

Параллельные тексты в задаче дешифровки древнерусских знаменных песнопений*

И. В. Бахмутова, В. Д. Гусев, Л. А. Мирошниченко, Т. Н. Титкова
gusev@math.nsc.ru, luba@math.nsc.ru

Новосибирск, Институт математики им. С. Л. Соболева СО РАН

Предложен новый компьютерно-ориентированный подход к проблеме нотолинейной реконструкции (дешифровки) древнерусских церковных песнопений XVI–XVII вв., представленных в знаменной форме записи. В основе подхода лежит анализ двознаменников — певческих книг, записанных в виде трех параллельных (синхронизованных между собой) текстов: знаменного, нотолинейного и стихотворного (старославянского). Введены понятия инвариантов и квазиинвариантов знаменного распева. Разработан алгоритм выделения их из обучающего материала (двознаменники) и использования для целей дешифровки. Получены оценки эффективности подхода на независимом контрольном материале. Основным достоинством подхода является ориентация его на общий случай беспометной нотации.

Ключевые слова: *параллельные тексты; знаменные песнопения; двознаменники; дешифровка; инварианты; квазиинварианты*

DOI: 10.21469/22233792.1.13.08

Parallel texts in the problem of deciphering of ancient Russian chant*

I. V. Bakmutova, V. D. Gusev, L. A. Miroshnichenko, and T. N. Titkova
Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences,
4 Acad. Koptyug avenue, Novosibirsk, Russia

The ancient Russian chants of XII–XVII centuries are presented in the neume writing form. The problem of chant translation into modern note writing is of deciphering character and, in the general case (chants without special marks that explain their singing value), is not yet solved. The number of “unreadable” ancient hymnals runs into the hundreds. The main difficulties of deciphering are connected with the polysemy of correspondence “neume–note.” The known examples of deciphering are few in number, made by manually, and refer to the separate hymn. The authors develop a new computer-oriented approach to the solution of this problem using the dvoeznamenniks of the end of XVII – beginning of XVIII centuries where the chants are written in three (synchronized between each other) parallel text: in neumes, in notes, and in old Slavonic verses. The emphasis places on revealing in texts of dvoeznamenniks not very long repeating chains of neumes that are interpreted either unambiguously (invariants) or with admissible deviations (quasi-invariants). On the basis of rather extensive learning material, the electronic dictionaries of invariants and quasi-invariants were constructed. The algorithm of deciphering of neumatic notation using these dictionaries was developed. The experiments on the control material have shown that at this stage (without appellation to the structural organization of neumatic hymnals), these dictionaries provide the deciphering of 60%–70% of neumatic text. The main features of the presented approach are: use of dvoeznamenniks of the golden age period of Russian chant in different genres and orientation, in general, toward the neumatic notation without special marks from XVI–XVII centuries.

*Работа выполнена при финансовой поддержке РФФИ, проект № 13-07-00400.

Keywords: *parallel texts; the ancient Russian chants; dvoeznamenniki; deciphering; invariants; quasi-invariants*

DOI: 10.21469/22233792.1.13.08

1 Введение

Древнерусские церковные песнопения XII–XVIII вв. представлены преимущественно в знаменной форме записи. Проблема перевода их в современную нотолинейную форму носит дешифровочный характер и в общем случае (песнопения раннего и среднего периодов) остается нерешенной. Знамена (или крюки) — специальные знаки, служащие для передачи мелодии. Они интерпретируются цепочками нот разной длины (обычно от одного до пяти нотных знаков). Общие сведения о знаменном (или крюковом) пении можно найти в [1]. По оценкам, приведенным в [2], количество известных певческих рукописей измеряется на данный момент более чем тысячью экземпляров, из них читаемыми (с существенными оговорками [3, 4]) считаются лишь около половины.

Особый интерес в плане дешифровки представляют так называемые *двознаменники* (см. [1, гл. 12]). Это певческие книги конца XVII–начала XVIII вв., содержащие песнопения, записанные в виде трех синхронизованных *параллельных текстов*: знаменного, нотолинейного и стихотворного (старославянский язык). Количество известных двознаменников невелико (порядка десятка [1]). Именно они лежат в основе предлагаемого авторами подхода к дешифровке. Часть из них используется для обучения, другая — для контроля. Этому предшествовала достаточно объемная и кропотливая работа по переводу двознаменников в цифровую форму, поскольку программ распознавания знаменной нотации на данный момент, насколько известно авторам, не существует.

