

Правительство Российской Федерации

Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
"ВЫСШАЯ ШКОЛА ЭКОНОМИКИ"»

Факультет филологии
Направление «Фундаментальная и прикладная лингвистика»

Курсовая работа

**ПОРОЖДЕНИЕ СЛОВАРЯ ДЛЯ ИСКУССТВЕННОГО ЯЗЫКА КАК
КОМПЬЮТЕРНО-ЛИНГВИСТИЧЕСКАЯ ЗАДАЧА**

Студент группы 23/л
2 курса подготовки бакалавров
Анастасия Олеговна Максимова

Научный руководитель
к.ф.н. доцент Б. В. Орехов

Москва
2013/2014 учебный год

1. Введение

В мире существует более шести тысяч языков, и помимо обыкновенных естественных языков также существуют искусственные, специально сконструированные (так называемые конланги). Их можно разделить на две группы: первая включает в себя языки, предназначенные для общения (например, эсперанто), вторая – языки, имитирующие естественные, создающиеся в первую очередь как часть художественной реальности и соотносящиеся с придуманными мирами. Примерами второго типа языков можно считать клингонский язык, разработанный лингвистом Марком Окрандом для одной из инопланетных рас из сериала «Звёздный путь», или квенья, разработанный Джоном Р. Р. Толкиеном и представляющий собой один из эльфийских языков для вселенных «Властелина Колец», «Хоббит, или Туда и Обратно» и «Сильмариллион».

Искусственные языки конструируются на основе существующих языковых данных. В первом случае создатели языка выбирают простые, интуитивно понятные будущему пользователю языка правила и явления: важно, чтобы создаваемый язык был прост и легок в изучении. Во втором же случае основным критерием для выбора того или иного языкового явления служит не простота или понятность, а его «экзотичность». Причём создатель интуитивно определяет экзотичность каждого явления. Важно отметить, что в обоих случаях заимствования явлений в создаваемый язык происходят бессистемно.

Целью настоящей работы является создание компьютерной программы, способной породить словарь для искусственно создаваемого языка. Что касается лексики конлангов, она создается вручную, и, как правило, каждое слово придумывается отдельно. Очевидно, что такой способ конструирования языка занимает много времени, и, что более важно, из-за этого такие придуманные словари редко бывают полностью продуманными. Обычно в языках, соотнесённых с художественными мирами, может быть подробно проработано одно или несколько актуальных семантических полей (поле оружия, магии, etc.), и при этом бытовая или просто нерелевантная лексика бывает вовсе не создана. Из-за этого язык теряет свою правдоподобность и схожесть с естественными языками. Итак, для того, чтобы словарь языка выглядел правдоподобно, должны быть выполнены два важных условия. Во-первых, порождение лексики должно быть систематичным, и, во-вторых, основываться на словаре уже существующего языка. Было решено, что только типологически редкие, но обязательно существующие, языковые явления будут заимствоваться в структуру лексики и грамматики. Лексика создаваемого языка, в свою очередь, будет ставиться в соответствие с лексикой русского языка. Сама компьютерная программа, которая будет порождать словарь, должна была быть написана на языке программирования Python с использованием модуля random, позволяющего получать

случайные последовательности букв, а также с использованием заранее готового алфавита и ряда ограничений на случайное порождение.

Эта работа, в основном, направлена на изучение лексикологических явлений в различных языках мира и их систематизацию, направленную на создание словаря, поэтому фонология, морфология и синтаксис конструируемого языка будут освещены очень кратко. Подробную информацию о грамматике можно найти в работе Ивана Левина [1].

Стоит отметить, что настоящая работа является очень актуальной в своей области, прежде всего потому что в лингвистике не очень распространены работы, посвященные изучению типологически редких явлений, а среди работ, направленных на изучение и создание искусственных языков, практически нет тех, которые рассматривают систематичное строение грамматики и словаря. Краткий обзор когда-либо существовавших конлангов находится в работе Умберто Эко [2].

2. Исследование по теоретическим данным

Для того, чтобы начать порождать словарь искусственного языка, нужно было для начала определить, как будут выглядеть входящие в него слова. Как правило, в большинстве языков длина слова варьируется от одного до четырёх слогов и для большей схожести с естественной структурой была выбрана средняя длина в три слога. Также было решено, что наиболее частотные не служебные лексемы, как, например, «мать», «отец», «человек», должны быть более короткими, чем остальные самостоятельные лексемы. Это решение было принято вследствие работы с частотными словарями русского [3] и английского [4] языков, в которых наиболее частотные лексемы имеют, в основном, не более двух, и, например, лексемы «год» (“year”), «время» (“time”), «день» (“day”) оказались в числе первых десяти наиболее частотных существительных. Средняя длина слов в обоих упомянутых языках равна примерно двум, и как видно, более частотные лексемы являются всё же более короткими, нежели менее частотные. Кроме того, с точки зрения удобства более частотные слова также имеет смысл сделать короче.

