирного времени, скрывать военные тайны и привлекать внимание к товарам, идеям и программам, которые публика должна поддерживать, хвалить или приобретать» [Thomas et al 1965: 7].

Описывая процесс становления инициальных аббревиатур, авторы по-разному оценивают их роль в языковой иерархии. Так, испанский лингвист Паула Лопес Руа пишет, что анализ инициальных аббревиатур демонстрирует их многогранность и сложность, а структура таких аббревиатур не поддается описанию в классическом формате [Rua 2002: 33]. С другой стороны, британский исследователь Лори Бауэр приуменьшает значение инициальных аббревиатур, считая их не только непредсказуемыми, но и слишком зависимыми от орфографии [Bauer 1983: 237-38]. В частности, он выделяет две причины непредсказуемости таких аббревиатур. Во-первых, они могут свободно образовываться любым образом от исходной фразы. Во-вторых, не все инициальные аббревиатуры произносятся как одно слово. К тому же, как пишет английский ученый Джордж Юл, некоторые новые акронимы входят в повседневный обиход настолько быстро, что многие носители языка даже не задумываются о значении составляющих их частей. Новообразованные слова типа *ATM* (Automatic Teller Machine) «банкомат» и *PIN* (Personal Identification Number) «ПИН-код» регулярно используются с повторением одного из компонентов их полного названия, как в предложении *I sometimes forget my PIN number when I go to the ATM machine* «Иногда я забываю свой ПИН-код, когда подхожу к банкомату» [Yule 2006: 58].

Акронимы употребляются не только в английском и других европейских языках, но и в арабском. Они окончательно оформились в арабском только в 1960-х гг., однако в целом аббревиация в качестве высокопродуктивного способа словообразования зародилась в арабском языке в глубокой древности. Буквенные сокращения впервые появились здесь еще в VII в. в первом письменном памятнике арабского языка – Коране. В начале 29 из 114 сур Священного Писания, представляющего собой ниспосланное откровение Аллаха, были размещены различные аббревиатуры, в которых, как считают арабы, зашифрован скрытый смысл, известный только Пророку и его последователям<sup>1</sup>. В арабской традиции они получили название *muqatta'at* 'at «разорванные» или *fawātiḥ* «открывающие». Несмотря на слитное написание таких аббревиатур, каждая из составляющих их букв произносилась отдельно (см. более подробно [Блинов 2020]).

Несомненно, что использование аббревиатур в тексте Священного Корана вдохновило многих арабских ученых, стремящихся следовать «Слову Божию», поэтому употребление различного рода усечений, контрактур, сигл и аббревиатурных символов стало неотъемлемой частью творчества средневековых мусуль-

---

<sup>1</sup> Среди арабских ученых бытует мнение, что они обозначают имена и атрибуты Аллаха.

манских авторов. Правда, по большей части составляющие их буквы произносились отдельно, за исключением некоторых примеров слитного чтения, как в случае хадисов, где их переписчики использовали следующие упрощенные аббревиатуры для обозначения часто повторяющихся глаголов: *nā*, *ṣanā*, *daṣanā* (*haddaṣanā* «он рассказал нам»); *'anā* (*'anba'anā* или *'ahbaranā* «он уведомил нас» или «он сообщил нам»); *'anā*, *'aranā*, *abanā* (*'ahbaranā* «он сообщил нам»). Кроме того, с данными аббревиатурами близко соотносились слова-гибриды типа *basmalah* (*bi-smi-llah*) «во имя Аллаха!» или *ḥamdalah* (*al-ḥamdu li-llah*) «во славу Аллаха!», которые, подобно акронимам, также произносились как отдельное слово, хотя и отличались от них способом словообразования.

Одним из первых арабских исследователей, затронувших тему акронимов в арабском языке, был выдающийся иракский лингвист Ибрагим ас-Самаррай. В своей статье 1987 г. «Сокращения и символы в арабском наследии» он впервые обратил внимание на новые акронимы, заимствованные из английского языка, такие как *al-yūniskū* «ЮНЕСКО», *al-yūnisīf* «ЮНИСЕФ», *al-'aliksū* «АЛЕКСО» и *al-fāu* «ФАО», сделав акцент на их успешном подчинении строю арабского языка благодаря приобретению определенного артикля и удачному завершению фонетической адаптации. В 1987 г. также вышла работа иорданского ученого Абда аль-Маджида Нуссейра «Акронимы», в которой он провел критическую оценку англоязычных акронимов, согласившись с зарубежными учеными, что *manḥūtātu l-budū'* «акронимы» и *muḥtaṣarāt* «аббревиатуры» образуются от начальных букв слов в фразе, при этом автор уточнил, что первые произносятся как отдельные слова, а вторые – как отдельные буквы. К сожалению, Нуссейр не представил арабские примеры для подтверждения своей теории, сосредоточившись на анализе англоязычных словоупотреблений.

Спустя 10 лет другой арабский ученый Абу Салим классифицировал акронимы в арабском языке, разделив их на две группы: первая включает исконные арабские акронимы, например *ramā* (*rābiṭatu l-mar'ati l-'urduniyyah*) «Иорданская ассоциация женщин», вторая состоит из акронимов, образованных на основе английских слов, например *sābik* – арабоязычный эквивалент английского акронима *SABIC* (**Saudi Basic Industries Corporation**) «Саудовская корпорация тяжелой промышленности». Тем не менее исследование Абу Салима главным образом сводилось к характеристике акронимов, заимствованных из английского языка, в то время как процент исконных арабских акронимов был минимальным. То же можно сказать о вышедшей в 1999 г. кандидатской диссертации Ахмеда Канаكري «Структура акронимов в английском и арабском языке», поскольку лишь 14 из 60 приведенных им арабских примеров являлись традиционными акронимами, например *ḥaṣd* (*ḥizbu š-ša'bi d-dīmuqrāṭiyyi l-'urduniyyi*) «Народная демократическая партия Иордании». Оставшиеся 46 примеров зачастую содержали слова-гибриды, образованные в результате лексической контаминации, типа *harwala* (*haraba wa wallā*) «убежать (бежать и скрыться)» или *ḡalmada* (*ḡalada wa ḡamada*) «застыть (замерзнуть и затвердеть)», которые существенно отличались от акронимов.

Из сравнительно недавних трудов выделяется несколько работ. Махмуд Хиджази в своей статье «Современные сокращения в СМИ: между арабским переводом и лексическим заимствованием» (2001) указывает, что процесс акронимиза-



ции исторически сложился в арабском языке в течение многих веков. В частности, он уверяет, что примеры акронимов можно найти в старинных арабских рукописях, например *'ilh* (*'ilā 'āhirihi*) «и т.д.». Также он добавляет, что с течением времени все большее количество акронимов начало появляться на страницах словарей. Иракские лингвисты Ахмед Хумейд и Аридж Алтай, исследование которых увидело свет в 2013 г. и было посвящено ошибкам при переводе акронимов в английском и арабском языке, сходятся во мнении с Хиджази и приводят такие примеры, как *kg* (*kilo gram*). Тем не менее нельзя не заметить, что многие лексические единицы, включенные вышеуказанными арабскими авторами в свои работы, по своей сути не являются акронимами, поскольку они не произносятся как отдельное слово, как в вышеупомянутом сокращении *kg* (*kilo gram*). В действительности их аббревиация исчерпывалась лишь отражением на письме.

Другие исследователи, занимавшиеся тематикой акронимизации в арабском языке, рассматривали механизмы формирования акронимов и ситуации, в которых они выявляются, а также семантические ограничения, накладываемые на данный процесс. В частности, Джихад Хамдан и Шахада Фарах в 2003 г. представили сравнительный анализ акронимов в арабском и английском языке, уделив основное внимание а) способам их образования в обоих языках и б) условиям их функционирования. Результаты анализа показали, что эти два языка имеют общие схемы словообразования (например, формируются на основе начальных букв составляющих их слов) и демонстрируют аналогичную тенденцию к опущению дополнительных служебных слов при употреблении акронимов, за исключением случаев, когда их применение необходимо для сохранения структуры семантически и/или фонетически значимых слов. В то же время авторы также обнаружили в каждом языке особенности, уникальные только для данного языка. Например, арабы часто добавляют гласные, отсутствующие в исходной фразе, к акрониму, чтобы найти ему адекватный эквивалент в действующем лексиконе. Рассуждая о контексте использования таких сокращений, Д. Хамдан и Ш. Фарах отмечают, что по сравнению с английским языком акронимы в арабском встречаются реже, не документируются в словарях и имеют более ограниченную область применения. По их наблюдениям, подобное положение вызвано целым рядом факторов – лингвистических, научных и культурных.

Последняя на сегодняшний день работа по актуальной проблематике «Лингвистические свойства акронимов в арабском языке» была выпущена ученым из ОАЭ Абдаррахманом ат-Тахейной в 2017 г. В ней автор обобщил все имеющиеся сведения об акронимах в арабском языке, всесторонне изучив их фонологические, семантические и прагматические черты.

### 3. Акронимы в современном арабском языке

Акронимы в современном арабском языке появились сравнительно недавно, приблизительно 50 лет назад, по сравнению с другими языками, например английским (см. раздел 2.2 «История инициальных аббревиатур и акронимов»). Большинство из них представляют собой названия различных учреждений, организаций, политических объединений, движений и информационных агентств. Становление процесса акронимизации в арабском языке происходило в двух направлени-

ях, которые развивались параллельно до тех пор, пока одно из них окончательно не вытеснило другое из употребления.

Изначально акронимы образовались здесь не от исконных арабских слов, а от иноязычных конструкций. Несомненно, что самым популярным источником таких единиц стал английский язык. Так, сокращение *sānā* (al-wikālatu l-'arabiyyatu s-sūriyyah li-l-'anbā') «САНА (Сирийское арабское информационное агентство)» образовалось от словосочетания *Syrian Arab News Agency*, сокращение *sūnā* (wikālatu l-'anbā'ī s-sūdāniyyah) «СУНА (Суданское информационное агентство)» – от словосочетания *Sudan News Agency*, сокращение *kūnā* (wikālatu l-'anbā'ī l-kuwaytiyyah) «КУНА (Кувейтское информационное агентство)» – от словосочетания *Kuwait News Agency*. Основываясь на информации, представленной на официальных сайтах вышеуказанных агентств, акроним *sānā* используется в арабском языке с 1965 г., *sūnā* – с 1970 г., *kūnā* – с 1976 г. Как мы можем убедиться, эти акронимы были составлены из начальных букв англоязычных названий агентств с внесением необходимых изменений, связанных с подчинением фонетическим законам арабского языка.

Другим интересным вопросом является фонетическая адаптация английских акронимов в арабском языке, таких как *laser* (Light Amplification by Stimulated Emission of Radiation) «лазер (усиление света посредством вынужденного излучения)», *radar* (Radio Detecting and Ranging) «радар (радиообнаружение и измерение дальности)», *AIDS* (Acquired Immune Deficiency Syndrome) «СПИД (синдром приобретенного иммунодефицита)» и *BASIC* (Beginners All-purpose Symbolic Instruction Code) «Бейсик (универсальный код символических инструкций для начинающих)». В результате этого в заимствованных акронимах произошли следующие фонетические преобразования: дифтонг /ei/ в словах *laser* и *radar* сменился гласными /i:/ (līzar) и /a:/ (rādār) соответственно (также в последнем примере удлинилась вторая гласная /a/), в начале слова *AIDS* появилась хамза /ʔ/ ('īdz), а в слове *BASIC* удвоился звук /s/ и удлинилась гласная /i/ (bīssīk). Также успешно завершилось их морфологическое освоение: например, акроним *rādār* приобрел категорию множественного числа *rādārāt* «радары» и определенное состояние *ar-rādār*. Использование акронимов, заимствованных из английского языка, в арабском может быть продиктовано несколькими причинами: а) английский считается языком науки; б) количество акронимов в арабском языке сравнительно невелико по сравнению с английским; в) значительно расширилось использование письменных сокращенных форм в Интернете.

При этом ни для кого не секрет, что не все регионы арабского мира в равной степени испытывают на себе влияние английского языка. Так, в регионе Магриба периодически регистрируются собственные акронимы, созданные на основе французского языка. Например, в Марокко в 2002 г. было учреждено новое государственное ведомство *hākā* (al-hay'atu l-'ulyā li-l-'ittiṣāli s-sam'iyyi l-baṣariyy) «Высшее управление аудиовизуальной связи». Акроним в данном случае образовался не от арабского названия, а от его французского аналога *Haute Autorité de la Communication Audiovisuelle*. Кроме того, подобный орган существовал во Франции еще ранее, в 1982-1986 гг., и также обозначался аббревиатурой *HACA*, поэтому марок-

канское ведомство, по всей видимости, было создано по его образцу. Таким образом, акроним *hākā*, заимствованный из французского языка, фонетически вписался в строй арабского языка более органично, чем его возможный эквивалент *h‘sb*, произнести который арабам не представлялось возможным.

Любопытно, что в соседнем с Марокко Тунисе можно встретить похожий акроним *hāykā* (*al-hay‘atu l-‘ulyā l-mustaqillah li-l-‘ittisālī s-sam‘iyyi l-baṣariyyi*) «Высшее независимое управление аудиовизуальной связи», который также образовался от французского слова *Haute Autorité Indépendante de la Communication Audiovisuelle* (НАИКА), что подтверждает приверженность франкоязычным заимствованиям в рамках процесса акронимизации в данном регионе.

Несмотря на вышеприведенные иноязычные акронимы, вторая группа акронимов – исконно арабские – является более многочисленной. Одним из первых акронимов такого рода стало сокращение *‘amal* (*‘afwāḡu l-muqāwamati l-lubnāniyyah*) «Группы ливанского сопротивления», использовавшееся для обозначения ливанской политической партии и имеющее следующее дословное значение: «надежда». Это движение было основано в Ливане в 1974 г., его военизированное крыло появилось в 1975 г. и активно участвовало в гражданской войне в стране. Именно тогда акроним *‘amal* привлек внимание СМИ, однако его первое подтвержденное использование в местных печатных изданиях случилось только в 1978 г. Впоследствии интервал между возникновением нового акронима и его проникновением на страницы газет и журналов будет только уменьшаться ввиду повсеместного развития средств массовой информации и усиления роли акронимизации в качестве заметного явления в арабском языке.

Тем не менее первым подлинным акронимом в современном арабском языке принято считать скорее не *‘amal*, а *fath* (*ḥarakatu t-tahrīri l-filasṭīniyyah*) «ФАТХ (Движение за освобождение Палестины)». Данное движение было создано в 1959 г. группой палестинцев, одним из которых был Ясер Арафат, однако сам акроним вошел в полноценное употребление только в 1968 г. Это уникальный пример так называемой «реверсивной акронимизации», где обратный порядок букв в акрониме *fath* досл. «открытие, освобождение, победа» был обусловлен негативным подтекстом прямого порядка букв, образующих слово *ḥatf* «смерть», что в очередной раз указывает на то, что акронимы в арабском языке семантически мотивированы.

Другое известное арабское движение *ḥamās* «ХАМАС» (досл. «воодушевление, энтузиазм») также представляет собой акроним, образовавшийся от словосочетания *ḥarakatu l-muqāwamati l-‘islāmiyyah* «Исламское движение сопротивления». Движение было основано 14 декабря 1987 г., а первое упоминание в прессе появилось 25 ноября 1988 г. Широкое распространение этого акронима в СМИ началось после 1993 г., когда его члены не признали палестино-израильский мирный процесс, результатом которого стало заключение Соглашений в Осло.

Большое количество акронимов появилось в арабском языке сравнительно недавно, в 2000-х гг., в особенности после событий Арабской весны, начавшейся в 2011 г. Самым известным из них, вне всякого сомнения, является слово *dā‘iṣ*. Акроним *dā‘iṣ* (*ad-dawlātu l-‘islāmiyyah fī l-‘irāq wa š-šām*) «ИГИЛ (Исламское государство Ирака и Леванта)» обозначает террористическую группу, основанную в

Ираке в 2011 г., и не имеет дословного перевода в отличие от других современных акронимов в арабском языке. Он впервые попал в поле зрения широкой общественности в 2014 г., когда эта группа оккупировала большие территории в Ираке и Сирии. Интересно, что акроним *dā'īs* начал использоваться на Ближнем Востоке позже западных СМИ, которые образовали от словосочетания *the Islamic State in Iraq and Syria* акроним ISIS (первой это сделало BBC), а арабы взяли за основу этот принцип, образовав *dā'īs*<sup>2</sup>. Оба термина появились во СМИ практически одновременно, быстрота формирования арабского термина не в последнюю очередь связана с популярностью его аналога на английском языке.

Несомненно, что название *dā'īs* на первых порах использовалось арабоязычными противниками ИГИЛ. Хотя *dā'īs* не имеет другого значения в арабском языке, будучи прямой калькой англоязычного акронима ISIS, в сознании арабов оно во многом перекликается с такими существительными, как *dā'is* «крушащий (или топчущий) что-либо под ногами» и *dāhis* «Дахис». Последнее слово напоминает об известном арабам легендарном эпизоде из истории междоусобных войн между племенами в период джахилии – спор из-за жеребца Дахиса и кобылы Габры, вызвавший сорокалетнюю войну. Подобный подбор семантических коннотаций исходит из установки, лежащей в мировоззрении арабов, что за любым акронимом должно крыться определенное буквальное значение.

Таким образом, слово *dā'īs* приобрело отчетливый негативный подтекст не случайно. Последователи ИГИЛ считали это название оскорбительным и угрожали вырвать язык всякому, кто употреблял его на публике, и наказывать клеветников палками. Кроме того, в отличие от *fath* досл. «открытие, освобождение, победа» или *hamās* досл. «воодушевление, энтузиазм» данный акроним казался им бессмысленным и оттого еще более уничижительным, поэтому они желали, чтобы их называли исключительно полным именем *ad-dawlatu l-'islāmiyyah* «Исламское государство». Многие интернет-пользователи, выступающие против ИГИЛ, в сетях «Твиттер» и «Фейсбук» специально рифмовали со словом *dā'īs* различные обидные слова и выражения типа *ḡāhiš* «осел» или *dā'ir* «проститутка».

Более того, согласно данным сразу из нескольких источников, в дальнейшем это слово стало нарицательным и приобрело множественное число *dawā'īs* в значении «религиозные фанатики, навязывающие свои взгляды другим».

#### 4. Семантическая обусловленность акронимов в арабском языке

При ближайшем изучении арабских акронимов становится понятно, что в большинстве случаев они семантически мотивированы. В частности, обычно они являются омонимами слов, уже существующих в арабском языке. В некоторых случаях к акронимам добавляются дополнительные буквы, чтобы сделать их семантически значимыми и похожими на устоявшиеся слова в арабском языке. Например, в случае со словосочетанием *al-munaẓẓamatu š-šuyū'iyyatu l-miṣriyyah* «Египетская коммунистическая организация» его официальным акронимом являлся алфавитизм *mīm šīn mīm*, однако в народе закрепилось неодобрительное прозви-

<sup>2</sup> По одной из версий, происхождение этого термина приписывают сторонникам президента Сирии Башара Асада.

ше *mišmiš* досл. «абрикос», образованное от начальных букв данного словосочетания с добавлением согласного звука šīp.

В результате анализа ряда арабских акронимов можно сделать вывод о наличии двух уровней семантической обусловленности. Первый уровень отражает внутреннюю значимость акронима, в то время как второй уровень подразумевает его обязательную положительную коннотацию. Например, акроним *tawād* досл. «гармония» обозначает одну из современных политических партий Иордании *at-taḡammu‘u l-waṭāniyyu l-‘urduniyyu d-dīmuqrāṭiyyu* «Национальное демократическое собрание Иордании», основанную в 2013 г. При этом значение этого акронима четко указывает на тот смысл, который закладывали в него основатели партии при ее создании, т.е. гармоничные отношения между членами партии и иорданцами. Среди других положительно окрашенных акронимов стоит выделить такие слова, как *maḡd*, *fātin* и *bāsim*.

Например, акроним *maḡd* досл. «слава, величие» обозначает объединение *munaẓẓamatu l-ḡabhati d-dīmuqrāṭiyyu* «Организация «Демократический фронт», основанное в Иордании в 1978 г., а акроним *fātin* досл. «привлекательный, очаровательный» обозначает организацию *mu‘assasatu filasṭīn li-t-taslīf wa t-tanmiyah* «Палестинский институт кредитования и развития», основанную в Палестине в 1999 г. для помощи бедным людям. Как мы можем убедиться, эти два акронима также призваны вызвать благоприятный отклик у реципиента. То же относится и к арабскому акрониму *bāsim* досл. «улыбающийся». Оно представляет собой сокращенную форму словосочетания *al-banku s-sa‘ūdiyyu li-l-muṣṭalahāt* «Саудовский банк терминов».

Целый ряд акронимов имеет более нейтральное значение. Например, слово *ḥašd* досл. «собрание» обозначает *ḥizbu š-ša‘bi d-dīmuqrāṭiyyi l-‘urduniyyu* «Народную демократическую партию Иордании», основанную в Иордании в 1993 г. По замыслу создателей партии, ее название символизирует группу лиц, собравшихся во имя всеобщего блага. Слово *tāḡ* досл. «корона» обозначает политическую партию *taḡammu‘u ‘amali l-ḡazā‘ir* «Собрание «Надежда Алжира», учрежденную в 2012 г. в Алжире. Слово *ḥaqq* досл. «право» обозначает суданскую партию *ḥarakatu l-qiwā l-ḡadīdati d-dīmuqrāṭiyyah* «Движение новых демократических сил», которая была учреждена в 2014 г.

Кроме того, в данном контексте нельзя не упомянуть достаточно востребованный акроним *wa‘d* «обещание», обозначающий сразу несколько политических образований в различных странах Арабского Востока: *al-ḥizbu l-waṭāniyyu li-l-‘adālah wa t-tanmiyah* «Национальная партия справедливости и развития» (Сирия), *al-ḥizbu l-waḥdawīyyu l-‘arabiyyu d-dīmuqrāṭiyyu* «Арабская объединительная демократическая партия» (Иордания) и *ḡam‘iyyatu l-‘amali l-waṭāniyyi d-dīmuqrāṭiyyu* «Национальное демократическое рабочее общество» (Бахрейн).

Отдельную группу акронимов составляют слова, лишённые отдельного буквального значения, такие как *wāḥ* (*wikālatu ḥabar li-l-‘anbā’*) «Информационное агентство «Хабар», *wās* (*wikālatu l-‘anbā’i s-sa‘ūdiyyah*) «Саудовское информационное агентство», *wā’* (*wikālatu l-‘anbā’i l-‘irāqīyyah*) «Иракское информационное агентство». Впрочем, как указывают арабские исследователи Джихад Хамдан и Шахада Фарах, хотя такие акронимы записываются как независимые слова, скорее

всего, они произносятся по буквам, т.е. *waw 'alif hā'*, *waw 'alif sīn* и *waw 'alif 'ayn* соответственно [Hamdan & Fareh 2003: 189].

Примечательно, что в других языках, например в английском, акронимы также могут передавать дополнительные аспекты значения, выражаемые референтом. Сюда относятся такие сокращения, как *JUST* досл. «справедливый» (Jordan University of Science and Technology «Иорданский университет науки и технологии»), *SMASH* досл. «удар» (Students Mobilized Against Saddam Hussein «Студенты, мобилизованные против Саддама Хуссейна»), *FIST* досл. «кулак» (The Federation of Inter-State Truckers «Федерация перевозчиков грузов между штатами»). Подобное сочетание букв зачастую объясняется стремлением некоторых организаций придумать красочный акроним, который бы привлек всеобщее внимание к их философии и ассоциировался в массовом сознании с их идеями или целями. Иногда это происходит в рекламных или информационных целях, как в следующих примерах: *CARE* досл. «забота» (Cooperative for Assistance and Relief Everywhere «Кооперативное общество по повсеместному оказанию помощи»), *SOLE* досл. «один» (Self-Organising Learning Environment «Самоорганизующаяся среда обучения»).

Таким образом, при составлении акронимов арабы стараются сделать их семантически значимыми, запоминающимися и заметными, а также избегать по возможности даже намеков на негативные коннотации. Для этого они готовы даже пожертвовать прямым порядком составляющих компонентов, как мы уже упоминали выше в разделе 3 «Акронимы в современном арабском языке», где отрицательный концепт *ḥatf* «смерть» был заменен положительным концептом *fath* досл. «открытие, освобождение, победа» благодаря обратной аббревиации словосочетания *ḥarakatu t-tahrīri l-filastīniyyah*.

Более того, поскольку практически за каждым акронимом скрыто дополнительное значение, это позволяет носителям языка ловко манипулировать ими в собственных целях ввиду гибкости их внутренней структуры и эквивалентности устоявшимся словам. Например, противники Движения за национальное освобождение Палестины нередко специально переименовывают акроним *fath*, превращая его в оскорбительное слово для сторонников ФАТХ<sup>3</sup>. Так, в 2015 г. в программе *al-'ittiḡāhu l-mu'ākis* «Обратное направление» на катарском телеканале «Аль-Джазира» один из оппонентов движения *fath* намеренно изменил его название на *mataff* «плевок», образовав его от словосочетания *munaẓẓamatu t-tahrīri l-filastīniyyah* (а не *ḥarakatu t-tahrīri l-filastīniyyah*), которое часто используют для обозначения организации ФАТХ, и прочитав его не в обратном порядке, а в прямом.

Стоит отметить, что акроним *fath* далеко не единственный пример реверсивной акронимизации в арабском языке. В этой связи можно вспомнить также такие акронимы, как *fidā'* досл. «жертва» (*al-'ittiḡādu d-dīmuqrāṭiyyu l-filastīniyyu* «Палес-

<sup>3</sup> Подобное «словотворчество» весьма характерно для арабов, которые часто изобретают новые акронимы в попытке выставить кого-либо в невыгодном свете. Так, противники Сил народной мобилизации (*al-ḥašdu š-ša'biyyu*) Ирака придумали для них акроним *ḡāhiš* досл. «осел», расшифровывающийся следующим образом: *ḡayšu l-ḥašdi š-ša'biyyu* «Армия народной мобилизации».

<sup>4</sup> Хотя акроним *fidā'* не полностью соответствует арабскому слову *fidā'*, он может считаться разговорным вариантом данной лексемы.

тинский демократический союз») и *da'm* досл. «поддержка» (**munaẓẓamatu l-'amali d-dīmuqrāṭīyyu** «Демократическая организация труда»). Тем не менее замена прямого порядка слов обратным здесь не призвана завуалировать негативный подтекст, а позволяет скорее подобрать омонимичное значение в тех случаях, когда при прямом прочтении акронима не образуется семантически значимое слово.

## 5. Механизмы формирования акронимов

### 5.1 Подбор буквенных сочетаний

Изучение арабских акронимов показывает, что в их формировании задействованы различные механизмы словообразования. Ниже приведены возможные варианты выбора буквенных сочетаний, составляющих основу акронима, с иллюстрирующими их примерами:

#### 1. Выбор начальных букв всех знаменательных слов из исходной фразы.

<i>dā'īš</i>	ad- <b>d</b> awlatu l-' <b>ī</b> slāmiyyah fi l-' <b>ī</b> rāq wa š-' <b>š</b> ām	«Исламское государство Ирака и Леванта»
<i>hamās</i>	<b>ḥ</b> arakatu l- <b>mu</b> qāwamati l-' <b>ī</b> slāmiyyah	«Исламское движение сопротивления»
<i>tawād</i>	at- <b>ta</b> ḡammu'u l- <b>w</b> aṭaniyyu l-' <b>u</b> rduniyyu d- <b>dī</b> muqrāṭīyy	«Национальное демократическое собрание Иордании»
<i>'amal</i>	' <b>a</b> fwāḡu l- <b>mu</b> qāwamati l- <b>l</b> ubnāniyyah	«Группы ливанского сопротивления»
<i>bāsim</i>	al- <b>b</b> anku s-sa' <b>ū</b> diyy li-l- <b>mu</b> ṣṭalahāt	«Саудовский банк терминов»
<i>maḡd</i>	<b>m</b> unaẓẓamatu l-ḡ <b>ā</b> bhati d- <b>dī</b> muqrāṭīyy	«Организация «Демократический фронт»
<i>tāḡ</i>	<b>ta</b> ḡammu'u ' <b>a</b> mali l-ḡ <b>ā</b> zā'ir	«Собрание «Надежда Алжира»
<i>wās</i>	<b>w</b> ikālatu l-' <b>a</b> n <b>ā</b> 'i s-sa' <b>ū</b> diyyah	«Саудовское информационное агентство»
<i>wā'</i>	<b>w</b> ikālatu l-' <b>a</b> n <b>ā</b> 'i l-' <b>ī</b> rāqīyyah	«Иракское информационное агентство»

Следует отметить, что определенный артикль *-al*, принимающий различные фонетические формы, не учитывается при образовании таких акронимов. Для удобства произношения между согласными, составляющими основу акронима, зачастую вставляются дополнительные гласные, которые не обязательно должны быть производными от исходной фразы. Также необходимо добавить, что согласный звук «хамза» и следующая за ним гласная в исходной фразе могут принимать вид разделительной «хамзы» в акрониме, как в слове *'amal*, либо «алифа», как в словах *dā'īš*, *hamās*, *tawād*, *tāḡ*, *wās* и *wā'*.

#### 2. Выбор начальных букв некоторых знаменательных слов из исходной фразы и опущение других слов.

<i>ḥašd</i>	<b>ḥ</b> izbu š-' <b>š</b> a' <b>bi</b> d- <b>dī</b> muqrāṭīyyi l-' <b>u</b> rduniyy	«Народная демократическая партия Иордании»
<i>wa'd</i>	al- <b>ḥ</b> izbu l- <b>w</b> aṭaniyy li-l-' <b>ad</b> ālah wa t- <b>tan</b> miyyah	«Национальная партия справедливости и развития»

<i>wa'd</i>	al-ḥizbu l-waḥdawīyyu l-'arabīyyu d-ḍīmuqrāṭīyyu	«Арабская объединительная демократическая партия»
-------------	---	---

Как мы видим в вышеуказанных примерах, родовые наименования (al-ḥizbu) и определения (al-'urduniyy) зачастую не представлены в акронимизированной форме. Причина опущения одного или нескольких знаменательных слов при образовании акронима, по всей видимости, кроется в желании его автора создать омоним арабского слова, значение которого отвечало бы цели или философии организации или политической партии такого лица.

3. Выбор начальных букв некоторых знаменательных слов из исходной фразы и любой буквы последнего знаменательного слова.

<i>ḥaqq</i>	ḥarakatu l-qiwā l-ḡadīdati d-ḍīmuqrāṭīyyah	«Движение новых демократических сил»
<i>fātin</i>	mu'assasatu filasṭīn li-t-tasliḥ wa t-tanmiyah	«Палестинский институт кредитования и развития»

В данном случае последняя буква акронима *ḥaqq* (qāf) взята из последнего слова исходной фразы *ḍīmuqrāṭīyyah*, а последняя буква акронима *fātin* (nūn) – из последнего слова исходной фразы *tanmiyah*.

4. Выбор начальных букв всех знаменательных слов и добавление к ним других, не присутствующих в исходной фразе.

<i>mišmiš</i>	al-munaẓẓamatu š-šuyū'īyyatu l-miṣriyyah	«Египетская коммунистическая организация»
---------------	---	---

5. Выбор начальных букв всех знаменательных слов в исходной фразе и изменение их порядка на обратный.

<i>fath</i>	ḥarakatu t-taḥrīri l-filasṭīniyyah	«Движение за освобождение Палестины»
<i>fidā</i>	al-'ittiḥādu d-ḍīmuqrāṭīyyu l-filasṭīniyy	«Палестинский демократический союз»
<i>da'm</i>	munaẓẓamatu l-'amali d-ḍīmuqrāṭīyy	«Демократическая организация труда»

6. Выбор начальных букв некоторых знаменательных слов и их перестановка для образования семантически значимого слова.

<i>wafā</i>	wikālatu l-'anbā'i l-filasṭīniyyah	«Палестинское информационное агентство»
<i>wāḥ</i>	wikālatu ḥabar li-l-'anbā'	«Информационное агентство «Хабар»
<i>wa'd</i>	ḡam'īyyatu l-'amali l-waṭaniyyi d-ḍīmuqrāṭīyy	«Национальное демократическое рабочее общество»

В отличие от полного изменения порядка компонентов в акрониме этот механизм предусматривает частичную перестановку букв для выполнения такой же цели. Например, вместо образования семантически немотивированного акронима *wāf* начальные буквы второго и третьего слова в конструкции *wikālatu l-'anbā'i l-filasṭīniyyah* меняются местами для получения семантически значимого слова *wafā*, разговорной формы лексемы *wafā'* «верность».



## 5.2. Словообразовательные модели

После правильного выбора буквенного сочетания соответствующий акроним оформляется на основе подходящей словообразовательной модели, как и большинство арабских имен. Ниже приведены самые популярные словообразовательные модели акронимов:

<i>fā'il</i> (CVVCVC)	<i>dā'is</i>	ad- <b>dawlatu</b> l-'islāmiyyah fi l-'irāq wa š-šām	«Исламское государство Ирака и Леванта»
	<i>fātin</i>	mu'assasatu <b>filastīn</b> li-t-taslif wa t-tanmiyah	«Палестинский институт кредитования и развития»
<i>fa'al</i> (CVCVVC)	<i>ḥamās</i>	<b>ḥarakatu</b> l- <b>muqāwamati</b> l-'islāmiyyah	«Исламское движение сопротивления»
	<i>tawād</i>	at- <b>taḡammu'u</b> l- <b>waṭaniyyu</b> l-'urduniyyu d- <b>dīmuqrāṭiyy</b>	«Национальное демократическое собрание Иордании»
<i>fa'l</i> (CVCC)	<i>maḡd</i>	<b>munaẓẓamatu</b> l- <b>ḡabhati</b> d- <b>dīmuqrāṭiyy</b>	«Организация «Демократический фронт»
	<i>ḥašd</i>	<b>ḥizbu</b> š-ša'bi d- <b>dīmuqrāṭiyyi</b> l-'urduniyy	«Народная демократическая партия Иордании»
	<i>wa'd</i>	al-ḥizbu l- <b>waṭaniyy</b> li-l-' <b>adālah</b> wa t-tanmiyah	«Национальная партия справедливости и развития»
		al-ḥizbu l- <b>waḥdawiyyu</b> l-'arabiyyu d- <b>dīmuqrāṭiyy</b>	«Арабская объединительная демократическая партия»
		ḡam'iyyatu l-' <b>amali</b> l- <b>waṭaniyyi</b> d- <b>dīmuqrāṭiyy</b>	«Национальное демократическое рабочее общество»
	<i>fath</i>	<b>ḥarakatu</b> t-tahrīri l- <b>filastīniyyah</b>	«Движение за освобождение Палестины»
	<i>da'm</i>	<b>munaẓẓamatu</b> l-' <b>amali</b> d- <b>dīmuqrāṭiyy</b>	«Демократическая организация труда»
<i>ḥaqq</i>	<b>ḥarakatu</b> l- <b>qiwā</b> l-ḡadīdati d- <b>dīmuqrāṭiyyah</b>	«Движение новых демократических сил»	
<i>fa'al</i> (CVCVC)	<i>'amal</i>	'afwāḡu l- <b>muqāwamati</b> l- <b>lubnāniyyah</b>	«Группы ливанского сопротивления»

Таким образом, на основании вышеизложенного можно сделать вывод, что фонетические комбинации акронимов и входящие в их состав гласные звуки подбираются не произвольно, а регулируются достаточно строгими правилами формирования моделей, среди которых выделяется модель *fa'l*.

## 6. Заключение

В настоящем исследовании представлен детальный анализ распространенных акронимов в современном арабском языке. Исходя из его результатов, можно сделать ряд выводов. Во-первых, акронимы в арабском языке появились сравнительно недавно в результате глобализации и развития контактов с различными языками. Большинство из них представляют собой названия различных учреждений, организаций, политических объединений, движений и информационных агентств. Во-вторых, акронимизация как словообразовательный процесс находилась под влиянием английского языка и в меньшей степени французского, на которых велись трансляции телепередач. В-третьих, после периода одновременного использования акронимов, образованных на основе англоязычных словосочетаний и исконных арабских лексем, первые были вытеснены из употребления. Как следствие, за последние несколько десятилетий главным источником создания акронимов является исконная арабская лексика, что свидетельствует о том, что акронимизация стала устоявшимся словообразовательным процессом.

Процесс акронимизации в арабском языке характеризуется следующими особенностями:

1. На первом семантическом уровне акронимы обычно омонимичны существующим словам, что делает их более запоминающимися и красочными. В некоторых случаях арабы могут использовать прием акронимизации в рекламных целях либо злоупотреблять им для изобретения шуточных или обидных акронимов. На втором семантическом уровне не допускается негативная коннотация акронимов – например, *fath* заменяет *hatf*. Многие акронимы, рассмотренные в данной работе, имеют отчетливую положительную окраску, например *mağd* «слава, величие», *fātin* «привлекательный, очаровательный» или *bāsim* «улыбающийся».

2. В основе акронимов лежат различные механизмы словообразования, которые достаточно многообразны и отличаются особыми правилами оформления в зависимости от выбора того или иного компонента исходной фразы. Сформированная классификация показывает богатство вариантов буквенных сочетаний, используемых для создания акронимов.

3. По аналогии с другими языками, например английским, фонетические комбинации акронимов формируются на базе арабских словообразовательных моделей; следовательно, они должны соответствовать законам арабской фонотактики. Так, в случае стечения согласных между ними добавляются гласные с целью их разделения либо используются уже имеющиеся гласные в компонентах акронима для разбивки кластера согласных.

## Литература

1. Блинов А.А. Развитие процесса аббревиации в арабском языке: от Корана до средневековых рукописей // *Труды Института востоковедения РАН. Вып. 29: Проблемы общей и востоковедной лингвистики. Язык меняющийся: семантические и грамматические изменения в языках Азии и Африки*, 2020. С. 20–41.

2. Abu Humeid, A. & Altai, A. Errors in English and Arabic Acronyms Made by Iraqi University Students of Translation // *International Journal of Humanities and Social Science*. 2013. Vol.3, № 6. P.169–182.
3. Abu Salim, I. al-Muḥtaṣarātu l-luġawīyyatu l-ḥadīṣah fī l-luġati l-‘arabīyyah // *Maġallatu maġma‘i l-luġati l-‘arabīyyati l-‘urduniyyi*. ‘Ammān: maġma‘u l-luġati l-‘arabīyyati l-‘urduniyyi, al-‘adad 52. 1997. Š. 255–272.
4. Akmajian, A., Demers R. & Hamish R. *An Introduction to Language and Communication*. Cambridge: MIT Press, 1984.
5. Altakhainah, A. The Linguistic Features of Acronyms in Arabic // *SAGE Open*. 2017. Vol. 7, Issue 4. URL: <https://journals.sagepub.com/doi/10.1177/2158244017746710>.
6. Bauer, L. *English Word-Formation*. Cambridge, UK: Cambridge University Press, 1983.
7. Bolinger, D. *Aspects of Language*. New York: Harcourt Brace Jovanovitch, 1968.
8. Booij, G. *The Grammar of Words*. Oxford, UK: Oxford University Press, 2007.
9. Cannon, G. Abbreviations and Acronyms in English Word-Formation // *American Speech*. 1989. Vol. 64, №2. P. 99–127.
10. Ehrmann, M., Rocca, L., Steinberger, R. & Tanev, H. Acronym Recognition and Processing in 22 Languages // *Proceedings of the 9th Conference Recent Advances in Natural Language Processing*. Bulgaria: Hissar, 2013.
11. Finegan, E. & Besnier, N. *Language: Its Structure and Use*. New York: Harcourt Brace Jovanovitch, 1989.
12. Foster, B. *The Changing English Language*. London, England: Macmillan, 1972.
13. Hamdan, J. & Fareh, S. Acronyms in English and Arabic // *Dirasat: Human and Social Sciences*. 2003. № 30. P. 183–193.
14. Hartmann, R. & Stork, F. *Dictionary of Language and Linguistics*. London, England: Applied Science Publishers, 1976.
15. Haspelmath, M. & Sims, D. *Understanding Morphology* (2nd ed.). London, England: Hodder Education, 2010.
16. Hiġāzī, M. al-‘Iḥtiṣārātu l-ḥadīṣah fī wasā‘ili l-‘i‘lām bayna t-tarġamati l-‘arabīyyah wa l-‘iqtirādi l-mu‘ġamiyy // *Maġallatu maġma‘i l-luġati l-‘arabīyyah bi-l-qāhirah*. Al-Qāhirah: maġma‘u l-luġati l-‘arabīyyah, al-‘adad 92. 2001. Š. 95–111.
17. Hornby, A. (ed.). *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press, 1974.
18. Kanakri, A. *The Structure of Acronyms in English and Arabic*. Irbid, Jordan: Yarmouk University, 1999.
19. Kiparsky: *Word-Formation and the Lexicon*. The University of Kansas: USA, 1982.
20. Lieber, R. *Introducing Morphology*. Cambridge, UK: Cambridge University Press, 2010.

21. Mencken, H. *The American Language*. New York, NY: Knopf, 1919.
22. Montgomery, A. *Acronyms and Abbreviations in Library and Information Work: A Reference Handbook of British Usage*. London: The Library Association, 1975.
23. Mossman, J. *Acronyms, Initialisms and Abbreviations Dictionary*. Detroit: Gale Research Company, 1994.
24. Newmark: *A Textbook of Translation*. New York: Prentice Hall, 1988.
25. Nusseir, A. Manḥūtātu l-budū' // *Mağallatu mağma' i l-luğati l-'arabiyyati l-'urduniyy*. 'Ammān: mağma'u l-luğati l-'arabiyyati l-'urduniyy, al-'adad 32, 1987, Ş. 115–121.
26. Plag, I. *Word-Formation in English*. Cambridge, UK: Cambridge University Press, 2003.
27. Pugh, E. *Pugh's Dictionary of Acronyms and Abbreviations*. Phoenix: Oryx Press, 1981.
28. Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. *A Comprehensive Grammar of the English Language*. London: Longman, 1985.
29. Robbins, R. Acronyms and Abbreviations from Aviation // *American Speech*. 1951. № 26. P. 67–70.
30. Rua: On the Structure of Acronyms and Neighbouring Categories: A Prototype-Based Account // *English Language & Linguistics*, 6.1. Cambridge, UK: Cambridge University Press, 2002. P. 31–60.
31. Russell, I. Among the New Words // *American Speech*. 1946. № 21. P.137–145, 220–226.
32. Samarrā'ī, I. al-Muḥtaşarāt wa r-rumūz fī t-turāşī l-'arabiyy // *Mağallatu mağma' i l-luğati l-'arabiyyati l-'urduniyy*. 'Ammān: mağma'u l-luğati l-'arabiyyati l-'urduniyy, al-'adad 32, 1987. Ş. 105–114.
33. Thomas, R., Ethridge, J. & Ruffner, F., Jr. *Acronyms Dictionary*. Detroit: Gale Research Company, 1960.
34. Thomas, R., Ethridge, J. & Ruffner, F., Jr. *Acronyms and Initialisms Dictionary*. Detroit, MI: Gale Research Company, 1965.

**Специфика дистрибуции агглютинативных аффиксов  
в глагольной словоформе турецкого языка**

**Замира Касымбековна Дербисева**

Кыргызско-Турецкий университет Манас, Бишкек, Киргизия,  
zamira.derbisheva@manas.edu.kg, ORCID 0000-0003-4333-4425

***Аннотация.*** Глагольная словоформа турецкого языка является единицей, в которой отражаются наиболее характерные особенности агглютинативного конструирования слова. Основная цель статьи состоит в том, чтобы соотнести логику сочетаемости грамматических аффиксов в турецкой словоформе с принципами дополнительной дистрибуции. Многообразие модально-временных форм побудило выявить причины двойного-тройного наслонения аффиксов, представляющих одну и ту же категорию.

***Ключевые слова:*** грамматика порядков, дистрибуция морфем, агглютинативная словоформа, модально-временные формы

***Abstract.*** A verbal form in Turkish is the unit that represents the most typical features of agglutinative word construction. The main goal of the paper is to correlate the logic of combining grammatical affixes in the Turkish word-form with the principles of additional distribution. The diversity of modal-temporal forms necessitates inquiry into the reasons for the double and triple co-occurrence of affixes that represent one and the same category.

***Keywords:*** grammar of orders, distribution of morphemes, agglutinative word-form, modal-temporal forms

Турецкий язык представляет собой классический образец агглютинативного языка, где ведущую роль играют аффиксальные морфемы. Для русского языка доминантной единицей является само слово, а именно его частеречный статус, тогда как для турецкого языка конститутивной единицей является аффикс с его агглютинативными свойствами. В русском языке слово как единица словаря и как единица морфологии, как правило, совпадают, и частеречный статус слова закрепляется его морфологическими признаками, тогда как в турецком языке слова в исходном виде трудно дифференцируются по частям речи, поскольку формально не маркированы, а при наращивании аффиксов частеречная принадлежность слова может варьироваться. Если в грамматике русского языка в создании грамматических форм превалирует флективная модификация, то есть словоизменение, то в турецком языке в грамматике ведущим процессом является формообразование, в основе которого лежит аффиксальное конструирование слова.

Во флективных языках проблема сочетаемости грамматических показателей в силу синкретичности флексий неактуальна, тогда как в агглютинативных языках сочетаемость грамматических морфем выполняет конститутивную функцию. Важ-

нейшей особенностью морфологической структуры слова в тюркских языках является проблема взаиморасположения однородных агглютинативных аффиксов относительно основы и относительно друг друга. Исследуя позиционные отношения аффиксов, тюркологи выясняли логику следования грамматических показателей в словоформе. Вопросы грамматики порядка рассматривались в трудах Г. Глисона, И. И. Ревзина, Г. Д. Юлдашевой, Н. А. Баскакова, В. Я. Пинеса, М. С. Джикки, В. Г. Гузева, А. Н. Барулина и др.

Н.А. Баскаков показал связь позиции аффикса со степенью его грамматичности, то есть указал вектор следования аффиксов от менее абстрактных (лексичных) грамматических категорий к более абстрактным. Им в частности установлено, что в тюркских языках наиболее близкое к корню положение занимают аффиксы лексико-грамматического словообразования, например, аффиксы, образующие глаголы от именных частей речи, залоговые показатели и т.п. Далее следуют аффиксы функционально-грамматического словообразования, которые, не изменяя семантики слова и его принадлежности к той или иной части речи, определяют использование слова в словосочетании и предложении. В частности, здесь имеются в виду показатели наклонения и времени, а также некоторые категории имен. Еще дальше от основы располагаются показатели категории числа, лица, принадлежности, которые наименее связаны с лексической семантикой слова и выражают синтаксические отношения в словосочетании и предложении. При этом установлено, что «в структуре тюркского слова наиболее абстрактные грамемы выражены в аффиксах, наиболее удаленных от корня. Чем ближе отстоят от корня те или иные форманты, тем больше вещественного, лексического заключено в категориях, ими выражаемых» [Баскаков 1979, с.137].

Наряду с этим в турецкой грамматике встречаются явления, которые не вписываются в каноны общей лингвистики. О том, что многие факты турецкого языка не поддаются описанию с помощью традиционного европейского теоретико-лингвистического аппарата, говорил А.Н. Барулин. Он подчеркивал необходимость учета типологических особенностей турецкого языка в создании общей морфологической теории: «...к описанию его морфологии невозможно напрямую применить принципы, выработанные европейской лингвистикой. Требуют отдельного осмысления и такие европейские понятия, как грамматическая категория, словоформа, лексема и др.» [Барулин 2019]. Действительно, единицы формообразования в турецком языке демонстрирует иные закономерности сочетаемости грамматических показателей, идущие вразрез с теорией флективной морфологии. В.Г. Гузев агглютинативные конструкции турецкого языка называет функциональными формами и отмечает их окказиональный характер: «Семантика функциональных форм более подвижна, имеет преходящий, окказиональный характер и вместе с тем обладает более сложной структурой. По-видимому, за каждой функциональной формой стоит некая разработанная языковым мышлением типизированная семантическая операция, которая заключается в том, что одно явление, (например, местонахождение предмета) истолковывается как другое, служит средством обозначения другого связанного с ним явления (например, признака предмета)» [Гузев 2019: 154].

Окказиональность турецкой словоформы заведомо предполагает отсутствие закономерности в ее образовании, ее случайный, непредсказуемый характер. Однако последователи дескриптивной теории считают, что и для агглютинативных языков действуют определенные законы дистрибуции. Под дистрибуцией морфемы Г. Глисон понимает сумму всех контекстов, в которых она может встречаться, в отличие от тех контекстов, в которых она встречаться не может [Глисон 1959: 96]. Исследуя правила дистрибуции на материале разных языков, в том числе и турецкого языка, Г. Глисон отмечает, что языку присуща известная степень свободы и в то же время некоторые конструкции, составленные из морфем, имеют строго установленный порядок. «Значение слова зависит не только от состава наличных морфем, но также и от порядка их следования» [Глисон 1959: 98-99].

В теории дескриптивной лингвистики установлены закономерности сочетаемости конструктивных элементов слов и словоформ, которые определяются следующими типами дистрибуции:

1) дополнительная (комплементарная) – два элемента никогда не встречаются в одинаковой позиции; такой тип дистрибуции называют контрастной дистрибуцией функционально различных единиц, принадлежащих к одному классу;

2) свободное чередование факультативных вариантов одной и той же единицы, когда два элемента встречаются в одинаковой позиции.

«Два элемента могут считаться одной морфемой, если они: 1) имеют какое-либо общее значение и 2) стоят в отношении дополнительной дистрибуции, обусловленной каким-либо фонологическим фактором... О двух элементах можно сказать, что они находятся в отношении **дополнительной дистрибуции** (complementary distribution), если один из них встречается в таком окружении, в каком никогда не встречается другой, то есть не существует контекстов, в которых они встречались бы оба» [Глисон 1959: 125].

Джон Лайонз также отмечает значимость комбинаторики морфем: «...каждая лингвистическая единица имеет как контрастивную, так и комбинаторную функцию. Понятно, что две единицы не могут противопоставляться, если они не являются по крайней мере частично эквивалентными по дистрибуции (для единиц, находящихся в отношении дополнительной дистрибуции, вопроса о контрастности не возникает). Единицы, встречающиеся в некотором данном контексте, но не контрастирующие одна с другой, находятся в отношении свободного варьирования» [Лайонз 1978].

Исследуя характер сочетаемости аффиксов в агглютинативном слове, Г. Глисон пишет: «В ряде случаев число аффиксов, используемых в пределах одной парадигмы, бывает очень большим; иногда и отдельное слово может состоять из длинной цепочки морфем. Необходимо найти простой способ констатации всех возможных сложных комбинаций. Часто этого можно добиться, классифицируя морфемы по группам, известным как **порядки** расположения, которые удобнее всего обозначать цифрами... Таким образом, **порядки** – это взаимоисключающие классы морфем, которые занимают определенное место в последовательности морфем, образующих слово» [Глисон 1959: 164].

Логика построения глагольной словоформы турецкого языка в терминах грамматики порядков впервые была описана в работе Г. Глисона, а позже эта тео-

рия была применена для изучения узбекского глагола в статье И. И. Ревзина и Г. Д. Юлдашевой. Г. Глисон предлагает 10 порядков для описания турецкого глагола:

Порядок	Аффиксы
Порядок 1	Аффиксы залога (пассивный, взаимный, возвратный)
Порядок 2	Аффиксы понудительного залога (каузатива)
Порядок 3	Аффикс отрицания
Порядок 4	Аффиксы времени (наст., буд. II, аффикс долженств. накл.)
Порядок 5	Аффикс прошедшего времени (на <i>-di</i> )
Порядок 6	Аффикс 3 лица ( <i>-lar-</i> )
Порядок 7	Аффикс условного наклонения ( <i>-sa-</i> )
Порядок 8	Аффиксы лица (1 л. ед., 1 л. мн., 2 л. ед., 2 л. мн.)
Порядок 9	Вопросительный аффикс ( <i>mi</i> )
Порядок 10	Аффиксы лица (1 л. ед., 1 л. мн., 2 л. ед., 2 л. мн.)

Таблица 1.

В статье И. И. Ревзина, Г. Д. Юлдашевой дается критический анализ по поводу предлагаемой Глисоном классификации, где приводятся доводы относительно несогласия с распределением аффиксов по порядкам, с представленной последовательностью аффиксов, с отнесением аффикса *-lar-* к категории лица (3 л.) и т. д. [Ревзин, Юлдашева 1969: 43].

Предлагая свою классификацию, И. И. Ревзин и Г. Д. Юлдашева выделяют семь порядков распределения глагольных аффиксов. При этом они указывают, что порядки устанавливаются в соответствии со значениями аффиксов. «Становится возможным поставить каждому порядку в соответствие определенное значение. Тогда систему глагольного слова можно представить не только как последовательность аффиксов, но и как последовательность значений, которые будут расположены в следующем порядке: интенсивность (многократность) + возвратность + каузативность, взаимность, совместность, пассивность + отрицание + время, наклонение + лицо, число + вопросительность» [Ревзин, Юлдашева 1969: 49].

В схеме И. И. Ревзина и Г. Д. Юлдашевой выдерживается принцип семантического соответствия порядков и глагольных аффиксов. И вместе с тем дается детализация 3, 5 и 6 порядков: третий порядок включает три подпорядка из залоговых аффиксов, пятый порядок насчитывает 12 подпорядков, содержащий аффиксы всех модально-временных форм; шестой порядок представлен шестью подпорядками, состоящими из аффиксов лично-числовой парадигмы.

Порядок	Аффиксы
Порядок 1	Аффикс интенсивности ( <i>ла/қила</i> )
Порядок 2	Аффикс возвратности ( <i>ин/н</i> )
Порядок 3	Аффикс залога (состоит из 3 подпорядков: каузативности ( <i>тир</i> ), взаимности ( <i>ши/и</i> ), пассивности ( <i>л</i> )).
Порядок 4	Аффикс отрицания ( <i>ма</i> )



Порядок 5	Аффиксы времени и наклонения (состоит из 12 подпорядков)
Порядок 6	Аффиксы лица и числа (состоит из 6 подпорядков)
Порядок 7	Аффикс вопросительности

Таблица 2.

В обеих таблицах представлены типичные модели сочетания аффиксов, представляющих разные порядки, что свидетельствует о категориальном назначении каждого порядка, уточняющего глагольное слово с точки зрения деривационных и грамматических признаков. Надо отметить, что в данных таблицах основное внимание уделяется последовательности порядков между собой, в соответствии с которой устанавливается «очередность» следования морфем в потенциальной глагольной словоформе. Вместе с тем не менее важной является проблема сочетаемости аффиксов внутри порядка, представляющего определенную грамматическую категорию. С этой точки зрения можно выделить порядки, в которых аффиксы вступают в свободные сочетания друг с другом, к примеру, к ним относится 3 порядок, представляющий залоговые аффиксы. В пределах этого порядка залоговые аффиксы могут свободно сочетаться между собой, не противопоставляясь между собой, не взаимоисключая друг друга (*кий+ин+тир+ма+й+ман*), что свидетельствует об их неграмматическом, то есть деривационном характере. Напротив, в 5 и 6 порядках отношения подпорядков существенно отличаются от характера взаимодействия морфем 3 порядка. В рамках 6 порядка лично-числовые аффиксы при формировании глагольной словоформы находятся в отношениях дополнительной дистрибуции, то есть они могут взаимодействовать друг друга, но не встречаются в словоформе совместно. Такое свойство лично-числовых аффиксов также подтверждает грамматический статус сказуемости. Характер сочетаемости аффиксов 5 порядка, включающего модальные и временные аффиксы, также соответствует критериям грамматичности: случаев совместного употребления аффиксов времени между собой или сочетания аффикса наклонения и времени в узбекских глагольных формах не зафиксировано.

