

ИССЛЕДОВАНИЕ ТИПОЛОГИЧЕСКОЙ БЛИЗОСТИ ПАРАМЕТРИЧЕСКИХ ЯДЕР КАРАЧАЕВО-БАЛКАРСКОГО И ДРУГИХ ТЮРКСКИХ ЯЗЫКОВ

И. Д. Семенова

Воронежский государственный университет

Поступила в редакцию 26 января 2017 г.

Аннотация: в статье исследуется типологическая близость параметрических ядер карачаево-балкарского и других тюркских языков: тофаларского, турецкого и узбекского.

Ключевые слова: лексико-семантическая типология, параметрические ядра лексики карачаево-балкарского, тофаларского, турецкого и узбекского языков.

Abstract: the paper presents results of the research of typological proximity of the parametric cores of Karachay-Balkar and other Turkic languages: Tofalar, Turkish and Uzbek.

Key words: lexical-semantic typology, parametric cores of Karachay-Balkar, Tofalar, Turkish and Uzbek vocabulary.

В настоящем исследовании для сравнения с карачаево-балкарским языком, представляющим *китчакскую* группу, взяты параметрические ядра (ПЯ) лексики трех тюркских языков: из *огузских* – турецкого [1], из *карлукских* – узбекского [2], из *саянских* – тофаларского [3].

Для обеспечения максимальной полноты данных методика сравнения была нами модифицирована по сравнению с использовавшейся ранее [4]: дефиниции значений слов сравниваемых языков были разделены на полнозначные составляющие, что обеспечило учет всех метаслов, независимо от порядка их следования и количества слов в дефиниции (метаязыковой части тюркско-русского словаря).

Эту методику уместно назвать *упрощенной* (и, если здесь позволительна метафора, – семантически *укрупненной*).

При сравнении значений нами не рассматривались служебные слова и несловарные словоформы, а также такие дискурсивные и грамматикализованные значения, как ‘быть’, ‘делать’, ‘содержащий’, фазовые метаслова ‘кончать-кончить’, родовое мериологическое значение ‘часть’ и т. д.

Поскольку частота появления метаслова в нашем материале может рассматриваться как показатель грамматичности, считаем полезным привести эти «грамматические» (или «технические») метаслова с их частотами: ‘в’ – 9216, ‘о’ – 5929, ‘на’ – 4624, ‘быть’ – 2601, ‘из’ – 1296, ‘для’ – 1024, ‘делать’ – 900, ‘с’ – 784, ‘часть’ – 576, ‘вид’ – 529, ‘или’ – 484, ‘не’ – 441, ‘и’ – 400, ‘становиться’ – 400, ‘давать’ – 361, ‘имеющий’ – 225, ‘род’ – 196, ‘без’ – 121, ‘к’ – 121, ‘у’ – 100, ‘за’ – 81, ‘по’ – 64, ‘от’ – 49, ‘во’ – 36, ‘как’ –

36, ‘при’ – 36, ‘об’ – 25, ‘весь’ – 16, ‘до’ – 16, ‘кончать’ – 16, ‘кончить’ – 16, ‘со’ – 16, ‘что’ – 16, ‘всех’ – 9, ‘каждый’ – 9, ‘так’ – 9, ‘то’ – 9, ‘вниз’ – 4, ‘где-л.’ – 4, ‘навстречу’ – 4, ‘после’ – 4, ‘равно’ – 4, ‘части’ – 4.

Для слияния было взято 6807 карачаево-балкарских слов с максимальным параметрическим весом, 695 тофаларских, 1057 турецких и 1535 узбекских. Для карачаево-балкарского ядра взято в 6–7 раз больше слов, чем для остальных языков.

Если каждый из родственных языков представлен параметрическим ядром, близким к 1000 слов, то для карачаево-балкарского языка взято ПЯ Большого словаря. Сделано это для обеспечения максимально полного сравнения карачаево-балкарской лексики с лексикой ПЯ других тюркских языков.

