

Мультимедийная база данных жестов русского жестового языка в трехмерном формате

© 2020

Ильдар Амирович Кагиров[@]
Дмитрий Александрович Рюмин
Александр Александрович Аксёнов
Алексей Анатольевич Карпов

Санкт-Петербургский институт информатики и автоматизации РАН,
Санкт-Петербург, Россия; [@]kagirov@iiias.spb.su

Аннотация: Настоящая статья посвящена описанию мультимедийной базы данных (БД) “TheRuSLan” русского жестового языка (РЖЯ) и принципов ее фонетической аннотации. Особенностью заявленной БД является трехмерное (3D) представление жестов, что дает новые возможности лексикографического описания лексики и исследования особенностей артикуляции РЖЯ. Трехмерное представление жеста достигается за счет использования камеры Kinect 2.0, записывающей данные в высоком разрешении (FullHD), в режиме карты глубины и в инфракрасном диапазоне. БД представляет собой запись 164 лексических единиц и клауз в исполнении 13 информантов минимум в 5 итерациях. Общий размер файлов составляет 3,8 Тб в оригинальном формате, общая длительность видео — более 8 часов. В статье дается обзор существующих БД по РЖЯ, раскрывается содержание описываемого словаря, обрисовываются принципы лингвистической аннотации жестов и обосновывается набор признаков, используемых при аннотировании настоящей БД. Показано, что некоторые параметры жеста могут существенно варьировать от информанта к информанту, а именно: форма руки и локализация. Однако причины такой вариативности, ее масштабы и значение для описания РЖЯ являются делом дальнейших исследований.

Ключевые слова: базы данных, жестовые языки, компьютерная лингвистика, русский жестовый язык
Благодарности: Данное исследование проводится частично в рамках Соглашения с Минобрнауки России № 075-15-2019-1295 (идентификатор проекта RFMEFI61618X0095).

Для цитирования: 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: 104–123.

DOI: 10.31857/S0373658X0008302-1

Multimedia database of Russian Sign Language items in 3D

Ildar Kagirov[@]
Dmitry A. Ryumin
Alexander A. Axonov
Alexey A. Karpov

St. Petersburg Institute for Informatics and Automation of RAS, St. Petersburg, Russia;
[@]kagirov@iiias.spb.su

Abstract: The present paper describes a multimedia database (MMDB) of the Russian Sign Language (RSL) and briefly sketches the outlines of its phonological annotation. This MMDB is unusual thanks

to 3D format, which provides new opportunities for the lexicographical description of the vocabulary and enhances the study of RSL articulation features. The three-dimensional representation of the gesture is achieved by using a Kinect 2.0 camera that records data in high resolution (FullHD), in the depth map mode and in the infrared range. The presented DB comprises 164 lexical items and clauses, demonstrated by 13 signers and repeated 5 times at least. The total size of all the files is 3.8 Tb, the video duration is more than 8 hours. The article provides an overview of the existing RSL databases, reveals the content of the described dictionary, outlines principles of linguistic annotation of gestures, and gives an explanation the set of features used for annotation of the present MMDB. It is noted in the paper that some parameters of signs can vary significantly from informant to informant, e.g. handshape and localization. However, the reasons for such variability, as well as its scope and significance for describing RSL, are a matter of further research.

Keywords: computational linguistics, databases, Russian Sign Language, sign languages

Acknowledgements: The research is supported by the Ministry of Science and Higher Education of the Russian Federation, under the agreement No. 14.616.21.0095 (ID Number: RFMEFI61618X0095).

For citation: Kagirov I. A., Ryumin D. A., Axyonov A. A., Karpov A. A. Multimedia database of Russian Sign Language items in 3D. *Voprosy Jazykoznanija*, 2020, 1: 104–123.

DOI: 10.31857/S0373658X0008302-1

Введение

С 2018 г. в лаборатории речевых и многомодальных интерфейсов СПИИРАН ведутся междисциплинарные исследования по созданию многомодального человеко-машинного интерфейса, поддерживающего распознавание как звучащего русского языка¹, так и жестового языка. Предполагается, что подобным пользовательским интерфейсом можно будет оснастить роботизированные тележки² для супермаркетов и продуктовых магазинов, и именно это обстоятельство определяет предметную область словаря подобной роботоплатформы. Настоящая статья посвящена описанию электронной лексической базы данных³ русского жестового языка (РЖЯ) по теме «продукты питания», составляющей ядро жестового интерфейса роботизированной тележки; также предложены принципы для ее фонологической аннотации. База данных называется TheRuSLan (*Thesaurus Russian Sign Language*). Следует подчеркнуть, что работа по созданию базы данных, сбору материала, а также по разработке специального программного обеспечения для внедрения в жестовый пользовательский интерфейс еще не завершена. Тем не менее авторы исследования полагают, что получившаяся база данных уже сейчас может представлять определенный интерес с точки зрения компьютерной лингвистики и лингвистики жестовых языков и использоваться для решения ряда прикладных и теоретических задач, таких как исследования аллофонического варьирования или машинного распознавания РЖЯ.

¹ В данной работе речь пойдет исключительно о данных, с которыми работает модуль жестового языка; с принципами распознавания русской речи, применяемыми для создания интерфейса, можно ознакомиться, например, в наших работах [Кипяткова, Карпов 2017; Karpov et al. 2014], принципы обработки и распознавания изображений описаны в [Ryumin et al. 2019].

² Имеется в виду роботизированная платформа (тележка для покупок), которая могла бы по запросу пользователя провести его в нужный отдел, следовать за ним и дать справку по ассортименту товаров магазина, см. [Кагиров 2018].

³ Авторы статьи хотели бы искренне поблагодарить сотрудников и студентов Межрегионального центра реабилитации лиц с проблемами слуха города Павловска, любезно согласившихся принять участие в записи БД. Также авторы отдельно выражают свою признательность Кристине Моисеенковой, Валентине Тупикиной и Елене Голубевой за проделанную организационную работу и помощь во время сессий. Все фотоматериалы публикуются с согласия информантов.

Статья построена следующим образом: во введении приводятся сведения о РЖЯ и дается краткий обзор лингвистических работ по соответствующей тематике. В следующем разделе перечислены основные базы данных по РЖЯ, существующие на сегодня. Далее описывается база данных, являющаяся предметом настоящей статьи: излагаются сведения об информантах, лексике, принцип ее отбора для словника, а также раскрываются технические характеристики БД: формат и способ записи, показан интерфейс пользователя и режимы съемки (в том числе и 3D). Заключительный раздел статьи посвящен фонологической разметке жестов, вошедших в БД: общим принципам фонологического уровня аннотации жестовых языков, особенностям артикуляции⁴ записанной лексики РЖЯ и признакам, на основе которых была проведена аннотация представленного материала.

Русский жестовый язык (РЖЯ) — это язык общения глухих и слабослышащих на территории Российской Федерации и ряда стран — бывших советских республик. Согласно данным международного каталога Ethnologue (www.ethnologue.com), число носителей РЖЯ на 2010 г. составляло 122 тыс. человек, при этом 99 % из них проживает на территории РФ [RSL; Карпов 2011]. По некоторым данным, число людей, пользующихся РЖЯ, достигает двух миллионов [Воскресенский 2002]. В 2012 г. РЖЯ был признан языком общения в РФ «при наличии нарушений слуха и (или) речи, в том числе в сферах устного использования государственного языка Российской Федерации» [Правовые основы 2016: 43]. Следует различать жестовый язык как самостоятельную систему коммуникации и спонтанную жестикуляцию, сопровождающую устные коммуникативные акты [Федорова 2018].

Научных работ по диалектной вариативности РЖЯ немного. Практически всеми исследователями отмечается, что существуют определенные различия фонологического (т. е. касающегося особенностей жестикуляции) и лексического характера между идиомами РЖЯ, что, в принципе, позволяет предполагать наличие диалектов. Однако в [Буркова, Варинова 2012] показано, что реальный уровень лексических различий в территориальных разновидностях РЖЯ значительно ниже, чем принято полагать. Исследований же по грамматическим различиям в диалектах РЖЯ практически не существует. В качестве исключения можно назвать работу [Kimmelman 2009], в которой, в частности, затрагиваются вопросы вариативности возвратных местоимений в диалектах РЖЯ.

Помимо РЖЯ, выделяется еще так называемая «калькирующая жестовая речь» (КЖР) [Зайцева 2000], принципиально отличная от естественной системы жестового общения — РЖЯ. КЖР копирует линейные структуры русского языка, являясь, по сути, жестовыми субтитрами звучащей речи. «Калькирующая жестовая речь — вторичная знаковая система, которая усваивается в процессе изучения глухим ребенком словесной речи. Жесты здесь являются эквивалентами слов, а порядок их следования такой же, как в обычном предложении» [Там же: 14]. Именно КЖР используется для официального перевода звучащего русского языка, при обучении глухих и слабослышащих, и большинство носителей РЖЯ рассматривает эту форму жестового общения как престижную [Grenoble 1992], однако редко пользуются ей в повседневном общении в чистом виде.