При рассмотрении структуры глагольной словоформы в турецком языке, напротив, распространены такие факты, как: 1) сочетаемость однородных грамматических показателей времени в пределах одной единицы; 2) совместимость аффиксов времени и наклонения. Логика сочетаемости грамматических аффиксов в турецкой словоформе не всегда отвечает указанным принципам определения грамматичности, в соответствии с которыми показатели с однородным грамматическим значением не могут встречаться в пределах одной формы. О невозможности «соседства» двух однородных элементов в одной словоформе А.А. Зализняк писал следующее: «Основная зависимость, связывающая грамматические значения словоформ и грамматические категории, состоит в том, что грамматическое значение словоформы может содержать не более одной граммемы одной и той же грамматической категории. ...Так, например, в грамматическое значение словоформы **не могут** (выделено – *ЗД*) одновременно входить граммемы «Д. падеж» и «П. падеж», «ед. число» и «мн. число» и т. п. [Зализняк 1967: 27].

В.А. Плуноян также отмечает один из конститутивных признаков флективной словоформы, связанный с синтагматической несовместимостью элементов одной грамматической категории: «...в одну грамматическую категорию объединяются взаимоисключающие значения – т.е. такие, которые не могут быть выражены одновременно в одной и той же словоформе данного языка. Так, в русском языке имя не может быть одновременно в единственном и множественном числе, глагол – одновременно в настоящем и прошедшем времени и т. п. [Плуноян 2003: 113].

Напротив, морфемы, представляющие неграмматические значения, могут реализоваться в рамках одной словоформы. Например, префиксы и суффиксы, выражающие однородные аспектуально-акционсартные значения, синтагматически могут быть совместимы в одной словоформе. Ср.: *no/na/открывать, no/rac/кидать* или *притани/ов/ыва/ть, выкристаллиз/ов/ыва/ться* и др. Правомерно предположить, что принцип дополнительной дистрибуции может быть применен как критерий дифференциации формообразующих и словообразующих значений морфем.

Для системы временных и модальных показателей турецкого глагола характерна несколько иная логика связей. Основу временной системы турецкого языка составляет пять форм, каждая из которых маркирована определенным временным показателем (*yor-*, *-di-*, *-miş-*, *-acak-*, *-ar-*):

1. Формы настоящего времени на *-yor-*: *yazıyorum* «я пишу»
2. Формы очевидного прошедшего времени на *-di-*: *yazdım* «я написал»
3. Формы неочевидного прошедшего времени на *-miş-*: *yazmışım* «я писал»
4. Формы категорического будущего времени на *-acak-*: *yazacağım* «я напишу»
5. Формы неопределенного будущего времени на *-ar-*: *yazarım* «я напишу»

Данные формы составляют ядро темпоральной системы турецкого языка и имеют широкую сферу функционирования. Указанные временные формы реализуются в трех модальных сферах:

1. Объективная модальность (*Hikaye*),
2. Субъективная модальность (*Rivayet*),
3. Условная модальность (*Şart*).

В результате взаимодействия этих двух систем образуются многочисленные модально-временные формы, которые представляют периферию данной системы.

### **I. Временные формы объективной модальности (*Hikaye birleşik kipleri*)**

Выделяется шесть форм объективной модальности:

1) **Форма определенного имперфекта** (*Şimdiki Zamanın Hikayesi*). Ср.: *Ona annesi çok güzel yemekler yap/ıyor/du* «Мама ему готовила очень вкусные блюда»<sup>1</sup>.

2) **Форма будущего-прошедшего времени** (*Gelecek Zamanın Hikayesi*). Будущее-прошедшее (иначе будущее намерения). Ср.: *Deniz'le bu kadar samimi olma/yacak/t/n* «С Дениз на этот раз я не был откровенным».

3) **Форма неопределенного имперфекта** (*Geniş Zamanı Hikayesi*).

Неопределенный имперфект (прошедшее незаконченное) указывает на несовершенное действие, не связанное с определенным моментом. Ср.: *Bu hatayı ben*

<sup>1</sup> Примеры взяты из работы Эйюп Гениша «Турецкий язык в упражнениях». М: Либроком, 2011. – 444 с.

*уарmış olsaydım özür dil/er/di/m* «Если бы я допустил эту ошибку, я бы принес свои извинения».

4) **Форма давнопрошедшего времени** (*Duyulan (Miş'li) Geçmiş Zamanın Hikayesi*). Ср.: *Hastalıktan çok zayıfla/miş/ti* «Он сильно ослабел от болезни».

5) **Форма прошедшего времени долженствовательного наклонения** (*Gereklilik Kipinin Hikayesi*). Ср.: *Anneni böyle küçük şeyler yüzüne üzme/meli/y/di/n* «Ты не должен был расстраивать маму такими пустяками».

6. **Форма прошедшего времени условно-желательного наклонения** (*Dilek şart Kipinin Hikayesi*). Ср.: *Beni bir saat bekletme/se/y/di/n toplantıya gecikmezdim* «Если бы ты не заставил меня целый час ждать, я бы не опоздал на собрание».

## II. Временные формы субъективной модальности (*Rivayet birleşik kipleri*)

Выделяется пять форм субъективной модальности:

1) **Форма настоящего времени субъективной модальности** (*Şimdiki Zamanın Rivayeti*). Ср.: *Reyhan çokgüzel yemekler pişir/iyor/miş* «Рейхан, оказывается, приготовила очень вкусное блюдо».

2) **Форма будущего времени субъективной модальности** (*Gelecek Zamanın Rivayeti*). Ср.: *Erkan kendisine özel hoca tut/acak/miş* «Эркан, говорят, нанял личного учителя».

3) **Форма расширенного времени субъективной модальности** (*Geniş Zamanın Rivayeti*). Ср.: *Eskiden bu gölde insanlar yüz/er/miş* «Раньше, говорят, здесь купались люди».

4) **Форма долженствовательного наклонения субъективной модальности** (*Gereklilik Kipinin Rivayeti*). Ср.: *Bir daha onu bu kadar bekletme/ meli/y/miş/im* «Я не должен был его тогда заставлять ждать».

5) **Форма условно-желательного наклонения субъективной модальности** (*Dilek şart Kipinin Rivayeti*). Ср.: *Keşke bunca yıl onun değişmesini bekleme/se/y/ miş/im* «Эх, не надо было мне ждать столько времени, что он изменится».

## III. Временные формы условной модальности (*Şart birleşik kipleri*)

1) **Форма настоящего времени условной модальности** (*Şimdiki Zamanın şartı*) Ср.: *Annen çağır/iyor/sa gitmelisin*. «Если мать зовет, надо идти».

2) **Форма будущего времени условной модальности** (*Gelecek Zamanın şartı*). Ср.: *Eğer televizyon seyretme/yecek/se/n boşu boşuna açta* «Если не смотришь телевизор, не надо зря его включать».

3) **Форма расширенного времени условной модальности** (*Geniş Zamanın şartı*). Ср.: *Televizyonu yakından seyred/er/se/n gözlerin bozulur* «Если будешь близко смотреть телевизор, испортишь глаза».

4) **Форма неочевидного прошедшего времени условной модальности** (*Duyulan (Miş'li) Geçmiş Zamanın şartı*). Ср.: *Selçuk'un morali bozul/muş/sa kimse onu düzeltmez* «Если у Сельчука испортится настроение, то никто его не поднимет».

5) **Форма очевидного прошедшего времени условной модальности** (*Görülen (Bilinen, Belirtili) Geçmiş Zamanın şartı*). Ср.: *Ödevini bitirme/di/y/se/n dışarı çıkmazsın* «Пока не закончишь работу, на улицу не выйдешь».

Агглютинативное формообразование в турецком языке не ограничивается только двойным усложнением, есть и регулярные формы с трехступенчатым аффиксальным наслоением, которые называются *katmerli birleşik çekim*. Формы *hikayenin şartı (gel/iyor/duy/sa/n, gel/miş/tiy/se)* как бы придают значение *обусловленности* всем 3 формам *hikaye*. Формы *rivayetin şartı (bil/iyor/muş/sa/n, gel/ecek/miş/se/n)* придают значение *обусловленности* основным формам *rivayet*.

Как видим, диапазон варьирования взаимосочетаемости аффиксов в глагольной словоформе чрезвычайно широк. Категории времени и наклонения тесно взаимодействуют друг с другом, находясь в сложном семантическом переплетении. Время, предположение, намерение, желание, условие, долженствование преобразуются во всех временных планах, и в разных вариациях эти семантические аспекты составляют сложное модально-временное содержание глагольного слова. Время и модальность в турецком языке, действительно, представляют сложный синтез: каждый временной план вторгается в сферу модальности, и происходит аффиксальное конструирование глагольной формы в виде формальных преобразований, представляющих в итоге сложную комбинацию модально-временных аффиксов. Нам предстоит разобраться, какова логика отношений категории наклонения и времени в турецком языке: считать их одной категорией или рассматривать их как две разные категории. Так, В. Г. Гузев, ссылаясь на утвердившуюся в тюркском языкознании традицию, относит к наклонению «...лишь такие финитные глагольные формы, в которых модальное значение выступает как первичное по отношению к временному значению и как общее, присутствующее во всех временных формах данного наклонения. Эта семантическая особенность наклонений находит отражение в порядке аффиксации в формах тех из них, которые имеют специальный показатель – носитель модального значения: в них показатель с временным значением следует за показателем, имеющим модальное значение» [Гузев 2019: 217]. Согласно точке зрения В.Г. Гузева, речь идет о единой модально-временной категории, где категория наклонения выступает как ведущий конститутивный компонент при выражении временных значений. Трактовка наклонения (*Kip*) как основополагающей категории является общепринятой и в работах многих турецких ученых (Хайдар Эдискун, Мухамед Эргин, Нуреттин Коч и др.).

Типологическая специфичность агглютинативной словоформы турецкого глагола заключается в следующем.

1. В подавляющем большинстве указанных форм аффиксы наклонения и времени сопровождают друг друга и ведут себя как показатели двух разных категорий. При этом не все аффиксы наклонения вступают в связь с временными показателями, среди них следует отметить аффиксы условного, долженствования и желательного наклонений.

2. Другая группа представляет формы, где наблюдается свободная совмещаемость двух временных аффиксов в рамках одной словоформы. К примеру, совместное употребление граммем настоящего и прошедшего времени в глагольной словоформе: *oku-yor-du-m*; граммем будущего и прошедшего времени: *oku-yacak-ti-m*. Вместе с тем нельзя не отметить тот факт, что при наличии двух временных аффиксов в глагольной словоформе отсутствуют аффиксы наклонения.

Семантическое обоснование употребительности глагольных форм с двойными темпоральными аффиксами убедительно было дано В.Г. Гузевым: «...специальные формы, служащие для сигнализации о хронологических соотношениях действий с каким-либо моментом в прошлом, имеют наряду с первичной точкой ориентации (грамматический настоящий момент) также и вторичную – грамматический прошлый момент. Это – грамматизованный образ такого временного момента, который мыслится как предшествующий обобщенному образу настоящего момента. Категории времени тех языков, в которых сформировались две грамматические отправные точки, представляют собой системы двойной ориентации. Грамматизованный образ прошлого момента обладает, естественно, той же подвижностью, что и грамматический настоящий момент, и так же способен использоваться прямо, т.е. ассоциироваться с образами реально прошлых событий, или относительно, т.е. передавать, например, события реального будущего. Вместе с тем анализ функционирования временных форм, значения которых ориентированы на грамматический прошлый момент (например, давнопрошедших времен), убеждает в том, что каждая из них, по сути дела, имеет двойную ориентацию, т.е. соотносит действия не только с грамматическим прошлым моментом, но и через него с настоящим моментом (в силу соотнесенности обоих моментов между собой) [Гузев 1990, с.69].

Однако с точки зрения правил дополнительной дистрибуции факты двойного употребления временных аффиксов ставят под сомнение их грамматический характер, и, следовательно, их совместная сочетаемость сближает их с залоговыми аффиксами, имеющими лексико-грамматический статус.

### **Выводы**

С точки зрения теоретических канонов индоевропейской языковой традиции категория времени в турецком языке не соответствует требованиям грамматической категории в том отношении, что в рамках глагольной словоформы широко распространено совместное употребление двух однородных временных аффиксов; при описании таких фактов следует рассматривать категорию времени отдельно от категории наклонения, трактуя временные аффиксы как деривативные, а аффиксы наклонения как грамматические.

В случае рассмотрения данных фактов турецкого языка в плоскости агглютинативной типологии, можно говорить о категории времени как неотъемлемой части категории наклонения, основной функцией которой является аспектуально-темпоральная конкретизация осуществления действия.

### **Литература**

1. Барулин А.Н. *Грамматические категории турецкой именной словоформы и ее структура*. 2019.  
URL: [https://www.researchgate.net/publication/333675242\\_Grammaticeskie\\_kategorii\\_tureckoj\\_imennoj\\_slovoformy\\_i\\_ee\\_struktura](https://www.researchgate.net/publication/333675242_Grammaticeskie_kategorii_tureckoj_imennoj_slovoformy_i_ee_struktura) Дата обращения 19.03.2021

2. Баскаков Н. А. *Историко-типологическая морфология тюркских языков: структура слова и механизм агглютинации*. Институт языкознания АН СССР. М.: Наука; 1979. 276 с.
3. Глисон, Генри Алан. *Введение в дескриптивную лингвистику*. Перевод с англ. Е.С. Кубряковой и В.П. Мурат. М.: Изд-во иностранной литературы; 1959.
4. Гузев В. Г. *Очерки по теории тюркского словоизменения: глагол: на материале староанатолийско-тюркского языка*. Л.: Изд-во Ленинградского университета; 1990.
5. Гузев В.Г. *Избранное: К 80-летию*, СПб: НП-Принт; 2019. 638 с.
6. Лайонз Д. *Введение в теоретическую лингвистику*. М.: Прогресс; 1978.
7. Зализняк А.А. *Русское именное словоизменение*. М.: Наука; 1967.
8. Плунгян В.А. *Общая морфология: Введение в проблематику: Учебное пособие*. Изд. 2-е, исправленное. М.: Едиториал УРСС; 2003. 384 с.
9. Ревзин И.И., Юлдашева Г.Д. Грамматика порядков и ее использования // *Вопросы языкознания*, 1969. № 1. С. 42–57.

## Графическое слово в зеркале японского языка: к уточнению понятия

**Алексей Сергеевич Зверев**

Институт востоковедения РАН, Москва, Россия,

zvereval@inbox.ru, ORCID 0000-0001-9059-4057

**Аннотация.** Статья уточняет понятие графического слова, раскрывая наличие у таких письменных единиц внутренней графической структуры, основанной на линейных иерархиях классов графем. Развиваемый подход может быть применен как к японскому, так и к европейским языкам типа русского, причем в последних усматривается оппозиция двух алфавитов, а пробел графически связывается с предшествующим словом. Также поднимается вопрос о качественной интерпретации линейных иерархий, и в качестве параметра, к которому могут быть чувствительны японская и русская (европейская) иерархии, предлагается количество графического материала.

**Ключевые слова:** графическое слово, орфографическое слово, пунктуация, знаки препинания, пробел, письменный язык, японский язык, русский язык

**Abstract.** The article refines the concept of graphic word arguing such written units to have an internal graphic structure based on linear hierarchies of grapheme classes. The proposed approach can be applied to Japanese as well as to European languages like Russian. The latter are claimed to feature the opposition of two alphabets while a space is graphically attached to the previous word. The issue of qualitative interpretation of linear hierarchies is also raised and the amount of graphic material is suggested as a parameter both Japanese and Russian (European) hierarchies might be sensitive to.

**Keywords:** graphic word, orthographic word, punctuation, space, written language, Japanese, Russian

### 1. Введение

Обсуждение фундаментального понятия «слова» и критериев словоделения обычно не обходится без упоминания единиц, именуемых графическими (орфографическими) словами [Яхонтов 1963; Haspelmath 2011: 4; Алпатов 2018: 45-46]. Графическое слово определяют как фрагмент текста между двумя пробелами. По своей форме этот термин претендует на высокую степень обобщения и в принципе должен описывать реализацию «слов» в условиях визуально-графической коммуникации – как некий базовый феномен синтагматической организации письменной речи.

Как правило, обсуждение феномена графических слов ведется в довольно критическом ключе, однако демонстрируемый при этом уровень проработки проблематики остается достаточно поверхностным, в особенности на фоне тщательного изучения других сторон «слова». Среди прочего обращают на себя внимание следующие два распространенных недостатка.

Во-первых, понятие графического слова не адаптируется к письменным языкам другого устройства, где пробелы не используются или их употребление значительно отличается. Вместо этого языки типа японского только кратко упоминаются как свидетельства неуниверсальности модели – и этим даже мотивируют далеко идущие выводы о нерелевантности самого понятия [Мельчук 1997: 198-199]. Подобный подход исходит из необоснованного требования применимости «среднеевропейского стандарта» к любым «экзотическим» языкам и резко контрастирует с практикуемым в лингвистике отношением, в частности, к аналогичным явлениям устной речи (в плане определения фонетического слова ср. [Spencer, Luís 2012: 74-77] о сопоставимости разных типов словесной просодии).

Между тем опыт изучения японского языка заставляет предполагать, что при отсутствии пробелов «словесный» уровень графического плана выражения может строиться на других принципах, пусть даже выделяемые в такой системе единицы не вполне точно соответствуют графическим словам европейских языков [Шалипина 2007: 68-69]. Как и в случае с фонетическими словами и ударением, японский материал может служить базой для расширения, переосмысления понятия, а не отказа от него.

Во-вторых, в этой простейшей модели графического слова не учитываются знаки препинания, так что текст полностью сводится к последовательности букв и пробелов. Так или иначе знаки препинания выносятся за рамки изучаемого объекта-слова и чаще всего даже не упоминаются, как если бы вовсе отсутствовали в языке. Такое упрощение выглядит вполне допустимым в морфологических исследованиях типа [Зализняк 1967], но в целом едва ли может претендовать на адекватное представление реальной письменной речи.

Вероятно, не последнюю роль в утверждении такого подхода к пунктуации играет распространение фоноцентричной идеологии: устный язык фактически отождествляется с языком вообще, а письменный представляет интерес в той мере, в которой отражает явления звучащей речи. В таких условиях собственное устройство визуально-графического сигнала оказывается на периферии внимания исследователей, что позволяет систематически игнорировать знаки препинания как единицы без (сегментных) устных коррелятов, которые трудно соотнести с «приоритетными» устными словами или хотя бы их частями.

В настоящей статье делается попытка уточнить понятие графического слова с опорой на материал японского языка и результаты исследователей-японистов. Помимо японского, мы рассмотрим и русский язык, имея в виду, что сделанные относительно него заключения должны быть в своей основе применимы и к другим языкам, использующим пробелы (и пунктуацию) по тому же «европейскому» образцу. В традиционную модель мы намерены добавить знаки препинания и в целом более полно описать графическое устройство европейских текстов, ориентируясь главным образом на печатную или электронную форму письменного языка.

При этом графическое слово понимается нами прежде всего как феномен графического означающего и в таком качестве наделяется самостоятельной ценностью как объект лингвистического исследования. Вместо того чтобы принципиально рассматривать этот феномен сквозь призму грамматики или фонетического устройства ассоциированной устной речи, мы принимаем последние во внимание



как отдельные важные аспекты слова, соотношение которых с графикой может носить нетривиальный характер. Не следует ожидать, что деление текста на графические слова будет в точности отражать какой-то другой аспект синтагматической организации – во всяком случае, не в большей мере, чем, скажем, фонетическое членение совпадает с грамматическим, а другие известные критерии словоделения сами согласуются друг с другом.

## 2. Японская графика и иерархии линейного порядка

Использование самостоятельной графемы-пробела в японском языке ограничено рядом редких специфических случаев, поэтому японское письмо традиционно считается беспробельным. Однако отсутствие привычных для европейцев межсловных пробелов компенсируется особым образом упорядоченной графической неоднородностью.

Как известно, в графической структуре японского текста ведущая роль отводится противопоставлению графем разных классов, которые организуются в исчислимый набор шаблонов. В составе таких шаблонов учитывают иероглифы (И), знаки слоговых азбук хираганы (Х) и катаканы (К), реже латиницу и даже такие периферийные жанрово-специфичные единицы, как арабские цифры [Tanaka 1980: 320-321; Mizutani et al. 1987: 109-173]. В то же время нам не известно описаний, которые бы включали в такие шаблоны знаки препинания (ЗП), – несмотря на то, что этот класс графем давно вошел в японскую письменную речь и хорошо представлен практически в любом тексте.

С точки зрения грамматики графически выделяемые сегменты в основном соответствуют синтаксемам-*бунсэцу* – единицам, состоящим из одного знаменательного слова, и, возможно, примыкающих к нему служебных. По объему информации и другим свойствам такие единицы вполне сопоставимы со словами русского языка, и в рамках настоящего исследования достаточно упрощенно понимать этот традиционный термин японской грамматики как один из синонимов «слова». Впрочем, безотносительно к применимости обоих терминов (с учетом всего разнообразия неграфических критериев) к данному уровню членения, нас будет интересовать сама сущность принципа, который позволяет с известной регулярностью разграничивать подобные единицы в японском тексте.

Сравнительно четкая выделяемость сегментов-*бунсэцу* на письме основывается на характерном единообразии грамматической структуры. Письменный японский – последовательно левостветящийся язык; в клаузе имена чередуются с постпозитивно оформляющими их служебными показателями, а завершает ее вершинный глагол с суффиксальной флексией – ср. схематичное представление простого японского предложения по [Tanaka 1980: 320]:

(1)  $N_1J_1 N_2J_2 \dots V$ .

( $N_i$  – существительное,  $J_i$  – падежный показатель при  $N_i$ ,  $V$  – глагол)

При этом с давних пор в японском языке получила развитие тенденция использовать иероглифы (позднее также катакану) для записи знаменательных основ, а грамматические элементы (суффиксы, энклитики) передавать при помощи слоговых азбук (в этой функции утвердилась хирагана). Наиболее строго такое распределение соблюдалось в первой половине XX в., после чего было во многом

нарушено – но все еще продолжает играть ключевую роль в японской письменности [Алпатов 2003: 44-45; Алпатов и др. 2008: 27-28, 61-62]. В соответствии с этим принципом, графика может выступать средством различения служебных показателей и их лексических источников (ср. *とき toki* ‘когда’ и *時 toki* ‘время’): если грамматикализованные лексемы сохраняют прозрачные знаменательные аналоги, первые предпочтительно записывать одной хираганой, а для вторых применяется иероглифика [Маевский 2000: 124].

В современном языке употребление катаканы ограничивается т.н. *гайрайго* (новыми заимствованиями, прежде всего из английского языка) и рядом других особых случаев<sup>1</sup>; в контексте данной работы можно кратко характеризовать катакану как специализированный заместитель иероглифики. Стоит подчеркнуть, что сближение иероглифов и катаканы как специализированных средств записи знаменательных основ для наших целей является более важным, чем сближение катаканы с хираганой и их противопоставление иероглифам (как фонетическое vs. нефонетическое письмо), которое обычно выходит на первый план в описаниях японской письменности.

Таким образом, постпозитивный характер японских служебных показателей предопределил главный принцип линейной организации текста: как правило, бунсэцу начинается с иероглифики (или катаканы), т.е. знаменательной части, а заканчивается грамматически нагруженной хираганой. Поэтому там, где хирагана сменяется иероглификой (катаканой), проходят границы между сегментами. Соответственно, базовый шаблон для графической репрезентации бунсэцу имеет следующий вид: **И/(К/Х)-Х**. Проиллюстрируем эту идею простым примером:

- (2) わたしはリンゴを食べた  
*watashi wa ringo o tabe-ta*  
 я TOP яблоко ACC есть-PST  
 «Я съел яблоко».

Приведенное предложение состоит из трех синтаксем-бунсэцу: *わたし* «я», *リンゴ* «яблоко», *食べた* «съел». Первая единица записана одной хираганой (*わたし* – X-X, или просто X), вторая – катаканой и хираганой (*リンゴ* – K-X), а третья демонстрирует наиболее распространенную комбинацию иероглифики и хираганы (*食べた*, И-X<sup>2</sup>). Можно заметить, что переход от хираганы к иероглифике или катакане в линейной последовательности графем сигнализирует о границе между бунсэцу.

Разумеется, речь здесь не идет о жестком правиле, а только о сильной тенденции. Но стоит указать, что различные отклонения, сложные переходные случаи (аналитические конструкции, композиты, сцепления) зачастую имеют нетривиальный статус и по другим критериям словоделения. Например, именные сцепления могут демонстрировать признаки как слов, так и словосочетаний, а с точки

<sup>1</sup> Среди прочего указывают, что к катакане обращаются с целью избежать графически однородных последовательностей из одной хираганы, затрудняющих восприятие текста [Алпатов 2003: 66].

<sup>2</sup> Здесь хираганой записано окончание глагольной словоформы, которое включает часть основы (*べ*) и флексию (*た*).

зрения графики обычно представляют собой довольно протяженные последовательности иероглифов (типа 新和英大辞典 *shin-waei-dai-jiten* «Новый большой японо-английский словарь»). Отдельные именные компоненты сцеплений остаются графически не разграниченными, но степень их самостоятельности и семантико-синтаксической интеграции может варьировать, и относительно синтагматического статуса этих образований ведутся дискуссии (ср. [Алпатов и др. 2008: 538-544]). Вспомогательные глаголы типа いる или みる, хотя традиционно и трактуются как относительно самостоятельные единицы и даже выделяются в отдельные бунсэцу, как правило, записываются хираганой и потому графически примыкают к предшествующему смысловому глаголу: конструкции типа 取ってみる *totte miru* ‘попробовать взять’ удовлетворяют шаблону И-Х как единое целое. В статистическом исследовании японской графики [Mizutani et al. 1987: 113] особо оговаривается включение таких вспомогательных глаголов в состав предшествующих бунсэцу.

В целом можно утверждать, что и в таких спорных случаях поведение графики едва ли бывает совершенно произвольным или нелогичным, но скорее поддерживает одну из конкурирующих (по разным критериям) синтагматических интерпретаций. Вместе с тем мы оставляем за графикой и право следовать собственным путем, не повторяя другие виды членения, но реализуя по-своему осмысленный подход к структурированию текста, особенности которого могут быть предметом специальных исследований.

Описанная модель становится более простой и гибкой, если от перечисления шаблонов перейти к иерархиям линейного порядка, которые задают разрешенные последовательности графем разных классов в пределах одного сегмента<sup>3</sup>. Так, три варианта записи слова *warai* ‘смех’ – 笑 (И), わらい (Х), 笑い (И-Х) – должны характеризоваться тремя разными шаблонами (указанными в скобках), но одинаково удовлетворяют порядку И > Х: хирагана может следовать за иероглификой, но не наоборот.

Как показывают данные [Mizutani et al. 1987: 113] об употреблении иероглифики и хираганы (бунсэцу, содержащие в своем составе знаки других классов, были исключены из рассмотрения; их оказалось менее 1%), иерархия И > Х покрывает 96,8% материала. Эта иерархия описывает структуру, например, таких единиц, как 昨夜 *sakuya* «прошлой ночью» (строится по шаблону И, которому следуют 7,7% бунсэцу), しかし *shikashi* «однако» (Х, 28,7%) и 作った *tsukutta* «сделал» (И-Х, 60,4% – наиболее частотный случай).

Если добавить в иерархию катакану – в качестве второстепенного заместителя иероглифики, – то она принимает следующий вид: И/(К) > Х. Эта иерархическая модель формализует популярные представления об основаниях выделяемости бунсэцу в японском письме, которые излагались выше, но не исчерпывает всего графического состава текста: для достижения удовлетворительной полноты описания в него следует включить знаки препинания.

---

<sup>3</sup> Хотя в литературе регулярно отмечается, что переход от хираганы к иероглифике указывает на границу бунсэцу, прямой формулировки правила в иерархических терминах нам не встречалось. Иерархический подход предлагался нами в статье [Зверев 2018: 34-35].

В отличие от иероглифов и знаков слоговых азбук, знаки препинания не служат для передачи на письме сегментного состава устных бунсэцу (сторона вопроса, которая, видимо, занимает больше внимания исследователей), но присутствуют в тексте наравне с ними и тоже участвуют в его графической организации. В частном случае знаки препинания могут разбивать протяженные последовательности графем одного класса (например, хираганы), играя решающую роль в снятии неоднозначности. Употребления такого рода отдельно отмечаются в описаниях японской запятой [Данилов, Сыромятников 2001: 50]. Ср. пару примеров, которые различаются только пунктуацией:

(3) ここで、はきものを脱いでください。

*koko de, hakimono o nui-de kudasai.*  
здесь INS, обувь ACC снимать-CNV IMP.

«Здесь снимите, пожалуйста, обувь». [Данилов, Сыромятников 2001: 50]

(4) ここでは、きものを脱いでください。

*koko de wa, kimono o nui-de kudasai.*  
здесь INS TOP, одежда ACC снимать-CNV IMP.

«А теперь – раздевайтесь!» [Данилов, Сыромятников 2001: 50]

Оба предложения состоят из трех синтаксем-бунсэцу. Последняя из этих единиц выделяется без затруднений благодаря регулярному сочетанию иероглифики и хираганы (脱いでください), но граница между первой и второй, целиком записанных хираганой, определяется исключительно положением запятой. Различия в поверхностно-графическом членении текста закономерным образом сказываются на грамматической структуре, лексическом составе и, в конечном итоге, смысле двух высказываний.

Знаки препинания составляют особый класс графем, и сильной стороной развиваемой модели является ее открытость для включения новых классов. Остается установить, какое место занимает пунктуация в той иерархии, которая определяет линейное членение японского текста на графические сегменты. Иными словами, где обычно проходят искомые графические границы – до или после знака препинания?

Ключ к ответу на этот вопрос находится в поведении пунктуации с точки зрения параметров графической связности – в характере ограничений на перенос, асимметричной конфигурации пустого пространства вокруг знака. Как уже было нами показано в специальном исследовании, абсолютное большинство знаков препинания графически аффилировано с преконтекстом<sup>4</sup> [Зверев 2018]. Именно такое поведение демонстрируется частотным ядром пунктуации – точками и запятыми. Поэтому класс пунктуационных графем должен занять место в конце иерархии, а в линейной графической структуре текста они в большинстве своем входят в состав предшествующих сегментов.

Примечательно, что асимметрия графических связей пунктуационных единиц чаще всего прозрачным образом согласуется с другими аспектами синтагма-

<sup>4</sup> Исключения в основном сводятся к поведению открывающих парных знаков, скобок и кавычек – (, 「, 『 и др. – единиц не вполне самостоятельных, сопряженных со своими закрывающими парами.

тики. Например, знаки конца предложения не только примыкают к нему графически, но и принадлежат этому предложению синтаксически (как завершающий компонент), характеризуют именно его (а не следующее предложение) в плане модально-экспрессивной окраски и т.д. Этой стороны вопроса мы немного касались в том же исследовании [Зверев 2018: 34-35].

Таким образом, базовая иерархия, характеризующая устройство единиц, которые можно условно именовать графическими бунсэцу, принимает трехуровневый вид: **И(К) > X > ЗП**. В соответствии с этой иерархией, например, предложение (3) делится на следующие сегменты (в скобках указаны представленные уровни иерархии): **ここで、(X > ЗП), はきものを (X), 脱いでください。(И > X > ЗП)**.

Предложенная иерархия формализует и уточняет известный принцип, который определяет, пусть не без исключений, деление японских текстов на сегменты, в какой-то мере подобные «европейским» графическим словам. Более того, помимо положения границ (т.е. сегментации как таковой), иерархическая модель одновременно описывает и внутреннюю структуру выделяемых единиц, накладывая на нее определенные ограничения. С нашей точки зрения, однако, данная модель представляет больший интерес не как анализ собственно японской графики, а как типовое описание некоторого графического явления, которое в той или иной форме может наблюдаться и в других письменных языках.

### **3. Иерархические структуры в русской графике и оппозиция двух алфавитов**

Сформированное на японском материале представление о графических словах как о единицах, которые выделяются на основе иерархий классов графем, определяющих их внутреннюю линейную структуру, дает возможность по-новому посмотреть на устройство таких единиц в языках «среднеевропейского стандарта». Обнаруживается, что упорядоченная графическая неоднородность «японского» образца в действительности присутствует и в привычных сегментах между пробелами в письменной русской речи.

Преобладающий «пробелоцентричный» подход в принципе не ставит вопроса о внутренней структуре графических слов – разнообразии среди всех прочих графических классов, противопоставленных сегментирующим пробелам, фактически игнорируется. Взгляд на русский текст сквозь призму японского языка позволяет выделить в составе графических слов три основных класса графем – прописные (Б) и строчные (б) буквы, а также знаки препинания (ЗП).

Хотя представление о прописных и строчных буквах с одним фонетическим значением как о разных графемах и получило некоторое развитие в русистике (ср. [Зализняк 2002: 561, 563-564]), множества прописных и строчных букв при этом все же не принято рассматривать в качестве двух отдельных алфавитов. В русской традиции парные парадигматические единства прописных и строчных букв, фонетически эквивалентных друг другу, считаются более тесными – и первичными по отношению к последующему объединению таких единств, которое и трактуется как единственный алфавит.

В то же время наличие сразу двух сосуществующих алфавитов – хираганы и катаканы – считается редкой особенностью японского письма [Алпатов 2003:

41]. Вероятно, такое расхождение в интерпретациях во многом обусловлено тем, что прописные и строчные буквы воспринимаются как противопоставленные в первую очередь по размеру, т.е. большей частью как геометрические варианты, хотя в русской кириллице имеются и заметные различия в начертании, особенно в курсиве (а в латинице различия еще более значительные). Это сближение даже мотивирует вместо различения сегментных графем оперировать несегментными, как сам признак размера, понятиями типа супрасегментной «капитализации». Напротив, знаки хираганы и катаканы с одним фонетическим значением<sup>5</sup>, за несколькими исключениями, по начертанию мало похожи друг на друга, даже если исторически восходят к одному иероглифу.

Тем не менее, с точки зрения линейной графической структуры текста и тех принципов, по которым она строится, классы прописных и строчных букв принципиально аналогичны катакане и хирагане в японском. В рамках обсуждаемой графической проблематики связь между фонетически эквивалентными прописными и строчными буквами оказывается не такой существенной, как распределение прописных и строчных букв в два класса графем и противопоставление между этими классами. Чтобы подчеркнуть этот аспект, в данном контексте можно даже именовать соответствующие классы графем прописным и строчным алфавитами.

Как известно, прописной алфавит может употребляться изолированно, т.е. не комбинируясь в словах со строчным, для записи целых текстов (всевозможных указателей, объявлений и т.п.), чаще относительно коротких, или их фрагментов (в частности, заголовков, аббревиатур), например, в целях коммуникативного выделения. Строчный алфавит можно считать основным, немаркированным средством, которым записывается большинство слов в тексте, хотя тексты, набранные одним строчным алфавитом, без прописного, по-видимому, не столь распространены (но ср. практику неформальной интернет-коммуникации).

Сочетание прописного и строчного алфавитов в пределах одного слова можно наблюдать главным образом в именах собственных (см. правила [Розенталь 1994: 12-30]) или в первых словах предложений (длиннее одной буквы). В подобных случаях прописная буква строго предшествует строчным, тогда как обратный порядок в русском языке практически нигде не встречается. Очевидно, что в иерархии линейного порядка прописной алфавит занимает первую позицию, а строчный должен следовать за ним: **Б > б**.

Необходимо, конечно, признать: несмотря на глубинное присутствие иерархического принципа, обычно из двух алфавитов русские графические слова действуют только один (скорее строчный), тогда как в японском, как отмечалось выше, в большинстве сегментов представлено по меньшей мере два класса графем (типично – иероглифика и хирагана). В этом смысле русские тексты выглядят намного более графически однородными.

Что касается знаков препинания, то их положение в иерархии определяется в целом аналогично японской пунктуации, хотя состав этого класса графем и их графическое поведение несколько отличается от японского. Как правило, знак

---

<sup>5</sup> Мы даже не можем указать японского термина для такого рода единства; в русском языке, как отметил А.А. Зализняк, в подобной функции используют слово «буква» [Зализняк 2002: 564].

препинания примыкает к предшествующему слову, образуя его «пунктуационное окончание». Соответственно, базовая линейная иерархия, определяющая графические слова русского языка с точки зрения их внутренней структуры, предстает в следующем виде: **Б > б > ЗП**. В первом слове предыдущего предложения можно наблюдать все три ее уровня: *Соответственно*, (**Б > б > ЗП**).

Стоит обратить внимание, что установленная таким образом иерархия указывает на ряд дополнительных границ между графическими словами, которые не проводятся в рамках традиционного подхода, ограниченного пробельным членением. Отметим несколько таких случаев:

- граница после дефиса (который также относится к классу пунктуации): *Ростов-на-Дону, врач-стоматолог-терапевт, кресло-кровать, красно-белый, кто-то*;
- граница после косой черты: *и/или, чай/кофе, сутки/двое, км/ч, б/у* (= бывший в употреблении);
- граница после точки в беспробельных сокращениях: *т.е.* (= то есть), *и.о.* (= исполняющий обязанности), *д.и.н.* (= доктор исторических наук), *у.е.* (= условная единица), *И.П.* (= Иван Петрович);
- граница перед прописной буквой в напоминающих композиты или сокращения образованиях, записанных в т.н. «верблюжем регистре» (англ. CamelCase): *ЭпиВакКорона, СберМегаМаркет, МегаФон, ВкусВилл, АвтоКредитБанк*.

Оставляя детальный разбор каждого из таких случаев за рамками данной статьи, можно сказать, что предлагаемое моделью членение в основном выглядит интуитивно приемлемым и поддерживается теми или иными семантико-синтаксическими соображениями, а вес исключений и спорных случаев едва ли заметно превышает таковой в японском. Как минимум, значительная часть таких конструкций представляет собой свободно образуемые сочетания, отдельные компоненты которых в разных отношениях сближаются со словами. В большинстве своем они представляют известную «серую зону», в пределах которой носители испытывают затруднения при обучении грамотному письму (словоделению) и регулярно ошибаются, вынужденные делать неочевидный выбор, скажем, между слитным, дефисным и раздельным написанием. По-видимому, сама норма письменного словоделения здесь недостаточно устойчива или не вполне сформирована.

Как и в японском языке, такие выражения не разбиваются на части пустым пространством, а сегментируются на основе иерархического противопоставления классов материальных графем<sup>6</sup>. Наблюдаемая вариативность может не только свидетельствовать о происходящих изменениях в синтагматической интеграции компонентов конструкции, но и отражать функциональную близость этих способов членения. Так, для сокращений вида *т.е.* распространена и запись через пробел (*т. е.*) – который на практике может опускаться благодаря тому, что необходимая сегментация успешно обеспечивается оппозицией материальных единиц.

---

<sup>6</sup> По-видимому, в языках типа русского эти способы должны соответствовать разной степени графической интеграции, однако от этого обстоятельства мы в данный момент отвлекаемся.

Таким образом, при внимательном анализе можно обнаружить, что определенные правила (тенденции), задающие линейную иерархию разных классов графем в составе сегментов-слов, действуют и в языках типа русского. В графической организации японского и «европейских» письменных языков есть больше сходства, чем представляют обычно, а различия в рассмотренной области носят скорее количественный, чем качественный характер<sup>7</sup>.

#### 4. Пробел как часть графического слова

До сих пор наше внимание было сосредоточено на поведении материальных единиц. Как следствие, за рамками описания русской (европейской) письменной речи – за пределами иерархически организованных сегментов – фактически был оставлен пробел – пустое пространство, которое принято ассоциировать с особой графемой, пусть не имеющей материального выражения, но занимающей отдельную позицию в знаковой последовательности<sup>8</sup>.

Между тем отличительной чертой развиваемого «японского» подхода является то, что он не предполагает введения каких-то добавочных служебных единиц (граничных маркеров, которые сами не считаются частью сегментов, образующих текст) – сегментация обеспечивается элементами в составе самих сегментов. Тем самым, не возникает и необходимости в специальных правилах обработки этих разделителей – весь сегментный графический состав текста описывается в рамках единой иерархической модели.

Для того, чтобы довести до логического завершения применение указанного принципа и добавить в модель пробел, требуется определить его место в иерархии линейного порядка.

Решение этого вопроса значительно усложняется тем, что пробел, в отличие от материальных знаков, не может быть подвергнут тем же тестам на графическую связность (в терминах [Зверев 2018]), которые позволяют выявить асимметрию графических связей точек или запятых, чтобы обосновать их положение в иерархии. Критерии запрета на перенос, асимметричной конфигурации пустого пространства вокруг знака к пробелу неприменимы.

Очевидно, что таким тестам в принципе можно подвергнуть материальные словоразделители типа латинского интерпункта и по аналогии распространить решение на пробел, однако достаточным материалом для таких заключений мы не располагаем. Современный пробел может рассматриваться если не как прямой этимологический потомок, то как «продолжатель традиции» таких материальных разделителей в европейском письме – ср. предположения о переходе от сплошных вертикальных черт к пунктирным линиям из трех, двух точек и, наконец, к точкам-интерпунктам в монографии [Wingo 2011: 14-16].

---

<sup>7</sup> Ср. в этом плане идею об использовании в европейских языках, как в японском, смешанного фонетико-иероглифического письма – в силу присутствия иероглифов типа арабских цифр [Алпатов 2003: 41].

<sup>8</sup> Пустое пространство при знаках препинания в беспробельных японских текстах (как после запятой в примерах (3), (4)) понимается как увеличенный интервал, ассоциированный с соответствующим знаком, а не выделяется в самостоятельную графему; можно считать это пустое пространство частью знака препинания.



Развивая аналогические соображения, необходимо также принять в расчет нередкое – и небезосновательное – сближение пробела с пунктуацией в целом и возможное включение его в пунктуационную периферию (ср. трактовку пробела у некоторых авторов в обзоре [Петриченко 2006: 41–45]). По аналогии с основной пунктуацией, рассмотренной нами выше, пробел должен связываться с концом иерархии. Если пробел относится к пунктуации, то его взаимное расположение с другими (материальными) пунктуационными знаками регулируется иерархическими правилами внутри класса знаков препинания – наподобие того, что определяет размещение закрывающей кавычки строго после запятой или точки в американском английском. Например, правила этой группы требуют, чтобы в печатном тексте пробел следовал за запятой, и запрещают обратный порядок.

Можно также утверждать, что само существование множества правил, которые описывают взаимодействие пробела с «пунктуационными окончаниями» графических слов (точками, запятыми и другими знаками, графическая связь которых с преконтекстом не вызывает сомнений), синтагматически связывает пробел именно с этими окончаниями, а значит, указывает на его ассоциацию с концом графической иерархии<sup>9</sup>. Отметим, что это наблюдение имеет силу независимо от того, признается ли в таких правилах сам пробел знаком препинания.

Другой аргумент, который можно было бы назвать «семиотическим», обращается к статусу пробела как своего рода нулевого сигнала: пустое пространство может нести информацию только на фоне собственно текста, т.е. в случае, если при нем имеются и какие-то материальные графические сигналы. Поэтому необходимым синтагматическим условием для информативности пробела (приписывания ему какой-то значимости) является считывание/порождение какого-либо материального текста; в ситуации преимущественного продвижения по строке слева направо таковым в первую очередь выступает преконтекст, расположенный слева от пробела<sup>10</sup>. Такой порядок также говорит в пользу привязывания пробела к концу графического слова, соответственно, его размещения в конце иерархии.

Все приведенные выше соображения так или иначе указывают на более тесную связь пробела с концами графических слов и его место в конце европейской иерархии, причем каких-либо убедительных свидетельств в пользу обратного мы предложить не готовы. В дополнение к этим наблюдениям, как кажется, следует обратить внимание еще на одно обстоятельство – совсем другого рода, связанное с качественной интерпретацией установленных иерархий.

---

<sup>9</sup> Редким исключением выглядит положение пробела относительно открывающих парных знаков (скобок, кавычек), аффилированных с постконтекстом, но правило для таких случаев, видимо, является только вторичным отражением правила для закрывающих пар и может даже не оперировать позициями сегментных знаков, а отсылать к симметрии как таковой.

<sup>10</sup> Приоритет преконтекста в принципиальном наделении значимостью не отрицает важности постконтекста, который позволяет уточнить размер пустого промежутка и его значение, например, отличить межсловный пробел от двойного пробела между предложениями в английской печати.

## 5. Качественная интерпретация иерархий

Если лингвистические представления о графеме формируются во многом с опорой на более разработанное понятие фонемы (ср. [Зализняк 2002: 559-561]), то для описанных иерархий линейного порядка ближайшей аналогией из области фонетики могут служить развитые подходы к моделированию структуры слога и такие результаты, как известный закон восходящей звучности (см. о таких результатах [Князев 2000: 72-74]). Весьма существенно, что указанный закон не просто задает иерархию классов фонем, фиксируя наблюдаемые закономерности в расположении сегментных компонентов слога. Этому порядку дается определенная качественная интерпретация, более того – закон в первую очередь позиционируется как описание тенденции в распределении некоторого признака (звучности, сонорности), и каждый класс фонем получает оценку по соответствующему параметру.

Эта аналогия заставляет поставить вопрос о качественной (признаковой) интерпретации выявленных графических иерархий. Она не обязательно должна ориентироваться на объяснение каких-то исторических изменений (типа появления заглавных букв или упомянутого перехода от поперечной черты к интерпункту по [Wingo 2011: 14-16]), а может строиться как описание чисто синхронной данности безотносительно к происхождению. Такое описание может быть задействовано в экспериментальном изучении процессов чтения/письма или в работе по оптимизирующему конструированию письменного языка – подобно исследованиям, оценивающим визуальную сложность знаков в разных письменностях (ср. [Chang L. et al. 2018]).

Кроме того, изучение этого вопроса может не замыкаться в пределах графики, но обращаться к качествам разного плана, связанным с любыми аспектами функционирования письменных единиц. Однако такая всесторонняя оценка является слишком масштабной задачей, и в данной статье мы ограничимся только поверхностно-графическими свойствами, которые могут иметь значение для низкоуровневой обработки визуального сигнала.

Для того чтобы интерпретировать иерархические оппозиции графических классов в терминах распределения признаков, требуется содержательно охарактеризовать графемы этих классов с точки зрения их существенных визуально-графических особенностей. Присутствие в иерархиях более двух уровней означает градуальность оппозиции, так что рассматриваемые признаки должны практически допускать несколько степеней интенсивности проявления.

По всей видимости, общепринятые представления о русской и японской графике в значительной мере уже содержат самую основную информацию, позволяющую предложить некоторое приемлемое решение этой задачи.

Как уже было отмечено, в русском языке оппозиция прописных и строчных букв воспринимается главным образом как различие по размеру (прежде всего высоте). Важность этого признака подчеркивается тем, что выражения «большие буквы» и «маленькие буквы» даже утвердились в качестве альтернативных обозначений для классов, соответственно, прописных и строчных букв [Ахманова 1969: 69].

Японские иероглифы характеризуют как знаки сложные, темные и угловатые, знаки хираганы – как простые и округлые [Taylor, Taylor 1983: 70; Маевский 2000: 121; Kessler, Treiman 2015: 14]. Дополняя эти характеристики, можно описать знаки катаканы как простые и угловатые, а также прямо назвать знаки обеих слоговых азбук светлыми – по аналогии с «темной» иероглификой. Все эти визуально-графические особенности объясняются тем, что знаки каны происходят из упрощенной записи иероглифов, при этом катакана была получена путем выделения из иероглифов отдельных частей, а хирагана – из скорописного написания целых иероглифов [Frellesvig 2010: 158-160].

Изложенные сведения, как кажется, уже позволяют выделить для рассмотрения некоторый параметр, причем общий для обеих графических систем. Этот параметр можно назвать *количеством графического материала*, определив его как совокупную площадь линий, точек и других элементов знака (образованного ими «пятна») в зрительном восприятии читающего; с точки зрения пишущего, в неформальном выражении, это количество чернил, необходимое для записи знака.

Наблюдаемые тенденции таковы, что темные, сложные иероглифы обычно превосходят по этому параметру светлую, простую кану (классы хираганы и катаканы в этом отношении примерно равны), а большие прописные буквы – маленькие строчные. Увеличение размера или заполненности (прямоугольного блока, в который вписан знак) можно расценивать как два альтернативных пути к увеличению количества графического материала в составе знака, реализованные, соответственно, в русском и японском письме.

Что касается пунктуации, то малое количество графического материала уже было отмечено нами в качестве одной из ключевых характеристик этого класса языковых единиц: прототипические знаки препинания – это мелкие значки [Зверев 2016]. Этот признак особенно ярко выражен у самых частотных знаков – точек и запятых, составляющих прототипическое ядро пунктуации, как в русском, так и в японском языках.

Таким образом, интерпретируя выведенные в предшествующих разделах иерархии классов графем (**И/К**) > **X** > **ЗП** – в японском, **Б** > **б** > **ЗП** – в русском), для обоих языков можно констатировать тенденцию к убыванию или, более точно, невозрастанию количества графического материала в составе сегментных компонентов от начала к концу графического слова.

Среди других параметров, которые можно было бы также предложить для рассмотрения (например, высота знака<sup>11</sup>), количество графического материала выделяется своей простотой и, в некотором смысле, фундаментальностью. Это базовое свойство противопоставляет текст его отсутствию – даже пустое пространство может быть естественным образом охарактеризовано по данному параметру. Это обеспечивает применимость оценки равно к японскому и русскому письму.

Можно заметить, что пробел тоже вписывается в указанную тенденцию, если рассматривать этот пустой сегмент как завершающую часть графического

---

<sup>11</sup> Вероятно, в русском высота (без выносных элементов) может считаться своего рода доминантой признака количества графического материала; подобные нюансы требуют дополнительной аккуратной проработки.

слова. Пробелу соответствует предельно малое – нулевое – количество графического материала. Такая интерпретация согласуется с перечисленными ранее изображениями, которые с разных сторон указывают на приоритетную ассоциацию пробелов с концами графических слов.

Принимая во внимание уникальные графические свойства пробела, которые во многом и обусловили трудность его анализа, можно выделить эту графему в качестве четвертого уровня русской (европейской) иерархии: **Б > б > ЗП > п**.

Применение на письме пробела – единицы с предельно возможным (нулевым) значением признака – повышает эффективность использования доступного диапазона значений, что, вероятно, и находит отражение в том, что в европейских языках типа русского можно видеть не три, как в японском, а четыре уровня иерархии. Но вообще число таких уровней едва ли может быть слишком велико – ср. в этом плане ограничения, скажем, на количество подъемов гласных<sup>12</sup>.

В свете обозначенной тенденции размещение пробела в конце графического слова вполне органично противостоит употреблению прописной буквы в его начале. Однако пробел, по-видимому, едва ли мыслится как «предельное конечное сокращение количества графического материала», тогда как практика начальной капитализации в известной мере осознается как увеличение в размерах начальной части слова (Его Первой Буквы). И распространенные представления о капитализации в целом как о супraseгментном явлении могут быть частично соотнесены с общей идеей распределения признака количества графического материала по сегментным компонентам слова.

Вхождение пробела в состав соседних графических слов переворачивает картину, сформированную нами прежде: получается, что в русском практически во всех графических словах представлено как минимум два уровня иерархии (класса графем). В такой необычной перспективе русский текст предстает даже графически неоднороднее японского. Такая иерархическая модель, включающая пробелы, охватывает всю существенную (в плане словоделения) графическую неоднородность европейских текстов, раскрывая общепринятое суждение о том, что в европейской письменности, вообще говоря, деление на слова выражено более отчетливо, чем сегментация подобного уровня в японском.

Предложенная качественная интерпретация позволяет переформулировать выявленные ранее закономерности (тенденции) в области русского и японского словоделения в единой – и при этом удивительно простой – признаковой форме: в обоих языках новый сегмент обычно начинается с увеличением количества графического материала. С точки зрения фундаментальных принципов построения графической структуры, это обстоятельство может расцениваться как одно из проявлений рекурсивности в организации текста: одно и то же графическое явление отмечает не только начало всего текста в целом, но и начало его структурных компонентов определенного уровня членения.

---

<sup>12</sup> Доступные значения признака должны ограничиваться и сверху, хотя бы в связи с ростом ресурсоемкости слишком крупных и темных единиц. Возможный путь обхода таких ограничений – использование оппозиции по другому параметру (например, цвету), с формированием иерархии нового уровня.

Обе описанные иерархии строятся как механизм более плавного (ступенчатого) понижения количества графического материала в текстовой последовательности (в направлении чтения/письма), в то время как его увеличение, ассоциированное с границами сегментов, в среднем должно быть более резким. Такая асимметрия должна способствовать более четкому, контрастному выделению границ. Впрочем, практическая значимость этого следствия сильно зависит от особенностей конкретной реализации иерархий в письменной речи. Так, графическая структура значительных фрагментов русских текстов сводится к чередованию пробелов и (групп) строчных букв, т.е. всего двух классов графем, а в таких условиях указанный фактор действовать не может. Но, скажем, в немецких текстах, которые изобилуют прописными буквами в начале существительных, роль этого фактора, вероятно, выше.

Переосмысливая известные межъязыковые различия в рамках развиваемой модели, можно сказать, что в японском языке пространство между несущими основную информацию знаками – иероглифами – не остается пустым, подобно пробелам между буквенными словами европейских текстов (хотя и там в промежутках встречаются знаки препинания), а заполняется материальными графемами другого класса – преимущественно уточняющей грамматику и в среднем более светлой хираганой. Как и пробелы, хирагана характеризуется пониженным количеством графического материала – сегментирующая «пустота», на которую привык ориентироваться современный европейский читатель, может быть относительной.

В то же время наполненный пробелами европейский текст больше уподобляется японскому и предстает, по крайней мере в пределах строки, *непрерывной* последовательностью графем и образованных ими более крупных единиц (графических слов), выделение которых обусловлено не «разрывами» в речевом потоке, а только неоднородностью в распределении признака по сегментным графемам. В таком представлении письменная речь становится несколько ближе к устной, имеющей более «континуальную» репутацию.

Словоделение в письменных языках «европейского» образца обычно понимают как сегментацию пробелами, но, видимо, более точным было бы утверждать, что в них, как и в японском, в этой задаче задействованы графемы разных классов, представленных – и противопоставленных друг другу – в тексте. Пробел – выделяющийся элемент с крайним значением признака, но он может рассматриваться и как часть более сложной графической структуры. В этом смысле его статус сопоставим с положением ударных слогов в устной речи, позиция которых тоже может быть жестко привязана к границам слов.

## 6. Заключение

Предложенная модель графического слова, основанная на иерархиях линейного порядка, вероятно, не должна претендовать на равную ценность для анализа любых письменных языков, но представляется заведомо более гибким и полным решением, чем широко применяемый подход, который обращает внимание только на наличие пробелов.