Такая диспропорция не должна оказать существенного влияния на результаты исследования, потому что *Коэффициент Типологической Близости* (КоТиБ) параметрических ядер тюркских языков вычисляется по формуле, учитывающей размеры сравниваемых ядер:

$$\text{КоТиБ} = \frac{L_1 \cap L_2}{0,5(L_1 + L_2)},$$

где L_1 – сравниваемое ПЯ первого языка; L_2 – сравниваемое ПЯ второго языка. Символ \cap в данной формуле имеет не совсем традиционное значение: он указывает на количество общих для сравниваемых множеств лексико-семантических пар.

Поскольку единицей измерения типологической близости параметрических ядер тюркских языков является *лексико-семантическая пара*, есть необходимость раскрыть содержание этого термина.

К примеру, у нас есть изосемантический, т. е. соотнесенный с одним и тем же русским словом (мета-

словом), ряд: карабалк. *кёк_II* 'небо', тофалар. *бұт-_{II}* 'небо', турецк. *gök* 'небо' и узбек. *falak* 'небо'. Эти 4 слова образуют 6 пар: 1) карабалк. *кёк_II* – тофалар. *бұт-_{II}*; 2) карабалк. *кёк_II* – турецк. *gök*; 3) карач.-балк. *кёк_II* – узбек. *falak*; 4) тофалар. *бұт-_{II}* – турецк. *gök*; 5) тофалар. *бұт-_{II}* – узбек. *falak*; 6) турецк. *gök* – узбек. *falak*. Эти пары и подсчитываются для каждой из 6 пар исследуемых языков, фигурируя в виде величины $L_1 \cap L_2$ в указанной формуле.

Упрощенная методика вычисления КоТибл включает следующие операции:

- 1) совмещение ПЯ четырех сравниваемых языков на одном листе электронных таблиц Эксель;
- 2) расщепление метаязыкового столбца, содержащего русские эквиваленты и дефиниции по числу

тестовых слов (с тем, чтобы каждый столбец содержал одно слово);

- 3) создание новой версии больших ПЯ четырех языков с однословным русским эквивалентом;
- 4) совмещение новых версий ПЯ на одном листе электронных таблиц Эксель;
- 5) сортировка русских эквивалентов по алфавиту;
- 6) подсчет посредством функции «СЧЕТЕСЛИ» слов с одинаковым эквивалентом;
- 7) фильтрование материала: удаление неполнозначных эквивалентов и несловарных форм;
- 8) разбиение рядов на языковые пары и подсчет связей в каждой паре языков (табл. 1);
- 9) суммирование связей в каждой паре языков (табл. 2);

Т а б л и ц а 1

Абсолютное количество семантических связей в парах тюркских языков

Кол-во слов Пары	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Всего
К.-балк.-Тоф.		234	319	267	213	147	202	232	154	143	81	62	35	28	26	67	33	59	96	42	20	67	2631
К.-балк.-Тур.	119	217	315	267	309	161	200	143	128	104	71	97	29	63	56	43	26	28	18	50	106	71	2621
К.-балк.-Узб.	236	475	540	442	355	196	308	273	178	141	177	84	156	63	98	98	92	43	61	98	136	109	4359
Тоф.-Тур.	16	39	52	40	72	48	44	32	13	15	20	19	6	13	7	12	4	2	9	5	8	13	489
Тоф.-Узб.	45	73	122	65	68	57	65	94	23	36	44	31	18	13	10	11	14	5	30	11	14	59	908
Тур.-Узб.	65	120	157	97	132	95	82	119	69	53	62	78	36	40	43	21	19	4	24	25	64	19	1424

Примечание. В верхней строке таблицы указано количество слов (лемм) во всех четырех языках.

Т а б л и ц а 2

Абсолютное количество семантических связей в параметрических ядрах тюркских языков

Язык	К.-балк.	Тоф.	Тур.	Узб.
К.-балк.		2631	2621	4359
Тоф.	2631		489	908
Тур.	2621	489		1424
Узб.	4359	908	1424	
Всего:	9611	4028	4534	6691

10) вычисление КоБл по указанной выше формуле;

11) представление результата в табличной форме.