Следующей задачей стало определение состава слогов, составляющих слова. На основе уже придуманной фонологии была выбрана следующая модель слога – (C)V, а именно простой открытый слог с возможным отсутствием начального согласного звука. Несмотря на то, что структура языка должна была опираться на типологически редкие явления, слова всё-таки должны были быть привычными для взгляда пользователя, поэтому был введён ряд ограничений на возможные сочетания звуков внутри слова. Фонемный состав создаваемого языка состоит из 34 звуков, включая 8 гласных, из которых 4 являются назализованными ([ĩ], [ẽ], [ũ], [ã] и [ĩ], [ẽ], [ũ], [ã]), поэтому, во-

первых, гласный в конце слога не должен быть назализованным во избежание возникновения подобия закрытого слога. Исключением в таком случае может служить последовательность нескольких гласных (“fueēuka” разрешено, “fueēka ” быть не должно), из которых любой кроме последнего может быть назализованным. С этим связано второе ограничение: в слове не должно быть больше трёх гласных подряд, что сделано для упрощения произношения слов (“fivēaa” разрешено, “fivēaa ” быть не должно). В-третьих, разные группы звуков должны иметь различную частотность: так, лабиализованные, палатализованные ([pʷ], [tʷ], [kʷ], [kʷ], [tsʷ], [tʃʷ], [xʷ], [ɣʷ], [ŋʷ]) и назализованные ([ĩ], [ẽ], [ũ], [ã]) фонемы должны встречаться реже фонем без дополнительной артикуляции ([pʰ], [tʰ], [kʰ], [tsʰ], [tʃʰ], [f], [v], [s], [z], [ʃ], [ʒ], [x], [ɣ], [χ], [ʁ], [ħ], [ʕ], [h], [m], [n], [ŋ], [r], [ɽ], [j], [w] и [i], [e], [u], [a]).

Итак, когда состав и внешний вид отдельных слов был определён, можно было переходить к общему порождению. Чтобы сделать словарь наиболее правдоподобным и похожим на настоящие словари, нужно было выделить и описать комплекс факторов, придающих всей лексике, а не только отдельным словам, эту правдоподобность. Первоначальной гипотезой о том, что может являться одним из таких факторов, стала гипотеза о схожести слов по их внешнему облику. Так, предполагалось, что лексемы, принадлежащие к одной части речи или к одному именному классу, должны выглядеть похоже, например, иметь одинаковое окончание или аффикс.

Для того, чтобы доказать или опровергнуть эту гипотезу, было проведено исследование на базе словарей русского [5, 6] и английского [7, 8] языков, а также были учтены данные таких типологических и лексикологических работ, как работа О.П. Фроловой про глаголы звука в китайском [9] или работа В.А. Виноградова по именным классам в африканских языках [10]. В ходе исследования мы выделяли схожие слова сначала на основе словарей естественных языков, и находили более подробную и подтверждающую гипотезу информацию по конкретным классам, к которым принадлежат эти схожие слова.

В результате был получен следующий список схожих внешних черт слов принадлежащих одному классу:

- одинаковый аффикс
- одинаковые морфологические изменения в основе слов (например, переход сочетания “CoC” в “CaC”)
- одинаковые фонологические или тоновые изменения в основе слов

Следует теперь рассмотреть подробнее теперь каждую из черт. Одинаковый аффикс оказался самым частотным способом выражения принадлежности тому или иному классу

слов. Во-первых, часто общим аффиксом объединяются заимствования (например, “а-“ в русском языке). Во-вторых, общий аффикс могут иметь существительные из одного именного класса, как, например, имена небесных тел в грузинском языке имеют общий префикс “m`-” [11]. В-третьих, общим аффиксом могут обладать лексемы одной части речи: так, глаголы обычно имеют одинаковые окончания, что встречается в английском (например, “-ate”), русском (например, “-ть”), французском (например, “-er”) и в других языках. Следующим способом выражения принадлежности стали схожие морфологические изменения, (ср. gu-ranka “нога”, ma-tanka “ноги” в языке фула [10]). Фонологические или тоновые изменения в основе слов оказались самым редким способом выражения принадлежности слов к одному классу, и, как и морфологические изменения встречались чаще всего в африканских языках. [10]

Помимо заметных с первого взгляда особенностей слов, стоило также обратить внимание и на другие объединяющие факторы. Одним из таких факторов стало одинаковое словоизменение зависимых слов, что так же можно назвать склонением или родом, если речь идёт о существительных. Так, если в русском существительные принадлежат одному склонению, сочетающиеся с ними прилагательные и глаголы изменяются одинаково, в то время как сами существительные могут быть не схожи внешне. Вследствие изучения склонений в русском языке, возникла следующая гипотеза, требовавшая подтверждения. Возможно, не только элементы одного грамматического класса, но и элементы одного лексического класса могут иметь общие особенности. Внимание к этой теории было привлечено названиями цветов (растений) в русском языке. На первый взгляд кажется, что они ничем не похожи и не объединены ничем, кроме общего семантического поля, однако, в результате сравнения большинства названий цветов была выявлена их принадлежность женскому роду. Исключений совсем немного, и в основном это заимствования из греческого или латинского языка, где имена собственные приобрели статус нарицательных и стали использоваться для обозначения цветов (например, пион, ирис, гиацинт, нарцисс и др.). Для определения рода и наличия заимствований использовался этимологический словарь П. Я. Черных [12] и в ходе исследования выяснилось, что цветы часто принадлежат одному роду. Так, в большинстве славянских языков, как и в русском, семантическое поле цветов соотносится скорее с женским родом, а в западноевропейских, таких, как французский и немецкий, скорее с мужским родом. Получается, что семантическое поле может выделяться не только внешними признаками, но и их свойствами в склонении и при непосредственном написании и переводе текстов на конструируемый язык.

