

ДАННЫЕ ИНТЕРНЕТА В ИССЛЕДОВАНИИ ЯЗЫКОВЫХ ИЗМЕНЕНИЙ: АНАЛИЗ ЧЕРЕДОВАНИЙ В РУССКИХ КОМПАРАТИВАХ И ПРОГРАММА ДЛЯ РАБОТЫ С ТАКИМИ ДАННЫМИ

Магомедова В. Д. (varya.magomedova@gmail.com)

Санкт-Петербургский государственный
университет, Санкт-Петербург, Россия

Слюсарь Н. А. (slioussar@gmail.com)

Национальный исследовательский университет
Высшая школа экономики, Москва, Россия
Санкт-Петербургский государственный
университет, Санкт-Петербург, Россия

Ключевые слова: компаратив, согласные, исторические чередования, оптимизация поиска

INTERNET DATA IN THE STUDY OF LANGUAGE CHANGE: A CASE STUDY OF ALTERNATIONS IN RUSSIAN COMPARATIVES AND A PROGRAM TO WORK WITH SUCH DATA

Magomedova V. D. (varya.magomedova@gmail.com)

Saint Petersburg State University, St. Petersburg, Russia

Slioussar N. A. (slioussar@gmail.com)

National Research University
Higher School of Economics, Moscow, Russia
Saint Petersburg State University, St. Petersburg, Russia

The Internet is a unique source of non-standard forms, which gives us a novel opportunity to analyze fine-grained dynamics of language change. We used this opportunity to study the decay of historic consonant

alternations in Russian. In standard Russian, these alternations are present in some verb forms and in comparatives (e.g. *suxoj* 'dry' — *sushe* 'drier', *ljubit'* 'to love' — *ljublju* 'I love'), as well as before certain derivational suffixes. Verb forms have been recently studied by Slioussar and Kholodilova (2013), and we looked at comparatives. Two groups of adjectives were selected: ones that have normative comparatives with alternations and ones that do not, but native speakers still try to generate such forms. In the first group, some adjectives like *ubogij* 'poky' have up to 30% of comparatives without alternations, but, unlike with verbs, no significant correlation with adjective frequency or its other characteristics was found. The second group consisted primarily of compound adjectives ending in *-gij*, *-kij*, *-xij*. Here, the most important factor is whether the second part of the compound is used as an independent adjective. If it is not (e.g. as in *dlinnorukij* 'long-armed'), most comparatives lack alternations. Searching for forms on the Internet, we faced many problems. The counts provided by search engines are extremely inaccurate, only the first thousand results are shown, they cannot be downloaded in a convenient format, contain a lot of typos and other irrelevant data etc. We present a program called *Lingui-Pingui* that we developed to solve these and some other problems.

Keywords: comparative, consonants, historical alternations, search optimization

1. Введение

Огромное количество неотредактируемых текстов в интернете дает лингвистам новые возможности для наблюдения за жизнью языка. С приходом массовой грамотности процессы языковых изменений значительно замедлились, и всё-таки удается найти интересные примеры таких явлений. В этой работе мы анализируем процесс разрушения исторических чередований согласных, система которых остается неизменной в нормативном русском языке¹, но явно расшатывается в целом ряде ненормативных форм. Мы рассматриваем это явление на примере форм компаратива.

Однако при использовании данных из интернета возникает целый ряд проблем, связанных с их сбором и обработкой. Например, статистика по результатам поиска, которую дают поисковые машины, очень неточная, что существенно осложняет оценку частотности различных форм. При этом просматривание результатов вручную затрудняется тем, что их невозможно сохранить в удобной для работы форме. Мы разработали программу «Lingui-Pingui», которая работает с Yandex.API и помогает собирать, сортировать и обрабатывать данные, полученные из Интернета. Цель данной работы двойка: представить результаты анализа форм компаратива и созданную нами программу «Lingui-Pingui».

