

ПОДХОДЫ К КОЛИЧЕСТВЕННОМУ АНАЛИЗУ ОЧИТОК В ЛЕКСИКОГРАФИИ

Шаповал Виктор Васильевич

кандидат филологических наук, доцент, Московский городской педагогический университет (ГАОУ ВО МПГУ)
shapovalvv@mgpu.ru

APPROACHES TO QUANTITATIVE ANALYSIS OF MISREADINGS IN LEXICOGRAPHY

V. Shapoval

Summary: We believe that an unskilled reader when trying to identify a handwritten record of a previously unheard and unknown word usually misreads its “dark” part or problematic segment of the word. For such a reader this “dark” segment is not a series of definite letters at the very beginning. The “dark” segment is seen as the series of elementary graphic elements. The last elements should be perceived as a few letters after the reader’s efforts to put them together and find acceptable (not necessarily correct) reading. Thus, the number of visual elements in the “dark” segment is a decisive factor for defining the limits of the reader’s imagination in such a difficult case.

Keywords: dictionary, dialect, rare word, typology of misreadings, frequency, trivial vs. exotic misreadings, Russian.

Аннотация: Мы полагаем, что неопытный читатель, пытающийся идентифицировать рукописную фиксацию ранее не слышанного слова, обычно неправильно читает его «темную» часть (проблемный сегмент) слова. Для такого читателя этот «темный» сегмент не является изначально рядом определенных букв. «Темный» сегмент рассматривается как ряд элементарных графических элементов. Эти элементы воспринимаются как ряд букв только после усилий неискушенного читателя сложить их вместе и найти приемлемое (не обязательно правильное) прочтение. Таким образом, количество визуальных элементов в «темном» сегменте является решающим фактором для определения пределов воображения читателя в таком трудном случае.

Ключевые слова: словарь, диалект, слабо документированное слово, типология очиток (ошибок прочтения), частота, тривиальность и экзотичность, русский язык.

Введение

Мы исходим из того, что неправильно опознанная запись слова в «темном»/проблемном месте не является для наивного читателя серией букв, но только рядом графических элементов, в которые он «вчитывает» (если воспользоваться метким глаголом Р.О. Якобсона) свою догадку.

Цели и задачи

Поскольку анализ диалектных словарей рядом исследователей позволил выявить значительное количество ошибок прочтения, актуальна задача анализа этого материала количественными методами.

Методы

Очевидно, что неправильный вариант прочтения может совпадать как с более, так и с менее частотными сочетаниями букв в данном языке [1]. Готовые данные о последних обычно ограничены верхней по рангу тысячью сегментов [2, с. 1051–1062]. Однако, учитывая в первую очередь наиболее очевидные контрасты, можно ответить на вопрос, не связан ли выбор неправильного прочтения с частотными характеристиками букв и их серий.

Результаты

Хотя невозможно охватить все очитки, они пока не вы-

явлены, но возможен анализ ограниченного материала. Так, частотность графем в небольшом словаре «блатной музыки» [3] можно сравнить с общеязыковыми показателями. В этом словаре нами выявлено 142 очитки [4], в которых исправления обоснованы материалом более ранних словарей [5] и др. Выяснилось, что в источнике вдвое чаще встречаются буквы **а, к, ь**, заметно реже – **е (+ё), о, я**. Эти отклонения противоречат положению о том, что «частотность не зависит от стилистической окраски текста» [6, с. 5]. При этом влияние очиток на выявленное отклонение ничтожно. Можно ли и далее считать жаргон именно стилистическим вариантом общего языка? Или же требуется пересмотреть такой статус социолекта, дискредитируемый даже на уровне употребительности букв?

Если при чтении рукописной фиксации малознакомого слова на место «тёмного» сегмента вместо правильного подставляется более распространенное в языке прочтение, то мы назовем это тривиализацией: «**Гамóка**... Ветка дерева» <курское *галюка* ‘то же’ [7, с. 187]. Например, в сетевом обратном словаре сегмент -мок- отмечен 163 раза, а -люк- – только 31. Вероятно, читающий интуитивно предпочел более тривиальное и частотное прочтение. Также и сегмент -медь явно чаще встречается, чем экзотический -хц(ъ): *беккемедь* ‘выпечка’ <(идиш) *беккехцъ [8, с. 849].

