

Двести лет санскритской лексикографии (1819-2019)

От Вильсона до Кёльна

1. Предыстория словарей
2. Генеалогия словарей
 - а) Плагиат
 - а. в СССР
 - б. на Западе
 - б) Парность
3. Размер словарей
 - а) Количество вокабул
 - б) Потребности текстов
4. География словарей
 - а) Петербургские словари
 - а. немецкие
 - б. русские
 - б) Пунские словари
 - а. санскритский
 - б. пракритский
 - с) Кёльнские словари
 - а. настоящее
 - б. будущее

(1) Предыстория словарей

На русском языке, кроме редко доступной и скорее вводной, чем новаторской статьи недавно ушедшей от нас В.А. Кочергиной, "Из истории лексикографии древнеиндийских языков", 1983 г., которая не упоминается даже в каталоге библиотеки Востфака, новейшая лексикография санскрита в целом и в на русском языке в частности, почти не описана. Разрешите предпринять попытку заполнения данной лакуны.

Предыстория вопроса. В 2004-м году в Москве прошел XXXVII Международный конгресс востоковедов. После одного из заседаний я узнал душещипательную историю о том, как китайцами сжигались тибетские монастырские библиотеки и потом отстраивались заново.

А потом, ныне покойный, Kamalleshvar Bhattacharya поведал про свой экземпляр MW. Его японский коллега высказался в том же духе, рассказав, что у него также уйма пометок и исправлений на полях словаря. И тогда я понял, что необходимо объединить усилия, что нельзя больше только чиркать бумажные книги. Словарь нужен всем, следовательно пометки нужно вносить удаленно в доступном для остальных виде. Но в 2004-м году не то, что электронных словарей, кроме

MW, мало что вообще было доступно в сети. Начать объединять усилия систематически удалось лишь спустя десять лет, когда исправление опечаток стало впервые досконально документироваться.

Личная переписка с М. Майрхофером привела к выложенным в свободном доступе этимологическим словарям, отсканированные в Иностранке. Тогда же Томас Мальтен, основоположник оцифровки словарей в Кёльне, выслал на DVD диске сканы Петербургских словарей и ксерокс «Синтаксиса» Дельбрюка. Петербургских словарей еще не было в Сети и я решил их обнародовать. Но Мальтен посчитал, что делать так не следует и написал жалобу проф. Ивбулису - файлы пришлось снять с сервера в Риге. В последствии инцидент исчерпал себя, но я привел эту историю лишь для того, чтобы показать, что даже в эпоху интернета далеко не все столь доступно сегодня, как было еще 10-15 лет тому назад. Уже шесть лет как все наработки доступны [публично](#). [02:40]

(2) Генеалогия словарей

Если классифицировать словари санскрита по эпохам, то периодов условно четыре и речь больше об используемых приемах и подходах, чем о подлинной хронологии. Пример несоответствия методам (устаревшие) и времени создания (XXI век) – Энциклопедический Пунский словарь:

- 1) до начала XIX века - компилятивный характер (больше 1000 лет)
- 2) от середины XIX в. - золотой стандарт
- 3) XX в. - накопление материалов (критические издания)
- 4) XXI в. - корпусная лингвистика и цифровые словари

(2a) Плагиат

На Западе. Тридцать лет тому назад чешский лексикограф Ладислав Згуста в своей хрестоматийной статье про плагиат в двуязычных словарях *Copying in Lexicography: Monier-Williams' Sanskrit Dictionary and Other Cases (Dvaikosyam)* на примере 35 словарных статей из MW, пишет: "Копирование информации касается проблем этического характера и авторского права. Случаев вопиющего плагиата в лексикографии не столь много"¹.

Стоит отметить, что издание словаря MW было обречено на успех изначально. Во-первых, потому что словарь на английском языке, для которого открыт индийский книжный рынок. Без цитат, однотомный на 1200 страницах он удобен и дешев и благодаря продуманной оригинальной макроструктуре в нем находить слова легче, чем, например, в Петербургских словарях, что, впрочем, уже не актуально для электронного словаря.

И лишь в Предисловии к IV-му т. своего словаря Бётлинг обрушивается сокрушительной критикой на Моньера, обвиняя его в "простом и механическом переводе, без проверки или критицизма", в самом низкопробном плагиате. Згуста

¹ и "overextensive, not acknowledged use of their dictionary by other scholars."

рассматривает приводимые Бётлингом примеры и распределяет их на следующие подгруппы, где MW:

- 1) пересказывает общую идею;
- 2) повторяет без изменений один в один;
- 3) видна некоторая самостоятельность;
- 4) не вполне убедительно предъявляет претензию.