Дактильная азбука является жестовым представлением букв русского алфавита и используется во вспомогательных целях: для передачи имен собственных, служебных слов, для уточнения, а также в том случае, если не существует общепринятого жеста для какого-либо понятия [Зайцева 2000; Карпов 2013].

Несмотря на вторичность калькирующей речи и дактильной азбуки, обе системы в той или иной степени присутствуют при общении на РЖЯ. Пропорции каждой из трех систем

⁴ Термин «артикуляция» по отношению к мануальным жестам в настоящей статье употребляется вслед за [Sandler 2012]; следует помнить, что в контексте изучения жестовых языков также выделяют артикуляцию как «движения рта, напоминающие произношение слова или его части, при произведении мануальных жестов» [Bauer 2018: 35]. Далее термин «артикуляция» используется исключительно в первом значении.

определяются множеством факторов: степенью образованности коммуникантов, коммуникативной ситуацией и т. п.

Собственно РЖЯ, несмотря на наличие большого количества коррелятов в сфере лексики [Зайцева 2000: 33 и сл.], принципиально отличен от КЖР (и, разумеется, от звучащего русского языка) в сфере грамматики. Одной из особенностей РЖЯ — как, впрочем, и других жестовых языков — является использование пространственных и визуальных средств, недоступных звучащим языкам. С основными особенностями грамматики РЖЯ можно ознакомиться, например, в работах [Зайцева 2000; Давиденко, Комарова 2006; Киммельман 2010; Львовская 2006; Прозорова 2007; Шамаро 2006; Буркова 2012–2015; Kimmelman 2014; Заварицкий 2015: 8–34]. Утверждение о том, что РЖЯ остается малоисследованным языком (в отличие от других жестовых языков, в первую очередь, американского жестового языка), не будет большим преувеличением: несмотря на то что в последнее время наметился определенный интерес как к изучению РЖЯ [Федорова 2012], так и к системам автоматического перевода с РЖЯ [Мануева и др. 2014; Karpov et al. 2016; Сысоев, Гриф 2018], в целом работ по РЖЯ пока что недостаточно.

1. Обзор существующих электронных баз данных РЖЯ

Как было замечено выше, описание РЖЯ по-прежнему находится на начальной стадии. То же утверждение окажется справедливым и в отношении корпусов и баз данных по РЖЯ. В то время как материал по другим жестовым языкам находится в сфере пристального внимания корпусной лингвистики (см., например, материалы проекта [SLCN]), то едва ли не единственным полноценным корпусом РЖЯ является «Корпус русского жестового языка», созданный Новосибирским государственным техническим университетом (НГТУ) [КРЖЯ]. Корпус содержит свыше 230 текстов от 43 носителей РЖЯ, аннотирован в среде ELAN [Language Archive] и снабжен фильтрами по поиску метаданных. В целом этот корпус представляет собой эффективный инструмент для изучения РЖЯ. Стоит отметить, однако, что большая часть информантов, принявших участие в записи корпуса, является выходцами из Новосибирска и Москвы, что ставит под вопрос применимость этого корпуса для исследований, например, петербургских идиом. На рис. 1 представлен графический пользовательский интерфейс корпуса НГТУ, демонстрирующий разметку, перевод и видеоданные.

Кроме корпуса НГТУ, существуют и другие электронные базы данных РЖЯ, например:

- 1) «Толковый словарь русского жестового языка — RuSLED». Представляет собой интерактивную оболочку к видеословарю РЖЯ, записанному в 2002 г. в Межрегиональном центре реабилитации лиц с проблемами слуха (МЦР) города Павловска [Воскресенский и др. 2009] (на рис. 2 приведен пример графического интерфейса этого словаря);
- 2) видеословарь «Тематический словарь русского жестового языка», созданный на основе словаря [Фрадкина 2001];
- 3) электронная обучающая система «Русский жестовый язык. Базовый курс» [РЖЯБК 2001];
- 4) интерактивный видеословарь «DigitGestus», созданный в 1996–1997 гг. (Новосибирск) [Digitgestus];
- 5) интерактивный видеословарь жестового языка европейского проекта Spreadthesign («Распространен жест») [Spreadthesign];
- 6) интерактивный видеословарь проекта «Сурдосервер 2.0» [Сурдосервер], представляющий собой расширенную версию словаря RuSLED.



Рис. 1. Пример пользовательского интерфейса корпуса РЖЯ, демонстрирующий аннотацию и перевод [КРЖЯ]

Все вышеперечисленные БД (кроме Корпуса РЖЯ от НГТУ) представляют собой видеословари и/или учебники, в которых лексика сгруппирована по темам. Во всех случаях реализована возможность поиска, однако разметка как таковая отсутствует. Использование любого из этих пособий (равно как и Корпуса НГТУ) для автоматического распознавания жеста требует предварительной работы по обработке и аннотации данных.

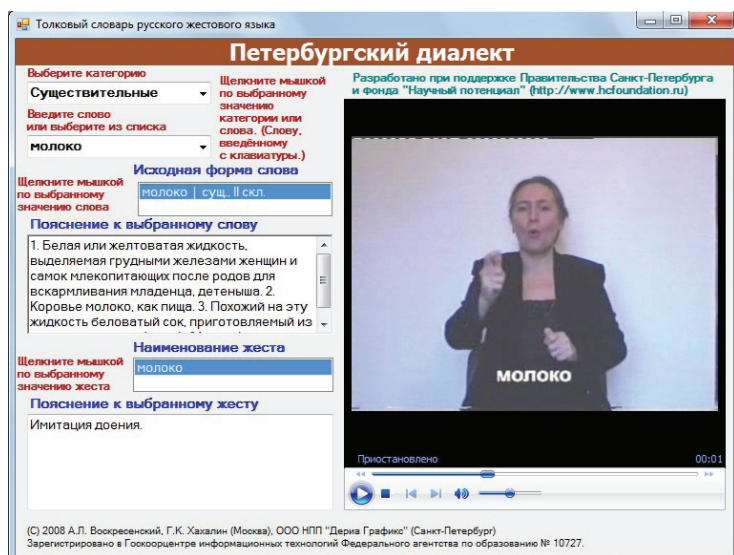


Рис. 2. Пример графического интерфейса словаря RuSLED [Воскресенский и др. 2009]

2. Собранная база данных РЖЯ по теме «продукты»

2.1. Лексический состав базы данных РЖЯ

Представленная в данной статье БД пока не является полноценным корпусом. Под корпусом текстов в современной лингвистике понимается «большой, представленный в электронном виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач» [Захаров 2005: 3]. Кроме того, сам термин «корпус текстов» подразумевает, что основной единицей БД является текст, а не лексема. Именно поэтому — ввиду отсутствия грамматической и лексической разметки, протяженных текстов и большого объема — корректным наименованием заявленного массива языковых данных в том виде, в котором он существует сейчас, будет «мультимедийная лексическая база данных», а не «корпус».

На настоящий момент БД представляет собой запись 164 лексических единиц и клауз в исполнении 13 информантов минимум в пяти итерациях. Общий размер файлов составляет 3,8 Тб в оригинальном формате, общая длительность видео — более 8 часов. Многократное повторение каждого стимула необходимо для обучения статистических моделей, способных к распознаванию жестов; в то же время многократное воспроизведение жестов создает благоприятные условия для анализа вариативности и разработки принципов аннотирования жестов. 10 информантов являются студентами МЦР (двое мужчин и девять женщин), их средний возраст составляет 20 лет, еще трое информантов — преподаватели МЦР от 30 до 40 лет. Все студенты являются глухими или слабослышащими. Поскольку практически все дикторы⁵ — выходцы из разных регионов России, во время «пилотных» записей предсказуемо наблюдалась определенная вариативность в показе жестов: «...сегодня некоторые жесты в одной школе-интернате, где дети проводят большую часть времени, могут отличаться от жестов в другой, так как ученики одной школы гораздо чаще и более тесно общаются внутри данной группы. Похожая ситуация складывается и в специализированных высших учебных заведениях для глухих. (...) Таким образом, из десятков близких регионов в одно учебное заведение съезжаются подростки, которые используют “школьные” и “семейные” жесты» [Филимонова, Душкина 2018: 11]. В связи с этим все переводы были предварительно стандартизированы специалистами из МЦР с опорой на словарь [Гейльман 2004]. Применение этого издания показалось авторам статьи удобным потому, что включенная в него лексика понятна большинству образованных носителей РЖЯ. Все жесты сопровождаются устной артикуляцией («проговариваются»), однако запись аудиоданных не велась.