Кроме упомянутых выше схожих изменений в зависимых и сочетающихся словах, иногда элементы одного класса могут объединяться с помощью присоединения к ним одинаковых частиц или послелогов. Этот способ объединения напоминает отчасти первый способ с одинаковыми аффиксами, однако, в этом случае не происходит словоизменения, а прибавляются разные служебные слова или клитики. Подобное явление можно наблюдать в языке суахили, где для каждого класса имеются отдельные изменённые формы для каждой морфемы [10].

Завершая описание способов, объединяющих слова одного именного класса, следует отметить, что помимо использования каждого из способов по отдельности встречается также сочетание нескольких разных способов. Это, например, происходит в языке банту, где совмещено аффиксальное и тоновое словоизменение, или в языке суахили, где префиксальное словоизменение сочетается с изменением отдельных морфем [10].

Однако правдоподобность словаря не может ограничиваться одними только внешними сходствами слов. Следующими важными пунктами, так же связанными с внешним видом слов, стали явления омонимии и полисемии. Как правило, и то, и другое возникает относительно случайно, часто связано с заимствованиями слов из соседних языков или какими-то диахроническими изменениями в языке. Также омонимия часто возникает за счёт устоявшихся фонологических правил, а потому является одним из важных факторов, влияющих на правдоподобность словаря любого естественного языка.

3. Принцип порождения словаря

Теперь, когда теоретическая база была освещена, а комплекс факторов, придающий лексике правдоподобность, выделен и описан, можно было переходить к его практическому применению. Так как именная морфология для языка $\text{ʉeik}^{\text{wák}}\text{'u}$ (так называется созданный конланг, по лексеме “язык”) была уже создана, выбирать подходящие способы нужно было в соответствии с ней. Поэтому заранее было определено, что внешних признаков, выделяющих элементы одной части речи (например, глаголы или прилагательные) или элементы одного именного класса, если речь идёт о существительных, в словообразовательной системе не будет. Однако в морфологии уже было прописано, что именных классов у существительных должно быть четыре, а значит, принадлежность существительных к ним должна была быть выражена каким-либо образом. Поэтому было принято решение распределить все лексемы по четырем именным классам, обозначив в словаре их принадлежность к одному из классов цифрой около лексемы. Этот способ совпадает с одним из описанных выше, когда внешне слова не схожи, но при этом принадлежность одному классу проявляется при порождении отдельных предложений и текстов. В дальнейшем предполагается создать автоматический

переводчик на создаваемый конланг, который будет выбирать нужный классификатор или местоимение в зависимости от цифры, стоящей в словаре рядом с лексемой.

Возвращаясь к системе именных классов, стоит подробнее рассказать о принципе распределения лексем. Как уже было сказано, групп должно было быть четыре, а распределение автоматическим. Как правило, деление на именные классы происходит по семантическому принципу, однако, в таком случае следовало проделать большую работу по выделению семантических полей вручную. Так как программа должна была самостоятельно порождать словарь и, следовательно, распределять существительные по группам, надо было задать максимально простые ограничения по семантике. Поэтому было решено, что два класса будут создаваться на основе небольших заданных словарей с русскими лексемами (не стоит забывать, что словарь конструируемого языка ставился в соответствие со словарём русского языка), в то время как два других класса будут включать в себя случайные лексемы. Соответственно, семантическая мотивация была только у первых двух групп.

Итак, первый именной класс покрывал семантическое поле, которому можно дать условное название «Человек». В этот семантический класс входили лексема “человек” и её супплетивная форма множественного числа “люди”, лексема “бог” и связанные с ней, а также лексемы со значением имен родства, например, “мать”, “отец”, “дочь” и другие. Для этой группы слов стоит дополнительное ограничение на количество слогов. Как было сказано выше, наиболее частотные самостоятельные лексемы часто бывают короче остальных, поэтому лексемы из семантического поля «Человек» будут иметь длину от одного до трёх слогов при случайном порождении. Во второй именной класс входят названия зверей и птиц. По коллективному решению язык является изолированным, а носители обитают в горной долине в климате, приближенном к району Непала или Индии, поэтому в семантический класс животных были включены лексемы домашнего скота (например, “коза”, “баран”), домашней птицы (“курица”, “петух”), а также горные птицы, не редкие для подобной местности (например, “ястреб”). В третий и четвёртый именные классы лексемы распределялись случайно с помощью функции `random.randint()`, однако, для большей правдоподобности и предотвращения равных объёмов этих именных классов, распределение происходило в соотношении 1:2.

Прилагательные же отделялись от существительных с помощью парсера по окончаниям и порождались вместе с ними с помощью функции `newword`, порождающей случайные последовательности слогов. Следующая техническая особенность созданного генератора словаря заключалась в отдельном порождении глаголов и отглагольных существительных. Из-за случайного порождения слов, хоть и с рядом ограничений,

однокоренные в русском языке слова практически никогда не были похожи друг на друга в конструируемом языке. Однако из-за того, что лексемы не могли быть абсолютно не похожи, было решено, что отглагольные существительные должны образовываться от исходных глаголов с помощью номинализационного префикса “ju” и соединительного гласного, случайно выбирающегося из гласных фонем без дополнительной артикуляции ([i], [e], [u], [a]) с помощью функции `random.choice()`. За счёт такого способа словообразования возникает омонимия среди имён действия, а также дополнительное сходство среди лексем. Стоит отметить также, что невозможность узнать, как во времени развивался создаваемый конланг принимается за аксиому, а потому полисемия и дополнительная омонимия будет возникать в порождаемом словаре случайным образом за счёт неравного распределения фонем с и без дополнительной артикуляции.