¹ В нормативном русском языке исторические чередования согласных (к // ч, к // ц, г // ж, г // з и др.) сохраняются в ряде глагольных классов (например, *писать* — *пишу*, *возить* — *вожу*), в некоторых формах компаратива (например, *сухой* — *суше*), а также перед определенными суффиксами (например, *скользить* — *скольжение*).

2. Предыдущие исследования

Н. А. Слюсарь и М. А. Холодилова [Slioussar, Kholodilova 2013] изучали исторические чередования согласных на материале глаголов. Они провели анализ ненормативных форм, которые можно в изобилии найти в интернете, и выяснили, что система чередований расшатывается. Это явление особенно интересно, так как парадигматическое выравнивание давно и активно обсуждается в рамках самых разных лингвистических направлений [например, Albright 2002, 2010; Anttila 1977; Benua 1997; Bybee 1985; Kiparsky 1982, 2002; Kuryłowicz 1949; Mańczak 1958; McCarthy 2005].

Слюсарь и Холодилова показали, что процесс парадигматического выравнивания зачастую идет в двух направлениях одновременно: избавление от чередований в тех формах, где они были, и появление чередований в тех, где их не было. Однако избавление от чередований более распространено. В этой работе также были выведены факторы, влияющие на сохранность чередований, на примере различных глаголов единственного продуктивного класса, где встречаются чередования, — 1 подкласса X класса. Это частотность, нормативность глагола (у редких и ненормативных глаголов больше проблем с чередованиями) и конечный согласный основы. Чередования сохраняются лучше всего у глаголов с основами на губные согласные и хуже всего у глаголов с основами, заканчивающимися группами шумных согласных. Кроме того, были найдены примеры чередований, не встречающихся в нормативном русском языке. В этом случае мы имеем дело с интересной ситуацией, когда носитель языка понимает, что в той или иной форме необходимо чередование, но уже не уверен, какое именно.

3. Интересующие нас формы компаратива и различные их свойства

1. Чередования согласных встречаются только в неизменяемых синтетических формах компаратива с суффиксом *-e*. Другие суффиксы, при помощи которых образуются эти формы, *-ee/eй* и *-ше*, чередований не требуют. Суффикс *-e* присоединяется ко всем прилагательным с основой на *-z*, *-k*, *-x*, у которых есть синтетические формы, и к нескольким основам на *-d*, *-t* (кроме того, есть единичные случаи вроде *дешевле*). При этом у некоторых прилагательных на *-кий* выпадает суффикс *-(o)к-*.
2. В данной работе мы рассмотрим два типа ненормативных форм: от прилагательных, у которых есть нормативный компаратив с чередованием, и от прилагательных, у которых его нет. Отсутствие синтетического компаратива может быть обусловлено как семантикой прилагательного (в этом случае не образуются никакие формы компаратива), так и его морфологическими характеристиками. Более подробную информацию об этом можно найти, например, в работах [Зализняк 1977; Князев 2007; Русская грамматика 1980 и др.]. Судя по нашим наблюдениям, в тех случаях, когда

образование компаратива допустимо с точки зрения семантики, носители русского языка имеют тенденцию употреблять не только аналитические, но и ненормативные синтетические формы. Следует заметить, что небольшая часть текстов, где встречаются такие формы, имеет шуточный характер. Так как однозначно отделить такие тексты от прочих не представлялось возможным (не всегда понятно, шутит ли пишущий), мы учитывали их вместе со всеми. В целом у нас создалось ощущение, что люди склонны чаще шутить с теми же формами, в которых чаще встречаются ошибки.