Т а б л и ц а 3

Коэффициент типологической близости параметрических ядер тюркских языков

Язык	К.-балк.	Тоф.	Тур.	Узб.
К.-балк.		0,70	0,67	1,05
Тоф.	0,70		0,56	0,81
Тур.	0,67	0,56		1,10
Узб.	1,05	0,81	1,10	
Всего:	2,41	2,07	2,32	2,96

Исходя из данных табл. 3 видно, что в отношении карачаево-балкарского ко всем тюркским языкам лексическая пара **КараБалк./Узбек.** параметрическим ядрам типологически наиболее близка; второе-третье место делят тофаларский и чуть-чуть отставший от него турецкий язык. Наглядно это представлено на рис. 1.

Как видим, карачаево-балкарский язык демонстрирует большую типологическую близость к карлукскому узбекскому языку и наименьшую – к огузскому турецкому языку. Справедливости ради отметим, что саянский тофаларский типологически не намного ближе к кипчакскому карачаево-балкарскому, чем турецкий.

Для тофаларского языка картина другая (рис. 2). Самым близким оказывается узбекский язык, а самым удаленным – турецкий.

Близость к турецкому языку представлена на рис. 3. И тут картина остается неизменной: ближе всего узбекский язык, дальше всего – тофаларский, между ними – карачаево-балкарский.

Близость к узбекскому языку представлена на рис. 4. Для узбекского языка, как и для турецкого, наиболее удаленный – тофаларский, средний – карачаево-балкарский. Самым близким к узбекскому оказы-

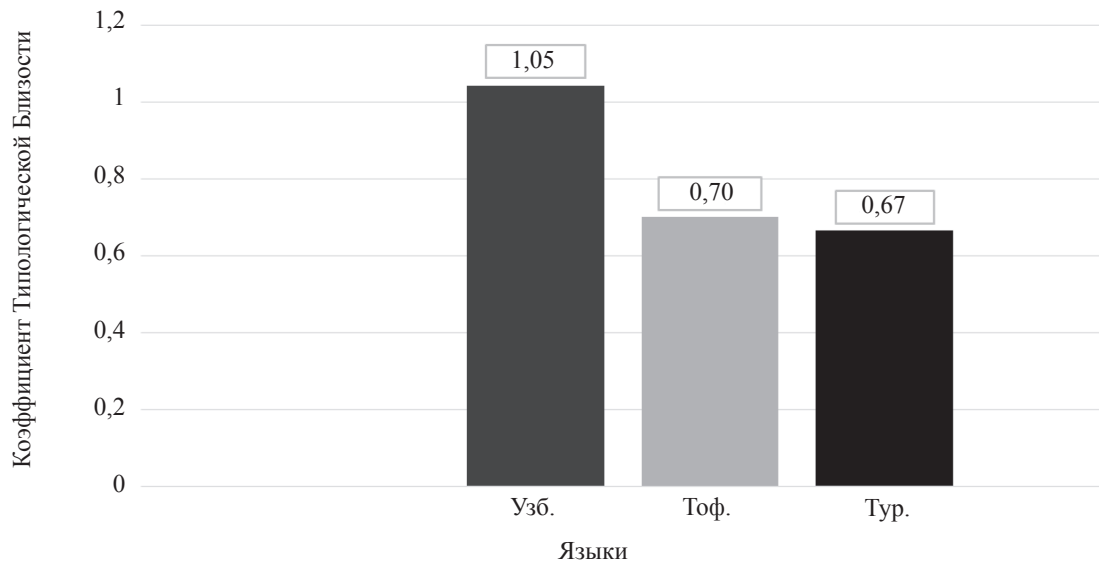


Рис. 1. Ранжирование лексики тюркских языков по упрощенному коэффициенту типологической близости относительно карачаево-балкарского ПЯ лексики