То есть Згуста не намерен очернить составителя словаря лишь потому, что он много копировал, пытается установить истину. Новый свет на происходившее тогда проливает недавно вышедший томик личной переписки Бётлинга к Роту.

в СССР. Вопрос о несоразмерно большом количестве заимствований, без сделанной подобающим образом оговорки, в санскритской лексикографии встает не впервые и у нас. Санскритско-русский словарь В.А. Кочергиной – согласно моим наблюдениям - переводной словарь санскритско-немецкого Капеллера. Оба ровно 29000 слов.

Другой пример, Рерих Ю.Н., Тибетско-русско-английский словарь — 1983-1987 содержит чуть больше 90 000 словарных статей (всего 2400000 слов «внутри»), но тоже скорее является переработкой труда индийского автора, чем оригинально расписанной картотекой. Впрочем требуется более детальный дополнительный разбор .

(2b) Парность

В истории санскритской лексикографии наблюдается определенная парность, удваивая и так порой монументальные достижения. Перечислю некоторые: 1) два словаря Böhtlingk'a в XIX в.²; 2) два словаря Монье-Вильямса – прижизненный (мало известен) и канонический-посмертный; 3) два словаря Апте - прижизненный одготомный и трехтомный 70 лет спустя; 4) два толковых санскритско-санскритских словаря - SKD на 40000 и VCP 48500 слов; 5) два словаря Майрхофера в XX-м. KEWA (краткий) и EWA (полный) и 6) два словаря до середины буквы «а» - PD 105000 и GST (Goldstücker) 7000. А в Российской Империи два брата - Коссовичи и Шерцли - издававшие книги по древним языкам, в том числе по санскриту.

(3) Размер словарей

(3a) Количество вокабул

Словари бывают карманными, среднего размера и полными. Карманных словарей санскрита в прямом смысле нет, но есть глоссарии от 1500 слов. «Большой» словарь Кочергиной среднего размера - содержит 29000 слов (~~BOP - 8500; GRA 10000; BEN - 17000; BHS - 18000; BUR - 19000; Macdonell - 20000; STC - 24000; CCS - 29000; AP90 - 32000, AP - 36000; CAE 38500; Wilson 44000~~).

² дополнение 60 лет спустя, Nachträge zum Sanskrit-Wörterbuch, публикация писем в 2007-м г., 870 стр.

Словарь Моньера полный и исчерпывающий – 194 000 слов (против MW72 – 51000 и PW (краткий) 132000). При этом толщина словаря Моньера равна толщине словаря Кочергиной. Количество статей уместается за счет формата и устройства книги.

(3b) Потребности текстов

Для перевода большей части корпуса санскритских текстов требуется до 72000 слов. Но при этом даже для перевода 17 томов Махабхараты - одновременно не больше 20000 слов. Хороший словарь от плохого отличается не количеством словарных статей, а скорее качеством подбора слов и полнотой описания. Для перевода Йогасутр нужно всего 500 слов, но доступные переводы на русский язык имеют известные недостатки. В три раза больше слов в Кавьядарше в переводе Алихановой (1600 слов), однако перевод великолепен, то есть дело не только в количестве. Вся Амаракоша, крупнейший санскритский словарь синонимов, всего 4000 слов.

Ригведа – 10000 слов, почти столько же в Чаракасамхите – 8900 и в Рамаяне – 11 000 слов. Но, из встречающийся в Ригведе, в Чараке - лишь 1400 слов, а в Рамаяне - 1700, то есть пересечений достаточно мало, хотя и количество схоже.

Если рассуждать вульгарно-статистическим методом, чем больше слов, тем больше должно быть и совпадений. Доля правды в этом есть. Махабхарата – 20000 разных основ и корней. Пересечений с Ригведой 2400 слов, а с Рамаяной уже все 8600. Для типологического сопоставления, «Словарь языка Пушкина» включает в себя свыше 21 000 слов.

Но если по параметру количества слов словарь Кочергиной, например, вполне вписывается в рамки серии и должен был бы удовлетворить потребности эпоса, то со значениями дело обстоит существенно хуже. Особо это касается технических терминов³. Ближе к требуемому для перевода количеству статей подходит санскритско-каталанский словарь на 61 000 слов, изданный в 2006 г.