Немногочисленные известные примеры дешифровки знаменной нотации сделаны вручную и касаются отдельных песнопений (или узких классов песнопений), эволюцию которых можно проследить по архивным материалам в течение достаточно длительного периода. Существенную роль при этом играет наличие графически близких византийских версий. Подготовленные авторами электронные версии двознаменников охватывают песнопения разных жанров и позволяют использовать для дешифровки новые подходы, превышающие по своей трудоемкости возможности исследователей, работающих вручную. Эти подходы основаны на введенных авторами понятиях *инвариантов и квазиинвариантов* знаменного распева.

Целью работы является описание алгоритма дешифровки, использующего параллельные тексты двознаменников, и оценка его эффективности. Проводится анализ ошибок и рассматриваются возможности дальнейшего продвижения.



2 Краткие сведения о знаменном распеве

В истоках знаменного распева лежит древнегреческое церковное пение на 8 ладов (гласов): дорийский, фригийский и т. д. (так называемая *система осмогласия* [1]). Начало системе положил обычай в каждый из восьми дней Пасхи исполнять песнопения на особый напев. Восьмидневный цикл напевов, которые хор исполнял в унисон, был распространен затем на 8 недель, а одноголосный (монодия) напев конкретного дня повторялся в течение соответствующей ему по порядку недели. Восемь недель составляли столп, который циклически повторялся в течение года.

В русском осмогласии понятие гласа как лада деформировалось, а средством мелодической характеристики гласа стала выступать *совокупность попевок*, т. е. мелодиче-

ских оборотов (в нашей терминологии — *элементарных структурных единиц* знаменного распева). Наиболее характерные для каждого гласа попевки представлены в подборке В. М. Металлова [5]. Но они не определяют однозначным образом гласовую принадлежность песнопения, поскольку некоторые попевки встречаются в песнопениях разных гласов.

Восемь гласов охватывают практически весь попевочный фонд церковного пения. Каждое песнопение, как правило, принадлежит одному из 8 гласов и строится на попевках данного гласа. Гласовая принадлежность песнопения обычно указывается в явном виде. Сказанное выше означает, что любой алгоритм дешифровки должен учитывать подчиненность песнопения системе осмогласия, поскольку одно и то же знамя может иметь разную интерпретацию в зависимости от гласа, типа структурной единицы, в состав которой оно входит, ее позиции в тексте и ряда других факторов.

Характер многозначности соответствия «знамя–нота» отражен в построенной авторами на основе двознаменников электронной азбуке знаменного распева [6]. Некоторые знамена, например такие, как  (статья закрытая малая) или  (статья простая с подверткой), имеют до 10 различных интерпретаций, отличающихся друг от друга интервально-ритмическими характеристиками. При этом каждая из интерпретаций может иметь несколько звуковысотных привязок. Именно с *многозначностью соответствия «знамя–нота»* связаны основные проблемы дешифровки.

Начиная с XVII в., знамена начинают снабжать *степенными и указательными пометами*. Первые уточняют высоту распева знамени, вторые — особенности его распева. Термин «читаемые» относится к пометным песнопениям, хотя и здесь однозначный результат не гарантирован (см. [3,4]). Беспометные рукописи XVI в. и более раннего периода практически нечитаемы.

Для устранения неоднозначности можно привлекать контекст, что эквивалентно переходу от отдельных знамен к более крупным структурным единицам, таким как *попевки, лица, фиты* и др. Однако известные подборки этих структурных единиц (азбуки, кокизники, фитники) малопригодны для целей дешифровки из-за формы их представления. Так, в подборке В. М. Металлова [5] попевки приведены только в нотолинейной форме, а в подборке М. В. Бражникова [7] около половины лиц и фит имеют лишь знаменное представление вместо двознаменного, требуемого для дешифровки.

Песнопения, представленные в двознаменниках, записаны в диапазоне *обиходного звукоряда* (рис. 1), включающего ноты *G, A, H* (малой октавы), *c, d, e, f, g, a, b* (первой октавы), *C, D* (второй октавы). В пометных рукописях ступеням звукоряда соответствуют обозначения **Г, Н, Ц, Г, Н, •, М, П, Ц, М, П, Ц**. Соответствующая знамени степенная помета указывает на наивысший звук в его распеве [1].