3. Наша задача заключалась в том, чтобы проанализировать распределение форм с чередованиями и без них от различных прилагательных. Мы искали как минимум десять форм для каждого прилагательного, проиллюстрированных ниже на примере прилагательного *упругий*: с чередованиями и без них, с суффиксами *-e* и *-ee/eй*, с приставкой *по-* и без нее. В принципе, утеря чередований сопровождается заменой суффикса *-e* на более новый продуктивный *-ей/ее*, но мы заметили, что эти процессы не всегда идут параллельно и включили промежуточные формы с чередованиями с суффиксом *-ee/eй*. Затем наше внимание привлек тот факт, что от некоторых прилагательных встречаются и формы без чередований с суффиксом *-e*, например, *круте* или *сухе*. Мы уже не успели включить их в основное исследование, поэтому в разделе 4 только перечисляем те (немногочисленные) прилагательные, от которых были найдены такие формы. Для прилагательных с суффиксами мы также искали формы, где эти суффиксы выпадают, например, *хлипее* наравне с *слипкее*.

- (1) *упруже, поупруже, упругее, упругей, поупругее, поупругей, упружее, упружей, поупружее, поупружей*

Важно заметить, что проблемы с чередованиями могут проявляться и в избегании синтетических форм за счет использования аналитических. Мы не анализируем соответствующие данные в этой статье, но планируем сделать это в будущем.

4. Распределение форм компаратива с чередованиями и без них

Согласно «Грамматическому словарю русского языка» [Зализняк 1977], в русском языке 113 бесприставочных прилагательных с основами на *-z*, *-k*, *-x*, образующих нормативные формы компаратива с чередованием. Ненормативные формы от этих прилагательных довольно редки², и результаты поиска этих форм, как правило, содержат много «мусора» — опечаток и других нерелевантных данных. Поэтому мы отобрали 23 прилагательных, у которых результаты

² Как показывают в своей работе Н. А. Слюсарь и М. А. Холодилова [Slioussar, Kholodilova in 2013], похожая ситуация наблюдается и с глаголами: ненормативные формы от глаголов, относящихся к литературному русскому языку, также редки.

первичного поиска нестандартных форм (одна форма без чередования на *-ee* для каждого прилагательного) содержали не менее 4% релевантных данных. К этим 23 прилагательным были добавлены девять с основами на *-т*, *-д* (все бесприставочные с чередованиями кроме *худой* в значении «плохой»). Также мы рассматривали прилагательные, которые не имеют нормативных синтетических компаративов: 13 сложных прилагательных с основами на *-з*, *-к*, *-х* (например, *близорукий*), несколько слов с суффиксами *-ск-* и *-цк-* и прилагательное *великий*.

Распределение различных форм у прилагательных первой и второй группы представлено ниже в Таблицах 1–4. Данные о частотности прилагательных взяты из «Частотного словаря современного русского языка» [Ляшевская, Шаров 2009], у тех прилагательных, которые не вошли в этот словарь, в соответствующей графе стоит прочерк. Количество результатов указывается после фильтрации.

Формы с суффиксами *-ee* и *-ей*, а также формы с приставкой *по-* и без нее для краткости приведены вместе. Заметим, что в среднем формы с суффиксом *-ee* примерно в два раза частотней, чем суффиксом *-ей*, но их соотношение очень отличается у разных прилагательных (в частности, в некоторых случаях форм с суффиксом *-ей* больше). Формы с приставкой *по-* в среднем встречаются в девять раз реже, чем без нее, но их соотношение также сильно варьирует. Так, у прилагательного *тугой* около четверти форм с приставкой *по-*.

Таблица 1. Распределение форм компаратива у прилагательных первой группы без суффикса *-(о)к-*

Прилагательное	Частотность	Кол-во результатов	Формы на <i>-е</i>	Формы на <i>-ee/ей</i>	Формы на <i>-ee/ей</i> с чередованиями
богатый	85,0	1873	99,7%	0,3%	0,0%
глухой	40,7	766	99,7%	0,0%	0,3%
густой	47,2	500	98,0%	1,6%	0,4%
крутой	43,4	1809	99,9%	0,0%	0,1%
молодой	414,1	1006	99,5%	0,5%	0,0%
плоский	29,1	686	99,3%	0,6%	0,2%
пологий	5,0	294	98,6%	0,7%	0,7%
простой	275,3	1622	97,1%	0,2%	2,7%
строгий	62,8	693	99,9%	0,1%	0,0%
сухой	83,0	475	95,2%	0,0%	4,8%
твердый	59,5	615	99,0%	1,0%	0,0%
толстый	84,4	1087	99,0%	0,5%	0,6%
тугой	11,9	492	100,0%	0,0%	0,0%
убогий	11,7	342	48,8%	32,8%	18,4%
упругий	10,7	588	68,9%	27,0%	4,1%
частый	114,6	1308	100,0%	0,0%	0,0%
чистый	159,0	1574	99,6%	0,3%	0,1%
яркий	93,9	1701	100,0%	0,0%	0,0%

