

Научная статья

УДК 81-139

DOI 10.25205/1818-7935-2022-20-3-90-108

База данных русского жестового языка поликлинического предназначения: лингвистические особенности материала и аннотирования

**Ильдар Амирович Кагиров¹
Дмитрий Александрович Рюмин²**

^{1,2}Санкт-Петербургский Федеральный исследовательский центр РАН (СПб ФИЦ РАН)
Санкт-Петербург, Россия

¹kagirov@iias.spb.su, <https://orcid.org/0000-0003-1196-1117>
²ryumin.d@iias.spb.su, <https://orcid.org/0000-0002-7935-0569>

Аннотация

Настоящая статья описывает цифровую базу данных русского жестового языка по тематической области «Первичный прием у врача-терапевта». Конкретной практической целью сбора базы данных является создание системы автоматического перевода с РЖЯ, основанной на технологиях глубокого машинного обучения. Вместе с тем, описываемая база данных представляет определенный интерес и с точки зрения лингвистики жестовых языков. Собранный материал позволяет проиллюстрировать ряд синтагматических явлений в слитной жестовой речи, аналогичных фонетической синтагматике, таких как эпентеза, ассимиляция, редукция и выпадение удержания. Данная работа проводится сотрудниками Лаборатории речевых и многомодальных интерфейсов СПб ФИЦ РАН в рамках междисциплинарного исследования по созданию системы машинного перевода с русского жестового языка.

Принципиальным отличием записанной базы данных от большинства датасетов, ориентированных на решение задач автоматического распознавания жеста, является наличие слитных высказываний и присутствие элементов развития собственно РЖЯ, а не калькирующей речи. К основным целям настоящей статьи относятся описание принципов создания базы данных, процесса ее формирования и записи, а также обоснование принятой системы аннотирования.

Автоматическая сегментация жестовой речи на минимальные компоненты является нетривиальной задачей, из-за вариативности формы (ср. плана выражения) жестов в слитной речи, серьезно затрудняющей четкое определение межжестовых границ.

В настоящей работе жест анализируется по модели движение-удержание, что позволяет выделить функциональное ядро жеста и отследить потенциальные вариации жестовых сегментов. Соответственно, в процессе аннотирования базы данных каждое высказывание на жестовом языке разбивалось на сегменты движения и удержания; полученные схемы жестов заносились затем в отдельный каталог, содержащий информацию об основных параметрах каждого удержания, в том числе и о возможности его контактной модификации. Результатом проделанной работы является полная декомпозиция записанных жестов, которая может быть использована при создании статистических моделей русского жестового языка.

К собственно лингвистическим особенностям базы данных относятся вариативность жестов, спонтанное переключение кодов (русский жестовый язык – калькирующая речь) и маусинг. Информанты, принявшие участие в записи базы данных, являются учащимися колледжа для глухих и овладели жестовой речью в разных регионах России, поэтому потенциально собранный материал может служить инструментом для исследования диалектной вариативности в РЖЯ.

Ключевые слова

русский жестовый язык (РЖЯ), жестовые языки, аннотирование языковых корпусов, вариативность в жестовых языках, структура жеста

© Кагиров И. А., Рюмин Д. А., 2022

ISSN 1818-7935

Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2022. Т. 20, № 3
Vestnik NSU. Series: Linguistics and Intercultural Communication, 2022, vol. 20, no. 3

Для цитирования

Кагиров И. А., Рюмин Д. А. База данных русского жестового языка поликлинического предназначения: лингвистические особенности материала и аннотирования // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2022. Т. 20. № 3. С. 90–108. DOI 10.25205/1818-7935-2022-20-3-90-108

Благодарности

Авторы статьи благодарят студентов и сотрудников ФКПОУ «Межрегиональный центр (колледж)» Минтруда России г. Павловска (Пушкинский район, Санкт-Петербург) за неоценимую помощь при записи базы данных. Исследование выполнено за счет гранта Российского научного фонда № 21-71-00141, <https://rscf.ru/project/21-71-00141/>.

Russian Sign Language Database for Clinical Use: Data and Annotation Peculiarities

Ildar A. Kagirov¹, Dmitry A. Ryumin²

^{1,2}St. Petersburg Federal Research Center of RAS (SPC RAS)
St. Petersburg, Russia

¹kagirov@iias.spb.su, <https://orcid.org/0000-0003-1196-1117>
²ryumin.d@iias.spb.su, <https://orcid.org/0000-0002-7935-0569>

Abstract

This work has been carried out by the members of the Laboratory of Speech and Multimodal Interfaces of the St. Petersburg Federal Research Center of the Russian Academy of Sciences within an interdisciplinary research project aimed at the creation of an automatic Russian sign language translation system. The paper presents the design of a Russian sign language digital database for a specific subject area, namely, «The first visit to doctor».

Our database is meant to be used first of all a dataset for training neural network-based systems of automatic translation from Russian sign language. But also, it can be of interest for linguistics of sign languages in general since the approach elucidates solution of a number of distorting phenomena typical for continuous sign language tracking, such as epenthesis, assimilation, reduction, and hold deletion.

The principal difference between the presented video data and other datasets developed for similar purposes is the use of continuous sign utterances and elements of Russian sign language proper instead of the so-called “signed Russian” (that is, the visual form of the Russian spoken language), popular in deaf schooling.

One of the most challenging problems, dealt with in the paper, is an automatic segmentation of sign speech into separate meaningful units with clearly defined boundaries. Its solution is hampered by signs’ deformations in continuous speech. In this paper, the segmentation is carried out within the framework of the movement-hold model. This approach allows extraction of their functional core as well as annotation of possible changes likely to appear in different signs’ segments. Accordingly, each utterance was subdivided into segments of motion and hold, and the resulting sign schemes were then entered into a separate directory containing information about the main parameters of each hold, including the possibilities of change due to immediate environment. The result of this work is a full decomposition of the signs, forming the database which can find its application in different statistical models of Russian Sign Language.

Among other linguistic peculiarities of the database should be noted lexical variability of signs, mouthing and code switching between Russian sign language and “signed Russian”. Also, our signers learned Russian sign language in different regions of Russia, thus the collected data is potentially a source for research in dialectal variability of Russian sign language.

Key words

Russian sign language, sign languages, linguistic corpora annotation, variability in sign languages, sign structure

For citation

Kagirov I. A., Ryumin D. A. Russian Sign Language Database for Clinical Use: Data and Annotation Peculiarities. *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2022, vol. 20, no. 3, pp. 90–108. DOI 10.25205/1818-7935-2022-20-3-90-108

Acknowledgements

The authors of the paper are deeply grateful to the students and professors of the Interregional Center (College) of the Ministry of Labor of Russia in Pavlovsk for their invaluable assistance in recording the database. This research was financially supported by the Russian Science Foundation (No. 21-71-00141, <https://rscf.ru/en/project/21-71-00141/>).