Помимо отглагольных существительных отдельно порождаться должны были также аффиксы, служебные слова и местоимения. Очевидно, что они должны были быть похожи на самостоятельные лексемы, но при этом из-за того, что аффиксы должны присоединяться к уже созданным словам, нужно было наложить дополнительные ограничения на их состав, чтобы избежать запрещённых сочетаний. Во-первых, служебные слова и аффиксы должны были быть короче самостоятельных лексем, и, во-вторых, они не должны были создавать закрытых слогов или слишком длинных сочетаний гласных. Помимо этого некоторые из служебных слов должны быть родственны прочим лексемам, что было уже прописано в грамматике. Так, например, прилоги должны быть родственны местоимениям, а возвратное местоимение “себя” должно было быть родственно лексеме “голова”. Последний аспект обосновывается данными из статей Wals.

Теперь, когда окончательное порождение словаря проработано, стоит обратиться к его использованию и самостоятельному существованию. Как говорилось выше, словарь конструируемого языка ставился в соответствие со словарём русского языка. В русском словаре содержалось около 120000 лексем и, очевидно, что некоторые из концептов, имеющих в нём, могут отсутствовать в создаваемом конланге. Именно это придаёт порождаемому словарю большую правдоподобность и более широкие возможности в области переводов текстов с других языков.

В приложении к этой работе помещены список служебных слов и часть словаря, с которыми можно ознакомиться, чтобы оценить их правдоподобность, и художественный текст, переведённый на конструируемый язык. Код программы, породившей этот словарь, можно будет найти и также оценить в Интернете на сайте github.com.

4. Заключение

Итак, у настоящей исследовательской работы было несколько важных целей: освещение и систематизации существующих явлений, присущих словарям различных естественных языков, описание комплекса факторов, придающих правдоподобность этим словарям, выделение из него типологически редких факторов и, наконец, попытка использовать полученные данные на практике в рамках проекта по созданию искусственного языка.

Обобщая результаты теоретической части исследования, можно разделить на три группы особенности словарей естественных языков, отличающие их от словарей искусственных языков. Во-первых, словари естественных языков имеют проработанную лексику, как правило, для всех семантических полей за редкими исключениями. Это позволяет облегчить коммуникацию между носителями разных языков и перенос концептов из одного языка в другой. Во-вторых, лексемы любого настоящего словаря обладают некоторыми внешними сходствами между собой, среди которых можно упомянуть одинаковые части слов (аффиксы или клитики), примерно одинаковую длину большинства лексем или их схожее словообразование и словоизменение, что также связано с их согласованием с прочими лексемами. В-третьих, помимо частичных внешних и грамматических сходств, ни один словарь не обходится без лексических схожих особенной, таких как полисемия или омонимия.

Закончив с обзором теоретических результатов, можно обратиться к результатам попытки создания программы-генераторы. В настоящее время программа, порождающая словарь, является практически универсальным генератором лексики. На вход принимается алфавит, количество слогов, набор морфем, порождающихся отдельно, а также словарь языка, ставящегося в соответствие с создаваемым конлангом. Однако, несмотря на свою полноту, сам по себе словарь не имеет прикладной ценности. Важно, что в программу на данный момент помещены ограничения, касающиеся конкретной грамматики создаваемого языка, такие, как разбивка по разным именным классам или отдельное специальное порождение отглагольных существительных. Поэтому в дальнейшие планы входит окончательная разработка автоматического переводчика, принимающего на вход текст на русском или любом другом языке и выдающего текст на создаваемом конланге. Так же в нынешней версии программы омонимии не так много, из-за чего словарь теряет свою правдоподобность и схожесть с естественными языками, и это предполагается дорабатывать в дальнейших версиях программы.

5. Список литературы

1. И. С. Левин, Конструирование фонологии и грамматики как лингвистическая задача (случай типологически редкого языка). 2014.

2. У. Эко, В поисках совершенного языка в европейской культуре. СПб.: Alexandria, 2007.
3. О. Н. Ляшевская, С. А. Шаров, Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009.
4. Б. Рудый. Частотный словарь английского языка. Киев (<http://langs.com.ua/movy/eng/freq1-500-r.htm>)
5. С.И.Ожегов, Словарь русского языка. М.: Рус.яз., 1986.
6. В.И.Даль, Толковый словарь живого великорусского языка. М.: Русский язык, 1978.
7. V. K. Muller, English-russian dictionary. М.: Moscow Russian Language Publishers, 1978.
8. Китайское языкознание: Изолирующие языки: IX международная конференция: Материалы. - М., 1998. - С. 183-185
9. В.А.Виноградов, Том 1. Основы африканского языкознания. Именные категории. М.: Аспект Пресс, 1997.
10. Б.Т. Руденко, Грамматика грузинского языка. Ленинград: АН СССР, 1940.
11. П. Я. Черных, Этимологический словарь. М.: Рус.яз.-Медиа, 2004.