Таблица 2. Распределение форм компаратива у прилагательных первой группы с суффиксом *-(о)к-*

Прилагательное	Частотность	Кол-во результатов	Формы на -е	Формы на -ее/ей	Формы на -ее/ей с чередованиями	Формы без чередований с выпадением суффикса
близкий	206,8	1505	98,24%	0,0%	0,10%	1,66%
высокий	483,3	1502	99,90%	0,0%	0,10%	0%
глубокий	137,3	1172	99,80%	0,0%	0,20%	0%
низкий	160,7	1197	98,50%	0,0%	0,08%	1,42%
веский	4,9	257	89,50%	0,0%	10,50%	0%
громоздкий	7,1	274	75,91%	14,6%	2,19%	7,30%
дерзкий	8,9	387	60,72%	0%	1,55%	37,73%
жестокий	37,5	581	84,50%	2,4%	13,10%	0%
жуткий	26,1	828	79,38%	1,43%	13, 23%	5,96%
одинокий	44,5	162	48,80%	29,0%	22,20%	0%
скользкий	11,8	325	13,85%	0,92%	3,38%	81,85%
стойкий	10,1	494	92,90%	0,2%	6,90%	0%
тяжкий	26,0	349	97,10%	1,7%	1,20%	0%
хлипкий	2,4	705	96,75%	0,44%	1,48%	1,33%

Анализ данных не выявил статистически значимой зависимости между количеством форм без чередований и такими факторами, как частотность прилагательного или конечный согласный основы, которые оказались значимыми при исследовании глагольных форм. Видно, что проблемы с чередованиями свойственны отдельным прилагательным, которые не объединяет какое-то общее свойство. Возможно, это связано с тем, что интересующая нас группа глаголов на данном этапе развития русского языка активно пополняется, а группа прилагательных — нет. Заметим при этом, что, хотя у многих прилагательных компаративов без чередований в процентном отношении совсем немного, например, 0,5% форм без чередований от такого частотного прилагательного, как *молодой*, — это несколько тысяч результатов.

Также встречаются формы компаратива с выпадением суффикса *-к-*. Данные об этих прилагательных приведены в Таблице 2. Кроме того, мы нашли единичные примеры форм с чередованием, где в противоположность норме выпадает или сохраняется суффикс (*хлиплее*, *скольже*, *громо(з)же*, *жуче*, *держе*, *мерже*, а также *близче*, *низче*). Изредка можно найти и формы с чередованиями, не встречающимися в нормативном русском языке, например, *богаще*, *круще*, *ярще*, а также крайне редкие примеры типа *скользже*. Наконец, были обнаружены формы *круте*, *сухе*, *ярке*, *близе* и *убоге*. Только первая из них встречается относительно часто (11 однозначных результатов после фильтрации и вычитки), другие крайне редки (1–3 формы на тысячу результатов). Перейдем ко второй группе прилагательных.