Введение

В настоящей статье представлена цифровая база данных русского жестового языка (далее – РЖЯ) по теме «Первичная консультация у врача-терапевта». Собранная языковая база данных (далее – БД) необходима для разработки машинного обучения искусственной нейронной сети, осуществляющей перевод с РЖЯ на русский язык. Так, известно, что при вербальном общении со слышащими людьми глухие «читают» артикуляцию губ («маусинг», беззвучная артикуляция соответствующих лексем). Однако в некоторых условиях (врач в медицинской маске) такое общение слабослышащего с врачом сильно затруднено. Разрабатываемая система машинного перевода может стать инструментом для решения проблемы общения глухих на приеме у врача в отсутствие сурдопереводчика. Основная цель данной статьи заключается в представлении собранной БД научному сообществу, описании явлений РЖЯ, отраженных в собранном языковом материале, и принципов аннотации языкового материала РЖЯ.

Статья организована следующим образом: после введения в разделе 1 представлена краткая информация о РЖЯ и существующих базах данных жестовых языков (в том числе медицинских). Следующие два раздела содержат описание сбора БД и входящей в нее жестовой лексики. В разделах 4 и 5 обсуждаются существующие способы анализа жеста и описываются основные принципы аннотирования настоящей БД, после чего следует раздел, перечисляющий лингвистические особенности собранной БД с точки зрения лингвистики жестовых языков. Наконец, в заключении сформулированы основные результаты проделанной работы.

1. РЖЯ и базы данных

Актуальность исследования жестовых языков, в том числе и РЖЯ¹, становится все очевиднее. Однако на сегодняшний день существует относительно небольшое количество баз данных РЖЯ, и чаще всего они ориентированы на учебные или строго лингвистические задачи. Одна из наиболее репрезентативных баз данных РЖЯ была собрана в 2014 году в Новосибирском государственном техническом университете [КРЖЯ] и насчитывает порядка 230 записей непрерывной жестовой речи, продемонстрированных 40 дикторами. Аннотация корпуса выполнялась в среде ELAN.

Остальные базы данных РЖЯ представляют собой видеоуроки, содержащие отдельные лексемы или словосочетания. Наиболее значимыми из них являются толковый словарь РЖЯ RuSLED (2002) и проект Сурдосервер 2.0. К сожалению, эти базы данных не аннотированы, поэтому данные из них трудно, практически невозможно использовать для лингвистических или компьютерных целей.

Не существует в настоящее время и хороших специализированных жестовых баз данных по медицинской тематике. Как правило, их создатели используют общие базы данных, из которых извлекается вся лексика, связанная с медициной и здравоохранением [Vidalon et al, 2015; Souza et al, 2021]. Лишь в работе Kosmopoulos et al [2020] был предложен другой подход: создание набора данных при помощи приглашенных экспертов. В эту БД включен 21 диалог, смоделированный на основе реального опыта взаимодействия лечащих врачей с пациентами. Сюда же можно отнести еще ряд языковых наборов данных, относящихся к теме здравоохранения и медицины [Necati et al, 2016; Medslet].

Ранее авторы данной статьи собрали мультимедийную базу данных русского жестового языка TheRuSLan, которая содержит лексические единицы по теме «Покупки в супермаркете». К особенностям базы данных [Кагиров и др., 2020] относится использование нескольких модальностей, а также принципы аннотирования, удобные как для машинного обучения, так и для лингвистического анализа.

¹ РЖЯ является языком общения глухих в России и странах ближнего зарубежья. По данным Ethnologue [<https://www.ethnologue.com>], в 2010 г. количество носителей РЖЯ составляло 122 тыс. человек. В 2012 г. РЖЯ был официально признан одним из языков общения в Российской Федерации.

2. Описание базы данных

На данный момент база данных состоит из 85 жестовых высказываний, записанных 12 дикторами в пяти итерациях. Наличие итераций продиктовано возможностью использования данной БД для задач машинного обучения и позволяет отследить появление вариативности в их реализации. Общий размер базы данных в исходном формате составил примерно 165,5 Гб, а общая продолжительность всех видеозаписей, включая итерации и паузы, около семи часов. Важно отметить, что видеозаписи жестов на РЖЯ в БД являются разноракурсными, что повышает их информативность с точки зрения плана выражения.

Запись производилась в учебных помещениях на неоднородном фоне с использованием мобильных телефонов (ОС Android) в формате FHD. Разрешение кадров – 1920 × 1080 пикселей, частота – 60 FPS. В среднем расстояние от камер до диктора составляло два метра. За камерами был расположен ноутбук, на экране которого информантам демонстрировались стимулы для перевода на РЖЯ (рис. 1). Следует отметить, что во время записи БД в помещении постоянно присутствовал специалист по РЖЯ для облегчения коммуникации между информантами и авторами БД.



Рис. 1. Общий вид массива камер для записи базы данных
Fig. 1. A general view of the camera array used for collecting DB

В записи приняли участие 11 студентов и один преподаватель ФКПОУ «Межрегиональный центр (колледж)» Минтруда России г. Павловска (МЦР). Все информанты глухие, использующие РЖЯ в качестве первого языка или же освоившие его в раннем возрасте (4–8 лет). Они являются уроженцами разных городов и населенных пунктов России (табл. 1). В последней графе указаны населенные пункты, где каждый из дикторов провел детство и освоил РЖЯ.

Таблица 1

Информация о дикторах

Table 1

Signers Data

Диктор	Пол	Возраст	Возраст овладения РЖЯ	Продолжительность записи (мин.)	Место постоянного проживания
Диктор 1	ж	21	первый язык	42	г. Йошкар-Ола
Диктор 2	м	18	первый язык	48	г. Ростов-на-Дону
Диктор 3	ж	24	5 лет	37	г. Шадринск
Диктор 4	ж	59	8 лет	47	Ейская обл.
Диктор 5	м	20	7 лет	27	г. Волгоград
Диктор 6	ж	19	первый язык	26	г. Уфа
Диктор 7	ж	21	первый язык	22	г. Киров
Диктор 8	м	19	4 года	33	Тульская обл.
Диктор 9	м	24	первый язык	31	Псковская обл.
Диктор 10	м	22	первый язык	42,5	Ленинградская обл.
Диктор 11	ж	19	7 лет	28,5	г. Иваново
Диктор 12	м	21	первый язык	33	г. Нижний Новгород
				Всего ≈ 417 мин.	

3. Лексический состав БД

Предметно-тематическая область БД разделена на следующие коммуникативные кластеры, представленные в таблице 2: «Общая информация и проведение консультации», «Анамнез», «Обследование», «Диагностика и рекомендации». Словарь кластеров создавался совместно с практикующими терапевтами на базе типичных диалогов между врачом и пациентом², за основу было взято пособие для терапевтов [Драпкина, 2019].

