

Слоговая структура адыгейского языка: от данных к обобщениям

© 2019

Георгий Алексеевич Мороз

Национальный исследовательский университет «Высшая школа
экономики», Москва, Россия; agricolamz@gmail.com

Аннотация: В данной работе приводятся результаты исследования слоговой структуры адыгейского языка. Слоговые структуры адыгейского языка варьируются от простой (V) до сложных (CCCVCC или CVCCC) с большими консонантными кластерами. В ходе предварительного эксперимента была обнаружена значительная вариативность в суждении носителей о слогоделении. Мы проанализировали слоговую структуру адыгейского языка на основе адыгейско-русского словаря Ю. А. Тхаркахо 1991 г. Ввиду вероятного «волнового» характера слога в адыгейском языке интервокальные кластеры в рассмотрение не включались, так как определить слоговую границу в таких кластерах не представляется возможным. В работе предложена общая схема структуры адыгейского слога и описаны исключения из нее. Оказалось, что инициали и финали адыгейских слогов имеют разную структуру, однако почти все возможные комбинации шумных и сонорных можно встретить и в инициали, и в финали. Вот основные выводы относительно слоговой структуры, к которым мы пришли: (а) в исконной адыгейской лексике не представлены кластеры сонорных; (б) в позиции инициали сонорные встречаются лишь поодиночке, в то время как шумные могут встречаться в кластере; (в) структура «гласный-сонорный-шумный» невозможна, но возможна структура «гласный-шумный-шумный»; (г) составные финали имеют структуру «сонорный-шумный-шумный», и не существует других трехкомпонентных финалей. Кроме того, в работе подсчитана ожидаемая частотность кластеров с теми или иными типами согласных в инициали и финали, как если бы появление того или иного типа согласных было независимым событием. Сопоставив ожидаемые и наблюдаемые значения, мы обнаружили, что (а) кластеры шумных в наших данных встретились чаще, чем ожидалось; (б) кластеры сонорных в наших данных наблюдаются реже, чем ожидалось; (в) кластеры согласных в наших данных были обнаружены чаще, чем ожидалось. Из всех этих наблюдений можно сделать вывод, что нулевую гипотезу о независимости появления тех или иных типов согласных в инициальных и финальных кластерах можно отвергнуть.

Ключевые слова: абхазо-адыгские языки, адыгейский язык, слог, фонология

Для цитирования: Мороз Г. А. Слоговая структура адыгейского языка: от данных к обобщениям.

Вопросы языкознания, 2019, 2: 82–95.

DOI: 10.31857/S0373658X0004303-2.

Благодарности: Статья подготовлена в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) и с использованием средств субсидии в рамках государственной поддержки ведущих университетов Российской Федерации «5-100».

Автор выражает глубокую признательность М. А. Даниэлю, Н. Р. Добрушиной, И. А. Зибер, И. В. Щурову и участникам семинара Международной лаборатории языковой конвергенции НИУ ВШЭ, а также двум анонимным рецензентам журнала «Вопросы языкознания» за обсуждение текста и результатов статьи. Ответственность за возможные ошибки и неверные интерпретации, естественно, лежит на совести автора.

Adyghe syllable structure: From empirical data to generalizations

George A. Moroz

National Research University Higher School of Economics, Moscow, Russia;
agricolamz@gmail.com

Abstract: This paper describes the syllable structure of Adyghe, a Northwestern Caucasian language spoken in Russia, Turkey and some other countries. The range of possible syllable structures in West Circassian varies from simple structures such as V or CV to complex ones such as CCCVCC or CVCCC. I discovered a considerable variability in interspeaker syllabification judgements during a preliminary experiment. In order to create a generalized model of Adyghe syllable structure I used only word-initial syllable onsets and word-final codas taken from words in Yu. A. Tkharakakho's Adyghe-Russian dictionary. I disregard word-medial consonant clusters because of unclear syllabification rules. I found that the structures of onsets and codas in Adyghe differ, but nearly all attested consonantal classes can occur both in onset and coda positions. Fricatives and plosives show similar distributional patterns, while sonorants differ from obstruents in the following: (a) there are no sonorant clusters; (b) sonorants always occur alone in onsets, so the only type of onset clusters consists of obstruents; (c) there are no syllables with the structure "vowel-sonorant-obstruent", but there are syllables with the structure "vowel-obstruent-obstruent"; (d) complex codas of three segments can only have the structure "sonorant-obstruent-obstruent". I also calculated expected cluster proportions on the assumption that the occurrences of all consonant classes in initial or final position are independent events (i. e. there are no mutual contextual restrictions). Having compared expected and actually observed proportions, I discovered the following: (a) clusters with obstruents are found more often than expected; (b) clusters with sonorants are less frequent than expected; (c) clusters appear more often than expected. These observations allow us to reject the null hypothesis that there are no mutual contextual restrictions on consonant types in initial or final positions.

Keywords: Adyghe, Northwest Caucasian, phonology, syllable

For citation: Moroz G. A. Adyghe syllable structure: From empirical data to generalizations. *Voprosy Jazykoznanija*, 2019, 2: 82–95.

DOI: 10.31857/S0373658X0004303-2.

Acknowledgements: The article was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and funded by the Russian Academic Excellence Project "5-100".

Введение

Адыгейский — это полисинтетический язык, который вместе с кабардино-черкесским, абхазским, абазинским и ныне вымершим убыхским составляет абхазо-адыгскую семью. Адыгские языки (адыгейский и кабардино-черкесский) были распространены на Северном Кавказе, однако в результате русско-кавказской войны достаточно много носителей были вынуждены покинуть родину. По этой причине в настоящее время носители адыгейского языка и их потомки проживают на территории России в Республике Адыгея и в Краснодарском крае, а также в Турции, Сирии, Иордании, Израиле и других странах.