Основная тематика лексики БД, как отмечалось выше, посвящена продуктам питания в супермаркете и может быть разбита на разделы «продукты», «отделы», «перемещение в пространстве», «поиск товара в магазине». Первичный список формировался экспортом текстовых файлов из меню навигации интернет-сайтов ряда крупных гипермаркетов Санкт-Петербурга. Формирование конечного словника проводилось путем отсеивания единиц, содержащих конкретные наименования (фирменное название продукта, производитель, состав). Кроме того, в конечный список не вошли товары, не пользующиеся, по личным ощущениям авторов данной работы, большой популярностью у потребителя (например, *тигровые креветки* или *помело*). Наконец, из словника были исключены лексические единицы, для которых в РЖЯ применяется дактилирование в силу отсутствия общепринятых жестов (например, *фасоль*, *парное мясо*, *чебуреки*). Одной из причин, побудивших

⁵ В англоязычной литературе используется термин «signer», для которого нет устоявшегося перевода на русский язык; в настоящей статье выбран термин «диктор».

авторов пойти по пути сокращения объема лексики, было стремление экономии вычислительных ресурсов, необходимых для хранения и обработки массива данных в контексте задачи машинного распознавания жеста.

Словарь команд включает в себя лексику, связанную с ориентацией (*вперед, назад, направо, налево*), ряд частотных глаголов перемещения (*поехали, пойдём*) и их модификаторы (*быстро, медленнее*), запросы местоположения (*где?, покажи мне*) и конкретных товаров (*мне нужен..., я хочу...*). В словарь отделов входят названия отделов магазина; фактически это группы товаров, объединённых по типу сырья, из которого они произведены, или по их назначению. На практике такие группы товаров часто выносятся в отдельные локации. Кроме того, в этот же список включены такие лексемы, как *касса, выход, туалет* и т. п. 85 % записей относятся к категории «продукты питания и прочие товары», остальные 15 % записей представлены командами пользователя, названием отделов, вопросами и ответами. Некоторые лексемы являются композитами, т. е. состоят из двух (и более) составляющих, способных функционировать как самостоятельные лексемы, например:

КЕФИР = КИСЛЫЙ.МОЛОКО

ВИНО = СЛАДКИЙ.АЛКОГОЛЬ

МЁД = ПЧЕЛА.СЛАДКИЙ

Все составляющие дополнительно вносятся в базу данных как отдельные единицы. Всего в нашей БД насчитывается 48 различных одноручных жестов и 116 двуручных. В дальнейшем речь пойдет преимущественно об одноручных жестах⁶.

2.2. Технические характеристики собранной базы данных

Запись проводилась при помощи разработанной авторами автоматической системы записи 3D видеопотока с использованием сенсора-дальномера Microsoft Kinect 2.0⁷. Стимулы перевода выводились на экран компьютера и демонстрировались информантам по пять раз минимум в целях дальнейшего обучения систем автоматического распознавания элементов жестового языка на основе вероятностных нейросетевых моделей. Камера Kinect 2.0 была установлена на расстоянии 1,5–2 метров от диктора, при этом дикторы записывались на однородном светлом фоне.

Отличительной особенностью представленной базы данных является запись жестов в трехмерном (3D) формате, что делает ее единственным в своем роде ресурсом по материалу РЖЯ. В самом деле, даже наиболее обширные базы данных из перечисленных выше представляют жест в виде двухмерного изображения. Применение карты глубины вводит в описание третье измерение, что позволяет более точно определять положение одного объекта относительно другого, в данном случае — рук относительно друг друга и корпуса говорящего. Расстояние между активной и пассивной рукой, удаленность рук от тела является средством выражения разнообразных лексических значений в жестовых языках мира. Формат 3D получен за счет того, что камера Kinect 2.0 имеет возможность записывать не только видеоданные в оптическом FullHD формате и инфракрасном диапазоне, но и в режиме «карты глубины» (рис. 3).

⁶ Машинное распознавание изображения является нетривиальной задачей, и отслеживание двух артикуляторов (рук) серьезно повышает объем работы (разметка данных или снижение размерности, извлечение признаков). На момент написания статьи задача обработки двуручных жестов не была решена в полном объеме, поэтому авторы решили ограничиться одноручными жестами. Будет ли подобное описание актуальным и для двуручных жестов — задача дальнейших исследований.

⁷ Возможно и использование других контроллеров, таких как LIPSedge® DL или Azure Kinect.

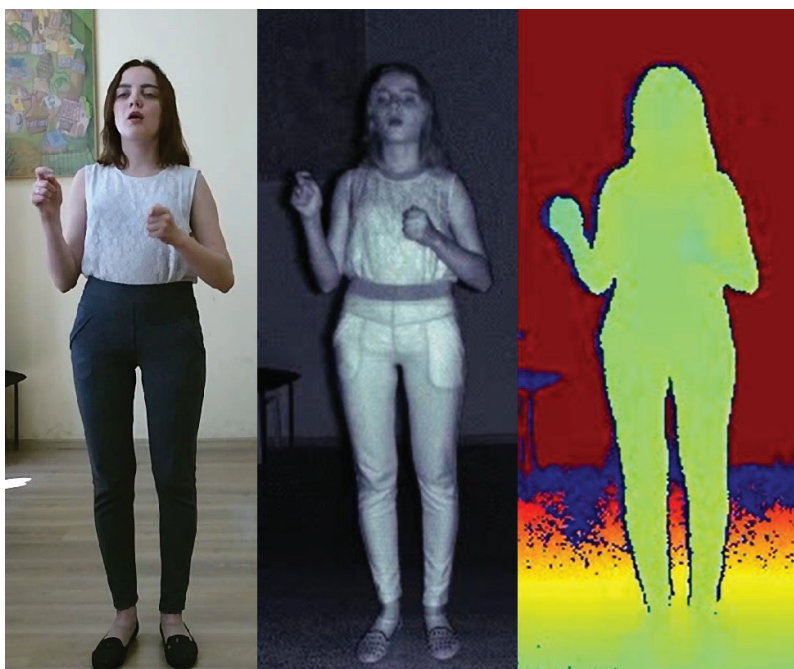


Рис. 3. Примеры записи жестов в цветном формате (слева), в инфракрасном диапазоне (в центре) и в режиме карты глубины (справа)

Цветовые обозначения на карте глубины соответствуют диапазонам спектра видимого излучения, иными словами, наиболее удаленные от камеры группы точек окрашены в красный цвет, наиболее приближенные — в фиолетовый. Объекты, находящиеся между этими точками, закрашиваются оттенками зеленого и желтого. Синяя область на рис. 3 соответствует «слепой зоне» камеры: из-за особенностей пространства и освещения (удаленный угол между полом и стеной) сенсор не смог определить глубину этого участка. Режим съемки в инфракрасном диапазоне позволяет проводить записи в затемненных или, наоборот, слишком освещенных местах.

Полученные в результате записи данные имеют следующие характеристики:

- 1) бинарные видеофайлы, содержащие запись цветной камеры в формате bin без сжатия (с оптическим разрешением 1920x1080 пикселей при 30 кадрах в секунду, цветность — 8 бит на пиксель);
- 2) бинарные видеофайлы, содержащие запись карты глубины и инфракрасной камеры в формате bin без сжатия (с оптическим разрешением 512x424 пикселей при 30 кадрах в секунду, цветность — 16 бит на пиксель);
- 3) текстовые файлы в формате XML с координатами скелетной модели диктора-носителя, разделенной на 25 суставов (элементов); каждая определенная точка — это пересечение двух осей (x, y) на координатной плоскости и дополнительное значение координаты с двойной точностью, обозначающее глубину точки, которая измеряется расстоянием от сенсора до точки объекта в диапазоне от 0 (1,2 метра) до 1 (3,5 метра).

Логическая структура базы данных может быть представлена в виде дерева каталогов, содержащих информацию для каждого показанного отдельным диктором жеста: а) видеозаписи показанного жеста в FullHD формате, в инфракрасном диапазоне и в формате

карты глубин; б) данные о положении скелета на видеозаписи; в) изображения в формате jpg, выделенные покادрово из видеозаписей и необходимые для разметки.

На рис. 4 представлен общий вид графического пользовательского интерфейса специализированного программного обеспечения для записи жестов. Нижняя часть интерфейса адресована диктору, правая панель и экран с видеоизображением предназначены для записывающего. Внизу отображаются стимул перевода и индикатор записи (зеленый цвет сигнализирует о том, что запись идет, красный — запись приостановлена). Оператор последовательно выводит на экран список стимулов для перевода (словник) и включает / выключает режим записи; цвет индикатора указывает информанту временные интервалы для демонстрации соответствующих переводных эквивалентов на жестовом языке.