Таблица 3. Распределение форм компаратива у сложных прилагательных второй группы и прилагательного *великий*

Прилагательное	Частотность	Кол-во результатов	Формы на -е	Формы на -е/ей	Формы на -е/ей с чередованиями
великий	276,3	87	6,9%	20,7%	72,4%
безрукий	1,3	18	0,0%	88,9%	11,1%
близорукий	3,5	26	15,4%	73,1%	11,5%
длинноногий	3,3	114	2,6%	80,7%	16,7%
длиннорукий	0,6	19	0,0%	73,7%	26,3%
жизнестойкий	0,4	16	81,3%	12,5%	6,3%
износостойкий	—	65	70,8%	15,4%	13,9%
легкоплавкий	—	24	83,5%	4,0%	12,5%
лопоухий	1,2	53	1,9%	84,9%	13,2%
морозостойкий	0,6	22	54,6%	18,2%	27,3%
термостойкий	—	21	42,9%	33,3%	23,8%
трудоемкий	4,7	287	74,6%	19,2%	6,3%
тугоплавкий	0,4	43	76,7%	16,3%	7,0%
энергоемкий	1,1	132	89,4%	6,8%	3,8%

Сразу бросается в глаза, что во второй группе значительно больше форм без чередований. Здесь значимое влияние на распределение форм оказал такой фактор: прилагательные *стойкий*, *емкий* и *зоркий* существуют и сами по себе, и от них образуются компаративы с чередованиями, а прилагательных *рукий*, *ногий* и *ухий* нет. Несмотря на то, что варианты соответствующих корней с чередованиями согласных высокочастотны и должны быть на слуху (*ножка*, *ручка*, *уши*, *ушко* и пр.), компаративы с чередованиями от этих сложных прилагательных встречаются редко. В Таблице 4 приведены данные по оставшимся прилагательным второй группы, обозначающим отнесенность к той или иной национальности. Они даются отдельно, так как у многих из них есть формы с выпадением суффикса.

Таблица 4. Распределение форм компаратива у прилагательных, обозначающих отнесенность к той или иной национальности

Прилагательное	Частотность	Кол-во результатов после фильтрации	Формы на -е	Формы на -е/ей	Формы на -е/ей с чередованиями	Формы с выпадением суффикса
итальянский	40,9	3	0%	57%	71,43%	14%
испанский	21,1	7	14%	64%	43,00%	0%
немецкий	123,6	27	0%	19%	92,59%	0%
азиатский	10,3	26	4%	57%	11,54%	77%
французский	101,5	244	42,62%	53%	39,34%	15%
русский	530,5	435	4,37%	7%	53,33%	40%

5. Сбор данных в интернете

5.1. Основные принципы и проблемы

Нашей основной задачей было оценить относительную частотность различных форм компаратива, встречающихся в интернете. Мы работали с поисковой машиной «Яндекса». Представим для начала, что мы сравниваем всего две формы, скажем, *суше* и *сухее*. Почему мы не можем просто задать поисковику два запроса и сравнить те числа, которые он выдает на первой странице? Во-первых, потому, что эти числа крайне неточные. Во-вторых, потому, что в результатах будет содержаться много нерелевантных данных, которые необходимо отсеять. Например, *суше* — это не только компаратив, но и форма от существительного *суша*.

Поэтому правильнее задать поисковику запрос, как в примере (2), предполагающий одновременный поиск двух форм, просмотреть какое-то количество полученных результатов и оценить, какова в них относительная частотность этих форм после отсева нерелевантных данных (этот метод изначально был предложен М. А. Холодиловой [2013]).

(2) "*суше*" | "*сухее*"

Проблема заключается в том, что просмотр результатов — крайне трудоемкое занятие. Их нельзя загрузить в удобном формате и затем подвергнуть первичной автоматической обработке. В связи с этим мы разработали программу «Lingui-Pingui»³ (<https://sites.google.com/site/varyamagomedova>) для облегчения задач лингвистического поиска, которая работает через API поисковой машины «Яндекса» (<http://api.yandex.ru/>) с использованием встроенного языка запросов.