Из таблицы 2 видно, что в БД включены только реплики врача (слышащего), а не глухого пациента. Такое решение продиктовано тем, что гипотетически разнообразие релевантных ответов пациента чрезвычайно велико, и создание такой БД – масштабная задача, требующая значительных временных затрат и существенно большего числа информантов. Потому гораздо более реалистичной стратегией является применение полученного языкового набора данных в интерфейсах, позволяющих не только распознавать, но и синтезировать жестовую речь при помощи 3D-аватара, переводя звучащие высказывания врача в жестовую форму. В будущем предполагается расширение БД за счет ответов пациента, и правила для РЖЯ, полученные на представленном в настоящей статье наборе данных, будут применены и к новому набору данных.

При формировании БД нужно также иметь в виду, что единого стандарта русской жестовой речи на данный момент нет. В сообществе носителей РЖЯ сосуществуют две формы жестовой речи: собственно РЖЯ и так называемая калькирующая жестовая речь (КЖР) [Grenoble, 1992; Зайцева, 2000]; такую языковую ситуацию можно, по аналогии с естественными языками, рассматривать как диглоссную. Фактически КЖР представляет собой идиом-посредник, калькирующий грамматические и лексические структуры звучащей речи [Буркова, Киммель-

² Их можно также назвать типичными речевыми образцами, паттернами или формулами общения в системе «доктор – пациент».

Таблица 2

Типичные речевые образцы общения на РЖЯ, вошедшие в БД

Table 2

RSL Typical Communicative Patterns Included in DB

Общая информация о пациенте и взаимодействие с ним	Анамнез	Обследование	Диагноз и рекомендации
1	2	3	4
Присаживайтесь.	На что жалуетесь?	Сейчас я вас осматрю.	Я назначу вам лечение, пожалуйста, выполняйте все рекомендации.
Покажите, пожалуйста, ваш номерок.	Сколько вам лет?	Ложитесь на кушетку.	Я напишу вам на бумажке, что нужно принимать.
Покажите, пожалуйста, вашу медкарту.	Кем вы работаете?	Ложитесь на живот.	Вам надо проверить зрение.
Вы по записи?	Вы беременны?	Ложитесь на спину	Если после курса лечения не будет улучшения, вас госпитализируют.
Сегодня я не могу вас принять.	Вы на диете?	Не двигайтесь.	Вам требуется срочная госпитализация.
Я приму вас сразу после этого пациента.	У вас есть хронические болезни?	Встаньте.	Вы согласны на госпитализацию?
Подождите в коридоре, пожалуйста.	У вас были серьезные заболевания?	Следите глазами за моим пальцем.	Вам нужно сдать мазок из носа и горла.
Как вас зовут?	У вас есть наследственные заболевания?	Покажите язык.	Диагноз я поставлю только после результата анализов.
Как ваша фамилия?	Сколько раз вы переболели коронавирусом?	Откройте рот широко и высуньте язык.	Это серьезное заболевание, вам надо лечиться.
Вам нужна справка?	От чего у вас были прививки?	Скажите: «А!»	Вы все еще больны, я не могу вас выписать.
Я выпишу вам справку.	Вы привиты от ковида?	Дышите. Не дышите. Дышите глубоко.	Ничего серьезного, но я напишу, что вам принимать.
Забирайте свою медкарту.	Вы курите?	Поднимите-опустите голову.	У вас учащенное сердцебиение.
Медкарту оставьте у меня.	Сколько сигарет вы выкуриваете в день?	Наклонитесь, пожалуйста.	У вас учащенный пульс.

Окончание табл. 2

1	2	3	4
Запишитесь на следующий прием через неделю.	Вы обычно быстро утомляетесь?	Разденьтесь, пожалуйста.	Не забудьте, перед анализом крови ничего есть нельзя!
Возьмите номерок в регистратуре.	Вы всегда так бледны?	Разденьтесь до пояса.	
Вы совершенно здоровы.	У вас есть проблемы с алкоголем?	Тут болит, когда я нажимаю?	
До свадьбы заживет.	Вы давно болеете?	Скажите, когда станет больно.	
Не занимайтесь самолечением.	Вы давно набрали избыточный вес?	Пожалуйста, встаньте на весы.	
	Какая у вас температура в течение дня?	Поверните голову вправо-влево.	
	К вечеру температура обычно поднимается?	Возьмите градусник и измерьте температуру.	
	Какой зуб у вас болит?	Давайте измерим давление.	
	Голова часто кружится?	Закройте глаза.	
	Бывает так, что вы испытываете удушье?	Разуйтесь, пожалуйста.	
	У вас руки дрожат?	Осмотр окончен. Одевайтесь.	
	Вы плохо спите?		
	Какой у вас стул?		
На что у вас аллергия?			
Какие лекарства вы принимаете?			
Вы сильно потеете?			

ман, 2019, с. 79–80]. Кроме того, следует иметь в виду, что жестовый язык, как и любой естественный язык, обнаруживает диалектную и социолектную вариативность [Kimmelman, 2009; Буркова, Варинова, 2012; Королькова, 2017].

В России КЖР – это вторичная жестовая система, которая усваивается в процессе обучения глухого ребенка устной речи в колледжах и спецшколах. В связи с тем, что она обычно используется для переводов на жестовый язык в формальных ситуациях, а также преподается

В настоящее время основные методы распознавания ЖЯ подразумевают применение искусственных нейронных сетей глубокого обучения. Прогресс в области обработки естественных языков, достигнутый за первые десятилетия XXI века, продемонстрировал перспективность нейронного подхода и обусловил отказ от систем машинного перевода, основанных на правилах. Однако эффективность нейронных сетей напрямую зависит от качества и количества обучающих наборов данных. Сегодня существует достаточно много языковых корпусов и наборов данных, аннотированных для задач распознавания параметров жеста, но наборов данных, созданных для работы со слитной жестовой речью (речевым потоком жестов), пока мало⁵. Между тем для успешного распознавания жестов в потоке речи и сегментации цепочки жестов необходимо учитывать специфические для любой слитной речи (включая жестовую) синтагматические и комбинаторные вариации жестов на «фонетическом» уровне, то есть возникающие под влиянием, с одной стороны, непосредственного «жестового окружения» и, с другой стороны, темпа речи.

При разметке БД РЖЯ авторы стремились выделить те параметры жестов, которые позволяют с наибольшей точностью идентифицировать как конфигурации рук диктора, так и границы жестов в потоке. Вместе с тем отсутствие в материале связанных текстов достаточной длины приводит к низкому уровню релевантности грамматической информации, и потому грамматическая разметка не проводится.