Панель в правой части экрана предназначена для изменения настроек записи: разрешения, количества кадров в секунду и т. п. «Скелет» на теле информанта построен на основе 25 точек (узлов), автоматически определяющихся сенсором Kinect 2.0 и представляющих собой пересечение двух осей (X, Y) в плоскости артикуляции. Эти координаты необходимы для вычисления глубины точки, которая определяется как расстояние от сенсора до самой точки. Визуализация точек в виде скелета необходима для быстрой оценки работоспособности программы записи жестов и сенсора Kinect 2.0.

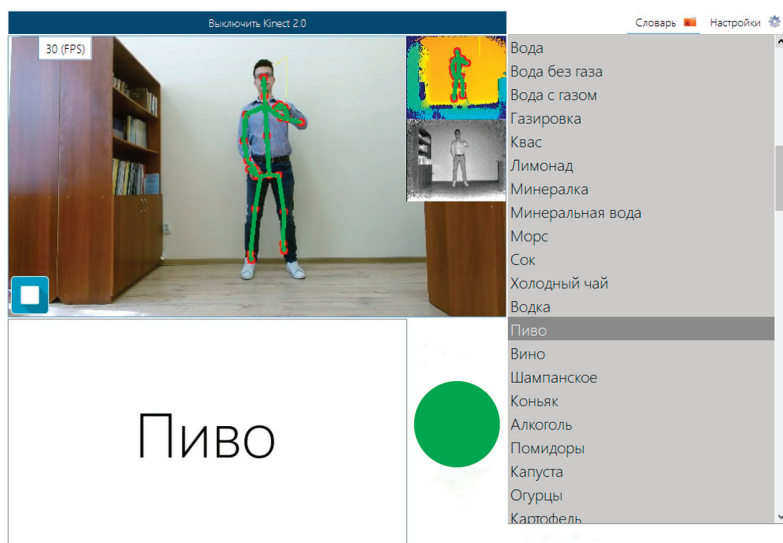


Рис. 4. Графический пользовательский интерфейс программы записи жестов

3. Аннотация жестовой БД

3.1. Фонологическая аннотация жестовых языков: общие принципы

Аннотация или разметка любого массива языковых данных — это способ представления некоторой информации о данных, выполненный при помощи метаязыка. В принципе, аннотаций одной и той же БД может быть бесконечно много, их количество будет

ограничено только количеством языковых уровней и элементов, выделяемых создателями разметки [Ляшевская 2016: 14].

Мультимедийная БД РЖЯ, описываемая в этой статье, создавалась как словарь РЖЯ, предназначенный для автоматического распознавания коротких, в одно-два слова, команд пользователя на жестовом языке. Именно по этой причине БД снабжена только фонологической⁸ разметкой. Во-первых, отсутствие пространственных текстов приводит к практически полному отсутствию синтаксических и морфологических явлений в БД, во-вторых, анализ элементов, из которых складывается жест, представляется весьма актуальным именно в контексте машинного распознавания жестовых языков. Обязательным условием аннотации должна быть ее пригодность для задач компьютерного зрения, то есть она должна быть ориентирована не только на специалистов, но и на компьютер. Иными словами, система аннотирования фонетического уровня БД должна основываться на признаках, которые можно было бы использовать не только для собственно лингвистической нотации, но и для компьютерного анализа и распознавания.

Особенностью предлагаемой системы аннотации является то, что параметры, заложенные в основу собственно фонологической интерпретации жеста, находят реализацию и в тех классах, которые выделяются для построения вероятностных моделей.

В работе [Stokoe 1960] была впервые предложена декомпозиция жеста на три составляющих: 1) форма и ориентация руки; 2) место выполнения жеста; 3) характер движения. Часто этот список расширяется за счет выделения ориентации руки в отдельный параметр и добавления мимического признака. Конкретные формы и ориентации руки, место и характер выполнения являются минимальными элементами жеста, не имея собственного значения, но обладая смыслоразличительной функцией. Набор реализаций этих признаков конечен (например, в [Battison 1978] выделяется 45 форм руки, 25 локализаций и 12 типов выполнения жеста в американском ЖЯ). Форма руки задается «выбранными» пальцами (т. е. пальцами, принимающими активное участие в конфигурации) и их положением (так, пальцы могут быть согнуты, выпрямлены, разведены и т. п.). Если форму руки, ориентацию и движение можно сравнить с фонемами звучащих языков, то выбранные пальцы и операции с ними логично сравнивать с дифференциальными признаками фонем.

Впоследствии анализ Стоуки был расширен в духе генеративной фонологии [Chomsky, Halle 1968]. Если модель Стоуки можно отнести к классификационным описаниям, то под влиянием генеративистов свет увидели «динамические» системы, позволяющие объяснить, как функционирует фонология конкретных жестовых языков. Так, в [Liddell, Johnson 1983] жест представлен как последовательность сегментов движения ('movement') и остановок ('hold'). Эта модель известна как Hold-Movement. В [Sandler 1986] конфигурация рассматривается как отдельный автосегментный параметр, т. е. принадлежащий самостоятельному уровню и ставящийся в соответствие сегментным единицам. Кроме того, были предложены подходы, оперирующие структурой слога в жестовых языках [Perlmutter 1992] и геометрией признаков [Brentari 1998]. Хороший обзор актуального положения вещей представлен в [Sandler, Lillo-Martin 2006].

Тем не менее принципы записи жеста, предложенные Стоуки, оказываются удобными в том случае, если мы имеем дело с задачей описания отдельных лексем, продиктованных изолированно, а не вычленившихся из потока речи. Одной из систем транскрибирования, основанной на нотации Стоуки, является Гамбургская система нотации, или HamNoSys [Prillwitz et al. 1989; Hanke 2004]. HamNoSys хорошо подходит для использования в электронных БД благодаря линейной структуре, проработанности и легкости конвертации в формат Unicode. Существуют работы по интеграции HamNoSys в общую систему аннотации корпусов жестовых языков [Hanke, Storz 2008]; также эта система нотации была применена

⁸ В жестовых языках по аналогии со звучащими языками выделяется фонологический уровень; фонология жестовых языков занимается анализом элементов, из которых состоит жест [Sandler, Lillo-Martin 2006: 113–119].

для аннотирования базы данных австралийского жестового языка [Johnston 2001], создания словарей новозеландского жестового языка [Kennedy et al. 1997] и немецкого жестового языка [AF 1996]. Другая система записи жестовых языков SignWriting [Sutton 1995; Thiessen 2011] достаточно сложна для использования в целях аннотирования из-за вертикальной ориентации строк и большей «идеографичности»: символы SignWriting скомпонованы так, что они отражают синхронность передачи нескольких потоков информации в высказываниях на жестовом языке. Если SignWriting удобен для повседневного бытового применения и по ряду критериев представляет собой систему письма [van der Hulst, Channon 2010], то HamNoSys несравненно более удобен для лингвистических аннотаций. Обзоры систем нотации для жестовых языков см., например, в [Карпов, Кагиров 2011: 130 и сл.; Frischberg et al. 2012: 1045–1054; Garcia, Sallandre 2013]

Насколько известно авторам данной статьи, систематического описания лексики РЖЯ с точки зрения артикуляции не существует. В большинстве пособий по РЖЯ и калькирующей речи, таких как [Жесты 1995; Зайцева 2000; Гейльман 2001; 2004; Фрадкина 2001; Заварицкий 2015], жесты представлены фотографиями и рисунками с комментариями общего характера. Например, в [Гейльман 2004, 1: 322] даны следующие пояснения: «Показывается образный жест ДООИТЬ, имитирующий это действие; артикулируется данное слово» (жест молоко); «кисть правой руки, сложенная как при показе буквы Ч, только обращенная ладонью влево, подносится стороной большого и указательного пальцев ко рту и чуть-чуть опускается; движение повторяется» (жест чай) [Там же, 2: 335] и т. п.

Особняком в этом списке стоит пособие Л. С. Димскис [2002], в котором предлагается нотация, основанная на принципах, сформулированных Стоуки. Вероятно, это единственная в своем роде попытка аннотирования лексики РЖЯ и калькирующей жестовой речи с четкой фиксацией таких признаков, как конфигурация руки, ориентация, локализация, активная рука и характер движения. Единственным недостатком пособия является ограниченность списка аннотированной лексики.

3.2. Фонологические характеристики жестов из БД: case studies

Любое различие в артикуляции снижает качество автоматического распознавания жеста; кроме того, несмотря на наличие карты глубины, компьютер работает только с фронтальной двумерной проекцией руки, не имея информации о полной конфигурации кисти. Демонстрация жестов информантами зачастую разнится в деталях: немного отличаются конфигурация кисти, локализация, однако характер выполнения жеста, т. е. движение, чаще всего остается неизменным. Поэтому нотация, позволяющая осуществлять поиск по корпусу и распознавание жеста компьютером, должна учитывать не все гипотетически возможные характеристики жеста, а те и только те, что позволяют отличить один жест от другого.