Кроме того, поисковая машина позволяет пользователю просматривать только первую тысячу результатов, однако М. А. Холодилова разработала алгоритм, который позволяет обойти это ограничение [Холодилова 2013: 20-21]. Это делается при помощи ограничения по дате изменения ресурса: для каждой тысячи результатов находится наиболее старый ресурс, и следующий запрос запускается с ограничением по дате так, чтобы результаты каждой следующей тысячи были полностью старше по дате изменения, чем результаты предыдущей. Текст самого запроса при этом остается без изменений. Этот алгоритм воплощен и в нашей программе. Тем не менее, зачастую первой тысячи результатов оказывается вполне достаточно для того, чтобы оценить относительную частотность различных форм.⁴

³ Мы выражаем благодарность Ивану Антонову за неоценимую помощь в разработке и отладке программы, а также Марии Холодиловой за предоставленный алгоритм и помощь в разработке требований к программе.

⁴ Как нам подтвердили в службе поддержки «Яндекса», все документы сразу сортируются по релевантности независимо от того, какое слово стоит в запросе первым, а какое последним. Поэтому соотношение найденного соответствует тому, что есть в сети и проиндексировано поисковой системой.

5.2. Программа «Lingui-Pingui»

Программа «Lingui-Pingui» автоматически формирует запросы по заданным параметрам, отправляет их поисковой машине, сохраняет полученные данные и представляет их в удобном для исследователя формате. Кроме того, производится грубая фильтрация данных и считается статистика (общее количество результатов, относительные частотности). Встроенные фильтры и сортировка не заменяют полностью ручной обработки данных, но существенно облегчают эту задачу.

До начала работы программа состоит из скриптов на языке Perl и папки «Input», содержащей файл настроек и текстовые файлы с входными данными. Перед использованием программы необходимо связаться со службой поддержки Yandex.API, чтобы получить возможность отправлять определенное количество поисковых запросов за день (более подробную информацию можно получить на странице <http://xml.yandex.ru/> и у службы поддержки). Более детальное описание программы можно найти в работе [Magomedova 2013] или получить, связавшись с нами по электронной почте.

5.2.1. Формирование запросов

«Lingui-Pingui» позволяет автоматизировать создание запросов из заданных списков приставок, корней и суффиксов/окончаний. Рассмотрим работу программы на примере поиска различных форм компаратива от прилагательных *сухой*. Сперва программа формирует по одному запросу на каждый корень, включая туда все возможные комбинации заданных морф. После формирования запроса к нему можно добавить исключения. Скажем, в нашем примере компаратив *суше* совпадает с формами существительного *суша*. Такие данные лучше заранее исключить из поиска. На этом этапе сформированный запрос будет выглядеть следующим образом:

(3) ("сухе" | "сухее" | "сухой" | "посухе" | "посухее" | "посухой" | "суше" | "сушее" | "сушей" | "посуше" | "посушее" | "посушей") ~ "на суше" ~ "по суше"

Пользователь может задавать запросы и вручную, а также редактировать то, что было создано автоматически. Скажем, в примере (3) комбинация морфов *посухе* не используется в качестве формы компаратива даже в ненормативном языке, и ее можно исключить.

5.2.2. Представление результатов поиска

Сбор данных при поиске ведется по заголовкам страниц, найденных «Яндексом», и по фрагментам текста с этих страниц, которые также выдаются поисковиком. Поэтому в первой тысяче результатов поиска иногда получается больше тысячи строк употреблений искомых форм (если они появляются и в заголовке и во фрагменте текста). Количество слов до и после искомой формы, которые мы берем из исходного фрагмента текста, предоставленного поисковиком, можно задать в файле настроек программы. По умолчанию все фрагменты сохраняются целиком.

Файл со статистикой появляется в подкаталоге «Summary». Там содержится следующая информация:

- количество результатов по версии «Яндекса» («Яндекс» выдает три таких оценки, выбранная нами является наиболее точной и, по крайней мере, в случае, когда количество найденных результатов меньше тысячи и можно посчитать их вручную, совпадает с номером последней найденной страницы);
- количество неповторяющихся фраз;
- количество результатов после отсеивания нестрогих соответствий (поскольку формы, которые мы ищем, — ненормативные, поисковик часто считает их опечатками и включает в результаты поиска соответствующие нормативные формы или другие близкие по написанию слова);
- количество результатов, где одновременно встречаются несколько искомых форм;
- сведения о каждой искомой форме (сколько раз она встретилась и какой это процент от общего числа форм).