В существующих описаниях адыгейского языка и его диалектов структура слога либо вообще не упоминается [Керашева 1957; Кумахова 1972; Smeets 1984; Höhlig 1997; Ситимова 2004], либо фрагментарно описывается в разделе фонотактики [Яковлев, Ашхамаф 1941: 427–429; Рогава, Керашева 1966: 43–46; Paris 1989: 162]. Единственная работа, в которой перечисляются разные типы слогов, представленные в адыгских языках, — это книга [Балкаров 1970: 50–53]. Однако в этой работе автор ставит своей целью описание не общей структуры слога, а правил сочетания слогов разной структуры внутри слова.

Данная статья решает задачу описания обобщенной структуры слога в адыгейском языке на основе эмпирических данных с использованием автоматического анализа

(см. аналогичный подход в [Žgank et al. 2005; Kishore et al. 2002]). Такое описание на основе обширного языкового материала позволит снизить вероятность того, что какая-то редкая слоговая структура выпадет из поля зрения исследования.

В разделе 1 приводится краткое описание фонологии адыгейского языка. В разделе 2 описаны некоторые сложности, с которыми может столкнуться исследователь при попытке описать слоговую структуру языка. В разделе 3 описывается концепция эмпирического подхода к решению фонологических задач. В разделе 4 предложена обобщенная структура адыгейского слога, выработанная на материале словаря [Тхаркахо 1991]. В разделе 5 мы рассматриваем возможное распределение разных типов согласных (сонорных, неносовых смычных и фрикативных) в инициали и финали, как если бы появление каждого из типов согласных было независимым событием. Мы показываем, что распределение, обнаруженное в наших данных, статистически значимо отличается от ожидаемых значений при гипотезе о независимом распределении перечисленных типов звуков.

1. Фонологическая система адыгейского языка

Адыгейский язык, как и другие языки абхазо-адыгской семьи, обладает богатой консонантной и бедной вокалической системами. В таблице 1 представлена консонантная система адыгейского языка, составленная на основе [Рогава, Керашева 1966: 57]. В таблице использованы знаки международного фонетического алфавита (они отсутствуют в оригинальной работе), в треугольных скобках приводятся соответствующие знаки современной адыгейской орфографии.

В вокалическую систему адыгейского языка традиционно включают три гласных: /a/ <a>, /e/ <э>, /ə/ <ы> [Рогава, Керашева 1966: 21–24]. Остальные гласные, представленные в адыгейском, обычно описываются как сочетание гласного и глайда в позиции после согласного: [i] /jə/ <и>, [e] /jə/ <е>, [æ] /ja/ <я>, [u] /wə/ <у>, [o] /wə/ <о>. После согласного реализация глайда ослабевает или пропадает. Это можно проиллюстрировать морфемой /jə/, обозначающей посессора: [jəfə] /jə-fə/ (POSS-лошадь) ‘его лошадь’ vs. [sifə] /s-jə-fə/ (1SG.IO-POSS-лошадь) ‘моя лошадь’.

2. Описание слоговой структуры

Существует несколько разных формальных методов репрезентации слоговой структуры языка (см., например, обзор [van der Hulst, Ritter 1999]). Мы будем применять классическую систему, используемую, например, в работах [Pike, Pike 1947; Fudge 1969; 1987], в рамках которой выделяется ядро слога (гласный, слоговой сонант или другие единицы), инициаль (согласные, предшествующие ядру) и финаль (согласные, следующие за ядром). Целью статьи является формулирование правил (или ограничений), регулирующих структуру слогового ядра, инициали и финали.

Слоговая структура и принципы деления на слоги значительно варьируются в языках мира. В некоторых языках допускаются лишь слоги одной определенной модели (например, в гавайском разрешены лишь слоги типа CV [Maddieson 2013]). В других языках возможна большая свобода устройства слога.

Принципы деления на слоги можно выявлять разными способами: попросить носителя посчитать слоги, провести между слогами границу, ответить, на каком слоге стоит ударение и др. Таким образом, задачи, которые обычно связывают с понятием слога, могут в принципе по-разному выделять слоги на одном и том же языковом материале. Кроме того, суждения носителей про слогоделение какого-то слова тоже могут различаться:

Таблица 1

Консонантная система адыгейского языка

Способ Место		смычные			фрикативные			аффрикаты			сонор- ные
		глух.	зв.	абр.	глух.	зв.	абр.	глух.	зв.	абр.	
губные	нелаб.	p <п>	b <б>	p' <пI>	f <ф>						w <у, о>; m <м>
	лаб.			p ^w <пIy>							
зубные	нелаб.	t <т>	d <д>	t' <тI>	s <с>	z <з>		ts̃ <ц>	d̃z <дз>	ts' <цI>	n <н>
	лаб.			t ^w <тIy>							
латеральные					l <ль>	ɮ <л>	l' <лI>				
альвеолярные					ʃ <ш>	ʒ <ж>		tʃ̃ <ч̃>	d̃ʒ <дж>	tʃ' <чI>	r <р>
шипяще- свистящие	нелаб.				ʃ̃ <ш̃>	ʒ̃ <ж̃>	ʃ' <шI>				
	лаб.				ʃ ^w <ш̃y>	ʒ ^w <ж̃y>	ʃ ^w <шIy>	tʃ̃ ^w <ч̃y>	d̃ʒ ^w <д̃зy>		
палатальные					ç <щ>	ʒ <жь>		t̃ç <ч̃>	d̃ʒ <дж>	t̃ç' <чI>	j <й>
велярные	нелаб.	k <к>	g <г>	k' <кI>	x <х>						
	лаб.	k ^w <кy>	g ^w <гy>	k ^w <кIy>							
увулярные	нелаб.	q <къ>			χ <х̃>	ʁ <г̃>					
	лаб.	q ^w <к̃y>			χ ^w <х̃y>	ʁ ^w <г̃y>					
ларин- гальные	нелаб.	ʔ <I>			h <х̃>						
	лаб.	ʔ ^w <Iy>									

например, в работе [Duanmu 2009: 1–2] приводятся примеры расхождений в суждениях разных лингвистов относительно числа слогов в английских словах типа *hour, flour, flower, shower*, а также того, где проходит слоговая граница в словах типа *happy* и др.