Существует несколько основных подходов к описанию собственно фонологической структуры ЖЯ, которую нужно, однако, отличать от фонетического уровня реализации системного набора жестов. В исследовании [Stokoe, 1960] впервые было дано описание внутренней структуры жеста и предложен набор из конечного множества дифференциальных признаков, образующих целостную систему. Так, среди основных признаков жеста Стоуки выделил всего лишь три базовых: конфигурацию, локализацию и движение. Конфигурация нормирует форму кисти руки (кистей рук) в том или ином жесте; локализация – место его исполнения, а движение – траекторию, его визуальный образ в пространстве, воспринимаемый адресатом. Если проводить аналогию с фонетикой звучащих языков, то признаки жеста функционально соответствуют дифференциальным признакам фонем.

В современном научном узусе именно эти признаки называются основными параметрами жеста⁶. Позднее стало ясно, что среди них имеет смысл выделять четвертый признак — ориентацию руки – и обязательно учитывать наличие немануального компонента, представленного в коммуникации прежде всего мимикой и – в меньшей степени – пластикой.

Несмотря на большой теоретический потенциал системы Стоуки, принципы, положенные в его основу, только отчасти подходят для аннотирования обучающих наборов данных. Первая и основная прагматическая проблема заключается в изменчивости жеста. Конфигурация рук и локализация во время жестикуляции не являются неизменными параметрами, варьируясь при выполнении жеста как одним и тем же отправителем, так и различными. Среди других явлений прагматического характера, затрудняющих определение границ между жестами в потоке и их функционального ядра, можно назвать:

1) эпентезу – межжестовые движения, возникающие в потоке речи [Grif, Kugaevskikh, 2020]. Так, при исполнении жеста БОЛЕЗНЬ/БОЛЕТЬ в разговорной речи активная рука в конфигурации \ni касается кончиками пальцев тыльной стороны ладони пассивной руки, однако в полной форме жеста это движение повторяется дважды (рис. 3):

⁵ Тем не менее работа в этом направлении ведется; в качестве примера см. [Гриф и др., 2018].

⁶ Другой термин, используемый в русскоязычной литературе, – «компоненты жеста» [Буркова, Киммельман, 2019, с. 121]. В настоящей статье используются синонимичные термины «параметры жеста» и «признаки жеста».

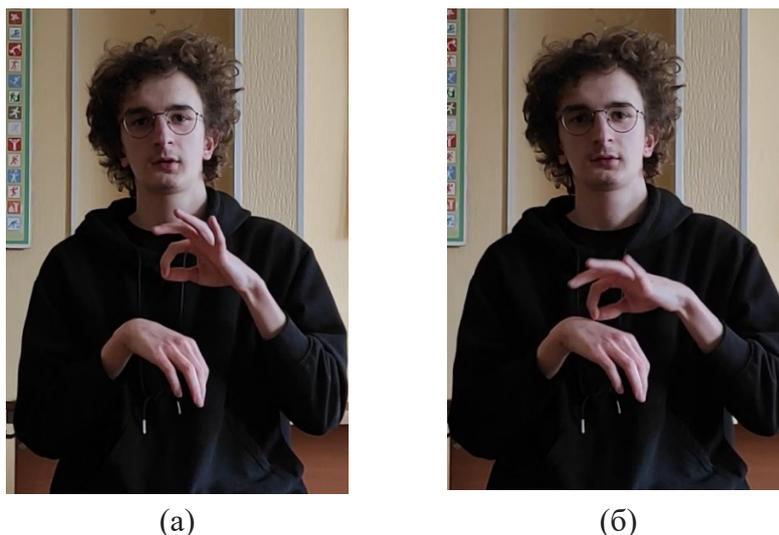


Рис. 3(а-б). Пример эпентезы в жесте «болезнь»
 Fig. 3. An example of epenthesis in the sign “illness”

2) ассимиляцию – результат взаимодействия жестов в потоке речи, приводящий к наложению параметров одного жеста на параметры другого, к их контаминации. Одним из частотных случаев ассимиляции, вошедших в описываемую БД, является так называемая антиципация⁷: в том случае, если двуручный жест следует за одноручным, пассивная рука в первом жесте иногда заранее принимает положение, свойственное ей во втором жесте (рис. 4). Схожим феноменом является персеверация жеста – наследование параметров пассивной руки одноручным жестом в том случае, когда он следует за двуручным.



Рис. 4(а-в). Антиципация и выпадение сегментов при переводе на РЖЯ фразы «Вопросы есть?»
 Fig. 4(а-в). Weak hand anticipation and hold deletion phenomena in RSL utterance “Any questions?”

3) эллипсис – исчезновение некоторых формантов из конфигураций и локализаций между жестами. Так, предыдущая иллюстрация явления антиципации (рис. 4) отображает также вы-

⁷ Здесь и далее (см. ниже термин «персеверация») употребляется терминология из [Буркова, Киммельман, 2019. с. 161]. Ср. с явлением прогрессивной ассимиляции в фонетике.

падение начального элемента жеста БЫТЬ: в своей кодифицированной словарной форме активная рука с конфигурацией \square опускается вниз, образуя, в конечном итоге, контакт с пассивной рукой. Из приведенного примера видно, что начальный сегмент этого жеста был опущен диктором для удобства артикулирования.

Среди прочих явлений, возникающих в потоке жестовой речи, можно перечислить редукцию движения, метатезу (взаимная замена начальной и конечной локализации) и редупликацию [Liddell, Johnson, 1989].

5. Аннотирование БД на принципе линейности

В ряде исследований ЖЯ, увидевших свет после публикации работы Стоуки, вводится еще один параметр жеста – линейность [Sandler 1986; Perlmutter 1992; Brentari 1998; Sandler, Lillo-Martin 2006]. К примеру, в модели Sandler'a [1986, 1989] конфигурация рассматривается как отдельный автосегментный параметр, то есть принадлежащий самостоятельному уровню и ставящийся в соответствие сегментным единицам.