Но и обратное предпочтение экзотического прочтения тривиальному (и, следовательно, «опозна-

ваемому» с большей легкостью) также встречается регулярно: рассмотрим призрачный жаргонизм с толкованием 'работник по дому, двору' *фалешо* [9, с. 57]. В данном случае, вероятно, сыграла свою поддерживающую роль подспудная ассоциация с *фальшью*, однако нами было доказано, что источником этого описания было общеевропейское название 'слуги' *валет*, оформленное общецыганским окончанием для заимствований мужского рода -о и прочитанное с начальным немецким *v-* [ф-]. Буква *m* у нас на 20-м месте по частоте, а буква *ш* встречается в 10 раз реже и находится аж на 26-м месте [2, с. 1051]. В общем, для выбора чтения *фалешо* вместо *валето* требовалось преднамеренно или интуитивно повышать уровень экзотичности. Если речь идет о десятикратном соотношении, как в паре *m – ш*, то контраст по употребительности довольно ощутим. Но, например, в случае замены *д* на *з* (ошибочное толкование 'зверь' вместо 'дверь' [10, с. 28]), контраст между 5-й по частоте буквой *д* и 9-й *з* не столь ярок (количественно это лишь двукратное понижение встречаемости в текстах).

Трудно произносимое тобольское слово «чрмáкъ – берестяная чашка» [11, с. 129] несет явные следы экзотизирующей интерпретации. Оно было опубликовано в таком виде, вероятно, благодаря фантазии наборщика и отсутствия у ссыльного автора возможности держать корректуру. Это пример высокого уровня экзотичности прочтения довольно ординарного названия берестяной посуды *чуман*. С другой стороны, принципы отбора материала в дифференциальный словарь таковы, что

можно надёжно принять колымское «Куропáшка – куропатка» [12, с. 73]. Оно могло бы быть искажением общенародного *куропáтка*, но в словаре диалектизмов последнему нет места.

Таким образом, у экзотического прочтения довольно высокие шансы на сохранение в словаре, и отличить его от реального диалектизма порой трудно. При решении этой задачи существенно учитывать количественные показатели всего списка выявленных очиток, см., например, исправления Даля [13].

Выводы

Нами получены количественные результаты ряда сравнений очиток и конъектур в словарях территориальных и социальных диалектов по частоте буквосочетаний, выбранных в качестве правдоподобных чтений в ситуации неопределенности. В плане тривиальности и экзотичности, учитывая общезыковые данные [2, с. 1051–1062], при всех разнонаправленных колебаниях (например, большей или меньшей тенденции к предпочтению экзотических трактовок) можно заключить, что при прочтении «тёмного» места примерно одинаково вероятен ошибочный выбор как более, так и менее частотного сегмента. Это важно знать при верификации очиток. Кроме того, важно отметить, что выявленные характеристики очиток не зависят от конкретного территориального или социального диалекта, к которому приписывается призрачное слово, а определяется исключительно условиями фиксации и копирования.

ЛИТЕРАТУРА

1. Шаповал В.В. Критика словаря и гипотетическое подстановочное прочтение // Человек и язык в коммуникативном пространстве: сборник научных статей. 2018. Т. 9. № 9. С. 180–185.
2. Ляшевская О.Н., С.А. Шаров. Частотный словарь современного русского языка. М.: Азбуковник, 2009. 1112 с.
3. Потапов С.М. Словарь жаргона преступников. М., 1927. 196 с.
4. Шаповал В.В. Цыганизмы в жаргонных словарях. Saarbrücken: LAP LAMBERT Academic Publishing, 2011. 192 с.
5. Попов В.М. Словарь воровского и арестантского языка. Киев: Т-во «Печатня С.П. Яковлева», 1912. 127 с.
6. Шеворошкин В.В. Звуковые цепи в языках мира. М.: Едиториал УРСС, 2004. 184 с.
7. Журавлёв А.Ф. Лексикографические фантомы. 1 // Dialectologia slavica. 4. М.: [ИСЛ РАН], 1995. С. 183–193.
8. Шаповал В.В. О двух механизмах накопления ошибок прочтения // Лексический атлас русских народных говоров: Материалы и исследования. 2019. № 13. С. 849–857.
9. Шаповал В.В. Новые «цыганизмы» в русских жаргонных словарях // Вопросы языкознания. 2008. № 6. С. 49–77.
10. Шаповал В.В. Текст источника как объект анализа для филолога и историка. Сб. научных статей. Москва, 2001. 40 с.
11. Бартенев В.В. О русском языке в Обдорском крае // Живая старина. 1894. Вып. I. Отд. IV. С. 126–129.
12. Богораз В.Г. Областной словарь Колымского русского наречия // Сборник Отделения русского языка и словесности ИАН. Т. 68. № 4. СПб., 1901. С. 19–163.
13. Шаповал В.В. В.И. Даль и критика словарей // Русский язык в научном освещении. М., 2009. № 1 (17). С. 158–181.

© Шаповал Виктор Васильевич (shapovalvv@mgsu.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»