В работе [Кодзасов, Муравьева 1980] было предложено выделять разные типы языков в зависимости от того, насколько важную роль слог играет в фонологии и фонетике данного языка. В одних языках слог характеризуется пограничными сигналами, маркирующими его границы («квантовые» языки), а в фонологии других языков слог не играет какой-либо важной роли, в связи с чем его граница часто никак не маркируется ни в звуковом сигнале, ни в фонологических правилах языка («волновые» языки). В волновых языках носители затрудняются определить границу слогов, в результате чего можно наблюдать вариативность в слогеделении. Согласно работе [Князев 1999], к таким языкам можно отнести русский: носители могут делить слово *кошка* на слоги двумя способами: *кош-ка* и *ко-шка*. По данным предварительного эксперимента с четырьмя носителями адыгейского языка (школьные учителя старше 40 лет), наблюдалась вариативность

суждений о слоговой границе. Например, слово /natrəf/ <натрыф> ‘кукуруза’ носители делили двумя способами: /nat-rəf/ и /na-trəf/. Аналогичное деление было предложено для слова /q^wəp^šhe/ <къупшъхъ> ‘кость’: /q^wə-p^šhe/ и /q^wəp-^šhe/. Тем самым, мы можем предположить, что адыгейский, как и русский, является языком, где граница слога может оставаться неопределенной («волновым» в терминологии [Кодзасов, Муравьева 1980]).

Принимая во внимание возможную «волновую» природу адыгейского языка, при построении фонотактической модели мы не можем опираться на интервокальные кластеры согласных, однако их можно использовать для проверки полученной модели. В связи с этим мы анализировали следующие типы данных:

- 1) инициали слогов, с которых начинаются слова;
- 2) финали слогов, на которые заканчиваются слова;
- 3) консонантные кластеры в середине слов.

Инициали первого и финали последнего слога были собраны на основе данных адыгейско-русского словаря [Тхаркахо 1991], доступного в электронном формате GoldenDict [Команда GoldenDict 2016].

3. Методология

В ходе исследования использовался разработанный нами алгоритм, который позволяет частично автоматизировать анализ слоговой структуры. Алгоритм состоит из следующих этапов:

1. На основании односложных слов выделить единицы, которые могут составлять ядро слога.
2. На основе всего имеющегося лексического материала выделить все возможные инициали первых слогов слов.
3. На основе всего имеющегося лексического материала выделить всевозможные финали последних слогов слов.
4. Составить список всех возможных комбинаций выделенных инициалей и финалей и проверить, какие из этих комбинаций представлены в односложных словах.
5. На основе полученных данных составить общую модель слога.
6. Протестировать полученную модель на всем материале, в том числе применительно к многосложным словам. Если в ходе тестирования обнаруживаются слова, в которых интервокальные кластеры нельзя объяснить полученной моделью, значит, нужно посмотреть на особые случаи, которые составляют исключения, и изменять модель с их учетом.
7. (Факультативный этап.) Проверить распределение разных типов слогов в разных позициях слова, выяснить, не наблюдается ли тяготение некоторых типов слогов к позиции в конце, в начале или в середине слова.

Данный алгоритм потенциально можно применить к любым языковым единицам, однако при использовании в анализе всего фонологического инвентаря языка получится слишком большое число потенциальных слогов. Чтобы избежать этой проблемы, исследователи обычно делят все звуки на гласные и согласные (например, такое можно видеть в работе [Балкаров 1970: 50–53]) и описывают слоговую структуру исключительно в этих терминах. Мы в нашей работе будем использовать более дробное деление согласных на шумные и сонорные.

Стоит отметить недостатки используемого метода получения данных. В словаре все лексемы записаны в некоторой начальной форме. В случае имен это форма абсолютива, которая не имеет окончания, а вот в случае глаголов это форма с модальным показателем

/-n/ <-н>, что дает в результате необычно большое по сравнению с другими типами сегментов количество сонорных в финали. Если бы глагольные леммы давались в словаре без окончания (например, в форме императива), то они могли бы иметь и консонантный, и вокалический исход. Таким образом, рассматривая данные, которые получены про финали слогов, нужно учитывать морфологически обусловленный перевес форм на /-n/ <-н>.

4. Результаты анализа адыгейско-русского словаря

В исследовании был использован материал адыгейско-русского словаря [Тхаркахо 1991], содержащего более 10 тыс. слов (некоторые заимствования, в основном русские, мы из рассмотрения исключили). Этот словарь в два раза меньше словаря [Шаов 1975], однако именно он был оцифрован и выложен в формате GoldenDict [Команда GoldenDict 2016], поэтому выбор словаря можно обосновать именно доступностью материалов в электронном виде.

Результаты представлены на рис. 1 и ниже в таблице 3. Как видно, чем меньше кластер, тем чаще он встречается в инициали или в финали слов. Односегментные финаль и инициаль встречаются чаще, чем двухсегментные, а трехсегментный кластер встречается реже всего. Также из рис. 1 понятно, что некоторые кластеры встречаются преимущественно в инициали, а некоторые — преимущественно в финали. Кластеры или одиночные сонорные тяготеют к финали, а остальные — к инициали. На рис. 1 видно, что одиночный шумный чаще встречается в инициалах, а одиночный сонорный — в финалях слогов. Кроме того, нулевые финали встречаются чаще, чем нулевые инициали.