Отталкиваясь от распространенных описаний японской графики, мы по-иному расставили акценты и перешли к формулировке правил в иерархических терминах, добавив в развиваемую модель пунктуацию. В остальном изложенные в соответствующем разделе сведения сами по себе хорошо известны японистам. Однако это знание практически не выходит за пределы японистики и не используется должным образом при обсуждении фундаментальной проблематики графического слова (редкое обращение к японским реалиям находим в работе [Шалапина 2007: 68-69]). Вместо этого исследователи предпочитают ограничиваться поверхностной оценкой японского языка по «среднеевропейскому стандарту»: есть ли в нем межсловные пробелы? Случаев применения обратного подхода – переноса японского опыта на европейский материал – нам вовсе неизвестно.

Между тем, как показано в настоящей статье, упорядоченное сочетание разных классов графем в одном тексте не является уникальной особенностью японского языка. Европейские письменные языки тоже имеют довольно сложную графическую структуру и задействуют похожие стратегии графической организации – обстоятельство, которое совершенно упускается из вида в традиционных «пробелоцентричных» описаниях.

Иерархический подход, мотивированный японским материалом, предлагает в новом свете увидеть привычное устройство европейских текстов. Такие нетривиальные, на первый взгляд, идеи, как сосуществование в русском языке двух алфавитов или принадлежность пробела предшествующему графическому слову, получают свое обоснование и вполне естественно вписываются в общую концепцию.

Но главное, что раскрывает такой подход, – это само наличие у графических слов внутренней графической структуры. В рассмотренных языках эта структура имеет линейно-иерархическую природу и, видимо, отличается особой чувствительностью к такому поверхностно-графическому параметру, как количество графического материала. Впрочем, не приходится ожидать, чтобы описанная тенденция к убыванию (невозрастанию) количества графического материала в пределах сегмента-слова одинаково проявлялась во всех языках – в этом отношении она едва ли выделяется на фоне сопоставимых фонетических феноменов. В целом вопрос о качественной интерпретации выявленных иерархий требует более тщательной разработки, а роль указанного признака – аккуратной проверки.

Таким образом, «словесный» уровень графической структуры предстает в виде последовательности сегментов, которые строятся из графем по единой схеме – в соответствии с иерархией линейного порядка, свойственной данному языку. Подобно сведениям об использовании тех или иных знаков письменности (иероглифов, букв определенных алфавитов, пунктуации), эта иерархия оказывается важным элементом знаний о графической системе конкретного языка и наравне с другими ее характеристиками может быть объектом разностороннего изучения.

### Грамматические сокращения

ACC – аккузатив  
CNV – деепричастие  
IMP – императив

INS – инструменталис  
PST – прошедшее время  
TOP – топик

## Литература

1. Алпатов В.М. *Япония. Язык и общество*. М.: Муравей, 2003.
2. Алпатов В.М., Аркадьев П.М., Подлесская В.И. *Теоретическая грамматика японского языка*. М.: Наталис, 2008. Том 1.
3. Алпатов В.М. *Слово и части речи*. М.: ЯСК, 2018.
4. Ахманова О.С. *Словарь лингвистических терминов*. М.: Сов. Энциклопедия, 1969.
5. Данилов А.Ю., Сыромятников Н.А. *Японский язык. Пунктуация, знаки повтора, вспомогательные пометы*. М.: Муравей-Гайд, 2001.
6. Зализняк А.А. *Русское именное словоизменение*. М.: Наука, 1967.
7. Зализняк А.А. О понятии графемы // *Русское именное словоизменение с приложением избранных работ по современному русскому языку и общему языкознанию*. М., 2002. С. 559-576.
8. Зверев А.С. Системоцентризм и антропоцентризм в изучении пунктуации: в поисках объективных критериев выделения знаков препинания // *Проблемы общей и востоковедной лингвистики 2015. Язык. Общество. История науки. К 70-летию члена-корреспондента РАН В.М. Алпатова. Труды научной конференции. Институт востоковедения РАН 22-23 апреля 2015 г.* Т.2. М.: ИВ РАН, 2016. С. 136–144.
9. Зверев А.С. Графическая связность: линейно-синтагматический статус японских пунктуационных показателей // *Вестн. Том. гос. ун-та*. Томск, 2018. №426. С. 27-38.
10. Князев С.В. О причинах некоторых звуковых изменений в праславянском языке // *Вестник МГУ. Сер. 9. Филология*. М.: 2000, №5. С. 72-97.
11. Маевский Е.В. *Графическая стилистика японского языка*. М.: Муравей-Гайд, 2000.
12. Мельчук И.А. *Курс общей морфологии*. М.: Прогресс, ЯРК; Вена: WSA, 1997. Том 1.
13. Петриченко М.А. *Очерки по пунктуации: Учебное пособие*. Кривой Рог: Видавничий дім, 2003.
14. Розенталь Д.Э., Джанджакова Е.В., Кабанова Н.П. *Справочник по правописанию, произношению, литературному редактированию*. М.: Московская международная школа переводчиков, 1994.
15. Шаляпина З.М. *Трехмерная стратификационная модель языка и его функционирования: к общей теории лингвистических моделей*. М.: Вост. лит., 2007.
16. Яхонтов С.Е. О значении термина «слово» // *Морфологическая структура слова в языках различных типов*. М.-Л., 1963. С. 165–173.
17. Chang L., Chen Y., Perfetti Ch.A. GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages // *Behavior Research Methods*. V50. 2018. P. 427–449.
18. Frellesvig B. *A History of the Japanese Language*. New York: Cambridge University Press, 2010.
19. Haspelmath M. The Indeterminacy of Word Segmentation and the Nature of Morphology and Syntax // *Folia Linguistica*. 2011. Vol.45. №1.

20. Kessler B., Treiman R. Writing Systems: Their Properties and Implications for Reading // Pollatsek A., Treiman R., eds., *The Oxford Handbook of Reading*. New York: Oxford University Press, 2015. P. 10–25.
21. Spencer A., Luís A.R. *Clitics. An Introduction*. New York: Cambridge University Press, 2012.
22. Tanaka T. Statistical Analysis of Japanese Characters // *COLING-80. Proceedings of the 8th International Conference on Computational Linguistics*. Tokyo, 1980. P. 315–322.
23. Taylor I., Taylor M.M. *The Psychology of Reading*. London: Academic Press, 1983.
24. Wingo E.O. *Latin Punctuation in the Classical Age*. The Hague: Mouton, 1972.
25. 水谷静夫 [Mizutani Shizuo] et al. 文字・表記と語構成 [*Moji hyooki to gokoosei* {Знаки, письмо и структура слова}]. Токио: Асакура, 1987



## Модель словоформы в чадских языках

**Ольга Валерьевна Столбова**

Институт востоковедения РАН, Москва, Россия,  
stolbova.o@yandex.ru, ORCID 0000-0002-7589-2854

**Аннотация.** В статье анализируются процессы, действующие в пределах словоформы типа  $C_1VC_2V$  в языках чадской ветви афразийской макросемьи. Высказано предположение, что для чадских языков актуальным является не сочетаемость согласных в пределах одной и той же словоформы (что существенно для семитских языков), а уровень их сонорности. Модель словоформы типа  $C_1VC_2V$  определяется следующим образом:  $C_1$  не может быть выше  $C_2$  по сонорности. На материале чадских языков, с привлечением внешних параллелей показано, что метатеза, ассимиляция, диссимиляция и лениция имеют одну цель – привести словоформу в соответствие с моделью. Проанализированный лексический материал подтверждает, что модель словоформы является чадской инновацией, возникшей в период распада чадской семьи.

**Ключевые слова:** словоформа, афразийские языки, чадские языки, сонорность, метатеза, ассимиляция, диссимиляция, лениция

**Abstract.** This paper discusses some processes typical of  $C_1VC_2V$  lexical forms in Chadic languages. It is claimed that metathesis, assimilation, dissimilation and lenition are aimed to bring lexical forms in accordance with the pattern. The latter in its turn follows the law of rising sonority, meaning that on the sonority scale  $C_2$  must be higher than (or at least, equal to)  $C_1$ . Viewing that for Semitic languages the compatibility of consonants is valid, the ‘law of rising sonority’ may well be a Chadic innovation. The supporting evidence includes Proto-Chadic lexical reconstructions, as well as isolated Chadic forms and numerous cognates in other Afro-Asiatic languages (mostly – in Semitic).

**Keywords:** Afro-Asiatic languages, Chadic Languages, sonority, metathesis, assimilation, dissimilation, lenition

Многие процессы, происходящие в чадских языках<sup>1</sup> на уровне словоформы вида  $C_1VC_2V$  (метатеза, ассимиляция, лениция, диссимиляция) вызваны необходимостью соответствовать одной и той же модели. Параметры модели определяются следующим правилом: в пределах одной словоформы менее звучная фонема должна предшествовать более звучной, но не наоборот. Допускается равенство  $C_1$  и  $C_2$  по звучности. Оценочная шкала для согласных выглядит следующим образом: фрикативные → аффрикаты → смычные → имплозивные → носовые → плавные.

<sup>1</sup> Семитские, древнеегипетский, берберские, кушитские, омотские и чадские языки образуют афразийскую макросемью.

Существенно отметить, что для наиболее изученных языков афразийской семьи – семитских – актуально правило сочетаемости согласных в позициях  $C_1$  и  $C_2$ . Эта проблема была исследована на материале нескольких семитских языков: арабского [Greenberg 1950], древнееврейского [Koskinen 1974], геез [Жвания 1972]. Как отметил Л. Коган: «На материале 3775 арабских корней Гринберг показал, что одинаковые или близкие по месту образования согласные не встречаются в составе корня в позиции первого и второго согласного» [Коган 2009, с. 46]. Однако ни в семитских, ни в других афразийских языках не было обнаружено никакой связи между комбинацией согласных и шкалой сонорности.

Поэтому, с одной стороны, можно предполагать, что анализируемая модель словоформы является чадской инновацией. Важно уточнить, что все описываемые ниже процессы не происходят на прачадском уровне, а затрагивают отдельные группы и даже языки внутри небольших подгрупп<sup>2</sup>.

С другой стороны, очевидно, что словоформы, унаследованные чадскими языками от праафразийского состояния, могут соответствовать модели<sup>3</sup>, а могут противоречить ей. Именно те словоформы, в которых  $C_1$  превышает по звучности  $C_2$ , наиболее часто претерпевают изменения, приводящие их в соответствие с моделью.

Самым простым способом достижения искомой цели является метатеза:  $C_1 - C_2 > C_2 - C_1$ .

Первый блок примеров (1.1 – 1.10) включает не соответствующие модели словоформы в АА языках и их чадские когнаты двух типов: с исходным порядком следования фонем ( $C_1 - C_2$ ) и с метатезой ( $C_2 - C_1$ ). Названия АА языков даются в латинице с заглавной буквы. Рефлексы в чадских языках приводятся в следующем порядке: западно-, центрально-, восточночадские.

1.1. Сем. \**lis-ān*- ‘язык’ [SED I №181]. На прачадском уровне и в подавляющем большинстве чадских языков сохраняется исходная структура: чад. \**lis-um*- ‘язык’: зап. Montol, Goemai, Mwaghavul *liis*, Bolewa *lisi-m*, Karekare *lūsə-m*, вост. Dangla *léésé*, Jegu *lèésó*, Birgit *liisi*, Mubi *lèésí*, Kwang *ki-listi*, *ke-les-úm<sup>4</sup>*, и т.д. Однако наблюдаются и несколько форм с метатезой: центр. Dzepaw *silé*, Peve *šele*; Masa *sinná* (< *siln* < *sil-m*-), вост. Kera *kə-səl* [CED №568]. Заметим, что Kera и Kwang (см. выше) – это близкородственные языки, они образуют небольшую группу восточной ветви. Тем не менее, в языке Kwang сохраняется исходная структура, а в языке Kera произошла метатеза.

1.2. Сем. Arabic *fazz-<sup>5</sup>* ‘лошадиная моча или жидкость из желудка верблюда’, берб. Ahagar *tafezza* ‘моча’, Air *tafezze*, Ghat *čifezza* ‘моча верблюда’ (арабизм) [AA №11]. Та же структура реконструируется на прачадском уровне \**pič’-r-<sup>6</sup>* ‘моча’ и сохраняется в нескольких чадских языках: зап. Hausa *fitsaa-rii*, вост. Dangla *pidyí*,

<sup>2</sup> О классификации чадских языков и истории вопроса см. [CED: 23-25].

<sup>3</sup> Это, например, афразийские корни вида \**p-r(r)*, \**p-l(l)*, см. [AA №№27-38].

<sup>4</sup> Соматическая лексика в чадских языках может маркироваться префиксами *k-* или *t-*, а также посессивным суффиксом *-m-*.

<sup>5</sup> С помощью знака *z* (в арабской графике – *ẓ*) обозначается рефлекс АА эмфатической палатальной аффрикаты *č’*.

<sup>6</sup> *-r-* – суффикс мн.ч. имен существительных.

Migama *piddɔ̀yì*, Birgit *fidɔ̀yì*. В подавляющем большинстве чадских языков произошла метатеза: \**pič* '-r- > \**č* 'Vp-r-: зап. Siri *č* 'ipá-ri, Warji *səpr-ái*, Kariya *ci-pí-r*, Miya *səpú-r*, Paa *ci-pù-rá*, центр. Dzeraw *čubūr*, *zúbūr*, Dari *zábūr*, вост. Kaba *čəbərə́*, Nanchere *čəbrə́*, Mawa *swari*, Ubi *čario*, Mabire *čora* (< *čVwVr* < *čVpar*). Метатеза произошла и в производном глаголе: зап. Tangale *sipe* 'мочиться'. Весь релевантный чадский материал представлен в [CED №609].

1.3. Сем. Arabic *rhd* 'растирать, дробить; froter, fractionner' [BK I: 936]. Исходная структура сохраняется в двух чадских языках: зап. Hausa *rèedaa*<sup>7</sup>, вост. Kwang *érdē* 'молоть'. Однако гораздо лучше представлены формы с метатезой: зап. Karekare *diiru* 'молоть', Bolewa *dir* 'молотить'; Ngamo *diirā* 'толочь, молотить', Warji, Kariya, Miya *dər*, Sagu *dar*, Jimbin *dira* 'молоть', центр. Musey *dura* 'толочь зерно'. К этому же корню восходят производные: зап. Sagu *iifə deren* 'жернов' (*iifə* 'камень'), Buli *adir* 'мука', центр. Ouldem *-dərə-ɬ* 'растирать'.

1.4. Сем. Geez *nafaya*, Akkadian *napū* 'просеивать', Geez *ma-nfē*, Akkadian *nappū* 'сито' [LSG: 390]. В чадских языках представлены только формы с метатезой: чад. \**pVn-* 'дуть, веять': зап. Mushere *pun* 'веять; сито', вост. Lele *piñy* 'дуть, веять', Dangla *pánné* 'дуть'.

1.5. Сем. Arabic *nkl*<sup>9</sup> 'выкапывать, копать' [BK II: 1320]. В центр.-чад. языках сохранилась исходная структура: Mofu *-kəč-* 'копать, выкапывать', Daba *k<sup>w</sup>áč* 'делая борозды', а в западной ветви произошла метатеза: Karekare *čak-*, Diri *čəkà*, Sagu *čək-* 'копать' [CLD II №848].

1.6. Сем. Arabic *bhs* 'уменьшать(ся), ослаблять' [BK I: 90]. В чадских языках засвидетельствованы только формы с метатезой: зап. Hausa *sábè* 'терять вес'; Tangale *saabe* 'становиться худым', центр. Mofu *-səbhw-* 'уменьшаться', Dzeraw *sésébé* 'худой', вост. Mawa *səba-ŋ* 'худеть' [CLD III №59].

1.7. Сем. Arab *kss* 'разбить, растолочь; ударить' [BK II: 894]. Непосредственные когнаты: зап. чад. Hausa *kíši* 'бить, колотить', Ngas *k<sup>w</sup>as* 'бить, ударять'; вост. Mokilko *kóssè* 'бить, ломать', *kókkisò* 'толочь'. Формы с метатезой: зап. Hausa *síssikaa* (< \**sik-sika*), *sússúkaa* 'молоть (в ступке)', Kirfi *šuyú-du* 'молотьба', центр. Munjuk *sukí* 'давить', вост. Mokilko *sókké* 'придавить' [CLD III №176].

1.8. Вост.-куш. Arbore *pača*, Tsamai *hoči* 'полный' [АТО: 9]. Исходная структура сохранилась в центр. ветви чадской семьи – Mbuko *pečereče* 'полный до краев', а форма с метатезой – в зап., а именно: Bade *čippà* 'быть полным'.

1.9. Сем. Arabic *bwk* 'лепить шары, катая куски глины между ладонями; former des boules d'argile en la roulant entre les deux mains' [BK I: 180], куш. Burji *bukk-ays-* 'месить'. Непосредственный рефлекс: зап.-чад. Курто *mbèk-téy* 'месить, перемешивать с помощью палки', центр. Mofu *-mbá-mbák<sup>w</sup>-* 'придавать куску глины шаровидную форму'<sup>10</sup>, Azumeina *mbukà* 'месить (тесто, глину)'. Формы с метатезой: зап. Hausa *k<sup>w</sup>áábà* 'месить глину', центр. Bura *kaba* 'придавать куску глины

<sup>7</sup> Список словарей и иных источников по чадским языкам см. в [CED, с. 6-18].

<sup>8</sup> С помощью символов *l* и *ʃ* обозначаются латеральные фрикативные фонемы (глухая и звонкая соответственно).

<sup>9</sup> Семитский интерденальный сибилант (*l*) восходит к АА палатальной аффрикате \**č*.

<sup>10</sup> Жители Нигерии и в настоящее время строят небольшие хижины не из кирпичей, а из шаровидных комков сырой глины. Именно подготовительная работа и описываются глаголами *b-k* / *k-b*.

шаровидную форму», Gude *kabə* ‘формовать круглые шары из строительной глины’.

1.10. Сем. *ʔVnas-* ‘человек, человечество; man, mankind’: Hebrew *ʔānōš*, Aramaic *\*ʔVnāšā* ‘человек, человечество’, Tigre *ʔanas* ‘человек’ [PSL: 232], Ugaritic *nš-m* ‘люди, человек’ [DUL: 640]. Возможной метатезой к сем. являются формы в центр.-чад. языках (группа маса), а именно: Musey *sana* ‘человек, мужчина’, Kulung *sánan* ‘человек’, Azumeina *sáná, sá* (стяжение из *sana*) ‘человек (мужчина)’. Сопоставление не вполне надежно, поскольку в некоторых публикациях  $C_2$  в языках той же группы (достаточно непоследовательно) трактуется то как корневой согласный, то как детерминатив<sup>11</sup>.

1.11. В следующем случае на исходный порядок согласных указывает египетская параллель, а именно: *hbs* ‘одевать, покрывать, заворачивать’, *hbs* ‘одежда, покрывало, ткань’ [EG: 64-66]. Исходная структура сохранилась в центр.-чад. языках: Gisiga *basay* ‘женская набедренная повязка’, Merey *bas-kwar* ‘набедренная повязка’, Gidar *buse* ‘ткань’, Daba *bās* ‘чехол, тент’, Lamang *mbusa* ‘покрывать, укрывать’ (производный глагол). Представлены и формы с метатезой: зап. Bolewa *sùbà*, Ngamo *sùbà* ‘одежда’, центр. Bachama *sùbwé-to* ‘одежда’, Mada *sbè-t*<sup>12</sup> ‘завернуть полностью в ткань’ [CLD III №51].

Приведенные примеры подтверждают гипотезу о «направленной» метатезе в чадских языках, цель которой – привести словоформу в соответствие с моделью.

Точно такая же трансформация ( $C_1 - C_2 > C_2 - C_1$ ) наблюдается и в следующей серии чадских словоформ (2.1 – 2.9.). Внешние параллели к примерам из этой серии пока не удалось обнаружить, однако есть все основания полагать, что и в этих случаях метатеза использовалась для устранения главного несоответствия модели –  $C_1$  выше  $C_2$  по сонорности. Исходная словоформа и ее предполагаемая метатеза приведены под а. и б. соответственно. По нашему предположению, формы под а. и б. восходят к одному и тому же прачадскому корню.

2.1а. Чад. *\*bVk-* ‘укрывать, крышка’: зап. Bolewa *bàakè* ‘плетеный поднос, крышка’, Duwai *biko* ‘крышка для горшка’, Раа *bakù-sù* ‘заворачивать (например, в пальмовые листья)»; центр. Daba *mbūk* ‘укрыть(ся)’, Fali Kiriya *bəkə* ‘чехол»; Dzeraw *bòk* ‘оболочка’, Masa *būka-d* ‘набедренная повязка’.

2.1б. Чад. *\*kVb-* ‘покрывать, крышка’: зап. Hausa *kube*; Ngas *kiiip* ‘чехол, оболочка’, Vade *kùbi* ‘покрывать»; центр. Bura *kubaha* ‘обернуть’, Mofu *-kub-* ‘покрывать’, Gizey *kúb* ‘закрыть покрывшей’, Даги *kūb* ‘крышка»; вост. Kera *kúbí* ‘покрывать’.

2.2а. Центр.-чад. *\*bVç-* ‘давить, топтать’: центр. Vana *básá* ‘давить, сжимать’, Pod *bičə* ‘сдавливать’, Lamang *əbica*, Vemgo *bicu* ‘сжимать, выжимать’, Lamang *bica* ‘отжимать (через сито)’, Mafa *mbác-* ‘топтать’, Muayang *ámàc* ‘раздавить’.

2.2б. Чад. *\*çVb-* ‘толочь, давить’: центр. Malgwa *səba* ‘молотить, толочь’, Mafa *cúb-* ‘кастрировать домашнее животное; to castrate by pounding’, зап. Hausa *cáabà* ‘обмолоченное зерно’, *cáabà-čee* ‘молотить (зерно)’ [CLD III №656].

<sup>11</sup> Например, Masa *sà* ‘l’homme’ [CC: 133], но в обратном словаре: ‘homme’ [sàná] [CC: 179].

<sup>12</sup> *-t* – суффикс комплетива.

2.3а. Центр. чад. (группа маса): Gizey, Masa, Ham gès ‘грязь, отбросы’, Musey gossa ‘пачкать, грязь’.

2.3б. Центр. Fali Kiriya *šig<sup>w</sup>a* ‘грязный’; вост. Lele *sùgè* ‘грязь, отбросы’.

2.4а. Центр. Mafa *dat-* ‘разрубать’, Muyang *dāt* ‘резать’. Надежная параллель есть в южно-куш. языке: Dahalo *dat-* ‘раскалывать, рубать’ [EEN: 31].

2.4б. Чад. \**IVd-* ‘рубить, резать’: зап. Buli *tòdà-n* ‘резать’, центр. Dghwede *tòdayà* ‘рубить, нарезать’, Hide *taday* ‘срубить дерево’, Mafa *hid-* ‘обрубать ветки дерева’.

2.5а. Центр. группа хиги Fali Kiriya *daŋi* ‘белый’.

2.5б. В той же группе языков представлены формы с метатезой: Higi Ghya *tì-βidi* ‘белый, свет’, Higi Futu *bir-βide* ‘свет’.

2.6а. Чад. \**kVl-*: зап. Murun, Mwaghavul, Goemai *kus* ‘тереть, растирать’, Goemai *k<sup>w</sup>as* ‘натирать, полировать’, центр. Ouldem *kulay* ‘тереть, растирать’.

2.6б. Зап. \**IVk-* ‘тереть, растирать’: Pero *čàkù*; Geji *tòkà*.

2.7а. центр. \**baŋ-* ‘убить’: Margi *βŋà* ‘убить (многих)’, Fali Kiriya *baŋà* ‘совершить массовое убийство, разрушить’, Meye *baβàŋ* ‘истребить’, Muyang *abàŋ* ‘убить (многих)’, *baŋabàŋà* ‘насилие’, Mafa *baŋ-* ‘быть уничтоженными’.

2.7б. Bura *ŋabu* ‘убить’ (метатеза засвидетельствована лишь в одном языке).

2.8а. Чад. \**pat-* ‘плоский’: зап. Ngizim *pátà* ‘макушка’, Duwai *patpát* ‘ладонь, flat of the hand’ (т.е., ‘плоскость руки’), центр. Gude *pátà* ‘большая плоская скала’, Mada *pta-tat* ‘плоский’.

2.8б. Центр. \**tap-* ‘плоский’: Podoko *tapə tapa*, Zulgo *tápát* ‘плоский’, Mafa *tapà<sup>a</sup>* ‘большая плоская скала’. Производный глагол: Muyang *tapà-t* ‘делать плоским’.

2.9а. Центр. \**gVz-* ‘чесотка’: Bura *g<sup>w</sup>azaza*, Gude *gəzá*, Daba *gèzìwè* ‘чесотка’, Mofu *-gəž-* ‘делать насечки, чертить’ (можно предположить семантический сдвиг: ‘чесаться, скрести’ > ‘делать насечки’).

2.9б. Зап. \**zag-* ‘чесаться, скрестись’: Karekare *zàgaa* ‘скрести’, Ngizim *zà-zgú* ‘скрести, тереть, чтобы избавиться от чесотки’, Duwai *zàzgo* ‘тереться, скрестись о стену (о животном)’ [CLD III №432].

Еще одним способом приведения словоформы в соответствие с моделью является ассимиляция по способу или месту образования фонем. С помощью ассимиляции удается достичь только равенства  $C_1$  и  $C_2$  по шкале сонорности.

Регрессивная ассимиляция по глухости – звонкости ( $C_1$  звонкий –  $C_2$  глухой >  $C_1$  глухой –  $C_2$  глухой) регулярно наблюдается в группе нгас (зап. ветвь).

3.1. Чад. \**gVwVp-* > \**g<sup>w</sup>ap-* ‘копье’: центр. Vana *g<sup>w</sup>ápà*, Glavda *gúpà*, Guduf, Hdi, Lamang *gupa* ‘копье’, группа нгас Ngas *k<sup>w</sup>op*, Goemai *k<sup>w</sup>ap*, Sura, Murun, Kofyar *kòp* ‘копье’.

3.2. Чад. \**buk-* ‘хижина’: зап. Hausa *búkkà* ‘хижина’, центр. Sukun *buk* ‘комната’, группа нгас Ngas *puk* ‘хижина’, Goemai *puk* ‘временная постройка, укрытие’.

3.3. Чад. \**bit-* ‘погасить огонь’: зап. Volewa *mbittu*, вост. Barein *biito* Migama *bittò*, Vidiya *bat*, *bit*, группа нгас Ngas *pit*, Mwaghavul, Murun, Goemai *pìt*.

Регрессивная ассимиляция по месту образования ( $C_1$  смычный –  $C_2$  фрикативный >  $C_1$  фрикативный –  $C_2$  фрикативный) отмечена только при  $C_1 = p$ .

3.4. Зап. (группа варджи) Раа *pita* ‘ломать’, Миа *pətə* ‘бить’, Diri *fətu* ‘ломать’.

3.5. Зап. (группа варджи) Warji *pəč-*, Paа *piču*, Diri *fəču* ‘распухать’.

3.6. Чад. \**pVs-* ‘копать, хоронить’: зап. Paа *pisù*, Miya *pəsə*, Warji *pəš-* ‘хоронить’, центр. Wandala *pásá* ‘копать’, Makari *fasi* ‘выкапывать’.

3.7. Центр. (группа тера): Tera *pəži-t*, Gaanda *fiža*, Gabin *fiče*, Boka *fúši* ‘зола’.

3.8. Зап. Ngas *pus*, Mwaghavul *pús*, Mushere *pús* ‘бросать, стрелять’, Tala *fəsi* ‘стрелять’.

3.9. Центр. Podoko *pəh<sup>w</sup>a*, Gaanda *foh<sup>w</sup>ə-ta*, Gabin *fuх<sup>w</sup>ə-tə* ‘мука’.

Прогрессивная ассимиляция по глухости – звонкости ( $C_1$  звонкий –  $C_2$  глухой >  $C_1$  звонкий –  $C_2$  звонкий) также уравнивает сонорность  $C_1$  и  $C_2$ . В нескольких чадских языках (Ngizim, Kilba, Ubi, Mubi) такое озвончение наблюдается регулярно.

4.1. Зап. Hausa *bábbákaa* ‘разводить огонь, жечь, жарить’, Volewa *bòkku* ‘жечь, жарить’, центр. Kilba *bágù* ‘готовить еду’, вост. Mawa *bəgə-η*, Ubi *bòogin*; Mubi *bàgé*, Jegu *biik* ‘жечь, жарить’.

4.2. Зап. Karekare *bàka* ‘рука, предплечье, крыло’, Galambu (*kà*)*bàgá* ‘плечо, крыло’, Duwai *bii-biyòk* ‘рука (от запястья до подмышки)’, центр. Bura *bubuku*, Chibak *bìbìkù*, Kilba *bàbùgu* ‘крыло’.

4.3. Вост. \**gVp-* ‘гор’: Dangla *gèepì*, Migama *gàpù*, Bidiya *geepinò*, gapi, Birgit *gèefó*, Mabire *gepo*, Ubi *gaba*, Mawa *gab*, Mubi *gèebi*.

4.4. Зап. Hausa *bìšè* ‘хоронить’, Bade *əbžù* ‘копать яму’, Ngizim *bəžù* ‘копать’, центр. Bura *bža* ‘хоронить’.

4.5. Вост. \**bis-* ‘москит’: Mawa *bisi*, Dangla *bìsì*, Migama *biisè*, Bidiya *biisa*, Jegu *biisó*, Birgit *bisiyó*, Ubi *bizyo*.

4.6. Зап. Hausa *bùusaa* ‘отделять отруби от зерна’, Ngamo *bùuzù* ‘отруби’, вост. Bidiya *bóssa* ‘отруби’.

Лениция ларингала в позиции  $C_2$  ( $C_1-h > C_1-w/y$ ) также приводит словоформу в полное соответствие с моделью. Падение ларингала компенсируется эмфатизацией согласного в позиции  $C_1$ . Первичность ларингала (а не эмфатического согласного) на афразийском уровне подтверждается внешними параллелями.

5.1. Куш. Igaqw *púx* (идеофон) ‘звук падения на землю’ [MQK: 83], берб. Rif *fuha* ‘падать’ [NZ: 544]. Чад. \**pVh-* ‘падать, ложиться’: центр. Ouldem *-pùh<sup>w</sup>* ‘падать’, зап. Mburku, Tsagu *píy*, Warji *píy-*, *píw* ‘ложиться’ (вторичная эмфатизация смычного), Bade *pàwu*, вост. Bidiya *paay* ‘спускаться’.

5.2. Сем. Arabic *d<sup>w</sup>* ‘кричать, звать’ [BK I: 704]. Центр. чад. Malgwa *dá* ‘звать’, Glavda *da* ‘звать, кричать’, Zulgo *de* ‘звучать, отзываться’, Makari *dì* ‘приветствовать’, Muayang *dwi* ‘крик’, Dari *dīē* ‘называть, звать’, вост. Mawa *dii-η-* ‘говорить’ (вторичная эмфатизация) [CED №128].

5.3. Сем. Arabic *byh* ‘резать (мясо) на кусочки’ [BK I: 182]. Зап.-чад. Ngas *bii*; Kariya *bíu*, *bəbí*, Siri *bíu* ‘резать, рубить’, Ngizim *biyú* ‘выпотрошить животное’ (вторичная эмфатизация).

Два типа рефлексов прачадских глухих смычных в позиции  $C_2$  в языке Tangale (зап. ветвь) также подтверждают актуальность модели словоформы для чадских языков. Если начальный согласный словоформы тоже глухой смычный, в Tangale наблюдается озвончение  $C_2$  (т.е. диссимилиация по глухости – звонкости). В следующих примерах рефлекс в языке Tangale стоит после знака ~.

6.1. Чад. \**kVp-* ‘бить, ломать’: зап. Hausa *káf* ‘удар’, центр. Mofu *-kup-* ‘дать пощечину, расплющить’, Dzeraw *kap, kep* ‘бить’, вост. Mokilko *kóppé* ‘бить, ударять’ ~ Tangale *kεεβε* ‘разбить, разломать’.

6.2. Чад. \**pVt-* ‘дуть, веять’: зап. Ngamo *àptá* ‘веять’, Bghom *paat* ‘дуть’, Guus *piriitá* ‘обмахиваться’, центр. Mofu *-pát-k-* ‘веять’ ~ Tangale *peede* ‘веять, обмахиваться’.

6.3. Чад. \**pVt-* ‘выходить’: зап. Hausa *fita*, Ngamo *fàta*, Karekare *fàtaa*, Bolewa *pàtaa*, Pero *pétò*, Bele *fátí-*, центр. Daba *pát* ‘выходить’, Mofu *pát* ‘выйти из дома’, Gude *pitá* ‘проходить мимо’ ~ Tangale *pódi* ‘выходить’ [CED №422].

6.4. Зап. Karekare *kùtu*, центр. Malgwa *kata* ‘заканчивать’ ~ Tangale *k<sup>w</sup>ade* ‘заканчивать’.

Однако если в позиции  $C_1$  стоит глухой сибиллянт, то озвончение не наблюдается, а именно:

6.5. Центр. Vanana *soto* ‘грязь’, вост. Sokogo *šíta* ‘грязный’ ~ Tangale *sitiŋ* ‘грязный’.

6.6. Зап. Goem *saat* ‘пролить’, Bolewa *šato* ‘капли дождя’ ~ Tangale *site* ‘лить воду’.

6.7. Зап. Pero *čéřò* ‘следить’ ~ Tangale *sope* ‘следовать за’.

6.8. Зап. Warji *čəpr-ái*, Kariya *čipi-r*, Miya *čəpú-r* ‘моча’ ~ Tangale *sipe* ‘мочиться’.

Такое «двойное» отражение глухих смычных в одной и той же позиции (*-p-* и *-b-*, *-t-* и *-d-*) нарушает основное положение компаративистики о регулярности фонетических рефлексов. Поэтому разумно предположить, что в данном случае озвончение смычных в интервокальной позиции – это не фонологический процесс. Изменения происходят под влиянием модели. В примерах 6.1 – 6.4 с помощью озвончения  $C_2$  словоформа приводится в полное соответствие с моделью:  $C_2$  становится выше  $C_1$  по шкале сонорности. В примерах 6.5 – 6.8 никакие изменения не требуются, поскольку словоформа изначально соответствует модели –  $C_2$  (смычный) выше  $C_1$  (сибилянта) по шкале сонорности. Именно поэтому не происходит озвончения смычного.

Приведенный лексический материал подтверждает гипотезу об актуальности модели словоформы с обозначенными выше параметрами ( $C_1$  не может быть выше  $C_2$  по звучности) для чадских языков. Различные и, на первый взгляд, никак не связанные между собой процессы – метатеза, ассимиляция, диссимиляция, лениция – находят объяснение в рамках выдвинутой гипотезы. А именно – все эти процессы имеют единую цель: привести словоформу в соответствие с моделью.

Актуальность данной модели словоформы для чадских языков подтверждается и статистически – наиболее частотной (и не подверженной метатезам) является структура с плавным или сонорным согласным в позиции  $C_2$ . Высокая частотность характерна и для структур с фрикативной фонемой в позиции  $C_1$  (особенно при звонком смычном в позиции  $C_2$  – т. е., *s-b-*, *s-d-*, *t-b-*, *t-d-*), при этом формы с метатезой (*b-s-*, *d-s-*) обладают низкой частотностью, а формы типа *b-t-*, *d-t-* почти не представлены в словарях.

Понимая, чем мотивированы изменения, мы можем объединить формы с метатезой и без метатезы в одно этимологическое гнездо, объяснить нерегулярность

фонетических соответствий и неравномерный характер изменения словоформ с течением времени (ведь все описанные выше трансформации не затрагивают словоформы, которые изначально соответствовали модели).

### Сокращения

АА – афразийский	куш. – кушитский
берб. – берберский	сем. – семитский
вост. – восточный	центр. – центральный
зап. – западный	чад. – чадский

### Источники

АА – Diakonoff, Igor, Belova, Anna, Militarev, Alexandr, Porkhomovsky, Victor and Olga Stolbova. *Comparative Historical Afrasian Dictionary //St-Petersburg Journal of African Studies*. 1994. № 2. P. 5–26.

АТО – *Sociolinguistic Survey Report of the Languages of the Gawwada, Tsamay and Diraasha Areas with Excursions to Ongota and Arbore*.  
URL <https://www.sil.org/resources/archives/9071>. Дата обращения 30.05.2021.

БК – Biberstein-Kazimirski, Albert De. *Dictionnaire arabe-français*. Paris: Maisonneuve et Co, 1860.

СС – Caitucoli, Claude. *Lexique Masa Tchad et Cameroun*. Paris: Agence de Coopération Culturelle et Technique, 1983.

CED – Столбова О.В. *Этимологический словарь чадских языков / Stolbova, Olga. Chadic Etymological Dictionary*. URL: <https://book.ivran.ru/f/ilovepdfmerged.pdf>. Дата обращения 30.05.2021.

CLD III – Stolbova, Olga. *Chadic Lexical Database. Issue III*.  
URL: <https://chadicnewsletter.wordpress.com/2010/03/18/chadic-lexical-database/>. Дата обращения 30.05.2021.

DUL – Olmo Lete, Gregorio del and Joaquin Sanmartin. *A Dictionary of the Ugaritic Language in the Alphabetic Tradition*. Leiden: Brill, 2015.

EEN – Ehret, Christopher, Elderkin, Edward and Derek Nurse. *Dahalo Lexis and its Sources // Afrikanistische Arbeitspapiere*. 1989. P. 5–49.

EG – Erman, Adolf and Hermann Grapow. *Wörterbuch der ägyptischen Sprache*. Bd. III. Berlin: Akademie, 1971.

LsG – Leslau, Wolf. *Comparative Dictionary of Ge'ez*. Wiesbaden: Harrassowitz, 1987.

MQK – Mous, Maarten, Qorro, Martha and Roland Kiessling. *Iraqw-English Dictionary*. Köln: Köppe, 2002.

NZ – Naït-Zerrad, Kamal. *Dictionnaire des racines berbères*. Paris – Louvain: Peters, 2002.



PSL – Kogan, Leonid. Proto-Semitic Lexicon // Stefan Weninger, ed., *The Semitic Languages. An International Handbook*. De Gruyter Mouton, 2011. P. 179–258.

SED – Militarev, Alexandr and Leonid Kogan. *Semitic Etymological Dictionary. Vol. I. Anatomy of man and animals*. Münster: Ugarit Verlag, 2000.

### Литература

1. Жвания Н. *Совместимость согласных в корнях языка геез*. Автореферат кандидатской диссертации. Тбилиси. 1972.
2. Коган Л.Е. *Семитские языки. Языки мира*. М.: Академия, 2009. С. 15–113.
3. Greenberg, Joseph. The Patterning of Root Morphemes in Semitic // *Word*. 1950. №2. P. 162–181.
4. Koskinen, Kalevi. Kompabilität in den dreikonsonanten Hebräischen Wurzeln // *Zeitschrift der Deutschen Morgenländischen Gesellschaft*. 1964. Bd. 114. P. 16–58.

## Сочетаемость праязыковых элементов и структурная трансформация иранских языков

Джой Иосифовна Эдельман

Институт языкознания РАН, Москва, Россия,

joy.edelman@gmail.com, ORCID 0000-0003-0056-6250

**Аннотация.** В реконструируемой праиранской морфо-синтаксической системе языковые единицы выявляли сочетаемость, или валентность, трех основных типов: 1) активные (корни и основы с именной и глагольной семантикой); 2) пассивные (часть корней и основ с наречной семантикой, частицы); 3) и активные, и пассивные (другая часть корней и основ с наречной семантикой). В статье анализируется роль этих единиц и их типов в дальнейшей структурной трансформации иранских языков..

**Ключевые слова:** семантика, сочетаемость, лексика, грамматика, структурная трансформация, иранские языки

**Abstract.** Language units of the Proto-Iranian reconstructed morpho-syntactic system reveal three main types of combinability, or valency: 1) active types (roots and stems with nominal and verbal semantics); 2) passive types (part of the roots and stems with adverbial semantics, and all particles); 3) active and passive types (the other roots and stems with adverbial semantics). The article analyses the role of these units, and their types in the further structural transformation of the Iranian languages..

**Keywords:** semantics, combinability, words, grammar, structural transformation, Iranian languages

Изменения структурных морфолого-синтаксических параметров иранских языков неразрывно связаны с историей их лексики, начиная с праязыкового периода. Тем самым этимология реконструированных праязыковых корней, основ и полных словоформ в их фонологическом облике и, что в данном случае особенно существенно, – в их семантике – является базой для дальнейших рассмотренной трансформации системы иранских языков в разных аспектах, включая структурно-грамматический. Реконструированное «ядро» слова, несущее основную семантическую нагрузку и обладающее активной сочетаемостью, или валентностью, предусматривает присоединение к нему «вторичных» языковых элементов, обладающих пассивной валентностью. Поэтому можно предположить, что соотношение грамматики со словарем при словоцентрическом подходе к языку актуально как для формализованного лингвистического описания языка, о котором писала З.М. Шаляпина [Шаляпина 1991], так и для исследования диахронических изменений в истории языка.

Связь некоторых структурных процессов в грамматике иранских языков с семантикой реконструируемых древних «ядерных» лексических единиц можно проиллюстрировать фрагментами «Этимологического словаря иранских языков» (далее ЭСИЯ).

Хотя отдельные реконструированные праиранские языковые единицы могли быть синхронны друг другу лишь относительно, они все же представляли собой элементы единой в целом системы, фонетико-фонологическая и грамматическая характеристика которой в принципе изучена и описана. Это был язык с богатой флективной системой, в котором словоизменительные парадигмы и словообразовательные модели (в составе: корень + аффиксы + флексии) были к тому же пронизаны системой морфонологических чередований. Таким образом мы можем рассматривать каждую реконструированную праиранскую словоформу как «комбинированную» языковую единицу – результат сочетаемости составляющих ее не только линейных, сегментных единиц, но и внесегментных элементов. При этом приходится учитывать, что реконструкция праязыковых словоформ на базе их рефлексов в более поздних языках может свидетельствовать о возможности их функционирования в праязыке, но у нас нет данных утверждать, что та или иная словоформа в тот период действительно существовала, или что ее не могло быть. Часть их могла возникнуть позднее, из исконного материала по аналогии с другими сходными словоформами.

В такой реконструированной системе естественно, что возможности сочетаемости языковых единиц зависят прежде всего от семантики ядра, то есть реконструируемого корня, точно так же, как от семантики зависят возможности сочетаемости единиц современных языков в их синхронном состоянии [Шаляпина 1992: 42]. Семантика же предопределяет и наиболее частотные словообразовательные и словоизменительные возможности продолжения данного корня, основы в дальнейшем.

Рассмотрим вначале праязыковые единицы, имевшие активную сочетаемость, или валентность, а именно корни и основы с именной и глагольной семантикой.

Уже в праиранском состоянии основы с семантикой предмета и/или понятия, относящиеся к именной подсистеме (включая и отглагольные имена и лексикализованные причастия) характеризуются продуктивным словообразованием: наблюдается большое число композитов, аффиксальных производных слов, включая относительные имена. Сохраняется унаследованная из позднего индоевропейского состояния система определенных именных категорий. Лексико-грамматическая категория рода (лексическая для субстантивов, словоизменительная для прилагательных и части числительных и местоимений), выражавшаяся сочетаемостью корня с различными посткорневыми элементами: суффиксами и падежными окончаниями определенного типа; часть субстантивов выражала род аналитически – синтаксической сочетаемостью с прилагательными и местоимениями того или иного рода. Две именные категории носили чисто грамматический характер, отражая ситуационную характеристику субстантивов и качеств: это категории падежа и числа, представляемые флективными падежными парадигмами. То есть они были выражены сочетаниями корней или вторичных основ с определенными флективными формантами, в части парадигм также с морфонологическими чередованиями презентации корня. Например, от праиранской основы имени *\*pitar-* «отец» выступает номинатив *\*pitā*, в косвенных падежах и производных словах основа представлена вариантами: *\*(p)tar-*, *piθr-*, *\*θr-*, *\*turi-*, *\*turiā-*, *\*pitruća-* и др.

Частично эти варианты сохранились в языке Авесты: авест. диалект Гат – ном. *pīā*, акк. *pataṛēm*, дат. *fədrōi*, *piθrē* и др. [Bartholomae 1904: 905-906; ЭСИЯ 6: 303].

Особую подсистему имен представляли личные местоимения 1-го лица, у которого двойственное и множественное число отличались от единственного лексически, а номинатив и косвенные падежи восходили к этимологически различным корням. Таким образом местоимение 1 л. ед. ч. ном. *\*azám* «я» из индоевропейского *\*eg(h)-óm* [ЭСИЯ 1: 294-295] не сочетается ни со словообразующими, ни с формообразующими элементами, это единственный отмеченный случай именной единицы с нулевой сочетаемостью.

Глагольная лексика праиранского состояния тоже характеризуется развитым словообразованием. Характерно употребление превербов (отделяемых и неотделяемых) и суффиксов, существенно меняющих семантику глагола. В этот период существует также небольшое число деноминативных глаголов, составляющих особый формализованный класс, входящий в разряд слов с семантикой действия и глагольной парадигмой (например сохранившиеся в Авесте: *arəza-* «борьба, битва, сражение» – *arəzayant-* прич. наст. вр. «сражающийся» [Bartholomae 1904: 201-202; ЭСИЯ 1: 230]; *baēšaza-* 1) «целебный»; 2) «средство для лечения; лекарство»; 3) «целебная сила» – *baēšazu-* – основа настоящего времени глагола «лечить, исцелять» [Bartholomae 1904: 915-916; ЭСИЯ 2, 62-3].

В ранней древнеиранской глагольной парадигме ведущей семантической категорией можно считать лексико-грамматическое членение основ на переходные / непереходные (и входящие в них менее крупные классы глаголов: каузативные, медиопассивные и др.), выраженное определенными суффиксами, огласовками основы, местом ударения и др., а также типами окончаний. Обозначения разных свойств действия (видо-временные характеристики, модальность, залог, версия) выражались суффиксацией и морфонологическими средствами, типами флективных личных окончаний. Отраженные окончаниями категории лица и числа характеризовали не само действие, а актантов.

Кроме корней, относительно жестко привязанных семантически (и потому – грамматически) либо к классу имен, либо к классу глаголов, в праиранском словнике имеются корни со значениями как предмета, понятия, так и действия, состояния. Возможность лексической и грамматической сочетаемости таких корней очень широка, что реализуется впоследствии и в словообразовании, и в формообразовании. Примером может служить корень *\*rauk-* : *\*rauč-* : *\*ruk-* «свет, светлый; светить(-ся), сверкать» и т.д., от которого до нас дошли в основном рефлексы имен, например *\*rauča(h)-* «светящийся, светлый; день» (> др.-перс. *raučah-* «день», совр. перс. *ruz* «день»); меньше дошло собственно глагольных основ, например, *\*rauča-* «светить, сиять» (> среднеперс. *rōz-* «светится, сиять» и каузатив *rōzēn-* «освещать»), причастие наст. вр. *\*raučana-* «светящий, освещающий» (> тадж. *равзан* [*rawzan*] «световое и дымовое окно в крыше» и др.), подробнее см. [ЭСИЯ 6: 413-424].

Дальнейшая трансформация именной парадигмы происходила путем утраты в первую очередь элементов, становившихся формальными, и переноса части морфологических противопоставлений на синтаксис, в частности, закрепление места главных членов предложения по отношению друг к другу (в основном (S)OV),

стандартизация порядка имен в определительных сочетаниях, усиление роли предлогов и послелогов и т.п. Все это способствовало сворачиванию падежных парадигм. В результате заударные окончания, становившиеся избыточными, отмирали (при более длительном сохранении информативно актуальных, в частности также заударного показателя генитива). При этом категория числа, информационно существенная, сохранялась в морфологии, но приобрела относительно единообразную для разных языков модель: формой единственного числа служит основа имени, множественного – имя с постфиксом и т.д.

В трансформации глагольной системы обращает на себя внимание различная продуктивность словообразования в разных иранских языках. В некоторых литературных языках наблюдается довольно большой процент деноминативных глаголов, часть из них образуется по древнеиранской модели на базе исконных имен, например от именных основ *\*pād-* «нога», *\*padá-* «стопа, ступня, след (ноги); шаг» – кл. перс. *pāyidan*, *pāyestan* «долго держаться, быть прочным, храниться», парф. *pādan* «стоять, останавливаться, ждать». Последующая продуктивность этой модели сказывается и в разновременных образованиях из заимствованных слов, например, кл. перс. *fahmīdan* «понимать», шугн. *fāmtow* «знать» от арабского масдара *fahm*, хорезм. *tafsīri(y)-* «толковать, комментировать (Коран)» от арабского *tafsīr* «толкование (Корана)». В дальнейшем при затухании собственно глагольного словообразования во многих языках развиваются сложноименные образования, состоящие из имени и глагола. Морфологическая парадигма глагола изменилась полностью, однако остался семантический принцип обозначения понятия действия: оформление соответствующими элементами категорий времени, лица и числа. Категория переходности / непереходности, весьма существенная, в подавляющем большинстве языков усиливается, приобретая ярко выраженное синтаксическое оформление (построение эргативных и эргативообразных конструкций предложения). Другие глагольные характеристики: видо-временные, модальные, залоговые, – выражаются различными по языкам средствами. Категории лица и числа, характеризующие не действие, а актантов, выражаются, как и в древности, глагольной флексией, в части языков с инновационными показателями (особенно для форм множественного числа).

Подробнее об этапах и о характере структурной перестройки древних морфологических парадигмах имени и глагола см. [Расторгуева 1975: 112 и сл.], о степени актуальности информации для ее сохранности при трансформации этих парадигм см. [Эдельман 1987: 159 и сл.].

Иные пути дальнейших трансформаций выявляют праязыковые единицы, имевшие пассивную сочетаемость, или валентность, а именно слова и частицы (в широком смысле) с локальной, направительной и темпоральной семантикой. К ним относится часть наречий (которые могли иметь и активную, и пассивную сочетаемость), а также некоторые наречия, предлоги, союзы, частицы (с пассивной сочетаемостью). Их дальнейшая судьба более тесно связана с синтаксисом, чем у слов с именной или глагольной парадигмой: здесь большую роль играют возможности сочетаемости языковых элементов не внутри словоформы, а на более широком пространстве – в синтаксисе.

Эти элементы (назовем их условно Adv, то есть «адвербы») уже в праиранском предложении с его относительно свободным порядком слов (о чем свидетельствуют тексты Авесты) могли размещаться в начале, середине и конце пред-

ложения и примыкать к словам с активной сочетаемостью, то есть с именной и глагольной семантикой. В дальнейшем их функционирование как самостоятельных слов или как «прилеп» к другому слову, характеризующемуся активной сочетаемостью, зависит от значения предложения в целом или отдельного синтаксического блока (синтагмы). Средствами служат элементы фразовой просодии: фразовая интонация, фразовое ударение, наличие или отсутствие пауз между блоками и т.п.

В подавляющем большинстве иранских языков установился порядок основных членов предложения (S)OV, в котором местами элементов типа Adv могут быть: Adv<sub>1</sub> – (S) – Adv<sub>2</sub> – O – Adv<sub>3</sub> – V (с вариантами наличия нескольких прямых и косвенных дополнений O<sub>1</sub>, O<sub>2</sub>, ... и окружающих их элементов типа Adv).

При более редком порядке (S)VO местами элемента Adv могут быть: Adv<sub>1</sub> – S – Adv<sub>2</sub> – V – Adv<sub>3</sub> – O (также с вариантами наличия O<sub>1</sub>, O<sub>2</sub> и окружающих их элементов Adv).

При этом обычно в повествовательном тексте обстоятельство времени, выраженное наречием или словосочетанием, ставится в начале предложения, занимая позицию Adv<sub>1</sub>. Элемент Adv<sub>2</sub>, стоящий перед именем объекта (Adv-O) или перед глаголом (Adv-V), может стать соответственно предлогом или словообразовательным префиксом при имени, либо префиксом-превербом при глаголе. Элемент Adv<sub>3</sub>, расположенный после имени объекта перед глаголом в предложении типа (S)OV, то есть в позиции (O – Adv – V), может стать послелогом при имени (O-Adv – V) или глагольным превербом, то есть занять место (O – Adv-V). Начало процесса вариативности такого «приращения» этого элемента, которое допускает двойное толкование, отмечено уже в языке Авесты. Например, праиранская лексема \*abi – наречие, предлог, преверб со значениями направленности, предназначённости и т.п. [ЭСИЯ 1: 72-74], отразившаяся в авестийском в виде *avi*, при употреблении в тексте авест. *asmanəm avi frašusāni, zəm avi ni-urvisyāni* «к небу мне пойти, к земле мне обратиться?» допускает двойную интерпретацию. Этот направительный элемент *avi* можно трактовать и как послелог к именам *asmanəm* «небу», *zəm* «земле», и как вторичный преверб к следующим за ним глаголам [Соколов 1979: 203].

В предложении типа (S)VO, редком для большинства иранских языков, элемент Adv<sub>3</sub>, оказавшийся в позиции (V – Adv – O), может стать приглагольным постпозитивным элементом – поствербом, или наречной энклитикой (V-adv – O). Такое явление характерно для хорезмийского языка, где благодаря относительно поздней перестройке структуры простого предложения с обычного для иранских языков порядка (S)OV на инновационный (S)VO развилась система приглагольных энклитических «прилеп» – поствербов или «послелогов», предвещающих последующий объект, выраженный именем или местоимением. Это явление в содержательном отношении является зеркальным отражением так называемой местоименной репризы, свойственной языкам Балканского языкового союза. Постглагольные цепочки имеют определенную аранжировку поствербов. См. например, хорезм. *p'cnwdy-w* (< *p'cnwd'-hy-w*) *'y wž f-nknc* «(он) вдел нитку в иголку», букв. «(он) вдел-ее-туда ту нитку в-иголку»; хорезм. *c'yt'hyw'byr* (< *c'yt'-hy-w'-byr*) «(он) взошел туда к нему», букв. «(он) вошел-ему-туда-на» [Henning 1977: 453;

Эдельман 2008: 55]. Здесь наречные энклитические поствербы *-w*, *-w'* [wā] «туда», указывающие на направление, ставятся после глагольной флексии и после энклитического местоимения *hy* [h(i)] «тот, он», указывающего на прямой объект в первом предложении и на косвенный объект – во втором, где послелог *-byr* [-bir] «на» замыкает цепочку. Таким образом, в первом предложении вся глагольная словоформа построена как (V-O-Adv) с последующей расшифровкой объектов в виде имен. Во втором предложении глагольная словоформа построена как (V-O-Adv-Adv) с постфиксами, указывающими на косвенный объект и наречие, и послелогом *-byr* [-bir], который, в свою очередь продолжает древнее наречие *\*upari* «наверху, вверху, сверху» [Bartholomae 1904: 394]; имя объекта отсутствует.

В качестве примера наречия, обладающего активной и пассивной сочетаемостью, можно привести разные типы трансформации праиранской лексемы *\*antar* «внутри, внутри, между, в», соответствующей др.-инд.наречию *antar* «внутри, внутри; между», из индоевропейского *\*enter* [ЭСИЯ 1: 160-163]. В древнеиранских текстах она зафиксирована как предлог и как наречие, рефлексы ее в живых и вымерших иранских языках дают представление о ее сочетаемости в предшествующие периоды.