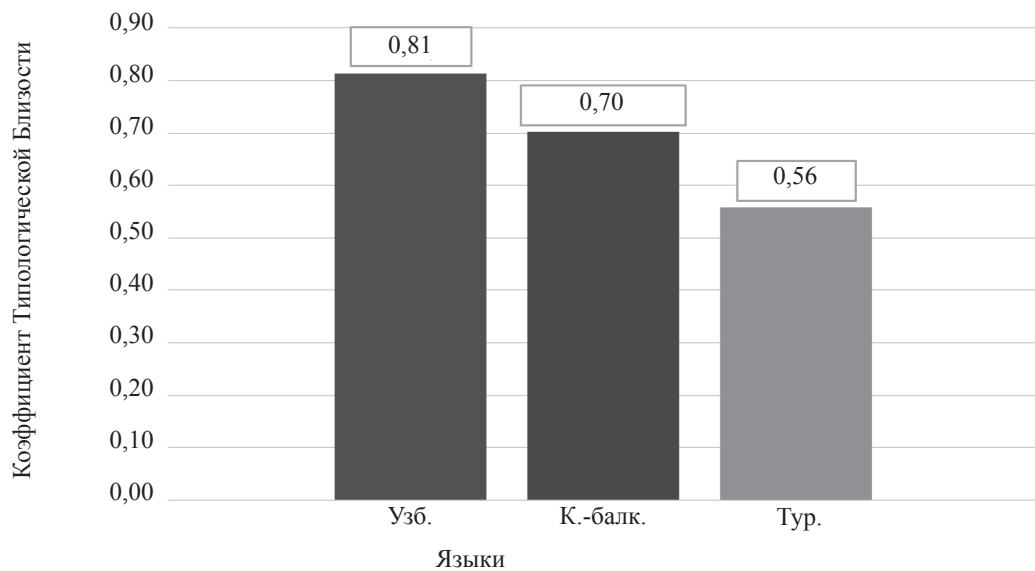


Рис. 2. Ранжирование лексики тюркских языков по упрощенному коэффициенту типологической близости относительно тофаларского ПЯ лексики

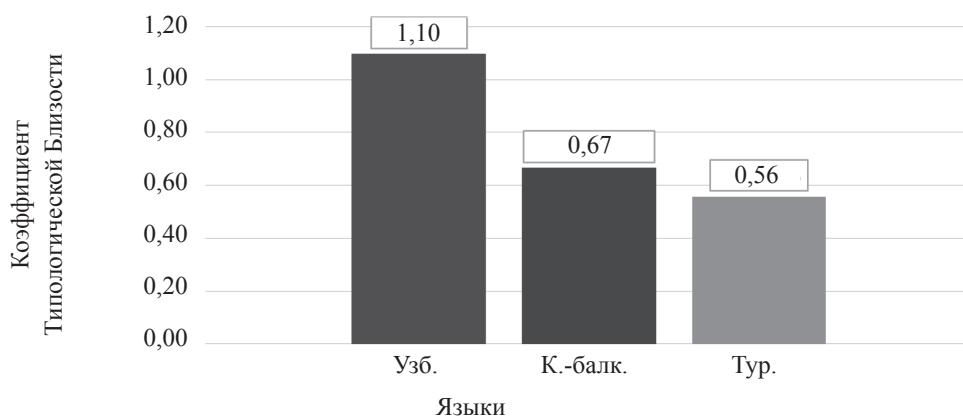


Рис. 3. Ранжирование лексики тюркских языков по упрощенному коэффициенту типологической близости относительно турецкого ПЯ лексики