6. Приложение

6.1. Список аффиксов и служебных слов

<u>Аффиксы</u>	CLF2 (suf) - k ^w i
POSS1 (pref) - ɬe	CLF3 (pref) - tʃ ^u
POSS2 (pref) - ʒa	CLF4 (suf) - ts ^h a
POSS3 (suf) - ŋ ^w ux ^w u	1SG.NFUT (pref) – ŋi
POSS4 (pref) - k ^w u	1PL.NFUT (pref) - ts ^h e
INDEF (pref) - ju	INCL.NFUT (pref) - ŋ ^w uk ^ʔ e
NMNLZ (pref) - ʃu	2SGINF.NFUT (pref) - t ^ʔ ije
PL (pref) - wi	2SGF.NFUT (pref) - hi
PL verbs (pref) - ŋ ^w u	2PL.NFUT (pref) - k ^ʔ a
1SG - iha	1SG.FUT (pref) - hi
1PL - se	1PL.FUT (pref) – wi
INCL - k ^h u	INCL.FUT (pref) - k ^w u
2SGF - k ^ʔ ua	2SGINF.FUT (pref) - t ^ʔ ey ^w a
2SGINF - t ^ʔ iu	2SGF.FUT (pref) - p ^ʔ iχu
2PL - xixe	2PL.FUT (pref) - yi
CLF1 (pref) - ha	3.FUT (pref) - ŋaak ^w u

3.NFUT (pref) - x^wu

MOD (suf) - t'ũŋi

PERF (pref) - t^ha

IMP - χe

SUBJ (pref) - t^hu

OPT (suf) - ŋ^weha

SENS (pref) - k'u

NEG (suf) - χu

Клитики

NOM - Ø

ERG (pref) - ʃa

GEN (pref) - x^wuni

DAT (pref) - p'at^hi

LOC/PREP (pref) - k^hĩa

PL (pref) - wi (для неодушевлённых)

Другие служебные слова

COM (pref) - t^hats'e

LAT (pref) - t'i

DISTR (pref) - γ^wi

QUOT (pref) - i

NEG (suf) - χa

Местоимения

1SG - iħa

1PL - se

INCL - k^hu

2SGF - k'ũa

2SGINF - t'iu

2PL - xixe

кто-то (suf) - t^henu

что-то (pref) - i

Dem (pref) - i

Dems (pref) - tʃ'e

Demr (pref) - t'a

Deml (pref) - u

сам (pref) - ts'a

себя (pref) - x^wet'aa

6.2. Отрывок из словаря

мавзолей 4 - wiħe 4

мавр 4 - k^wuetʃ'et^hit^he 4

мавра 4 - x^weuke 4

мавританка 3 - k'ix^wea 3

мавританский - γi

маврский - γ^wiγuniγ^wu

маг 3 - t'alizi 3

магазин 4 - iu 4

магазинный - juazenu

магазинщик 4 - awiaγ^wu 4

магистерско 3 - aŋ^wuiats'a 3

магистерский - ne

магистерство 3 - ħae 3

магистр 3 - zuγ^wu 3

магистраль 3 - ħuγi 3

магистральный - ts^hukeet^hu

магистрант 4 - utʃ'eru 4

магистрат 3 - ʃireħa 3

магистратский - k^hep'ux^wi

магистратура 4 - ŋ^wue 4

магически - ziγ^wuγ^wu

магический - uk^wutʃ'aup'u

магия 3 - ŋ^wuts'a 3

магма 4 - x^waħa 4

магматический - ts'ets^he

магнат 4 - juŋevezi 4

магнатский - iɪy^we
 магнезиальный - ʃeʉtʃ^hazek^{wu}
 магнезит 3 - ʃueɬup'a 3
 магнезия 3 - k^huip'axi 3
 магнетизация 4 - t'iu 4
 магнетизированный - tʃ^hunāifu
 магнетизировать - k'uɬi
 магнетизм 3 - jɛtʃ^hu 3
 магнетизёр 3 - t'ua 3
 магнетизёрский - eip'ɪenu
 магнетизёрство 4 - ɲaju 4
 магнетический - juts^haxisi
 магниевый - ts'eʃanaɣ^{wu}
 магний 3 - tʃ^hu 3
 магнит 4 - ɣesits^hevi 4
 магнитить - k'unuk'a
 магнитный - p'a
 магнито 4 - uk'aɲia 4
 магнитограмма 4 - wayiʃɛʃu 4
 магнитограф 3 - avazimi 3
 магнитола 3 - ts^haju 3
 магнитолог 4 - sumase 4
 магнитометр 3 - t'iak'ale 3
 магнитометрия 3 - t^hiheip'u 3
 магнитострикция 3 - jajāu 3
 магнитосфера 4 - ɣik'uhe 4
 магнитофон 4 - p'imāiwi 4
 магнитофонный - vuap'aale
 магнификат 4 - tʃ^hay^weʃiɲ^we 4
 магнолия 3 - хупагиву 3
 магометанин 4 - ɲ^wuheɬuk^{wu} 4
 магометанский - p'itʃ^haza
 магометанство 4 - k^weweɬaak'u 4
 мадам 4 - aɬu 4
 мадемуазель 3 - ts'eʒak'i 3