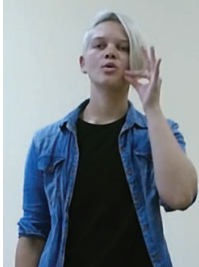
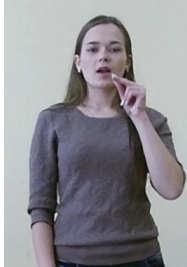
Артикуляция одной и той же лексической единицы у разных дикторов может различаться, например, по следующим параметрам (табл. 1, с. 115).

Из табл. 1 видно следующее.

1. Различия в конфигурации: все случаи из этой группы могут быть описаны как различия в положении пальцев: как выбранных, так и невыбранных. Выбранные пальцы могут быть загнуты в разной степени, а невыбранные пальцы могут принимать, теоретически, самую разную конфигурацию.
2. Различия в ориентации кисти: угол, формируемый кистью и одной из плоскостей артикуляции, может существенно отличаться.
3. Различия в локализации: в рамках подробной локализации, предлагаемой, в частности, в описаниях HamNoSys, единообразное описание одной и той же лексемы для

Таблица 1

Вариативность параметров жеста у разных информантов⁹

1. конфигурация	2. ориентация	3. локализация	4. контакт
			
(КИСЛЫЙ) $\hookrightarrow$	(СЛАДКИЙ) $\hookrightarrow \neg \circ$	(ПЕЧЕНЬЕ) $\hookrightarrow \overline{2} \ 3 \ }$	(ПЧЕЛА) $\hookrightarrow^x \ }$
			
(КИСЛЫЙ) $\hookrightarrow \overline{3} \ 4 \ 5$	(СЛАДКИЙ) $\hookrightarrow \wedge 0$	(ПЕЧЕНЬЕ) $\hookrightarrow \overline{2} \ 3 \ \cup \square$	(ПЧЕЛА) $\hookrightarrow \cup \cup$
5. амплитуда			
			
МОЛОКО $\ddot{\cdot} \hookrightarrow [\cup \neg \cup] \downarrow \uparrow +$		МОЛОКО $\ddot{\cdot} \hookrightarrow [\cup \neg \cup] \downarrow \uparrow +$	

разных информантов чаще всего невозможно, см. различное положение кисти для колонки 3.

- Частным случаем локализации является контакт активного органа артикуляции с пассивным. Колонка 4 показывает, что значения этого параметра могут отличаться.
- Различия в амплитуде движения: если сама по себе траектория движения является достаточно стабильным параметром, то амплитуда движения предсказуемо разнится от информанта к информанту.

Перечисленные выше различия в артикуляции зафиксированы только для изолированных лексем. Специальных исследований по вариативности жеста в потоке речи для записанной БД не проводилось. Все перечисленные отклонения от «словарной формы»

⁹ Транскрипции жестов здесь и далее даны в соответствии с нотацией HamNoSys 4.0.

не мешают восприятию текстов на РЖЯ человеком-адресатом, однако, очевидно, составляют проблему при машинной обработке. Представляется, что эта задача может быть решена двумя способами: максимальным расширением инвентаря визуальных образов, являющихся проявлением конкретного параметра, и, наоборот, подбором таких параметров, которые, с одной стороны, были бы подвержены минимальному изменению в зависимости от особенностей артикуляции носителя РЖЯ и, с другой стороны, все же позволяли бы с уверенностью отличать один жест от другого. Первый способ требует больших вычислительных ресурсов и поэтому кажется менее предпочтительным по сравнению со вторым.

3.3. Система признаков конфигурации и локализации

При описании жестовых языков мира выделяются десятки конфигураций пальцев и типов локализации. Для настоящего словаря далеко не все они оказываются актуальными. Ниже будут приведены только те конфигурации и локализации, которые актуальны для аннотации словаря. Для описания конфигурации руки было выделено семь основных форм кисти; остальные формы могут быть описаны как их модификации за счет комбинаций пальцев и дополнительной конфигурации. Можно выделить следующие базовые конфигурации кисти (рис. 5¹⁰).

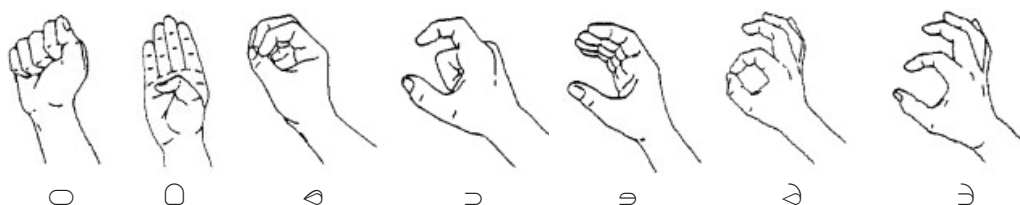


Рис. 5. Базовые конфигурации кисти

Форма □ (ладонь) может быть модифицирована за счет применения следующих операций (рис. 6).

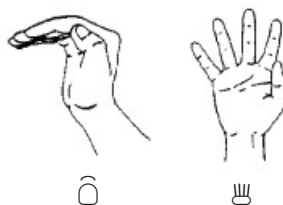


Рис. 6. Операции с ладонью

В первом случае ладонь полусогнута ($\hat{\square}$), во втором — пальцы разведены. Трактовка последнего случая как ладони, а не комбинации из четырех пальцев выбрана потому, что не всегда можно с точностью определить, демонстрирует ли диктор ладонь или просто недостаточно плотно сводит пальцы.

¹⁰ Изображения конфигураций кисти на рис. 5, 6, 7 взяты из документации HamNoSys 4: https://www.sign-lang.uni-hamburg.de/dgs-korpus/files/inhalt_pdf/HamNoSys_Handshapes.pdf.

Для базовых форм $\square$ и $\square$ возможна дальнейшая модификация за счет операций с пальцами (рис. 7).

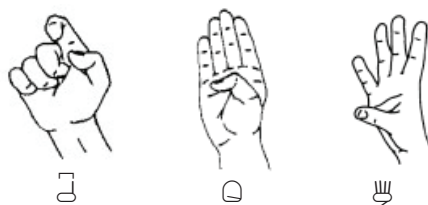


Рис. 7. Некоторые операции с пальцами

В первом случае палец согнут, во втором — большой палец прижат к ладони («спрятан»), в третьем — палец отведен в сторону перпендикулярно плоскости ладони.

Отдельной операцией над пальцами выступает контакт. Во всех случаях активным пальцем является большой, а палец или пальцы, с которыми он контактирует, являются пассивными. Нотация требует указания фаланги и номера пассивного пальца, контактирующего с большим. Всего было получено 24 конфигурации; для наглядности они сведены в табл. 2, определяющую каждую форму кисти по параметрам базы и операций, которым она подвергается.

Таблица 2

Признаки конфигураций кисти руки

ВЫБРАННЫЙ ПАЛЕЦ	операция с выбранным пальцем	БАЗА								
		$\square$	$\square$			$\square$	$\square$	$\square$	$\square$	$\square$
			$\square$	$\square$	$\square$					
нет		+	+	+	+	+	+	+	+	+
$\square^1$	без операции	+	—	—	—	—	—	—	—	—
	$\square$ (палец спрятан)	—	+	—	—	—	—	—	—	—
	$\square$ (палец согнут)	+	—	—	—	—	—	—	—	—
	$\square$ (палец отставлен)	—	—	—	+	—	—	—	—	—
контакт (с указанием пассивных пальцев)	$\chi \square_3$	—	+	—	—	—	—	—	—	—
	$\chi \square_5$	—	+	—	—	—	—	—	—	—
	$\chi \square_{2\ 3}$	+	—	—	—	—	—	—	—	—
$\square^2$	без операции	+	—	—	—	—	—	—	—	—
	$\square$ (палец полусогнут)	+	—	—	—	—	—	—	—	—
	$\square$ (палец согнут)	+	—	—	—	—	—	—	—	—
$\square^3$		—	+	—	—	—	—	—	—	—
$\square^{2\ 3}$	без операции	+	—	—	—	—	—	—	—	—
	$\square$ (пальцы разведены)	+	—	—	—	—	—	—	—	—
	$\square$ (палец согнут)	+	—	—	—	—	—	—	—	—
$\square^{2\ 5}$		+	—	—	—	—	—	—	—	—

Поскольку отслеживание положения и формы каждого пальца на руке является нетривиальной технической задачей, все выделенные 24 формы при распознавании трактуются как простейшие. Каждая конфигурация кисти представлена в нескольких проекциях и разных вариантах исполнения; все проекции и варианты соотнесены с конкретной конфигурацией из табл. 2.