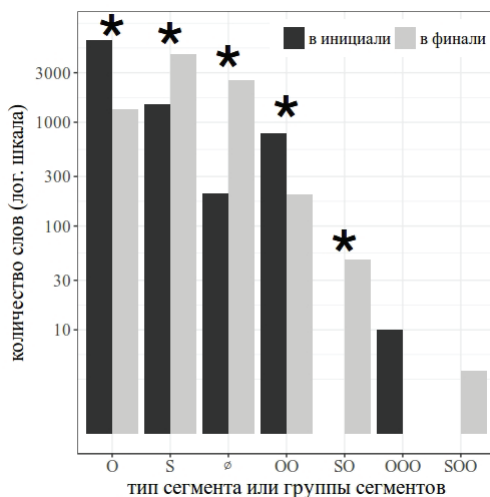


Рис. 1. Результаты анализа словаря [Тхаркахо 1991]. С — сонорные, О — шумные, Ø — отсутствие согласных (т. е. число неприкрытых слогов в случае инициали, и число открытых слогов в случае финали; звездочкой обозначено наличие статистической значимости различий)

Чтобы ответить на вопрос, является ли полученное распределение типов согласных между инициалами и финалями случайностью или оно отражает некоторую языковую тенденцию, мы проверили статистическую значимость таких предпочтений при помощи теста хи-квадрат (результаты каждого теста приведены в таблице 3). Мы исходили

из нулевой гипотезы, согласно которой распределение звуков каждого типа равномерно в инициалах и финалях. Обосновать такую нулевую гипотезу можно отсутствием в теории слога каких-либо предположений насчет распределения разных типов согласных в разных частях слога. Единственная теория, претендующая на описание законов, согласно которым разные типы сегментов появляются в слоге — иерархия сонорности (формулировалась много раз начиная с XIX в., историю вопроса см., например, в [Clements 1990]). Эта теория предсказывает распределение разных типов сегментов в инициали или финали (например, сонорные имеют тенденцию занимать позицию ближе к гласному, чем шумные), в то время как мы делаем утверждение про распределение сегментов одного типа в инициалах и финалях. Теория сонорности не предсказывает полученные нами тенденции шумных и их кластеров находиться в инициали слога, а сонорных в финали. Результаты тестов отражены на рис. 1. Наличие звездочки над соответствующим типом консонантного кластера указывает на статистически значимую разницу между числом случаев в инициали и финали.

Разные виды кластеров (шумный, комбинация шумных, сонорный, открытый/неприкрытый слог и т. д.) встречаются в наших данных с разной частотностью. На рис. 1 для каждого из типов кластеров приводится число вхождений в словаре. В большинстве случаев тест хи-квадрат подтверждает, что разница между частотностью одного и того же вида кластера в инициали и в финали статистически значима. Тест не применялся лишь для двух видов трехсегментных кластеров, у которых из-за малого числа наблюдений ожидаемые значения получаются меньше 5 (т. е. применение теста хи-квадрат к этим данным некорректно, а обычная альтернатива хи-квадрату — тест Фишера — неприменим к таблицам меньше, чем 2 на 2).

Особенно следует отметить поведение консонантных кластеров, содержащих сонорные согласные. Во-первых, сонорные не встречаются в кластерах согласных в инициали. Во-вторых, в тех случаях, когда сонорный встречается в консонантных кластерах в финали (т. е. в 50 примерах, около 0,006 % всех данных), он всегда находится ближе к ядру слога, чем шумные, что согласуется с иерархией сонорности.

Полученные наблюдения над типами кластеров и составляющими слога (инициаль или финаль) можно визуализировать при помощи графика «биplot», который строится на основе метода главных компонент (см., например, [Husson et al. 2017]). Данный метод обычно используют для уменьшения размерности многомерных данных путем поиска гиперплоскостей, описывающих максимальную дисперсию данных. Эта визуализация позволяет изменить оси координат и представить на одном графике и исследуемые переменные, и наблюдения. На рис. 2 визуальны представлены результаты анализа словаря [Тхаркахо 1991] при помощи метода главных компонент. По осям отложены значения в новой искусственной системе координат (формируемые первой и второй главными компонентами). Рядом с каждой осью подписан процент дисперсии, объясняемой данной компонентой. Так как в наших данных всего две переменных, сумма объясняемой дисперсии равна 100 %. Стрелки обозначают старые оси координат (конструируемые на основе переменных «инициаль» и «финаль»). Буквы обозначают позицию наших наблюдений в новой системе координат. При применении метода главных компонент значения, полученные в новой системе координат, нельзя интерпретировать. Поэтому при анализе результатов, полученных этим методом, следует опираться на расположение наблюдений (O, S и т. д.) относительно переменных (старые оси координат «инициаль» и «финаль»). Анализируя расстояния между наблюдениями и переменными, представленными на рис. 2, можно отметить то же самое, что было видно при анализе рис. 1. Одиночные шумные тяготеют к инициали, одиночные сонорные и части слога, не содержащие согласных, тяготеют к финали. Остальные типы кластеров встречаются значительно реже, так что нельзя выделить тенденцию тяготения к той или иной позиции. Также на графике можно видеть некоторую тенденцию кластеров из двух шумных к инициали. График «биplot» на рис. 2 имеет преимущество перед диаграммой на рис. 1, так как сводит

в одной системе координат и отношение между наблюдаемыми кластерами согласных, и отношение их к позиции в слоге.