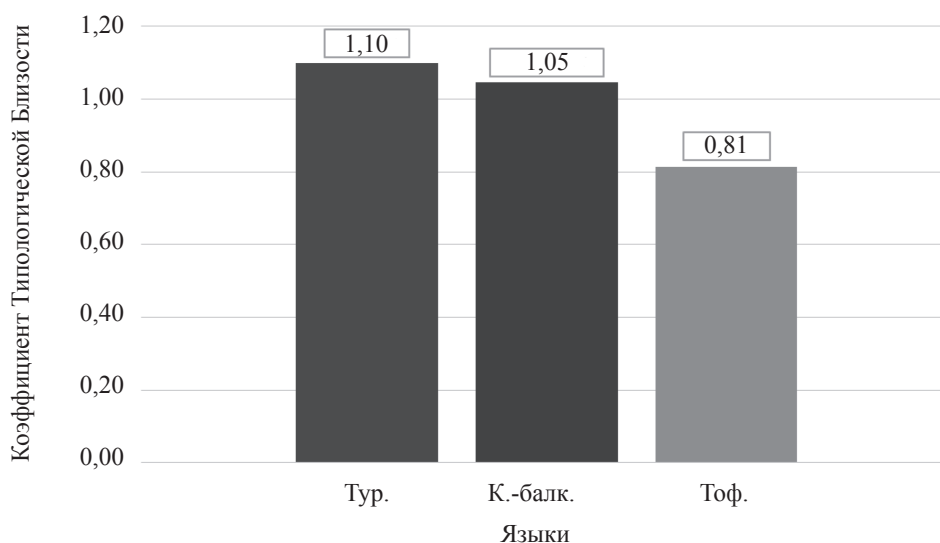


Рис. 4. Ранжирование лексики тюркских языков по упрощенному коэффициенту типологической близости относительно узбекского ПЯ лексики

вается турецкий язык. Тут узбекский, как и тофаларский, отвечает турецкому языку взаимностью.

А. В. Дыбо утверждает: «Язык, будучи средством общения, функционирует в некотором социуме, где это общение происходит. Как могут получиться с течением времени различные линии эволюции одного и того же языка? Простое предположение состоит в том, что языковые изменения (которые происходят в каждом языке постоянно из-за ряда факторов, связанных с особенностями его существования) происходят в основном однонаправленно внутри общающегося социума (просто поскольку язык должен сохранять общепонятность в этом социуме), а при разделении социума и прекращении общения между вновь образовавшимися социумами в разделившихся частях появляются разные изменения. С накоплением различий взаимопонимание между идиомами, развившимися в новых социумах, затрудняется и можно говорить о наличии нескольких отдельных языков. Тогда реконструкция праязыка, произведенная лингвистическими методами над материалом этих отдельных языков, должна свидетельствовать, в частности, о наличии в прошлом определенного социума, бывшего носителем этого реконструированного праязыка, и осмысленной задачей будет найти какие-то внешние свидетельства существования такого социума. Причем решение таких задач важно не только (и, в общем, не столько) для сравнительно-исторического языкознания (это последнее важно и как собственно лингвистическая дисциплина, в качестве более дальнего приложения своих усилий оно имеет еще и построение теории и типологии языковых изменений как части общей теории языка)» [5].

Применение метода параметрического анализа к лексике карачаево-балкарского языка и выделение его лексико-семантического ядра, сопоставление этого

ядра с ядрами узбекского, турецкого и тофаларского языков и выявление общих семантических связей между языками открывают перспективу исследования лексико-семантической близости языков в типологическом аспекте и построения типологии лексико-семантических изменений.

Применение параметрического анализа лексики (ПАЛ) к четырем тюркским языкам можно рассматривать как пилотный проект типологического исследования лексики (в идеале) всех или (в реальности) всех имеющих тюрко-русские словари языков. Это делает целесообразным получение хотя бы самых предварительных данных по типологической характеристике четырех тюркских языков.

Так, сумма КоТиБов, объединяющих данный язык с остальными сравниваемыми (СумКоТиБ), может рассматриваться как показатель центральности-периферийности ПЯ лексики данного языка в типологическом пространстве лексики тюркских языков.

На рис. 5. представлено ранжирование тюркских языков на шкале центральности-периферийности по величине СумКоТиБа.

Таким образом, по нашим данным, максимально центральным оказывается узбекский язык, а явно периферийным – тофаларский. Кипчакский карачаево-балкарский язык занимает промежуточную лексико-типологическую позицию между карлукским узбекским и огузским турецким языками.

Типологически информативным может также быть показатель доли семантически уникальной лексики в ПЯ.

К семантически уникальной лексике отнесем такие случаи, когда данному сочетанию данного слова с русским метасловом не находится эквивалента с таким же метасловом в другом тюркском языке.

Например, тоф. *айас* 'вёдро' и тоф. *айас* 'погожий' содержат уникальные для рассмотренных множеств слов метаслова: 'вёдро' и 'погожий'. В тофаларско-русском словаре [6] это слово представлено следующим образом: «*айас* 'вёдро, ясный, погожий'». Из трех метаслов в других тюркско-русских словарях использовано лишь метаслово 'ясный'. Этот случай поясняет, что мы имели в виду, когда говорили не только об «упрощенной», но и об «укрупненной» методике вычисления КоТиБа: дефиниции, состоящие из одного слова, сравнивать проще, но за это приходится платить ростом («укрупнением») семантической неопределенности метаслова, не ограниченной другими словами дефиниции.

