

*А. В. Коваленин, Е. Ю. Нечипоренко
(Фонд знаменных песнопений)*

Корпус знаменных песнопений как инструмент исследователя

Качественное отличие современных исследований текстов от их изучения в недавнем прошлом состоит в возможности опоры на накопленные в электронном виде массивы данных – корпуса, словники, разного рода грамматические таблицы. Всякое научное предположение проверяется на множестве примеров, по множеству источников. Наличие электронного корпуса позволяет и ускорить эту работу, и качественно изменить круг возможных вопросов к текстам. Поэтому в языкознании опора на электронные корпуса стала уже неотъемлемой чертой серьезной научной работы.

Сегодня есть все возможности для того, чтобы корпусные исследования пришли и в музыкальную медиэвистику для изучения не только словесных, но и крюковых и нотолинейных текстов знаменного распева. Чтобы такие корпуса становились максимально полезным орудием труда, а также основой для взаимодействия учёных, их устройство и теоретические основания должны стать предметом обсуждения специалистов. Для этого в своем докладе мы расскажем об устройстве созданного нами три года назад «Фонда знаменных песнопений»¹ (ЗФ) и покажем некоторые способы использования этого корпуса в работе исследователя.

§1. Постановка задачи

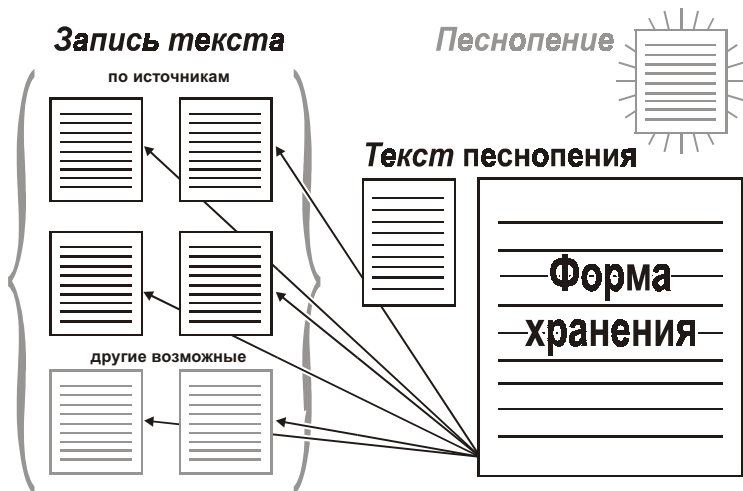
Проект изначально имел чисто практическую направленность – для возрождающейся традиции знаменного пения остро не хватало певческого материала. Предполагалось набирать и выставлять для общего доступа песнопения и их расшифровки, чтобы обеспечить необходимую для богослужения полноту. С другой стороны, необходимость ориентироваться в большом количестве рукописей заставила создавать такие описания рукописей, по которым можно

было легко находить требуемое песнопение. Имелось в виду (и эта надежда частично оправдалась), что Фонд знаменных песнопений делается вскладчину, то есть песнопения (а теперь можно говорить и о встроенных аналитических средствах) набираются и присылаются нам из разных источников разными людьми. Такие люди, как правило, разбираются в крюках, но не обязательно являются опытными пользователями компьютера. Отсюда возникло первое требование к способу набора песнопений – **требование простоты**. Набор в ЗФ ведется в простейших текстовых редакторах без использования специальных средств. В идеале, набрать песнопение должно стать не сложнее, чем переписать его от руки.

Практическая направленность проекта внесла своеобразие в постановку задачи о форме хранения песнопения, поскольку разные группы пользователей предъявляют разные требования к представлению текста. Если для исследовательской позиции важно точное соблюдение всех деталей оригинала, то для практического использования на клиросе текст должен подвергаться какому-то преобразованию, в зависимости от степени подготовленности клироса и приверженности той или иной графической и орфографической традиции. Одни хотят получить нотные расшифровки, другие нет; одним нужны крюковые разводы фит, другим они уже мешают. В оригинале могут быть отклонения от нормативной крюковой записи, которые для практики требуют устранения. Таким образом, запись, точно передающая оригинал, является лишь одним из возможных представлений текста, хотя и безусловно необходимым.

Это обуславливает **требование гибкости**: форма хранения песнопения должна быть такой, чтобы из нее можно было получать разные его записи. Для этого выбор интерпретации каждой детали, допускающей различное выражение, должен оставаться за пользователем, хотя сама деталь должна быть зафиксирована. Найти такую форму – значит на практике отличить текст (то, что записано) от его записи (то, как записано). Таким образом, само понятие текста, когда речь идет о техническом воплощении, приобретает для нас редуцированное, но конкретное содержание: *текст* понимается как форма, способная породить все возможные *записи текста*. На деле, двигаясь эмпирически, мы достигаем способности порождать записи текста только в известных традициях. С появлением необходимости учесть еще одну «традицию» (т. е. набор требований к представлению текста) мы должны быть готовы увидеть и отметить новые точки отличия записей текста.

Рассмотрим принятые в ЗФ конструктивные решения.

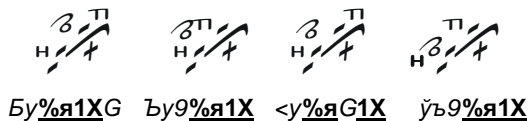


Текст и запись текста. Текст как форма хранения информации о возможных записях текста.

§2. Представление знаков письменности

Опыт непосредственного набора знамен специальными шрифтами показал свою бесперспективность для создания корпуса, обнаружив множество недостатков.

а) *при наборе*. Певческое знамя – составной знак, к основному элементу добавляется несколько дополнительных, некоторые из которых должны быть покрашены «киноварью». Поскольку большинство дополнительных элементов при наборе не сдвигают курсор, выбрать нужный из них для редактирования затруднительно. В хорошем шрифте многие элементы (в частности, текстовые ударения, киноварные пометы при певческих знаменах) имеют несколько позиционных вариантов, подобрать нужную комбинацию из которых трудно (*пример 1*).



Пример 1. Некоторые из вариантов набора шрифтом одной стрелы.