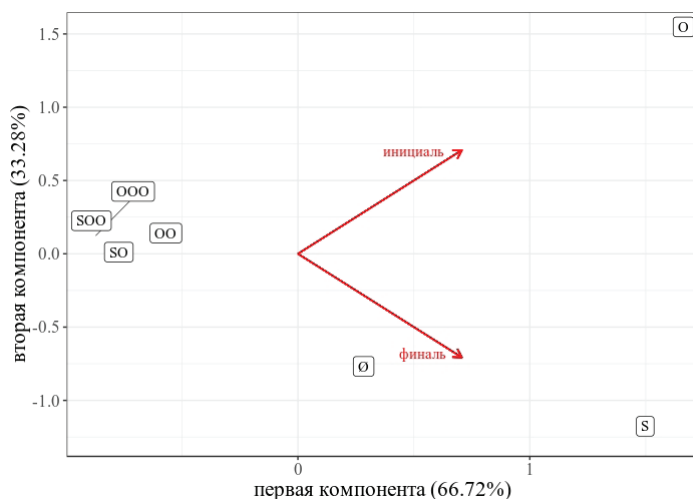


Рис. 2. Визуализация результатов анализа словаря [Тхаркахо 1991] методом главных компонент. Обозначения те же, что и на предыдущем рисунке.

Выше мы представили отдельные наблюдения относительно начальных и конечных фрагментов слога, не объединяя их в общую модель. При составлении такой модели мы можем ожидать два сценария: либо все типы инициалей и финалей могут сочетаться друг с другом, либо же существуют ограничения на сочетания каких-то типов финалей и инициалей. Предположим, что все типы инициалей могут сочетаться со всеми типами финалей, тогда все наши наблюдения об инициалах и финалях можно совместить в единой схеме:

$$(1) \quad (O)(O)(O) \begin{matrix} (S) \\ V \end{matrix} (S)(O)(O)$$

Схема разделена на три части: в первой части представлены все возможные инициали (в том числе без согласных). Скобки окружают факультативный элемент, т. е. запись «(S)» обозначает одновременно возможность и наличия одного сонорного элемента, и его отсутствия. Во второй части представлен единственный обязательный, по нашему мнению, элемент адыгейского слога — гласный. В третьей части представлено обобщенное описание финали (включающее пустую финаль).

Данной схеме соответствует 30 возможных типов слогов, однако не для всех нам удалось подобрать пример односложного слова в словаре. Не все инициали зафиксированы в сочетании со всеми финалями, а многосложные слова не допускают однозначной интерпретации их слоговой структуры.

Таблица 2

Порожденные моделью слоговые структуры с примерами слов

№	Пример			Слоговая структура		
	Орфография	Транскрипция	Глосса	Инициаль	Ядро	Финаль
1	а	а:	тот		V	
2	и	jə	восемь	S	V	
3	пэ	pɛ	нос	O	V	

№	Пример			Слоговая структура		
	Орфография	Транскрипция	Глосса	Инициаль	Ядро	Финаль
4	бгы	bgə	талия	OO	V	
5	пскэ	pskɐ	кашель	OOO	V	
6	ар	aːr	тот.ABS		V	S
7	ин	jə-n	мазать-MOD	S	V	S
8	тын	tə-n	платить-MOD	O	V	S
9	стын	stə-n	гореть-MOD	OO	V	S
10	стхъэн	stɣɛ-n	загубеть-MOD	OOO	V	S
11	аркъ	arq	водка		V	SO
12	оркъ	wɛrɣ	дворянин	S	V	SO
13	куанд	kʷaːnd	куст	O	V	SO
14	пклантI	pɬɛˈaːntʰ	пот	OO	V	SO
15				OOO	V	SO
16					V	SOO
17				S	V	SOO
18	пцэкъэнтф	pɬɛqɛntf	удочка	O	V	SOO
19				OO	V	SOO
20				OOO	V	SOO
21	ащ	aːɕ	тот.OBL		V	O
22	ежъ	jɛʒ	сам	S	V	O
23	кIыб	ɬɛˈəb	спина	O	V	O
24	тхылъ	tɣəl	книга	OO	V	O
25	пскищ	psk-i-ɕ	кашель-LNK-3	OOO	V	O
26	ахъщ	aːɣɕ	деньги		V	OO
27	нэпх	nɛpɣ	гипноз	S	V	OO
28	хъапс	ɬaːps	тюрьма	O	V	OO
29	щхэпс	ɕɣɛps	клей	OO	V	OO
30	пскэщт	pskɛ-ɕt	кашлять-FUT	OOO	V	OO

Как видно из таблицы 2, пять гипотетических типов слогов не встретились в адыгейском словаре (см. строки 15–17, 19–20). Все типы не встретившихся в словаре слогов, кроме одного случая (строка 15: OOOVSO), содержат сложную финаль -SOO. Значит, нам следует изменить схему в (1), чтобы модель не порождала структуры, которые не представлены в словаре. Это возможно, если три слова с кластером -SOO, которые все же встретились в адыгейском словаре, отнести к исключениям, а схему изменить:

$$(2) \quad \begin{matrix} (O) & (S) \\ (O)(O)(O) & V & (S) \\ & & (O) \end{matrix} \quad (O)$$

Под эту модель попадают все перечисленные в таблице 2 типы слогов за исключением слогов с финалью -SOO. Описание станет полным, когда мы добавим ограничение на логи структуры OOOVSO.

Важно отметить, что отсутствие какой-то слоговой структуры в словаре (сколь угодно большом) необязательно свидетельствует о том, что эта слоговая структура запрещена в языке. Возможно, соответствующие типы инициали и финали редки сами по себе, и уж тем более редко встречаются вместе. Рис. 1 и таблица 3 показывают, что это именно так: VSO встречается 47 раз из 8622, а трехконсонантная инициаль OOO встречается 9 раз

из 8620. Если бы появление в одном слоге инициали ООО и финали SO было независимым событием, мы бы ожидали, что оно встретится в нашем словаре значительно меньше одного раза ($47 : 8622 \cdot 9 : 8620 \cdot 8620 = 0,04906054$). Оказывается, что отсутствие слоговых структур ОООVSO можно объяснять не фонотактическим запретом адыгейского языка, а следствием низких частотностей финалей и инициалей, входящих в данные слоговые структуры. Наверное, невозможно показать, что первичнее: фонотактические ограничения или частотные распределения разных типов кластеров. Следует различать фонотактический запрет на некоторую комбинацию сегментов внутри слога, которая в языке будет сопровождаться фонологическими правилами, «исправляющими» эту структуру (например, в заимствованиях), и фонотактически редкие комбинации сегментов. Используемая нами методология никак не различает эти две ситуации.