Одним из «линейных» способов описания жеста является модель движения-удержания (англ. the Movement-Hold model, далее – модель МН) [Liddell, Johnson, 1989]. Согласно модели МН, каждый жест может быть разбит на сегменты удержания (фиксации) положения и движения (перехода к следующему), что естественным образом дестабилизирует в какой-то момент все основные параметры жеста. Эта модель позволяет разбить жесты на последовательности статических сегментов (с четко определенными параметрами) и динамических, которые представляют собой связующие элементы между статическими сегментами. В качестве примера рассмотрим описание жеста сидеть/садиться с учетом параметра «Движение» (рис. 5):

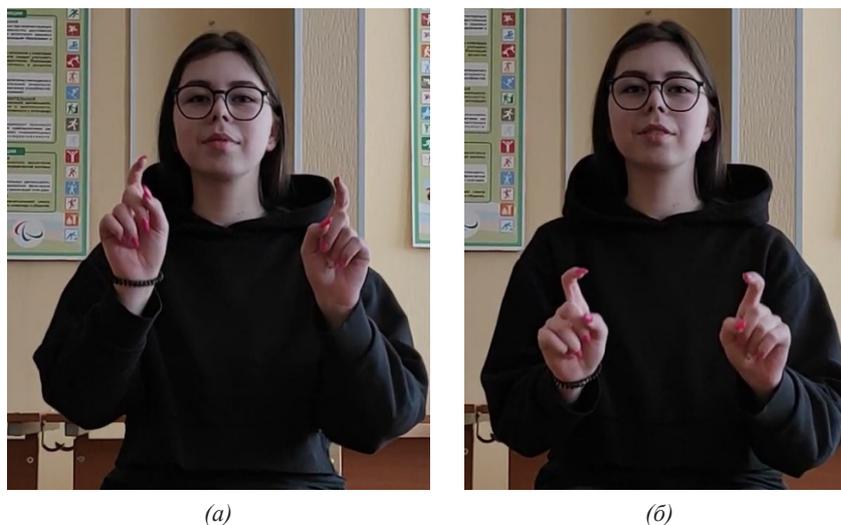


Рис. 5. Две фиксации (а, б) в структуре жеста сидеть/садиться
Fig. 5. Two holds in the structure of the sign “be.sitting/sit.down”

В табличной форме основные параметры удержаний (фиксаций положения) выглядят следующим образом (табл. 3.):

Таблица 3

Параметры сегментов жеста сидеть/садиться

Table 3

Segment Parameters of the Sign be.sitting/ sit.down

сидеть/садиться/be.sitting/sit.down			
	сегменты		
	Hold (H)	Movement (M)	Hold (H)
Правая рука			
конфигурация	☞		☞
локализация	☞		☞
Левая рука	СИММ.		

Среди преимуществ модели МН в контексте поставленных при аннотировании задач стоит упомянуть четкое выделение устойчивых параметров жеста (сегменты удержания) и представление высказывания в виде цепочки удержаний. Подобное представление решает проблему описания вариативности жеста, а также позволяет компактно описывать такие фонетические явления, как эпентеза, ассимиляция и эллипсис или редукция сегмента.

Все высказывания, входящие в нашу БД, были вручную разбиты на сегменты, соответствующие моментам фиксации, и сегменты, соответствующие движению артикуляторов. Вычленение сегментов удержания проводилось по двум критериям:

- 1) отсутствие «смазанных» кадров, вызванных движением артикуляторов диктора;
- 2) максимальная репрезентативность артикуляторов во фронтальной проекции.

Инструментом для сегментации и аннотирования послужила программа ELAN [https://archive.mpi.nl/tla/elan]. Разметка была осуществлена на трех уровнях: а) сегментация по удержаниям и движению, б) морфологическая аннотация и в) перевод. Пример такой аннотации представлен на рис. 6:

The screenshot displays the ELAN 6.3 interface. At the top, there's a menu bar with options like 'Файл', 'Правка', 'Аннотация', 'Слой', 'Тип', 'Поиск', 'Вид', 'Параметры', 'Окно', and 'Справка'. Below the menu is a video player showing a person sitting at a desk. To the right of the video are controls for volume and speed. Below the video is a timeline with various tracks. The 'default' track shows a 'Hold4' segment. The 'Annotation' track shows 'INDX' with markers for 'H', 'H', 'Ho', 'H', and 'Hoi'. The 'Translation' track shows the text 'Вы привиты от коронавируса?'. The timeline is marked with time intervals from 00:00:00.500 to 00:00:02.500.

Рис. 6. Пример «линейно» аннотированного высказывания из БД (ELAN).

Fig. 6. An example of DB annotation (ELAN)

При этом каждому сегменту Hold поставлен в соответствие файл с описанием основных параметров жеста: конфигурации, локализации и, при необходимости, немануального компонента (табл. 4а, 4б):

Таблица 4а

Представление информации о лексемах, вошедших в БД

Table 4a

Information on Lexemes Included in the DB

Hold12	Рука 1	Рука 2
	активная	пассивная
конфигурация	○	○
локализация	☞	☞
ориентация	↑	>
контакт	~ ~ X 1	
немануальный компонент	нет	

Hold13	Рука 1	Рука 2
	активная	пассивная
конфигурация	☞	○
локализация	☞	☞
ориентация	^	>
контакт	~ ~ X 1	
немануальный компонент	нет	

Таблица 4б

Представление информации о лексемах, вошедших в БД

Table 4б

Information on Lexemes Included in the DB

лексема	коронавирус
удержания	Hold12-Move-Hold13
структура	НМН

Таким образом, БД организована в виде структуры, включающей каталоги, содержащие размеченные видеозаписи с выделенными сегментами Hold и Move; отдельный каталог с описанием основных параметров всех сегментов Hold, каждому из которых присвоен порядковый номер; и каталог с лексемами, вошедшими в БД, представленными в виде последовательностей пронумерованных Hold и Move (рис. 7).

Всего при аннотировании БД зафиксировано 317 различных сегментов удержания (Hold) для активной и пассивной рук.

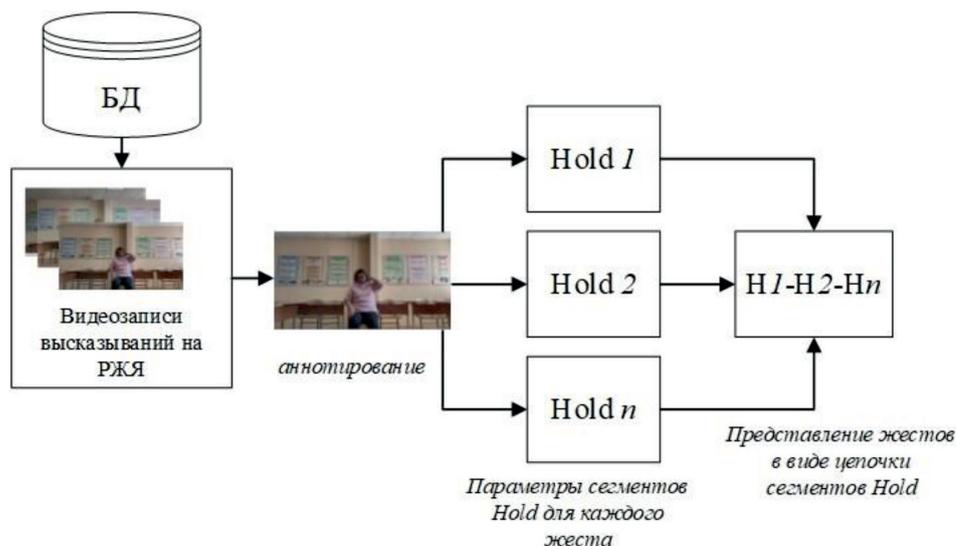


Рис. 7. Логическая структура БД
Fig 7. DB logical structure

Отдельно необходимо пояснить нашу трактовку параметров «конфигурация» и «локализация». В описании жестовых языков мира наблюдается их бесконтрольная пролиферация, исследователи оперируют десятками конфигураций и локализаций, но далеко не все из них релевантны для нашего материала. В работе Кагирова и др. [2020] был предложен компактный способ описания конфигураций и локализаций, следующий логике укрупнения зон жестового пространства. Этот способ показал свою эффективность и при разметке представленного в данном исследовании языкового материала. Конфигурации для одного и того же жеста, продемонстрированные различными дикторами, интерпретировались как аллофонические.