(4) География словарей

(4a) Петербургские словари

³ Строго филологический метод, т. е. метод дословного перевода, здесь совершенно неприемлем. В данном случае особенно явствует правильность точки зрения акад. Ф. И. Щербатского, который неоднократно подчеркивал, что слова, хотя заимствованные из обыденного языка, но получившие в известной научной системе положение технических терминов, никоим образом нельзя переводить в их обычном значении. Как во всех философских системах Индии, равно как и в индийской математике, астрономии, политико-экономических учениях, в грамматике и теории поэзии, — точно так же и в индо-тибетской медицине имеется определенная терминология. Правильно передать ее смысл — вот к чему мы должны стремиться в наших работах. Главное — смысл; буквальный же перевод всегда будет непонятным и создаст совершенно превратное представление об оригинале. Е.Е. Обермиллер. Пути изучения тибетской медицинской литературы / Библиография Востока. 1936. Вып.8-9, с.48-60

(4а, а) немецкие

Петербургские словари эпохальны. Мы продолжаем жить в эпоху после Петербургских словарей, «лучшим нашим источником санскритологии» словами Шерцля, и все англоязычные словари, в частности, вышедшие после - это разной степени вытяжки именно на основе семитомных малого и большого Петербургских словарей.

Улучшить их невозможно. Но эпоха корпусной лингвистики и цифровых словарей дает возможность сделать работу удобнее даже с 5000 страничным *magnum opus*. Четвертое поколение словарей пока лишь зарождается, но приблизить этот день можно лишь работая совместно, а не по отдельности. Забегая вперед, какие преимущества дает использование вычислительной техники в санскритологии?

- 1) Умный режим поиска (1), основанный на частотности корпуса Оливера (2), то есть 560 000 грамматически размеченных предложений на древнеиндийском языке, где, например, через два клика можно узнать, что *manas* в Ригведе встречается 275 раз из 170860 словоформ всего;
- 2) Сводный список слов в санскрите (с унифицированной орфографией, 4, и в оригинальном написании, 3), на основе которого построен мой Обратный словарь санскрита (5);
- 3) Параллельный корпус (6) Ригведы (2013) и Атхарваведы (2019), а также частично Махабхараты (2019) и Рамаяны (2019) с русским переводом и примечаниями. Что в частности дает возможность сделать:
- 4) Перелинкованные цитаты с оригинальными текстами, например, словарь Грассмана с переводом Т.Я. Елизаренковой (7), а также немецким и английским переводом (доступно с 2018 г.). Следующим этапом можно сделать, например, кликабельными 8000 отсылок к Панини у MW или связать упоминания флоры и фауны с современной базой данных биологической номенклатуры.

О Петербургских словарях сказано много, поэтому я сосредоточусь в основном на словарях Пунских и Кёльнских.

~~Быстрая карьера ученого объясняется вкладом (9/10), который он внес в индийскую филологию изданием (совместно с немецким индологом Р. Ротом) санскритско-немецкого словаря. Это издание составило целую эпоху в изучении древнеиндийской культуры и подвело итог всех индологических исследований предшествующего периода. В словарь была включена полная роспись всех изданных санскритских текстов. Каждое приведенное в словаре значение санскритского слова было подтверждено всеми контекстами его употребления. Словарь в семи томах издавался Петербургской Академией наук с 1852 по 1875 г. В 1879—1889 гг. была подготовлена краткая редакция этого словаря, тоже в семи томах. Хотя другое издание названо кратким, сокращение объема было вызвано уменьшением количества отсылок к первоисточникам. Количество же санскритских слов в нем значительно увеличено.~~

(4a, b) русские

Больше, чем напечатанных, санскритско-русских словарей запланированных, не изданных и не оконченных.

1. Отделение. 2 Санскритская Антология Петрова - 1846-1875 [istor-vostokoved 340], где "во второй части антологии приводился этимологический словарь к текстам" и он "собрал материал для санскритско-русского словаря, но закончить словарь так и не успел";

2. словарь Коссовича 1854-1856, труд остался незаконченным, недавно переиздан;

3. словарь Шерцля, "для командировки в Лондон с целью подготовки санскритско-русского словаря" - 1872 [7522-15059-1-SM];

4. словарь Кудрявского 1918, картотека утеряна в Воронеже во время гражданской войны;

5. картотека словаря В.А. Кочергиной, утеряна в архивах издательства «Русский язык».

~~В 2005-м году при личной беседе, отвечая на мою жалобу, как же много опечаток в словаре Кочергиной, А.А. Зализняк рассказал, что еще в 70-е передавал список опечаток на нескольких страницах, но что дело обстоит ровно так же, как и до этого.~~

(4b) Пунские словари

Я провел месяц в г. Пуне, знакомясь с работой Deccan College и BORI. Работа над словарями ведется в каждом из них, но при этом на совершенно разном уровне подготовки. Напомню, PD 105000 и GST 7000 - это не просто неполноценные, а заведомо незавершабельные словари.