мадера 4 - heʃa 4
 маджара 3 - ɲifeu 3
 мадригал 4 - tʃ^huvaami 4
 мадригальный - ɣ^wuuk^weɣ^wa
 мадьяр 4 - ts'ak'a 4
 мадьярский - k'eʒuk^hufe
 маета 4 - featʃ^hii 4
 маетность - he
 маетный - hiʒazats'i
 мажор 4 - aʃeuts'ihi 4
 мажордом 3 - hiɣ^wuɲ^waat'u 3
 мажоритарный - firazahe
 мажорно 3 - mix^wiɣ^wetʃ^hu 3
 мажорный - ts^haa
 мазальщик 4 - ɣ^wāiye 4
 мазанка 3 - t^hiɰuzu 3
 мазанковый - suts'eɲ^{wu}
 мазануть - ɣ^wasu
 мазать - p'ewuzi
 мазаться - t'ap^hu
 маздакизм 4 - p'ay^wi 4
 маздеизм 4 - hezi 4
 мазевый - tʃ^hafaak^hi
 мазеобразный - x^wexiup'a
 мазковый - ʃi
 мазнуть - mat'ахаха
 мазня 4 - x^wīik^hi 4
 мазок 4 - wanats'a 4
 мазохизм 3 - k^huki 3
 мазохист 3 - ʃats^hazaa 3
 мазохистский - ɣap'ɲ^wi
 мазурка 4 - vak^waka 4
 мазурничать - aɲaek^weɣ^wa
 мазурничество 4 - t'uik'eʃeatʃ^ha 4
 мазурочный - ɲ^wafax^we

мазурский - evaniʕi
 мазут 3 - vah^wemep^{ha} 3
 мазутный - t'u
 мазчик 3 - ɲ^wits^{hi}zetʃⁱ 3
 мазь 4 - k'isuŋa 4
 маис 4 - ɲep^wezaŋ^{we} 4
 майсовый 3 - zur'uŋ^wake 3
 май 4 - x^wuɣux^wia 4
 майка 3 - wiŋa 3
 майолика 4 - ux^wiat'aɣu 4
 майоликовый - suaɣ^wat^{hu}
 майонез 4 - ju^tʃ^{hu}vezi 4
 майонезный - fietʃⁱat'u
 майор 3 - uŋ^wa 3
 майоран 3 - k'afatʃ^{ha} 3
 майорановый - ueɣ^wi
 майорат 3 - hiŋu 3
 майоратный - mefu
 майоратство 4 - ɣaiha 4
 майордом 4 - hiitʃⁱuits'ip'a 4
 майоритет 3 - zak^withisu 3
 майорский - haŋ^wu
 майорша 4 - k^wuvi 4
 майский - ɣiek^wu
 майя 4 - juatʃⁱ 4
 мак 4 - eɣek^wie 4
 макака 3 - ip'uk'e 3
 макание 4 - ʃuuk^{hu}ʕa
 макаронный - i
 макароны 4 - aɣ^we 4
 макательный - fuhãa
 макать - ɣ^weajãi
 макаться - ɲinaɣetʃ^{he}
 македонец 4 - ep^wuŋ^{we} 4
 македонский - k^{hi}ux^{wa}

македонцы - ʕek^{hua}
 макет 3 - ts^{hu}ŋ^wi 3
 макетировать - ɣ^wex^wei
 макетный - p'i
 макетчик 4 - eɣ^wiɣaŋ^{wa} 4
 макиавеллизм 4 - x^wit^{hu} 4
 макияж 3 - ts'aku 3
 маклер 4 - p'uŋawe 4
 маклерство 4 - k^{hu}ut'ak^we 4
 макнуть - iusɲ^wu
 маковка 4 - nets'ex^we 4
 маковый - ɲ^we
 макраме 3 - x^wazex^wup^{hi} 3
 макрелевый - ui
 макрель 4 - t^{ha}zevi 4
 макро 4 - wesatʃⁱa 4
 макрокефал 3 - zetʃⁱets^{ha}awu 3
 макрокефалия 3 - p^{he}k'aɣise 3
 макроклимат 3 - ɣeɣa 3
 макрокосмос 4 - ɣaɣ^wũuke 4
 макромир 4 - fi^xuxi 4
 макромолекула 3 - t^{hi}fak'u 3
 макроподы 4 - muʃue 4
 макропроцессор 4 - iap'i 4
 макрорайон 4 - t'afã 4
 макрорельеф 4 - lek'iu 4
 макроскопический - hut^{he}it'e
 макросоциология 4 - ɣ^wietʃⁱats'u 4
 макроспорангий - ts^{he}euts'asa
 макроспориоз 4 - ɣik'ik'e 4
 макроспорофилл 4 - ts'et^{hi}k^wizu 4
 макроспоры 4 - uk'aɣa 4
 макроструктура 4 - vase 4
 макроуровень 3 - k'ihɪ 3
 макрофаг 4 - emik^wits'a 4