Естественно, что после расщепления дефиниций единиц, состоящих из леммы и одного метаслова, оказывается больше, чем лемм (и слов = лемма + дефиниция) в анализируемом множестве.

В лингвистической типологии вообще [7] и при использовании параметрическим анализом лексики, в частности, уже сложилась традиция пользоваться термином *уникалия* для обозначения уникальных явлений (в том числе и слов, не имеющих семантических эквивалентов в ПЯЛах других языков). Следуем этой традиции, используя долю уникалий в ПЯЛах как *Индекс периферийности (ИнП)* (табл. 4).

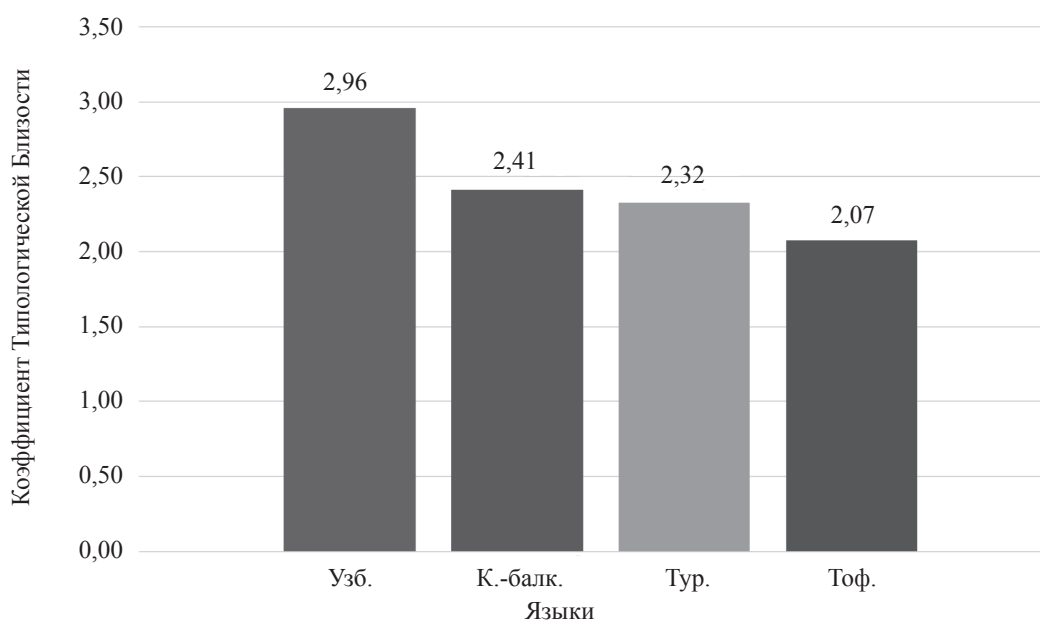


Рис 5. Ранжирование лексики тюркских языков по центральности-периферийности

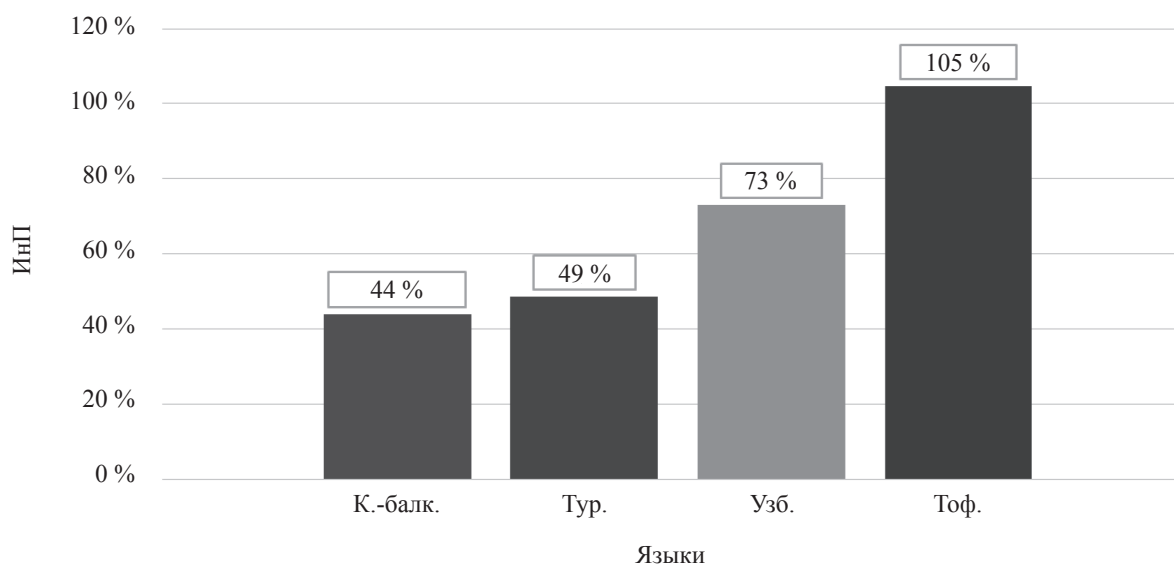


Рис. 6. Распределение тюркских языков по Индексу периферийности

Т а б л и ц а 4

Вычисление Индекса периферийности
по параметрическим ядрам тюркских языков

Язык	Уникалий	Ядро	ИнП, %
Карачаево-балкарский	2992	6807	44
Тофаларский	727	695	105
Турецкий	513	1057	49
Узбекский	1119	1535	73

На рис. 6. представлено ранжирование четырех тюркских языков по ИнПу.

Этот показатель дает несколько иную картину, помещая в центр карачаево-балкарский язык и отодвигая к периферии узбекский. Неизменной остается периферийная позиция тофаларского языка. Представленная картина, несмотря на учет размеров множеств, все же коррелирует с размерами последних. Возможно, в будущем ИнП потребуется рассчитывать с учетом количества не слов, а квази-ЛСВ – «расщепленных» единств «лемма+метаслово».

ЛИТЕРАТУРА

1. Бугаёв В. П. Параметрический анализ лексики турецкого языка : дипломная работа по специальности