В 2014 г. в предисловии к репринту Фриша я писал про 1800 лет, которые потребуются, чтобы завершить новый крупнейший в мире словарь санскрита, призванный сбросить с корабля современности петербургские словари. Я, пожалуй, ошибался.

(4b, a) санскритский

Пунский словарь не только никогда не будет закончен, но такой задачи никто, как выяснилось на месте, и собственно не ставил – завершаемость как критерий у Майрхофера (согласно Трубачеву) не в почете. Главный редактор, назначенный в 2017 г., Прасад Джоши (Dr. Prasad Joshi) ставит целью куда более скромную задачу: готовить по 200 страниц к печати в год. До него, с 1948 г. было уже восемь главных редакторов, включая двух женщин. Каждое из последних изданий содержит около 200 страниц. Последний том (vol. 31) вышел в 2015 г. Его

отношение к делу размеренно, поступь неспешная. С учетом того, что напечатано аж 5408 страниц в 11 томах, (от अ до अप्रमाणसमाधिचर्या, больше 105 000 словарных статей) это процесс, который можно созерцать, но который вряд ли будет завершен до конца нашей калпы.

В течение 25 лет (с 1948 по 1973) сорок, в том числе выдающихся санскритологов с мировым именем (М. Gokhale, D.G. Joshi, D.M. Deo, Palsule) делали выписки из 1476 книг. Эти полторы тысячи источников А.М. Гхатадже (1913-2003) изначально распределил по собственной небесспорной классификации, вводя 62 таксона. Деление это весьма искусственное: частично на жанровой основе, частично - по содержанию. Приведу несколько примеров, агрикультура, алхимия, аллегорические пьесы, антологии, араньяки, артхашастры, астрология, астрономия, но тут же, отдельно, Шанкара и джайнизм (догматы) и джайнизм (рассказы).

Норматив был 200 выписок в день. Единого стандарта не было, выписки (карточки) очень разные по полноте описания, но в основном не содержат даже базовых сведений, как, например, предполагаемое значение слова. Каждую выписку (цитату) после перекрестно проверял другой редактор. Получилась картотека, которая занимает 3082 крупных жестяных ящиков в огромном, пыльном зале, называемом Scriptorium, миллион вокабул.

~~Карточки бывают трех цветов: обычные, зеленые и розовые. Немногочисленные зеленые карточки - выписки из PW. Сделаны они поверхностно, без переноса значимых меток (отмеченное в словаре ошибочное прочтение слова, например).~~

Словарь с таким же успехом мог быть создан и в XVIII веке, в нем нет ничего даже от начала XIX-го. Даже желание упорядочить источники в хронологической последовательности - яркий признак словарей XIX века. Ввиду того, что после принципиально новые словари не создавались, а лишь копировались канонические, то никакой новизны, как в Пунском словаре, например, нет. Это даже собственно не словарь, а сборник цитат, некий прообраз корпуса текстов, только рукотворный. Ведь какой-либо строгой процедуры по определению или уточнению значений нет - при отсутствии, например, аюрведиста, максимум, на что можно рассчитывать это копипаст из двухтомного словаря аюрведических терминов Джоши. В течение полтора века выходили статьи, уточняющие отдельные словарные гнезда словарей, но с учетом того, что вторичная литература в Пуне намеренно не привлекается, снова тиражируются уже порой известные неточности. Образ лягушки в колодце, которая не подозревает о существовании окружающего мира весьма уместен.

Как устроена работа над словарем мне и Нараяну Прасаду показал молодой кандидат наук Бхав Шарма (Bhav Sharma). Программа для обработки словаря, специально разработанная для нужд местных лексикографов лет 20 тому назад,

так и называется – «Deccan College Dictionary Building» и данный суффикс «-ing» равен знаку бесконечности.

В зале, где работают лексикографы, намеренно отключен интернет. Спасибо, что не свет. Получить доступ извне невозможно - словарная база данных хранится на локальном сервере, на интранете. То есть перед нами вариант kīra-maṇḍūka-puṣya. Это изначально намеренно отрезанная от мира средневековая разработка, которая устарела морально еще до ее зарождения. Именно поэтому на каждом столе лежит по бумажному репринту кирпича MW. Редактора даже не могут воспользоваться электронным кельнскими словарями, впрочем, не совсем так. Они рассматривают, прячась, словарные статьи в телефоне. В случае Шармы это треснувший на мелкие осколки маленький экран, в котором он пытается разглядеть контур слова на деванагари. Одновременно с этим большой монитор на рабочем столе не включен, не востребован.