В компьютерных системах локализация и идентификация формы кисти человека может быть осуществлена при помощи платформы с открытым исходным кодом TensorFlow Object Detection API [Huang et al. 2017], которая упрощает создание, обучение и развертывание моделей обнаружения всевозможных объектов. Фактически в базе данных хранятся изображения выделенных нами конфигураций кисти в различных проекциях (рис. 8), с разной степенью отличий в расположении пальцев.



Рис. 8. Различные проекции и конфигурации 2 3 4

Следующий параметр — локализация — отличается нестабильностью. Как было указано выше, локализация жеста у разных информантов выглядит по-разному. Решением этой проблемы, с одной стороны, является «укрупнение» зон Y-плоскости артикуляции. Путем сравнения локализации жестов у разных дикторов были определены границы зон на лице и теле человека, позволяющие снять неоднозначность и непротиворечиво описать все имеющиеся жесты. Интересно, что не все полученные зоны можно описать символами HamNoSys. Например, в оригинальной нотации отсутствуют обозначения для нижней челюсти и подбородка. Ниже приведена карта зон с легендой (рис. 9, табл. 3).

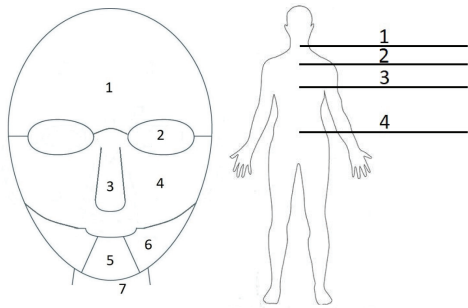


Рис. 9. Основные зоны на плоскости артикуляции Y, выделяемые при аннотации жестов

Таблица 3

Зоны артикуляции одноручных жестов БД

ГОЛОВА		ТЕЛО	
номер зоны	пояснение	номер зоны	пояснение
1	выше уровня глаз	1	голова
2	глаза	2	плечи
3	нос	3	грудь
4	рот и щеки	4	живот
5	подбородок		
6	челюсть		
7	шея		

С другой стороны, подавляющее большинство жестов на практике не обладает абсолютной локализацией; если укрупнение зон артикуляции способно решить проблему локализации по оси X, то точное определение места выполнения жеста по оси Y, по-видимому, невозможно. Так, жест кислый некоторые информанты начинают у подбородка и завершают у переносицы, другие начинают у рта и заканчивают в районе носа; еще в ряде случаев движение завершается выше уровня глаз (рис. 10).

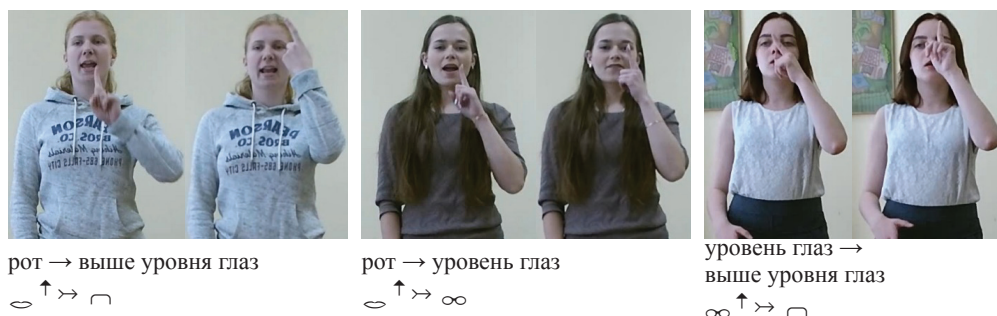


Рис. 10. Смена уровня выполнения жеста при различных начальной и конечных точках движения у разных информантов. В качестве примера используется жест кислый

Однако в подобных случаях конечная локализация всегда находится выше / ниже по оси Y, чем начальная, вне зависимости от абсолютной локализации в каждом конкретном случае. Поэтому для таких жестов имеет смысл ограничиться указанием того, что конечная точка движения находится выше / ниже определенного уровня.

Уверенное определение локализации позволяет отслеживать движение рук методом сравнения кадров. Предварительный анализ видеозаписей показывает, что параметр траектории движения обнаруживает стабильность. Совокупно с анализом глубины изображения это позволит в будущем с достаточно высокой уверенностью определять третий параметр жеста — движение и, таким образом, решить задачу распознавания жестов РЖЯ.

Заключение

В этой статье представлена лексическая мультимедийная база данных русского жестового языка TheRuSLan по предметной теме «продукты». Ценность подобной БД обусловлена двумя факторами: во-первых, исследований и, соответственно, баз данных и корпусов РЖЯ на сегодня достаточно мало. Исследование жестовых языков в России является сравнительно новой дисциплиной, что обуславливает сравнительную скудость источников по РЖЯ. Во-вторых, представленная БД представляет собой запись жестов в трехмерном (3D) формате, что предоставляет значительно больше возможностей для анализа и лексикографической обработки материала РЖЯ по сравнению с уже существующими БД.

Следует понимать, что настоящее исследование носит пилотный характер; в фокусе статьи, таким образом, оказывается не подробный анализ собранного языкового материала и сделанные на его основе выводы лингвистического характера, а собственно презентация базы данных лексики РЖЯ. Кроме того, в статье приведена фонетическая аннотация предметной лексики и очерчен круг проблем, с которыми столкнулись и к решению которых предложили подходы создатели настоящей БД. Особенный интерес представляет вариативность жестов у разных носителей. Напомним, что, согласно [Brentari et al. 2010: 284ff], частотность аллофонических чередований в жестовых языках, по-видимому,

гораздо ниже таковой в звучащих языках. В чем кроются причины подобной вариативности в представленной БД — с недостаточным знакомством информантов с [Гейльман 2004] или же с иными факторами — можно будет сказать только после специального исследования. Тем не менее авторам БД удалось, даже при столь высоком уровне различий в артикуляции, найти непротиворечивый способ аннотирования материала.

Описанная выше база данных на текущем этапе может быть отнесена к мультимедийным словарям РЖЯ, посвященным достаточно узкой предметной области, однако расширение словника и дальнейшая разработка системы аннотации могут привести к появлению еще одной лингвистической базы данных; разработанный инструментарий, как представляется, вполне может быть применен для аннотирования корпусов и баз данных жестовых языков.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Буркова 2012–2015 — Буркова С. И. *Русский жестовый язык: общие сведения*. [Burkova S. I. *Russkii zhestovyi yazyk: obshchie svedeniya* [Russian Sign Language: General information].] URL: <http://rsl.nstu.ru/site/signlang>.
- Буркова, Варинова 2012 — Буркова С. И., Варинова О. А. К вопросу о территориальном и социальном варьировании русского жестового языка. *Русский жестовый язык: Первая лингвистическая конференция*. Федорова О. В. (ред.). М.: Буки Веди, 2012, 127–143. [Burkova S. I., Varinova O. A. On geographical and social variation in Russian Sign Language. *Russkii zhestovyi yazyk: Pervaya lingvistsicheskaya konferentsiya*. Fedorova O. V. (ed.). Moscow: Buki Vedi, 2012, 127–143.]
- Воскресенский 2002 — Воскресенский А. Л. Непризнанный язык (Язык жестов глухих и компьютерная лингвистика). *Труды Международного семинара по компьютерной лингвистике и интеллектуальным технологиям «Диалог-2002»*. Т. 2. Протвино, 2002, 100–106. [Voskresenskii A. L. An unrecognized language (Sign language of the deaf and computational linguistics). *Trudy Mezhdunarodnogo seminar po komp'yuternoi lingvistike i intellektual'nykh tekhnologiyam «Dialog-2002»*. Vol. 2. Protvino, 2002, 100–106.]
- Воскресенский и др. 2009 — Воскресенский А. Л., Гуленко И. Е., Хахалин Г. К. Словарь RuSLED как инструмент семантических исследований. *Материалы Международной конференции «Диалог-2009»*. Вып. 8(15). М.: РГГУ, 2009, 64–68. [Voskresenskii A. L., Gulenko I. E., Khakhalin G. K. The RuSLED dictionary as an instrument for semantic research. *Materialy Mezhdunarodnoi konferentsii «Dialog-2009»*. Issue 8(15). Moscow: Russian State Univ. for the Humanities, 2009, 64–68.]
- Гейльман 2001 — Гейльман И. Ф. *Знакомьтесь: ручная речь*. М.: Зарпей, 2001. [Geyl'man I. F. *Znakom'tes': ruchnaya rech'* [Introducing the manual speech]. Moscow: Zagrei, 2001.]
- Гейльман 2004 — Гейльман И. Ф. *Словарь русского жестового языка*: в 2-х т. СПб: Прана, 2004. [Geyl'man I. F. *Slovar' russkogo zhestovogo yazyka* [A dictionary of Russian Sign Language]: In 2 vols. St. Petersburg: Prana, 2004.]
- Давиденко, Комарова 2006 — Давиденко Т. П., Комарова А. А. Краткий очерк по лингвистике РЖЯ. *Современные аспекты жестового языка*. Комарова А. А. (сост.). М.: БОГ, 2006, 146–161. [Davidenko T. P., Komarova A. A. A brief sketch of Russian Sign Language. *Sovremennye aspekty zhestovogo yazyka*. Komarova A. A. (comp.). Moscow: The All-Russian Deaf Society, 2006, 146–161.]
- Димскис 2002 — Димскис Л. С. *Изучаем жестовый язык: учебное пособие*. М.: Academia, 2002. [Dimskis L. S. *Изучаем жестовый язык: учебное пособие* [Learning the sign language: A manual]. Moscow: Academia, 2002.]
- Жесты 1995 — *Жесты: словарь-справочник*. М.: Зарпей, 1995. [*Zhesty: slovar'-spravochnik* [Signs: A reference dictionary]. Moscow: Zagrei, 1995.]
- Заварицкий 2015 — Заварицкий Д. А. *100 фраз на русском жестовом языке. Разговорник для священнослужителей*. М.: Лепта Книга, 2015. [Zavaritskii D. A. *100 fraz na russkom zhestovom yazyke. Razgovornik dlya svyashchennosluzhitelei* [100 phrases in Russian Sign Language. A phrasebook for priests]. Moscow: Lepta Kniga, 2015.]
- Зайцева 2000 — Зайцева Г. Л. *Жестовая речь. Дактилология*: учеб. для студ. высш. учеб. заведений. М.: Гуманит. изд. центр ВЛАДОС, 2000. [Zaitseva G. L. *Zhestovaya rech'. Daktilologiya* [Sign speech. Dactilology]. A textbook for higher school students. Moscow: VLADOS Publ., 2000.]