Здесь нижняя строка показывает последовательность знаков крюкового набора простым текстовым шрифтом. Как видно, отличается не только состав знаков, но и порядок их набора.

б) *при хранении*. Неодинаковый выбор порядка дополнительных элементов и их позиционных вариантов, необходимое использование шрифтов в различных кодировках приводят к невозможности самого простого поиска по тексту.

в) *при отображении*. Невозможно сменить шрифт на шрифт в другой кодировке.

Необходимость использовать разные шрифты умножает описанные трудности, заставляя исследователя осваивать новые кодировки. Без практики навыки теряются, и многие интересные шрифты пропадают для использования.

Таким образом, при наборе шрифтами исследователь вынужден заниматься преодолением технических трудностей, далёких от содержания его работы.

Причины этих неудобств кроются в очевидной, хотя не широко осознанной концептуальной ошибке: в цепочке действий «набор – хранение – отображение» шрифты и их кодировки должны по своей природе играть роль только на последнем этапе. Однако эти процессы обычно сливаются, как будто отображение является единственной задачей набора, что порождает множество проблем даже для не очень сложного набора с диакритикой. Сложность набора крюков только помогла увидеть неудачность этого подхода.

Сказанное о шрифтовых кодировках не может не относиться и к такой из них, как Unicode. Идея зафиксировать необходимый набор семантически различных начертаний становится иллюзорной, когда речь идет о рукописях, тем более в глубоком диахроническом срезе, тем более, когда семантика знаков сама по себе остается научной проблемой. Поэтому исследователи рукописей острее всего заинтересованы в том, чтобы вернуть шрифтам их правильное место и утвердить другие способы хранения знаков.

Общее решение, которое мы приняли для хранения знаков в корпусе, состоит в том, чтобы для обозначения знаков любой письменности использовать только такие символы, которые набираются непосредственно с клавиатуры компьютера (*терминальные символы*). Это русские и латинские буквы, знаки препинания и некоторые специальные символы кодировки ASCII. Такие представления письменности мы называем *терминальными представлениями* (ТП). В ТП всякий элемент набирается единственным способом, что сосредоточивает внимание на том, что важно для текста, и позволяет отвлечься от того, что важно только для его шрифтовой записи. Например, показанная в примере 1 стрела в ТП знамен выглядит так: **нпб_./+.** Отображение помет **т** и **з** здесь не зависит

от их точного места в знамени, крыж под оксией и во всех других знаменах обозначается +, одно очко у стрелы, палки, статьи, крюка и стопицы – точкой. (Обратное, впрочем, неверно: в отличие от набора шрифтом, за отдельным терминальным символом не следует видеть знак, за одинаковым – одинаковый. Например, ꙗ в ТП церковнославянской письменности обозначается ꙗ̑, но это не комбинация двух самостоятельных знаков ꙗ и ь.)

Решение об использовании ТП для крюков и нот имело прототип в виде представления церковнославянской письменности, разработанного интернет-сообществом «Славянская типографика»² (ССТ) под руководством М. Гринчука («стандарт НІР»³). В этом стандарте набран почти полный корпус богослужебных текстов («Библиотека святоотеческой литературы»⁴, больше 20 мегабайт), опыт использования которого позволил оценить преимущества использования в корпусе терминального представления:

1) Соблюдается исходное требование простоты – ни для набора, ни для чтения не требуются специальные редакторы.

2) Сохраняется точность передачи текста, не исключая передачи знаков, не предусмотренных в шрифтах. Например, в наших церковнославянских шрифтах нет надстрочной ꙗ, в крюковом шрифте нет признаѡв, но это не мешает фиксировать эти графические особенности в корпусе.

3) Следствием этого является неограниченная гибкость системы обозначений. При необходимости зафиксировать дополнительный знак или графический вариант, оказывается несложным ввести новое обозначение. Для исследований рукописей это является принципиально необходимым в силу имеющегося противоречия: чтобы набрать рукопись, надо исследовать состав знаков в ней, но чтобы исследовать, удобно сначала набрать текст, для чего требуется ввести хотя бы временные обозначения графических вариантов знаков, польза от различия которых находится ещё под вопросом.

Например, при составлении заявки ССТ на расширение церковнославянского диапазона Unicode, обсуждался вопрос о необходимости различия вариантов надстрочной ꙗ (без титла и с разными титлами). Наличие в корпусе отдельных обозначений для этих вариантов позволило подобрать примеры, которые, однако, не выявили полезность этого различия, то есть предусмотренное многообразие обозначений скорее всего следует признать избыточным. Противоположный пример – нехватки обозначений – возник с набором знамени «светлая стрела с сорочьей ножкой». В нашем формате знамен для неё имелось одно обозначение `_:/u`, так как

позиция сорочей ножки казалась неважной. Но при наборе пореформенного Октоиха конца XVII века⁵ квалификация наборщика⁶ позволила увидеть, что в данной рукописи разные положения сорочей ножки имеют разное певческое значение. Пришлось добавить обозначение *_u:!*.

4) Гибкость системы обозначений не противоречит её относительной наглядности. Хотя с технической точки зрения обозначения могут быть хоть словесными (т. е. можно было бы так и писать: *[ять]*, *[стрела мрачная полукрыжевая]*), для конкретной письменности удастся подобрать лаконичную систему обозначений, чтобы, например, острое ударение обозначать знаком ' , а не *[острое ударение]* или *&aigue;*. Такая система позволяет в повседневной работе с корпусом обходиться без перевода ТП в более наглядное (шрифтовое) представление.

5) В результате наборщик (исследователь) работает в более естественной для его работы знаковой системе, сосредоточиваясь на значимой для него информации, а не на технических проблемах её отображения.

6) Исчезает необходимость освоения различных кодировок и позиционных вариантов символов в каждой из них. Необходимо знать только одно ТП для каждой письменности.

7) Это знание позволяет упростить общение специалистов при электронной переписке. В письмах и на интернет-форуме⁷ мы уже не пишем *стрела мрачная полукрыжевая*, а пишем *_./+*. Это позволяет точно цитировать большие фрагменты текста. Таким образом, знание ТП становится необходимым элементом современной культуры исследователя.