макроцефал 4 - eʷĩexʷi 4	малоавторитетный - tʃʰutʰua
макроцефалия 3 - zekaletsʰu 3	малоазиатский - sunjeʃu
макроэкономика 3 - aeji 3	малоалкогольный - ɢa
макроэкономический - mitʰeŋʷaĩtsʰe	малоблагоприятный - tsʰazuʷexʷu
максимализм 3 - pʰufe 3	малобюджетный - kaekʰaxi
максималист 4 - xʷizi 4	маловероятный - xʷi
максималистский - ʷewautʰi	маловесный - ɲa
максимально 4 - miɦuʃii 4	маловместительный - kiŋitsʰetʰa
максимальный - xehi	маловодный - tʃʷuwakʰi
максимизация 4 - ʃaʷasa 4	маловодье 4 - uʃetʰaka 4
максимум - kʰunũa	маловразумительный - utsʰa
макулатура 4 - ʃewatʃʷi 4	маловыгодный - tʰe
макулатурный - tʰukʰa	маловыразительный - fuxũaxi
макушечный - ʃu	маловысотный - ɦukʷaxexʷe
макушка 4 - ua 4	малогабаритка 3 - wiɦitʃʷe 3
малайзийский - pʰe	малогабаритный - ɢiwaʒana
малайский - tʰa	малоговорящий - iɢe
маланья 3 - jewe 3	малоголовый - su
малахай 4 - iãi 4	малограмотность - tʰitʰia
малахит 4 - tʃʷitʰuʷa 4	малограмотный - tsʰetsʰaɦe
малахитовый - ɲexʷixavi	малодаровитый - tʰiʒawa
малевальный - ua	малодейственный - ietʰajazi
малеванье 4 - miʷiniʒe 4	малодействительный - xaɦe
малевать - ɲaʷeue	малодержанный - ɲafimiju
маленькая 3 - ʷixʷĩene 3	малодойка 4 - iŋatʃʰu 4
маленький - pʰe	малодойный - tsʰe
малина 4 - uzatʰuxʷu 4	малодоказательный - i
малинник 4 - itsʰexekʰi 4	малодостоверный - tʰiʷawu
малинный - tʃʷamãiukʰa	малодоступный - ɦeiva
малиновка 4 - tʃʷẽapʰu 4	малодоходный - ɢĩiatʃʷea
малиновый - tʰiʃukʷu	малодушество 4 - jiʃi 4
мало 4 - jutʰekʰu 4	малодушествовать - axʷaʃaitʃʷu
мало-мальски 3 - tsʰiʃãetsʰu 3	малодушие 4 - eʒa 4
мало-мальский - ɲakʰi	малодушничать - tʃʰietʃʷava
мало-помалу 3 - waxʷeʷu 3	малодушно - tsʰuʷa

малодушный - ɸip'atʃʰiri
 малоезженный - kʰek'eɣukʷa
 малоезжий - zuʰea
 малозавидный - zur'utʰee
 малозаметный - uketsʰeɣu
 малозаселенный - ets'i
 малозатратный - t'uŋʷelisi
 малоземелье 4 - una 4
 малоземельный - a
 малознакомый - wiʃezia
 малозначащий - ɲa
 малозначительный - tʃʷuχutʃʰi
 малоизведанный - vitʃʷekʷitʃʷu
 малоизвестный - sexiʰu
 малоизученный - ɣʷusiats'a
 малоимущий - tʃʰak'ixʷe
 малоинтеллигентный - wamu
 малоинтересный - kʷetʃʰitʰaɣʷe
 малоискусный - p'a
 малоисследованный - ʃetsʰaɣahu
 малокалиберный - tʃʰitʰaŋe
 малокалорийный - ɲia
 малоквалифицированный - k'a
 малоквартирный - ifeaxʷi
 малокомпетентный - hiŋʷu
 малокормный - k'exʷu
 малокровие 4 - jiɣaɣa 4
 малокровный - ee
 малокультурный - tʃʰuuetʃʰu
 малолетний - va
 малолетство 4 - ɲʷaʃekʰu 4
 малолитражка 3 - ɣive 3
 малолитражный - pʰaruts'u
 малолюдность - ʰuefiɣʷi
 малолюдный - xʷa

малолюдство 4 - tʃʰuhu 4
 малолюдье 4 - tsʰat'ats'i 4
 маломальский - ɣʷu
 маломерка 4 - tʃʷeti 4
 маломерный - xʷeneisa
 маломерок 3 - tsʰeʃeɲʷi 3
 малометражный - ts'iqʷuʰe
 маломощность - kuɣaha
 маломощный - χuiʰu
 малонадежный - xipʰik'e
 малонадёжный - ʃetep'e
 малонаезженный - ts'ik'a
 малонаезженный - tʰa
 малонаселенный - tʰe
 малонаселённый - kʰip'ezi
 малоношенный - kʷexʷi
 малообеспеченный - ɲʷihaliʃa
 малообжитой - zimuxiva
 малообитаемый - tʃʷiyʷa
 малооблачный - ɣʷiʰu
 малообоснованный - tʃʷatʰeap'iɣi
 малообразованный - tʃʷuts'aχutsʰa
 малообщительный - tʃʰakui
 малоопытный - heɣa
 малоосвоенный - kʷuxikʷi
 малоосновательный - miɣũure
 малоотходный - ɲeaja
 малопитательный - urʰeɲeɣu
 малоплодовый - tsʰetʃʰuup'a
 малоплодородный - t'umu
 малоподвижный - vakʷe
 малоподготовленный - tsʰa
 малоподходящий - vii
 малополезный - ʃiɲʰe
 малопоместительный - χip'u

малопоместный - χiуааа
 малопонятный - ts'an'w'ik'ааи
 малопочтенный - muwaχiχи
 малоприбыльный - k'a
 малопривлекательный - k^hijała
 малопригодный - wie
 малоприменимый - t'aу^{wu}
 малоприметный - fawi
 малоприспособленный - t^hiuη^{wu}
 малопристойный - eheek'a
 малоприятный - uye
 малопродуктивный - lauu
 малопроезжий - t^hit'уар'е
 малопроизводительный - x^{wi}hi
 малопьющий - t^hut'ucheja
 малоразвитой - t'e
 малоразвитый - jaiak'i
 малоразговорчивый - ize
 малоразмерный - sizufu
 малораспространенный - zats'auk'a
 малорентабельный - ruу^{wak}hi
 малоречивый - t^hut^hinu
 малорослый - fu
 малороссийский - fuη^{wik}ux^{wi}
 малороссия 4 - η^{wi}ji 4
 малороссиянин 4 - у^{wut}t^hayey^{wu} 4
 малороссиянка 4 - zujeke 4
 малоросский - k^{wi}
 малороссы 3 - χihe 3
 малосведущий - p^huxak^{wefe}
 малосемейный - ets^hujece
 малосильный - aaufa
 малосимпатичный - ηaxix^{wey}we
 малосмысленный - x^wa
 малоснежный - i