На лексическом уровне эта языковая единица участвует в словообразовании: 1) в качестве основы с активной сочетаемостью – в образовании имен и сложных наречий; 2) в качестве вспомогательного элемента с пассивной сочетаемостью – в роли преверба в глаголе; она может стать отдельной лексемой – предлогом (связанным уже с морфолого-синтаксическим уровнем, см. ниже).

В именах данная единица может выступать как центральное семантическое ядро слова, то есть основа имени, которая может сочетаться с суффиксами: с суффиксом *\*a-* – имя *\*antar-a-* «внутренний, внутренняя часть чего-либо», например авест. *antara-* «внутренний, находящийся внутри»; с суффиксом мужского рода *\*-ka* – имя *\*antara-ka-* букв. «внутренний», например согд. ман. *'ndryk*, согд. будд. *'ntryk* «евнух», то есть «допущенный во внутренние покои дома»; с суффиксом женского рода *\*-či-* – имя *\*antara-či-* букв. «внутренняя», например вахан. *andarč* «жены братьев по отношению друг к другу», то есть «(живущие) внутри единой большой семьи».

Рефлексы основы *\*antar* зафиксированы также в качестве первого элемента имени-композиата, с сохранением основного значения, то есть при обозначении внутренней части чего-либо, например авест. *antarə-naēta-* м.р. «внутренняя половина/сторона». Из реконструированного *\*antar-rauna-* «внутренняя сторона» происходит ср.-перс. *andarōn* 1) прилагательное «внутренний»; 2) наречие «внутри»; 3) предлог «внутри, в»; кл. перс. *andarūn* «интерьер, женская половина дома; внутренний мир, душа». Сходные значения имеют лексемы с отпадением начального слога, в форме *darūn*: кл. перс. *darūn*, тадж. *darun*, совр. перс. *darun* 1) «внутренность, внутренняя часть чего-либо»; 2) «внутренний, духовный мир; душа»; 3) «внутренний, расположенный внутри». Позднее по этой модели в разных языках образуются имена с рефлексам компонента *\*antar-*, например хорезм. *'ndr-n'c* «внутренняя часть ноздри», язг. *əndər-avdūst* «тонкие перчатки, надеваемые под рабочие рукавицы».

В наречиях данные основы могут выступать в сочетаниях с предлогами, например, в сочетании *\*hača-antar*, букв. «изнутри», где первым компонентом выступает *\*hača-* предлог, префикс «из» с частичной утратой значения, см. согд. будд. *čntr*, согд. ман. *čndr*, ягн. *čintir* «внутри», язык ванечи *zdare* «внутри» [ЭСИЯ 3: 301].

В словообразовании глаголов рефлексы *\*antar-* выступают как словообразующие элементы, однако уже не с активной, а с пассивной сочетаемостью: в глаголах движения выступают как превербы, указывающие на направление внутрь с разными оттенками значения, например: тадж. *даромадан* «входить, вступать; поступать [на работу, учебу в учреждение]; входить, помещаться [в узкое пространство]»; кл. перс. *daramadan* «выходить, появляться [откуда-то, сюда]; восходить [о светилах]».

На морфолого-синтаксическом уровне рефлексы *\*antar-* образуют вспомогательные синтаксические элементы при именах: 1) предлоги, закрепившиеся в результате употребления в позиции (Adv-O), например: кл. перс. *andar, dar*; совр. перс., тадж. *dar* «в, внутри»; 2) послелоги, закрепившиеся в результате употребления в позиции (O-Adv), например: таджикские диалектные *-anda, -ənda, -nda* «в, внутри, при»; шугн. *-andir*, язг. *-ənda* с разными оттенками значений «в, внутри, при (ком-либо)» и др.

Наречия, продолжающие *\*antar* и производные с этим элементом как основой, в дальнейшем могут выступать в функции основ для изафетных предлогов, например, тадж. *дарун* наречие «внутри, внутрь» используется также и как основа отыменного изафетного предлога *дарун-и* «внутри, внутрь». Отмечено также сочетание с предлогом *az* «из»: тадж. *az дарун-и* «изнутри».

Зафиксирован единичный случай лексикализации в виде сложного слова-фразеологизма, продолжающего целое предложение с элементом *\*antar-* в качестве предлога: кл. перс. *nēst-andar-jahān* «редкостная, единственная», букв. «не-есть (больше такой) в мире».

Подробнее о прототипе *\*antar* и его дальнейших рефлексах см. [ЭСИЯ 1: 160-163].

Из всего сказанного можно сделать следующий вывод. Соотношение грамматики со словарем при словоцентрическом подходе к языку важно для исследования диахронических изменений в разных языках, восходящих к единому языку. Рассмотрение реконструируемых праиранских корней, основ и слов с точки зрения их сочетаемости – активной и пассивной – дает возможность не только объемного анализа семантики древних языковых единиц, но и определения их принадлежности к тому или иному лексико-грамматическому классу и их роли в трансформации системы иранских языков в разных аспектах, включая структурно-грамматический.

## Сокращения

авест. – авестийский  
вахан. – ваханский  
др.-перс. – древнеперсидский  
кл.-перс. – классический персидский  
парф. – парфянский  
совр. перс. – современный персидский

ср.-перс. – среднеперсидский  
тадж. – таджикский  
хорезм. – хорезмийский  
шугн. – шугнанский  
ягн. – ягнобский  
язг. – язгулямский



согд. будд. – согдийский язык буддийских текстов  
согд. ман. – согдийский язык манихейских текстов

## Литература

1. Расторгуева В.С. Вопросы общей эволюции морфологического типа // *Опыт историко-типологического исследования иранских языков (в двух томах). Т. I. Фонология. Эволюция морфологического типа* / Отв. ред. В.С. Расторгуева. М.: Наука, 1975. С. 89–224.
2. Соколов С.Н. Язык Авесты // *Основы иранского языкознания. Древнеиранские языки* / Редколлегия В.И. Абаев, М.Н. Боголюбов, В.С. Расторгуева (отв. ред.). М.: Наука, 1979. С. 129–233.
3. Шаляпина З.М. Грамматика и ее соотношение со словарем при словоцентрическом подходе к языку (на опыте формализованного лингвистического описания) // *Вопросы языкознания* 1991. № 5. С. 42–54.
4. Эдельман Д.И. Взаимодействие языковых уровней в истории иранских языков // *Иранское языкознание. Ежегодник 1982* / Редколлегия: М.Н. Боголюбов (гл. ред.), А.Л. Грюнберг, В.А. Ефимов, В.С. Расторгуева. М.: Наука. 1987. С. 158–168.
5. Эдельман Д.И. Хорезмийский язык // *Основы иранского языкознания. Среднеиранские и новоиранские языки* / Редколлегия М.Н. Боголюбов, В.А. Ефимов (отв. ред.), Д.И. Эдельман. М.: Издательская фирма «Восточная литература», 2008. С. 6–60.
6. ЭСИЯ 1 – Расторгуева В.С., Эдельман Д.И. *Этимологический словарь иранских языков. Т. 1.* М.: Издательская фирма «Восточная литература» РАН, 2000. 327 с.
7. ЭСИЯ 2 – Расторгуева В.С., Эдельман Д.И. *Этимологический словарь иранских языков. Т. 2.* М.: Издательская фирма «Восточная литература» РАН, 2003. 502 с.
8. ЭСИЯ 3 – Расторгуева В.С., Эдельман Д.И. *Этимологический словарь иранских языков. Т. 3.* М.: Издательская фирма «Восточная литература» РАН, 2007. 493 с.
9. ЭСИЯ 6 – Эдельман Д.И. *Этимологический словарь иранских языков. Т. 6.* М.: Наука – Восточная литература, 2020. 463 с.
10. Bartholomae Chr. *Altiranisches Wörterbuch*. Strassburg: K. J. Trübner, 1904 (repr.: Berlin-New York: Walter de Gruyter & Co., 1979). 1999 Kol.
11. Henning W.B. The structure of the Khwarezmian verb // *Henning W.B. Selected papers. Vol. II.* Leiden: E.J. Brill, 1977. P. 449–455.

## Структура лексико-семантического поля соматических единиц на материале фразеологического фонда персидского языка

**Юлия Вадимовна Будкина**

Институт востоковедения РАН, Москва, Россия,  
ulchik08@rambler.ru, ORCID 0000-0001-8485-8853

**Аннотация.** В статье рассматривается лексико-семантическое поле соматизмов современного персидского языка в аспекте их участия во фразеологических сочетаниях. Уточняется классификация соматизмов, выделяются лексико-семантические группы, для каждой из которых приводятся примеры соматических фразеологизмов.

**Ключевые слова:** лексико-семантическое поле, лексико-семантическая группа, соматизм, семантика, фразеология

**Abstract.** This article discusses the lexico-semantic field of body parts and related vocabulary in Modern Persian in regard to their participation in phraseological units. Lexico-semantic groups within the field are reexamined, and examples of phraseologisms are provided for each group.

**Keywords:** lexico-semantic field, lexico-semantic group, semantics, body part vocabulary, phraseology unit

### 1. Лексико-семантическое поле в качестве объекта исследования

В настоящее время проблема изучения лексико-семантических полей является весьма актуальной, т.к. интерес к исследованиям подобного рода продиктован необходимостью всестороннего изучения слова как основной единицы языковой системы. На данный момент именно лексико-семантическое поле представляет собой универсальный «инструмент» для изучения понятийного, предметного и функционального сходства анализируемых лексических единиц.

Известно, что первые попытки объединения слов в семантические поля были предприняты еще при создании идеографических словарей. Наиболее же активное системное изучение лексики, представленной в виде семантических полей, было характерно для работ немецких ученых начала XX в. Большой вклад в развитие полевой теории внесли Р. Карнап и Л. Витгенштейн, также важное значение и на сегодняшний день имеют ставшие уже классическими исследования Й. Трира, В. Порцига, Л. Вайсгербера.

Среди всех видов объединения языковых единиц по содержанию лексико-семантическое поле выступает как наиболее обширное. Чаще всего оно представляет собой ономаσιологическую и семантическую «группировку» множества языко-

вых единиц. Объединяя слова разных частей речи, лексико-семантическое поле также характеризуется наличием интегральной семы.

В современной лингвистике существует несколько определений термина «лексико-семантического поля» (ЛСП). Например, Л.М. Васильев трактует это понятие следующим образом: «семантическими полями принято считать и семантические классы слов какой-либо одной части речи, и семантически соотнесенные классы слов разных частей речи, и лексико-грамматические (функционально-семантические) поля, и парадигмы синтаксических конструкций, различные типы семантико-синтаксических синтагм. Объединяет их, несмотря на существенные структурные различия, то, что все они имеют в своей основе ту или иную семантическую категорию или категориальную ситуацию» [Васильев 1990: 126].

И.В. Арнольд же считает, что «... его можно понимать как множество слов, покрывающих определенную область человеческого опыта, и, следовательно, связанных по значению. Множество это создается указанием общего свойства его компонентов» [Арнольд 1984: 3].

В данной работе мы будем опираться на определение И.М. Кобозевой, так как, по нашему мнению, оно является наиболее ёмким и, тем не менее, точно отражающим суть рассматриваемого языкового явления. С ее точки зрения, «семантическое поле представляет собой совокупность языковых (главным образом лексических) единиц, объединенных общностью содержания (иногда также общностью формальных показателей) и отражающих понятийное, предметное или функциональное сходство обозначаемых явлений» [Кобозева 2000: 99].

Для чего же необходимо изучать лексико-семантические поля? Считается, что подобные исследования показывают различия в картине мира носителей разных языков, помогают рассмотреть их систему ценностей через призму языкового материала. Каждый народ имеет свои принципы членения внешнего мира, свой взгляд на окружающую действительность, поэтому семантические системы разных языков не совпадают [Васильев 1971: 23].

Примерно с середины 70-х годов XX века заметно возрос интерес лингвистов к проблеме описания фразеологических полей. В первую очередь, это объясняется более пристальным изучением семантики фразеологических единиц. Среди отечественных ученых, занимавшихся исследованиями в этой области, можно выделить Н.Н. Амосову, А.В. Кунина, И.И. Чернышеву.

Как уже упоминалось, построение любого семантического поля невозможно без выделения некой семы, которая была бы общей для рассматриваемого ряда слов. А.В. Кунин определяет семы, как микроэлементы значения, обозначающие реальные или воображаемые признаки денотатов [Кунин 1996: 187]. В нашем исследовании, посвященном рассмотрению фразеологических единиц с компонентами-соматизмами, главной семой выступает сема «телесность» в значении «то, что относится к телу/организму человека или животного».

Следует отметить, что характерной особенностью, наблюдаемой при построении лексико-семантических полей для фразеологизмов с компонентами-соматизмами, являются процессы метонимизации и метафоризации. Метафоры здесь выступают в качестве способа образной номинации. При помощи данного средства выразительности появляется возможность иносказательно дать оценку какому-либо процессу либо явлению действительности. Метонимический перенос же состоит

в том, что части человеческого тела характеризуют человека в целом. Таким образом, метафора и метонимия используются с целью придания образности и эмоциональности высказыванию.

## 2. Что такое «соматизм»? Проблема определения и классификации

Компоненты-соматизмы находят широкое распространение во фразеологическом пласте лексики. Еще Ф.О. Вакк в своих исследованиях отмечал, что соматические фразеологизмы относятся к одному из древнейших пластов фразеологии и являются наиболее употребляемыми среди прочих разновидностей фразеологизмов [Вакк 1964: 23]. Наибольший интерес к соматической лексике в целом характерен для периода конца XX – начала XXI века.

Этимологически термин «соматизм» восходит к греческому слову *σῶμα* [*sōma*] – «тело». Под соматической фразеологической единицей (СФЕ) нами понимается фразеологизм, одной из составных единиц которого является слово, обозначающее не только внешние физические формы (такие, как *рука, нога, голова* и проч.), но и элементы внутренней системы организма человека либо животного (например, *кровь, сердце, печень, мозг* и т.д.), а также мимику и жесты.

Само понятие «соматическая фразеологическая единица» трактуется лингвистами по-разному. Так, некоторые ученые не включают в семантическое поле СФЕ выражения с компонентами, обозначающими внутренние органы, жесты и мимику. Например, О. Назаров полагает, что соматизмами являются только те единицы, которые называют части тела человека [Назаров 1973]. В статье это понятие будет рассматриваться в более широком понимании, т.к., по нашему мнению, соматическая лексика не ограничивается словами, называющими лишь части тела человека либо животного, поскольку это приводит к необоснованному сужению семантического объема соматизмов.

Продуктивность фразеологии с компонентами-соматизмами легко объяснить, так как данное языковое явление носит ярко выраженный антропоцентрический характер. В процессе познания мира и собственного тела люди начинают соотносить себя с окружающей действительностью, приписывая различным явлениям природы «человеческие» свойства. «Поскольку в центре внимания человека находится он сам, то отсюда его постоянное стремление описывать окружающий мир по образцу и подобию своему. Языковой антропоморфизм является не пережитком первобытного мышления, ..., но общим законом развития средств номинации в языке» [Кунин 2005: 7]. Подобным образом во многих языках появляются различные меры длины, равные среднему размеру стопы или руки, например, в мусульманских странах *зира* – это мера длины, соответствующая половине локтя (в Иране принято обозначение *зар*).

Проблема классификации соматизмов актуальна и в настоящее время. Многие лингвисты предлагали свои варианты, например, Н.М. Шанский разделил соматизмы на три группы [Шанский 1985: 104]:

- 1) общие обозначения тела либо частей человеческого тела и названия, которые связаны со спецификой мужского и женского тела;
- 2) обозначения внутренних органов, обеспечивающих жизнеспособность и жизнедеятельность человеческого организма;

3) обозначения материи (субстанции) тела человека и его органов.

В настоящее время данная классификация представляется слишком общей и в то же время недостаточно полной. На современном этапе превалирует подход по конкретизации обозначаемых соматическими единицами объектов. На этом основании соматизмы принято разделять на шесть групп согласно классификации Р.Ю. Мугу [Мугу 2003: 81-82]:

1) сомонимическая лексика – единицы, относящиеся к телу человека/животного;  
2) ангионимическая лексика – слова, обозначающие компоненты кровеносной системы;

3) остеонимическая лексика – лексемы, имеющие отношение к костям тела;

4) спланхонимическая лексика – слова, обозначающие внутренние органы;

5) сенсонимическая лексика – единицы, которые называют органы чувств;

6) также рассматривается группа слов, служащих для номинации болезней.

Нами дополнительно была выделена группа соматизмов, служащих для описания жестов и мимики, поз и выражения лица человека, таким образом, количество рассматриваемых групп увеличивается до семи:

7) лексика, обозначающая жесты, мимику, позы и выражения лица человека.

Расширение классификации мы полагаем целесообразным по причине того, что подобная специальная лексика не входит ни в одну из вышеперечисленных групп, однако широко применяется на практике; а также имеет отношение к телу человека, но не содержит в своем составе компонент, называющий конкретный орган, хотя косвенно и имеет к нему отношение (например, *кивнуть* можно только головой; *подмигнуть* – глазом и т.д.). В этом смысле нам близка позиция Е.М. Верещагина и В.Г. Костомарова, которые полагают, что жесты, мимика, позы, выражения лица наряду с иными симптомами душевных состояний являются важными составляющими соматического языка, к которому они причисляют «любой значащий признак, положение или движение лица или всего тела человека» [Верещагин, Костомаров 1980: 24].

Рассмотренная выше классификация соматических единиц стала основой для построения лексико-семантического поля соматических единиц на материале фразеологического фонда персидского языка.

### 3. О лексико-семантическом поле СФЕ современного персидского языка

Совокупность лексических единиц, которые обозначают части тела, органы человека/животного и прочие проявления ранее упомянутого «соматического языка», может быть систематизирована и представлена в виде лексико-семантического поля (ЛСП). В данном исследовании анализируется ЛСП соматических фразеологизмов персидского языка в рамках широкого понимания термина «лексико-семантическое поле», предусматривающего включение в него словосочетаний (в данном случае фразеологических выражений).

Современный персидский язык изобилует соматизмами, а фразеологические единицы с компонентами-соматизмами в нем особенно распространены. Это не удивительно, ведь большинство рассмотренных в настоящей работе соматизмов входит в стословный список Сводеша и относится к базовой лексике, имеет развитую семантическую структуру и обладает разнообразными символическими значениями. При проведении настоящего исследования было выделено и рассмотрено бо-

лее тысячи фразеологических единиц современного персидского языка с компонентами-соматизмами. В основном мы рассматривали выражения и способы их употребления, указанные в работах Г.С. Голевой и Сейфа Гафари.

В рамках настоящего исследования нами предложена следующая структура ЛСП соматических фразеологических единиц современного персидского языка: выделяется ядро, околядерная зона, ближняя и дальняя периферии. При построении данной модели мы принимали во внимание семантику рассматриваемых единиц, возможность их употребления в переносном значении, а также сопутствующие им ассоциации.

Следует отметить, что при разграничении ядра и периферии любого ЛСП неизбежна некоторая субъективность в оценке ввиду недостаточной разработанности объективных критериев отнесения единицы к той или иной зоне. Лексические единицы были объединены по функционально-семантическому принципу, их расположение носит градуированный характер. Ключевым параметром к отнесению лексической единицы к разряду соматических, как уже упоминалось ранее, является наличие и выраженность в ней архисемы «телесность» в значении «то, что относится к телу/организму человека/животного».

Всего нами было проанализировано 1010 соматических фразеологических единиц персидского языка, которые были объединены в лексико-семантические группы (ЛСГ). Под лексико-семантической группой в настоящем исследовании понимается объединение слов одной части речи с общим основным компонентом значения.

Таким образом, в результате анализа мы выделили 5 лексико-семантических групп соматических фразеологических единиц современного персидского языка:

- 1) ЛСГ «внешние органы/части тела человека»;
- 2) ЛСГ «внутренние органы и системы организма» + волосы и ногти;
- 3) ЛСГ «части тела животных» (отличающиеся от частей тела человека);
- 4) ЛСГ «жесты и мимика»;
- 5) ЛСГ «болезни и продукты жизнедеятельности».

**Ядро.** К ядерной зоне ЛСП соматизмов мы отнесем лексемы, номинирующие внешние органы/части тела человека (ЛСГ №1), потому как в ядре располагаются нейтральные слова, наиболее полно и емко передающие идею поля, а также характеризующиеся как наиболее частотные и семантически максимально простые (например, *sar* – *голова*, *dast* – *рука*, *pā* – *нога*). В данную группу входит наибольшее число рассмотренных нами фразеологических выражений с компонентами-соматизмами, а именно 597 единиц.

Фразеологическая активность соматизма и его частотность употребления зависят от степени закрепленности концепта в общественном сознании и понимания обществом необходимости того или иного органа или части человеческого тела. Так, фразеологизмы с компонентами-соматизмами, называющими наружные органы чувств, при помощи которых человек познает окружающую действительность и свое место в ней, особенно распространены. Это объясняется тем, что такие части тела более доступны для наблюдения и описания, тогда как внутренние органы остаются скрытыми для непосредственного обозрения при помощи органов чувств.

Традиционно наиболее «активными» соматизмами в составе ФЕ персидского языка являются:

- *dast* «рука»;
- *sar* «голова»;
- *češm* «глаз»;
- *pā* «нога».

Они обладают ярко-выраженным символическим значением, что характерно не только для персидского языка, но и для многих других.

Согласно исследованиям М.А. Пеклера [Пеклер 1967: 15] и А.Д. Рахштейна [Рахштейн 1980: 116], компонент «рука» является наиболее продуктивным и частотным среди соматизмов в составе фразеологических оборотов в русском языке. А. Блюм, в свою очередь, исследовал соматические фразеологизмы современного английского языка и пришел к выводу о том, что лексема *hand* «рука» также является наиболее часто встречающейся [Блюм 2000: 1].

Следует отметить, что персидский язык здесь не исключение, и соматизм *dast* «рука» лидирует по частотности употребления. В результате проведенного исследования было отобрано 142 фразеологические единицы, имеющие в составе вышеупомянутой соматизм. Среди символических значений данного соматизма в составе ФЕ мы выделяем:

1) труд и профессиональные качества человека: *dast dāštan dar kāri* «быть мастером своего дела» (букв. «иметь руку в работе»);

2) власть: *dast az āstin darāvardan* «обрести силу» (букв. «вытащить руку из рукава»);

3) характеристика человека: *dast be dahān* «бедный, неимущий» (букв. «рука во рту»);

4) недостаток: *dast-e kasi be jā-yi band ast* «кто-либо разбогател, у кого-либо появился недостаток» (букв. «чья-либо рука привязана к месту»);

5) безысходность: *dast az jān šostan* «махнуть рукой на свою жизнь/ничего не ждать от жизни» (букв. «смыть руку с жизни»);

6) отказ: *dast bardāštan az kasi* «отказаться от кого-либо» (букв. «забрать свою руку от кого-либо»);

7) промедление: *dast be dast mālidan* «медлить» (букв. «тереть руки»).

Из вышеперечисленных примеров можно сделать вывод о высокой продуктивности и разнообразии символических значений соматизма *dast* «рука» в составе фразеологических оборотов современного персидского языка. Также довольно часто данный соматизм употребляется в составе «формул вежливости», которые зачастую носят иносказательный и образный характер: например, *dast-e čap be dast-e rāst mohtāj našavad* «пусть левая рука не нуждается в правой». Данное выражение является добрым пожеланием человеку ни в чем не испытывать нужды.

В результате исследования были также проанализированы остальные соматизмы, входящие в ЛСГ №1. Таким образом, количество выражений с компонентом *sar* «голова» насчитывает 93 единицы; с соматизмом *češm* «глаз» – 80 единиц, с лексемой *pā* «нога» – 50.

Среди прочих особенностей персидского фразеологического фонда можно выделить обилие фразеологизмов, состоящих из пары соматизмов вида:

- *dast-o pā* «рука и нога»,
- *češm- o del* «глаз и сердце»,
- *češm-o guš* «глаз и ухо» и т.д.

Например, *dast-o pā-ye hod rā gom kardan* «растеряться» (букв. «потерять руки и ноги»).

Также отмечается наличие одностилевых соматизмов-синонимов, продуктивных для образования фразеологизмов: например, «лицо» в персидском языке выражается двумя лексемами – *surat* и *ru*. Рассматривая ФЕ с данными соматизмами, мы приходим к выводу, что лексема *ru* встречается в составе соматических фразеологизмов чаще, однако, нами были выявлены выражения, где компоненты *surat* и *ru* являются взаимозаменяемыми:

*Ru-eš (surat-aš) az seyli sorx ast* «Он сохраняет достоинство несмотря на беды и несчастья» (букв. «Его лицо красное из-за пота»).

Подобная ситуация нехарактерна для русского языка, где единственным нейтральным по стилю вариантом является лексема *лицо*.

**Околоядерная зона** включает наименования внутренних органов и систем организма, а также волос и ногтей, т.е. относится к ЛСГ № 2. Единицы, составляющие околоядерную зону, также характеризуются конкретностью и широкой распространенностью, они выражают существенные признаки ЛСП. В результате исследования было рассмотрено 302 фразеологических единицы данной группы.

Следует отметить высокую продуктивность лексемы *del* «сердце» в образовании фразеологических единиц – нами было выявлено 136 фразеологизмов. Можно сказать, что это один из наиболее часто употребляемых соматизмов в составе фразеологических выражений:

*Del-e hod rā barā-ye čizi sābun zadan* «готовиться к чему-либо приятному, предвкушать что-либо» (букв. «намыливать свое сердце»).

Также употребляется в составе «парных» соматизмов:

*Del az dast dādan* «терять терпение» (букв. «упускать сердце из рук»).

Многие ученые в своих исследованиях большое внимание уделяли рассмотрению исключительной роли соматизма *del* в персидском языке. Прежде всего, стоит отметить тот факт, что помимо значения «сердце», данная лексема также называет «органы брюшной полости, живот». Отмечается, что *del* часто обозначает эмоциональный центр, что и служит обоснованием семантического перехода от первоначального значения лексемы «сердце» к значению «живот» [Руссо 2010]:

*Az del beravad har ānke az dide beraft* «С глаз долой из сердца вон» (букв. «Уйдет из сердца так же, как и ушло из взгляда»).

Таким образом, принимая во внимание многообразие символических значений лексемы *del*, а также широкую распространенность ФЕ с этим компонентом, мы полагаем целесообразным отнести группу выражений с соматизмом *del* к ядру настоящего ЛСП.

На **периферии** лексико-семантического поля находятся более сложные по семантическому признаку единицы, нередко они стилистически окрашены. Также периферийные лексемы в отличие от ядерных и околоядерных не могут употребляться в неограниченных сферах и обладают низкой степенью метафоризации.

К **ближней периферии** нами были отнесены лексемы, которые называют части тела животных, отличающиеся от компонентов человеческого организма



либо имеющие иное название (например, *коготь*, *хвост*, *лапа* и т.д.), входящие в ЛСГ № 3. Всего рассмотрено 44 фразеологических единицы. Например:

*Dom darāvardan* «расхрабриться» (букв. «достать хвост»);

*Dom-e kasi rā tu-ye bošgāb gozāstan* «льстить кому-либо» (букв. «класть чей-либо хвост в тарелку»).

**Дальняя периферия**, по нашему мнению, включает в себя микрополе «жесты и мимика», а также микрополе «болезни организма и продукты жизнедеятельности».

ЛСГ № 4 «жесты и мимика» насчитывает 35 единиц, большая часть из которых является глагольными фразеологизмами. Следует отметить, что единицы этой группы не обладают высокой продуктивностью, то есть одна лексема образует в основном всего одно-два фразеологических выражения. Например:

*Ada va atvār darāvardan* «вести себя неестественно» (букв. «доставать ужимки и кривляния»).

ЛСГ № 5 «болезни организма и продукты жизнедеятельности» представлена 32 фразеологизмами. Наибольшей продуктивностью среди слов, обозначающих болезни, обладает лексема *dard* «боль», с которой нами было отобрано 17 фразеологических единиц, а среди продуктов жизнедеятельности – *ašk* «слеза», восемь единиц:

*Dardi rā davā nemikonad* «это делу не поможет» (букв. «это болезни не излечит»);

*Ašk-e temsāh* «крокодиловы слезы».

В целом периферия ЛСП отличается разреженностью признаков и ослаблением семантических связей с лексемами ядра.

#### 4. Выводы

Рассмотрев лексико-семантическое поле соматизмов современного персидского языка в аспекте их участия во фразеологических сочетаниях, мы пришли к нескольким выводам.

Во-первых, мы предлагаем дополнить уже существующую классификацию соматизмов Р.Ю. Мугу и выделить в ней группу единиц, служащих для описания жестов и мимики, поз и выражения лица человека, также являющимися компонентами «соматического языка» по Верещагину и Костомарову. Таким образом, классификация соматических единиц по конкретизации обозначаемых объектов будет насчитывать семь групп. Подобное расширение обосновывается достаточным распространением единиц, обозначающих позы, мимику, жесты и т.д., для выделения ее в качестве отдельной группы. К тому же, такие слова имеют отношение к телу человека, и поэтому могут быть включены в лексико-семантическое поле СФЕ, пусть даже и будут располагаться в зоне дальней периферии.

Во-вторых, на основании анализа частотности употребления и выраженности семы «телесность», была построена модель лексико-семантического поля соматических единиц персидского языка, где мы выделили пять лексико-семантических групп: «внешние органы/части тела человека», «внутренние органы и системы организма + волосы и ногти», «части тела животных» (отличающиеся от частей тела человека), «жесты и мимика», «болезни и продукты жизнедеятельности».

В результате исследования к ядру ЛСП соматических фразеологических единиц была отнесена ЛСГ «внешние органы/части тела человека». Принимая во внимание высокое распространение, а также чрезвычайно разнообразное символическое значение, мы также полагаем необходимым отнести группу соматизмов с компонентом *del* «сердце» к ядру ЛСП. В околядерной зоне располагается ЛСГ «внутренние органы и системы организма + волосы и ногти». К зоне ближней периферии нами отнесена ЛСГ «части тела животных» (отличающиеся от частей тела человека), а зона дальней периферии включает в себя 2 ЛСГ: «жесты и мимика» и «болезни организма и продукты жизнедеятельности». Таким образом, отмечаются некоторые особенности языковой картины мира носителей персидского языка. В целом по сравнению с русским языком наблюдается большее распространение одностилевых соматизмов-синонимов. Еще одной характерной чертой фразеологического фонда современного персидского языка можно назвать всеобщее употребление выражений с парами соматизмов, т.к. в результате исследования было выявлено более 100 подобных примеров.

Соматические фразеологизмы, отражая в своей семантике длительный процесс развития культуры народа, фиксируют и передают от поколения к поколению культурные установки и стереотипы, эталоны и архетипы. Анализируя именно такие единицы, можно наиболее явно увидеть межкультурные отличия и специфику того или иного народа, а полевой подход в изучении фразеологических единиц является весьма продуктивным, так как он помогает выявить системную организацию языка. В целом дальнейшее изучение семантики и функционирования СФЕ представляется нам актуальным и перспективным.

### Источники

Голева Г.С. *Фарси-русский фразеологический словарь: Наиболее употреб. общенародные фразеологизмы, пословично-поговорочные выражения и афоризмы соврем. фарси. Около 6 000 фразеол. единиц.* М.: Грааль, 2000. 626 с.

Рубинчик Ю.А. *Персидско-русский словарь*, тт.1–2. М., 1983. 1664 с.

Ghaffari, Seif. *Farhange Estelahate Farsi be Englisi (A dictionary of idioms: Persian – English)*, Paykan, 1997. 452 p.

### Литература

1. Арнольд И.В. Лексико-семантическое поле в языке и тематическая сетка текста // *Текст как объект комплексного анализа в ВУЗе*. Л: ЛГПИ, 1984. С. 3–11.
2. Блюм А. *Семантические особенности соматической фразеологии*. М.: АСТ-пресс, 2000.
3. Вакк Ф. О. *О соматической фразеологии в современном эстонском литературном языке*. Автореферат дис. ... кандидата филологических наук. М., 1964. 29 с.
4. Васильев Л.М. *Современная лингвистическая семантика*. М.: Высшая школа, 1990.
5. Васильев Л.М. Теория семантических полей // *Вопросы языкознания*. 1971. № 5. С. 105–113.

6. Верецагин Е.М., Костомаров В.Г. *Лингвострановедческая теория слова*. М., 1980.
7. Кобозева И.М. *Лингвистическая семантика*. М., 2000. 352 с.
8. Кунин А.В. *Курс фразеологии современного английского языка*. Дубна: «Феникс+», 2005.
9. Мугу Р.Ю. *Полисемантизм соматической лексики (на материале немецкого и русского языков)*. Майкоп, 2003.
10. Пеклер М.А. *Русско-немецкая идиоматичность (на материале глагольно-субстантивных сочетаний, образуемых частотными переходными глаголами)*. Автореферат дис. ... кандидата филологических наук. Тбилиси, 1967.
11. Рахштейн А.Д. *Сопоставительный анализ немецкой и русской фразеологии*. М.: Высшая школа, 1980.
12. Руссо М. М. Локализация эмоций в языках мира // *Лингвистика и методика преподавания иностранных языков: периодический сборник научных трудов ИЯз РАН* (электронное научное издание). 2010. Вып. 2.  
URL: [http://www.ilingran.ru/library/sborniki/-for\\_lang\\_2010\\_02](http://www.ilingran.ru/library/sborniki/-for_lang_2010_02).
13. Шанский Н.М. *Фразеология современного русского языка*. М.: Просвещение, 1985.

## Новое в санскритской компьютерной лингвистике

**Марцис Юрьевич Гасунс**

независимый исследователь, Краснодар, Россия,  
gasyoun@ya.ru, ORCID 0000-0003-4513-884X

**Аннотация.** Корпусная лингвистика больше десяти лет тому назад стала стандартом для описания русской грамматики в рамках «Русской корпусной грамматики». Хотя первый санскритский корпус увидел свет в том же далеком 2010 г., а «Параллельный санскритско-русский корпус» спустя еще три года, санскритологи далеки от написания санскритской грамматики на основе корпуса. Однако уже известно, какой части крупных словарей достаточно для покрытия нужд корпуса (от 25000 до 91000 лемм). Упомянуты два направления из области юго-восточных OCR: с подключением грамматики и без.

**Ключевые слова:** санскрит, лексикография, параллельный корпус, OCR, Google Vision, древнеиндийский язык

**Abstract.** Russian corpus linguistics became the standard for describing Russian grammar within the framework of the «Russian Corpus Grammar» more than ten years ago. Although the first Sanskrit corpus was published back in 2010, and the «Parallel Sanskrit-Russian Corpus» three years later, Sanskrit scientists are still far from writing a Sanskrit grammar based on the data from corpus. But they have already found out which part of the dictionary is enough to cover the needs of the corpus (from 25.000 to 91.000 words). Two future paths of South Asian OCR are mentioned: involving grammar and without it.

**Keywords:** Sanskrit, lexicography, parallel corpus, OCR, Google Vision Old Indian language

### 1. Взаимодействие словаря, грамматики и корпуса

На конференции «Востоковедные чтения 2021. Сочетаемость языковых единиц и языковые модели. Памяти З.М. Шаляпиной (1946-2020)» (ИВ РАН, Москва, 14-15 апреля 2021 г.) после доклада «Параметрический анализ «Санскритско-русского словаря» В.А. Кочергиной» состоялся обмен мнениями, достойный послужить введением к моему обзору событий последних тридцати лет в компьютерной санскритологии. Позвольте процитировать реплику С.А. Крылова: «словарь и грамматика представляют собой непосредственно данность, по которой мы закономерно судим о языке, но пора добавить третий вид источников для санскритологии, а именно корпуса. Мы судим о языке по его словарям, грамматике и корпусам. Последние годы роль корпусов заметно возросла, теперь некоторые теоретики и вовсе отрицают значимость словарей и грамматик как источников информации. Что, конечно, представляет собой отчасти полемическую крайность,

но имеет свои основания. Дело в том, что создатели «Словаря языка Пушкина» во главе с Г.О. Винокуром и предыдущие филологи тоже фактически имплицитно опирались на какие-то аналоги корпусов, но эти корпуса существовали в их психике как в психике начитанных филологов. Дело в том, что всякий начитанный филолог имеет своего рода корпус в своем мозгу, это отмечал в последнее время А.Д. Шмелев, который на основании этого впадает в обратную крайность, а именно в отрицание значимости корпусов. С точки зрения реалистичности надо признать, что словари, грамматики и корпуса скорее дополняют друг друга, но не могут заменить».

И следом слова А.А. Кротова: «Корпусом являлись те тексты, которые указаны и расписаны до одного слова составителями словаря языка Пушкина. Поэтому все произведения, которые они взяли, являются корпусом этого словаря. Все словари опираются на корпус каких-то расписываемых текстов. Поэтому корпус это не нечто новое, это – полуфабрикат, это – предсловарь, несделанный словарь. Для параметрического анализа лексики корпусá на сегодняшний день оказываются непригодны по одной важной причине – еще не научились компьютерным образом достоверно извлекать синонимы из данных корпусов и распределять значения. Когда это будет сделано, тогда можно будет опираться не только на словари, но и на корпуса. На сегодняшний день такой возможности пока нет».

Есть два принципиально разных пути. Есть намерение имитировать размах крыльев птицы, а есть путь создания самолета. Мы не пытаемся имитировать Панини, потому что на сегодняшний день компьютер этого сделать не может. Разработки последних десятилетий привели нас в тупик. В частности это видно по работам американского профессора Питера Шарфа, ученика величайшего паниниста современности Дж. Кардоны. Его код закрыт. Двадцать лет спустя он даже сам не смог воспроизвести то, что уже работало прежде – полуавтоматическая разметка морфологии «Повести о Раме» из «Махабхараты». Поэтому мы не пытаемся воссоздать древнеиндийскую грамматику так, как она была создана. Мы пытаемся идти путем застывшей парящей птицы, то есть путем самолета. Это принципиально разные вещи.

Зачем компьютеру санскрит? На этот вопрос впервые в 1985 году в журнале «Artificial Intelligence» пытался ответить американский ученый Рик Бригс [Briggs 1985], но дальше его предположений о пригодности санскрита как основы для языка программирования будущего дело не пошло. Способы архивации, задействованные внутри санскрита и для описания санскрита испокон веков, остались невостребованными вне воображения сотрудника NASA.

Тогда спросим иначе. А зачем санскриту компьютер? Есть такие показатели, которые мы можем измерить хоть в рамках выписок из текстов, хоть в рамках электронного корпуса, но с совершенно разными трудозатратами. Например «формульность» в эпосе можно посчитать руками, а можно – используя возможности регулярных выражений и сведений о сочетаемости слов, посчитанных компьютером. Может ли компьютер понимать, читать и говорить на санскрите? На этот вопрос до сих пор нет удовлетворительного и однозначного ответа. Для того чтобы определить количество морфем в санскрите, ученые, начиная с пионера американской санскритологии У. Уитни [Whitney 2000], давно привлекают квантитативные показатели.

Так, например, в мировой практике санскритской лексикографии применение способов автоматизации работы с данными встречается еще крайне sporadically, а в области лексикографической типологии мы и вовсе делаем первые шаги. Однако есть области, в которых российская санскритология опережает мировые достижения. Мнение о мнимой отсталости русской индологии часто ошибочно. Проиллюстрируем данное утверждение несколькими примерами, добытыми в России.

## 2. Количество слов в санскрите

Сколько слов в санскрите? Известно, например, что в «Сводном списке слов словарей санскрита» 385 000 слов. Первоначально в 2014 г. мы составили конкорданс глагольных корней, а теперь завершили работу над сводным указателем санскритских суффиксов. И мы готовы ответить на столь простой, но тем не менее интересный вопрос. Соединяя 20 превербов и их сочетания с 850 корнями и 305 первичными и вторичными суффиксами, мы получаем почти все 385 000 слов. Столько «нормализованных», приведенных к единому способу записи слов можно найти во всех современных двуязычных и толковых санскритских словарях вместе взятых.

Рассмотрим частеречный срез на одном примере, на глаголах. Если опираться на словарь, то в санскрите около 10500 глаголов (согласно составленному на его основе «Сводному списку санскритских глаголов» с примерами из картотеки корпуса). С учетом сочетаний корней и приставок в словаре Monier-Williams (MW 2021) из пяти теоретически мыслимых сочетаний превербов самый длинный нам известный пример – это глагол *samabhīvyāhāra* «упоминание вместе» = *sam+abhi+vi+āñ+hr*, где *sam*, *abhi*, *vi* и *āñ* – четыре преверба, *hr* глагольный корень «брать».

Каждая словоформа в «Сводном списке санскритских глаголов» подкреплена санскритской цитатой. Например, начало словарной карточки корня *āp* выглядит следующим образом:

√āp 1. P.

to come to any one, to enter, to equal, to fall, to fall upon, to gain, to meet with, to obtain, to occupy, to overtake, to pervade, to reach, to suffer, to take possession of, to undergo

āpnomi – *āpnomi saphalāmllokāmstat karma brūta māciram* / MBh, 1, 220, 10.2 (Ind. Pr. 1. sg. √āp 1. P.)

āpnoti – *bhogena mokṣam āpnoti katham vadasi yogabhṛt* // MBhT, 3, 1.3 (Ind. Pr. 3. sg. √āp 1. P.)

āpsgplaḥ – *ātmanam ātmanā hatvā kiṃ dharmaphalam āpsgplaḥ* // MBh, 12, 7, 4.2 (Ind. Pr. 1. pl. √āp 1. P.)

āpnuvanti – [...] *te balavīryayuktāḥ samāḥ śataṃ jīvitam āpnuvanti* // AHS, Utt., 39, 162.2 (Ind. Pr. 3. pl. √āp 1. P.)

āpnuyām – *sarvajñānītyanutsargād bodhiṃ kiṃ nāpnuyāmaham* // BoCA, 7, 19.2 (Opt. Pr. 1. sg. √āp 1. P.)

āpnuyāḥ – [...] *no 'smān nirayānmokṣas tvam ca putrāpnuyāgatim* // MBh, 3, 94, 14.2 (Opt. Pr. 2. sg. √āp 1. P.)

āpnuyāt – *bhogena siddhim āpnoti bhogena mokṣam āpnuyāt* // MBhT, 3, 2.3 (Opt. Pr. 3. sg. √āp 1. P.)

āpnuyāma – *puṇyam evāpnuyāmeha pāpaṃ pāpopasevanāt* // MBh, 3, 1, 26.2 (Opt. Pr. 1. pl. √āp 1. P.)

āpnuyuh – *athāpnuyuh padaṃ śākram iṣtvā kratuśataṃ vidheḥ* // MrgTī, Vidyāpāda, 2, 11.2, 15.1 (Opt. Pr. 3. pl. √āp 1. P.)

āpnuhi – *dharmeṇa caturo varṇān pālayan kleśam āpnuhi* // Rām, Ay, 98, 57.2 (Imper. Pr. 2. sg. √āp 1. P.)

āpnotu – *nṛpo 'bravītsāñjalirāgatasprho yatheṣṭamāpnotu bhavānavighnataḥ* / BCar, 11, 72.1 (Imper. Pr. 3. sg. √āp 1. P.)

Список упорядочен по санскритскому алфавиту, и каждый приставочный глагол рассматривается как отдельная запись без указания на связь с исконным корнем. DCS 2021 опирается на словник словаря MW, отсюда и отсутствие связей приставочных корней с «голыми» корнями. Сбросить словарь с корабля современности пока остается недостижимой мечтой яростных сторонников чистых корпусов.

Полученная цифра вполне сопоставима с количеством глаголов в русском или современном немецком языке. Интересно, что не только количество глаголов сопоставимо, как показано выше, но, в продолжение [Гасунс 2014], по общей оценке А.А. Зализняка, «число корней может быть в разных языках различно, но чаще всего порядка 2 – 3 тысяч» [Зализняк 2008]. Списки корней, прилагаемые к грамматикам, в счет не идут – их количество обычно исчисляется несколькими сотнями, изредка одной или двумя тысячами. Исключением среди грамматик, понимаемых в широком смысле, может послужить, например, пятитомный справочник *Urasargārtha Candrikā* [Śāstī 1981], посвященный санскритским превербам (*urasarga*). Он заведомо полнее корпуса, даже через двести лет.

Даже корпуса? Далеко не все многообразие возможных сочетаний превербов представлено в корпусе и вряд ли когда-либо в обозримом времени будет. «Дополнительный том пополняемого санскритского словаря» (NWS 2021), завершённый в 2016 г., где под словом «словарь» подразумеваются Петербургские словари О. Бётлинга, «Большой» и «Малый», основан на 2590 источниках. Из них 171, в том числе вторичные, опубликованы за последние 140 лет и, собственно, составляют часть пополнений, расширений языка. Подобное количество источников в несколько раз превосходит количество привлечённых источников в любом полностью завершённом санскритско-санскритском или санскритско-двуязычном словаре и почти в два раза – в задуманном, но «незавершённом» словаре из Пуны (EDSHP 2021) с его 1500 заявленными первичными источниками, из которых тридцать лет сорок ученых делала выписки вплоть до 1976 года (года выхода первого тома) и которые еще три тысячи лет предстоит обрабатывать, таким образом сделав его самим долгим заведомо провальным словарным начинанием в истории человечества.

Корпуса имплицитно содержат сведения о частности и продуктивности, чего почти нет в грамматиках и о чем лишь косвенно можно судить из словаря. Заметим, что частотность, в частности, достаточно последовательно, хотя и не всегда верно, прописана в «Учебнике санскрита» В.А. Кочергиной, чего нет даже в некоторых куда более полных грамматиках. А полноценного частотного санскритско-

го словаря на основе корпуса все еще не существует. А что доступно уже сейчас? Мы знаем многое о глаголах, нам доступны десятки тысяч контекстов. Так, например, наибольшее количество примеров употребления из оцифрованной санскритской литературы и разных словоформ, засвидетельствовано для следующих глаголов: *kr*, *vac*, *bhū*, *as*, *gam*, *drś*, *da*, *brū*, *sthā*, *jan*, где корень *kr* не только самый частотный, но у него засвидетельствованы целых 250 финитных словоформ (из около 800 возможных) в отличии, например, от *añj* с 1 формой:

*karomi*, *kurmi*, *kurve*, *kṛṇomi*, *kṛṇve*, *kṛṇoṣi*, *karṣi*, *kuruṣe*, *kṛṣe*, *kṛṇuṣe*, *karoṣi*, *kṛṇute*, *kurute*, *kṛṇoti*, *karoti*, *kurvaḥ*, *kṛṇuthaḥ*, *kṛthaḥ*, *kurvāte*, *kurutaḥ*, *kṛṇutaḥ*, *kurmaḥ*, *kṛṇmaḥ*, *kurmahe*, *kṛṇmasi*, *karavāmahe*, *kṛṇmahe*, *kṛṇutha*, *kṛtha*, *kurutha*, *kṛṇvate*, *kurvate*, *kṛṇvanti*, *kurvanti*, *kurvīya*, *kuryām*, *kuryāḥ*, *kurvūthāḥ*, *kurvīta*, *kṛṇvīta*, *kriyāt*, *kuryāt*, *kuryātam*, *kuryātām*, *kuryāmaḥ*, *kurvīmahi*, *kuryāma*, *kuryāta*, *kurvīran*, *kuryuḥ*, *karavai*, *karavāni*, *kuru*, *kṛṇuhi*, *kuruṣva*, *kṛṇu*, *kṛṇutāt*, *kurutāt*, *kṛṇuṣva*, *kurutām*, *karotu*, *kṛṇutām*, *kṛṇotu*, *karavāva*, *karavāvahe*, *kurutam*, *kṛṇvāthām*, *kṛṇutam*, *karavāmaḥ*, *karavāmahai*, *karavāma*, *kuruta*, *kṛṇuta*, *kurudhvam*, *kṛṇotana*, *kṛṇota*, *kṛṇudhvam*, *kurvatām*, *kṛṇvantu*, *kurvantu*, *akaravam*, *akaroḥ*, *akṛṇoḥ*, *akṛṇuta*, *akuruta*, *akṛṇot*, *akarot*, *akurvahi*, *akṛṇutam*, *akurvātām*, *akṛṇutām*, *akurutām*, *akurmahi*, *kurma*, *akṛṇotana*, *akṛṇudhvam*, *akṛṇota*, *akṛṇvan*, *akṛṇvata*, *akurvan*, *akurvata*, *akurvanta*, *kariṣyāmi*, *kariṣye*, *kariṣyase*, *kariṣyasi*, *kariṣyate*, *kariṣyati*, *kariṣyāva*, *kariṣyāvah*, *kariṣyāvahe*, *kariṣyathaḥ*, *kariṣyataḥ*, *kariṣyete*, *kariṣyāmahe*, *kariṣyāmaḥ*, *kariṣyatha*, *kariṣyanti*, *akarīṣyam*, *akarīṣyah*, *kariṣyāḥ*, *akarīṣyata*, *akarīṣyat*, *kariṣyāma*, *kariṣyadhvam*, *kartāham*, *kartāsmi*, *kartāsi*, *kartā*, *kartārau*, *kartārah*, *akri*, *akah*, *akṛthāḥ*, *akṛta*, *askṛta*, *akṛtam*, *akartām*, *akrātām*, *akratām*, *akarma*, *akarta*, *akran*, *akrata*, *akaram*, *akarah*, *akarat*, *akaran*, *acīkṛthāḥ*, *acīkarat*, *acikaran*, *akārṣam*, *akārṣīḥ*, *akārṣīt*, *akārṣuḥ*, *kriyāsam*, *kriyāsmā*, *kriyāsuḥ*, *cakre*, *cakara*, *cakāra*, *cakriṣe*, *cakṛṣe*, *cakartha*, *cakrathuḥ*, *cakrāthe*, *cakrāte*, *cakratuḥ*, *cakṛma*, *cakra*, *cakruḥ*, *cakrire*, *kriyate*, *kriyete*, *kriyante*, *kṛṇavā*, *kṛṇavai*, *karavaḥ*, *kṛṇavaḥ*, *kṛṇavase*, *karavāt*, *kṛṇavat*, *kṛṇavate*, *karavat*, *kṛṇavāva*, *kṛṇavāvahai*, *karavāvahai*, *kṛṇavāma*, *kṛṇavāmā*, *kṛṇavāmahai*, *kṛṇavatha*, *kṛṇvaite*, *kṛṇavan*, *kṛṇavanta*, *karavan*, *kriyatām*, *kriyantām*, *akriyata*, *akriyanta*, *akāri*, *kriyeta*, *kriyeran*, *karam*, *kārṣīḥ*, *kṛṣīṣthāḥ*, *karoh*, *kārṣīḥ*, *kṛthāḥ*, *kah*, *kṛta*, *karot*, *karat*, *kārṣīt*, *kṛtam*, *karma*, *kārṣta*, *kartana*, *kṛdhvam*, *kurvan*, *kārṣuḥ*, *kran*, *kriyāma*, *kṛdhī*, *kṛṣva*, *kṛdhi*, *kartam*, *karatam*, *karatām*, *karta*, *cakriyāḥ*, *kṛṇavam*, *kuruthāḥ*, *kṛṇvata*, *cakaram*, *acakrat*, *acakriran*, *karā*, *karāni*, *karah*, *karasi*, *karase*, *karate*, *karati*, *karāthah*, *karathah*, *karataḥ*, *karāmahe*, *karāma*, *karan*, *karanti*.

Невооруженным глазом заметна некоторая корреляция между количеством словоформ и частотностью употребления в целом. Но более глубокие корреляции, как, например, про то, что глаголы распадаются на «две большие группы по способу образования презенса и перфекта» [Старостин 1985] компьютерным вычислениям недоступны, для обобщений требуется человеческий гений.

Независимо от того, кто именно автор словаря, важнейшие глаголы санскрита представлены во всех словарях и глоссариях, то есть субъективный фактор составителя словаря – фактор отнюдь не определяющий и порой заметно переоценен. Иначе бы словник П.С. Палласа не совпадал на две трети с базовым списком М. Сводеша (см. по этому вопросу [Кретов, Титов 2006], который, в свою очередь, опираясь опять же на совершенно субъективные факторы, имеет три редакции,



а не одну, единственно возможную. Не бывает бесстрастной, безличностной лексикографии.

Но «вольность» составителя словаря, согласно двадцатилетним наблюдениями над более чем шестьюдесятью двуязычными словарями разных семейств языков воронежского языковеда проф. А.А. Кретьова, сводится в рамках словника всего к 1/3 вокабул. То есть даже если человек, который составил словарь, вам, например, как личность глубоко неприятен, то он смог бы испортить в лучшем случае вряд ли заметно больше трети словаря. А так как в словаре В.А. Кочергиной прослеживается преемственность от канонических немецко-санскритских словарей, то следует признать, что часть претензий к нему совершенно надуманны, *ad hominem*.

Заметим, что работы с корпусом все равно ведутся через призму словарей. Иначе как знать, на какие единицы членить тексты? Не бывает девственных корпусов, где нет или никогда не было следов ни одного словаря.

В мире согласно скромным оценкам современных ученых насчитывается 30 миллионов санскритских рукописей, большинство из которых никогда не были не то, что изданы или отредактированы, но и никогда не будут оцифрованы. Соответственно работы хватило бы еще как минимум до конца следующего столетия. Нам нужен компьютер не столько потому, что мы, предположительно, стали глупее предшественников и наши познания уже не энциклопедичны. Компьютер не заменяет санскритолога, но без компьютера современный санскритолог будет тратить неоправданно, избыточно много времени на достаточно тривиальные операции. И, собственно, нередко вынужденно тратит, совершая мартышкин труд.

Однако вернемся к «Сводному словнику словарей санскрита», который в обработанном виде должен войти в разрабатываемый мной двухтомный «Обратный словарь санскрита» (7 и 8 тт. серии «Bibliotheca Sanscritica»). Если брать не «нормализованную» запись, а в том виде, в каком вокабулы записаны в самих словарях, получается всего 430 000 лемм в санскрите. Если же снять разности орфографии, то остаются уже озвученные 385 000 вокабул. За последние 14 лет в Кёльнских словарях группой энтузиастов с моим участием было исправлено порядка 10 000 опечаток, существенная доля из которых приходилась на заглавные слова. Лишь в 2021 г. нами начата планомерная сверка содержания самих статей, опираясь на разные вспомогательные алгоритмы. И этот список расписан по словарям. Семь лет как доступен сводный список санскритских слов, но кроме горстки разработчиков он, конечно, остается совершенно недооцененным. Помимо этого, лет двадцать как существует «Сводный список словоформ санскрита» – несколько сотен тысяч ячеек таблицы, калькуляция знаменитого парижского математика Жерара Юэ (HWNORM).

Конечно, словоформы Юэ порождены компьютером (на основе несколько тысяч строк кода) и соответственно содержат разные неточности, например, не учитываются известные даже из элементарных курсов грамматики исключения. Есть надежда, что еще через 25 лет ошибок будет значительно меньше, если генератор не будет вовсе заброшен, впрочем, доступ к исходному коду на этот раз приоткрыт. Жаль, что разные мимолетные университетские проекты в той же Индии, изобретая Yet Another Sanskrit Sandhi Engine, не учитывают разработки кол-

лег старшего поколения и набивают все те же шишки, занимаясь, в погоне за мнимой новизной, регенерацией несуществующих санскритских слов из года в год.