Таблица 3

Сравнение разных типов инициалей и финалей по частотности

Тип сегмента / конс. кластера	В инициали	В финали	Результат теста хи-квадрат	Остатки
неприкр. / откр. слог	204	2522	$\chi^2 = 1971,1$; $df = 1$; $p\text{-value} < 2,2e-16$	31,39319
О	6130	1338	$\chi^2 = 3074,9$; $df = 1$; $p\text{-value} < 2,2e-16$	39,21025
S	1498	4511	$\chi^2 = 1510,8$; $df = 1$; $p\text{-value} < 2,2e-16$	27,4842
ОО	512	133	$\chi^2 = 222,7$; $df = 1$; $p\text{-value} < 2,2e-16$	10,55223
SO	0	47	$\chi^2 = 47$; $df = 1$; $p\text{-value} = 7,099e-12$	4,84768
ООО	9	0	неприменим	—
SOO	0	3	неприменим	—
Итого	8353	8554		

Таким образом, от наблюдений о составе инициалей и финалей мы перешли к обобщенной модели адыгейского слога, которая покрывает 25 типов слоговой структуры. При этом все типы слогов, представленные в адыгейском языке (кроме трех исключений с финалью -SOO), данной схемой описываются, и лишь один тип слога (ООOVSO), который возможен согласно этой модели, в исследуемом материале не встретился.

В ходе морфологической деривации полученные формы следуют указанной слоговой структуре, но есть несколько аффиксов, которые состоят лишь из согласных. От них можно было бы ожидать усложнения слоговой структуры. Например, в бесленеевском диалекте кабардино-черкесского языка возможно присоединить непосредственно к корню аффикс темпорального сдвига, образуя кластер из трех гоморганных сегментов в финали:

(3) сә-t-t'-t

LOC-1PL.ERG-копать-TMP

‘Мы это там копали’¹.

Однако в случае адыгейского все аффиксы либо имеют вид CV, либо состоят из одного / двух согласных (абсолютив /-r/ <-p>, эргатив /-m/ <-m>, модальный суффикс /-n/ <-n> и показатель будущего времени /-сt/ <-шт>), но тогда в позиции после согласного требуют перед собой вставки эпентетического гласного. Следовательно, усложнения финали при присоединении морфем, состоящих из согласных, не происходит. Из всего этого следует,

¹ Автор благодарен П. М. Аркадьеву за этот пример.

что выделенная модель описывает все потенциальные формы адыгейского языка. Как говорилось выше, все полученные частоты финалей смещены из-за того, что при подсчете использовались не все морфологические формы, а глагольная форма с модальным показателем /-п/ <-н> использовалась, наоборот, необычно часто. Если бы при анализе использовалась не эта форма, а глагольная основа, то часть финалей, которые на рис. 1 попали в категорию S, оказались бы в категориях O и Ø.

5. Случайны ли распределения согласных в кластерах?

Из приведенных в предыдущем разделе данных видно, что разные типы согласных встречаются в кластерах с разной частотой. Из того, что кластер ОО встретился в инициали 512 раз и лишь 133 раза в финали, можно ли, например, сделать вывод о том, что он тяготеет к позиции в инициали? Шумные сами по себе заметно реже встречаются в финали (ср. таблицу 4), так что свойство кластера ОО чаще встречаться в инициали, чем в финали, может вытекать из свойств составляющих его согласных. Чтобы исключить эту гипотезу и показать, что речь идет именно о дистрибутивных свойствах кластеров, начнем с того, что выясним, сколько раз единицы каждого типа встретились в инициали и финали.

Таблица 4

Таблица сопряженности для параметров типа согласного и позиции в слоге

	Кол-во в инициали	Доля в инициали	Кол-во в финали	Доля в финали
Неприкрытый / открытый слог	204	0,02	2522	0,29
S	1498	0,17	4561	0,53
O	7181	0,81	1208	0,19

Рассмотрим нулевую гипотезу, согласно которой появление каждого типа сегмента в инициали / финали является независимым. Так как число наблюдений достаточно велико, будем считать, что наблюдаемые доли являются приближениями вероятностей соответствующих независимых событий. Т. е. если доли, представленные в таблице 3, являются вероятностями встретить соответствующие сегменты в инициали или финали единиц словаря, то мы предполагаем, что появление того или иного сегмента в инициали или финали не зависит от появления другого сегмента в той же инициали или финали. Тогда мы можем посчитать вероятности совместного появления двух сегментов в одной инициали или финали, перемножив соответствующие вероятности и вероятность встретить кластер соответствующего размера.

В таблице 5 приведены ожидаемые и наблюдаемые значения в инициали и финали, а также результаты теста хи-квадрат, сравнивающего эти значения. Для инициалей и финалей тест применялся по отдельности. Полученные вероятности встретить те или иные два сегмента вместе были нормализованы (столбец «нормализованные вероятности») и использовались для подсчета ожидаемых соотношений всех типов кластеров (столбец «ожидаемые значения»), если считать появление того или иного типа сегмента независимым событием. Таким образом, мы ожидали бы, что кластеры SO или OS встретились бы нам в инициали приблизительно 163 раза, кластеры, в которых представлен исключительно S, встретились бы 10 раз и т. д. В столбце «наблюдаемые значения» приведены значения, которые мы получили из данных. Кроме того, в таблице приводятся значения теста хи-квадрат.