малоснежье 3 - zuts'ats'i 3
 малосодержательный - vits^ha
 малосознательный - ge
 малосольный - p'уay^{wi}χи
 малосостоятельность - u
 малосостоятельный - у^{wu}χεje
 малоспособный - va
 малостоящий - χik^wanemi
 малость - k^{wik}ux^weri
 малосущественный - p'i
 малотиражный - p'a
 малотоннажный - p'i
 малоубедительный - hip'u
 малоуглеродистый - k^{wup}'uu
 малоуглеродный - x^{wu}
 малоудойливый - ts'afeyuhi
 малоудойный - p'ap'uk^{hee}
 малоумие 4 - p'iziłu 4
 малоупотребительность - zuhu
 малоупотребительный - zu
 малоурожайный - ts^hi
 малоусидчивый - у^wa
 малоуспевающий - χiviу^{wuts}ha
 малоуспешный - t^hik^{he}gu
 малоутешительный - уey^wуар'е
 малоформатный - t^hhe
 малохлебный - η^{wi}
 малохоженный - zama^hu
 малохольный - χiets^ha
 малохудожественный - i
 малоценный - fiakaи
 малочисленность - у^wuna
 малочисленный - x^wa
 малочувствительный - ts^he

6.3. Перевод басни с русского на язык геикваку

Овца и кони (по А. Шлейхеру)

овца с LOC=PL-конь

k'eetʃ'eʃa k'h̥a=wi-jukʷiʃaʎu

Стриженная овца увидела коней, везущих тя желую повозку с большим грузом, быстро несущих человека.

DAT=PL-конь 3.NFUT-PERF-говорить ERG=овца [[[тяжёлый=mod indef-повозка с loc=[большой=mod indef-груз]] 3.nfut-везти]=mod [indef-человек быстрый 3.nfut-нести]=mod pl-indef-конь] 3.nfut-perf-видеть erg=[3.nfut-perf-стричь=mod indef-овца] ts'ats'ēu=t'uni ju-p'usi t'ats'e k'h̥a-haʃimuʃu=t'uni ju-xip'huts'i xwu-heʒaukʷuʃu=t'uni ju-k'e p'ets'h̥i xwu-ʃēaxaet'e=t'uni wi-ju-jukʷiʃaʎu xwu-tʰa-ŋāip'h̥etʃ'a ʃa=xwu-tʰa-ħukʰap'h̥eni=t'uni ju-k'eetʃ'eʃa

Овца сказала коням:

DAT=PL-конь 3.NFUT-PERF-говорить ERG=овца p'atʰi=wi-jukʷiʃaʎu xwu-tʰa-zaʃuʷuxwu ʃa=k'eetʃ'eʃa

“Горит моё сердце, когда вижу коней, везущих человека”.

[GEN=ISG POSS2-сердце] 3.NFUT-гореть когда [[INDEF-человек 3.NFUT-везти]=MOD PL-INDEF-конь] 3.NFUT-SUBJ-видеть ERG=ISG QUOT xwuni=iha ʃa-ts'uts'i xwu-phits'ik'a zeip'e ju-k'e xwu-heʒaukʷuʃu=t'uni wi-ju-jukʷiʃaʎu xwu-tʰa-ŋāip'h̥etʃ'a ʃa=iha i

Кони сказали:

PL-конь 3NFUT-PERF-говорить wi-jukʷiʃaʎu xwu-tʰa-zaʃuʷuxwu

«Слушай, овца, наше сердце то же горит, потому что мы знаем:

слушать(IMP) овца [GEN=PL-1PL POSS2-сердце] тоже 3.NFUT-гореть потому.что PL-1PL 3.NFUT-знать it'h̥i k'eetʃ'eʃa xwuni=wi-se ʃa-ts'uts'i fenuie xwu-phits'ik'a wu wi-se xwu-rits'a

человек из овечьей шерсти делает себе новую тёплую одежду,

[новый=MOD тёплый=MOD одежда] из LOC=[GEN=INDEF-овца POSS2-шерсть] DAT=POSS1-голова 3NFUT-делать ERG=INDEF-человек umusūe=t'uni vuŋʷi=t'uni hut'eʃusa hiŋʷe k'h̥a=xwuni=ju-k'eetʃ'eʃa ʃa-tʃ'iap'aʃue p'atʰi=le-xwet'aa xwu-iʃisith̥e ʃa=ju-k'e

а у овец не остаётся шерсти».

[POSS2-шерсть GEN=овца] 3NFUT-оставаться-NEG QUOT . ʃa-tʃ'iap'aʃue xwuni=k'eetʃ'eʃa xwu-p'eakʷeʒaʃuʷi-χu i

Когда овца услышала это, она убежала в поле.

когда DEMS 3NFUT-SUBJ-PERF-слышать ERG=овца DEMS в LOC=INDEF-поле 3NFUT-PERF-бежать

zeip'e tʃ'e xwu-tʰu-tʰa-nak'ara ʃa=k'eetʃ'e ʃa tʃ'e ŋʷaχakʷuxʷa kʰɿa=ju-ŋemi xwu-tʰa-xʷejəŋeh