Последние двадцать строк «Сводного списка слов в санскрите» в транслитерации SLP1 выглядят следующим образом (заглавными буквами набраны аббревиатуры словарей согласно CDS 2021):

*hvAtavya: hvAtavya: AP, CAE, MW, PW, PWG*

*hvAtf: hvAtf: MW, MW72*

*hvAna: hvAna: BEN, CAE, MW, MW72, PW, PWG, SHS, STC, VCP; hvAnaM: AP90*

*hvAnam: hvAnam: AP*

*hvAnIya: hvAnIya: MW, MW72*

*hvAya: hvAya: MW, PW, PWG*

*hvAyaka: hvAyaka: AP, MW, PW, PWG*

*hvAyakIy: hvAyakIy: PW*

*hvAyakIya: hvAyakIya: MW*

*hvAyayitavya: hvAyayitavya: MW*

*hvAyin: hvAyin: MW, MW72*

*hvAra: hvAra: CAE, CCS, GRA, MD, MW, MW72, PW, PWG*

*hvArya: hvArya: CAE, CCS, GRA, MW, MW72, PW, PWG; hvAryaH: AP*

*hvAla: hvAla: CAE, CCS, MW, PW, PWG; hvAlaH: AP*

*hvf: hvf: AP, AP90, BEN, BOP, BUR, CAE, GRA, KRM, LAN, MD, MW, MW72, SHS, SKD, STC, VCP, WIL, YAT*

*hvfta: hvfta: MW, MW72*

*hvF: hvF: MW*

*hve: hve: AP, AP90, BEN, BOP, BUR, MW, MW72, SHS, SKD, VCP, WIL, YAT*

*hveY: hveY: KRM*

*hveya: hveya: MW, MW72*

SLP1 показала себя как наиболее надежная система транслитерации для хранения данных для санскрита, так как в отличие от НК (= «киотицы»), например, в SLP1 можно верно отразить редкие случаи зияния в санскрите без дальнейшей потери данных при обратной конвертации, например, в деванагари. Другой такой транслитерации нет.

### 3. Технологии распознавания санскритского текста

На сегодня самым продвинутым OCR алгоритмом для работы с брахми-производными письменностями является алгоритм Google Vision, сменивший на этом посту несколько лет тому назад Tesseract. Чем он хорош? Тем, что в Google в США работают много индийских программистов, около 2016 г. они сочли возможным улучшить степень распознавания санскрита в целом и деванагари в частности. Сайт SCR 2021 примечателен тем, что его основатель Мартин Глюкман (Martin Gluckman) из Южной Африки, послуживший инициатором данного углубленного обучения OCR алгоритмов, обычно платный сервис Google Vision представляет бесплатно в рамках своей образовательной лицензии. Иначе 3370000 распознанных картинок обошлись бы нам в сумму около 5000 американских долларов. То есть сперва обучил до того отстающий алгоритм, а потом сделал общедоступным, как Прометей.

Помимо этого у нас есть незаслуженно отодвигаемые отечественные разработки, которые можно было бы применить, например, к двадцати двум разновидностям всевозможных южноазиатских письменностей в Сериндийском фонде ИВР РАН. Это достаточно нетривиальная задача даже для более зрелых алгоритмов ввиду малого количество одинаково оформленного материала для обучения, но первичная помеха этому, помимо отсутствия финансирования, именно закрытость фондов.

Подходов несколько. Подход отечественного исследователя А. Строганова принципиально отличается от подхода Google Vision, так как он опирается на вспомогательные инструменты для достижения более высокого процента точности распознавания. То, что распознается, корректируется на сервере с применением грамматического анализа, но пока лишь удовлетворительно для тибетского языка.

Дальнейшая судьба компьютерной санскритологии связана с тем, что мы должны объединить усилия добровольцев с возможностями искусственного интеллекта. По отдельности ни те, ни другие не могут ничего или могут ничтожно мало. В 1985-м еще не было значимого компьютерного ресурса. Ресурс теперь появился, но мы его не используем должным образом, используя компьютер часто лишь как печатную машинку.

Относительно недавно с помощью добровольцев мы начали распознавать разные санскритско-русские глоссарии и словари. Шесть лет назад, например, был издан репринтом словарь О. Фриша, словарь малого объема, чуть больше 8150 слов, а теперь он доступен электронно. Издаваемые мною в рамках серии «*Bibliotheca Sanscritica*» словари полностью или частично оцифрованы и доступны публично (например, помимо изданных у нас, KEWA 2021). Партизанской предыстории индологии пора поставить точку. Есть задачи, где нужно действовать сообща, и благоприятное время уходит невозвратно.

Что такое частичная оцифровка? Разрезанным картинкам (= словарным статьям) присваивается словесный заголовок или выписываются их выборочные грамматические характеристики. Даже в такой компьютерной среде есть определенные преимущества, и это может дать нам принципиально новые инструменты, невозможные в бумажных оригиналах. Через двадцать пять лет после появления первой оцифровки словаря Monier-Williams мы лишь приближаемся к реализации высказанных еще в конце 80-х годов идей. Те самые «современные методы научного исследования, включая вычислительные машины», о которых востоковеды прежних поколений не смели и мечтать. Но кто из современных санскритологов может, например, составить регулярное выражение и обработать полученные данные?

Еще одна наша разработка, о которой хотелось бы упомянуть – режим *simple*, или упрощенный поиск, по словарям санскрита на сайте Кёльнского университета (CDS 2021). Если человек не знает, что он ищет, алгоритм *simple* пытается понять, что же на самом деле он мог иметь в виду. Не только в южной Индии люди записывают санскрит диковинными способами, но и у нас в русскоязычной среде находятся умельцы. Наша задача – придумать набор алгоритмов, чтобы поле ответа никогда не пустовало. И, соответственно, алгоритм поправит лишние буквы

или подставит недостающие. Как и прежде, сосуществуют различные системы транслитерации, некоторые, например, остались только в Индии, но, несмотря на это, довольно распространены – алгоритм *simple* поможет отгадать и их.

Мы делаем первые шаги, чтобы словарь, корпус и грамматику срастить вместе. Удачный пример воплощения озвученного подхода – Параллельный санскритско-русский корпус (PSRK 2021). Словарь получает совершенно новый функционал. Мы впервые можем добавить не только санскритско-русские отсылки, а через немецкий словарь Грассмана к Ригведе, например, выйти на английский перевод столетней давности, заново связав лемму и контекст, откуда она взята – добавив все доступные нам на сегодня оцифрованные переводы. А когда-то, возможно, сможем даже дополнить его французским переводом Л. Рену – во славу отечественного востоковедения.

Последнее достижение отечественных разработок было начато еще в 2013 г. RKG 2021 – это целиком Ригведа, большая часть Атхарваведы, Махабхараты и половина Рамаяны, целиком запланирован «Океан сказаний» Сомадевы. То есть впервые в мировой науке представлены крупные пласты разных хронологических срезов древнеиндийского языка вместе с переводами: санскрит ведийский, эпический и классический. Корпус столь крупного языка как санскрит (Grosskorpus-Sprache словами мэтра мировой санскритологии М. Майрхофера) не создается в одиночку, у нас в «Обществе ревнителей санскрита» десятки добровольцев и координирует их усилия наш бессменный Анатолий Артеменко из Одессы. Такого корпуса нет нигде больше.

### Источники

CDS 2021 – *Cologne Digital Sanskrit Dictionaries*

URL: <https://www.sanskrit-lexicon.uni-koeln.de/>. Дата обращения 15.07.2021.

EDSHP 2021 – *An Encyclopedic Dictionary of Sanskrit on Historical Principles*.

URL: <https://sanskrit-lexicon.uni-koeln.de/scans/PDScan/2020/web/>. Дата обращения 15.07.2021.

MW 2021 – *Monier-Williams Sanskrit-English Dictionary*.

URL: <https://sanskrit-lexicon.uni-koeln.de/scans/MW72Scan/2020/web/>. Дата обращения 15.07.2021.

PSRK 2021 – *Параллельный санскритско-русский корпус*.

URL: <http://samskrta.ru/parallel-corpus/>. Дата обращения 15.07.2021.

KEWA 2021 – *Kurzgefasstes etymologisches Wörterbuch des Altindischen*.

URL: <http://samskrta.ru/sanskrit-lexicon/KEWA/>. Дата обращения 15.07.2021.

SLS 2021 – *Сводный список санскритских глаголов*.

URL: <https://samskrta.ru/sanskrit-lexicon/sopasarga/>. Дата обращения 15.07.2021.

NWS 2021 – *Kumulatives Nachtragswörterbuch des Sanskrit*.

URL: <https://nws.uzi.uni-halle.de/>. Дата обращения 15.07.2021.

SCR 2021 – *SanskritCR*.

URL: <https://ocr.sanskritdictionary.com/>. Дата обращения 15.07.2021.

DCS 2021 – *Digital Corpus of Sanskrit*.

URL: <http://sanskrit-linguistics.org/dcs/>. Дата обращения 15.07.2021.

RKG 2021 – *Русская корпусная грамматика*.

URL: <http://rusgram.ru>. Дата обращения 15.07.2021.

HWNORM - *Headword normalization for Cologne dictionaries*

URL <https://github.com/sanskrit-lexicon/hwnorm1> Дата обращения 15.07.2021.

## Литература

1. Гасунс М.Ю. *Состав и строй древнеиндийских корней: история изучения*. Автореферат дис. ... кандидата филологических наук. М., 2014. 23 с.
2. Зализняк А.А. *О профессиональной и любительской лингвистике*. 2008.  
URL <https://goldrussian.ru/o-professionalnoj-i-ljubitelskoj-lingvistike.html>. Дата обращения 01.08.2021.
3. Кочергина В.А. *Санскритско-русский словарь*. / Под ред. В.И. Кальянова. С приложением «Грамматического очерка санскрита» А.А. Зализняка. М.: Русский язык; 1987 (1978).
4. Кочергина В.А. *Учебник санскрита*. М.: Филология; 1994. 336 с.
5. Кретов А.А., Титов В.Т. Алгоритм позиционного выявления синонимии // Вестник ВГУ. Серия «Системный анализ и информационные технологии». 2006. № 1.
6. Старостин С.А. Парадигматические типы древнеиндийского глагола // *Древняя Индия. Язык. Культура. Текст*. М., 1985. С. 272
7. Фриш О. *Санскритская хрестоматия*. В 2 т. Т. 2. Словарь / сост. О. Фриш; редактор В. Поржизка; автор предисл. и биогр. М. Ю. Гасунс. – М.: «АБВ»; 2015. (Репр. воспроизведение изд. / сост. О. Фриш. – Прага: Изд. АН Чехословакии; 1956). 355 с.
8. Briggs R. Knowledge Representation in Sanskrit and Artificial Intelligence // *AI Magazine*. Vol. 6. № 1. 1985. P.32
9. Śāstī cārūdeva. *Upasargārtha-candrikā* (= Лунный свет значений упасарг (превербов)). Delhi: Bhāratīya vidyā prakāśana, 1981. (In Sanskrit)
10. Whitney W.D. *The roots, verb-forms and primary derivatives of the Sanskrit language*. Dheli: Motilal Banarsidass; Reprint edition; 2000. 250 p.

**Структурно-семантические особенности  
глаголов «тянуть, тащить», «курить» и их деривативов  
в таджикском и шугнанском языках**

**Лейли Рахимовна Додыхудоева**

Институт языкознания РАН, Москва, Россия,  
leiladod@yahoo.com, ORCID 0000-0002-4567-9454

**Аннотация.** В статье рассматриваются примеры таджикских и шугнанских сложных и производных (префиксальные) глаголов, играющих ведущую роль в современных процессах словообразования и словоизменения, анализируются двучленные сочетания, компонентами которых исходно являлись или являются знаменательные слова, именная часть и глагол. Наряду с этим рассмотрению подвергаются устойчивые сочетания и фразеологизмы на основе образованных аналитическим способом сложноименных глаголов с опорным глаголом «тянуть, тащить», их структурный состав и семантическая составляющая, а также особенности их функционирования в современном таджикском и шугнанском языках. В данных сложноименных глагольных образованиях прослеживаются наиболее употребительные модели сочетаемости в зависимости от исходной лексической семантики и формальных свойств опорного глагола, взаимодействие конструкции с акциональными свойствами глагола, а также отчасти морфосинтаксические особенности сложных глаголов.

**Ключевые слова:** иранские языки, таджикский язык, шугнанский язык, сложноименные глаголы, производные (префиксальные) глаголы, глаголы курения

**Abstract.** The article examines examples of Tajik and Shughnani compound, derivative (prefixed) and causative verbs that play a leading role in modern word formation and inflection processes. We analyze compounds consisting mainly of two components (a nominal component and a verb) which were originally content words, or have remained so. In addition, we examine phraseological units based on the compound verbs 'to pull' and 'to smoke', their structural composition and semantic components, as well as the peculiarities of their functioning in the modern Tajik and Shughnani languages. In these compound verb formations, we trace the most common combinations of their components, depending on the initial lexical semantics and formal properties of the main verb. We also examine the interaction of the construction with the functioning and application of the verbs in question, as well as, in part, their morphosyntactic features.

**Keywords:** Iranian languages, Tajik, Shughnani, compound nominal verbs, derivative (prefix) verbs, lexicon, smoking

## **1. Сложноименные глаголы таджикского и шугнанского языков**

При структурном анализе таджикские и шугнанские глаголы дифференцируются на простые, производные (префиксальные) и сложные. Последние – аналитические формы – играют ведущую роль в современных процессах словообразования и словоизменения.

Объектом нашего исследования являются сложные глаголы, образованные аналитическими способами. Так, в таджикском языке сложный глагол состоит из двух и более лексических единиц, образующих одну синтаксическую единицу. В таджикском языке имеется достаточно широкий набор глаголов, которые могут выступать в качестве вершины сложного глагола. Это глаголы движения (*омадан* 'приходить', *даромадан* 'входить', *баромадан* 'выходить', *рафтан* 'идти, уходить' и др.), глаголы позиции (*нишастан* 'сидеть', *хобидан* 'лежать', *истодан* 'стоять'), а также глаголы *кардан*, *намудан* 'делать', *будан* 'быть', *шудан*, *гардидан* 'становиться', *дидан* 'смотреть', *гуфтан* 'говорить' и многие другие.

Данные глагольные образования представляют собой в основном двучленные сочетания, составной частью которых исходно являлись или являются знаменательные слова. Такими компонентами выступают в основном именная часть и глагол (иные сложные типы глаголов, такие как деепричастие и глагол и др., в данной статье специально не рассматриваются). В качестве первой части могут выступать различные части речи (существительные, прилагательные, наречия, а также имена с предлогами и междометия). Глагольные компоненты выполняют грамматические функции, при этом участвуют и в формировании общего значения. Наряду с этим во многих случаях функционировать самостоятельно вне состава сложного глагола может как именная часть (как семантически не привязанная к глаголу), так и глагол.

При анализе мы опираемся на структурно-семантическую типологию сложного глагола, разработанную для персидского языка Ю.А. Рубинчиком, который на основе семантических взаимоотношений компонентов (именной части и компонирующего глагола), а также сочетательных свойств глагола (валентность) выделил моделируемый и немоделируемый типы<sup>1</sup> [Рубинчик 1981: 153].

В статье рассматривается ряд сложноименных глаголов, образованных аналитическим способом, их структурный состав и семантическая составляющая, особенности их функционирования в современном таджикском и шугнанском языках, а также устойчивые глагольные сочетания и фразеологизмы на их основе.

К анализу лексики были привлечены данные словарей (Карамшоев 1988; 1991; 1999; ТРС 2006; ТСТЯ 2010 и др.), а также материалы, полученные в ходе полевой работы с носителями в Шугнани, Таджикистан.

## 2. Глаголы со значением «тянуть, тащить»; «курить» в таджикском и шугнанском языках

При структурно-семантическом разборе в таджикском и шугнанском языках глагола «курить, втягивать в себя (дым с помощью специального приспособления), в нос (растертый в порошок табак и др.), сосать (жевательную табачную смесь и др.)», было выявлено, что при передаче данного значения в обоих языках употребляется сложноименный глагол, в котором в качестве глагола-модификатора используется простой глагол «тянуть, тащить». Однако в таджикском и шугнанском языках это разные глаголы: в таджикском – *кашидан*, в шугнанском –

---

<sup>1</sup> Моделью в сфере фразеологии признается регулярность строения, образования, функционирования и понимания фразеологизмов.

*tīzdōw*. На этой основе в каждом из языков при включении именных и прочих компонентов образуется целый комплекс многообразных значений. Мы рассмотрели «словарные гнезда» (термин О. Семереньи) (1959) обоих простых глаголов в таджикском и шугнанском языках, включив большую часть образуемых на их основе на основе сложноименных глаголов.

При этом с таджикским *кашидан* ‘тянуть, тащить’ связаны глаголы, используемые в ряде памирских языков. Это ваханский, ишкашимский, сангличский, мунджанский и йидга: вах. *xaš-*: *xašt* ‘тянуть, тащить; носить; курить’, ишк. *čəlim xaš-* ‘курить’, *xaš-*: *xašəd* ‘тащить, тянуть, перетаскивать; стаскивать, снимать (одежду, обувь)’, сангл. *xā/aš-*: *xašə(d)* ‘тянуть, тащить; носить, курить’, мундж. *xaš-*: *xišk’* ‘тянуть, тащить; курить’, йд. *xoš-*: *xišč* ‘тянуть, тащить; курить’; тот же глагол использован и в ягнобском языке: *xaš-*: *xašta* ‘тянуть, тащить; курить; снимать (одежду, обувь)’ (Стеблин-Каменский 1999: 406). Ср. также: тадж. Вандж *nəxaš/kaš kardan* ‘тянуть, тащить волоком’ (Розенфельд 1982).

Однокоренной с таджикским глагол представлен и в шугнанском языке, однако с отчасти иным набором значений: *kirəxtow* ‘тащить, волочить; двигать, передвигать, раздвигать’. При этом данный глагол имеется и в других языках шугнано-рушанской группы, а также в язгулямском: *xəraχ-*: *xəraχt* ‘тянуть, тащить; волочить; курить (табак); класть под язык (табачную смесь)’. См. древнеиранское *\*<sup>1</sup>karš-*: *kṛš* ‘тянуть, тащить; волочить; проводить черту, борозду; пахать, обрабатывать землю’, ср.-перс. *krš-* ‘тащить, тянуть, натягивать’ (Эдельман 2011: 297-301). С этим же глагольным корнем связаны в шугнанском языке лексемы: *čəχt* (а также вах. *kyšt* и таджикско-бадахшанское *koum*) ‘заплечные вертикальные носилки для снопов, сена на основе двух деревянных жердей’ и шугн. *jīχc* ‘завязки в заплечных носилках; наплечники для укрепления корзины на плечах’ (там же: 305), хотя в данном приспособлении и применяется способ, отличный от волочения. А также шугн. *kigīxtow* ‘трудиться, возиться, принуждать, мучить’ (Эдельман 2009: 237-238).

Однако в значении *курить* в шугнанском языке употребляется глагол *tīzdōw* ‘тянуть, тащить’. Данный глагол – *tāž-*: *tīžd* ‘тянуть, тащить, натягивать, растащить’ – Г. Моргенстиерне возводит к *\*taga-* и связывает с перс. *tanjidan* (Morgenstierne 1974: 82). Ту же основу предполагает Дж. Чен: *\*tanč-* ‘to twist (together), become narrow, dense, constrict’, ср.-перс. *tnz-* ‘to strain, press’, *tng-* ‘to narrow’ (Cheung 2007: 377-378), ср. в этой связи шугнанско-баджувское *taxs/-*: *tixt* ‘дымить, чадить, коптить; вонять, дурно пахнуть’ (Карамшоев 1999: 50, 54) и руш. *taxs-*: *taxst* ‘то же’. По-видимому, к этому же гнезду примыкает и вах. *tanž/z-*: *tanž/zd* ‘укреплять основу (при тканье), натягивать паутину (о пауке)’ (Стеблин-Каменский 1999: 354)<sup>2</sup>.

Мы продемонстрируем структурные особенности и становление круга значений образований на базе данных глаголов в таджикском и шугнанском языках (в каждом из языков отдельно). При этом мы опираемся на отправное теоретичес-

<sup>2</sup> Ср. также бур. *task-’, task-íc-*, Йасин *tásk-* / *-ltásk-* (Tiffou 2014: 198, also *tásk-/-ltásk-*) ‘to draw, pull, drag; to draw a bow, to smoke (tobacco, opium)’ и рус. *taskat’, taščit’* ‘to draw, pull’, чешское *tasiti*, Pl *taskáč, taszczyč* id (Čašule 2017: 14)



кое положение, что валентность есть явление глубинного, логико-семантического плана, являющееся основой для отражения некой ситуации («содержательная» валентность) и имеющее конкретную языковую реализацию («формальная» валентность) (о разработке понятия см. [Домброва 2017: 24], а также [Шаляпина 2007, гл. 5]).

### 3. Глагол *кашидан* в таджикском языке

Простой глагол тадж. *кашидан* ‘тянуть, тащить’ используется в качестве глагола-модификатора при образовании группы сложных глаголов и устойчивых сочетаний.

На этой основе при включении именных компонентов образуется целый комплекс значений, начиная от ‘тянуть, тащить, протягивать’, ‘притягивать, привлекать (внимание, к кому-/чему-л.)’, ‘снимать / надевать, натягивать (одежду)’, ‘взвешивать, впитывать, всасывать’, ‘накладывать (пищу)’, ‘разливать (чай)’, ‘извлекать, сбивать (масло)’, ‘рисовать, чертить’ до ‘длиться, продолжаться; испытывать, претерпевать (разного рода эмоции и состояния)’ и ‘курить’ (ТРС 2006). С этим же глаголом связана традиционная в Средней Азии народная конная игра, представляющая собой спортивное состязание, борьбу за тушу козла – *бузкаши*, из тадж.-перс. *буз* ‘козел’ и *каш* осн. наст. вр. глагола *кашидан* ‘тянуть, тащить’, тадж.-юж. *буз кашидан*, *бузкаши* (*кӯбкорӣ*) *кардан* ‘играть в конную игру «перетягивание козла», «козлодрание», *пийодакашиак* ‘пешее перетягивание’ (пешее состязание за тушу козла) (СЮГТЯ 2012).

Приведем примеры глагольных образований, выделяя группы на основе отдельных характерных для них структурно-семантических признаков.

3.1 В первой группе приведены глагольные образования, выделенные на основе вхождения в их состав соматизмов и понятий, связанных с жизнедеятельностью человека, см. Таблицу 3.1.

Таблица 3.1. Образования с глаголом *кашидан* на основе соматизмов

Таджикский язык	Русский перевод
(тан) дароз кашидан	лежать, вытянувшись; затягиваться (о времени)
кад кашидан	расти
даст кашидан	отдернуть руку; <i>пер.</i> отказываться от чего-л., отступаться; тадж.-юж. крутить (СЮГТЯ 2012)
тангдастӣ кашидан	испытывать нужду
(дар) ба бар кашидан	обнимать
пой кашидан (аз чое)	прекращать посещение чего-л.
аз по кашидан	разделить вино (с кем-л.)
гӯш кашидан	тащить за уши; прислушиваться, раскаяться
аз гӯш(и касе) кашидан	тащить, тянуть кого-л. за уши; склонять
сар кашидан	подниматься, тянуться вверх, расти, колоситься (о злаках)
(аз чизи) сар кашидан	превысить пределы, вытянуться вверх, <i>пер.</i> выйти из повиновения
сар ба фалак кашидан	упираться вершиной (головой) в небо; <i>пер.</i> заноситься

рӯ ба пӯстак кашидан	<i>букв.</i> натягивать на лицо кожу, <i>пер.</i> бесстыдно смотреть в глаза другим после проступка
ба рӯи худ чарм кашидан	<i>букв.</i> натягивать на лицо кожу, <i>пер.</i> преодолевать, преодолевать стыдливость
чарм ба рӯи худ гарм кашидан	<i>букв.</i> натягивать на лицо кожу, <i>пер.</i> преодолевать стыдливость, не иметь стыда
забонкӯтоҳӣ кашидан	держать язык за зубами, молчать
худро чет кашидан (аз чизе)	отстраняться, самоотстраняться от чего-л.
ба худ кашидан	притягивать, привлекать к себе; <i>пер.</i> впитывать, втягивать в себя
дил кашидан	тадж.-юж. иметь склонность, желание (СЮГТЯ 2012)
дили маймун гуштбирён кашидан	иметь несбыточное желание

3.2 Следующая группа сложноименных глагольных образований выделяется на основе передачи значений разного рода произвольных и произвольных физиологических действий (активно производимых многоактных, состоящих из нескольких однородных актов), состояний и рефлекторных реакций живых существ, см. Таблицу 3.2.

Таблица 3.2. Словоименные сочетания с глаголом *кашидан*, передающие многоактные физиологические действия

<b>Таджикский язык</b>	<b>Русский перевод</b>
нафас кашидан	дышать; передохнуть
нафаси чукур кашидан	глубоко вдохнуть
оҳ кашидан	вздыхать
дам кашидан; як дам кашидан	пить глотками; выпить залпом
хӯрт кашидан	глотать, пить глотками
осо кашидан	зевать
хамёза кашидан	зевать; потягиваться
хуррок кашидан	храпеть, всхрапывать
наъра/сайха кашидан	реветь, рычать; вопить, кричать
нула/уллос кашидан	выть, завывать
шиха кашидан	ржать (о коне)
чор кашидан	кричать изо всех сил, кричать во всеуслышание
хуштак/ чуррак кашидан	свистеть, посвистывать
ханчол кашидан	царапать, скрести
куллоб кашиди нӯшидан	выпить одним глотком, залпом
дам фурӯ кашидан дам ба дарун кашидан	замолкать; затаить дыхание

3.3 Еще одна группа сложноименных глагольных образований выделяется на основе передачи значений разного рода эмоциональных и физиологических проявлений состояний с оттенком длительности, которую придает глагол *кашидан*, см. Таблицу 3.3.

Таблица 3.3. Сложноименные сочетания с глаголом *кашидан*, передающие многоактные физиологические действия

Таджикский язык	Русский перевод
дарл/азият <i>кашидан</i>	мучиться, переносить страдания, боль
касал <i>кашидан</i>	болеть, переносить боль, страдания, мучиться
забунї <i>кашидан</i>	изнемогать от слабости, терпеть притеснения
хавогир <i>кашидан</i>	беспокоиться, тревожиться, волноваться
ташвиш <i>кашидан</i>	беспокоиться
фиروق <i>кашидан</i>	переносить разлуку
бори ор <i>кашидан</i>	испытывать муки совести
ваҳшат <i>кашидан</i>	бояться, пугаться, испытывать страх
азоб <i>кашидан</i>	испытывать трудности; мучиться, страдать
азоби рӯҳї <i>кашидан</i>	испытывать муки, страдать, нервничать
озор <i>кашидан</i>	терпеть обиду, испытывать боль, мучение
танҳой <i>кашидан</i>	быть одиноким
абру боло <i>кашидан</i>	поднимать брови (при изумлении)
пешонї дарҳам <i>кашидан</i>	хмуриться; дуться; выражать неудовольствие; проявлять недовольство
дар гӯшаи <i>кашидан</i>	удалиться в уголок, <i>пер.</i> дистанцироваться, воздержаться

3.4 В следующей далее подгруппе сложноименных глагольных образований выделяется эмоциональная составляющая, длительные по времени эмоциональные состояния с переносным значением «переносить, испытывать, претерпевать, терпеть», см. Таблицу 3.4.

Таблица 3.4. Образования с глаголом *кашидан*, передающие значение «длиться, продолжаться» – «терпеть, переносить, испытывать, претерпевать»

Таджикский язык	Русский перевод
тангї/тангдастї <i>кашидан</i>	испытывать нужду, находиться в тяжелом материальном положении
мушкилї <i>кашидан</i>	испытывать трудности
захмат <i>кашидан</i>	испытывать трудности; беспокоиться, утруждать себя
захмати беҳуда/беҳуда <i>кашидан</i>	понапрасну утруждать себя; заниматься бесполезным делом; трудиться напрасно
ранч <i>кашидан</i>	претерпевать, испытывать трудности
ситам <i>кашидан</i>	терпеть притеснение, угнетение; переживать несчастье, испытывать горе
чафо <i>кашидан</i>	испытывать несправедливость и притеснение; мучиться, страд

хорй кашидан	быть униженным, оскорбленным, испытывать притеснение
изо кашидан	стыдиться, смущаться
ҳичолат кашидан	стыдиться, испытывать смущение
дахшат кашидан	тадж.-юж. стыдиться
ташвиш кашидан	беспокоиться, волноваться; заботиться;
ҳасрат кашидан	печалиться, страдать, тосковать; сожалеть; <i>разг.</i> завидовать

3.5 В данной подгруппе глагольных образований выделяется эмоциональная составляющая, при которой подчеркнуты значения эмоционального сближения, привлечения, очарования, побуждения к действию, реакции, пробуждения заинтересованности, см. Таблицу 3.5.

Таблица 3.5. Словоименные сочетания с глаголом *кашидан*, передающие многоактные физиологические действия

Таджикский язык	Русский перевод
ба худ кашидан	впитывать, втягивать в себя; притягивать, привлекать к себе
диққати касеро ба худ кашидан	привлекать к себе чье-л. внимание
ба ... кашидан; ба чазо / чавобгарӣ кашидан	привлекать к кого-л. к чему л.; привлекать к ответственности, наказывать

К этой подгруппе примыкают фразеологизмы на основе опорного слова *кашол* ‘длительный, продолжительный’ (образования от глагола *кашидан*): *бо дили кашол* ‘охотно, с радостью’, а также сочетания на основе словоименного глагола *кашол будан* ‘иметь склонность к чему-л.’ – *дил кашол будан* ‘иметь желание, склонность к чему-л.’ и *дил(и касе) кашол шудан* ‘чье-л. сердце предрасположено’ в значении ‘очень желать чего-л.’. Сюда же примыкают новообразования *дилкашол* ‘склонный, расположенный к кому-л., чему-л., испытывающий склонность, влечение’ и *дилкашолӣ* ‘склонность, влечение’, *дилкашолӣ доштан* ‘испытывать склонность’.

3.6 При ближайшем рассмотрении понятия *куруение* в целом в этнокультурном аспекте и глагола ‘курить’ в таджикском и шугнанском языках выяснилось, что данный феномен можно разбить на несколько составляющих, базирующихся на способе употребления табака или курительной смеси. И для таджикского, и шугнанского языков можно выделить такие традиционные этнокультурные феномены, как *курить калян* или *курить сигареты*, а также *сосать*, помещая за щеку, под язык (табачную жевательную смесь (*нос*) и т.п. Кроме того, в понятие *куруение* в целом включается и такой способ, как *нюхать табак*. При этом в других иранских языках, например, в осетинском, в значении *куруить* отмечается глагол *damun* : *dymd* : *dund* ‘дуть, надувать, исчезать, курить’ (Абаев 1958: 382), связанный с др.-ир. <sup>10</sup>*dau-*: *dū* ‘дымить, окуривать, воскурять (благовоние)’ (Расторгуева, Эдельман 2003: 396, см. также 317).

Традиционный курительный прибор, в прошлом и ныне широко распространенный от Китая до Африки (персидское *чилип*), состоит из сосуда, наполняемого водой (*нарғил* ‘кокос(овый орех)’), укрепленной в его горлышке трубки (*чубу/ак*)

с табаком (нижний конец которой опускается в воду) и из шланга или деревянной трубки, введенной в сосуд выше уровня воды (дым охлаждается, и курильщик через нее тянет дым, фильтруемый сквозь воду). Основание кальяна (наргиле) делают из плодов тыквы, кокосового ореха и др., что послужило названию прибора, а способ употребления курительной смеси стал основой для сложения глаголов, обозначающих данное действие, во многих языках.

Было выявлено, что при передаче данных значений в обоих языках – при всех способах употребления табака (или иных курительных смесей) – используется сложноименной глагол, в котором в качестве глагола-модификатора выступает простой глагол «тянуть, тащить», передающий значение «тянуть, втягивать (в себя), вбирать (через нос), впитывать (через рот) и т.п.» (тадж. *кашидан*, шугн. *tīzdow*) без различия специфического способа употребления (курить, нюхать, жевать), см. Таблицу 3.6.

В разговорном варианте таджикского языка и южных говорах таджикского языка представлены и другие обозначения процесса курения, не связанные с глаголом *кашидан*: *чокидан* (ТРС 2006; СЮГТЯ 2012).

Таблица 3.6. Сложноименные глаголы с опорным глаголом *кашидан*, передающие значение «курить»

Таджикский язык	Русский перевод
тамоку <i>кашидан</i> тамокукаш тамокукашӣ	курить табак курильщик курение
тамокуро <i>қуллоб</i> <i>кашидан</i>	курить <i>взатыжку</i> , <i>затыгиваться</i> ( <i>қуллоб</i> <i>затыжка</i> )
папирос <i>кашидан</i> папироскаш <i>папироскашӣ</i>	курить папиросы; курить курильщик курение
сигор <i>кашидан</i> сигоркаш сигоркашӣ	курить сигареты; курить курильщик курение
қалён <i>кашидан</i>	курить кальян ( <i>қалён</i> <i>наргиле</i> ; кальян)
чилиим <i>кашидан</i> чилимкаш чилимкашӣ чилимкаду	курить кальян курильщик чилима, кальяна курение чилима, кальяна тыква-горлянка, табакерка
нос <i>кашидан</i> носкаш носкашӣ	употреблять <i>нос</i> (табак, закладываемый под язык) человек, употребляющий <i>нос</i> употребление <i>носа</i>
тарёк <i>кашидан</i> тарёккаш тарёккашӣ	курить опиум наркоман, курильщик опиума курение опиума, наркомания
нашға <i>кашидан</i> нашғакаш нашғакашӣ	принимать ананшу; <i>пер.</i> пьянеть наркоман; ананшист, опьяненный, хмельной наркомания; опьянение
кашиданӣ	порция табака на одну <i>затыжку</i>

#### 4. Производные и понудительный глаголы на основе *кашидан* в таджикском языке

В дополнение мы рассмотрим также производные (префиксальные) глаголы *баркашидан* ‘взвешивать, измерять, извлекать, изображать’, *даркашидан* ‘втягивать (жидкость); натягивать, вынимать’, в которых семантическая составляющая уточняется при помощи префиксов с пространственно-направительным значением *бар-* и *дар-*.

В качестве глагольного префикса *бар-* указывает в целом на движение вверх, изнутри наружу (*баркашидан* ‘вытаскивать’), в сторону, а также на сталкивание, сближение предметов [Расторгуева, Керимова 1964: 16]<sup>3</sup>. В современном таджикском языке префикс *бар-* по большей части используется в значении установления, введения чего-либо и считается непродуктивным. Впрочем, при образовании компьютерных терминов префикс *бар-* проявляет весьма высокую продуктивность. Глаголообразующий префикс *дар-* в таджикском языке указывает в основном на направление движения внутрь. Однако В.С. Расторгуева и А.А. Керимова отмечают, что в настоящее время эти префиксы являются неотделимой частью глагольных основ и их префиксальный характер не проявляется (там же).

К этой группе примыкает производный глагол *фуру кашидан* ‘глотать, поглощать, засасывать, затягиваться при курении’, образованный при помощи префиксальной составляющей, однако здесь автономность морфемы все еще осознается, что передается на письме отдельным написанием составляющих компонентов, см. Таблицу 4.1

Таблица 4.1. Производные префиксальные глаголы и фразеологизмы на их основе на базе глагола-модификатора *кашидан*

Таджикский язык	Русский перевод
баркашидан	взвешивать, измерять; извлекать (из чего-л.), закрыть, изображать
баркаш	взвешивание; отвешивание
дар тарозу баркашидан	взвешивать на весах
дар тарозуи акл баркашидан	<i>пер.</i> хорошо обдумать (взвесить на весах разума), продумать

<sup>3</sup> В качестве глагольного префикса *бар-* указывает: 1) на движение вверх: *бархостан* «вставать», *бардоштан* «поднимать», в то же время в переносном значении *барангехтан* «подстрекать», «возбуждать», *барафрухтан* «зажигать»; 2) на движение изнутри наружу: *баркашидан* «вытаскивать», *баркандан* «вытаскивать», «выдергивать». Оба значения – движения вверх и наружу – объединяются в глаголах: *баромадан* «выходить» и «подниматься вверх», «восходить», *баровардан* «выносить», «выводить» и «поднимать вверх»; 3) на движение в сторону или назад, обратно: *баргаштан* «поворачиваться» (в сторону за угол), «возвращаться», *баргардондан*, *баргардонидан* «возвращать», *бардодан* «отдавать», «возвращать» (взятое); 4) на сталкивание, сближение предметов: *бархурдан* «сталкиваться», «наталкиваться», «встречаться», *барбастан* «связывать», «привязывать» [Расторгуева, Керимова 1964: 16]. Префикс *бар-* проявляет весьма высокую продуктивность при образовании компьютерных терминов.

гапро ба тарозу баркашидан	взвешивать слова на весах, <i>пер.</i> говорить обдуманно, взвешенно
даркашидан	выпивать, втягивать (жидкость); натягивать, вынимать; преподносить (подарок)
дарёдаркаш	<i>миф.</i> великан, силач
курта даркашидан	надеть платье
дам даркашидан	хранить молчание, промолчать
фуру кашидан	глотать, поглощать, засасывать, затягиваться (при курении)

Здесь же отметим наличие в таджикском языке понудительного глагола *кашон(и)дан* ‘переносить, перетаскивать; нанести, натащить; навезти’ и новых сложноименных глаголов на его основе – *кашон-кашон бурдан* ‘перемещать, переносить, таскать’, *кашонда даровардан* ‘ввозить (продукцию)’, а также термин на основе деепричастия *оби кашонда* ‘откачиваемая вода’.

О сложении поздних суффиксальных каузативов в таджикском: *kašon-* : *kašond-* ‘перетаскивать; заставлять таскать’, связанных с классическим перс. *kaš-* : *kašīd-* ‘тащить, тянуть’ см. (Эдельман 2011: 299). Ср. однако точку зрения, представленную в современном таджиковедении, сложившуюся, вероятно, на основе синхронного подхода к деривации<sup>4</sup>.

### 5. Глагол «тянуть, тащить; курить» в шугнанском языке

В данном разделе мы подвергаем разбору сходный по значению шугнанский глагол *tīždōw* ‘тянуть, тащить’, на основе которого аналитически при включении именных компонентов образуется целый комплекс значений ‘носить, таскать, возить; снимать/ натягивать (одежду); втягивать, впитывать; чертить, рисовать; взвешивать, вешать; извлекать (масло)’, а также ‘курить’; он же выступает как вспомогательный глагол *tīždōw* ‘тянуть, тащить’ (Карамшоев 1999). В шугнанском языке представлен практически весь набор значений таджикского сложноименного глагола, см. Таблицу 5.1.

Таблица 5.1. Сложноименные глаголы с опорным глаголом *tīždōw* ‘тянуть, тащить’

Шугнанский язык	Русский перевод
arā tīždow	пилить
o(h) tīždow	охать, вздыхать
o(h)i sar tīždow	испускать глубокий вздох, переживать

<sup>4</sup> «Производные глаголы в современном таджикском языке образуются: а) путем прибавления суффиксов *-идан*, *-онидан* к основам существительных или прилагательных (суффикс *-он* образует основу настоящего времени глагола, суффикс *-ид* образует основу прошедшего времени глагола, суффикс *-ан* – суффикс инфинитива): *кучидан*, *паридан*, *расидан*, *кашондан*, *кашидан*, *хазидан*, *давидан*, *гузарон(и)дан*, *гуселонидан*, *кучон(и)дан*» (Шарипова 2019: 50).

nafas tīždow	делать вздох, передыхать
dām tīždow	делать вздох, глоток, пить залпом
qīžak tīždow	ныть, плакать, <i>пер.</i> раскаиваться
lajūm tīždow	сдерживать удила, <i>пер.</i> сдерживать порыв
tīrāng tīždow	затягивать подпругу, оказывать давление на кого-л.
peḥ tīždow	гнать(ся), преследовать
daw tīždow	передвигаться перебежками, скачками, бегать
tasmā tīždow	вырезать ремень
čārḫ tīždow	точить
tangdasti tīždow	голодать, букв. влачить бедность, жить в бедности
foqā tīždow	голодать, жить в бедности, нищете
ranj tīždow	испытывать трудности, мучиться
rā(g) tīždow	мучить, тянуть жилы
woḥ tīždow	переносить, возить сено
woḥ-tīžd	переноска, перевозка сена
woḥ-tāžij	переносчик, перевозчик; предназначенный для перевозки
đūy/ruyan tīždow	сбивать пахтанье, извлекать масло
molā tīždow	боронить
ḫac tīždow	носить/возить воду
ḫac tāžij	водонос, поливальная машина
wīz tīždow	переносить, перевозить груз
wīz tīžd čōḫ(t)	переноска, перевозка груза
wīz-tāžij/kaḫ	носильщик, перевозчик груза, грузовая машина
čōḫ(t)-wīz tīždow	переносить снопы на носилках на спине
čōḫ(t)-wīz tīžd	переноска снопов на носилках на спине
čōḫ(t)-wīz tāžij/ čōḫ(t)-wīz kaḫ	переносчик снопов на носилках на спине
kaḫt tīždow	возить зерно(вой хлеб)
počtā tīždow	доставлять почту
počtā-tāžij	почтальон
līlen tīždow	снимать/ натягивать (одежду)
kurta ar xu tīždow	надевать на себя платье
abrā tīždow	обтягивать
ḫāt(ak) tīždow	проводить, черту, чертить, рисовать
naqšā tīždow	чертить, рисовать; готовить проект, планировать
sūrat tīždow	рисовать, писать (картину)
sūrat-tāžij	художник
wīz tīždow	взвешивать, вешать
put tīždow	взвешивать, вешать
qiloč tīždow	измерять маховой саженью (в обхват, в размах двух рук)
ḫac(-ta pi xu) tīždow	впитывать воду
tar piro xu tīždow	продвигаться, выступать вперед, <i>пер.</i> брать на себя ответственность
tar zibo xu tīždow	отступать назад, <i>пер.</i> не брать на себя ответственность



jûn tîždow	тянуть, вытягивать душу, <i>пер.</i> приставать, мучать
purtak tîždow	насмехаться, <i>пер.</i> приставать
may tîždow	вырываться вперед
guj tîždow	играть в игру «перетягивание козленка», <i>пер.</i> набрасываться, «разрывать»
guj-tîžd	«перетягивание козленка» (игра)
subtanob tîždow	исчезать, удирать ( <i>subtanob</i> ‘рано утром’)
duy tîždow	извлекать, сбивать масло

Большая часть примеров на сложноименные глаголы здесь является сложениями из заимствованных из таджикского языка имен и исконного опорного глагола, причем все сочетание образовано по таджикской модели, например: *nafas tîždow*, *tangdasti/foqā tîždow*, *naqšā tîždow*, *qiloč tîždow* и ряд других, а также имена деятеля: *sûrat-tāžij* ‘художник’; здесь же можно упомянуть сочетания, образованные по таджикской модели с русскими заимствованиями, например, *počtā tîždow* ‘доставлять почту’, *počtā-tāžij* ‘почтальон’.

В шугнанском языке наблюдаются и случаи частичного калькирования сложноименных глаголов таджикского языка, когда сложноименной глагол в целом складывается по таджикской модели, а именная часть представлена переводом таджикского существительного: *ħac tîždow*, *woħ tîždow*, *peħ tîždow*, *tar piro/zibo xi tîždow* и др. Встречаются случаи калькирования сложноименных глаголов таджикского языка, при котором глагольное образование строится по таджикской модели, но на основе исходных составляющих со сдвигом в значении именной части, ср., например, шугн. *guj-tîžd* и *guj tîždow* и тадж. *бузкашиӣ*, тадж.-юж. *буз кашидан*. В шугнанском языке именная часть представлена именем молодого животного – *guj* ‘козленок’ при наличии лексемы для половозрастной особи – *buč* ‘козел’, *vaz* ‘коза’, ср. тадж. *буз* ‘козел’.

Один из случаев употребления в шугнанском языке одного из сложноименных глаголов достаточно ранний. Он зафиксирован Р.Б. Шоу в шугнанской сказке еще в середине XIX века: 61. ... “*Mu muzā tāž.*’ *Bād yi pōd-i di tîžd, lîv...* «... “Сними мой сапог”. И как станет она снимать сапог с ноги...» (из сказки) (прив. по тексту: Додыхудоева 2020а: 72, 82). Ср. в той же сказке на таджикском языке: *Вазирзода ... кордчаро ба соқи мӯзааш андохт, то ки вақти кашидану пӯшидани мӯза... бародари* «Сын вазира... положил ножик за голенище своего сапога, чтобы ... не забывать его (своего названного брата – Л.Д.)» (прив. по тексту: Додыхудоева 2020б: 48, 65).

К этому примеру можно добавить еще один устойчивый оборот и даже несколько его модификаций, которые сложились на основе сложноименного таджикского глагола *мӯза кашидан* ‘снимать/надевать сапоги’: *обро надида мӯза кашидан* ‘букв. снимать сапоги, не видя воды; *пер.* приступать к делу, не подготовившись; проявить поспешность при выполнении дела’ (ТСТЯ 2010б: 6), а также: *обро дида мӯза кашидан* ‘букв. снимать сапоги, увидев воду; *пер.* приступать к делу, подготовившись, уяснив ситуацию’ и пословичное выражение *обро надида, мӯза накаш* ‘букв. не видя воды, не снимай сапог’, что соответствует русскому ‘не зная брода, не суйся в воду’. Ср. к последнему случаю выражение: *Обро нади-*

да мӯза накаш. Дарди худ пеши дардмандон бигӯ. ‘Не видя воды, не снимай сапог. Делись своей болью с тем, кто поймет’, то есть «не делай ничего поспешно, не прояснив ситуации» [Зарб.=Посл.] (Кенджазода 2021: 169).

Еще несколько шугнанских фразеологизмов отмечены в «Кратком словаре фразеологических единиц шугнанского языка и их эквиваленты в русском языке»: Устойчивое выражение *nafas tīžd naboft* букв. ‘нельзя сделать вздох’, пер. о большом скоплении кого/чего-л; *xi yūy-um tīžd* букв. ‘свои уши натянул, наострил’, пер. ‘принял к сведению, намотал на ус’ (Мирзоев, Карамова 2014: 114, 122).

Там же приведен относительно новый для шугнанского языка фразеологизм *pūst pi kāl tīždow* букв. ‘шкуру на голову натянуть’. Можно полагать, что данная ФЕ построена по модели, имеющей место в таджикском языке *рӯ ба нӯстак кашидан* букв. ‘на лицо кожу/шкуру натягивать’ (см. также др. сходные выражения в Таблице 3.1). Данное выражение отмечено у Садриддина Айни в романе «Дохунда» в виде *ба рӯи худ чарм кашидан* букв. ‘натягивать на лицо кожу’, пер. ‘преодолевать стыдливость, не иметь стыда’. Однако в шугнанском языке данная ФЕ применяется в переносном значении ‘наказ(быв)ать, спустить шкуру’: *yakbor-ga di jinow kor-jāt tu pūst pi tu kāl tāžum* ‘другой раз за такую работу я шкуру с тебя спущу’ (Мирзоев, Карамова 2014: 31), ср. в том же значении ‘наказывать’ в (Аламшоев 2018: 134, 379)..

Некоторое число фразеологических единиц различного происхождения с данным опорным глаголом приводится в работе Ш.М. Аламшоева. Среди них можно отметить фразеологизмы, созданные под влиянием таджикского языка и вошедшие в шугнанский язык в последнее время: *firū tīždow* ‘глотать, проглатывать’, *dar bar tīždow* ‘обнимать, заключать в объятия’, *dar-dam tīždow* ‘гнаться, за кем-то’, *ba jawobgari tīždow* ‘привлекать к ответственности, наказывать’, а также единицы с именным компонентом из русского языка: *kilo tīždow* ‘взвешивать, определять вес’ и *pilān tīždow* ‘составлять план’. В той же работе приведены новые значения устойчивых единиц. Так, например, для фразеологизма *jūn tīždow* приводится значение ‘бескрайняя жадность’, а для выражения *qiloč tīždow* – ‘прыгать туда-сюда; шагать быстро, торопиться’ (2018: 147, 487, раздел «Материалы»).

5.2. Сложноименной глагол, передающий значение *курить*, который выражается в обоих языках, шугнанском и таджикском, на основе близких по семантике и грамматической функции (а также грамматическому значению) компонентов (см. 3.6 выше), вынесен в отдельную таблицу и рассматривается особенно подробно, см. Таблицу 5.2

Таблица 5.2. Сложноименной глагол «курить» в шугнанском языке

Шугнанский язык	Русский перевод
tamoki tīždow tamoki-tāžij/kaḥ tamoki-tīžd	курить табак, <i>нос</i> (жевательный табак) курильщик, употребляющий <i>нос</i> курение, употребление <i>носа</i>
yī tamoki-tīžd (tūr)	в одно мгновение; мигом
čilim tīždow čilim-tāžij/kaḥ čilim-tīžd	курить кальян ( <i>čilim</i> кальян, папироса) курильщик, употребляющий <i>нос</i> курение (папиросы, кальяна, табака), употребление <i>носа</i>

papiros tīždow	курить папиросы
papiros-kaḥ	курильщик, употребляющий нос
sigor tīždow	курить сигареты
xu tīždow	курить табак
naswor tīždow naswor-kaḥ	курить нос курильщик носа
maxorka/ā tīždow	курить махорку
yi tīžd	порция (табака) на одну затяжку
taryok tīždow	курить опиум
afyūn tīždow	курить опиум
afyūn-kaḥ	курильщик опиума

В настоящее время в шугнанском языке широко используются композиты, построенные на базе основы настоящего времени таджикского глагола *кашидан*, *каш*, выражающие имя деятеля: *papiros-kaḥ*, *čilim-kaḥ* ‘курильщик (папирос, сигарет, кальяна); *naswor-kaḥ* ‘курильщик носа (носа)’. Они были включены в систему шугнанского языка позднее на этапе активного двуязычия.

В начале 1920-х Г. Моргенстиерне зафиксировал в шугнанском форму *čay-č* ‘тащить; курить’, но, как считает Д.И. Эдельман, это либо ошибка, либо «могло быть ранним заимствованием таджикского *kaš*- < \**karš*-» (Morgenstierene 1974: 41; Эдельман 2011: 301).

И в таджикском, и в шугнанском языках элементы данного сложноименного глагола в значении *курить*, как правило, не употребляются самостоятельно, однако могут употребляться таким образом в высказывании, где ситуацию задает контекст, например: шугн. *Vaxtibegim sa yān wev tamokiyeu xu, vegā-ra tāžām* ‘эй, Бахтибегим, сходи, размельчи те листья табака, и вечером (я) покурю’ (Карамшоев 1999: 55). При отсутствии личного местоимения глагольный показатель присоединяется к первому члену синтагмы и таким образом помещается между именной и глагольной частями: *čilim-um tīžd-xu tu ra kāl dūd* ‘Я курил сигарету, и у меня закружилась голова’ (там же: 55).

В ваханском языке зафиксировано выражение: *wuz-əm čilim botāu* ‘я бросил курить’ (Стеблин-Каменский 1999: 112), (букв. ‘бросил папиросу’), при этом здесь имеется сложноименной глагол *čilim xaš*- : *xašt* ‘курить’ (*čilim* ‘кальян, сигарета, папироса’) (Грюнберг, Стеблин-Каменский 1976).

Приведем таблицу таджикских и шугнанских соответствий при переводе выражений, связанных с запрещением курения, см. Таблицу 4.3. Как следует из примеров, как в таджикском, так и в шугнанском языках усеченные инфинитивы данных сложноименных глаголов имеют тенденцию к преобразованию в композиты и переходу в категорию имен.

Таблица 5.3. Сложноименной глагол «курить» и композит «курение» (тадж. *сигоркашиӣ*, шугн. *čilim-tīžd*) (Мамадасламов 2019: 30)

Шугнанский язык	Русский перевод
Tar čīd darūn čilim natāž	Внутри дома не кури!
Дар даруни хона сгор накаш(ӣ)	Внутри дома не кури!
Tar čīd darūn čilim-tīžd mumkin nist	Внутри дома курение запрещено

Дар даруни хона сгор кашидан манъ аст	Внутри дома курение запрещено
Dars waxt-andir čilim-tižd mumkin nist	Во время занятий курение запрещено
Дар вақти дарс сигоркашӣ манъ аст	Во время занятий курение запрещено
Duxturen-en ziddi čilim-tižd	Врачи – против курения
Духтурон ба муқобили кашидани сигор астанд	Врачи выступают против курения сигарет

## 6. Заключение

Итак, мы рассмотрели сложноименные глаголы, передающие значение курения в таджикском и шугнанском языках в связи со связанными с ними базовыми (вспомогательными) глаголами – тадж. *кашидан*, шугн. *tīždow*). Анализ материала, связанного с этими сложноименными глаголами, показал, что:

- Количество сложных глаголов (в том числе префиксальных) в современных иранских языках, в частности таджикском и шугнанском, интенсивно растет.

- В данных языках данные сложноименные глаголы – различные по происхождению – имеют близкий базовый круг значений. Однако лексико-семантическая сфера таджикского глагола значительно богаче. В нем выявлены сочетания с глаголом *кашидан*, передающие многоактные физиологические действия, психологические и физиологические состояния, эмоциональные и физиологические проявления состояний с оттенком длительности, а также такие значения, как «привлекать внимание, побуждать к действию». При этом образования с глаголом *кашидан*, передающие значение «длиться, продолжаться» – «терпеть, переносить, испытывать, претерпевать», имеют скорее временное, конечное, нежели постоянное наполнение. Указанные лексико-грамматические разряды выделяются с учетом разных сторон семантики глагола и отражения характер протекания действия.

- У таджикского глагола *кашидан* имеются также аффиксальные и понудительные глаголы.

- Круг значений сложноименного глагола с шугнанским опорным глаголом *tīždow* в целом значительно беднее по сравнению с таджикским глаголом.

- В сложноименных образованиях с опорными глаголами тадж. *кашидан* и шугн. *tīždow*, передающими значение «курить», зона курения лексикализуется следующим образом: в обоих языках выделяется только один доминантный глагол (периферийные отсутствуют), который сопровождается соответствующими лексико-семантическими элементами и грамматическими средствами для каждого типа курения.

Структурно-семантические особенности сложноименного глагола «курить» и его деривативов в таджикском и шугнанском языках связаны в первую очередь с особыми лексическо-грамматическими средствами передачи каждого из основных типов курения, а также традиционными приемами и способами употребления табака и его смесей. При этом в обоих языках в лексикализации ситуаций курения задействован параметр способа и характера протекания действия – процессуальности и итеративности.

Анализ материала позволяет продемонстрировать, что сложноименный глагол «курить», выраженный в таджикском и шугнанском языках именем и глаголом, с точки зрения семантики описывает одно событие. При этом различается вклад в итоговое значение как главного именного члена, так и зависимого глагола, который придает всему выражению значение протяженности, процесса. Наряду с этим можно считать, что в составе данного сложноименного глагола опорный глагол в обоих случаях в той или иной степени грамматикализован.