8) Набранное легко может быть переведено в запись шрифтом, выбор которого не обусловлен какой-либо одной кодировкой, и изменение которого не требует изменения единицы хранения.

9) Над записью в ТП удобно производить содержательные преобразования, создание которых становится доступным специалисту в соответствующей письменности, а не в программировании.

10) Фрагменты разных письменностей могут легко встраиваться в одну единицу хранения.

Для корпуса знаменных песнопений мы, во-первых, расширили имевшееся представление церковнославянского текста для удобства набора словесного текста рукописей и, во-вторых, разработали свои ТП для знамен (пример 2) и нот.

<p>знамена</p>  <p>МК: Ч Б/: Г=\ НТ_/:</p>	<p>ноты</p>  <p>г1*н1 ц1г1 н!г!н1 *н4</p>
<p>старославянский</p> <p>раждаѡшти б_ыль благѡ</p> <p>раждаѡшти б_ыль благу</p> <p>церковнославянский</p> <p>ГДН, ѠЧНСТН ГРЪХН НАША:</p> <p>гд\си, w=чи'сти гръѡхи`на'ша</p>	<p>греческий</p> <p>ἀπὸ κεωφύλλων βοτάνης</p> <p>a=по` kewfv'ллwn бота'ни_с</p> <p>a=po` kewfv'llwn bota'nh_s</p> <p>хирагана</p> <p>し、あ わ れ め よ</p> <p>шу a-ва-ре-ме ё</p>

Примеры терминальных представлений
разных письменностей

*Г *Н Ц Г Н С М П В *М *П *В Р # З Л И Б Т У Ш
Г Н Ц Г Н · М П Ш Ѡ П Р 3 Л Ѡ Ѡ У Ѡ
А К К. К: К.: К! К- К? Кч Кч? Щ Кх ^х

Т Т! Т? Т. Т.! Т: ТЭ Э Э! Э+ Э: Э- тЭ!- \ \? л\ \:

= =Э =+ =. =: =\ =\ =ч =ч Д Д' Д- Дд Дч тДЭч

/ч /ч :/ч :/ч /: т!/: / /о /шз /ч т_/ :/?

_:/о _:/- _:/-о ,, /- /+ /: ,+:/ */ **/ /:о

З З: ЭЭЧ ЭЭ\ У У: О_ О=Э Ч Ч. Р Ф +


Пример 2. Терминальное представление знамен столбового распева.
Строгое описание формата знамен можно найти на сайте
«Фонда знаменных песнопений»: <http://znamen.ru/fmt-znam.php>.

Терминальное представление знамен разработано для столпового и демественного распева. Оно нуждается в придирчивом рассмотрении специалистами с точки зрения его полноты и возможности расширения (например, для путевой нотации, или для более древних знамен), поэтому на его строении остановимся подробнее.

Обозначение знамени состоит из опорного знака и знаков вспомогательных элементов (очки, облачки, подчашие, оттяжки, отсеки и пр.), конкретное значение которых определяется опорным. Вспомогательные элементы разбиты на три группы. Ближе всего к опорному знаку выписываются тушевые элементы, затем вокруг знамени – киноварные указательные, в последнюю очередь – киноварные степенные: [³ м [² т [[К] :ч ¹] ? ²]].

Обозначения знаков выбраны из соображений а) наглядности, б) мнемоничности. Примеры опорных знаков:

- а) = (статья), 1 (палка), + (крыж), Э (запятая), Ф (фита), У (паук);
основания стрел: _ (простая), „ (громная), ,+ (крыжевая);
б) К (крюк), Ч (чашка), А (параклит), З (змийца), Р (рог),
Т (стопица), Х (хамило).

Эти два принципа читаемости проводились, однако, не в ущерб разделению сущностей.

1. *Различные вещи обозначаются различно*, даже если в рукописи выглядят одинаково. В частности а) основание простой стрелы получило отличное от статьи обозначение; б) киноварный крыж используется в рукописях в трех смыслах – как знак нижнего согласия, как обозначение странного голоса и как синоним пометы Г, поэтому получил в записи текста три разных обозначения – соответственно *, # и з; в) тихая скамейца получила обозначение, производное не от крюка, с которым графически совпадает, а от скамейцы: !/: (!: – обозначение скамейцы).

2. *Одинаковые вещи обозначаются одинаково*, даже если в рукописи выглядят по-разному. Так, есть одно обозначение для всех подверток (у крюка, палки, статьи), независимо от их вида и расположения; одно обозначение для очков у крюка, стрел, палки, стопицы, статьи и голубчика. Одно обозначение применяется для пометы Г, которая в разных рукописях может записываться по-разному: Г, Гн, г. Это является графической особенностью рукописи, поэтому информация о ней отмечается в особом поле описания рукописи, что позволяет, когда требуется выдача в оригинальном виде, точно отразить и эту особенность. Так же отражается различие и тождество \wedge и ГП.

Терминальное представление нот устроено просто. Обозначение ноты складывается из двух частей: обозначение ступени звукоряда (как в формате знамен: *г, *н, ц, г, н, с, м, л, в, *м, *п, *в) и обозначение длительности (! – восьмая, 1 – четверть, 2 – половинка, 4 – целая).

§3. Разметка

Запись песнопения, как это видно в примере 3, соединяет фрагменты различных письменностей. Каждый фрагмент предваряется особым словом – меткой (они обычно начинаются со знака %), которая только именуется способ интерпретации фрагмента, чтобы содержание интерпретации можно было задать в отдельном, «стилевом» файле. Разные пользователи (группы исследователей) используют разные стилевые файлы, что позволяет преобразовывать единицу хранения в разные формы – например, в наглядный вид или в вид с аналитической разметкой. В частности, %з, %н, %т – метки знаменного, нотного и словесного фрагментов. В этом примере в каждом из них помечены также участки фиты и ее развода, чтобы пользователь мог по своему желанию отказаться от них или придать им требуемое оформление. Особой интерпретации также подвергаются обозначения в скобках (эти скобки, таким образом, также являются своего рода метками) – в крюковой строке это киноварные подобны, в словесной – огласовки. Кроме фрагментов, отображающих все три текста песнопения, в файле содержатся и фрагменты, содержащие внешнюю и служебную информацию.