Кроме того, все выявленные эпентезы аннотируются как сегменты движения (Move), а сегменты удержания, подверженные выпадению (эллипсису и редукции), получили особую помету. Персеверация и антиципация (прогрессивная ассимиляция) трактовались как комбинаторные изменения параметров жеста, в связи с чем все зафиксированные примеры этих явлений отмечаются особо. Предполагается, что расширение и дальнейший анализ материала позволят создать статистическую модель структур РЖЯ, с высокой точностью определяющую эти явления⁸.

6. Лингвистические особенности авторской БД

Собранная БД обладает рядом особенностей, которые могут быть интересны с точки зрения изучения РЖЯ. В отличие от ряда существующих словарей РЖЯ, настоящая БД содержит высказывания, а не отдельные лексемы, что позволяет считать собранный материал источником данных по всем синтагматико-комбинаторным явлениям спонтанной жестовой речи.

Была также предпринята попытка оценить влияния диглоссного характера. Дело в том, что практически все дикторы сопровождали жесты РЖЯ маусингом. Для оценки потенциал-

⁸ Завершая параграф, следует упомянуть работу Клезович [Klezovich, 2019], в которой представлена программная реализация алгоритма, автоматически вычлняющего основные формы руки в жестах РЖЯ. Теоретической базой для нее также послужила модель НМ. Основная идея предложенного А. Г. Клезович подхода состоит в том, что для составления списка конфигураций достаточно аннотировать их только в позициях удержания. Разработанный алгоритм извлекает из видео только фиксации (то есть позиции без движения). Несмотря на то, что автор работает с изолированными жестами, проблема вычленения отдельных конфигураций близка к проблемам, затронутым при аннотировании настоящей БД.

ного влияния условий записи на появление маусинга каждому диктору после сеанса записи предлагалось создать спонтанный нарратив в форме рассказа о своих хобби, ярких случаях из жизни и т. п. Эксперимент показал, что все дикторы, кроме шестого и седьмого, сопровождали нарратив маусингом. В качестве рабочей гипотезы было принято предположение о том, что для включенных в БД информантов практически обязательный маусинг вызван, вероятнее всего, интерференцией со звучащим русским языком, обязательным в образовательных стандартах. Однако не стоит исключать и фактор языкового престижа, когда информант, вынужденный работать с языком-посредником (в данном случае письменное представление русского языка), бессознательно меняет свои речевые привычки (см., например, [Кибрик, 1972, с. 81]).

Другой лингвистической особенностью БД является вариативность жестов на лексическом и фонологическом-фонетическом уровне. Как отмечалось в разделе 4, информанты не были ограничены стилистическими и/или диалектными установками, свободно выбирая лексические средства для перевода стимулов. Поэтому, несмотря на искусственность лабораторных условий, полученный речевой материал можно считать репрезентативным и с точки зрения изучения вариативности единиц и структур в РЖЯ.

В первую очередь, это касается различных вариантов перевода одной и той же лексемы. Всего в БД представлено 147 различных лексем и устойчивых словосочетаний, 27 из которых были переведены на РЖЯ по-разному, то есть с использованием принципиально различных параметров жеста (конфигурация, локализация, движение). Это особенно актуально для лексики, специфичной для заданной темы (*диета, анализ, удушье, измерять температуру, медкарта, учащенный* (пульс), *регистратура, избыточный вес* и т. п.).

Другая группа лексем, подверженных переводческой вариативности, представлена частотной бытовой лексикой, представленной ее ключевыми словами *язык, курить, разуваться, год, потеть* и т. п. Эта вариативность предположительно связана с диалектными различиями РЖЯ и/или социолектными особенностями учебного сообщества, к которому принадлежат все информанты.

Зарегистрирована и артикуляционная вариативность одной и той же лексической единицы у разных дикторов. Артикуляция, как правило, отличается по всем основным параметрам: конфигурация кисти, локализация, ориентация и характер движения. Это явление, функционально схожее с аллофоническим варьированием в звучащих языках, естественно в ЖЯ мира [Rozelle, 2003] и хорошо знакомо специалистам в этой области. Тем не менее фиксация жестовых вариантов реализации одних и тех же «фонем» (аллофонов) и лексем в рамках одной БД может быть полезна как для исследования собственно явления вариативности жеста в РЖЯ, так и для создания систем автоматического перевода с РЖЯ, базирующихся на методах машинного обучения [Кагиров и др., 2019].

Стандартная типология варьирования представлена в анализируемой БД следующими позициями:

- 1) выбранные пальцы – варьируется положение выделенных пальцев (они могут быть согнуты в разной степени);
- 2) ориентация ладони – угол, образуемый рукой и одной из жестовых плоскостей, может существенно различаться;
- 3) локализация – в рамках систем детальной нотации, таких как HamNoSys, часто невозможно единообразное описание одной и той же лексической единицы для разных дикторов;
- 4) амплитуда движения – этот параметр изменяется от одного диктора к другому.

Специальный анализ языковых данных на предмет выявления диалектальных особенностей в речи информантов авторами данной статьи не проводился. Не исключено, что вариативность жестов, наблюдаемая при записи БД, связана с тем обстоятельством, что все информанты являются уроженцами разных регионов России, однако определенные выводы по этой теме будут возможны только в результате дальнейших исследований.

Заключение

В данной статье представлена база данных РЖЯ, содержащая лексику по теме «Первичный осмотр у врача-терапевта». Особенности предложенной в статье разметки БД обусловлены описанной выше практической задачей: аннотация отражает прежде всего устойчивые и неизменные параметры жеста, что позволяет использовать представленный набор данных для обучения искусственных нейросетей, работающих с РЖЯ. В то же время инвентарь выделенных жестовых сегментов базируется на системе параметров, применяемых в лингвистике ЖЯ, и может быть интересен с точки зрения лексикографического описания РЖЯ и исследований по структуре жестов в ЖЯ мира.

Среди непосредственных результатов проделанной работы можно перечислить создание (пусть и ограниченного) инвентаря сегментов удержания для РЖЯ, документацию лексической вариативности и контактных межжестовых изменений параметров (эпентеза, ассимиляция, редукция и выпадение удержания). Наконец, БД фиксирует лексику по предметной области «Медицинское обслуживание», что можно рассматривать как вклад в «копилку» электронных ресурсов по РЖЯ.