5.2.3. Сортировка и последующая обработка результатов

Вкрапления нерелевантных данных в результаты поиска практически неизбежны даже при самом тщательном подходе к формированию запросов. Отфильтровывать такие данные вручную — крайне трудоемкая задача. Автоматическая сортировка существенно ее облегчает.

Для сортировки используются элементы, совместная встречаемость с которыми (в рамках одного фрагмента) практически гарантирует, что найденная форма относится к нужным нам результатам, и элементы, которые почти наверняка являются индикаторами нерелевантных данных. В нашем случае это слова *чем, тем, гораздо, еще* и пр., которые указывают на компаратив, и слова *более, менее, совсем* и т.д., которые указывают на форму положительной степени с опечаткой, а также, например, буква *i* и слово “*як*”, которые позволяют отсортировать большую часть данных на украинском языке. Программа также содержит инструмент для пересчета статистики для данных, просмотренных вручную, и инструмент, позволяющий отбирать или исключать из файлов строки, содержащие определенные слова и комбинации, что тоже существенно облегчает ручную обработку данных, так как почти у каждой конкретной формы или конкретного типа опечатки есть свои «маркеры» вместе с которыми именно эта форма(опечатка) часто встречается в результатах.

6. Заключение

Разрушение исторических чередований согласных, исследованное Н. А. Слюсарь и М. А. Холодиловой на материале глаголов, явно затрагивает и формы компаратива. Так как группа прилагательных, от которых образуются нормативные компаративы с чередованиями, не пополняется новыми словами, эти чередования в общем меньше подвержены разрушению, чем у глаголов продуктивного X класса, о которых говорилось в разделе 2. Тем не менее,

у некоторых прилагательных до трети найденных в интернете форм не имеют чередований (это *убогий* и *упругий*, в меньшей степени *громоздкий*, *лихой*, *одинокий* и т.д.). Но в полную силу проблемы с чередованиями проявляются, когда люди пытаются образовать синтетический компаратив от прилагательных, у которых нет соответствующих нормативных форм. Мы выбрали для анализа группу таких прилагательных с основами на заднеязычные, так как в русском языке нет ни одной нормативной формы на *-хее*, *-гее* или *-кее*. Тем не менее, встречается множество ненормативных форм такого рода.

В группе сложных прилагательных типа *длинноногий* и *трудоёмкий* решающим фактором оказалось то, существует ли вторая часть в виде самостоятельного слова. В этом случае в Интернете находится больше компаративов с чередованием, в то время как от прилагательных на *-рукий*, *-ногий* и *-ухий* они образуются с трудом, несмотря на частотность вариантов соответствующих корней с чередованием (*ручка*, *ножка*, *уши*, *ушко*). Таким образом, на несколько ином материале, чем с глаголами, мы снова можем сделать вывод о том, что всё упирается в доступность конкретной формы от конкретного слова.

Литература

1. Зализняк А. А. Грамматический словарь русского языка: Словоизменение. М., 1977.
2. Князев Ю. П. Грамматическая семантика: Русский язык в типологической перспективе. М., 2007.
3. Ляшевская О. Н., Шаров С. А. Частотный словарь современного русского языка. М., 2009.
4. Русская грамматика. М., 1980.
5. Холодилова М. А. Позиционные свойства местоимений в русском языке. Магистерская диссертация. СПбГУ, 2013.
6. Albright A. Base-driven leveling in Yiddish verb paradigms // Natural Language and Linguistic Theory. Vol. 28. 2010. P. 475–537.
7. Albright A. The identification of bases in morphological paradigms. Doctoral dissertation, University of California, 2002.
8. Anttila R. Analogy. The Hague, 1977.
9. Benua L. Transderivational Identity: Phonological relations between words. Doctoral dissertation, University of Massachusetts, 1997.
10. Bybee J. Morphology: A study of the relation between meaning and form. Amsterdam, 1985.
11. Kiparsky P. Explanation in phonology. Dordrecht, 1982.
12. Kiparsky P. Paradigmatic effects. Stanford, 2002.
13. Kuryłowicz J. La nature des procédés 'analogiques' // Acta Linguistica. Vol. 5. 1949. P. 15–37.
14. Magomedova V. 2013. Lingui-Pingui a script collection for linguistic search Ms., St. Petersburg State University. (https://sites.google.com/site/varyamagomedova/lingui_yandex/user-guide)