- Захаров 2005 — Захаров В. П. *Корпусная лингвистика: учеб.-метод. пособие*. СПб.: СПбГУ, 2005. [Zakharov V. P. *Korpusnaya lingvistika* [Corpus linguistics]. A teaching guide. St. Petersburg: St. Petersburg State Univ., 2005.]
- Кагиров 2018 — Кагиров И. А. Концепция роботизированной продуктовой тележки с поддержкой русскоязычного речевого и жестового интерфейсов пользователя. *Региональная информатика и информационная безопасность: сборник трудов*. Вып. 5. Советов Б. Я., Юсупов Р. М., Касаткин В. В. (ред.). СПб.: СПОИСУ, 2018, 273–277. [Kagirov I. A. A concept of an automatical shopping cart supporting spoken Russian and Russian Sign Language user interfaces. *Regional'naya informatika i informatsionnaya bezopasnost': sbornik trudov*. No. 5. Sovetov B. Ya., Yusupov R. M., Kasatkin V. V. (eds.). St. Petersburg: The St. Petersburg Society for Information Technology, Computation, Communication and Management Systems, 2018, 273–277.]
- Карпов 2011 — Карпов А. А. Компьютерный анализ и синтез русского жестового языка. *Вопросы языкознания*, 2011, 6: 41–54. [Karpov A. A. Computer analysis and synthesis of Russian Sign Language. *Voprosy Jazykoznanija*, 2011, 6: 41–54.]
- Карпов 2013 — Карпов А. А. Машинный синтез русской дактильной речи по тексту. *Научно-техническая информация. Серия 2: Информационные процессы и системы*, 2013, 1: 20–26. [Karpov A. A. Automatic synthesis of Russian fingerspelling from the text. *Nauchno-tehnicheskaya informatsiya. Seriya 2: Informatsionnye protsessy i sistemy*, 2013, 1: 20–26.]
- Карпов, Кагиров 2011 — Карпов А. А., Кагиров И. А. Формализация лексикона системы компьютерного синтеза языка жестов. *Труды СПИИРАН*, 2011, 1(16): 123–140. [Karpov A. A., Kagirov I. A. Formalizing the lexicon of an automatical sign language synthesis system. *Trudy SPIIRAN*, 2011, 1(16): 123–140.]
- Киммельман 2010 — Киммельман В. И. *Базовый порядок слов в русском жестовом языке*. Дипломная работа. М.: РГГУ, 2010. [Kimmelman V. I. *Bazovyi porядok slov v russkom zhestovom yazyke* [Basic word order in Russian Sign Language]. Graduation thesis. Moscow: Russian State Univ. for the Humanities, 2010.]
- Кипяткова, Карпов 2017 — Кипяткова И. С., Карпов А. А. Исследование нейросетевых моделей русского языка для систем автоматического распознавания слитной речи. *Автоматика и телемеханика*, 2017, 78(5): 110–122. [Kipyatkova I. S., Karpov A. A. A study of neural-network models of Russian for ASR systems. *Avtomatika i telemekhanika*, 2017, 78(5): 110–122.]
- КРЖЯ — *Корпус русского жестового языка* [Russian Sign Language Corpus]. URL: <http://rsl.nstu.ru/>.
- Львовская 2006 — Львовская А. *Анализ видовой и временной системы глагола русского жестового языка (РЖЯ) на основании работы Ö. Dahl «Tense and Aspect Systems»*. Курсовая работа. М.: МГУ, 2006. [L'vovskaya A. *Analiz vidovoi i vremennoi sistemy glagola russkogo zhestovogo yazyka (RZhYa) na osnovanii raboty Ö. Dahl «Tense and Aspect Systems»* [Analysis of aspect and tense system of the verb in Russian Sign Language based on Ö. Dahl's *Tense and Aspect Systems*]. Term paper. Moscow: Moscow State Univ., 2006.]
- Ляшевская 2016 — Ляшевская О. Н. *Корпусные инструменты в грамматических исследованиях русского языка*. М.: Издательский дом ЯСК, 2016. [Lyashevskaya O. N. *Korpusnye instrumenty v grammaticheskikh issledovaniyakh russkogo yazyka* [Corpus instruments in grammatical studies of Russian]. Moscow: YaSK Publishing House, 2016.]
- Мануева и др. 2014 — Мануева Ю. С., Гриф М. Г., Козлов А. Н. Построение системы компьютерного сурдоперевода русского языка. *Труды СПИИРАН*, 2014, 37(6): 170–187. [Manueva Yu. S., Grif M. G., Kozlov A. N. Developing a system of automatic translation of Russian speech into sign. *Trudy SPIIRAN*, 2014, 37(6): 170–187.]
- Правовые основы 2016 — *Правовые основы реализации Конвенции ООН о правах инвалидов в Российской Федерации*. Сборник нормативных правовых актов, принятых Российской Федерацией в целях выполнения положений Конвенции о правах инвалидов. Шестаков В. П. и др. (сост.). СПб.: НЦЭПР им. Г. А. Альбрехта Минтруда России, 2016. [Pravovye osnovy realizatsii Konventsii OON o pravakh invalidov v Rossiiskoi Federatsii] [Legal foundations of the realization of the UN Convention on the Rights of Persons with Disabilities in Russian Federation]. Shestakov V. P. et al. (comp.). St. Petersburg: G. A. Albrecht Federal Scientific Center of Rehabilitation of the Disabled, 2016.]
- Прозорова 2007 — Прозорова Е. В. Российский жестовый язык как предмет лингвистического исследования. *Вопросы языкознания*, 2007, 1: 44–61. [Russian Sign Language as a subject of linguistic study. *Voprosy Jazykoznanija*, 2007, 1: 44–61.]