### Сокращения

бур. – бурушаски	руш. – рушанский
вах. – ваханский	сангл. – сангличский
др.-инд. – древнеиндийский	сар. – сарыкольский
др.-ир. – древнеиранский	ср.-перс. – среднеперсидский
ишк. – ишкашимский	тадж. – таджикский
йд. – йидга	тадж.-перс. – таджикско-персидский
кл. перс. – классический персидский	тадж.-юж. – южные диалекты таджикского
мундж. – мунджанский	шугн. – шугнанский
перс. – персидский	ягн. – ягнобский
рус. – русский язык	язг. – язгулямский

### Источники

- Аламшоев Ш.М. *Фразеология шугнанского языка (структурно-грамматический и семантический аспекты)*. Дисс. на соискание уч. ст. канд. филол. наук. Душанбе, 2018.
- Мирзоев Ш., Карамова И. *Краткий словарь фразеологических единиц шугнанского языка и их эквиваленты в русском языке* / Под ред. Ш.П. Юсуфбекова. Хорог, 2014.
- Русско-таджикский словарь / сост. С.Д. Арзуманов, Х.А. Ахрори, М. Бегбуди и др.; под ред. М.С. Асимова. М.: Рус язык, 1985
- Словарь народных идиоматических выражений (*Фарҳанги ибораҳои халқӣ*) / собр. и сост. Р. Абдуллозода; под ред. Г. Джураева. Душанбе: Адиб, 1988. (На тадж. яз.)
- СЮГТЯ – Словарь южных говоров таджикского языка (*Фарҳанги гушиҳои ҷанубии забони тоҷикӣ*) / сост. М. Махмудов, Г. Джураев, Б. Бердиев; отв. ред. А. Мирбобоев. Душанбе: Пайванд, 2012. (На тадж. яз.)
- ТРС 2006 – Таджикско-русский словарь / Фарҳанги тоҷикӣ ба русӣ. Под ред. Д. Саймиддинова, С.Д. Холматовой, С. Каримова. Институт языка и литературы им. А. Рудаки. 2 изд. дополненное. Душанбе: Пажӯҳишгоҳи фарҳанги форсӣ-тоҷикӣ, 2006.
- ТСТЯ – Толковый словарь таджикского языка. Под ред. С. Назарзода и др. (*Фарҳанги тафсирии забони тоҷикӣ*). Т. 1–2. Зери таҳрири С. Назарзода, А. Сангинов, С. Каримов, М. Ҷ. Султон). Душанбе, 2010. (На тадж. яз.)

### Литература

1. Абаев В. И. *Историко-этимологический словарь осетинского языка*. Т. I. М.–Л.: Изд-во АН СССР, 1958.

2. Грюнберг А.Л., Стеблин-Каменский И.М. *Языки Восточного Гиндукуша. Ваханский язык: тексты, словарь, грамматический очерк.* М., 1976.
3. Додыхудоева Л.Р. Волшебная сказка «Водарег» в иранском фольклоре Центральной Азии: тексты и структура повествования. Часть 1 // *Родной язык.* 2020. № 1. С. 49–85.
4. Додыхудоева Л.Р. Волшебная сказка «Водарег» в иранском фольклоре Центральной Азии: этнолингвистические сведения. Часть 2 // *Родной язык.* 2020. № 2. С. 6–75.
5. Домброва Т.И. Валентность в системе грамматических категорий английского глагола // *Scientific journal of the Lviv State University of Life Safety "Philological Periodical of Lviv".* 2017. № 2. С. 22–31.
6. Карамшоев Д. *Шугнанско-русский словарь.* М.: Восточная литература. Наука. Т. I (1988), II (1991), III (1999).
7. Кенджазода М.С. *Структурно-грамматические особенности простых предложений в современном таджикском литературном языке (на основе материалов современной прозы).* Дисс. на соискание уч. ст. докт. филол. наук. Душанбе, 2021.
8. Мамадасламов М.С. *Выражение пространственно-временных отношений в системе предлогов и послелогов в разносистемных языках (на базе шугнанского, немецкого и таджикского языков).* Дисс. на соискание уч. ст. доктора филол. наук. Душанбе, 2019.
9. Пахалина Т.Н. Ваханский язык. М., 1975.
10. Пейсииков Л.С. К типологии служебных глаголов в персидском языке // *Уч. зап. кафедры иностранных языков ИМО. Вып. 1. Вопросы языка и литературы стран Востока.* М., 1958.
11. Пейсииков Л.С. *Вопросы синтаксиса персидского языка.* М., 1959.
12. Пейсииков Л.С. *Лексикология современного персидского языка.* М., 1975.
13. Пейсииков Л.С. К истории словообразования отыменных глаголов в персидском языке // *Иранское языкознание.* М., 1976, 190-198.
14. Расторгуева В.С. *Краткий очерк грамматики персидского языка.* М., 1953.
15. Расторгуева В.С. Иранские языки // *Языки мира: Иранские языки. I. Юго-западные иранские языки.* М., 1997.
16. Расторгуева В.С., Керимова А.А. *Система таджикского глагола.* М.: Наука. 1964.
17. Расторгуева В.С., Эдельман Д.И. *Этимологический словарь иранских языков.* Т. II. М.: Вост. лит., 2003.
18. Розенфельд А.З. *Таджикско-русский диалектный словарь (Юго-восточный Таджикистан).* Л.: Изд-во ЛГУ, 1982.
19. Рубинчик Ю.А. Структурно-семантические свойства многочленных глагольных фразеологизмов // *Труды СамГУ им. А. Навои. Новая серия. Вып. 219. Вопросы фразеологии.* Самарканд, 1972.
20. Рубинчик Ю.А. Некоторые закономерности в образовании рамочной конструкции персидскими сложными глаголами // *Труды СамГУ им. А. Навои. Новая серия. Вып. 269. Вопросы востоковедения.* Самарканд, 1975.

21. Рубинчик Ю.А. Персидский сложный глагол и вопросы образования рамочной конструкции // *Индийская и иранская филология. Вопросы грамматики*. М., 1976. С. 292–310.
22. Рубинчик Ю.А. *Основы фразеологии персидского языка*. М.: Наука, 1981.
23. Рубинчик Ю.А. *Грамматика современного персидского литературного языка*. М., 2001.
24. Стеблин-Каменский И.М. *Этимологический словарь ваханского языка*. СПб., 1999.
25. Турсунова Д.Ш. Семантические особенности сложного глагола в персидском языке // *Вестник Челябинского государственного университета*. 2012. № 12 (266). Политические науки. Востоковедение. Вып. 12. С. 132–133.
26. Шаляпина З.М. *Трехмерная стратификационная модель языка и его функционирования: к общей теории лингвистических моделей*. М.: Восточная литература, 2007.
27. Шарипова С.С. *Приставочные глаголы движения в русском языке в сопоставлении с таджикским языком*. Дисс. на соискание ученой степени канд. филол. наук. Душанбе, 2019.
28. Эдельман Д.И. *Сравнительная грамматика восточноиранских языков. Лексика*. М., 2009.
29. Эдельман Д.И. *Этимологический словарь иранских языков*. Т. 4. М.: Восточная литература, 2011.
30. Čašule I. Burushaski and unique Slavic isoglosses // *Himalayan Linguistics*. 2017. Vol. 16 (2). P. 1–25.
31. Cheung J. *Etymological Dictionary of the Iranian Verb*. Leiden, Boston: Brill. 2007.
32. Szemerényi O. Iranian studies. I // *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 1959. Bd. 76, 1-5. P. 60–77.
33. Morgenstierne G. *Etymological vocabulary of the Shughni Group*. Wiesbaden, 1974.

## Приграничный русско-китайский пиджин 90-х годов XX в.

Герман Борисович Дудченко

Невский институт языка и культуры, Санкт-Петербург, Россия,  
dudchenkog@gmail.com, ORCID 0000-0002-1137-3799

**Аннотация.** Статья посвящена языковому явлению, которое получило наименование «пиджин». В основном на примере города Суйфэньхэ автор рассматривает пиджин, возникший на границе России и Китая в начале 90-х годов XX в. Граждане двух стран совершали передвижения через границу, вели торговлю, но испытывали трудности в общении из-за языкового барьера. Между ними возник контактный язык, включающий несколько слов и фраз. Появление и развитие данного пиджина представляет интерес с точки зрения креолистики.

**Ключевые слова:** пиджин, международное общение, китайский язык, русский язык, приграничье, Суйфэньхэ, Хуньчунь

**Abstract.** The article is devoted to the linguistic phenomenon, which has received the name “pidgin”. Mainly on the example of the city of Suifenhe, the author considers the pidgin that arose on the border of Russia and China in the early 90s of the XX century. Citizens of the two countries traveled across the border, conducted trade, but experienced difficulties in communication because of the language barrier. A contact language arose between them, including several words and phrases. The emergence and development of this pidgin is of interest from the point of view of creolistics.

**Keywords:** pidgin, international communication, Chinese language, Russian language, borderlands, Suifenhe, Hunchun

В изучении языковых явлений традиционно особый интерес представляет российско-китайская граница, точнее, приграничные территории двух стран, которые в настоящее время обозначены на карте мира как Российская Федерация и Китайская Народная Республика. Одно из государств занимает первое место в мире по площади территории, другое – по численности населения. Граница между ними тоже очень протяжённая, она насчитывает более 4200 километров. Широко известны разные события в истории, происходившие вдоль границы, во взаимоотношениях между странами складывались целые эпохи.

Были времена, когда между представителями народов, населявшими данные территории с обеих сторон границы, контакты развивались настолько активно, что вырабатывались свои языковые средства общения. Достаточно упомянуть Кяхтинский язык, применявшийся в XVIII – первой половине XIX вв. Данный контактный язык, который часто определяют как китайско-русский торговый пиджин, своё название получил из-за того, что использовался в городах Кяхта и Маймачин (ныне город Алтинбулаг в Монголии), располагавшихся в то время вблизи русско-китайской границы [Перехвальская 2008: 50].



В научной литературе с применением термина «пиджин» рассмотрены языковые явления, которые на данных приграничных землях распространились в более поздние времена. Выделяется русский пиджин в его дальневосточном варианте, служивший основным средством общения представителей коренных народов с русскими. На территории нынешнего Приморского края он использовался с середины XIX в. по 1930-е гг. [Гурьян, Карпенко 2019: 49]. Китайско-русские контактные идиомы возникали по обе стороны границы России и Китая [Жданова 2016: 41; Гурьян 2021: 151]. Встречались такие контактные языки не только в приграничных районах, в рассматриваемый нами период они использовались повсюду на вещевых рынках.

Нетрудно предположить, что ситуации, в которых возникает новый контактный язык, складываются каждый раз, когда массовые международные контакты возобновляются после их длительного отсутствия. Так до конца 80-х годов XX в. граница между СССР и КНР была практически закрытой в течение более четверти века. Вскоре граждане двух стран стали приезжать на сопредельные территории и предлагали купить тот или иной товар. При этом из соответствующего языка они не знали, как правило, ни слова. В общении между ними выработался определённый контактный язык.

Есть основания, чтобы рассматривать его именно как пиджин. В нём ярко выражены общепринятые критерии данного понятия. Например, социолингвистические критерии – он использовался преимущественно в торговле, применялся в равной степени носителями обоих языков, контактирующие не изучали языки друг друга, но стремились к тому, чтобы их понимали. Кроме того, в формировании спонтанно возникавшего средства общения были задействованы и подвержены изменениям слова из обоих языков – и китайского, и русского. Происходило редуцирование лексики и грамматики языков, в словарном составе оказалось отчётливо выражено явление полисемии, активно применялись устойчивые словосочетания.

Предлагаемый вниманию материал собран нами в городе Суйфэньхэ в 1996-1997 гг. и ранее не публиковался. Общение между гражданами двух стран при помощи данных языковых средств велось и в других городах, расположенных вдоль государственной границы со стороны провинции Хэйлуцзян – Дуннин, Мишань, Хулинь, Жаохэ, Фуюань, Тунцзян, Хэйхэ, а также со стороны Автономного района Внутренняя Монголия – в городе Маньчжоули (Маньчжурия). Начиная формироваться рассматриваемый вариант русско-китайского пиджина, по всей видимости, несколько раньше – в самом начале 90-х годов, когда массовые передвижения через границу стали реальностью.

Необходимо сразу сделать несколько пояснений. Состав основных языковых единиц пиджина представлен ниже в виде записей, расположенных в алфавитном порядке. Очевидно, что в представляемом словнике собрана не вся лексика данного русско-китайского пиджина. Необходимо учесть, что он получил распространение не только в Суйфэньхэ. В других местах и в узких кругах могли возникать свои слова и выражения.

Латинскими буквами без квадратных скобок обозначены слоги китайского языка при помощи пиньинь. Рядом с иероглифами они указаны с обозначением тонов. В иных случаях записи в пиньинь приведены без надстрочных знаков то-

нов, поскольку по-китайски они могли записываться разными иероглифами и, соответственно, произноситься разным тоном. Слова из языков, в письменности которых задействованы латинские буквы, переданы в их оригинальном написании. В квадратные скобки заключены знаки международного фонетического алфавита.

АВА́НС – (русск. аванс < фр. avance) задаток; аванс.

ДОЙ́ЧЕ ЦЫ́ГА – (искаж. кит. 多少 *duōshǎo zhège* букв. сколько это, в русской транскрипции передается как «дошао чжэзэ», но в северных диалектах вместо «чжэ» часто произносится «цз») сколько стоит? Обычный вопрос покупателя продавцу с указанием на товар. В ответ продавец набирает цифры на калькуляторе или записывает на бумаге.

ДУЛУ́ГА – (искаж. русск. друг) *то же, что* КОРЕФА́НА.

КАМБЭ́Й – (кит. 干杯 *gānbēi*) до дна! поехали! *англ. cheers!*; *возглас перед тем, как выпить спиртное.* Слово, часто звучащее в ресторанах города Суйфэньхэ.

КАПИТА́НА – (русск. капитан < фр. capitaine < лат. caput букв. голова, глава) 1) начальник; старший; главный; шеф; босс; 2) полицейский (*вне зависимости от ранга*); 3) Председатель КНР; Президент РФ.

КОРЕФА́НА – (русск. корефан) 1) хороший человек; 2) друг; 3) мужчина; парень; мальчик; *любой человек мужского пола, не вызывающий у говорящего неприязненного отношения*; 4) возлюбленный; муж; супруг; 4) прохожий; 5) продавец; официант; водитель; рабочий. Слово, которое в городе Суйфэньхэ часто используется и как обращение.

КУ́НЯ – (искаж. кит. 姑娘 *gūniáng* – девушка) 1) женщина; дама; девушка; девочка; *любая представительница прекрасной половины человечества вне зависимости от возраста*; 2) возлюбленная; жена; супруга; 3) прохожая; 4) продавщица; официантка; уборщица; горничная. Слово, которое выкрикивают в коридорах гостиниц города Суйфэньхэ, вызывая кого-нибудь из обслуживающего персонала. Повсеместно в городе используется в качестве обращения.

ПАДУЛУ́ГА – (искаж. русск. подруга) *то же, что* КУ́НЯ.

ПОМОГА́Й – (русск. помогать) 1) грузчик; носильщик; 2) уличный гид; 3) переводчик; 4) человек, навязчиво предлагающий свои услуги.

ПУХА́О – (кит. 不好 *bùhǎo* – плохой) 1) плохой; плохо; 2) дефект (*товара*); недостаток; изъян; 3) *выражение несогласия (в том числе с ценой в процессе торга), возражения, отказа*; 4) *то же, что* ХУЛИГА́НА (1).

СЕСЕ – (кит. 谢谢 *xièxiè*) спасибо.

СУ́ПЕР МИ́НИМУМ – (русск. супер < лат. super букв. сверху, над; русск. минимум < лат. minimum – наименьшее) ниже (данной цены) не торгуюсь. Слова, с которыми продавец называет наименьшую цену, давая знать о неуместности дальнейшего торга.

ХА́О – (кит. 好 *hǎo* – хороший, хорошо) 1) хороший; красивый; качественный; 2) хорошо; договорились; 3) *выражение одобрения, восхищения, лучших чувств*; 4) *выражение приветствия*; привет!

ХУЛИГА́НА – (русск. хулиган < *англ. hooligan < ирл. Houlihan* – фамилия семьи, известной своей грубостью) 1) плохой (нехороший) человек; 2) представитель криминального мира; бандит; мафиози.

ЧИФА́ – (искаж. кит. 吃饭 chīfàn букв. есть рис) есть; кушать; принимать пищу. Используется также и как выражение желания поесть, и как приглашение в ресторан.

ЮАНИ МЙЯ – (кит. 元 yuán Юань – денежная единица КНР; искаж. кит. 没有 méiyǒu нет, не имеется) денег нет.

ЮАНИ НА́ДА – (кит. 元 yuán Юань – денежная единица КНР; русск. надо) 1) платный; платно; 2) выражение требования заплатить.

Краткий анализ приведённого словаря показывает, что лексические единицы примерно поровну подразделяются на те, которые взяты из китайского языка, и те, которые имеют русскоязычное происхождение. Этот момент лишний раз подчёркивает, что данное спонтанно возникшее языковое явление представляет собой именно пиджин. Люди, вступившие в международное общение, изначально не владели языками друг друга и, как правило, не стремились их дальше осваивать. Но потребность в общении с проявлением эмоций у них была, поэтому «на паритетных началах» произошёл отбор лексических единиц из двух языков.

Возникает вопрос, почему оказались востребованы именно эти слова? Некоторые из критериев отбора понятны и объяснимы – в словнике имеются наиболее распространённые слова из общей лексики, а также слова, способные образовывать синекдохи. В каких случаях выбиралось русское слово, а в каких китайское? Ответы на эти и другие вопросы можно найти, если разобрать разные нюансы в отношении каждого слова в отдельности. Для этого сгруппируем лексические единицы по признаку того, из какого языка они вошли в пиджин. Начнём с русских слов.

Произнесение слова *аванс* не вызывает трудностей у носителей китайского языка. Они могут расслышать в нём три слога, которые состоят из звуков, похожих на те, что имеются в китайской фонетике. Если *аванс* записать с помощью пиньинь, то оно будет выглядеть, к примеру, следующим образом: awansi. При этом в последнем слоге такое сочетание звуков представляет «особый гласный». Будучи произнесённым нейтральным тоном, он едва заметен, и в итоге слышится слово, артикулируемое при помощи китайской фонетики, но поразительно близкое к «оригиналу» из русского языка – слову *аванс*.

Это слово употреблялось, когда граждане России, приезжавшие в Суйфэньхэ, заказывали у мастеров какую-либо услугу – пошить одежду, вставить стёкла в оправу очков, починить обувь и т.д. Поняв, что нужно сделать, мастер мог запросить предоплату. Но произносить это слово, используя китайскую фонетику, затруднительно – в нём содержатся согласные звуки, идущие подряд, среди них звук [r], в слове четыре слога. Более того, если попытаться преобразовать звучание слова *предоплата* на манер силлабических конструкций китайского языка, то получится что-то вроде puleidupulade, слогов в нём будет ещё больше – не четыре, а шесть, а главное – слово меняется до неузнаваемости для носителей русского языка.

Китайский эквивалент слов *предоплата, задаток* – 预付款 yùfùkuǎn – тоже не очень подходит для использования в русско-китайском пиджине. В нём есть придыхание и дифтонг – явления, не характерные для фонетики русского языка,

а также звук [ü]. Это слово можно записать при помощи букв русского алфавита. Например, палладицей – *юйфукуань*. Но если по этим буквам его воспроизвести вслух, да ещё и без артикуляции тонов, то носитель китайского языка совершенно не поймёт, о чём речь. Таким образом, выбор для использования в данном пиджине слова *аванс* легко объясним.

Также объясним выбор слов *друг* и *корёфан*. В рассматриваемый период возникла потребность использовать их в пиджине в качестве универсального обращения. В более ранние времена для этого могло бы применяться слово *товарищ*, которое имеет эквивалент в китайском языке – 同志 *tóngzhì*. Но поскольку ко времени начала массового приезда в Китай в среде российских граждан это обращение уже практически вышло из употребления, то оно и для пиджина было бы устаревшим. Иногда слово *товарищ* в это время использовалось при обращении к россиянам на русском языке местными жителями не в Суйфэньхэ, а в более крупных и отдалённых от границы городах, в частности в Харбине.

Вхождение в русско-китайский пиджин слова *друг*, образованного из эквивалента данного слова на китайском языке – 朋友 *péngyǒu*, столкнулось бы с трудностями фонетического характера для носителей русского языка. В этом слове есть и придыхание, и отсутствующий в фонетике русского языка заднеязычный звук [ŋ], и дифтонг [ou]. Можно добавить, что артикуляция звука [e] в фонетике китайского языка не соответствует русскоязычным аналогам данного звука. При передаче русскими буквами он записывается то как «ы», то как «э». Насколько подходит для пиджина эквиваленты слова 朋友 *péngyǒu*, имеющиеся в русском языке?

В рассматриваемом контексте слова *друг* и *корёфан* означают одно и то же, являются взаимозаменяемыми. Однако в Суйфэньхэ несоизмеримо чаще использовалось слово *корёфан*. Особенности фонологии китайского языка таковы, что его носители в произнесении [drug] с наименьшими трудностями столкнулись бы только со звуком [u]. Во-первых, в этом слове есть закрытый слог, в котором последний звук [g]. Слоги, в конце которых произносятся велярные звуки [g] и [k] – явление, распространённое во многих языках мира (славянских, тюркских и т.д.), но совсем несвойственное китайскому языку. В китайской фонетике это звуки инициалей, то есть за ними следует гласный звук. Во-вторых, в начале слова *друг* проговариваются два согласных звука подряд, один из которых [r]. В таких случаях, при произнесении иноязычных слов, первый согласный артикулируется как инициаль, а звук [r] переходит в [l]. В итоге получается что-то похожее на [dɛlugɛ]. Буквами русского алфавита это можно записать как «дулуга», примерно так оно и слышалось, но не всегда сразу было понятно, что имеется в виду слово *друг*.

Если учесть, что звуки [d] и [g] в фонетике китайского языка произносятся без придыхания, но не с настолько ярко выраженным озвончением, как в русском языке, то [dɛlugɛ] может быть произнесено и как [tɛlukɛ]. В этом случае россиянину, приехавшему в Суйфэньхэ, узнать слово «друг» было бы весьма проблематично.

Что касается слова *корёфан*, то с его произнесением для носителей китайского языка трудностей возникает меньше. Слог [fan] широко представлен в силлабическом строе китайского языка и артикулируется достаточно близко к такому же слогу в словах русского лексикона – *фантастика*, *сарафан* и т.д. Два предыду-

щих слога в слове *корёфан* являются открытыми. В китайском есть похожие слоги, которые в пиньинь записываются *kuo* (либо *guo*) и *ge*. И хотя звуки [к] и особенно [г] в этих слогах артикулируются не так, как в фонетике русского языка, слово *корёфан*, произносимое жителями города Суйфэньхэ, легко узнавалось приезжими россиянами и прочно вошло в лексикон русско-китайского пиджина.

Аналогичным образом мог произойти отбор слов *капитан* и *хулиган*. Нетрудно заметить, что в словнике все эти слова завершаются открытым слогом – *корёфана*, *капитана*, *хулигана*. Для воспроизведения данных слов русского языка при помощи китайской фонетики возникал выбор. В одних случаях можно использовать в конце слова слоги, записываемые как *fan*, *tan*, *gan* или похожие на них. Тогда сохраняется длина слова, количество слогов, но не совсем ясно, какой звук лучше произнести в конце – [н], [ŋ] или какой-то ещё. В других случаях один закрытый слог разбивается на два открытых, к примеру *fa* и *na* (либо *ne*). То есть звук [н] передаётся не как финаль, а как инициаль. В результате этот согласный звук более отчётливо выражен, а удлинение не мешает понять носителю русского языка, какое слово имеется в виду. Поэтому местные жители произносили эти слова по-разному, а приезжие россияне так же, как в русском языке. Иногда они добавляли звук [а], обыгрывая артикуляцию собеседников.

Особое внимание в этом ряду обращает на себя слово *капитана*. Оно присутствовало в языковом явлении, обозначаемом в публикациях как сибирский пиджин, где имело значение «господин, сударь», в том числе при обращении, а также просто «мужчина-европеец» [Перехвальская 2008, с. 20]. Кроме того, употребление слов *хулиган* и *капитан* отмечено в международном общении в Харбине в первой половине XX в. В частности, слово *gebidan* обозначало всех чиновников [Ma Na 2015, с. 351].

Интересны метаморфозы, связанные с появлением в пиджине слова *помогай*. Оно образовано от инфинитива русского глагола *помогать*. Неопределённой формой глагола в русско-китайских контактных языках могло передаваться намерение говорящего. Но приходилось эти слова как-то преобразовывать, чтобы они, с одной стороны, соответствовали силлабическому строю китайской фонетики, с другой – были легко узнаваемы носителями русского языка, и для них сразу становилось понятным, о чём речь. Поэтому глаголы часто ставились в повелительное наклонение, которое могло заменить любые формы глаголов. Хотя возникали случаи, когда носителям китайского языка русские глаголы приходилось произносить в инфинитиве.

Достаточно привести пример с другим глаголом – *убирать* в значении «делать уборку». Когда горничные в гостиницах заходили в номера к российским туристам, то проговаривали это слово как *wubilaqi* или *wubilatie*. Было заметно, что затруднение в произнесении у них вызывал не только звук [г], но и мягкий [т], да ещё и в конце закрытого слога. Тем не менее, если произнести *wubilai*, то непонятно, что делать постояльцу – то ли самому произвести уборку, используя принёсённые тряпки, тазы и т.д., то ли убрать свои вещи на время уборки помещения. На самом деле горничные сообщали о своём приходе для уборки, чтобы их действия не вызывали у постояльцев удивления или недоумения. Поэтому в данном случае подходил инфинитив глагола *убирать*.

Что касается глагола *помогать*, то в пиджине оказалась вполне приемлемой его форма в повелительном наклонении – *помогай*. Есть вариант, когда носителями китайского языка использовался слог не *gai*, а *ga*. Получалось *помога* с ударением на последнем слоге. Так могли говорить носильщики на вокзале, подбегавшие к россиянам и их тяжёлым сумкам. В данном контексте смысл этого слова не вызывал сомнений, но в других случаях для россиян более понятным и привычным было слово *помогай*. Далее заметно явление перехода одной части речи в другую, и уже не в пиджине, а в лексиконе русского языка появились существительные *помогай* и *помогайка*.

Немало вопросов возникает в связи с появлением в пиджине выражения *супер минимум*. Почему отобраны были эти иноязычные по происхождению слова, и почему не просто *минимум*? Как смысл выражения согласуется со смыслом слов в русском языке, и насколько удобно оно воспроизводится с помощью фонетики китайского языка? Возможно такое объяснение – эти слова являются ярко выраженными интернационализмами. То есть они были известны в Китае в виде записей латинскими буквами ещё до активизации в начале 90-х годов XX в. массовых контактов в приграничье граждан двух стран. Особенно это касается морфемы *супер*, которая благодаря глобализационным процессам получила распространение во многих языках мира. В частности, в английском – *superstar*, *superman*, *superliner* и т.д. В русском языке она стала не только приставкой во многих словах – *супергерой*, *суперзвезда*, *супердержава*, *супертанкер* и т.д. В молодёжном сленге морфема перешла в отдельное слово, наречие – «*всё было супер*», «*это супер*», «*он одевается не хорошо, а супер*». При произнесении этого слова жителями Суй-фэньхэ не возникало больших фонетических трудностей, так как звук, похожий на [r], в данном случае артикулировался достаточно легко благодаря наличию в китайской фонетике эризации, к тому же распространённой в северных диалектах. Этот звук передавался через слог *er*.

Передача значения «наименьшая цена» без слова «минимум» привела бы к применению степени сравнения, а это усложнило бы подбор слова, подходящего для пиджина. Поэтому *минимум* оказался приемлемым вариантом – и в словарном составе русского языка весьма распространённое слово, и с точки зрения фонетики китайского языка не вызывает трудностей, образуется сложением обычных открытых слогов. Исключение составляет последний звук [m], он артикулировался отдельно от всех слогов, не образуя нового открытого слога. Таким способом слово *минимум* носителями китайского языка и проговаривалось, а слово *супер* перед ним использовалось, очевидно, для подчёркивания, для передачи категоричности своей позиции.

Слово *надо*, которое легко произносилось носителями китайского языка *nada* или *nađe*, использовалось не только во фразе *юани надо*. Китайские продавцы переходили от пиджина к овладению русским языком, довольно быстро пополняли словарный запас. *Надо* явилось одним из наиболее востребованных ими слов. Продавец мог по-русски назвать предлагаемый товар и тут же добавить слово *надо*, образовав вопрос. Но это речь уже не о пиджине.

При вхождении в русско-китайский пиджин слов и фраз из китайского языка также происходили некоторые их изменения. Для того, чтобы быть использован-

ными носителями русского языка, они получали соответствующее фонетическое оформление.

Слово *сколько*, разумеется, постоянно используется в торговле, и китайские продавцы в Суйфэньхэ, осваивавшие русский язык, его разучили одним из первых. Но в середине 90-х годов оно было в ходу наряду с фразой из пиджина [doɪʃe siɡə] в разных вариантах – [doʃe dzigə], [doʃao siɡa] и т.д. В пиджин данное сочетание слов было отобрано из китайского языка, возможно, из-за краткости. Дело в том, что слово *сколько* при передаче слогами фонетики китайского языка оказывалось бы длинным, из четырёх слогов – например, sikelike, и было бы не сразу различимым для россиян. Другое объяснение – в русском варианте, как правило, данная фраза звучит как *сколько стоит*, а во втором слове имеются два согласных звука подряд, а также закрытый слог с глухим звуком [t] в конце. Данные явления были непривычны для носителей китайского языка, вызывали у них некоторые трудности при произнесении иноязычных слов.

Для носителей русского языка слово 多少 *duōshāo* также вызывало трудность в артикуляции, а именно – дифтонг [ao]. В данной позиции он воспринимался ими как один слабо различимый звук. Поэтому дифтонг переходил в звук [e]. Трудно рассудить, помогало или препятствовало созвучие 多少 *duōshāo* слову *дешево*. Возможно, во избежание этого созвучия в пиджине и произошёл переход звука [ʃ] в [ʒ].

Желая привлечь внимание приезжих россиян, китайские продавцы указывали рукой на разложенный товар и при этом повторяли слово 这个 *zhège*. В пиджине оно преобразовалось в [dzigə], ещё чаще [siɡa], возникали и другие варианты. В силу фонетических особенностей диалектов китайского языка, распространённых в том числе и на приграничных территориях, слог zhe самими носителями китайского языка артикулировался как ze. Поэтому вся фраза произносилась россиянами в разных вариантах, местные жители обыгрывали их артикуляцию, но каждый раз продавцы и покупатели легко понимали друг друга.

Возглас «камбэй» в развёрнутых комментариях по поводу его появления в пиджине, в международном общении, не нуждается. В нём обращает на себя внимание восприятие россиянами первого звука как [k], а не [g]. Также в этом слове заметен переход звука [n] в [m], что в подобной позиции характерно для японского языка, в котором тоже есть этот возглас – «*кампай!*».

Нетрудно заметить, что слово *куня* в данном русско-китайском пиджине является «гендерным аналогом» слова *корефан*. Одно из этих слов взято из китайского языка, другое из русского. В сравнении со всеми возможными вариантами, которые могли войти в лексикон пиджина из русского языка, преобразование слова *guniang* в *куня* оказалось наиболее подходящим вариантом. Если не проговаривать последний звук – заднеязычный [ŋ], то не возникает трудностей при артикуляции этого слова россиянами. Восприятие и воспроизведение ими звука [g] как [k] соответствует особенностям фонетики китайского языка, если не учитывать тоны. В итоге получается слово, похожее на обычную лексическую единицу – существительное женского рода, которое имеет окончание «я» и относится к первому склонению. Таким оно вошло и в лексикон русского языка. Образовались даже такие слова как *кунечка*, *кунька* и другие.

Ранее слово китайского происхождения *куня* в значении «дочь, девушка» неоднократно фиксировалось в сибирском пиджине [Перехвальская 2006, с.43]. Отмечено оно и в контактных языках в рассматриваемый нами период в Забайкалье [Жданова 2016, с. 60].

Не возникает вопросов по поводу появления в русско-китайском пиджине слова 谢谢 *xièxiè*. Вариант на русском языке – *спасибо* – растянулся бы до четырёх слогов *sibaxibo* или *siraxibo*, что было бы не очень привычно для носителей китайского языка. В то же время слово *xièxiè* (*сесе*) не только отличалось краткостью и простотой для запоминания россиянами, но и приблизительно соответствовало фонетике русского языка, в которой есть звуки и их сочетания, воспринимаемые на слух похожим образом.

Настолько же удобным для вхождения в пиджин оказалось слово 好 *hǎo*. Оно короткое, при этом легко различимое и даже созвучно слову *хорошо*. Разумеется, носителями русского языка это слово произносилось при помощи не дифтонга, а зияния. То есть формировалось два разных слога с ударением на первый. Возможно, свою роль в популяризации этого слова среди приезжих россиян сыграла и его схожесть с итальянским возгласом при приветствии и прощании *сiao* [ʃao], который им был хорошо известен. Более того, слово «чао» использовалось в русской разговорной речи и популярных песнях, фильмах и т.д.

В связке со словом *хао* в пиджин вошёл и его антоним, то есть «плохой, плохо» – 不好 *bùhǎo*, произносимый как *пухао*. Слог *bu* передаёт языковую единицу со значением отрицания. Звук [b] в этом слог в пиджин вошёл в глухом варианте – [p], как и в примере в слове *куня* со звуками [g] и [k].

Надо заметить, что в процессе формирования русско-китайского пиджина также применялось слово *хорошо*. Даже при том, что очевидны трудности для носителей китайского языка – длина слова в три слога и звук [r]. *Хорошо* стало использоваться для эмоционально окрашенного приветствия китайскими гражданами россиян. Кроме того, возник вариант антонима к слову *хорошо* – *харанушо*, что значит «плохо». Отрицательная частица 不 *bù* вставлялась в середину наречия, обыгрывая оба языка.

Слово 吃饭 *chīfàn*, произносимое по-русски как *чифа*, было относительно известно на российском Дальнем Востоке ещё до активизации приграничных контактов. Сложился и глагол *чифанить* со значением «есть, принимать пищу», применявшийся в определённой стилистике языка. В Суйфэньхэ слово произносилось россиянами без тонов, без придыхания в первом слог и без последнего звука [n]. Отчасти это объясняется тем, что оставалось неясным, каким следует произносить этот звук – твёрдым или мягким, в палладице этот слог записывается как «фань». Тем не менее в Суйфэньхэ слово *чифа* было понятно всем местным жителям. Нескольким позже, когда во Владивостоке стали открываться небольшие китайские кафе и закусочные, они получили название *чифанька*.

Слово 元 *yuán* закрепилось в русском языке как *юань* задолго до начала массовых контактов в приграничье и того, как данные денежные знаки стали использоваться для оплаты товаров и услуг приезжими россиянами в Суйфэньхэ и других городах на территории КНР. В русском языке это было обычное слово муж-



ского рода, второго склонения, с зиянием в корне, в исходной форме состоящее из двух слогов. В рассматриваемый период оно стало употребляться заметно чаще. В пиджине использовалось, как правило, во множественном числе, в том числе и носителями китайского языка, то есть с открытым слогом [pi] в конце. Также в пиджине заметна синекдоха – слово *юани* приобрело такие значения, как «деньги» и «платный».

Несмотря на существенные различия в фонетике языков, слово 没有 *méiyǒu* появилось в пиджине и звучало как «мия» или «мио»), будучи вполне узнаваемым носителями китайского языка. Как правило, оно употреблялось в связке со словом *юани*. Учитывая, что *мия* в данном случае означает «нет, не имеется», россияне могли образовать соответствующую форму этого слова – *юаней*. Данное словосочетание – *юани мия* или *юаней мио* – употреблялось в тех случаях, когда говорящий хотел отвергнуть навязчивые предложения покупки, поторговаться, выразить сожаление, что не может приобрести предлагаемый товар или воспользоваться услугой.

В дальнейшем китайские продавцы, переходя от пиджина к русскому языку, освоили слово *нет*, хотя звучало оно с ярко выраженным влиянием особенностей китайского языка. Слово *нет* могло обрести звучание *netule*. *Мия* стало употребляться реже.

Интересно проследить дальнейшую судьбу русско-китайского пиджина, столь неожиданно возникшего в первой половине 90-х годов XX в., и в середине десятилетия имевшего хождение наряду с другими языковыми явлениями. Жители Суйфэньхэ и приезжавшие в город на заработки граждане КНР из других мест провинции Хэйлунцзян, вступая в контакт с россиянами, переходили от пиджина к русскому языку. Не владея им на уровне носителей и при этом допуская в речи некоторые слова из пиджина, они формировали своеобразный идиолект или «интерязык». Так обозначается самостоятельная языковая система, которая является результатом попыток тех, кто изучает иностранный язык, воспроизвести нормы целевого языка. Говорящий на «интерязыке» (идиолекте) в отличие от пользователя пиджина стремится овладеть не родным для него языком и говорить на нём правильно [Перехвальская 2008, с.20].

Как развивался данный «интерязык» (идиолект) – это обширная тема отдельных исследований. С ней связаны многие специфические языковые явления. Например, редупликация слов русского языка. Активное употребление составных наименований, таких как *папа-мама* – «родители» или *Китай-человек* – «китаец», отмечено в идиолектах Забайкалья [Жданова 2016, с.74]. В Суйфэньхэ и других городах на востоке провинции Хэйлунцзян в общении с россиянами местные жители использовали такие редупликации как *водка-пиво* и *хулиганы-мафия*.

В свою очередь, у россиян не оставалось потребности продолжать использовать пиджин. В тех местах в городе, где они чаще всего бывали, со всеми могли говорить по-русски. Многие из россиян, как и граждан КНР, продолжали использовать слова *корёфан*, *куня*, *хулиган*, *супер минимум* и т.д. Однако не приходится вести речь о начале формирования креольского языка или иной формы контактных языков, близкой к этому понятию.

По прошествии лет поток российских туристов, приезжавших специально в Суйфэньхэ, заметно снизился. Для многих из них город стал восприниматься

как промежуточный пункт на пути в Харбин, Далянь и другие места в КНР. По нашим наблюдениям, относящимся к 2018-2019 гг., русско-китайский пиджин в Суйфэньхэ уже практически не применялся.

К данному периоду достаточно массовым стал приезд в другой приграничный город – Хуньчунь, расположенный не в провинции Хэйлуунцзян, а в Яньбянь-Корейском автономном округе провинции Цзилинь. В нём сложилась ситуация ярко выраженного языкового многообразия. Достаточно упомянуть, что все вывески в городе выполнены на трёх языках – китайском, корейском и русском. Но на этом фоне пиджинов не возникало. К тому времени, когда в Хуньчуне активизировался трансграничный туристский поток, уже был накоплен опыт международного общения в приграничье. Граждане КНР, которые занимались приёмом российских туристов, торговали, выполняли функции переводчиков, владели русским языком на уровне, достаточно высоком для их работы.

Тем не менее русско-китайский пиджин 90-х гг. XX в., распространившийся на приграничных территориях, в международном общении сыграл весомую роль. Благодаря ему люди имели возможность приезжать в приграничные города, вести куплю-продажу, эмоционально проявлять лучшие чувства друг к другу. В завершение хочется выразить надежду на то, что пример данного пиджина найдёт своё место в теории и истории контактных языков, будет задействован в дальнейших исследованиях.

### Литература

1. Гурьян Н.В., Карпенко С.Ю. Сино-русские идиолекты Приморского края: к постановке проблемы // *Китайская лингвистика и синология: тезисы докладов международной конференции*. М.: РГГУ, 2019. С. 49–54.
2. Гурьян Н.В. Сино-русские идиолекты и русско-китайские пиджины // *XXXI Международный конгресс по источниковедению и историографии стран Азии и Африки: Россия и Восток. К 100-летию политических и культурных связей новейшего времени*. СПб.: Изд-во Студия «НП-Принт», 2021. Т.2. С. 151–153.
3. Жданова Н.А. *Современный русско-китайский пиджин Забайкалья в структурно-системном и коммуникативном аспектах (на материале речи китайского этнолекта)*. Дис. ... канд. филол. наук. Улан-Удэ, 2016. 147 с.
4. Ма На. Харбинский русско-китайский пиджин первой половины XX века и его языковые особенности // *Преподаватель XXI век*. 2015. № 1. С. 345-354.
5. Перехвальская Е.В. *Русские пиджины*. СПб.: Алетейя, 2008. 363 с.
6. Перехвальская Е.В. *Сибирский пиджин (Дальневосточный вариант)*. Автореф. дис. ... докт. филол. наук. СПб., 2006. 51 с.

**Феномен японских заимствований в китайском языке:  
проблемы упорядочения терминологии**

**Елена Николаевна Емельченкова**

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия,  
ekolpachkova@gmail.com

**Аннотация.** В китайском языке лексические единицы, заимствованные из японского или через японский, где они были записаны иероглификой, представляют собой отдельный лексический пласт, отличаясь от заимствований из других языков. Фокусируясь на анализе этой достаточно гетерогенной лексики и сформировавшегося вокруг нее терминологического аппарата, автор рассуждает о необходимости пересмотра русскоязычной синологической терминологии в этой области с опорой на имеющиеся в науке классификации заимствований и учетом достижений современной лексикологии.

**Ключевые слова:** японские заимствования в китайском языке; обратные заимствования; вторичные заимствования; графические заимствования; лексическая деархаизация; лексикология китайского языка

**Благодарности:** Исследование выполнено при поддержке гранта JTI 2019: A Study of Japanese borrowings in the Chinese language in the Meiji era. Автор выражает свою благодарность prof. Mareshi Saito (University of Tokyo, Japan) и И.С. Ибрахим (СПбГУ) за помощь в организации сбора материала, а также А.В. Костыркину и К.И. Демичу (ИВ РАН, Москва) за ценные комментарии при обсуждении материалов исследования.

**Abstract.** In Chinese, lexical units borrowed from/ through Japanese, where they were written in characters, form a specific lexical layer, differing from other languages loanwords. Focusing on the analysis of this rather heterogeneous vocabulary and the terminological apparatus, the author discusses the need to revise the Russian-language terminology in this area and rely on the classification of Japanese borrowings, as well as taking into account the recent achievements of modern lexicology.

**Keywords:** Japanese loanwords in Chinese; Japanese borrowings; japanism; reborrowing; returning loanwords; secondary borrowing; graphic loanwords; dearchaizing vocabulary; Chinese lexicology

**Acknowledgments:** the work was supported by JTI, grant JTI 2019: A Study of Japanese borrowings in the Chinese language in the Meiji era. The author is grateful to prof. M. Saito (University of Tokyo, Japan) и I. Ibrahim (Saint-Petersburg university) for their help with data collection, and to A.Kostyrkin and K.Demich (Institute of Oriental studies, Moscow) for valuable discussion.

Терминологический аппарат, который имеется сегодня в синологии применительно к заимствованиям в китайский язык из японского, формировался на протяжении XX в., когда изучение этого пласта лексики не отличалось системностью,

глубиной или широтой охвата, что однако присуще не только отечественной науке, аналогичная ситуация наблюдалась в китаеведении многих стран. На современном этапе, когда на глазах стало усиливаться взаимодействие между двумя соседствующими языками – китайским и японским, объединяемыми еще и типом используемой письменности, и ощущается всплеск интереса к исследованию активизировавшихся теперь уже на рубеже XX–XXI вв. процессов заимствования лексики, встала проблема ревизии сложившейся в российской синологии терминологии и упорядочения понятийного аппарата, используемого востоковедами для описания лексического влияния японского языка на китайский.

Языковые контакты часто сопровождаются развитием способности языка выступать источником при переходе лексической единицы из него в какой-то другой язык с последующим его «приспособлением» к новым условиям функционирования в языке-реципиенте. При прямом лексическом заимствовании сохраняется материальная сторона языкового знака, который начинает жить по законам принявшей его новой системы. Пример подобного заимствования представляет русское слово *суфле*, заимствованное из французского языка, где оно представлено как *soufflé*. Поскольку заимствоваться может не только форма, но и содержание, в отечественной лингвистике существует две трактовки понятия «заимствование»: в узком понимании учитывается только факт перехода материального экспонента (плана выражения) иноязычного языкового знака, при таком подходе за рамки заимствований выводятся кальки, возникающие при оформлении заимствуемого из языка-донора элемента средствами языка-реципиента. В широком понимании термина, когда учитывается план содержания иноязычного элемента, калькирование также будет относиться к одному из типов заимствований.

Именно такой подход был закреплен в традиционном определении этого феномена как «образовании нового фразеологизма, слова или нового значения слова путем буквального перевода» [ЛЭС: 211]. Для целей настоящего исследования важно, что при таком подходе заимствуемый иноязычный элемент получает выражение средствами языка-реципиента, в качестве которых могут выступать как исконные, так и ранее заимствованные, но уже полностью освоенные элементы. Китай и Япония относятся к странам так называемого «иероглифического клуба», члены которого используют иероглифическую систему письма. Если до второй половины XIX в. языковое влияние носило преимущественно односторонний характер и шло из Китая в Японию, включая заимствование письменности, то на рубеже XIX – XX вв. вектор движения сменился на обратный. Связано это было с политической, экономической и социальной ситуацией в обеих странах. Долго находившийся в самоизоляции Китай отстал в своем социально-экономическом и научно-техническом развитии, тогда как Япония в эпоху Мэйдзи (明治時代 1868–1912) сделала серьезный рывок вперед, главным образом за счет целенаправленной политики привлечения с Запада передовых технологий, а также модернизации политической и экономической систем страны. Эти процессы имели своим следствием активное пополнение словаря японского языка неологизмами, из которого, в свою очередь, за счет оформления новых лексических единиц в иероглифике началось их дальнейшее распространение в соседних языках по всему дальневосточному ареалу.

Изменения в словарном фонде японского языка оказались столь кардинальными, что в лингвистической науке об этих процессах стали говорить как о «лексической революции Мэйдзи» [Маевский 2005: 248]. И хотя в середине XX в. китайские ученые не признавали довольно значительный по объему пласт неологизмов, попавший через японский язык в китайский в качестве заимствований, ссылаясь на то, что слова, созданные из элементов китайского языка по китайским словообразовательным моделям, таковыми считаться не могут [Ван 1958], в конце XX в. китайская лингвистика от этой позиции отошла. И, как представляется, применительно к процессам иноязычных заимствований, имевшим место на рубеже XIX–XX вв. в Китае, можно говорить о «лексической революции» и в китайском языке.

Начавшиеся в первой половине XIX в., еще до начала «реставрации Мэйдзи», и занявшие несколько десятилетий процессы обновления представляли собой словотворчество отдельных языковых личностей, за которым следовал длительный процесс «вживания» неологизма в ткань языка. Вестернизация японской лексической подсистемы в тот период шла как синоцентрический по форме процесс, поскольку иностранные понятия облекались в форму *канго* 漢語 ‘китайские слова; китаизмы’ как новые языковые единицы, составленные из элементов китайского происхождения и отражаемые на письме иероглифическими знаками. Это отличает ход освоения иноязычных слов периода Мэйдзи от заимствований ксенолексики в японский язык на современном этапе, когда пополнение словарного фонда идет в значительной мере за счет *гайрайго* 外来語 ‘иностранные слова’ [Маевский 2005: 247].

Прежде всего бросается в глаза резкий количественный скачок неологизмов в эпоху Мэйдзи на фоне постепенного и поэтому не столь заметного прироста *канго* до середины XIX в. Словотворчество в анализируемый период, помимо массовости, характеризовалось рядом специфических черт, главная из которых состояла в том, что новые лексические единицы создавались целенаправленно в иероглифической форме, ранее доступной лишь образованным слоям японского общества. Это придавало новым словам тот книжно-ученый пафос, который всегда был в почете у интеллектуалов, воспитанных в духе конфуцианства, и ксенолексики, облеченная в иероглифику, воспринималась как более изысканная, престижная и высокоинтеллектуальная.

Подобный путь пополнения терминологического аппарата, призванного обслуживать модернизацию Японии, был аналогичен тем процессам, которые имели место в западных странах, в которых формирование новой узкоспециальной лексики в какой-либо терминологически слабо освоенной области науки и техники часто идет за счет материала латинского и древнегреческого языка (путем сочетания корневых и/или присоединения аффиксальных морфем), а не использования лексики живых языков. Цель у словотворцев в обоих случаях совпадала – придать особый статус вновь создаваемым терминам как важным элементам языка науки.

В японоведении с целью размежевания *канго* домэйдзийского периода и тех неологизмов, что были созданы во второй половине XIX в., последние стали называть *синканго* – 新漢語 ‘новые канго; новые китайские слова’. Эти ранее отсутствовавшие в ходу лексические единицы, возникнув в прогрессивной студенческой

кой среде, с взрослением и карьерным ростом соответствующих носителей японского языка постепенно вошли в язык государственных чиновников, а затем расширили сферу бытования и до всего языкового сообщества Японии, включая тех, кто не мог позволить себе получение образования в принципе. В результате такое студенческое поветрие, как мода на *канго* со временем привело к признанию этих единиц стандартом японского языка, вплоть до приписывания им статуса более приоритетных, чем исконно японские слова [Аракава 2018: 43].

Мэйдзийские неологизмы (в терминологии Е. В. Маевского), будучи «изобретаемыми» *ad hoc* при переводе западной литературы в традиционно китайской иероглифической форме из подручных средств языка, могли какое-то время существовать в нескольких синонимичных вариантах. Со временем «неудачные, нетерминологичные, дезориентирующие, требующие от читателя такого радикального переосмысления привычных слов, к которому он не готов и о необходимости которого даже не догадывается» [Маевский 2005: 249], отбрасывались или уходили на лексическую периферию. Поскольку западная литература переводилась на китайский язык и напрямую с европейских языков, какое-то время (вторая половина XIX – первая половина XX вв.) конкурировать могли сразу несколько вариантов перевода, предложить которые могли как китайские, так и японские ученые.

Сам процесс появления новых слов в лексической подсистеме языка носил творческий характер, при выборе переводных эквивалентов зачастую сами переводчики прибегали к разным переводным эквивалентам и несовпадающему набору иероглифических знаков для фиксации ранее не имевшегося в языке термина. При семантическом калькировании лексическая единица языка-реципиента получает новое, прежде не присущее ей значение по аналогии с иноязычным элементом в языке-доноре. При этом выбор средств фиксации соответствующего значения осуществляется не произвольно, а на основании определенных ассоциаций, возникающих у переводчика в связи с исходной лексической единицей и тем элементом, который он предлагает в качестве переводного эквивалента. Субъективность восприятия иноязычных единиц приводила к появлению целого ряда вариантов перевода. Так, известен пример подобных «словообразовательных метаний» видного японского философа, переводчика и просветителя Ниси Аманэ (西周), который для обозначения европейской философии наряду с транскрипцией *хиросохи* ヒロソヒ в разные периоды предлагал как минимум четыре иероглифических варианта термина – *ригаку* 理学, *рирон* 理論, *китэцугаку* 希哲学 и *тэцугаку* 哲学 [Демич, Костыркин 2020].

В отличие от современных заимствований, добрую долю которых составляют стилистически маркированные коллоквиализмы и отдельные группы узкоспециальной лексики, в том числе и арготическая лексика, словообразовательные процессы второй половины XIX – начала XX вв. завершились появлением слов, большинство которых мы сегодня относим к основному словарному фонду и сочли бы устойчиво актуальной, общей лексикой в любом языке мира. Значительная часть неологизмов того периода была призвана обслуживать процессы модернизации страны и возникла в ходе дальнейшего развития промышленной революции, пополнив словарь значительным количеством ранее не представленных лексических единиц, называющих новые реалии того времени, типа 自転車 ‘велосипед’, 政府 ‘правительство’, 電話 ‘телефон’, 消毒 ‘дезинфицировать’ и т.п.

## Классификация японских заимствований в китайском языке

Сформировавшийся в японском языке пласт лексики, отражавший процесс рецепции понятий западной культуры, науки и техники, через прогрессивных китайских студентов, обучавшихся в то время в Японии, постепенно проник и в китайский язык. Этот пласт был довольно гетерогенен и включал лексику практически всех областей жизни человека, не ограничиваясь научно-технической терминологией. Первую типологию заимствований в китайский язык предложили в 1958 г. Гао Минкай и Лю Чжэнтань [Гао, Лю 1958]. Начиная с этого момента последующие исследователи этого явления в китайском языке расходились как в общем числе типов заимствований, так и в подходах к их трактовке, связано это было со серьезными расхождениями в определении критериев классификации и значительной семантической размытостью самих заимствований.

Опишем типы языковых единиц, которые, как представляется, включает пласт японских заимствований в китайском языке.

Прежде всего, выделяются «**чистые японизмы**», описывающие элементы традиционной японской культуры. При отсутствии лексического эквивалента для реалии, когда не имеется более или менее полных соответствий в языке-реципиенте, имеет место прямое лексическое заимствование имеющейся в исходном языке единицы с фиксацией ее в иероглифической форме: 寿司 *shòusī* ‘суши’; 和服 *héfú* ‘кимоно’; 武士道 *wǔshìdào* ‘путь воина», кодекс самурая’; 柔道 *róudào* ‘дзюдо, вид японского боевого искусства’; 盆栽 *pénzāi* ‘บอนсай, миниатюрное растение, выращенное на подносе’.

Среди «чистых японизмов» имеется незначительное число прямых лексических заимствований, так, для традиционного мата из тростника был использован вариант записи 榻榻米 *tàtāmǐ* ‘татами’ с фиксацией оригинальной звуковой формы иноязычного элемента. Чаще всего этим способом заимствуются имена собственные, например, названия известных компаний: 夏普 *Xiàpǔ* ‘Sharp’, 雅马哈 *Yǎmǎhā* ‘Yamaha’, по частотности превышающий калькированный вариант 山叶 *Shānyè*, и др.

Японизмы, записываемые в иероглифике, заимствовались в исходном виде, но произносились уже по нормам китайского языка. В китайской лингвистике для этого типа единиц использовался термин 日語漢讀 *Rìyǔ Hàndú* ‘японские слова, читаемые по-китайски’; в отечественной синологии В. И. Горелов для них ввел в научный оборот термин «японизм» [Горелов 1984:147].

Огромный пласт неологизмов был заимствован в китайский из европейских языков через язык модернизированной Японии, выступивший в данном случае языком-посредником. В тот период на японский переводилось огромное количество западной научной литературы, в результате чего в язык вошло множество новейших терминов практически из всех областей знания, включая весь цикл точных и гуманитарных наук, военное дело, медицину и, конечно, новые реалии быта. Поскольку для записи новых терминов в японском языке использовалась иероглифика, идентичная форма фиксации неологизма упростила и ускорила дальнейшую рецепцию соответствующего иноязычного элемента китайским языком. Эти слова оказались формально неотличимы от исконно китайской лексики.