Воронежский государственный университет
Семенова И. Д., соискатель кафедры теоретической
и прикладной лингвистики
E-mail: semjonova-i@yandex.ru
Тел.: 8-926-054-34-69

«Теоретическая и прикладная лингвистика» / В. П. Бугаёв. – Воронеж, 2006. – 99 с.

2. Хайиткулова Н. Р. Параметрическое ядро узбекской лексики на фоне родственных языков : дипломная работа по специальности «Теоретическая и прикладная лингвистика» / Н. Р. Хайиткулова. – Воронеж, 2009. – 113 с.

3. Воржева И. В. Параметрическое ядро тофаларской лексики на фоне родственных языков : дипломная работа по специальности «Теоретическая и прикладная лингвистика» / И. В. Воржева. – Воронеж, 2008. – 131 с.

4. Кретов А. А. Проблемы количественной лексикологии славянских языков / А. А. Кретов, И. А. Меркулова, В. Т. Титов // Вопросы языкознания. – 2011. – № 1. – С. 52–65.

5. Дыбо А. В. Семантическая реконструкция в алтайской этимологии : лексика конкретного словаря / А. В. Дыбо // Слово и язык : сб. ст. к 80-летию акад. Ю. Д. Апресяна. – М. : Языки славянских культур, 2011. – 736 с.

6. Рассадин В. И. Тофаларско-русский словарь. Русско-тофаларский словарь / В. И. Рассадин. – Иркутск : Восточно-Сибирское книжное издательство, 1994. – 288 с.

7. Шилихина К. М. Основы лингвистической типологии : учеб.-метод. пособие для студ. вузов / К. М. Шилихина. – Воронеж, 2007. – 63 с.

Voronezh State University
Semyonova I. D., Post-graduate Student of the Theoretical and Applied Linguistics Department
E-mail: semjonova-i@yandex.ru
Tel.: 8-926-054-34-69