Такая разметка, как видно, теоретически минимальна. В ней ровно столько меток, сколько поводов для смены интерпретации.

Между фрагментами нет и не должно быть предписанных форматом иерархических отношений, ибо всякая иерархия возникает уже как следствие взгляда на текст, определяемого аспектом исследования, т. е. является частью интерпретации. С одной точки зрения текст может быть разбит на праздники и стихиры, с другой – на страницы и строки, с третьей – на оригинальный и восстановленный текст, и т. п. Таким образом, в тексте можно усмотреть разные, независимо сосуществующие иерархии, внутреннее устройство которых должно задаваться не в единице хранения, а там, где описывается интерпретация разметки.

Дисциплину работы с требующими различной интерпретации текстовыми фрагментами мы назвали *Технологией смешанного набора* (ТСН). Метки в ТСН играют ту же роль, что «стили» в обычных текстовых процессорах (например, MS Word), но их интерпретацию отличают две принципиальные особенности.

1. Если интерпретация метки фрагмента не задана, то фрагмент уничтожается. Это решение было принято как вынужденное, ибо если просто проигнорировать, например, метку %T в примере 3, то словесный фрагмент станет интерпретироваться как крюковой. Но у него оказалось исключительно полезное следствие.

Убирая описание интерпретации метки из стилевого файла, можно легко отказаться от ненужных конкретному пользователю фрагментов. Так, один пользователь (*пример 4 а*) отказался от нотной расшифровки, другой (*пример 4 б*) – от знамен. Также никому из них не отобразились содержащиеся в файле сведения о наборщике.

Этот приём имеет и другую сторону: если исследователь впишет свои рабочие замечания, предварив их собственной меткой, которой не придаст никакой интерпретации в стилевом файле, то это никак не повлияет на получаемое им представление песнопения. В примере таким замечанием является пагинация.

В результате разные группы пользователей могут работать со своими фрагментами текста (например, отвечая за свои аспекты исследования); стилевой файл становится носителем взгляда этой группы на текст: что в нем для них важно, а что нет.

2. Вторая важная особенность интерпретации стиливых меток – это возможность связать с ними не только внешние (оформительские) параметры для их фрагментов, но и **набор содержательных преобразований**, применяемых к этим фрагментам. Перевод терминального представления в кодировку требуемого шрифта – отнюдь не единственное из них. Большинство преобразований (такие как смена регистра, раскрытие титл, снятие ударений, выбор орфографического варианта) оставляют фрагмент в терминальном представлении.

Чтобы такие преобразователи могли находиться в ведении специалистов-непрограммистов, создана возможность их реализации не только в виде скриптов на алгоритмическом языке (RHP), но и в виде списков производимых специальным образом текстовых замен. Это позволяет специалистам самостоятельно переопределять или дополнять системные преобразователи, создавая собственные таблицы замен для чисто лингвистических алгоритмов, таких как, например, выделение приставок или простановка ударений. На практике большинство пользователей использует в работе готовые преобразователи и предписанные техническим редактором корпуса метки.

§4. Описание рукописи

В Технологии смешанного набора набирается не только песнопение, но и описание рукописи. В описании, кроме служебных фрагментов, описывающих внешние характеристики рукописи, используются три особых фрагмента:

- помеченные %ред – для замечаний редактора корпуса;
- помеченные %т – для текстов из рукописи, не входящих в песнопения (заголовки, уставные указания и др.)
- помеченные %% – специальные строго организованные фрагменты-ссылки на песнопения. Такой фрагмент содержит шифр песнопения по структуре службы, номера начальной и конечной страницы, инципит.

Чтобы песнопение можно было найти в корпусе по инципиту или систематическому указателю, достаточно, чтобы в описание рукописи был включен о нем %%-фрагмент. В полном виде описание включает в себя не только %%-фрагменты о каждом песнопении, но также все %т-фрагменты, что позволяет воспроизвести рукопись целиком, если все песнопения из неё уже набраны в отдельных файлах.

Различные интерпретации файла с описанием рукописи можно наглядно увидеть в нашей информационной системе «Певческие рукописи Красноярска».⁹ Система, по-разному интерпретируя файл описания рукописи, представляет пользователю разные её стороны. В одном случае ссылки на песнопения перемежаются имеющимися в рукописи надписями, в другом — автоматически вставляемой рубрикацией. Каждое из этих представлений допускает по усмотрению пользователя варьирование: текст показывается либо церковнославянским шрифтом (по выбору из шрифтов в двух разных кодировках), либо в рабочем представлении, либо в упрощенной (современной) письменности; в нормативной или оригинальной орфографии; с разной степенью детализации сведений о песнопениях.

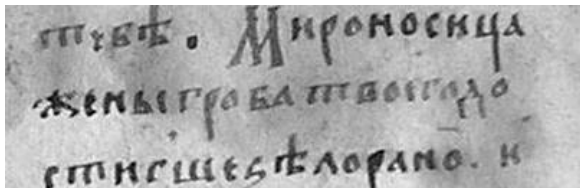
Система шифров песнопений (то есть систематическая классификация богослужебных песнопений) разработана нами для автоматической вставки рубрикации при интерпретации файла с описанием. Наличие этих шифров позволяет организовывать запросы к описаниям из внешних программ.

§5. Обозначения вариантов. Учёт орфографических традиций

Чтобы понять, каким образом достигаются заявленные в §1 цели, рассмотрим процесс фиксирования в корпусе фрагментов. Мы увидим, как, по мере углубления в текст (понимания самого текста или его орфографической традиции) рабочее представление

меняется, буквальная запись уступает место всё более обобщённой, которая даёт больше возможностей для операций с текстом в корпусе (например, при поиске, при внешнем представлении).