В ближайшем будущем авторами настоящей статьи планируют расширить БД, провести углубленный анализ фонологико-фонетических явлений с лингвистическими экспериментами, а также создать и разметить отдельную БД из записей спонтанных нарративов, собранных в процессе работы над представленной в статье БД.

Ценность этой БД видится в том, что создание подобных наборов данных помогает расширить знания о РЖЯ. Предполагается использование описанной в настоящей статье базы для создания автоматической системы распознавания РЖЯ, что будет способствовать значительному улучшению качества жизни людей с нарушениями слуха.

Список литературы

- Буркова С. И., Варинова О. А.** К вопросу о территориальном и социальном варьировании русского жестового языка // Федорова О. В. (ред.). Русский жестовый язык: Первая лингвистическая конференция. М.: Буки Веди, 2012. С. 127–143.
- Буркова С. И., Киммельман В. И., Филимонова Е. В., Кюсева М. В., Варинова О. А., Заварицкий Д. А., Приставко К. В., Кадыргулова Р. С.** Введение в лингвистику жестовых языков. Русский жестовый язык. Учебник. Новосибирск: НГТУ, 2019. 356 с.
- Гейльман И. Ф.** Словарь русского жестового языка. В 2 т. СПб: Прана, 2004. 363 с.
- Гриф М. Г., Королькова О. О., Мануева Ю. С.** Разработка алгоритмического и программного обеспечения выбора значения многозначного слова и омонима в системе компьютерного сурдоперевода русского языка на основе семантической модели // *Вестн. Новосиб. гос. ун-та. Серия: Лингвистика и межкультурная коммуникация*, 2018. Т. 16. № 3. С. 32–44.
- Драпкина О. М.** Краткие алгоритмы ведения пациентов на этапе оказания первичной медико-санитарной помощи. Пособие для врачей-терапевтов / под ред. О. М. Драпкиной. М.: Видокс, 2019. 20 с.
- Зайцева Г. Л.** Жестовая речь. Дактилология: Учеб. для студ. высш. учеб. заведений. М.: Гуманит. изд. центр ВЛАДОС, 2000. 192 с.
- Кагиров И. А., Рюмин Д. А., Аксенов А. А., Карпов А. А.** Мультимедийная база данных жестов русского жестового языка в трехмерном формате // *Вопросы языкознания*, 2020. № 1. С. 104–123. DOI: 10.31857/S0373658X0008302-1
- Кибрик А. Е.** Методика полевых исследований (К постановке проблемы). М.: Издательство Московского университета, 1972.

- Королькова О. О.** Изучение функциональной стороны русского жестового языка // *Вестн. Новосиб. гос. ун-та. Серия: Лингвистика и межкультурная коммуникация*, 2017. Т. 15. № 3. С. 67–75.
- Brentari D.** *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press, 1998.
- Camgöz N. C., Kindiroğlu A. A., Karabüklü S., Keleşir M., Özsoy A. S., Akarun L.** Bosphorus Sign: A Turkish Sign Language Recognition Corpus in Health and Finance Domains. *Proc. of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, Portoroz, Slovenia, 2016, pp. 1383–1388. URL: <https://aclanthology.org/L16-1220>
- de Souza R. S., de Martino J. M., Temoteo J., Rodrigues I.** Automatic Recognition of Continuous Signing of Brazilian Sign Language for Medical Interview. *Proc. of the Sixth International Conference on Informatics and Assistive Technologies for Health-Care, Medical Support and Wellbeing HEALTHINFO*, Barcelona, Spain, 2021, pp. 41–46. URL: https://www.researchgate.net/publication/355208551_Automatic_Recognition_of_Continuous_Signing_of_Brazilian_Sign_Language_for_Medical_Interview
- Grenoble L.** An overview of Russian Sign Language. *Sign Language Studies*, 1992, no. 77, pp. 321–338.
- Grif M. G., Kugaevskikh A. V.** Recognition of Deaf Gestures Based on a Bio-Inspired Neural Network. *Journal of Physics: Conference Series*, 2020, 1661 012038. DOI:10.1088/1742-6596/1661/1/012038
- Kimmelman V.** Reflexive Pronouns in Russian Sign Language and Sign Language of the Netherlands: MA Thesis in Linguistics. Amsterdam: Universiteit van Amsterdam, 2009.
- Klezovich A.** Automatic Extraction of Handshapes Inventory in Russian Sign Language. *Linguistics. WP BRP. NRU HSE*, 2019, no. 86. Higher School of Economics Research Paper No. WP BRP 86/LNG/2019.
- Kosmopoulos D., Oikonomidis I., Constantinopoulos C., Arvanitis N., Antzakas K., Bifis A., Lydakis G., Roussos A., Argyros A.** Towards a Visual Sign Language Dataset for Home Care Services. *Proc. of the 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 2020, pp. 520–524. DOI: 10.1109/FG47880.2020.00099
- Leech G.** Introducing Corpus Annotation. In *Corpus Annotation* / Ed. by R. Garside, G. Leech, and A. McEnery. London, U.K.: Longman, 1997, pp. 1-18.
- Liddell S., Johnson R. E.** *American Sign Language: The Phonological Base*. Washington, D.C.: Ms. Gallaudet University, 1989.
- Perlmutter D. M.** Sonority and Syllable Structure in American Sign Language. *Linguistic inquiry*, 1992, no. 23(3), pp. 407-442.
- Rozelle L.** The structure of sign language lexicons: inventory and distribution of handshape and location: PhD Dissertation. Washington: University of Washington, 2003.
- Sandler W.** The Spreading Hand Autosegment of American Sign Language. *Sign Language Studies*, 1986, no. 50, pp. 1–28.
- Sandler W., Lillo-Martin D.** *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press, 2006.
- Stokoe W. C.** *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. Buffalo: Dept. of Anthropology and Linguistics, University of Buffalo. 1960.
- Vidalón Y. J. E., de Martino J. M.** Continuous Sign Recognition of Brazilian Sign Language in a Healthcare Setting. *Journal of Communication and Information Systems*, 2015, no. 30(1), pp. 82–89. DOI:10.14209/jcis.2015.10

References

- Brentari, D.** *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press, 1998.