- РЖЯБК 2001 — *Русский жестовый язык. Базовый курс. Компьютерная обучающая система*. М.: Истина, 2001. [*Russkii zhestovyi yazyk. Bazovyi kurs. Komp'yuternaya obuchayushchaya sistema* [Russian Sign Language. Basic course. A digital learning system.]. Moscow: Istina, 2001.]
- Сурдосервер — *Проект «Сурдосервер 2.0»* [Project “Surdoserver 2.0”]. URL: <http://surdoserver.ru>.
- Сысоев, Гриф 2018 — Сысоев М. Д., Гриф М. Г. Разработка системы удаленного сурдоперевода для глухих. *Разработка системы удаленного сурдоперевода для глухих. Инновационные инфокоммуникации XXI века*: 19 межвуз. Всерос. студен. науч.-практ. (очно-заоч.) конф. Клепиков С. И. и др. (ред.). Хабаровск: ХИИК, 2018, 97–99. [Sysoev M. D., Grif M. G. Developing a remote sign-translation system for the deaf. *Razrabotka sistemy udalennogo surdoperevoda dlya glukhikh. Innovatsionnye infokommunikatsii XXI veka*: The 19th student conf. Klepikov S. I. et al. (eds.). Khabarovsk: Khabarovsk Institute for Telecommunication and Information Technology, 2018, 97–99.]
- Федорова 2012 — Федорова О. В. (ред.). *Русский жестовый язык: Первая лингвистическая конференция*. М.: Буки Веди, 2012. [Fedorova O. V. (ed.). *Russkii zhestovyi yazyk: Pervaya lingvistsicheskaya konferentsiya* [Russian Sign Language. The 1st linguistic conference]. Moscow: Buki Vedi, 2012.]
- Федорова 2018 — Федорова О. В. О русской жестикуляции с лингвистической точки зрения (к выходу монографии Е. А. Гришиной). *Вопросы языкознания*, 2018, 5: 114–123. [Russian gestures from a linguistic perspective (dedicated to the publication of Elena A. Grishina's monograph). *Voprosy Jazykoznanija*, 2018, 5: 114–123.]
- Филимонова, Душкина 2018 — Филимонова Е. В., Душкина В. А. *Учебно-методическое пособие для изучающих русский жестовый язык*. М., 2018. [Filimonova E. V., Dushkina V. A. *Uchebno-metodicheskoe posobie dlya izuchayushchikh russkii zhestovyi yazyk* [A teaching guide for Russian Sign Language learners]. Moscow, 2018.] URL: https://www.academia.edu/38047551/Учебно-методическое_пособие_для_изучающих_русский_жестовый_язык.
- Фрадкина 2001 — Фрадкина Р. Н. *Говорящие руки: Тематический словарь жестового языка глухих России*. М.: Рефл-бук, 2001. [Fradkina R. N. *Govoryashchie ruki: Tematicheskii slovar' zhestovogo yazyka glukhikh Rossii* [Speaking hands: Thematical dictionary of Russian Sign Language].]
- Шамаро 2006 — Шамаро Е. Ю. Некоторые факты видо-временной системы РЖЯ. *Современные аспекты жестового языка*. Комарова А. А. (сост.). М.: БОГ, 2006, 180–191. [Shamaro E. Yu. Some facts from the tense-aspect system of the Russian Sign Language. *Sovremennye aspekty zhestovogo yazyka*. Komarova A. A. (comp.). Moscow: The All-Russian Deaf Society, 2006, 180–191.]
- AF 1996 — Arbeitsgruppe Fachgebärdenlexika. *Fachgebärdenlexikon Psychologie*. Hamburg: Signum, 1996.
- Battison 1978 — Battison R. *Lexical borrowing in American Sign Language*. Silver Spring (MD): Linstok Press, 1978.
- Bauer 2018 — Bauer A. Артикуляция слов в русском жестовом языке (РЖЯ) [Word articulation in Russian Sign Language]. *Deutsche Beiträge zum 16. Internationale Slavistenkongress. Belgrad, 2018*. Kempgen S. et al. (eds.). Wiesbaden: Harrassowitz Verlag, 2018, 35–44.
- Brentari 1998 — Brentari D. *A prosodic model of sign language phonology*. Cambridge (MA): MIT Press, 1998.
- Brentari et al. 2010 — Brentari D. (ed.). *Sign languages*. Cambridge: Cambridge Univ. Press, 2010.
- Chomsky, Halle 1968 — Chomsky N., Halle M. *The sound pattern of English*. New York: Harper & Row, 1968.
- Garcia, Sallandre 2013 — Garcia B., Sallandre M.-A. Transcription systems for sign languages: A sketch of the different graphical representations of sign language and their characteristics. *Body — language — communication. An international handbook on multimodality in human interaction*. Vol. 1. Müller C. et al. (eds.). Berlin: De Gruyter Mouton, 2013, 1125–1140.
- Digitgestus — *Digitgestus*. URL: <http://www.digitgestus.com>.
- Frischberg et al. 2012 — Frischberg N., Hoiting N., Slobin D. I. Transcription. *Sign language. An international handbook*. Pfau R. et al. (eds.). Berlin: De Gruyter Mouton, 2012, 1045–1075.
- Grenoble 1992 — Grenoble L. An overview of Russian Sign Language. *Sign Language Studies*, 1992, 77: 321–338.
- Hanke 2004 — *HamNoSys — Representing sign language data in language. Resources and language processing contexts*. URL: https://www.sign-lang.uni-hamburg.de/dgs-korpus/files/inhalt_pdf/HankeL-RECSLP2004_05.pdf.
- Hanke, Storz 2008 — Hanke T., Storz J. Ilex — a database tool for integrating sign language corpus linguistics and sign language lexicography. *Proc. of LREC 2008, Sixth International Conf. on Language*

- Resources and Evaluation*. Calzolari N. et al. (eds.). Paris: European Language Resources Association, 2008, 64–67.
- Huang et al. 2017 — Huang J., Rathod V., Sun Ch., Zhu M., Korattikara A., Fathi A., Fischer I., Wojna Z., Song Y., Guadarrama S., Murphy K. Speed/accuracy trade-offs for modern convolutional object detectors. *Proc. of 30th IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos (CA): Institute of Electrical and Electronics Engineers, 2017, 3296–3297.
- Johnston 2001 — Johnston T. The lexical database of Auslan (Australian Sign Language). Wilbur R. (ed.). *Sign Language and Linguistics*, 2001, 4(1/2): 145–169.
- Karpov et al. 2014 — Karpov A., Markov K., Kipyatkova I., Vazhenina D., Ronzhin A. Large vocabulary Russian speech recognition using syntactico-statistical language modeling. *Speech Communication*, 2014, 56: 213–228.
- Karpov et al. 2016 — Karpov A., Kipyatkova I., Zelezny M. Automatic technologies for processing spoken sign languages. *Procedia Computer Science*, 2016, 81: 201–207.
- Kennedy et al. 1997 — Kennedy G., Arnold R., Dugdale P., Fahey S., Moskovitz D. (eds.). *A dictionary of New Zealand Sign Language*. Auckland: Auckland Univ. Press; Bridget Willam Books, 1997.
- Kimmelman 2009 — Kimmelman V. *Reflexive pronouns in Russian Sign Language and Sign Language of the Netherlands*. MA thesis. Amsterdam: Univ. of Amsterdam, 2009.
- Kimmelman 2014 — Kimmelman V. *Information structure in Russian Sign Language and Sign Language of the Netherlands*. Ph.D. diss. Amsterdam: Univ. of Amsterdam, 2014.
- Language Archive — *The Language Archive*. URL: <https://tla.mpi.nl/forums/software/elan>.
- Liddell, Johnson 1983 — Liddell S., Johnson R. E. *American Sign Language: The phonological base*. Washington (D.C.): Ms. Gallaudet Univ., 1983.
- Perlmutter 1992 — Perlmutter D. M. Sonority and syllable structure in American Sign Language. *Linguistic Inquiry*, 1992, 3(23): 407–442.
- Prillwitz et al. 1989 — Prillwitz S., Leven R., Zienert H., Hanke T., Henning J. *HamNoSys. Version 2.0; Hamburg Notation System for sign languages. An introductory guide*. Hamburg: Signum Verlag, 1989.
- Ryumin et al. 2019 — Ryumin D., Kagirow I., Ivanko D., Axyonov A., Karpov A. A. Automatic Detection and Recognition of 3D Manual Gestures for Human-machine Interaction. *ISPRS — International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2019, XLII-2/W12: 179–183.
- RSL — “Russian Sign Language” at *Ethnologue*. URL: <https://www.ethnologue.com/language/rsl>.
- Sandler 1986 — Sandler W. The spreading hand autosegment of American Sign Language. *Sign Language Studies*, 1986, 50: 1–28.
- Sandler 2012 — Sandler W. The Phonological organization of sign languages. *Language and Linguistic Compass*, 2012, 6(3): 162–182.
- Sandler, Lillo-Martin 2006 — Sandler W., Lillo-Martin D. *Sign language and linguistic universals*. Cambridge: Cambridge Univ. Press, 2006.
- SLCN — *Sign Linguistics Corpora Network*. URL: <https://www.ru.nl/slcN/>.
- Spreadthesign — *Spreadthesign*. URL: <https://www.spreadthesign.com/ru.ru/search>.
- Stokoe 1960 — Stokoe W. C. *Sign language structure: An outline of the visual communication systems of the American deaf*. Buffalo: Dept. of Anthropology and Linguistics, Univ. of Buffalo, 1960.
- Sutton 1995 — Sutton V. *Lessons in sign writing. Textbook and workbook*. La Jolla (CA): Deaf Action Committee for Sign Writing, 1995.
- Thiessen 2011 — Thiessen S. *A grammar of sign writing*. MA thesis. Grand Forks (ND): Univ. of North Dakota, 2011.
- Van der Hulst, Channon 2010 — van der Hulst H., Channon R. Notation systems. *Sign languages*. (Cambridge Language Surveys.) Brentari D. (ed.). Cambridge: Cambridge Univ. Press, 2010, 151–172.