Кроме того, признаком калькирования принято считать наличие нехарактерных, нетипичных для языка-реципиента черт в формальной стороне вновь созда-

ваемого слова, когда при конструировании новой лексической единицы используются не только иноязычные структурные элементы, но и «чужеродная» для языка-реципиента словообразовательная модель. В период Мэйдзи образование новых слов и выражений в японском шло по словообразовательным моделям китайского языка, для которого, например, типична глагольно-именная модель словообразования, отражающая закрепившуюся на уровне структуры слова SVO тенденцию в базовом порядке элементов в предложении, которая не совпадает с базовым SOV порядком в японском [Кульнева 2018:26]. К этому типу относятся слова: 復員 *fùyuán* ‘демобилизация’, 任命 *rènmìng* ‘назначить на должность’, 自由 *zìyóu* ‘свобода’, 風琴 *fēngqín* ‘орган’, 商標 *shāngbiāo* ‘торговая марка’, 反感 *fǎngǎn* ‘неприязнь, антипатия’ и др.

Этот тип единиц создавался с целью перевода иностранных терминов путем комбинирования уже имевшихся в языке элементов как структурные и этимологические кальки соответствующих слов английского, реже других европейских (немецкого, французского, итальянского и латинского) языков. Для него в китайской лингвистике предлагался термин 新造借詞 *xīn zào jiècí* ‘вновь созданные заимствования’.

Примерами подобных заимствований из других европейских языков являются: 感性 *gǎnxìng* от (нем.) *Sinnlichkeit* ‘чувствительность’; 悟性 *wùxìng* от (нем.) *Verstand* ‘понимание’; 小夜曲 *xiǎoyèqǔ* от (фр. < итал.) *Serenade* ‘серенада’; 素描 *sùmiáo* от (фр.) *dessin* ‘рисунок, набросок’; 二重奏 *èrchóngzòu* от (итал.) *duetto* ‘дуэт’; 初夜權 *chūyè quán* от (лат.) *jus primae noctis* ‘право первой ночи’.

В этом типе обнаруживается небольшое число фонетических заимствований с разной степенью звукового соответствия исходному элементу в языке-доноре: для передачи термина ‘антимоний, сурьма’ от (англ.) *antimony*; 瓦斯 *wǎsī* для передачи термина ‘газ’ от (фр.) *gaz*; 加答兒 *jiādá'ér* для передачи термина ‘катар’ через (фр.) *catarrhe* от (др.-греч.) *κατάρρη*; 基督 *jīdū* для передачи термина ‘Христос, помазанник’ через (лат.) *Christus* от (др.-греч.) *Χριστός*; 沃度 *wòdù* для передачи термина ‘иод’ от (фр.) *Jod*.

Третий тип заимствованных единиц представляет собой результат лексической деархаизации, вернувшей в активное употребление слова или отдельные их элементы (морфемы), когда-то бытовавшие в китайских текстах и получившие в Японии конца XIX – начала XX вв. новую жизнь и иное, ранее не имевшееся значение, в котором они и вернулись обратно в китайский язык. Эти единицы из древне- и среднекитайского языков были использованы носителями японского языка для передачи западных терминов, через который они попали в родной язык уже как семантические неологизмы. Этот тип заимствований в китайской науке получил название 回归借詞 *huíguī jiècí* ‘вернувшиеся заимствования’. Пример подобного возвращения в активное употребление когда-то в той или иной степени устаревших и отодвинутых на периферию лексической подсистемы слов или их частей представляют такие языковые единицы, как:

革命 *gémìng* встречается в тексте «И цзин», в древнекитайском языке имело значение «сменить волю неба», описывая свержение династии. Данная единица



была заимствована японским языком для перевода английского слова *revolution* ‘революция,’ после чего вернулась в язык-источник с измененной семантикой;

悲观 *bēiguān* изначально буддийский термин «смотреть на все живое с состраданием, стремление спастись от горя и избавлять от бед» был выбран для передачи значения *pessimistic* ‘пессимистичный’;

分子 *fēnzǐ* изначально «потомство по внебрачной линии» стало использоваться для перевода термина *molecule* ‘молекула’;

物理 *wùlǐ* изначально «законы вещей» – для передачи *physics* ‘физика’ и др.

Многие памятники древнекитайской письменной традиции стали источниками деархаизованных в эпоху Мэйди лексических единиц. Так, в одних «Исторических записках» Сыма Цяня содержится свыше десятка этимонов, к которым возводят такие мэйдзийские неологизмы, как: 共和 *gònghé* для передачи *republic* ‘республика’; 教授 *jiàoshòu* для передачи *professor* ‘профессор’; 遺傳 *yíyùn* для передачи *heredity* ‘наследственность’; 經費 *jīngfèi* для передачи *expenses* ‘расходы’; 交通 *jiāotōng* для передачи *traffic* ‘дорожное движение’; 市長 *shìzhǎng* для передачи *mayor* ‘мэр, городской глава’; 主義 *zhǔyì* для передачи *principle* ‘принцип’; 侵犯 *qīnfàn* для передачи *violation* ‘нарушение’; 侵略 *qīnlüè* для передачи *invasion* ‘вторжение’; 生產 *shēngchǎn* для передачи *production* ‘производство’; 天主 *tiānzhǔ* для передачи *Deus* ‘Бог’ и др.

Трактат «Мэн-цзы» послужил источником для деархаизованных слов 會計 *kuàijì* для передачи *accounts, book-keeper* ‘бухгалтерия; бухгалтер’, 具体 *jùtǐ* для передачи *concrete* ‘конкретный’, 交際 *jiāoji* для передачи *association, social intercourse* ‘общение, социальные контакты’ и др.

### Формирование терминологического аппарата японских заимствований в синологии

Описание ксенолексики – в нашем случае японских заимствований в китайском языке – невозможно без тщательного разграничения смежных понятий, которое, с одной стороны, должно учитывать лингвоспецифические черты соответствующих единиц в системе, с другой стороны, опираться на уже сложившийся в языкознании понятийный и терминологический аппарат. Лексические единицы, пришедшие в китайский язык через японский, длительное время были довольно скудно описаны в отечественной и западной синологии. Основной причиной этого, как уже отмечалось выше, являлся слабый интерес к этой теме со стороны самих носителей языка вообще и китайских лингвистов в частности. В XX в. изучением японских заимствований в родной язык занимались лишь немногие китайские ученые, чем, по-видимому, и можно объяснить крайнюю малочисленность работ по данной проблематике и за пределами Китая.

В отечественной синологии заимствования из японского языка попадали в фокус научного интереса считанное число раз, в частности в работах В. В. Иванова, В. И. Горелова, Н. В. Перфильевой, О. А. Бехтеревой, П. В. Кульневой, Е. Д. Кругловой и некоторых других исследователей, в том числе носителей китайского языка, пишущих по-русски.

В. В. Иванов ввел в научный оборот синологии термин «вторичное заимствование», охарактеризовав пласт японских заимствований следующим образом: «В китайском языке встречается довольно много слов, пришедших из японского языка, но в самом японском языке большинство из них представляет собой не исконно японские слова, а иностранные заимствования. В данном случае имеет место вторичное освоение иноязычной лексики, и поэтому лексические единицы этого рода можно назвать «вторичными заимствованиями» [Иванов 1973: 45].

В. И. Горелов, следуя терминологии В. В. Иванова, применял термин «вторичные заимствования» и к собственно чистым японизмам, и к словам, пришедшим из японского языка, но в самом японском языке представляющим собой не исконно японские слова, а иностранные заимствования. Характеризуя данную лексику, синолог подчеркивал, что «будучи записаны китайскими иероглифами, они по своему внешнему виду, по словообразовательной структуре ничем не отличаются от исконно китайских слов. Их иногда называют 日语汉读 (японские слова, произносимые по-китайски)» [Горелов 1984: 147].

Бытующий в китаеведческой науке термин «обратное заимствование» некоторыми исследователями позиционируется как синонимичный «вторичному заимствованию». Так, О. А. Бехтерева относит его к диахроническому процессу, в ходе которого какой-либо элемент заимствуется другим языком в одну эпоху, а затем возвращается в него уже в другую эпоху из заимствовавшего его когда-то языка [Бехтерева 2016].

Ряд отечественных исследователей следует традиции западной синологии и считает японские заимствования исключительно графическим типом. Так, Н.В. Перфильева и Ху Пэйпэй рассматривают понятие графического заимствования в узком и в широком смысле. В узком понимании графические заимствования в китайском языке – это слова, у которых форма написания (иероглифы) были полностью заимствованы из японского языка. В широком понимании речь идет о сохранении исходной графики заимствованного слова в китайском языке [Перфильева, Ху 2018].

Этот подход занимает основное место в западном китаеведении, где для анализируемого типа единиц предлагались термины – graphic loans (графические заимствования) у Ф. Масини (Federico Masini) [Masini 1993], symbolic loans, emblematic loans (символьные заимствования) у Э. Кук (Angela Cook) [Cook 2018], и противопоставляются эти единицы фонематическим (phonemic loans) и фонетическим (phonetic loans) заимствованиям соответственно. Аналогичной позиции придерживается Чэнь Хайцзинь, которая японские заимствования (japanese loan-words) относит к графическому (graphic) подтипу, противопоставляя их фонематическим (phonemic) заимствованиям [Chen 2014: 17].

В Японии процесс калькирования европейской терминологии с использованием *канго* в условиях насущной потребности в номинации того периода принято рассматривать как особый тип заимствований, который Мориока Кэндзи назвал «замещением», поскольку речь шла о вытеснении исходной семантики языковой единицы новым значением. Согласно Мориоке, «на первом этапе большая часть переводных слов зависела от китаизмов канго периода Эдо. Затем постепенно увеличивалось число новых канго, и появилась тенденция к избеганию «замеще-

ния» (цит. по [Демич, Костыркин 2020]). По мнению японоведов, в дальнейшем эту функцию взяло на себя «традиционное и осложненное словообразовательное калькирование» [Демич, Костыркин 2020: 169].

Неустоявшаяся в лексикологии китайского языка терминология имеет следствием серьезные расхождения с тем аппаратом, который сложился в отечественной лексикологической традиции, не вполне совпадает она и с видением данной проблематики у современных китайских исследователей.

Во-первых, материальное заимствование в естественных языках может идти устно, то есть происходить «на слух», зачастую без учета специфики письменной фиксации в языке-источнике, это типично для индоевропейских языков и практически неактуально для анализируемых в данном исследовании языков Дальнего Востока. Материальным будет и заимствование слов в графической форме, когда рецепция идет через письменные тексты, что с учетом графического облика лексической единицы является принципиальным моментом для языков-членов «иероглифического клуба». В китайском языке новые лексические единицы, в эпоху Мэйдзи пришедшие из Японии, осваивались сугубо через письменность: общность систем письма минимизировала потребность в рецепции японских заимствований, которая на современном этапе сводится к замене полных форм иероглифических знаков на упрощенные, когда в ряде случаев могут наблюдаться изменения в написании соответствующих иероглифов, но не их полной замене.

Во-вторых, для материала китайского языка, как и для других естественных языков, представляется релевантным разграничивать прямое (непосредственное) и опосредованное заимствование. По своей траектории заимствование может быть прямым или опосредованным, когда из языка-донора в язык-реципиент слово заимствуется через второй или даже третий языки-посредники. Если учитывать этот фактор, то большая часть заимствований эпохи Мэйдзи будет относиться к опосредованным заимствованиям.

Такой подход обнаруживается у китайских исследователей, в частности Цэнь Цисянь (岑麒祥) считает мэйдзийские неологизмы непрямыми заимствованиями (间接外来词 *jiànjiē wàiláicí*) [Цэнь 1990: 3], поскольку здесь имеет место результат опосредованной рецепции лексики из европейских языков, при котором японский язык выступил лишь языком-посредником, но никак не собственно донором. Для заимствований опосредованного типа – из западных языков через японский в китайский – можно говорить о заимствовании из исходного языка-донора путем калькирования и перенимания графической формы в языке-посреднике.

В-третьих, в отечественной лексикологии с подачи Л. П. Крысина термином «вторичное заимствование» принято обозначать лексические единицы, которые появились в языке «наряду с ранее заимствованным и ассимилированным в языке словом ..., по форме совпадающее с ранее заимствованным, но имеющее иное значение, вплоть до полной омонимии» [Крысин 2004: 143]. Пример вторичных заимствований в русском языке являют слова *адрес, пилот, паста, винчестер* и многие другие, которые когда-то вошли в словарный фонд извне, были полностью освоены и зафиксированы в лингвистических словарях и справочниках как ксенолексика, а затем оказались заимствованы еще раз, но уже в ином, ранее не известном значении.

Анализируемая лексика китайского языка представляет собой феномен иного типа, поэтому применительно к заимствованиям из японского языка в китайский следует отказаться от неудачного в данном контексте термина «вторичные заимствования». За исключением лексической деархаизации, во всех остальных типах заимствований в языке отсутствуют совпадающие по форме, но отличные по значению лексические единицы, а именно так понимается термин «вторичное заимствование» в отечественной лингвистике XXI в., предполагая неоднократный перенос различных значений на одну и ту же материальную форму, чего не наблюдается у большей части заимствований в анализируемый период. И даже при деархаизации, с учетом того разрыва, который имеется между современным китайским языком и древним (или даже среднекитайским), вряд ли можно сопоставлять эти языковые единицы как существующие в рамках одной лексической подсистемы. Все-таки это слишком далеко разошедшиеся системы.

Более адекватным представляется термин «обратное заимствование», который предполагает ситуацию, когда какое-то слово образуется в данном языке из иноязычных элементов, а затем оно заимствуется обратно в тот язык, из элементов которого была сконструирована единица, или в его современный потомок. Ярким примером обратного заимствования является существительное *шапка* в русском языке, заимствованное опосредованно из французского (через польский и немецкий языки) в значении «головной убор европейского образца» и вернувшееся через несколько столетий во французский язык как *шарка* для обозначения ушанки или зимнего головного убора на меху в русском стиле [Плунгян 2010: 16].

Как представляется, применительно к анализируемому материалу к «обратным заимствованиям» в строгом смысле можно отнести только вышеописанные случаи лексической деархаизации, тогда как калькирование понятий западной науки и техники, которые вошли в китайский язык опосредованно через японский, где они были зафиксированы в иероглифической форме, может считаться «обратным заимствованием» в широком понимании. Использование соответствующей графической оболочки можно рассматривать как процесс графического заимствования, как это принято в западной китаеведческой науке, однако вряд ли уместно в данном случае говорить о «вторичном» или «повторном» заимствовании. Второй тип рассмотренных нами японских заимствований в китайском языке точнее было бы определить как «обратные графические заимствования», поскольку при семантическом калькировании, предполагающем расширение смыслового потенциала уже имеющихся в системе элементов за счет значений, присвоенных/переданных им от их переводных эквивалентов в языке-доноре, наблюдается значительный сдвиг или даже полное несовпадение смыслов между тем, что было, и тем, что стало, часто вызывая у носителей языка когнитивный диссонанс, поскольку знакомые им элементы оказываются в непривычном сочетании или в нетипичном значении.

\*\*\*

Таким образом, проведенный нами анализ подтверждает необходимость пересмотра имеющейся в отечественном востоковедении терминологии для заимствований из японского языка в китайский. Классификация этих единиц по семанти-

ческому критерию позволяет выделить среди них те типы, которые неправомерно фигурируют в синологической литературе как «вторичные заимствования». Учет накопленного в современной лексикологии понятийного аппарата может помочь в определении заимствований в китайский из европейских языков через японский как опосредованных и реализуемых путем семантического калькирования. При этом только один из выделенных нами типов, а именно лексических единиц, возвращенных к жизни путем лексической деархаизации слов и выражений древне- и среднекитайского языков, будет относиться к «обратным заимствованиям в узком смысле». Оставшийся тип, реализованный за счет типичного для естественных языков механизма расширения семантического потенциала уже имеющихся в системе элементов за счет «чужеродных» значений из другого языка, будет описан как «обратные графические заимствования» с указанием лингвоспецифических черт данного явления в китайском языке.

### Литература

1. Аракава Ёсико. Неологизмы эпохи Мэйдзи как зеркало общества // *ЯПОНИЯ: 150 лет революции Мэйдзи. Исследования российских и зарубежных учёных, посвящённые 150-ой годовщине революции Мэйдзи. Issues Of Japanology*. 2018. № 7. СПб.: Art-xpress. С. 42–53.
2. Бехтерева О.А. *Японские заимствования в китайском языке: периодизация и адаптация*. Иркутск: ИГЛУ, 2016 .
3. Горелов В.И. *Лексикология китайского языка*. М.: Просвещение, 1984.
4. Демич К.И., Костыркин А.В. Особенности становления современной японской философской терминологии в эпоху Мэйдзи по лексикографическим данным // *Труды Института востоковедения РАН. Проблемы общей и востоковедной лингвистики. Язык меняющийся: семантические и грамматические изменения в языках Азии и Африки*. № 29. 2020. С. 167–178.
5. Иванов В.В. *Терминология и заимствования в современном китайском языке*. М.: Наука, 1973.
6. Крысин Л.П. Вторичное заимствование и его описание в толковом словаре // *Русский язык сегодня*. №3. М., 2004. С. 143–148.
7. *Лингвистический энциклопедический словарь (ЛЭС)* / В. Н. Ярцева (гл. ред). М.: Советская энциклопедия, 1990.
8. Маевский Е.В. Лексическая «Революция Мэйдзи» // *Актуальные вопросы японского и общего языкознания: памяти И. Ф. Вардуля*. М., 2005. С.245–255.
9. Перфильева Н.В. Различные трактовки заимствования в работах китайских лингвистов // *Современные исследования социальных проблем (электронный научный журнал)*, 2012. № 8.  
URL <http://sisp.nkras.ru/e-ru/issues/2012/8/hu.pdf>. Дата обращения 21.02.2021.
10. Перфильева Н.В., Ху Пэйпэй Способы ассимиляции английских заимствований в китайский язык // *Вестник РУДН. Серия: Теория языка. Семиотика. Семантика*. 2018. №2. С. 287–301.
11. Плуноян В.А. *Почему языки такие разные*. М.: АСТ-пресс книга, 2010.

12. Семенас А.Л. Особенность лексических заимствований в китайском языке // *Вопросы языкознания*. 1997. № 1. С. 48–57.
13. Чжан Кэ. Понимание термина заимствованное слово в русской и китайской лингвистике // *Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика*. 2017. Т. 8. № 3. С. 724–734.
14. Chen, Haijing *A Study of Japanese Loanwords in Chinese*. University of Oslo: Master's Thesis. 2014. 125 p.
15. Cook, Angela A typology of lexical borrowing in Modern Standard Chinese // *Lingua Sinica*. 2018. Vol. 4, article number 6.
16. Masini, Federico. The Formation of Modern Chinese Lexicon and Its Evolution Toward a National Language: The Period from 1840 to 1898. // *Berkeley: Journal of Chinese Linguistics Monograph Series* № 6. 1993. 295 p.
17. 岑麒祥 《汉语外来语词典》北京商务印书馆. 1990 年. (Цэнь Цисянь *Словарь заимствованных слов в китайском языке*). Пекин: Коммерческая пресса, 1990. 497 с.
18. 高名凯、刘正琰 《现代汉语外来词研究》文字改革出版社. 1958 年. (Гао Минкай, Лю Чжэнтань. *Исследование заимствованных слов в современном китайском языке*). Пекин: Реформа иероглифической письменности, 1958. 189 с.
19. 王力 《漢語史稿》科學出版社, 1958 年. (Ван Ли. *История китайского языка*). Пекин: Наука, 1958. 613 с.

**Параметрический анализ  
«Санскритско-русского словаря» В.А. Кочергиной**

**Алексей Александрович Кретов**

Воронежский государственный университет, Воронеж, Россия,  
kretov@rgrph.vsu.ru, ORCID 0000-0002-1474-3177

**Марцис Юрьевич Гасунс**

независимый исследователь, Краснодар, Россия,  
gasyoun@ya.ru, ORCID 0000-0003-4513-884X

**Владимир Владимирович Леонченко**

независимый исследователь, Долгоруково, Россия  
lymagic@ya.ru, ORCID 0000-0002-0334-7507

***Аннотация.*** Цель анализа – выявить параметрическое ядро санскритской лексики по данным словаря-источника. Вводится понятие лексико-семантической системы. Системообразующими считаются синтагматические (А и Б), парадигматические (А или Б), эпидигматические (если А, то Б) связи слов. Функциональная нагрузка на слова косвенно оценивается по длине леммы. Обосновывается преимущество длины слова как показателя его функциональной активности. Выбор объекта исследования мотивируется необходимостью единого метаязыка (русского языка) для типологических сопоставлений. Осуществлено «взвешивание» каждой леммы словаря по четырём параметрам (синтагматическому, парадигматическому, эпидигматическому и функциональному), а также по интегральному параметрическому весу. Проведена стратификация словаря на малое, среднее, большое ядра, на ядро и периферию.

***Ключевые слова:*** санскрит, параметрический анализ лексики, лексико-семантическая система, ядро санскритской лексики, корпус текстов, санскритская лексикография

***Abstract.*** The purpose of the analysis is to identify the parametric core of the Sanskrit vocabulary according to the data of the source dictionary of V.A. Kochergina. The concepts of the lexico-semantic system are introduced. Syntagmatic (A and B), paradigmatic (A or B), epidigmatic (if A, then B) connections of words are considered system-forming. The functional load on words is indirectly estimated by the length of the lemma. The advantage of the word length as an indicator of the functional activity of the word is justified. The choice of the object of research is motivated by the need for a single metalanguage (Russian) for typological comparisons. Each lemma of the dictionary is «weighed» by four parameters (syntagmatic, paradigmatic, epidigmatic and functional), as well as by an integral parametric weight. The stratification of the Sanskrit dictionary into small, medium, large kernels, into the core and periphery of the dictionary is obtained.

***Keywords:*** Sanskrit, parametric analysis of vocabulary, lexico-semantic system, the core of Sanskrit vocabulary, the corpus of texts, Sanskrit lexicography

## 1. Теоретические основания исследования: словарь или корпус

Как известно, в непосредственном наблюдении нам дана только речь, являющаяся результатом использования языка. В этом смысле позиция лингвиста средни позиции охотника, восстанавливающего происходившие события по оставленным следам, или позиции палеозоолога, по окаменевшему отпечатку ископаемого ящера реконструирующего его облик, или позиции медика-диагноста, по наблюдаемым проявлениям болезни диагностирующего заболевание.

Результатом реконструкции языка по оставленным им следам-текстам является комплект из грамматики и словаря данного языка, в связи с чем предложенная Ю.Д. Апресяном [Апресян 1995] идея «интегрального описания языка» (словарь и грамматика образуют единство и «настроены» друг на друга, так что грамматические правила формулируются с учётом ограничений, налагаемых словарными свойствами единиц или классов единиц, а в словарном описании указываются те типы грамматических правил, в которых данное слово может участвовать) обретает особое значение.

Таким образом, в непосредственном наблюдении язык дан в виде реконструкта – словаря и грамматики, являющихся отражением представлений лингвистов, сформировавшихся, с одной стороны, под влиянием их теоретических взглядов, а с другой стороны – под влиянием *результата функционирования языка*, каким является больший или меньший корпус *текстов*. Поскольку тексты являются *отражениями* языка, а словари – *отражениями* текстов, то лексико-семантическая система языка в словаре является «отражением отражения».

Корпус текстов ни в коей мере не рядоположен словарю и грамматике. Это всё равно что кусок сырой баранины считать люля-кебабом или окаменевший след ископаемого ящера приравнять к трёхмерной реконструированной фигуре этого существа.

Взгляд на лингвистические корпусы как на своего рода лингвистическую панацею, сложившийся у некоторых лингвистов под впечатлением от возможностей новых информационных технологий, сопровождаемый недооценкой словарей и переоценкой корпусов, является своего рода «детской болезнью новизны» в лингвистике, основанной либо на недостаточным знакомстве с технологией создания словарей, либо – на сознательном игнорировании этой технологии. «Традиционные» словари производны от своих картотек, а многомиллионные картотеки появились в результате расписывания огромного множества текстов, образующих *корпус данного словаря*. Отдавать предпочтение корпусу перед словарём – всё равно что отдавать предпочтение набору сырых продуктов перед готовым блюдом. Да, в корпусе набор карточек с контекстами нужного слова можно получить почти мгновенно. (Впрочем, извлечение употреблений данного слова из картотеки также занимает не много времени). А дальше – всё то же самое: ручная работа лексикографа по выделению значений и фразеосочетаний, дефинированию их, поиску синонимов, антонимов, гиперонимов и других близких по форме, значению и происхождению слов.

Практика лексикографирования, насчитывающая многие века, представляет собой процесс накопления – кумуляцию – знаний о лексиконе данного языка, ос-



нованный на наблюдении над всё возрастающим множеством текстов. Каждый последующий словарь данного языка является очередной ступенькой на лестнице познания его лексико-семантической системы.

На рубеже 1950-1960-х гг. лингвистика переживала сходное увлечение частотными словарями. Это увлечение позволило понять, что частотность слова зависит от корпуса текстов, на котором производится подсчёт. В.В. Морковкин, сопоставив частоты русских слов в восьми частотных словарях, установил, что у многих слов частота оказалась неподтверждённой данными других словарей: «Все слова, характеризующиеся... индексами от 1 [первая сотня самых употребительных слов – здесь и далее выделено нами – А.К., М.Г., В.Л.] до 35 [тридцать пятая по употребительности сотня], обнаружены среди 2500 самых употребительных лексических единиц не менее чем у двух частотных словарей, что позволяет считать их словами с подтверждённой высокой употребительностью. Индекс «36-60» дается при словах (их около 2400), которые встретились среди первых 2500 лексических единиц только в одном из частотных словарей-источников. Лексические единицы с индексом 36-60 считаются словами с неподтверждённой высокой употребительностью» [Морковкин 1985: 18-19].

Разнобой частотных словарей в определении частоты слов – следствие различия корпусов, на которых определялась частотность слов (если оставить в стороне возможные ошибки в лемматизации корпусных словоформ).

Ни один даже самый большой корпус не может претендовать на равенство генеральной выборке, содержащей все тексты на данном языке. Следовательно, центральным в оценке лингвистических корпусов является далеко не самый обсуждаемый и не самый популярный в корпусной лингвистике вопрос – вопрос их репрезентативности, их близости к генеральной выборке. Насколько нам известно, критерии определения и оценки репрезентативности корпусов так и не сформулированы. Поэтому претензии, предъявляемые энтузиастами корпусов к словарям, могут быть с большим основанием предъявлены самим лингвистическим корпусам.

У корпуса есть преимущество перед словарем: он содержит все слова, а академический словарь – только те, которые встретились не менее, чем у трёх авторов. Это при том, что слова с частотой 1, например, в «Частотном словаре русского языка» под ред. Л. Н. Засориной (М., 1977) составляют 34% словника, а слова с частотой 1-2 – почти 49%. Корпус превосходит словарь в охвате периферии. Поскольку нас интересует ядро лексики, это преимущество корпуса перед словарём несущественно.

Ещё одно преимущество корпусов перед словарями в том, что создание словарей требует годов, а то и десятилетий. На это время корпус и картотека словарей (пусть и частично) консервируются, и словари отстают от речевой практики, что может давать видимость основания для упреков словарей в неактуальности, нерепрезентативности.

Всё, что лингвистике известно про скорость протекания лексико-семантических процессов, не позволяет согласиться с таким упреком. Молодёжные жаргоны, по свидетельству социолингвистов, действительно, обновляются каждые пять лет. Литературные же языки, в силу своей нормированности и кодификации, изменяются медленнее, чем разговорная речь. Один из факторов, ограничиваю-

щих скорость изменений в языке, – продолжительность человеческой жизни: и в детстве, и в старости человек должен понимать людей всех возрастов. Средняя продолжительность жизни в России – 73 года<sup>1</sup>. Но человек на протяжении жизни должен понимать и тех, кто старше его на 70 лет, и тех, кто моложе его на 70 лет. Следовательно, русский язык в принципе не может претерпеть существенных изменений на протяжении как минимум (в действительности – значительно более) 140 лет, иначе представители разных поколений не смогут понимать друг друга, а язык не сможет выполнять своей главной функции – коммуникативной. Самые быстрые существенные изменения в языке, известные науке, занимали не менее 200 лет.

Глоттохронологией накоплен богатый материал о скорости изменений в лексике. Например, в исландском языке 100-словный список М. Сводеша за последнюю тысячу лет изменялся со скоростью 0,04 [Bergsland 1962]. С.А. Старостин, исследовав языки с достаточной письменной историей, установил, что на протяжении последних 1000-1500 лет их базисная лексика изменялась со скоростью 0,05. [Старостин 2007: 416-419]. Если за последние 1000 лет базисная лексика языков менялась со скоростью 5% в тысячелетие, то за 100 лет в ядре лексики из 1000 слов должно было измениться пять слов. Соответственно, за последние 200 лет, по С.А. Старостину, ядро русской лексики в 1000 слов могло измениться на десять слов. Вряд ли это основание считать «нерепрезентативным», например, «Словарь языка Пушкина» и любой другой словарь русского языка сопоставимого или большего объёма.

Таким образом, единственным источником информации об устройстве лексико-семантической системы языка являются словари (в идеале – толковые и созданные на их базе двуязычные).

Далее необходимо обсудить два вопроса: 1) что такое лексико-семантическая система и 2) при каких условиях возможна типология лексико-семантических систем. Начнём с первого.

## 2. Что такое лексико-семантическая система?

Любая система представляет собой единство элементов и структуры (связей, объединяющих элементы в целое).

**Элементом** лексико-семантической системы (ЛСС), отражённой в словаре, условимся считать словарную статью, представленную леммой (формой, представляющей слово во всей его тотальности).

Что касается типов связи между леммами, то их три.

1) **Синтагматические** связи, соответствующие логическому оператору «И»: *А и Б – синтагма* (пара элементов, объединённых синтагматическими связями). Для единиц ЛСС синтагматическими связями являются синтаксические связи. Эти связи (в отличие от парадигматических) даны в непосредственном наблюдении.

2) **Парадигматические** связи соответствуют логическому оператору «ИЛИ» *А или Б – парадигма*. Парадигматические отношения – это отношения выбора, конкуренции. Например,

---

<sup>1</sup> [https://ruxpert.ru/Статистика:Продолжительность\\_жизни\\_в\\_России](https://ruxpert.ru/Статистика:Продолжительность_жизни_в_России).

Мы гуляли по Неглинной,  
Заходили на бульвар,  
Нам купили синий-синий  
Презелёный красный *шар!* (С. Михалков, А что у вас?)

Понятно, что шар не может быть одновременно синим, зелёным и красным. Он или синий, или зелёный, или красный. Прилагательные тут находятся в отношениях конкуренции, т.е. в парадигматических отношениях, а узнаём мы об этом потому, что все они сочетаются с одним и тем же существительным – шар. Так синтагматика оказывается проявлением парадигматики. В парадигматических отношениях находятся те слова, у которых сходная, близкая синтагматика. Синтагматика и парадигматика – разные стороны одного феномена, соотносящиеся как явление (синтагматика) и сущность (парадигматика), форма (синтагматика) и содержание (парадигматика) во всей диалектической сложности их взаимоотношений, сознавать которую необходимо, а преувеличивать и абсолютизировать – едва ли. Чтобы подчеркнуть единство синтагматики и парадигматики, можно по аналогии с термином М. М. Бахтина «хронотоп», предложить термин «синдигматика».

3) **Эпидигматические** (термин Д. Н. Шмелёва) связи – это «деривационные в широком смысле» связи, соответствующие логическому оператору импликации «Если..., то...». *Если А, то В.* Если есть слово *берёзовый*, то есть и *берёза*. Если есть *берёза*, то (в системе языка, являющейся системой возможностей) есть и *берёзовый*.

Это относится не только к производным формам (леммам), но и к производным значениям (многозначности). В двуязычных словарях в явном виде представлена *эпидигматика* (внутренняя деривация), поэтому термины *эпидигматический* и *деривационный* в параметрическом анализе лексики употребляются как синонимы.

4) Тремя описанными типами связей исчерпывается система языка в состоянии покоя. Но в таком состоянии находится только мёртвый язык – язык, на котором не создаётся новых текстов. Для живого языка необходима ещё одна характеристика – **функциональная**. Эта характеристика не дискретна: 0/1, а континуальна: 0→1. В непосредственном наблюдении такой параметр слов, как функциональная активность не дан. Он измеряется косвенно. Самый известный (часто – единственный известный) способ косвенной оценки функциональной активности слов – их частота в том или ином корпусе, которая может быть переведена в относительную частоту (например, *УнаМ* – количество *употреблений на миллион*) или ранг частоты (чем меньше ранг слова в частотном словаре, тем выше его функциональная активность).

Хотя показания частотных словарей сильно зависят от выборки (корпуса текстов), на которой они получены, это не мешает большинству лингвистов экстраполировать их показания на язык (точнее – генеральный корпус) в целом.

Более надёжным представляется второй способ косвенной оценки функциональной активности слов. Ещё Дж. К. Ципф установил зависимость между длиной слов и их встречаемостью в тексте: с возрастанием средней длины слов убывает их частота [Zipf 1949].

Существует оптимум между длиной слова и его частотностью. Если употребительность слова вступает в противоречие с его длиной, то оптимум достигается употреблением сокращённой формы слова. Так, например, *Санкт-Петербург* сначала сократился до *Петербурга*, а затем – до *Питера*. Название столицы Мадагаскара – *Антананариву* сократилось в обиходе малагасийцев до *Тана*. Предельной формой реализации этой тенденции являются аббревиатуры.

Для того, чтобы произошла компрессия формы слова, необходим длительный период аномально высокой (при такой длине) употребительности слова. Поэтому малая длина слова (в общем случае) свидетельствует о его высокой употребительности или, как минимум, о том, что в прошлом это слово очень долго было употребительным. Исключения затрагивают преимущественно заимствования (такие, как *ас, до, ре, ми, фа, соль, ля, си, су, сэр, хан, чай* и т.п.), которые не бывают одновременно и короткими, и многозначными, и богатыми синонимами и сочетаемостью.

У второго способа косвенной оценки функциональной активности слов, кроме независимости от выборки, есть и другое важное для лексико-семантической типологии достоинство: частотные словари имеются для незначительного количества языков, тогда как длина слова в любом языке – совершенно объективная и верифицируемая характеристика, которая всегда при нём.

Таким образом на вопрос: «Что такое лексико-семантическая система?», мы даём ответ: «ЛСС – это множество словарных статей, описывающих слова данного языка, объединённое синтагматическими, парадигматическими и эпидигматическими связями», а на вопрос: «Что такое функционирующая лексико-семантическая система?», добавляем: «характеризующееся неравномерной функциональной нагруженностью элементов, косвенно оцениваемой по их длине (в словаре – по длине лемм)».

### **3. Возможна ли типология лексико-семантических систем?**

На этот вопрос дают два ответа. Первый из них: «...в сравнении с фонологией и грамматикой, лексика представляет собой более слабую и нечеткую, в большей мере вероятностную систему, значительно меньше структурированную. Поэтому, в отличие от фонетико-фонологической и грамматической типологии, в лексической типологии отсутствуют и, по-видимому, невозможны классификации языков» [Мечковская 2000: 106].

Мнение Н.Б. Мечковской, разделяемое многими, не учитывает иерархической и полевой организации словаря, в котором важнейшей является оппозиция «ядро : периферия». Ядро лексики устойчиво и структурировано. Периферия – изменчива и нечётка. Ядро мало по размеру, периферия – безгранична. Ядром владеет каждый индивид, периферией – все общество, пользующееся данным языком. Поэтому в лексической типологии классификации языков возможны, но в основу их должны быть положены ядра лексико-семантических систем. Таким образом, важнейшим шагом на пути к лексико-семантической типологии является выделение ядер лексико-семантических систем языков мира.

Познавать лексику – значит минимизировать её. Параметрический анализ позволяет представить словарь в виде конуса, вершиной которого является одно-

единственное слово. Такое слово вбирает в себя весь словарь: его позиция определяется всеми остальными словами словаря, взятыми по четырем системообразующим параметрам. Поэтому доминанта словаря несет не меньше информации о нем, чем пик Джомолунгмы о горе в целом.

Имея на карте координаты одних пиков гор, мы можем в общих чертах представить себе очертания горной системы. Аналогично, и слова-доминанты дают представление о своеобразии лексико-семантической системы. Пик – точка, вобравшая в себя всю гору. **Доминанта** – слово, вобравшее в себя весь словарь. Более богатую информацию дают малое, среднее и большое ядра лексико-семантической системы, но не нужно думать, будто бы «большие словари, подобно большим батальонам, всегда правы». В больших словарях системное ядро лексики заведомо присутствует, но оно тонет в море периферийной (заимствованной и производной) лексики. Более того, на периферии лексико-семантической системы границы между языками размыты – именно из-за обилия заимствованной лексики. Самый наглядный пример – английский язык с его германским лексико-семантическим ядром и романской периферией. Следовательно, выбирая в качестве источника большой словарь, мы обрекаем себя на сизифов труд: выделить 1000 слов из 23 000 легче, чем из 80 000 или 125 000. Использование больших словарей вместо средних/малых подобно использованию карты крупного масштаба со множеством подробностей, когда надо составить представление об общем виде объекта, взглянув на него с высоты птичьего полёта.

Выделению системного ядра лексики и её стратификации по системообразующему весу помогает применение параметрического анализа лексики, подробно описанного в работах В.Т. Титова и примененного им к материалу романских языков [Титов 2002; 2005]. На сегодняшний день параметрический анализ лексики применен к 60 языкам Евразии [Кретов 2016; Кретов 2017].

Кумулятивная организация лексики по принципу матрёшки (ядро в ядре) хорошо известна преподавателям иностранных языков: «Коллекцию градуальных лексических минимумов составляют семь списков: 1-й – 500 самых употребительных русских слов, 2-й – 1000 самых употребительных русских слов, <...> 7-й – 3500 слов, причём каждый предыдущий список целиком включается в последующий» [Морковкин 1985: 10]<sup>2</sup>.

«Лексический минимум элементарного уровня составляет 780 единиц. Лексический минимум базового уровня составляет 1300 единиц<sup>3</sup>».

Вот другой источник. «1500-2000 слов. Активный словарный запас. Такого количества слов вам хватит, чтобы общаться в течение всего дня. Единственный минус – общаться сможете только на повседневные темы»<sup>4</sup>.

---

<sup>2</sup> То, что в данном случае речь идёт о ранжировании по частотности, непринципиально: важна формулировка кумулятивного принципа организации словаря.

<sup>3</sup> Государственный стандарт по русскому языку как иностранному. Элементарный уровень / Владимирова Т.Е. и др. – 2-е изд., испр. и доп. – М., СПб.: Златоуст, 2001, С.16. [http://gct.msu.ru/docs/A1\\_standart.pdf](http://gct.msu.ru/docs/A1_standart.pdf)

<sup>4</sup> <https://enguide.ua/magazine/skolko-slov-neobhodiom-znat-na-angliyskom-dlya-normalnogo-obshcheniya>

Не противоречит предыдущим и мнение шведского полиглота Эрика В. Гуннемарка: «400-500 слов – активный словарный запас для владения языком на базовом (пороговом) уровне. 800-1000 слов – активный словарный запас для того, чтобы объясниться; или пассивный словарный запас для чтения на базовом уровне. 1500-2000 слов – активный словарный запас, которого вполне хватит для того, чтобы обеспечить повседневное общение в течение всего дня; или пассивный словарный запас, достаточный для уверенного чтения. <...> овладев 1500 правильно выбранными словами, при условии некоторой дополнительной тренировки вы сможете общаться практически свободно»<sup>5</sup>.

В обоснование 1000 как размера параметрического ядра лексики В.Т. Титов приводит следующие аргументы: «в общенациональном китайском языке – *путунхуа* – 1324 разных слога... каждый слог – звуковая оболочка морфемы или **простого слова**» [ЛЭС 1990:225]», «первоначальное количество корней в эсперанто – 927 [ibid:594]; <...> «...в тезаурусе Роже [Roget's 1986] <...> 990 семантических инвариантов (рубрик словаря)» [Титов 2002: 8].

Как видим, 1000 слов – вполне достаточное множество для характеристики ядерной лексики языка. Понятно, что 1000 слов – величина условная, но очень полезная в качестве ориентира. Так, например, полевая лингвистика рекомендует для словаря бесписьменного языка собрать 10 000 слов, что соответствует размеру идиолекта взрослого носителя языка. Слов может быть больше или меньше, но ориентир задан и работает. Именно этот ориентир в 1000 слов (и содержащиеся в ней ядра лексики – от меньшего к большему) и следует положить в основу лексико-семантической типологии.

Заметим, что компаративисты судят о родстве языков по лексическому множеству на порядок меньшему – по 100-словному списку М. Сводеша, обоснованность которого, насколько нам известно, не ставилась под сомнение – быть может, напрасно. За полвека пользования списком Сводеша обнаружено немало недостатков. Есть серьёзные основания полагать, что список Сводеша составлен интуитивно, а не с опорой на сколь-либо верифицируемую процедуру. «Вспомним хотя бы словарь П.С. Палласа, изданный в двух отделениях (первое – в двух частях в Санкт-Петербурге в 1787 и 1789 гг.) при Екатерине II и по ее инициативе [Паллас, 1787 – 1789]. Первое отделение этого словаря содержало 285 словарных статей (понятий, «логических атомов»), отражавших (с большей или меньшей полнотой) материал 200 языков и наречий.

Это несколько больше, чем список базисных понятий М. Сводеша, который столь близок к словнику П.С. Палласа, что делает предположение о случайном совпадении маловероятным. Из 200 понятий первого списка М. Сводеша, к которому можно добавить семь понятий, которыми второй список (в 100 понятий) отличается от первого, общими со словарем П.С. Палласа оказываются 126, что составляет 61% ... (от 207)» (Список 126 совпадающих понятий см. в [Кретов 2006: 109-110]).

Второй ответ исходит из того, что в лексической типологии возможны лишь **локальные** лексико-семантические классификации языков (ср. работы [Толстой 1963], [Гудавичюс 1985], [Рахилина, Прокофьева 2004], [Майсак, Рахилина 2007],

---

<sup>5</sup> <https://www.native-english.ru/articles/words>

[Рахилина 2015] и др.). Правда, это путь к *микротипологии* (типологии фрагментов словаря, типологии лексико-семантических подсистем), а не путь к целостной лексико-семантической типологии языков (*макротипологии*): «Узкий сектор экстралингвистического мира... не стоит класть к основу выделения индивидуальных семантических типов. Таких классификаций можно было бы сделать много, но они не отражали бы существенных характеристик данного языка» [Городецкий 1969: 303].

При обилии учебников, называющихся «Лексикология английского (немецкого, русского, китайского и т.д.) языка», не существует описания лексико-семантической системы ни одного языка мира. А следовательно не может быть и лексико-семантической типологии языков мира.

#### **4. Идеографические словари – описание лексико-семантической системы?**

Некоторые полагают, что описаниями лексико-семантической системы являются идеографические словари. В этом есть доля правды: там, действительно, есть информация об эпидигматике (многозначности) и парадигматике (семантической близости) единиц языка, благодаря чему идеографические словари представляют, действительно, систему означаемых, объединённых тождеством означающих.

Но это не вся правда: идеографические словари не содержат информации о синтагматических связях слов и об их функциональной нагрузке. Главный же недостаток идеографических словарей, не позволяющий им претендовать на статус модели лексико-семантической системы, состоит в произвольности их синопсиса (синоптической схемы в терминах [Караулов 1976: 20, 35, 59], определяемого на высших уровнях исключительно произволом лексикографа, и только на низших – самим языком. Синопсис подобен сетке параллелей и меридианов, накидываемой на земной шар. С его помощью можно отождествлять те или иные фрагменты внеязыковой действительности, находящие отражение в разных языках, но он так же ничего не говорит о типе языка, как географические координаты не несут информации о том, что к ним привязано: гора, равнина или впадина; тундра, тайга, джунгли, пустыня или море. В идеографическом словаре нет информации о синтагматике и функциональной активности слов в их единстве с парадигматическими и эпидигматическими характеристиками. Поэтому идеографический словарь не является моделью лексико-семантической системы языка.

Итак, единственным реалистичным путём к лексико-семантической типологии является анализ словарей. Не утратило своей актуальности мнение Б. Ю. Городецкого: «Часто в семантических описаниях мы вынуждены пользоваться описаниями языков (словари [здесь и далее выделено нами – А.К., М.Г., В.Л.], очерки и т.п.), а не материалом самих языков (не продуктами речевой деятельности и не языковым знанием информантов). Специфика типологической работы с описаниями языков требует разработки особой методики инвентаризации. <...> Хотя типологическое использование существующих описаний языков всегда связано с возможностью ошибок, это не дает основания для отказа от таких работ. Имеющиеся (в том или ином зафиксированном виде) сведения о семантике языков мира

предоставляют нам ценнейший материал для извлечения знаний более высокого порядка, знаний о семантических типах. Умелое использование этого материала надо, видимо, предпочесть простому ожиданию «лучших времен», когда семантика всех языков будет описана единым и последовательным образом. Насущные потребности теории и практики, а также естественный профессиональный интерес лингвистов диктуют необходимость скорейшего максимального использования тех данных о семантике языков мира, которые имеются в нашем распоряжении» [Городецкий 1969: 316].

Сбылось и ещё одно предсказание Б.Ю. Городецкого: «При соответствующих условиях некоторые процедуры инвентаризации могут быть изложены в виде алгоритма, оперирующего с источниками инвентаризируемых данных (такими источниками могут быть двуязычные словари, толковые словари, заполненные анкеты, зафиксированные результаты экспериментов с информантами). Это создаст предпосылки для создания механизированных и автоматизированных систем инвентаризации семантических данных с использованием различного рода машин» [Городецкий 1969: 317].

Как справедливо полагал Б.Ю. Городецкий, «важную роль на всех этапах работы играет понятие метаязыка семантической типологии» [Городецкий 1969:326]. «Пока не ясно, на какой основе лучше всего строить метаязык, можно использовать в качестве такой основы естественные языки. Предполагается, что человеческие понятия могут быть описаны на любом языке. Поэтому при описании связей семантических единиц с понятиями можно использовать любой язык» [Городецкий 1969: 328].

К этому следует добавить, что в качестве метаязыка целесообразно использовать родной язык исследователей, что обеспечивает более качественную семантическую интерпретацию материала словарей-источников. В нашем случае это означает, что метаязыком описания должен быть русский язык, а источниками данных – иноязычно-русские словари.

Что касается «разработки особой методики инвентаризации», то ею является параметрический анализ лексики (ПАЛ), предложенный В.Т. Титовым [Титов 2002, 2004].

Весьма распространённое и приобретшее прочность догмы положение о единстве лексической и грамматической семантики должно быть оспорено и оставлено. Ему противопоставляется следующее положение: лексическая семантика относится к грамматической, как содержание относится к форме. Иными словами, лексическая семантика подобна жидкости, а грамматическая семантика подобна сосуду. Одно и то же лексическое (т.е. внеязыковое) содержание может быть грамматически оформлено в языке как признак – *белый*, как предмет – *белизна*, как процесс – *белеть* или как признак признака – *бело*.

Случаи представления одного и того же содержания средствами лексики или грамматики в разных языках мира не опровергают, а подтверждают это положение. Если лексика и грамматика взаимодействуют, то это является лучшим доказательством того, что они представляют собой разные сущности. Может ли одна *рука* взаимодействовать сама с собой? Нет. А левая рука с правой может. И то, что обе называются *руками*, несущественно. Точно так же экстралингвистическая лек-



сическая семантика Универсума взаимодействует с интралингвистической грамматической семантикой Языка. И то, что обе называются *семантиками*, ни в коей мере не свидетельствует об их тождественности.

В связи с этим при прочих равных предпочтительнее выбирать наименее маркированную языковую форму, какой является форма имени существительного: «Динамический подход к частям речи на основе теории нейтрализации позволил В.Г. Руделёву [Руделёв 1996] определить значение существительного отрицательно: <...> существительное – это то, что не обозначает действие, признак, количество, пространство и время. <...> Мы сочли необходимым определить положительно абстрактную лексическую семантику существительного посредством понятия **субстанции**. Понятие субстанции (материи) оказывается противопоставленным различного рода признаковым характеристикам субстанции (материи), а также формам и свойствам её существования, которые и перечислены В.Г. Руделёвым: действие (процесс), собственно признак, количество, время и пространство» [Шарандин 2001: 134].

### **5. Параметрический анализ лексики санскрита**

После этих предварительных замечаний перейдём к вопросу об изучении лексико-семантической системы санскрита с помощью параметрического анализа лексики.

Вопрос о параметрическом анализе лексики санскрита прежде не ставился, возможно, в следствии отсутствия (до начала XXI века) метода решения этой задачи. Таким методом является параметрический анализ лексики (ПАЛ).

В 2008 г. предварительные итоги ПАЛ славянских языков были доложены И.А. Меркуловой и А.А. Кретовым на XIV Международном съезде славистов в Охриде (Македония). Следующим этапом признания ПАЛа научным сообществом явилась публикация журналом «Вопросы языкознания» (№1, 2001) совместной статьи А.А. Кретова, И.А. Меркуловой и В.Т. Титова «Проблемы количественной лексикологии славянских языков». 15 февраля 2018 г. в диссертационном совете Д 002.006.01 при Институте языкознания РАН была защищена написанная под руководством А. А. Кретова канд. дисс. И. Дж. Семеновой «Параметрический анализ лексики карачаево-балкарского языка на тюркском фоне» (ведущая организация – Институт востоковедения РАН).

Методом параметрического анализа на сегодня исследована лексика 60 языков, представляющих следующие языковые семьи: индоевропейская 39, алтайская 10, уральская 4, кавказская 4, австронезийская 1, сино-тибетская 1, тай-кадайская 1.

Данные по санскриту (в том виде, в каком он нашёл отражение в словаре-источнике) вводятся в научный оборот впервые и носят предварительный характер, нуждаясь в серьёзном анализе и конструктивно-критическом отношении.

*Цель исследования – с помощью параметрического анализа санскритско-русского словаря выделить ядро санскритской лексики.*

*Объектом исследования является «Санскритско-русский словарь» [Кочергина 1987] (далее – Источник).*

Выбор санскритско-русского словаря двуязычного словаря обусловлен потребностью в едином метаязыковом основании для сравнения лексики языков, без которого лексико-семантическая типология невозможна, а также тем, что мета-

язык должен быть родным для исследователя, что обеспечивает более качественный семантический анализ лексики.

На лексику словаря наложены следующие ограничения: из рассмотрения исключены периферийные слова: онимы, местоимения, наречия, числительные, союзы, частицы, предлоги, междометия и звукоподражания. В результате реальный объект исследования составил менее 24 000 слов (23 576) вместо изначальных 29 007 в словнике словаря.

*Метод исследования* – Параметрический анализ лексики, состоящий во «взвешивании» каждого слова по четырем системообразующим параметрам на фоне всего словаря посредством вычисления для него четырех частно-параметрических весов: *эпидигматического (деривационного), парадигматического, синтагматического и функционального*. По каждому из параметров выделяется не менее 1000 слов с наибольшим частно-параметрическим весом. Сложение частно-параметрических весов дает Интегральный параметрический вес каждого слова, по которому и осуществляется системная стратификация лексики Источника.

Каждый частно-параметрический вес варьирует в интервале от 0 до 1 и вычисляется по формуле:

$$Pr_i = \frac{\Sigma r - R_{1-i}}{\Sigma r}$$

где

$\Sigma r$  – сумма слов всех рангов,

$R_{1-i}$  – сумма слов от первого ранга до данного,

$Pr_i$  – вес слов данного ранга [Титов 2004:15].

Логика формулы – спортивная: чем меньше участников показало такой же или лучший результат, тем выше место (больше вес) данного участника. Каждое слово в словаре влияет на вес остальных слов по каждому из четырех параметров, что резко сужает исследовательский произвол и повышает объективность исследования.

Принимается кумулятивная модель, в соответствии с которой словарь организован по принципу «ядро в ядре». Словарь делится на ядро и периферию. В **ядре словаря** входят слова, вошедшие в первую тысячу наиболее «весомых» слов хотя бы по одному из четырех параметров. Таких слов в Источнике – 27% (периферия составила 73% лексики словаря). В ядре словаря находится **большое ядро**, включающее слова, вошедшие в первую тысячу по двум параметрам. В большое ядро входит **среднее ядро** – леммы, вошедшие в ядро словаря по трём параметрам. Малое ядро составляют леммы, вошедшие в ядро словаря по всем четырём параметрам. В малом ядре представлена максимальная концентрация системных свойств словаря, поэтому системный анализ лексики целесообразно начинать с **малого ядра**.

Рассмотрим *распределение лемм по функциональному*. Самые короткие слова по длине приближаются к длине морф, а самые длинные слова – к длине словосочетаний. Собственно слова находятся в интервале между морфемами и словосочетаниями. Самым коротким полнзначным словом санскрита является лемма *i* ‘идти; отправляться’ и *e* ‘приходить из, от’. В качестве функционального ядра лексики взято 2755 лемм длиной в четыре фонемы и менее.

Стратификация лемм *по деривационному весу*. В качестве ядра выбрано 1245 лемм с пятью и более дефинициями. Самыми многозначными словами являются три слова: *da* ‘давать, дарить’, *sarj* ‘освобождать, отпускать, позволять’, *sadhana* ‘приводящий к, создающий, направляющий’.

*Парадигматическое* ядро Источника составляет 1192 синонимических ряда, состоящие из шести и более лемм. *Парадигматической доминантой* является значение ‘солнце’, образующее синонимический ряд из 61 леммы, *парадигматической вице-доминантой* является значение ‘сильный’, образующее ряд из 42 синонимов. Значение ‘идти’ образует ряд из 40 синонимов, ‘огонь’ – из 36 синонимов, ‘смерть’ – из 35 синонимов, ‘сила, мощь’ – из 34. Количество синонимов в Источнике является показателем (и мерой) существенности значения для хранителей древнеиндийских традиций.

*Синтагматическая стратификация лексики санскрита*. К сожалению, информация о сочетаемости лемм в Источнике почти отсутствует. Это заставило нас обратиться к корпусу санскритолога Дюссельдорфского университета им. Г. Гейне Оливера Хельвига. В нем со словарем Источника совпало 13 682 лемм, для которых была с помощью программных средств определена сочетаемость в крупно-корпусной базе. Одна из причин неполного совпадения списков – различное написание лемм, другая – низкая совместимость Источника и базы вследствие их ориентации на разные словари.

В синтагматическое ядро Источника включены 1104 леммы, сочетающиеся с 200 и более разными леммами. *Синтагматической доминантой* оказалась лемма глагольного корня *as* ‘быть’, с которой сочетаются 5894 основы, *вице-доминантой* *adi* ‘начало’, на третьем месте *vas* ‘говорить’.

*Интегральная стратификация лексики Источника*. В малое ядро вошло 96 слов, в среднее – 548, в большое – 2416 и в ядро словаря – 8825. *Интегральная доминанта* санскрита по данным Источника это слово *go* ‘бык’ – у него максимальный параметрический вес – почти 3,977, *вице-доминанта* *dhi* ‘мысль’, на третьем месте *gam* ‘идти’.

### Итоги и перспективы

Мы нашли самое короткое слово Источника – *i* ‘идти’, самые многозначные – *da*; *sarj*; *sadhana*, самое богатое сочетаемостью – *as* ‘быть’. По данным Источника, самое главное слово санскрита – *go* ‘бык’ и вице-доминанта – *dhi* ‘мысль’. Самой омонимичной лексемой Источника является *vas* – пять омонимов.