Для примера используем словесный фрагмент из рукописи XV века:



Первая стадия фиксирования текста – это буквальная запись текста оригинала. В таком виде текст обычно поступает к ответственному редактору корпуса от наборщиков-энтузиастов. Знаки рукописи записываются один за другим в терминальном представлении:

Мироносица жены гроба твоего достигъше

Это минимально необходимая стадия, отражающая только запись оригинала. Все последующие стадии меняют это представление, но таким образом, чтобы обязательно оставалась техническая возможность получить это буквальное представление. Их задача – дать ключ к получению других вариантов записей, в том числе важного для практиков «современного нормативного».

Поэтому на второй стадии добавляем в эту форму хранения информацию, необходимую для этого варианта. Места различия между нормативной и оригинальной записями записываются в следующей общей форме: [оригинал|норма]. Сравнив нашу запись с современной (Мѣроно̀си́ца жены́ грѣба твоегѣ достигѣше) и соблюдая эту общую форму, получим:

М[и|v̄]рон[o|w']сиц[а|ы] ж[е|_e]ны[|'] грѣ[|']ба твоег[o|w'] дости[|']г[|]ъше

Конечно, эта форма записи выглядит громоздко. Следующие шаги были предприняты первоначально только ради преодоления этого недостатка.

Во-первых, для наиболее частых различий введены сокращения. Вместо [и|v̄] пишется [v], вместо [е|_e] – [e], вместо [о|w] – [o], и т. п. Во-вторых, запись можно сократить, если учесть орфографические особенности рукописи. В частности, поскольку для этой рукописи в целом характерно отсутствие ударений, различие в наличии/отсутствии ударений можно не отмечать. Запись существенно сокращается:

М[v]рон[o]сиц[а]ы ж[e]ны грѣба твоег[o] дости'г[ъ]ше

Это не означает, что ударения, расставленные теперь и в оригинальном тексте, претендуют на то, чтобы правильно отражать ударения XV века – они служат только получению современного представления. С другой стороны, они не затрудняют восстановление написания оригинала, так как убрать их автоматически несложно. Для этого понадобится применить преобразователь *УбратьУдарения*. Чтобы он включался при восстановлении записи оригинала, его нужно упомянуть в специальном фрагменте, в котором отмечаются общие орфографические особенности:

%Орфография <:УбратьУдарения>

Так же, если, например, оказывается, что во всей рукописи не используется *v*, а на её месте всегда используется *и*, то вместо записи альтернативы [*v*] можно прямо использовать терминальную запись *v* (sic), заметив это в таком же поле:

%Орфография [*v*"|и]

Ещё одну деталь придётся добавить в текст для обеспечения возможности получения текста в упрощённой (т. е. современными русскими буквами) записи. Поскольку слово «твоего» надо будет в этой «традиции» писать с прописной буквы, а алгоритмически установить этот факт затруднительно, вносим специальную помету **^**. В итоге получаем:

Mv"p[o]но'сиц[а|ы] ж[е]ны' гро'ба ^твоег[o]` дости'г[ъ]ше.

Эта наша запись буквально не соответствует никакой орфографии: ни принятой в оригинале, ни современной нормативной, ни современной упрощённой. Но она позволяет получить запись этого текста во всех этих трёх нужных нам видах, для чего в соответствующих стилевых файлах будут назначены требуемые преобразователи. Для обеспечения выбора и такого варианта, как «запись оригинала», в необходимую для восстановления его записи цепочку преобразований будут вставлены упомянутые выше преобразования ударений и ижицы.

Если появится еще какая-то традиция, которая не получается из этой записи, то понадобится отметить и другие места разночтений. Этот риск уменьшится, если мы сможем за разночтением увидеть явление, и уже для явления ввести обобщённое, содержательное обозначение. В нашем примере за сокращением [*o*], означающим «в оригинале *o*, а в норме *o*» и выражающим механическую замену одной буквы на другую, стоят два разных грамматических явления: в слове *мироносица* это способ выражения множественного числа, в слове *Твоего* – способ выражения родительного падежа. Если мы вместо них будем использовать содержательные обозначения

(например, *твоег[о]{Gen}*, *мироносица{Pl.Nom}*), то добавление новой традиции не потребует внесения изменений в саму единицу хранения, а потребует только уточнения преобразований наших обозначений в записи конкретной традиции. При этом можно разбивать корпус двумя способами.

Один путь – двигаться чисто эмпирически, то есть брать на себя ответственность только за воплощения реально встретившихся в корпусе фактов. Тогда в преобразователи, обеспечивающие каждую традицию можно будет только добавить по одной строчке: для получения записи оригинала –

| *мироносица{Pl.Nom}* | *мироносица* | ,

для получения нормативного варианта –

| *мироносица{Pl.Nom}* | *mv`ронw'сицы* | .

При такой стратегии понимание традиции будет накапливаться в этих преобразователях по мере пополнения корпуса.

Другой путь возможен, когда для какой-то из традиций удастся создать лингвистические преобразователи (для нашего примера нужна процедура склонения). Тогда имя такого преобразователя можно было бы упомянуть в цепочке преобразований, необходимых для отображения текста в этой традиции.

Можно привести крюковой пример. В рукописях (особенно поздних старообрядческих наречных) встречаются явные отклонения от нормативного использования знамен. Для практиков нужно их исправлять, для исследователей – фиксировать оригинал. Например, в конце строки вместо статьи иногда используется стрела простая. Можно воспользоваться вариантной записью, записав это как [/ |=], этим уже обеспечив возможность получения как нормативного, так и оригинального вида в конкретном случае. А можно, для регулярного случая, ввести новое обозначение, например, [=] или [*постояти*], прописав в поле %*Орфография* замену [=] на /. Этим и все такие случаи смогут быть отмечены, и информация о крюковой орфографии данной рукописи пополнится новой строчкой.

Подведём некоторые итоги. Прежде всего, у нас появилось **два понятия нормы**.

Во-первых, можно говорить о *технической норме* – той форме записи, из которой порождаются записи различных «традиций», поскольку она подчиняется некоторой внутренней инструкции для ответственного редактора корпуса.