В библиотеке, в которой по протоколу положено работать 45 лексикографам, видно два аспиранта - девушку и парня. Вокруг по периметру бумажные книги, из которых в течении предшествующих десятилетий делались выписки, сотни тысяч выписок от руки. Это было в доисторические времена и каторжный труд, который оказался не востребованным.

Грубо определено соответствие для 85% карточек в отсканированных текстах. Это наиболее сильная сторона и настоящее достижение словарного факультета (а точнее программистов), однако, как и другие современные индийские, это закрытая неюникодная разработка и воспользоваться ею извне, как уже говорилось, невозможно, а внутри – она не дойдет даже до конца буквы «а», пока первые тома не превратятся в труху. Но, ввиду того, что лягушка в колодце, никто об этом не узнает. Это строго засекреченный проект и скопировать даже сканы книг - запрещено. Но скажу более того - до совсем недавнего времени и остальные разработки по санскриту и не только в Пуне были тоже с закрытым кодом. Партизанить, так до победного конца.

На вопрос, что хочет успеть сделать Шарма в связи со словарем, он отвечает уклончиво. Мол, ему осталось 27 лет до пенсии, при этом совершенно очевидно, что мы с ним одного возраста, и уж, не судите строго, сколько успеет, столько успеет. ~~Изначально задумка составителей словаря заключалась в том, что не будет учтена никакая вторичная литература. То есть сотни уточняющих статей и комментариев здесь никого не интересуют? Имеет ли смысл тогда такой словарь, который устарел до того, как зародился? На мой вопрос, а что вы делаете, когда встречается аюрведическая терминология, он говорит – открываю разные словари, включая двухтомный Джоши. Не держать же аюрведиста, не так ли? То есть это исключительно уместная, изначально не рассчитана быть прикладной работа на уровне конца XIX в.~~

~~Принципы работы над словарем, как и многое тут, устарели еще до того, как к нему приступили. Непосредственно в картотеке (Scriptorium) работа как таковая уже не ведется, она заперта на ключ. Все карточки (slips) давно отсканированы, правда как и большинство индийских сканов – в отвратительном качестве. То есть не совсем так – работая над словарем редакторы забирают себе на стол небольшие стопки карточек, перевязанных обувными шнурками, поглядывая туда для перестраховки. Перед тем, как карточка станет частью черновика словаря, она проходит еще два круга корректуры.~~

Хотя картотека оцифрована (как обычно в Индии в сомнительном качестве), составители продолжают пользоваться бумажными карточками. На столе у каждого перевязанная шнуром стопка карточек. Из 48 рабочих мест сейчас над словарем гипотетически работают 14 студентов. Однако во время моего посещения составителей, помимо главного редактора Прасада Джоши (Dr. Prasad Joshi) и Шилпы Сумант (Dr. Shilpa Sumant), на месте было не больше трое, включая Бхав Шарму (Dr. Bhav Sharma).

(4b, b) пракритский

Принципы работы пракритского словаря в BORI мало отличаются от санскритского, так как у них один и тот же основоположник. В плане технологической отсталости тут все еще хуже, хотя оба заведения находятся в часе езды друг от друга (в BORI все делается исключительно в ручную). При этом пракритский словарь продвинулся гораздо дальше санскритского соседа и дошел до начала «ка» (до вокабулы «kava-vaya»).

(4c) Кёльнские словари

Индийская лексикография не может продолжать делать вид, что компьютер еще не изобретен. То, на что раньше потребовалось месяцы работы, теперь можно сделать за считанные секунды. Это не означает, что корпус сам делает словарь, но рутинная работа теперь иная.

Приведу пример. С помощью компьютерного алгоритма нечеткого поиска Fuzzy составлен, например, список из сотен слов-кандидатов (636 шт.) на наличие опечаток в словаре Кочергиной, который потом проверялся нами в ручную. Речь о 2% словарных статей, которым нет аналогов ни в одном из 33 предшествующих словарей, что, согласитесь, странно: *aduḥka* в м. *aduḥkha*, *bāṣpasālila* в м. *bāṣpasālila*, *vipratopa* в м. *vipratīpa*.

Эпоха корпуса пока не наступила. Те, кто используют его достижения, делают это робко, остальные же даже не подозревают о существовании или о предполагаемой пользе корпусной лингвистики для древнеиндийской лексикографии. Премного благодарен!