- Burkova, S. I., Kimmelman, V. I., Filimonova, E. V., Kuseva, M. V., Varinova, O. A., Zavaritsky, D. A., Pristavko, K. V., Kadyrgulova, R. S.** Introduction to the Linguistics of Sign Languages. Russian Sign Language. Textbook. Novosibirsk: NGTU, 2019, 356 p. (in Russ.)
- Burkova, S. I., Varinova, O. A.** On Territorial and Social Variation in Russian Sign Language. In Russian Sign Language: The First Linguistic Conference / Ed. O. V. Fedorova. Moscow: Buki Vedi, 2012, pp. 127–143. (in Russ.)
- Camgöz, N. C., Kindiroğlu, A. A., Karabüklü, S., Kelepir, M., Özsoy, A. S., Akarun, L.** Bosphorus Sign: A Turkish Sign Language Recognition Corpus in Health and Finance Domains. *Proc. of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, Portoroz, Slovenia, 2016, pp. 1383–1388. URL: <https://aclanthology.org/L16-1220>
- De Souza, R. S., de Martino, J. M., Temoteo, J., Rodrigues, I.** Automatic Recognition of Continuous Signing of Brazilian Sign Language for Medical Interview. *Proc. of the Sixth International Conference on Informatics and Assistive Technologies for Health-Care, Medical Support and Wellbeing HEALTHINFO*, Barcelona, Spain, 2021, pp. 41–46. URL: https://www.researchgate.net/publication/355208551_Automatic_Recognition_of_Continuous_Signing_of_Brazilian_Sign_Language_for_Medical_Interview
- Drapkina, O. M.** Brief Algorithms of Management of Patients at the Stage of Providing Primary Healthcare. Manual for therapists; Ed. O. M. Drapkina. Moscow: Vidox, 2019, 20 p. (in Russ.)
- Geilman, I. F.** Russian Sign Language Dictionary, in 2 Volumes. St. Petersburg: Prana, 2004, 363 p. (in Russ.)
- Grenoble, L.** An overview of Russian Sign Language. *Sign Language Studies*, 1992, no. 77, pp. 321–338.
- Grif, M. G., Kugaevskikh, A. V.** Recognition of Deaf Gestures Based on a Bio-Inspired Neural Network. *Journal of Physics: Conference Series*, 2020, 1661 012038. DOI:10.1088/1742-6596/1661/1/012038
- Grif, M., Korolkova, O., Manueva, Y.** A New Algorithm and Other Software for Disambiguation of Polysemy and Homonymy for Computer Translation into Russian Sign Language Based on Semantic Principle. *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2018, vol. 16, no. 3, p. 32–44. (in Russ.)
- Kagirow, I. A., Ryumin, D. A., Axyonov, A. A., Karpov, A. A.** Multimedia Database of Russian Sign Language items in 3D. *Voprosy Jazykoznanija*, 2020, no. 1, pp. 104–123. (in Russ.) DOI: 10.31857/S0373658X0008302-1
- Kibrik, A. E.** The Methodology of Field Investigations in Linguistics (Setting up the Problem). Moscow: Moscow University Press, 1972. (in Russ.)
- Kimmelman, V.** Reflexive Pronouns in Russian Sign Language and Sign Language of the Netherlands: MA Thesis in Linguistics. Amsterdam: Universiteit van Amsterdam, 2009.
- Klezovich, A.** Automatic Extraction of Handshapes Inventory in Russian Sign Language. *Linguistics. WP BRP. NRU HSE*, 2019, no. 86. Higher School of Economics Research Paper No. WP BRP 86/LNG/2019.
- Korolkova, O. O.** Functional Aspect of the Russian Sign Language: Defining a Research Approach. *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2017, vol. 15, no. 3, p. 67–75. (in Russ.)
- Kosmopoulos, D., Oikonomidis, I., Constantinopoulos, C., Arvanitis, N., Antzakas, K., Bifis, A., Lydakakis, G., Roussos, A., Argyros, A.** Towards a Visual Sign Language Dataset for Home Care Services. *Proc. of the 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 2020, pp. 520–524. DOI: 10.1109/FG47880.2020.00099
- Leech, G.** Introducing Corpus Annotation. In Corpus Annotation / Ed. by R. Garside, G. Leech, and A. McEnery. London, U.K.: Longman, 1997, pp. 1–18.
- Liddell, S., Johnson, R. E.** American Sign Language: The Phonological Base. Washington, D.C.: Ms. Gallaudet University, 1989.

- Perlmutter, D. M.** Sonority and Syllable Structure in American Sign Language. *Linguistic inquiry*, 1992, no. 23(3), pp. 407-442.
- Rozelle, L.** The structure of sign language lexicons: inventory and distribution of handshape and location: PhD Dissertation. Washington: University of Washington, 2003.
- Sandler, W.** The Spreading Hand Autosegment of American Sign Language. *Sign Language Studies*, 1986, no. 50, pp. 1–28.
- Sandler, W., Lillo-Martin, D.** Sign Language and Linguistic Universals. Cambridge: Cambridge University Press, 2006.
- Stokoe, W. C.** Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. Buffalo: Dept. of Anthropology and Linguistics, University of Buffalo. 1960.
- Vidalón, Y. J. E., de Martino, J. M.** Continuous Sign Recognition of Brazilian Sign Language in a Healthcare Setting. *Journal of Communication and Information Systems*, 2015, no. 30(1), pp. 82–89. DOI:10.14209/jcis.2015.10
- Zaytseva, G. L.** Sign Speech. Dactilology: A Textbook for Higher School Students. Moscow: VLADOS, 2000, 192 p. (in Russ.)

Список источников

- Multimodal Medslset (Medical Sign Language Set) [Medslet]. URL: <https://ieee-dataport.org/open-access/display-multimodal-medslset-medical-sign-language-set> (дата обращения: 01.05.2022).
Корпус русского жестового языка [КРЖЯ]. URL: <http://rsl.nstu.ru/> (дата обращения: 01.05.2022).

List of Sources

- Multimodal Medslset (Medical Sign Language Set) [Medslet]. URL: <https://ieee-dataport.org/open-access/display-multimodal-medslset-medical-sign-language-set> (accessed: 01.05.2022).
Russian Sign language corpus. URL: <http://rsl.nstu.ru/> (accessed: 01.05.2022).

Сведения об авторах

- Кагиров Ильдар Амирович**, научный сотрудник лаборатории речевых и мультимодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН (СПб ФИЦ РАН)
- Рюмин Дмитрий Александрович**, кандидат технических наук, старший научный сотрудник лаборатории речевых и мультимодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН (СПб ФИЦ РАН)

Information about the Authors

- Kagirov Ildar Amirovich** – researcher at the Speech and Multimodal Interfaces Laboratory of the St. Petersburg Federal Research Center of RAS
- Ryumin Dmitry Alexandrovich**, PhD in Engineering, senior researcher at the Speech and Multimodal Interfaces Laboratory of the St. Petersburg Federal Research Center of RAS

*Статья поступила в редакцию 02.05.2022;
одобрена после рецензирования 10.07.2022; принята к публикации 27.07.2022
The article was submitted 02.05.2022; approved after reviewing 10.07.2022; accepted for publication 27.07.2022*