Во-вторых, у нас появилось *формальное описание орфографической традиции* (своего рода «локальная норма» или «норма конкретной рукописи») – это перечень преобразований, которые

ший метод: и образец, и тексты, по которым производится поиск, преобразуются в упрощённую форму – в данном случае в современную гражданскую орфографию, после чего все перечисленные формы выглядят одинаково и находятся за один проход процедуры поиска. Преобразованная форма, таким образом, служит обобщающим обозначением для целого множества различных форм. Так, в нашем примере: буква о служила обобщающим обозначением множества форм {*сѡ, ѡ, Ѡ, ѡ̇, ѡ̆, ѡ̇* и др.}:

{ *_w', w, _o=, _o=', o', w', ...* } : **o**

В знаменном тексте, так же как и в словесном, поиск в режиме *Строго по образцу* приносит мало пользы ввиду вариантности крюковой записи. Приведём примеры, когда показанный приём приходит на помощь.

1. Трудность для поиска создаёт нерегулярность использования помет в рукописях. Существуют рукописи с избыточной опомеченностью, в лицах-попевках могут помечаться все знамена, а может только начальное и т. п. Поэтому при строгом поиске по образцу с пометами многие случаи окажутся не найдены. Решает проблему преобразование, удаляющее пометы из поискового образца и из текста. Можно сказать при этом, что беспометная запись послужила обобщением вариантов записи с пометами.

2. Другой вид вариантности связан с тем, что некоторые знамена, наряду с певческим значением, несут структурную функцию (*†* или *††* в конце песнопения, *†* – в начале). Из-за этого, например, при буквальном поиске фрагмента по образцу со статьёй не будут найдены аналогичные фрагменты, где конечный звук записан крыжем. Выход из положения обеспечивает преобразование, заменяющее крыжи статьями. Обозначение статьи при этом работает как обобщенное обозначение группы знамен {*†, ††, †*}:

{ *=, =+, +* } : **=**

3. Так же, певческое знамя, выражающее две или три тонемы¹⁰, имеет мелодически эквивалентную комбинацию однотонемных знамен. Эту эквивалентность выразим в тех же обозначениях на примере стрелы светлой:

{ *_/, Э: К:* } : **Э: К:**

4. Подобная п. 2 взаимозаменяемость знамен, однако, проявляется обычно не во всём тексте, а только в определённых комбинациях знамен. Например, в записи *↗ ↘ ↙ ↚* каданса попевки Руза 7 гласа вместо *↘* в рукописях встречается и *↗*, и *↙*, и *↚* – но это не значит, что мы можем заменять *↗* на *↘* и во всех других случаях.

Обобщающее обозначение в таком случае должно вводиться не для отдельного знамени, но для целой комбинации знамен. Им может служить как одна из равнозначных комбинаций (а), так и произвольная запись, например, название этой комбинации знамен (б):

а) { /: \ Э: =: , /: Ч Э: =: , /: Т. Э: =: , /: /4 Э: =: } : /: Ч Э: =:

б) { /: \ Э: =: , /: Ч Э: =: , /: Т. Э: =: , /: /4 Э: =: } : Руза

Как и в примере со словесным текстом, если мы по такой формуле заменим конкретный вариант записи Рузы на его обобщенное обозначение как в поисковом образце, так и в просматриваемом тексте, то мы найдем все места употребления Рузы, а не только точно совпадающие с образцом.

§7. Анализ знаменного текста

1. Преобразование 4б (§6) есть уже первый шаг мелодического анализа. На месте любого варианта записи мотива появляется имя мотива. По силе обобщения этот шаг вбирает в себя все преобразования 1–3 (§6). Если его сумеет применить ко всем мотивам, причём как в просматриваемых при поиске знаменных текстах корпуса, так и в образце поиска, то мы получим возможность находить в корпусе все варианты записи искомой мелодии, независимо от разнообразия её записи знаменами – так же, как находили все варианты записи слова «Иосиф».

Другим поводом для попыток анализа знаменного текста явилась практическая задача. В нашем издательстве готовилось переиздание синодального Ирмология. Однако, при вычитке выявились следующие недостатки этого издания: 1) одинаковые в крюковых текстах попевки здесь по-разному записаны нотами; 2) во многих местах нарушено нормативное распределение слогов текста в попевке; 3) интонационный облик некоторых попевок размыт.

При этом мы не могли непосредственно сверяться с крюковыми источниками, так как крюковыми ирмологиями с пореформенным текстом не располагаем. В то же время, представления о формульном строении распева позволили предположить, что для сравнения можно использовать структурные элементы самого распева, даже без полного совпадения словесного текста – во всяком случае, если их удастся пронаблюдать в сходной мелодической (соседние мотивы) и просодической (слоговая структура подлежащего словесного текста) ситуации. Но чтобы само это наблюдение стало возможным, необходимо уметь выделять одни и те же структурные элементы в исправляемом и в контрольном (крюковом) тексте.

Таким образом, две практические задачи – поиск и корректура – поставили перед нами задачу *анализа*, то есть расчленения записи мелодии на структурные части.

2. Решение этой задачи не может не опираться на *каталог графических выражений мотивов* (кратко – **свод мотивов**, или просто **мотивник**), объединяющий формулы типа 4б. Однако решение не сводится к механическому объединению таких формул для каждого мотива в единую таблицу замен, поскольку один графический комплекс может встречаться в попевах в разной роли (например, ♯ / / может встретиться и в кадансе Рузы, и в её начале). Выделение мотивов должно, таким образом, опираться на знание структуры попевки, в которую они включены. Это определяет структуру свода мотивов, с помощью которого мы на материале ирмосов 7-го гласа проделали первые опыты мотивного анализа. Она отражает следующую рабочую модель строения распева.

Мелодия песнопения состоит из попевок. Попевка представляет собой цепочку структурных звеньев (*схему попевки*), некоторые из которых могут быть вырождены (опущены). Каждое звено может иметь множество вариантов записи, выбор которых не зависит от других звеньев (кроме, может быть, влияния соседнего звена на стыках звеньев). Попевка имеет характерное интонационно яркое звено, как правило, конечное. (Иногда словом «попевка» называют только это звено – кадансовый мотив попевки. В своде мотивов мы ему даём имя, одинаковое с именем попевки.)