Таблица 5

**Ожидаемые и наблюдаемые распределения классов сегментов
в инициали и финали, а также значения тестов хи-квадрат**

В инициали ($\chi^2 = 209,73$; $df = 2$; $p\text{-value} < 2,2e-16$)				
	Вероятность	Нормализованная вероятность	Ожидаемое значение	Наблюдаемое значение
P(S, O)	0,02 (0,17 · 0,81 · 0,06 · 2)	0,32 = 0,02 : (0,02 + 0,001 + 0,04)	163,84 (512 · 0,32)	0
P(S, S)	0,001 (0,17 · 0,17 · 0,06)	0,02 = 0,001 : (0,02 + 0,001 + 0,04)	10,24 (512 · 0,02)	0
P(O, O)	0,04 (0,81 · 0,81 · 0,06)	0,66 = 0,04 : (0,02 + 0,001 + 0,04)	20,48 (512 · 0,66)	512
В финали ($\chi^2 = 203,74$; $df = 2$; $p\text{-value} < 2,2e-16$)				
	Вероятность	Нормализованная вероятность	Ожидаемое значение	Наблюдаемое значение
P(S, O)	0,004 (0,53 · 0,19 · 0,02 · 2)	0,37 = 0,004 : (0,004 + 0,006 + 0,0007)	66,6 (180 · 0,37)	47
P(S, S)	0,006 (0,53 · 0,53 · 0,02)	0,56 = 0,006 : (0,004 + 0,006 + 0,0007)	100,8 (180 · 0,56)	0
P(O, O)	0,0007 (0,19 · 0,19 · 0,02)	0,07 = 0,0007 : (0,004 + 0,006 + 0,0007)	12,6 (180 · 0,07)	133

Тест хи-квадрат показал, что наблюдаемые значения статистически значимо отличаются от ожидаемых значений. Полученные результаты говорят о том, что нулевую гипотезу о независимости появления каждого типа согласных можно отвергнуть. Шумные согласные в адыгейском языке тяготеют к совместному появлению в инициали или финали больше, чем к появлению вместе с сонорным. Большой объем исследуемого словаря позволяет не только распространить полученные результаты на весь словарный материал языка, но и с некоторыми оговорками сделать выводы об особенностях адыгейской фонотактики в целом.

6. Заключение

В данной статье на основе словаря [Тхаркахо 1991] мы рассмотрели слоговую структуру адыгейского языка и предложили схему ее описания (см. (2)). Мы показали, что сонорные входят в состав консонантных кластеров реже, а шумные — чаще, чем ожидалось бы, если бы это были независимые события. В нашей работе мы предлагаем отдельно исследовать инициали, финали и их сочетаемость, таким образом выделяя подзадачи при исследовании слоговой структуры. Проведенное исследование показывает, что структура слога может изучаться и в языках «волнового» типа (без пограничных сигналов структуры слога) — за счет анализа структуры абсолютных инициалей и финалей.

Кроме того, наше исследование показало, что слоговую структуру можно описывать, автоматически извлекая материал из словарей. Это важно, так как часто в работах не указывается объем и природа обработанного материала и не приводятся списки слов, на основе которых получено то или иное фонотактическое описание. Конечно, примененный в статье метод содержит ряд недостатков. Во-первых, нужно делать поправки на интервокальные кластеры и следить, чтобы они также описывались полученной моделью. Во-вторых, необходимо учитывать морфологические формы, в которых леммы даются в словаре. Обилие

определенных форм в словаре может давать завышенное значение для какого-то типа слов. В-третьих, нужно делать поправки на отсутствующие в словаре формы и проверять, не получится ли незасвидетельствованная слоговая структура при образовании той или иной морфологической формы. В-четвертых, стоит уделить внимание частотности: те частотности, которые мы получим при анализе словаря, могут отличаться от тех, которые могут быть получены при анализе корпуса текстов.

Деление на типы сегментов, которое мы использовали, позволяет получить описание адыгейского слога. Однако стоит отметить, что при описании фонотактики можно использовать и более дробную систему признаков, которая позволила бы исследовать большее разнообразие слоговых структур.

СПИСОК СОКРАЩЕНИЙ

ABS — абсолютив	O — шумный согласный
ERG — эргатив	OVL — косвенная форма
FUT — будущее время	PL — множественное число
IO — не прямой объект	POSS — посессивный показатель
LNK — линкер (соединительный элемент между именем и числительным)	S — сонорный согласный
LOC — локативный преверб	SG — единственное число
MOD — модальный суффикс	TMP — временной сдвиг
Ø — отсутствие согласного	V — гласный