15. *Mańczak W.* Tendences generals des changements analogiques // *Lingua*. Vol. 7. 1958. P. 298–325, 387–420.
16. *McCarthy J.* Optimal paradigms // *Paradigms in phonological theory* / ed. by Downing L. J., Hall T. A., Raffelsiefen R. Oxford, 2005. P. 170–210.
17. *Slioussar N., Kholodilova M.* Paradigm leveling in non-standard Russian // *Proceedings of the 20th FASL conference*. AnnArbot, MI, 2013.

References

1. *Albright A.* Base-driven leveling in Yiddish verb paradigms // *Natural Language and Linguistic Theory*. Vol. 28. 2010. P. 475–537.
2. *Albright A.* The identification of bases in morphological paradigms. Doctoral dissertation, University of California, 2002.
3. *Anttila R.* *Analogy*. The Hague, 1977.
4. *Benua L.* *Transderivational Identity: Phonological relations between words*. Doctoral dissertation, University of Massachusetts, 1997.
5. *Bybee J.* *Morphology: A study of the relation between meaning and form*. Amsterdam, 1985.
6. *Kholodilova M. A.*, 2013. *Pozicionnye svoystva mestoimenij v russkom jazyke*. Master Dissertation, St.Petersburg State University.
7. *Kiparsky P.* *Explanation in phonology*. Dordrecht, 1982.
8. *Kiparsky P.* *Paradigmatic effects*. Stanford, 2002.
9. *Knjazev, Ju. P.* 2007. *Grammaticheskaja semantika: Russkij jazyk v tipologicheskij perspective* ('Grammatical semantics: Russian language in a typological perspective'). Moscow: Jazyki Slavjanskikh Kul'tur, 2007.
10. *Kuryłowicz J.* La nature des procèsdits 'analogiques' // *ActaLinguistica*. Vol. 5. 1949. P. 15–37.
11. *Ljashevskaja O. H., Sharov S. A.*, 2009. *Chastotnyj slovar' sovremennogo russkogo jazyka (na materiale Nacional'nogo korpusa russkogo jazyka)*. (Modern Russian language frequency dictionary (based on data from Russian National Corpus)). M.: Azbukovnik.
12. *Magomedova V.* 2013. *Lingui-Pingui: a script collection for linguistic search*. Ms., St. Petersburg State University. (https://sites.google.com/site/varyamagomedova/lingui_yandex/user-guide).
13. *Mańczak W.* Tendences generals des changements analogiques // *Lingua*. Vol. 7. 1958. P. 298–325, 387–420.
14. *McCarthy J.* Optimal paradigms // *Paradigms in phonological theory* / ed. by Downing L. J., Hall T. A., Raffelsiefen R. Oxford, 2005. P. 170–210.
15. *Shvedova, N.* (ed.). 1982. *Russkaja grammatika* ('Russian grammar'). Moscow: Nauka.
16. *Slioussar N., Kholodilova M.* Paradigm leveling in non-standard Russian // *Proceedings of the 20th FASL conference*. AnnArbot, MI, 2013.
17. *Zaliznyak, A. A.* 1977. *Grammaticheskij slovar' russkogo jazyka. Slovoizmenenie* ('The grammatical dictionary of the Russian language. Inflection'). Moscow: Russkij Jazyk, 1977.