Структурно эта модель следует теории знаменного распева, построенной Б.Карастояновым¹¹, в которой попевка рассматривается как структурный элемент 3-го уровня, а в качестве её звеньев (структурных элементов 2-го уровня) рассматриваются мотивы в конкретно обозначенном смысле¹². Однако, из теории не следует, что полная схема попевки может быть описана простой цепочкой таких мотивов. Может оказаться, что какое-нибудь звено будет допускать множество возможных цепочек мотивов, не сводимое к одной цепочке с пропусками; то есть понадобится не одноуровневая, а, как минимум, двухуровневая модель. Наша модель, таким образом, является упрощённой, но анализ по этой модели необходим как первый шаг, для последующего её уточнения.

Анализ по этой модели по сложности и методам напоминает задачу морфемного анализа: в услышанном слове *рассказчик* (фонетически *раскащик*) надо увидеть выражения приставки *раз-*, корня *-сказ-* и суффикса *-чик-* (приставка употреблена в варианте *рас-* из множества вариантов {*раз-*, *рас-*, *разо-*}; а *-щ-* – результат слияния согласных на стыке корня и суффикса). Но в нашем случае сложность состоит ещё в том, что если система морфем языка изучена, то состав и границы мотивов ещё предстоит выяснить.

Мотивник, построенный по этой модели, представляет собой набор таблиц, по одной для каждой попевки, заполняемых по Технологии смешанного набора (пример 5). Столбцу таблицы соответствует структурное звено попевки, строкам соответствуют источники, из которых выписаны варианты записи звеньев.

ПЕРЕКЛАДКА					Описание попевки:
	сРечь	Горка	сРечь	Перекладка	
Q.I.6					%попевка Перекладка %Структура сРечь Горка сРечь Перекладка %из Q.I.6 %часть сРечь %= Т. %= Т Т. %часть Горка %= Э: К: Т %= Э: =: Т %= _./+ К. %= _./+ Т %часть сРечь %= Т %= Т Т %= Т Т Т %= Т Т Т Т %= Т Т Т Т Т %часть Перекладка %= \ = \ =
Q.III.44					%из Q.III.44 %часть сРечь %= сТ. %= рТ рТ. %часть Горка %= Э: пК: МК: %= Э: =: МК: %= Т: пК: МК: %= Э: пК: мТ. %= пб _./ МК:
Krit					

Пример 5. Страница Мотивника для попевки Перекладка (фрагмент) и её набор в Технологии смешанного набора

3. **Алгоритм анализа** мелодического текста при помощи этих таблиц состоит из нескольких этапов. На первом этапе заменяются своими именами из схемы попевки только конечные (ядерные) звенья. То есть в последовательность знамен (обозначений структурных элементов распева 1-го уровня) вводятся обозначения 2-го структурного уровня, подобно тому, как в естественном языке сложное понятие встаёт в один ряд с простыми словами:

Э: _./ Т Т. | Руза | _./+ К. Т Т | Перекладка |

На следующем этапе таким же образом выделяются графические выражения предпоследнего структурного звена. При этом, чтобы исключить ложное узнавание звена в других местах попевок, в формулу преобразования включается и имя уже выделенного звена:

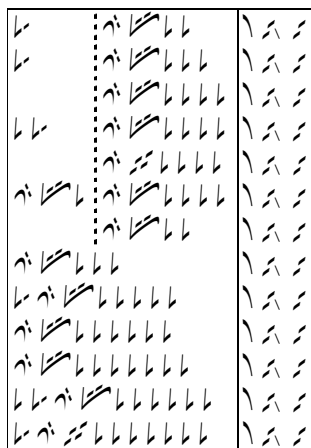
{ *Т. Руза* , *Т Т. Руза* , *Т Т Т. Руза* , ... } : *нРечь Руза*

Теперь уже выделено по два звена в каждой попевке, и можно снова применять этот же приём для третьего с конца звена – и так пока вся схема попевки не будет исчерпана. Если после этого в знаменном тексте остаются участки, изложенные знаменами, а не именами мотивов, то это говорит о несовершенстве мотивника: а) неправильности или неполноте схем попевок; б) неполноте свода мотивов каждой попевки.

4. Составление мотивника – выделение схем попевок и составление каталога выражений звеньев – в свою очередь потребовало инструментальной поддержки. Эта работа состоит в следующем.

Сначала мелодические тексты источников разбиваются на строки. При этом принимаются во внимание: колонны словесного текста, подтверждаемые явными цезурами в крюковой записи (например, стихами); имеющиеся рукописи-строчники; разбиение уже проработанных источников. Для полученных строк по имеющимся теоретическим руководствам устанавливаются имена записанных в них попевок¹⁵. Когда попевка встречается первый раз, в мотивнике для неё заводится новая страница со схемой попевки из одного «корневого» звена. Затем строится гипотеза о том, какая часть этой строки выражает это звено, и эта часть вносится в соответствующую клетку мотивника. После этого к тексту источника применяется вышеописанная процедура анализа, которая теперь обнаруживает в нем и только что записанные варианты. Работа по определению перечня попевок заканчивается, когда процедура анализа в каждой строке выделяет один корневой мотив.

Имея, таким образом, в зачатке будущий мотивник, мы сразу же начинаем применять его, проделывая один и тот же шаг, пока в результатах анализа остаются невыделенные участки: сделав выборку строк по имени попевки, мы пытаемся усмотреть и выделить в этих участках новые звенья (пример б).



Пример б. Выделение очередного звена в процессе выявления структуры попевки. Справа – уже выделенное звено

При этом берутся во внимание как формальные критерии (лаконичность записи разбиения), так и содержательные (теоретические представления о мотивном строении попевки, о функции знамени в мотиве и пр.). Гипотеза о том, что является следующим звеном, закрепляется изменением схемы попевки в мотивнике и внесением в клетку об этом звене всех его выделенных вариантов.