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Балкаров 1970 — Балкаров Б. Х. *Фонетика адыгских языков (синхронно-диахронное исследование)*. Нальчик: Книжное изд-во “Эльбрус”, 1970. [Balkarov B. Kh. *Fonetika adygskikh yazykov (sinkhronno-diakhronnoe issledovanie)* [Phonetics of Adyghean languages (A synchronic-diachronic study)]. Nalchik: El’brus Publ., 1970.]
- Керашева 1957 — Керашева З. И. *Особенности шапсугского диалекта адыгейского языка*. Майкоп: Адыгейское книжное изд-во, 1957. [Kerasheva Z. I. *Osobennosti shapsugskogo dialekta adygeiskogo yazyka* [Features of the Shapsug dialect of Adyghe]. Maykop: Adyghe Publ., 1957.]
- Князев 1999 — Князев С. В. О критериях слогоделения в современном русском языке: теория волны сонорности и теория оптимальности. *Вопросы языкознания*, 1999, 1: 84–102. [Kniazev S. V. On the criteria of syllabization in contemporary Russian: the sonority wave and the theory of optimal value. *Voprosy Jazykoznanija*, 1999, 1: 84–102.]
- Кодзасов, Муравьева 1980 — Кодзасов С. В., Муравьева И. А. Слог и ритмика слова в алюторском языке. *Публикации отделения структурной и прикладной лингвистики МГУ. Филологический факультет*. Вып. 9. М.: МГУ, 1980, 103–127. [Kodzasov S. V., Murav’eva I. A. Syllable and word rhythemics in Alutor. *Publikatsii otdeleniya strukturnoi i prikladnoi lingvistiki MGU. Filologicheskii fakul’tet*. No. 9. Moscow: Lomonosov Moscow State Univ., 1980, 103–127.]
- Команда GoldenDict 2016 — Команда GoldenDict. *Компьютерная программа AdygeDict*, v. 1.1. 2016. [The GoldenDict Team. *AdygeDict Software*, v. 1.1. 2016.]
- Кумахова 1972 — Кумахова З. Ю. *Абадзехский язык и его место среди других адыгских диалектов*. Нальчик: Книжное изд-во “Эльбрус”, 1972. [Kumakhova Z. Yu. *Abadzekhskii yazyk i ego mesto sredi drugikh adygskikh dialektov* [The Abadzekh language and its place among other Adyghean dialects]. Nalchik: El’brus Publ., 1972.]
- Рогав, Керашева 1966 — Рогав Г. В., Керашева З. Ю. *Грамматика адыгейского языка*. Краснодар; Майкоп: Краснодарское книжное изд-во, 1966. [Rogava G. V., Kerasheva Z. Yu. *Grammatika adygeiskogo yazyka* [Adyghe grammar]. Krasnodar; Maykop: Krasnodar Publ., 1966.]
- Ситимова 2004 — Ситимова С. С. Особенности бжедугского диалекта адыгейского языка. Дис. ... докт. филол. наук. Майкоп: Адыгейский государственный университет, 2004. [Sitimova S. S. *Osobennosti bzhedugskogo dialekta adygeiskogo yazyka* [Features of the Bzhedug dialect of Adyghe]. Doctoral diss. Maykop: Adyghe State Univ., 2004.]

- Тхаркахо 1991 — Тхаркахо Ю. А. *Адыгейско-русский словарь*. Майкоп: Адыгейское книжное изд-во, 1991. [Tkharakho Yu. A. *Adygeisko-russkii slovar'* [Adyge-Russian dictionary]. Maykop: Adyge Publ., 1991.]
- Шаов 1975 — Шаов Ж. А. *Адыгейско-русский словарь*. Майкоп: АНИИ ЭЯЛИ, 1975. [Shaov Zh. A. *Adygeisko-russkii slovar'* [Adyge-Russian dictionary]. Maykop: Adyge Research Institute for Economics, Language, Literature and History, 1975.]
- Яковлев, Ашхамаф 1941 — Яковлев Н. Ф., Ашхамаф Д. А. *Грамматика адыгейского литературного языка*. М.; Л.: Изд-во АН СССР, 1941. [Yakovlev N. F., Ashkhamaf D. A. *Grammatika adygeiskogo literaturnogo yazyka* [Standard Adyge grammar]. Moscow; Leningrad: Academy of Sciences of the USSR Publ., 1941.]
- Clements 1990 — Clements G. N. The role of the sonority cycle in core syllabification. *Papers in laboratory phonology*, 1990, 1: 283–333.
- Duanmu 2009 — Duanmu S. *Syllable structure: The limits of variation*. Oxford: Oxford Univ. Press, 2009.
- Fudge 1969 — Fudge E. Syllables. *Journal of Linguistics*, 1969, 5: 253–287.
- Fudge 1987 — Fudge E. Branching structure within the syllable. *Journal of Linguistics*, 1987, 23: 359–377.
- Höhlig 1997 — Höhlig M. *Kontaktbedingter Sprachwandel in der adygeischen Umgangssprache im Kaukasus und in der Türkei: vergleichende Analyse des russischen und türkischen Einflusses in mündlichen adygeischen Texten*. München: Lincom Europa, 1997.
- Husson et al. 2017 — Husson F., Lê S., Pagès J. *Exploratory multivariate analysis by example using R*. Boca Raton: CRC Press, 2017.
- Kishore et al. 2002 — Kishore S. P., Kumar R., Sangal R. A data driven synthesis approach for Indian languages using syllable as basic unit. *Proc. of Intl. Conf. on NLP (ICON)*, 2002: 311–316.
- Maddieson 2013 — Maddieson I. Syllable Structure. *World Atlas of Language Structures Online*. Dryer M. S., Haspelmath M. (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: <http://wals.info/chapter/12>. Accessed on 2019-01-31.
- Paris 1989 — Paris C. Esquisse grammaticale du dialecte abzakh (tcherkesse occidental). *The indigenous languages of the Caucasus*. Vol. 2. Hewitt B. G. (ed.). Delmar: Caravan Books, 1989, 157–260.
- Pike, Pike 1947 — Pike K., Pike E. Immediate constituents of Mazateco syllables. *International Journal of American Linguistics*, 1947, 13: 78–91.
- Smeets 1984 — Smeets R. *Studies in West Circassian phonology and morphology*. Leiden: The Hakuchi Press, 1984.
- van der Hulst, Ritter 1999 — van der Hulst H., Ritter N. A. Theories of the syllable. *The syllable: Views and facts*. Vol. 45. Van der Hulst H., Ritter N. A. Berlin: Mouton de Gruyter, 1999.
- Žgank et al. 2005 — Žgank A., Horvat B., Kačić Z. Data-driven generation of phonetic broad classes, based on phoneme confusion matrix similarity. *Speech Communication*, 2005, 47(3): 379–393.

Получено / received 19.06.2017

Принято / accepted 17.04.2018