*Перспективы исследования* – проверка и уточнение полученных данных, их оценка на фоне других языков, выявление национально-специфической лексики санскрита, определение индекса семантической удаленности санскрита от русского языка.

### Литература

1. Апресян Ю.Д. Интегральное описание языка и системная лексикография // *Избранные труды*. Т. II. М.: Языки русской культуры, 1995.
2. Городецкий Б.Ю. *К проблеме семантической типологии*. М., 1969.

3. Гудавичюс А. *Сопоставительная семасиология литовского и русского языков*. Вильнюс, 1985.
4. Караулов Ю.Н. *Русская и общая идеография*. М.: Наука, 1976.
5. Кочергина В.А. *Санскритско-русский словарь. около 30.000 слов* / Под ред. В.И. Кальянова. С приложением «Грамматического очерка санскрита» А.А. Зализняка. 2-е изд. испр. и доп. М. Русский язык, 1987 (1978). 944 с.
6. Кретов А.А. В.Т. Титов и параметрический анализ лексики // *Вестник ВГУ. Серия Лингвистика и межкультурная коммуникация: Научный журнал*. 2017, № 4. С.5–9.
7. Кретов А.А. *Единство Европы по данным лексики: коллективная монография* / А.А. Кретов, О.М. Воеводская, И.А. Меркулова, В.Т. Титов. Воронеж: Издательский дом ВГУ, 2016. 412 с.
8. Кретов А.А., Меркулова И.А. Динамика и устойчивость в лексической системе славянских языков // *Динамика та стабільність лексичних і словотвірних систем слов'янських мов – тематичний блок. XIV Міжнародний з'їзд славістів 10.09.–16.09.2008*, Охрид, Республіка Македонія. 2008.
9. Кретов А.А., Титов В.Т. Алгоритм позиционного выявления синонимии // *Вестник ВГУ. Серия «Системный анализ и информационные технологии»*. 2006. № 1. С. 62–65
10. Майсак Т.А., Рахилина Е.В. (ред.). *Глаголы движения в воде: лексическая типология*. М., 2007.
11. Мечковская Н.Б. *Общее языкознание. Структурная и социальная типология языков: Учеб. пособие*. Минск, 2000.
12. Морковкин В.В. *Лексические минимумы современного русского языка* / В.В. Морковкин, Ю.А. Сафьян, Е.М. Степанова, И.В. Дорофеева; Под ред. В.В. Морковиной. М.: Русский язык, 1985. 608 с.
13. Рахилина Е.В., Прокофьева И.А. Родственные языки как объект лексической типологии: русский и польские глаголы вращения // *Вопросы языкознания*, 2004. № 1. С. 60–78.
14. Рахилина Е.В. (составитель). *Глаголы звуков животных: типология метафор*. М.: Языки славянской культуры. 2015. 400 с.
15. Руделёв В.Г. Динамическая теория частей речи русского языка // *Вестник Тамбовского университета*. Серия: Гуманитарные науки, 1996. № 1. С. 83–89.
16. Старостин С.А. Сравнительно-историческое языкознание и лексикостатистика // *С.А. Старостин. Труды по языкознанию*. М.: Языки славянских культур. 2007. С. 407–447.
17. Титов В.Т. *Общая количественная лексикология романских языков*. Воронеж, 2002.
18. Титов В.Т. *Частная количественная лексикология романских языков*. Воронеж, 2004.
19. Толстой Н.И. Из опытов типологического исследования славянского словарного состава I // *Вопросы языкознания*. 1963. № 1. С. 29–45. 1966. № 5. С. 16–37.

20. Фриш О. *Санскритская хрестоматия*. В 2 т. Т. 2. Словарь / сост. О. Фриш; редактор В. Поржизка; автор предисл. и биогр. М. Ю. Гасунс. М.: «АБВ», 2015. (Репр. воспроизведение изд. / сост. О. Фриш. Прага: Изд. АН Чехословакии, 1956). 355 с.
21. Шарандин А.Л. *Курс лекций по лексической грамматике русского языка: Морфология*. Тамбов: Изд-во Тамбовского ун-та. 2001. 313 с.
22. Bergsland K., Vogt H. On the validity of glottochronology // *Current Anthropology*. 1962. № 3. P. 115-153
23. Zipf G. K. *Human Behavior and the Principle of Least Effort*. Cambridge, Mass.: Addison-Wesley Press, 1949.

**Моделирование семантического перехода:  
обзор направлений исследования  
(на материале словарной статьи *vīzu* ‘падать’ в словаре Tamil Lexicon)**

**Анна Александровна Смирнитская**

Институт востоковедения РАН, Москва, Россия,

nyushas@gmail.com, ORCID 0000-0001-7814-8716

**Аннотация.** Рассматриваются современные направления исследования семантических переходов: работы по колексификации и построение семантических карт, теория семантических переходов и этнолингвистическое направление, фреймовый подход и исследования по дистрибутивной семантике. Проводится анализ различных подходов на примере словарной статьи тамильской лексемы *vīzu* ‘падать’. Показано, что представление словарной статьи полисемичной лексемы как совокупности случаев семантических переходов и их дальнейшая визуализация в виде графа может дать новую информацию для осмысления явлений полисемии и семантической деривации.

**Ключевые слова:** полисемия, семантический переход, глаголы падения, тамильский язык, семантическая деривация, семантика, лексикография, семантическая типология

**Abstract.** The article discusses the current trends in the study of semantic derivation (semantic shifts): papers on colexification and construction of semantic maps, the theory of semantic shifts and ethnolinguistic direction of research, frame-based approach in lexical typology and papers on distributive semantics. Various research methods are analyzed using the example of the Tamil polysemous lexeme *vīzu* ‘to fall’. It is shown that the presentation of a dictionary entry of a polysemous word as a set of cases of semantic shifts and further, their visualization as a graph, can provide new information for understanding the phenomena of polysemy and semantic derivation.

**Keywords:** polysemy, semantic shift, verbs of falling, Tamil, semantic derivation, semantics, lexicography, semantic typology

## 1. Введение

Понятие «семантический переход» стало в последние годы одним из центральных для семантической типологии (см. [Vanhove 2008], [Juvonen, Koptjevskaja-Tamm 2016], [Gast, Koptjevskaja-Tamm 2021] и др.). Настоящая работа придерживается теоретического осмысления этого понятия, предложенного в работе Анны А. Зализняк [Зализняк 2001]. Вслед за автором мы понимаем под «семантическим переходом» наличие «концептуальной смежности между двумя языковыми значениями, проявляющейся в совмещении этих значений в пределах одного “слово в широком смысле»» [Зализняк 2013, с. 37]. Такое «слово в широком смысле» определяется при этом в двух временных планах: как синхронное единство слова с его полисемией и как единство слова, взятого в историческом развитии. Семантичес-

кий переход обнаруживает себя в форме «реализаций, различаемых в зависимости от того, какие языковые единицы являются носителями сопоставляемых значений: синхронная полисемия, диахроническая семантическая эволюция, морфологическая деривация, когнаты и заимствование» [там же].

В настоящее время группой исследователей под руководством Анны А. Зализняк в Институте языкознания (Москва) разрабатывается Каталог семантических переходов в языках мира; создана и находится в процессе пополнения база данных Каталога — DatSemshift 2.0 (<https://datsemshift.ru>). Каждый семантический переход в Каталоге представлен в виде пары значений: значения-источника и значения-цели, соединенных вектором семантического развития, и снабженных указанием некоторых дополнительных параметров:

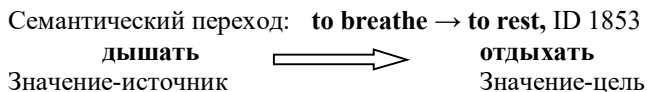


Схема 1. Структура описания семантического перехода в Каталоге Datsemshift. (Для семантического перехода и составляющих его значений указываются дополнительные параметры: таксономическая категория значений, тип реализации, связанные переходы и др., приводятся примеры употреблений).

В то же время явления, связанные с этим понятием, находятся в фокусе нескольких других лингвистических дисциплин, а именно, исторической лексикологии, компаративистики, этнолингвистики, дистрибутивной семантики. Каждое из этих теоретических направлений имеет свои цели и методы и свой специфический «интерес», предмет исследования, в рамках явления семантического перехода; соответственно с этим, разные подходы к проблеме используют различные теоретические модели для ее описания.

Настоящая работа делится на две части. В первой части проводится обзор современных подходов к моделированию лексико-семантических сдвигов с примерами некоторых конкретных исследований – как чисто аналитических, так и ориентированных на компьютерную обработку. Вторая часть посвящена рассмотрению применения этих подходов к анализу одной словарной статьи – а именно, словарной статьи полисемичной тамильской лексемы *vizu* ‘падать’ из словаря Tamil Lexicon.

## 2. Моделирование лексико-семантических сдвигов

В настоящее время существует несколько теоретических подходов, развивающих схожие представления. Каждое из них из них разрабатывает собственную методику исследования, некоторые используют современные компьютерные методы, составляют базы данных случаев изменения лексических значений. Поскольку фокус исследовательского интереса разный для разных направлений, понятия, используемые для изучения этого явления, также различаются. Приведем список направлений (возможно, неполный) с примерами.

## 2.1. Аналитические подходы к моделированию лексико-семантических сдвигов

(1) Вопрос о том, как и почему происходит семантическая эволюция, исследуется как в рамках одного языка, так и типологически. Теоретическому осмыслению проблем семантической деривации посвящены, в частности, работы П. Коха (Peter Koch). В статьях [Koch 2001, 2016] ставятся вопросы о методах исследования путей семантической эволюции в рамках традиционной парадигмы (начиная с античной риторики и кончая структурализмом) и в свете достижений современной когнитивной семантики и прагматики, выделяются типы семантических изменений: генерализация, специализация, когипонимический переход, метафора, метонимия, пейоративный и мелиоративный сдвиг и др.; предлагаются уточнения описаний метонимического и метафорического переноса в современных когнитивных теориях. Изучению конкретных случаев лексико-семантических сдвигов посвящаются многочисленные исследования в рамках лексической типологии, как например, статья А. Шеппер, Л. Сан Рок и Р. Хендери из сборника [Juvonen, Kortjevskaja-Tamm 2016], в которой описываются особенности коллексификации понятий «дерево», «дрова» и «огонь» в языках Австралии и Папуа-Новой Гвинеи.

(2) Изучение семантического перехода может использоваться как средство понять, что же происходило в лексической системе языка на ранних, труднодоступных для исследования этапах ее развития. К этой группе работ относятся диахронические исследования – в рамках лексикологии, этнолингвистики или семантической типологии. Так, исследователь М. Фортескью разрабатывает семантическое поле эмоций в эскимосском языке и языках проживающих рядом этносов; рассматривает лексическое выражение эмоций в языках эскимосского ареала на территории США, Канады и России; с помощью диахронического анализа когнатов реконструирует картину древней категоризации внутреннего эмоционального опыта, см. [Fortescue 2016]. В работах Т. А. Михайловой метод изучения семантических переходов используется при анализе исторических памятников, в том числе на древнеирландском, галльском и других кельтских языках и эволюции связанных с ними представлений, в частности, о цветообозначениях, о понятиях «путь» и «дорога», «муж», «воровство», «солнце», о теонимах и др., см. [Михайлова 2013, 2019 и многие другие].

В работах этнолингвистической школы Н.И. и С.М. Толстых исходной точкой исследования становится синхронное состояние языка, т.е. значение-цель, но само исследование направлено в прошлое: ставится вопрос о том, какие исходные значения привели к данному в процессе исторического развития. По этой причине основной реализацией семантического перехода здесь становится диахроническое развитие, дополненное морфологической деривацией, см. такие работы, как [Толстая 2002, 2009].

(3) Главным предметом внимания исследователя может быть именно семантический переход как особое явление. В рамках изучения семантических переходов в понимании Анны А. Зализняк проводятся конкретные исследования, такие как работы М. М. Руссо, М. В. Орловой, И.А. Грунтова и др., например, [Грунтов 2007], [Руссо 2019]. В частности, в статье М.М. Руссо «О направлении семантических переходов» [Руссо 2019] показано, что направление перехода устанавлива-



ется в каждом случае на основании направлений в языковых реализациях. Для определения направления перехода используются четыре критерия: синхронный, исторический, сравнительно-исторический и экстралингвистический. В частности, синхронный критерий в переходе *to break* (tr.) → *to confess*, ID 0180 (ср. рус. *расколоть орех* и *расколоться, что в подвале спрятан самогонный аппарат*) и сравнительно-исторический при диахроническом развитии, как в переходе *tomorrow* → *tomorrow*, ID 0068 (фр. *demain* из лат. *mane*).

(4) Исследование семантических сдвигов может быть ограничено одним или несколькими видами их проявления. К этому разделу можно отнести работы по типологии колексификации – одного из подходов к явлению полисемии; в таких проектах не учитываются другие манифестации семантических переходов: морфологическая деривация, заимствования, когнаты, семантическая эволюция. Авторы используют теоретический аппарат, основанный на классической работе [François 2008] и развиваемый в современных статьях, как [François 2021]. Здесь можно рассматривать работы с участием Александра Франсуа (Alexandre François), как например COLEX (начат в 2018) и работы других авторов, его подходом вдохновлены такие проекты, как CLICS (см. далее).

Рассматриваются случаи семантического сближения в синхронной перспективе, исследуется полисемия как основной вариант проявления этого сближения. Вводятся понятия «колексификация» [François 2008], «колекс», «дизлексификация», «дизлекс», изучаются паттерны колексификации. Авторы предлагают «панхроническую» перспективу, подчеркивающую внутреннюю связь между «*synchronic variation*» и «*long-term historical change*». База данных COLEX разрабатывается в виде таблицы, в левой части которой описывается паттерн колексификации (значение-источник, значение-цель и некоторые параметры), и в правой части приводится список лексических единиц разных языков, которые этот паттерн реализуют. В отличие от базы данных *DatSemShift*, здесь не учитывается морфологическая деривация, когнаты и заимствования. Одно из достоинств подхода – возможность проследивать сходство паттернов колексификации, например, в креольских языках Океании по сравнению с субстратом – меланезийскими языками и в противовес английскому.

(5) Изучение семантической деривации может быть одним из аспектов исследования лингвоспецифичности в синхронных направлениях исследований; фокус изучения при этом скорее направлен на особенности перехода и его цель. К этому разделу можно отнести различные работы по лексической семантике и лексической типологии, в том числе работы московской семантической школы под рук. Ю. Д. Апресяна; фреймовый подход Московской лексико-типологической группы; близкие к данному направлению исследования языковой картины мира.

Так, фреймовый подход разрабатывается авторами Московской лексико-типологической группы (MLexT) под руководством Е.В. Рахилиной. Восходя к традициям Московской семантической школы (см., в частности, [Апресян и др. 2006]), подход предполагает «сравнительный анализ внутри- и межъязыковых квазисинонимов через призму их сочетаемостных свойств» [Рыжова 2018: 21]. Типы ситуаций, т.е. непересекающиеся понятийные фрагменты, на которые можно разбить семантику слова для анализа, или «...ситуации, которые чаще всего противополо-

ставляются лексическими средствами в языках мира» [Рахилина, Резникова 2013: 12], в этом подходе называют фреймами. Особенностью фреймов является то, что они типологически ориентированы, их набор составляется по результатам подробного анализа материала нескольких языков.

Основным предметом исследования в направлениях этого типа является синхронное состояние языка, привлекаются диахронические сведения по мере необходимости; в частности, при рассмотрении метафорических значений.

## 2.2. Подходы, ориентированные на компьютерную обработку

С развитием технологий в последние годы стали возможны новые методы изучения семантики, в том числе изучения сближений и изменений значения слова. Эти методы касаются как представления самих значений, так и представления связей между ними, а также источников данных. Лексико-типологические исследования на материале языковых корпусов дают существенно больше информации по сравнению с традиционными подходами.

(1) Если значения представлены в виде абстрактных метаязыковых «ярлыков», как в теории семантических переходов, новые компьютерные модели могут использоваться для представления связей между ними. Связи визуализируются в виде сетей, графов; широко применяются семантические карты значений. Эти методы сближают лексическую типологию с методами грамматической семантики, грамматической типологии. Так, бельгийский исследовательский проект LeDiaSEMa (2016–2018), авторами которого являются T. Georgakopoulos и S. Polis, изучает возможность создания семантических карт, иллюстрирующих семантические изменения значения в диахронии. Одной из целей проекта является создание автоматизированной системы создания диахронических карт. Для их построения проект использует готовые данные по изменению значений в процессе семантической эволюции. Данные при этом могут быть получены из разных источников – как из собственной работы исследователей, так и из других баз данных, в которых сопоставлен материал, относящийся к разным моментам развития языка. Фактором, ограничивающим возможности проекта, является слабая разработанность области и недостаток достоверных данных о семантической эволюции [Georgakopoulos, Polis 2018].

Связи в таком графе могут быть направленными и ненаправленными. При построении карт может указываться «вес» узлов (plotting weighted maps): в лексической матрице проверяется, сколько раз встретился тот или иной паттерн полисемии, и соответственно увеличивается «вес» узла. Составляются как традиционные семантические карты «первого поколения», показывающие связь значений (link-based), так и карты «второго поколения», учитывающие их близость или удаленность друг от друга (proximity-based), что может показывать, какие значения в языках чаще реализуются одной конструкцией [там же].

(2) Ограничение данных семантического развития данными колексификации, извлекаемыми из словарей, расширяет возможности типологического исследования явления по разным языкам. Автоматизацию извлечения значений из существующих словарей и лингвистических баз данных использует группа, разрабатывающая базу данных колексификаций в языках мира CLICS (Database of Cross-Lin-

guistic Colexifications) – агрегатор данных о колексификациях в языках мира<sup>1</sup>. Над проектом работает международная группа исследователей, первоначально только из немецких университетов Max Planck Institute for the Science of Human History (Jena), Philipps-University (Marburg) и Heinrich Heine University (Düsseldorf), позже она стала интернациональной. База данных развивает теоретические положения Александра Франсуа (см. раздел 2.1. выше) и извлекает из существующих словарей случаи «колексификации», то есть полисемии или объединения значений в рамках одной лексической единицы.

В системе отчасти преодолена проблема сопоставимости словарных данных, поскольку базовым источником ее являются данные серии «Межконтинентальных словарей» (Intercontinental Dictionary Series). При этом авторы проекта сознательно не проводят различие полисемии и омонимии, почти невозможное при автоматизированном подходе; это приводит к некоторому количеству неясных данных на данном этапе работы – «шума», затрудняющего дальнейшее исследование. Кроме того, не рассматриваются, в отличие от DatSemshift, случаи семантической эволюции, когнаты, морфологическая деривация, заимствования. Другой особенностью этого ресурса является «далекий от идеального список концептов, лежащий в его основе. В большинстве случаев он достаточно хорош для предметной лексики, но глагольные и признаковые зоны в нем представлены довольно бедно» (см. подробное обсуждение этой системы в [Рыжова 2018: 16]).

CLICS является агрегатором информации из открытых лингвистических корпусов и баз. На настоящий момент эта база данных – наибольшая по охвату языков и лексических единиц среди источников такого рода. С 2014 года, когда появился первый вариант базы данных, было выпущено три версии. В последней версии CLICS<sup>3</sup> охвачено около 3156 языков мира и 2906 концептов. Для каждого значения в базе устанавливаются колексификационные связи с другими значениями в языках, на основании этих связей строится сеть с центральным концептом в вершине. Интернет-функционал снабжен визуальным представлением, демонстрирующим сети взаимосвязей для каждого значения [Rzymiski et al 2020, CLICS<sup>3</sup> 2021].

(3) Компьютерные методы позволяют использовать большие данные для анализа языковых явлений (большие языковые корпуса); однако этими данными можно воспользоваться, только отказавшись от традиционного лингвистического анализа, поскольку экспертная обработка к таким объемам неприменима. Такого рода работы относят к направлению дистрибутивной семантики и векторного анализа, используется обучение с помощью нейросетей, применяется квантитативный подход. Нейросети используются как для представления значения, так и для анализа его связи с другими значениями и его изменений во времени. Само лексическое значение в этом случае представляется как результат соотносительности со своим окружением, как вектор своих связей. Нейросеть обучается на большом объеме данных при помощи специальных программ, таких как Word2Vec, Bert. Таким способом можно получать информацию как о связях концепта с другими в рамках исследуемого корпуса, так и информацию об историческом изменении значения, ср. [Dubossarsky et al 2016], [Hamilton et al 2016], [Tahmasebi et al 2021].

---

<sup>1</sup> На данный момент актуальна версия CLICS<sup>3</sup>: <https://clics.clld.org/>

(4) Создаются новые комплексные подходы, включающие различные применения автоматизации для работы с семантикой лексической единицы. Так, разработке комплексного метода автоматизации в рамках фреймового подхода посвящена диссертация Д. А. Рыжовой [Рыжова 2018]. Автор использует модели дистрибутивной семантики и кластерного анализа векторной сочетаемости для составления анкеты заданного семантического поля. Затем анкета переводится на другие языки с помощью параллельных корпусов, машиночитаемых словарей и онлайн переводчиков, и заполняется с помощью одноязычных корпусов в выбранном языке; для визуализации полученных результатов используются решетки формальных понятий. В исследовании показана работа предложенных алгоритмов на материале семантических полей 'острый', 'толстый', 'гладкий'. Анализ понятий, разрабатываемый в диссертации, предназначен для автоматизации описания семантического поля без разделения прямых и переносных значений [Рыжова 2018: 35], и тем самым, может быть использован также в описании случаев семантического сдвига.

### 2.3. Ограничения методов автоматизации

Однако, в наличии аннотированных корпусов или просто цифровых данных на конкретном языке кроется и ограничение подобных методов. Поскольку только небольшой процент языков мира может гордиться наличием данных в форме, пригодной для автоматической обработки, только эти языки могут быть таким способом исследованы.

Среди ограничений всей этой группы методов надо указать три. Во-первых, автоматическая кросс-идентификация лексических значений по формулировкам в толковых словарях с неизбежностью включает в выдачу некоторые неверные данные («шум») и, наоборот, упускает нужные. Те же неточности в данных возникают в результате работы векторного анализа больших корпусов. Во-вторых, автоматизированные методы ограничивают классы явлений, поскольку не все разновидности материала можно обработать таким автоматическим способом. Разработка семантических связей, реализованных когнатами, заимствованиями, морфологической деривацией, семантической эволюцией, даже близкими к рассматриваемым примерам случаями грамматикализации – чрезвычайно затруднена и во многих случаях невозможна. В-третьих, компьютерные методы имеют существенные ограничения в плане языков – для анализа доступна только их малая часть, поскольку отнюдь не для всех языков мира имеются представительные лексические корпуса. Еще серьезнее данные ограничивают метод нейросетей, потому что для обучения нейросетей требуются корпуса большего размера, а такие корпуса созданы на сегодняшний момент только для самых изученных языков. И самые большие ограничения накладываются на использование нейросетей в диахронической семантике, потому что здесь требуется наличие корпусов с описанием разных исторических состояний языка. Но, несмотря на все эти ограничения, очевидно, что будущее за подобными видами автоматизированного или полу-автоматизированного анализа, по крайней мере, в качестве дополнения к традиционным методам работы лингвиста.

### 3. Словарная статья полисемичной лексики как поле изучения лексико-семантических сдвигов

Интересно провести анализ того, как именно можно использовать разные теоретические подходы, разные способы моделирования лексико-семантических сдвигов при изучении одного объекта. В качестве объекта в настоящем исследовании выбирается словарная статья слова с высокой полисемией. В такой статье с большой долей вероятности будут собраны значения, представляющие разные случаи семантического «расширения» – как случаи реального диахронического развития, которые лексикограф мог бы записать с пометой «устаревшее» или, наоборот, «сленг», так и случаи когнитивного сближения значений и подзначений, отражающие разные направления развития, разные возможные грани семантики значения, взятого исходным.

В данной работе мы выбираем словарную статью полисемичной тамильской лексики *viṣu* с основным значением ‘падать’. Для анализа выбран словарь *Tamil Lexicon* [1924-1936], как наиболее полный и разнообразный в интересующем нас плане. Этот словарь сам по себе чрезвычайно интересен как источник данных. Дело в том, что его составители ставили своей целью собрать наиболее представительный словарь, используя данные более ранних лексикографов и учитывая максимально полно, насколько это было возможно, употребления слов, начиная с древней тамильской литературы «эпохи санги» (ранее 3 в н.э.) на старотамильском языке, и далее из литературы на средневековом и современном языке (см. периодизацию в [БРЭ 2021]). Таким образом, в словаре представлены и современные значения, и значения, давно вышедшие из употребления; в одной словарной статье можно найти целую историю лексического поля. Именно такая словарная статья может дать много информации как объект изучения семантических сдвигов разными методами.

Конечно, никакой словарь не может быть свободен от установок и теоретических представлений, которыми был связан его составитель. Но в данном случае установка составителей *Tamil Lexicon* на максимальную подробность для наших целей, наоборот, полезна. Однако также необходимо понимать, что с этой установкой может быть связано более дробное выделение значений, признание отдельным значением случая, который другой лексикограф отнес бы к контекстно зависимому употреблению лексики.

#### 3.1. Словарная статья глагола *viṣu* ‘падать’ в словаре *Tamil Lexicon*

Приведем словарную статью глагола *viṣu*<sup>2</sup> ‘падать’ в том виде, в котором она представлена в словаре *Tamil Lexicon*<sup>3</sup>:

**viṣu-**, *v. intr.* To fall, fall down; to descend; to flow down; 2. To lie prostrate, as in reverence; 3. To fall down exhausted; 4. To fall sick or ill; to be bed-ridden; To be defeated or overthrown; 6. To fall low, decline, as prices; to sink to a lower level; 7. To

<sup>2</sup> Символом *z* принято обозначать ретрофлексный аппроксимант, другой способ его записи *ḷ*, в МФА он указывается как *ɻ*: [ɻɐm̩ɨ] *tamil* ‘тамилский язык’.

<sup>3</sup> В передаче словарной статьи сокращена грамматическая информация, записи тамильской графикой, комментарии к значениям и примеры.

be destroyed, ruined; 8. To die; 9. To set, as the sun; 10. To fall into desuetude; 11. To settle; to be fixed; 12. To crowd together; to swarm; 13. To be allotted; to fall to one's lot or share; 14. To be imprinted; 15. To appear; to be formed; to come out; 16. To take place, happen; 17. To pass; to pass away; 18. To rush, as against a person; 19. To exhibit desire; 20. To fall into position; 21. To devolve; 22. To be born, used in contempt; 23. To issue, proceed from, as speech; 24. To become detached; 25. To swoop; 26. To discharge, empty, as a river; 27. To hang down, droop; 28. To fall into a rage; 29. To lose animation, as face; 30. To be locked up, as capital.

Как можно увидеть, после грамматических помет перечисляются многочисленные значения этой полисемичной лексемы, в том числе значения, которые можно отнести к «прямым», как *fall, fall down*, затем многочисленные значения, которые можно назвать «переносными», но которые выглядят привычными для русскоязычного читателя, как *to decline, as prices* (ср. рус. *цена упала*); *to rush, as against a person* (ср. рус. *он напал на человека*), *to die* (ср. рус. *герои пали в бою*), а также значения, семантическая мотивация которых с первого взгляда не так понятна, как *to set, as a sun; to crowd together; to take place, happen; to discharge, as a river* и другие.

Для целей данного исследования эта словарная статья будет представлена в виде графа семантических связей (семантических переходов), в котором значением-источником мы будем считать «заглавное» значение этой лексемы ‘падать’, а значением-целью – все остальные значения, которые эта лексема описывает. Действительно, полисемия, как известно, является одной из неизбежных стадий семантического перехода, и такое представление согласуется с концептуальным аппаратом Каталога семантических переходов *Datsemshift 2.0*, в рамках которого работает автор (см. [Зализняк 2001, 2013, Zalizniak 2018], Смирнитская 2016а). Каждая пара значений будет представлять один семантический переход, один случай когнитивной смежности с исходным значением ‘падать’. Структуру полисемичной словарной статьи *vizu* ‘падать’, таким образом, можно представить в виде графа, см. схему 2 (граф включает основные значения, оставшиеся рассматриваются в тексте далее).

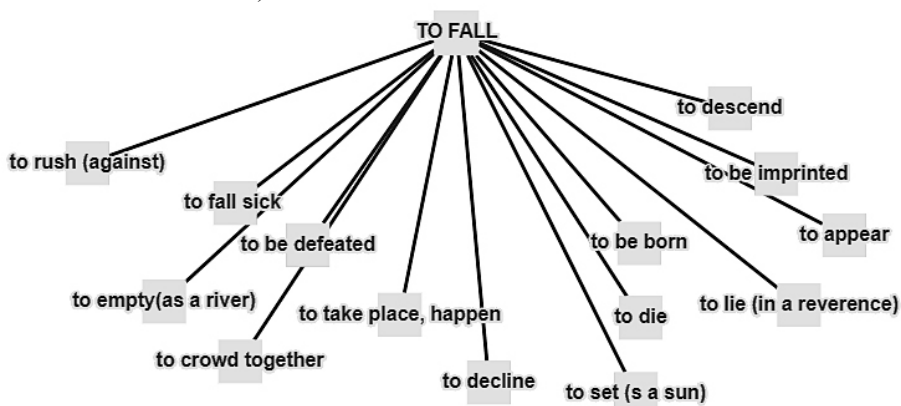


Схема 2. Графическое представление структуры значения словарной статьи полисемичной лексемы *vizu* ‘падать’.

При этом графическое представление приводится на основании той информации, которая есть в словаре. В частности, значения, указываемые в позициях «узлов» графа, мы можем указать только в таком виде, в каком семантическое пространство, описываемое этой лексемой, было разделено на значения составившими словарь лексикографами. Номенклатура «ярлыков» значений, их описание – все это мы в данной работе не изменяем, поскольку следуем данным словаря. Соответственно, некоторая дополнительная работа может потребоваться, если выделенные ярлыки между собой пересекаются или неочевидны.

#### **4. Анализ полисемичной словарной статьи *viṣi* ‘падать’ в разных теоретических направлениях**

Как в рассмотренных теоретических направлениях можно провести анализ интересующей нас словарной статьи? Не все направления разрабатывают проблему полисемии. Если бы мы исследовали полисемичную лексему в рамках этнолингвистического подхода школы Н.И. и С. М. Толстых, теоретическая постановка задачи была бы связана с определением заключенных в форме мотивационных моделей. Если бы мы рассматривали ее в рамках проекта LeDiaSema, необходимо было бы для составления семантической карты диахронического развития сопоставить исторически более раннее состояние языка, т.е. включить в исследование данные исторического корпуса (или корпусов) раннего тамильского языка, что в данный момент неосуществимо, поскольку такие корпуса еще не созданы. Для использования компьютерно-ориентированных методов был бы необходим доступ к большим тамилоязычным корпусам, что также сейчас не представляется возможным.

Для целей исследования полисемии в данной статье используется подход теории колексификации (François, CLICS), а также метод Каталога семантических переходов Datsemshif и фреймовый подход Московской лексико-типологической группы.

##### **4.1. Анализ в системе CLICS**

Как мы можем проанализировать эту словарную статью с помощью базы данных колексификаций в языках мира CLICS3? Зададим поиск по значению FALL ‘падать’ на странице ресурса: <https://clics.clld.org/>. База данных показывает нам граф, в котором встречаются связанные значения, часть из тех, что перечислены в словарной статье «падать». Для дравидийских языков при этом зарегистрирована только одна семантическая связь FALL – TUMBLE, других данных по дравидийским языкам на настоящий момент в базе данных CLICS нет.

Однако типологически зарегистрированные в других языках семантические соотношения действительно встречаются также и в данной статье из тамильского словаря. Это связи FALL – DESCEND, FALL – TUMBLE, FALL – BE BORN, всего три примера. Про некоторые другие случаи колексификации можно предположить опосредованную связь, так, колексификация FALL – KILL, возможно, может быть связана с семантическим переходом to fall – to die, имеющемся также и в структуре полисемичной словарной статьи *viṣi* ‘падать’. И остается достаточно большое количество семантических связей, для которых соответствий в сло-

варной статье *vizu* нет. Перечислим эти связи: FALL – EAT, FALL – GO, FALL – GO DOWN, FALL – CUT, FALL – CUT DOWN, FALL – SINK, FALL – DROP (SOMETHING), FALL – DRIP (EMIT LIQUID), FALL – SLIP, FALL – CLIMB, FALL – LAND. Вероятно, эти случаи колексификации встречаются в других языках и отображают семантические связи, свойственные этим языкам. Мы оставим эти случаи за рамками рассмотрения.

Следует заметить, что поскольку целью CLICS является агрегация данных о случаях колексификации в языках мира, можно было бы надеяться увидеть в БД полное совпадение, найти все случаи уже описанными. Однако это оказывается лишь отчасти так.

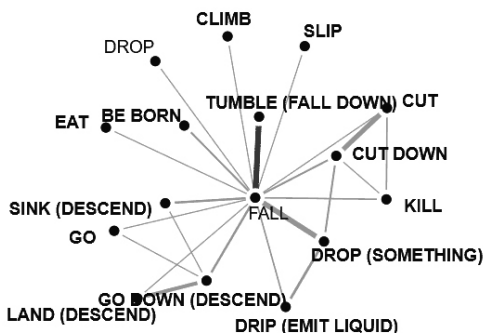


Схема 3. Колексификации с исходным значением FALL в базе данных CLICS 3.

#### 4.2. Анализ с помощью Каталога семантических переходов Datsemshift

Рассмотрим, каким образом можно проанализировать структуру полисемичной словарной статьи статьи *vizu* ‘падать’ с помощью Базы данных семантических переходов в языках мира Datsemshift 2.0. Для этого соотнесем граф словарной статьи *vizu* со списком семантических переходов с участием значения *to fall* в базе данных Datsemshift. Всего на настоящий момент в БД описано 30 переходов с участием значения *to fall*. Мы обнаруживаем совпадения – так, в структуре словарной статьи реализуются переходы:

- *fall down* – *to decrease drastically*, ID 0342, значение (6) *to fall low, as prices*;
- *to fall down* – *to fail*, ID 0592;
- *to fall down* – *to die*, ID 0975 (значение (8), также полисемия *utir*);
- *to fall down* – *to set (of heavenly bodies)*, ID 1319, значение (9), ср. англ. *sunfall* ‘закат’;
- *to fall down* – *to happen*, ID 0934;
- *to fall down* – *to get (in games)*, ID 5148;
- *to fall down* – *autumn*, ID 2793, ср. тамильский дериват *ilaiyutirkālam* ‘осень’ от *utir* ‘опадать (о листьях)’ и др.

Однако, по-прежнему нет возможности соотнести с описанием в рамках Datsemshift следующие значения *vizu* ‘падать’: *fall* – *to sink* (6); *fall* – *to settle, to be fixed* (11); *fall* – *to issue, proceed from. As speech* (23); *to fall* – *to be locked up, as a capital* (30) (ср., *станция попала в окружение*) и другие.



### 4.3. Анализ с помощью теоретического аппарата фреймов Московской лексико-типологической группы

Семантические переходы как таковые не являются главным объектом изучения в рамках фреймового подхода. Они исследуются при изучении переносных значений лексемы, ее метафорических проявлений. Как пишут авторы, «метафорическое разнообразие поля укладывается в ограниченный набор деривационных моделей. Источником для этих моделей являются фреймы — прототипические ситуации, которые регулярно в языках используют специальные средства лексического кодирования» [Рахилина, Резникова, Рыжова 2020: 105 и далее]. Каждая из этих ситуаций профилирует свой аспект в общей семантике семантического поля – и именно этот аспект получает метафорическую интерпретацию при семантическом сдвиге.

Преимуществом фреймового подхода является совмещение в себе трех основных подходов к изучению семантики: «и работы со словарями, с корпусами (в случае, если они доступны) и с носителями-информантами» [Рыжова 2018: 20], однако, трудоемкость подобной работы ограничивает возможности использования метода для широкого круга лексики.

В рамках фреймового метода был осуществлен проект «Типология падения» на материале 42 языков (см. сборник [Типология глаголов падения 2020]), в котором также участвовал автор настоящей статьи. Приводим фрагмент таблицы с перечислением основных описанных авторами деривационных моделей метафорического переноса<sup>4</sup>.

<b>Исходный фрейм</b>	<b>Производное значение</b>	<b>Примеры метафорической конструкции</b>
Падение сверху	LESS/BAD IS DOWN. Начало состояния. Наблюдатель в конечной точке и др.	<i>Температура упала. Свет падает на него. Ему упал выигрыш.</i>
Потеря вертикальной ориентации	Потеря функциональности (о человеке). Превращение и др.	<i>Проигрши. Падение в бою или от болезни. Превращение воды в лед.</i>
Падение-разрушение	Потеря функциональности (неодуш). Капитуляция.	<i>Система рухнула.</i>
Открепление (выпадение из контейнера)	Пропуск элемента. Отставание и выпадение из группы. Утеря качеств, утрата благ.	<i>Выпало из головы. Лекция выпала. Выпал из группы и т.д.</i>

*Схема 4. Некоторые модели семантических сдвигов глаголов падения по [Типология глаголов падения 2020: 105]*

<sup>4</sup> Полный список фреймов падения, выделенных участниками проекта «Типология падения» в процессе работы над этим полем в языках различных семей, можно найти в статьях Рахилиной, Резниковой, Рыжовой из того же сборника [Типология глаголов падения 2020].

Значение падения с высоты:

- (1) *tiḥireṅru pantu eṅkaḷ kāl-aṭi-y.il viḥu-nt-atu*  
вдруг мяч наш нога-подножие-Лос падать-Pst-3Sgn.  
'Вдруг нам под ноги упал мяч'.<sup>5</sup>

Значение падения высокого объекта в сторону:

- (2) *veṭṭa-ppa.ṭṭ-a maram tavarāṇa ticaiy-il viḥu-nt-atu*  
пилить-Pass.Pst-Птср дерево неверный направление-Лос падать-Pst-3Sgn  
'Спиленное дерево упало не в ту сторону'.

Значение выпадения из контейнера:

- (3) *paīyil-iruntu maṇi pars kīzē viḥu-nt-atu*  
сумка-Abl деньги кошелек вниз падать-Pst-3Sgn  
'Кошелек выпал из сумки'.

Значение падения-отсоединения:

- (4) *maratt-iṅ kiḷai tiḥireṅa muṛi-ntu cālaiy-il viḥu-nt-atu.*  
дерево-Gen ветка вдруг отломиться-Vbp дорога-Лос падать-Pst-3Sgn  
'Ветка дерева внезапно отломилась и упала на дорогу'.

А также некоторые переносные значения:

- (5) *cūriyaṅ malai vāyil viḥu-ntu koṅṭiru-nt-atu*  
солнце гора подножие падать-Vbp продолжать-Pst-3Sgn  
'Солнце садилось у подножия горы'.
- (6) *piṭkāy-iṅ matippu \$37-āka viḥ-nt-atu*  
биткойн-Gen оценка 37-Adv падать-Pst-3Sgn  
'Цена биткойна упала на 37 долларов'.
- (7) *avaṅ vākkuvātatt-il viḥu-ntu-viṭṭ-āṅ*  
он спор-Лос падать-Vbp-оставлять-Pst-3Sgm  
'Он проиграл в споре'.
- (8) *Inta maṇitaṅ viḥu-ntu pōṇāṅ*  
этот человек падать-Vbp уходить-Pst-3Sgm  
'Этот человек умер' (букв. «упал»).

#### 4.4. Анализ с помощью объединения подходов

Как нам представляется, наилучший результат можно получить, если рассматривать значения в словарной статье с помощью объединения разных подходов. Рассмотрим некоторые примеры.

---

<sup>5</sup> Часть примеров взята из статьи автора в сборнике «Типология глаголов падения», 2020, данные собраны с помощью анкетирования носителя языка, а также из текстов.

- (9) *piṭkāy-iṅ matippu Ṣ37-āka vīz-nt-atu*  
 биткойн-Gen оценка 37-Adv падать-Pst-3Sgn  
 ‘Цена биткойна упала на 37 долларов’.

Значение ‘уменьшаться’ – результат семантической деривации значения ‘падать’, фрейм ‘**падение сверху вниз**’. Представляет регулярный семантический переход **fall down – to decrease drastically**, ID 0342, регистрируемый на данный момент в 15 языках, начиная с латинского *cado* (*pretium cadit* ‘цена падает’) и русского *Давление падает*, также в коми, адыгейском, турецком, французском, венгерском языке и др.

- (10) *avaṅ vākkuvātatt-il viṣu-ntu viṭṭ-āṅ*  
 он спор-Loc падать-Vbp оставлять-Pst-3Sgm  
 ‘Он проиграл в споре.’ (букв. ‘он упал в споре’)

Значение ‘проиграть’ – результат семантической деривации значения ‘падать’, фрейм ‘**потеря вертикальной ориентации**’. Представляет регулярный семантический переход **to fall down – to fail**, ID 0592, регистрируемый также в русском (ср. *Он провалился на экзамене*), исландском, японском, французском, лезгинском и других языках.

- (11) *Inta maṇitaṅ viṣu-ntu pōṅāṅ*  
 этот человек падать-Vbp уходить-Pst-3Sgm  
 ‘Этот человек умер’ (букв. «упав, ушел»).

Значение ‘умирать’ – результат семантического сдвига значения ‘падать’, фрейм ‘**потеря вертикальной ориентации**’. Представляет регулярный семантический переход **to fall down – to die**, ID 0975, регистрируемый на данный момент в 25 языках, – в т.ч. рус. *пасть в бою*, в греческом, японском, испанском, армянском, геэз, тагалог, коми, литовском, суахили, санскрите, шведском и др.

- (12) *cūriyaṅ malai vāyil viṣu-ntu koṅṭiru-nt-atu*  
 солнце гора подножие падать-Vbp продолжать-Pst-3Sgn  
 ‘Солнце садилось у подножия горы’ (букв. ‘продолжало падать’)

Значение ‘опускаться (о небесном теле)’ связано семантическим переходом со значением ‘падать’, фрейм ‘**падение сверху вниз**’. Представляет регулярный семантический переход **to fall down – to set (of heavenly bodies)**, ID 1319, регистрируемый на данный момент в БД *Datsemshift* в восьми языках – японском, пушту, эвенкийском, латыни и др.

## 5. Выводы

В настоящей статье были рассмотрены разные подходы к изучению семантических переходов (лексико-семантических сдвигов). Были рассмотрены такие направления исследований, как теория семантических переходов и этнолингвистический подход, фреймовый подход в семантической типологии и исследования по дистрибутивной семантике, проекты по колексификации и разработке семантических карт и др. Было показано их соотношение на примере анализа полисемичной та-

мильской словарной статьи. Рассмотрев словарную статью *vizi* ‘падать’ в словаре Tamil Lexicon, мы смогли проанализировать с помощью объединения исследовательских подходов некоторые из значений полисемичной лексемы и высказать предположения о семантических переходах, связывающих эти значения со значением ‘падать’, принятым за исходное.

Словарную статью полисемичной лексемы, как было показано, можно представить как совокупность случаев семантических переходов, реализованных полисемией, и далее, визуализировать это представление в виде графа семантических переходов. Изучение такого графа с разных теоретических позиций может дать новую информацию и показать преимущества и недостатки современных способов теоретического осмысления полисемии.

Как нам представляется, наилучших результатов в изучении структуры полисемичной словарной статьи как источника данных о семантических сдвигах можно достичь именно с помощью объединения теоретических подходов. Как именно, по каким принципам должно происходить это объединение – нам представляется значимым направлением будущих исследований.

### Литература

1. БРЭ: Тамильский язык. // Гл. ред. Ю. С. Осипов. *Большая российская энциклопедия*, 2004–2021. URL: <https://bigenc.ru/>. Дата обращения: 08.07.2021
2. Грунтов И. А. «Каталог семантических переходов» – база данных по типологии семантических изменений // *Компьютерная лингвистика и интеллектуальные технологии. Труды международной конференции Диалог-2007*. М.: 2007
3. Зализняк Анна А. Семантическая деривация в синхронии и диахронии: проект создания «Каталога семантических переходов» // *Вопросы языкознания*. 2001. № 2. С. 13–25.
4. Зализняк Анна А. Семантический переход как объект типологии. // *Вопросы языкознания*, 2013. № 2. С.32–51
5. Михайлова Т.А. О логике семантической деривации в возникновении и эволюции терминов свойства: др.ирл. *aithech tige* vs. др.англ. *husband* // *Вестник МГУ. Серия 9. Филология*. 2013. № 4. С. 126–137.
6. Михайлова Т.А. Дороги, которые мы выбираем: о реализации понятия № 67 «дорога» списка Сводеша в кельтских языках // *Вопросы языкознания*. 2019. № 5. С. 101–119.
7. Рахилина Е.В., Резникова Т.И. Фреймовый подход к лексической типологии. // *Вопросы языкознания*. 2013. №2. С. 3–31.
8. Руссо М. М. О направлении семантических переходов // *Лингвистика и методика преподавания иностранных языков*. М., 2019. С. 181–190.
9. Рыжова Д.А. *Автоматизация лексико-типологических исследований: методы и инструменты*. Дисс. ... кандидата филологических наук. М., 2018.

10. Типология глаголов падения. // *Acta linguistica Petropolitana. Труды Института лингвистических исследований РАН*. Глав. ред. Е. В. Головки. Т. XVI. Ч. 1. *Типология глаголов падения*/ Ред. Е.В. Рахилина, Т.И. Резникова, Д.А. Рыжова, С.Ю. Дмитренко. СПб.: ИЛИ РАН, 2020. 1096 с.
11. Толстая С.М. Мотивационные семантические модели и картина мира // *Русский язык в научном освещении*. М., 2002. № 1. С. 113–127.
12. Толстая С. М. Категория родства в этнолингвистической перспективе // *Категория родства в языке и культуре*. М., 2009. С. 7–22).
13. CLICS: *Database of Cross-Linguistic Colexifications*. Marburg: Forschungszentrum Deutscher Sprachatlas (Version 1.0). List, Johann-Mattis, Thomas Mayer, Anselm Terhalle, and Matthias Urban (2014). URL: <https://clics.cld.org/>. Accessed on: 11.07.2021
14. DatSemshift 2.0: Анна.А. Зализняк, М.А. Бибаева, М.А. Бобрик-Фремке, М.К. Воронов, И.А. Грунтов, О.В. Дереза, Т.А. Михайлова, М.В. Орлова, М.М. Руссо, А.А. Смирнитская. *Каталог семантических переходов в языках мира, DatSemshifts 2012–2019*. URL: <http://datsemshift.ru/>
15. Dubossarsky H., Weinsall D., Grossman E. Verbs change more than nouns: a bottom-up computational approach to semantic change. // *Lingue e linguaggio*. 2016. Vol. XV.1. P. 5–25.
16. Fortescue M. The semantic domain of emotion in Eskimo and neighbouring languages // Juvonen P., Koptjevskaja-Tamm M. (eds.). *The lexical typology of semantic shifts*. Berlin: De Gruyter Mouton. 2016. P. 285–335.
17. François A. Lexical tectonics: mapping structural change in patterns of lexification. // Thanasis Georgakopoulos & Stéphane Polis (eds), *The future of mapping: New avenues for semantic maps research*. Special issue of *Zeitschrift für Sprachwissenschaft* 2021 (в печати).
18. Gast V., Koptjevskaja-Tamm M. Patterns of persistence and diffusibility in the European lexicon. // *Linguistic Typology*. 2021 (to appear in) <https://doi.org/10.1515/lingty-2021-2086>
19. Georgakopoulos A., Polis S. The semantic map model: State of the art and future avenues for linguistic research. // *Language and Linguistics Compass*. 2018. Vol. 12. № 2. P. 1–33.
20. *Graph Online*: инструмент создания графов. URL: <https://graphonline.ru/#> Accessed on: 06.06.2021.
21. Juvonen P., Koptjevskaja-Tamm M. (eds.). *The lexical typology of semantic shifts*. Berlin: De Gruyter Mouton, 2016.
22. Koch P. Lexical typology from a cognitive and linguistic point of view. // M. Haspelmath, E.König, W.Oesterreicher, W.Raible (eds.) *Language typology and language universals: An international handbook*. Vol. 2. Berlin & New York: Walter de Gruyter, 2001.
23. Koch P. Meaning change and semantic shifts. // Juvonen P., Koptjevskaja-Tamm M., eds. *The Lexical Typology of Semantic Shifts*. Berlin; Boston: Mouton de Gruyter, 2016.

24. Russo M. Differences and interactions between scientific and folk biological taxonomy // Juvonen P., Koptjevskaja-Tamm M. (eds.). *The lexical typology of semantic shifts*. Berlin: De Gruyter Mouton. 2016. P. 493-532.
25. Rzymiski, C., Tresoldi, T., Greenhill, S.J. et al. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. // *Sci Data* 2020. Vol. 7, 13.
26. Schlechtweg, D., McGillivray, B., Hengchen, S., Dubossarsky, H., & Tahmasebi, N. (2020). Semeval-2020 task 1: Unsupervised lexical semantic change detection. // *Proceedings of the Fourteenth Workshop on Semantic Evaluation*. ACL, 2020.
27. Tahmasebi N., Borin L., Jatowt A., Xu Y., Hengchen S. (eds.). *Computational approaches to semantic change*. Berlin: Language Science Press. 2021.
28. *Tamil lexicon*. Madras: University of Madras, 1924-1936.
29. Vanhove M. (ed.) *From polysemy to semantic change. Towards a typology of lexical semantic associations*. (Studies in language companion series, 106.) Amsterdam, 2008.
30. Zalizniak Anna A. The catalogue of semantic shifts: 20 years later. // *Russian journal of linguistics*. 2018. Vol.22. №. 4. P. 770–787.

## СВЕДЕНИЯ ОБ АВТОРАХ

**Владимир Михайлович Алпатов** – доктор филологических наук, академик РАН, профессор, заведующий отделом языков Восточной и Юго-Восточной Азии, руководитель Научно-исследовательского центра по национально-языковым отношениям. <https://orcid.org/0000-0003-4323-2832>

**Ксения Владиславовна Антонян** – кандидат филологических наук, старший научный сотрудник Отдела языков Восточной и Юго-Восточной Азии Института языкознания РАН. <https://orcid.org/0000-0002-8531-3641>

**Станислав Борисович Белецкий** – кандидат филологических наук, преподаватель кафедры иностранных языков и литературы Университета Додомы (Танзания), преподаватель школы филологии НИУ «Высшая школа экономики» (Москва). <https://orcid.org/0000-0003-3372-1480>

**Анна Григорьевна Белова** – доктор филологических наук, главный научный сотрудник Отдел языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0001-6748-0342>

**Андрей Андреевич Блинов** – кандидат филологических наук, старший научный сотрудник Отдел языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0002-9694-8706>

**Елена Георгиевна Борисова** – доктор филологических наук, профессор, профессор кафедры германистики и лингводидактики Института иностранных языков Московского городского педагогического университета. <https://orcid.org/0000-0003-3878-5344>

**Юлия Вадимовна Будкина** – аспирант Института востоковедения РАН. <https://orcid.org/0000-0001-8485-8853>

**Марцис Юрьевич Гасунс** – кандидат филологических наук, независимый исследователь. <https://orcid.org/0000-0003-4513-884X>

**Лилия Михайловна Горелова** – кандидат педагогических наук, доктор филологических наук. <https://orcid.org/0000-0001-6911-0118>

**Павел Валерьевич Гращенков** – доктор филологических наук, ведущий научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН, доцент кафедры теоретической и прикладной лингвистики филологического факультета Московского государственного университета им. М.В. Ломоносова. <https://orcid.org/0000-0001-9754-2452>

**Кирилл Игоревич Демич** – научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН

**Замира Касымбековна Дербешева** – доктор филологических наук, профессор Кыргызско-Турецкого университета Манас. <https://orcid.org/0000-0003-4333-4425>

**Лейли Рахимовна Додыхудоева** – кандидат филологических наук, старший научный сотрудник Сектора иранских языков Института языкознания РАН. <https://orcid.org/0000-0002-4567-9454>

**Герман Борисович Дудченко** – кандидат исторических наук, доцент кафедры регионоведения и экономики Невского института языка и культуры. <https://orcid.org/0000-0002-1137-3799>

**Елена Николаевна Емельченкова** – кандидат филологических наук, доцент кафедры китайской филологии Санкт-Петербургского государственного университета. <https://orcid.org/0000-0001-5090-3011>

**Алексей Сергеевич Зверев** – научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0001-9059-4057>

**Александр Вячеславович Костыркин** – кандидат филологических наук, старший научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН.

**Алексей Александрович Кретов** – доктор филологических наук, профессор, зав. кафедрой теоретической и прикладной лингвистики Воронежского государственного университета. <https://orcid.org/0000-0002-1474-3177>

**Елена Михайловна Лазуткина** – кандидат филологических наук, ведущий научный сотрудник Института русского языка им. В. В. Виноградова РАН. <https://orcid.org/0000-0003-0135-8277>

**Юрий Александрович Ландер** – кандидат филологических наук, доцент Школы лингвистики факультета гуманитарных наук Национального исследовательского университета «Высшая школа экономики», ведущий научный сотрудник Международной лаборатории языковой конвергенции. <https://orcid.org/0000-0003-1798-7174>

**Владимир Владимирович Леонченко** – магистр математики и информатики, независимый исследователь. <https://orcid.org/0000-0002-0334-7507>

**Светлана Дмитриевна Малышева** – старший лаборант Отдела языков народов Азии и Африки Института востоковедения РАН.

**Софья Андреевна Маннапова** – кандидат филологических наук, доцент кафедры китайского языка Института иностранных языков Московского городского педагогического университета. <https://orcid.org/0000-0002-5314-6009>



**Тамара Григорьевна Погибенко** – кандидат филологических наук, старший научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0002-6667-1929>

**Елена Леонидовна Рудницкая** – доктор филологических наук, ведущий научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0002-4455-262X>

**Анна Александровна Смирнитская** – научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0001-7814-8716>

**Ольга Валерьевна Столбова** – доктор филологических наук, главный научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0002-7589-2854>

**Джой Иосифовна Эдельман** – доктор филологических наук, профессор, главный научный сотрудник Сектора иранских языков Института языкознания РАН. <https://orcid.org/0000-0003-0056-6250>

**Илья Сергеевич Якубович** – доктор филологических наук, ведущий научный сотрудник Отдела языков народов Азии и Африки Института востоковедения РАН. <https://orcid.org/0000-0002-7576-7930>

*Научное издание*

**ПРОБЛЕМЫ ОБЩЕЙ  
И ВОСТОКОВЕДНОЙ ЛИНГВИСТИКИ.**

**Сочетаемость языковых единиц и языковые модели  
Памяти З. М. Шаляпиной (1946-2020)**

Сборник статей

Верстка: А. С. Панина

Подписано 22.12.2021  
Формат 60x90x16. Бумага офсетная.  
Усл. печ. л. 20,125. Уч.-изд. л. 21  
Тираж 500 экз. Зак. №

Федеральное государственное бюджетное учреждение науки  
Институт востоковедения РАН.  
107031 Москва, ул. Рождественка, 12  
Научно-издательский отдел.  
Зав. отделом И. В. Федулов.  
E-mail: [izd@ivran.ru](mailto:izd@ivran.ru)

Отпечатано в ООО «Белый ветер»  
115093, г. Москва, ул. Щипок, дом 28  
тел. (495) 651-84-56