Для этой работы было важно, что попевочный состав каждого ирмоса оставался в рукописях разных веков стабильным. Это позволило при выделении структурных звеньев вести работу сразу по нескольким источникам. Более ясную, четкую и неизменную форму попевки дают беспометные книги, пометный поморский Ирмологий позволил уточнить высотное положение мотивов, а рассмотрение наречного и нотолинейных Ирмологиев показывают картину постепенного отхода от нормативного вида попевки. В итоге свод мотивов седьмого гласа был сделан по пяти ирмологиям от начала XVII до XIX века: раздельноречному беспометному начала XVII в.¹⁴, поморскому пометному XVIII в.¹⁵, наречному XVIII в.¹⁶, рукописным «топорикам» начала XVIII в.¹⁷ и синодальному изданию¹⁸. Пользователь может включать и отключать использование тех или иных источников, по которым ведётся анализ.

§8. Следствия и перспективы

Имена мотивов, возникшие как обобщающие обозначения в задачах поиска и анализа (§6–7), существенно отличаются от тех обозначений, которые мы применяли в технической норме (§5), тем, что по ним не нужно восстанавливать запись оригинала. Но если бы такое восстановление было возможно, то их можно было бы использовать и в технической норме записи песнопения. Другими словами, если бы у нас были преобразователи, которые восстанавливают запись мотива над некоторым текстом (крюковую или нотную) по одному лишь имени этого мотива (и, разумеется, словесному и мелодическому контекстам), то, например, участок песнопения с колоном «и всадники три стояще» мог бы храниться, например, так:

РузыНачало [И вса'-] нРечитатив [дники] Руза [три' стоя'ще]

или даже просто

РУЗА [И вса'дники три' стоя'ще] ,








что представляло бы инвариантную запись этого фрагмента, так как остальные детали обуславливались бы при этом характеристиками требуемой традиции.

К созданию таких преобразователей мы сегодня не готовы. Не говоря уже о нередкой неустойчивости «крюковой орфографии» в рукописях, на пути решения *проблемы синтеза* стоит множество нерешённых проблем. Так, для такого использования первой формы необходимо установить способ выбора записи мотива в зависимости от (а) подлежащего текста и (б) соседних мотивов; для второй формы также (в) критерии выбора того или иного мотивного варианта попевки, (г) способ сегментации колона под мотивы попевки. Вполне вероятно, что однозначности достигнуть всё-таки окажется невозможно, и задачу синтеза попевок придётся рассматривать в более слабой постановке – не в контексте корпусной проблемы точного восстановления записи оригинала, а имея целью только получение одного из возможных вариантов записи – например, для перевода с одной традиции в другую.

Однако, как бы далеко нам ни было до решения задачи синтеза в той или иной постановке, применение описанной процедуры анализа к текстам представительного корпуса знаменных песнопений при наличии полного свода вариантов мотивов по разным источникам способно дать богатый материал для изучения поставленных (а–г) и других проблем.

В частности, становится несложно создавать инструменты, которые по каждому тексту из корпуса позволяли бы получать:

1. для изучения строения попевки и учебно-методических целей:
 - а) для каждого варианта каждого мотива – все варианты связанного с мотивом текста (*пример 7*); б) справочник сочетаний мотивов; в) полный строчник, то есть перечень всех графических выражений всех или отдельных попевок;

-  • ма(ннем) • по(нмо) • на(крестък) • пѣ(ннеми) • спа(са) • свѣтъъ • со(быше) • о(тча) • тѣмѣ • ма(ти) • бы(ше) • ра(дѣисла) • пой(тѣ) • тварѣ • ...
-  • сокрѣ(шешеомѣ) • и бл(дники) • истрѣ(шешеомѣ) • и л(довомѣ) • покры(ла) • оутѣр(днсла) • плотъско(ѣ) • ктѣвѣк • оутѣре(нюуще) • етраха • отеца • ...
-  • перви(ѣ) • ниже ѡт(цѣмѣ) • и о(троки) • и бо(площенна) • кезѣк(менна)
-  • б' нача(тцѣк) • оуѣлы(шахо)
-  • оутѣрди(сла) • утѣрже(и) • небеса • ниже впецѣ • нешпа(ана)
-  • добродѣ(тели) • и нетлѣ(нна)
-  • бо глѣби(нѣ) • небеса

Пример 7. Варианты текста, соответствующего звену «Начало Рузы».

ника могут быть полезны только с опорой на представительный в систематическом и жанровом отношении электронный корпус. Представляется поэтому, что эти две задачи – создание Корпуса знаменных песнопений и создание Свода мотивов знаменного распева являются сегодня важнейшими в деле обеспечения технической поддержки современной музыкальной медиавистики.

¹ Фонд знаменных песнопений. – <http://znamen.ru>.

² Сообщество «Славянская типографика». – <http://cslav.ru>.

³ Проект «Печатный двор». Рабочие документы. – <http://www.pechatnyj-dvor.su/docs.html>. Все обозначения букв одной таблицей – <http://znamen.ru/fmt-slav.htm>.

⁴ Библиотека святоотеческой литературы. – <http://orthlib.ru>.

⁵ ГПНТБ СО РАН, Q.Ш.33.

⁶ А. А. Лукашевич из Москвы.

⁷ Форум любителей древнерусского пения и богослужения. – <http://talk.canto.ru>.

⁸ Новосибирск, ГПНТБ СО РАН, Q.I.22.

⁹ Информационная система «Певческие рукописи Красноярска». – <http://ruk.kraslib.ru>.

¹⁰ Б. Карастоянов. Тонемите в знамения разпев. – Българско музикознание, 1980, кн.4, с.50–66.

¹¹ Б. Карастоянов. Попевки и фиты в центонах знаменного роспева и их основные формулы. – Българско музикознание, 1991, кн.1, с.

¹² Б. Карастоянов. Мотивите в руския знаменен разпев. – Българско музикознание, 1986, кн.1, с.41–50.

¹³ Мы использовали пособие: Е. Григорьев. Пособие по изучению церковного пения и чтения. – Рига, 2001.

¹⁴ Новосибирск, ГПНТБ СО РАН, Q.I.6.

¹⁵ Новосибирск, ГПНТБ СО РАН, F.I.5.

¹⁶ Новосибирск, ГПНТБ СО РАН, Q.Ш.44.

¹⁷ Красноярск, ККГУНБ, № Р-1115706.

¹⁸ Ирмосы. Синодальное издание, 1898 г.

¹⁹ Обзор темы – <http://znamen.ru/bvr.php>.