Рис. 1 Обиходный звукоряд

Особенности нотолинейной интерпретации знамен поясняются с помощью указательных помет: — (или т) — тихая; ° — ломка; б — борзая; у — ударка; « — качка (или купно); з — зевок; р — равно. Знамена с указательной пометой и без нее трактуются как разные, поскольку могут иметь отличающиеся распевы. Длительности звуков обозначаются следующим образом: ○ — 1 (целая); ♩ — 2 (половинная); ♪ — 4 (четвертная); ♫ — 8 (восьмая). Для обозначения высоты и длительности звука используем комбинацию буквы и цифры (например, H4 — это четвертная нота «си» малой октавы). Знак $(*)$ используется в качестве разделителя между нотолинейными интерпретациями разных знамен, а (\sim) — как символ эквивалентности, отделяющий знаменную цепочку от ее нотолинейного представления (например, запись $\text{H}^{\text{°}}\text{ш}^{\text{°}}\text{а}^{\text{°}} \sim d4c4d2 * c4e4$ означает, что цепочка из двух знамен, стоящая слева от (\sim) , интерпретируется в тексте двознаменника, соответственно, цепочками из трех и двух нот, отделенными друг от друга знаком $(*)$).

3 Описание подхода

На данный момент мы располагаем электронными версиями трех двознаменников конца XVII – начала XVIII в. Это разные редакции многожанровой певческой книги «Октоих» из Соловецкого собрания Российской национальной библиотеки, г. С.-Петербург (шифры 618/644, 619/647, QI188). Число песнопений в разных гласах двух первых Октоихов варьирует в диапазоне от 25 до 29, для QI188 этот показатель на треть ниже. Длины песнопений составляют от нескольких десятков до двух–трех сотен знамен. На основе этого (обучающего) материала была построена электронная азбука знаменного распева [6], выгодно отличающаяся от известных авторских азбук по многим показателям, в частности наличием информации о частоте встречаемости и звуковысотной привязке различных интерпретаций каждого знамени в каждом гласе.




Анализ электронной азбуки показал, что при всей вариативности знаменного распева в каждом гласе существует некоторое количество знамен, не меняющих своей интерпретации. Обычно их доля не превышает 10% от размера алфавита. В связи с этим возникло предположение о наличии в общем случае повторяющихся¹ цепочек знамен произвольной длины L ($L \geq 1$), однозначно или почти однозначно (с допустимыми отклонениями) интерпретируемых в пределах одного гласа. Цепочки первого типа естественно называть *внутригласовыми инвариантами* (ВИ), а второго — *квазиинвариантами* (КВИ). Их можно трактовать как особый тип структурных единиц, своего рода «островки стабильности», характеризующиеся существенным снижением уровня неоднозначности в интерпретации знамен.

В работах [8, 9] гипотеза о наличии в текстах двознаменников цепочек знамен со свойствами ВИ и КВИ подтверждена экспериментально на указанном выше материале. Для каждого из 8 гласов построены словари ВИ и КВИ, разбитые на подсловари в соответствии со значениями $L = 1, 2, \dots$. Алгоритмы выделения ВИ и КВИ основаны на следующих ображениях:

- 1) элементами словарей ВИ и КВИ могут быть, как уже говорилось выше, лишь цепочки знамен, повторяющиеся в конкретном гласе, так как понятие инвариантности подразумевает сохранение нотолинейной интерпретации (в строгом смысле или с допустимыми вариациями) при повторении цепочки. Требование повторяемости в неявной

¹Требование повторяемости является принципиальным, поскольку оно исключает из рассмотрения однократно встречающиеся цепочки знамен. Такими, начиная с некоторого L , являются практически все достаточно длинные цепочки

форме ограничивает длины ВИ и КВИ, так как с увеличением длин цепочек частота F их встречаемости в гласе падает. Число просматриваемых цепочек можно ограничить, введя пороговое значение F_r . При $F < F_r$ цепочка выбывает из рассмотрения. Будем использовать пороговое значение $F_r = 3$;

- 2) цепочка, составленная из многозначных знамен, может оказаться однозначно или почти однозначно интерпретируемой в гласе. Поэтому при формировании словарей ВИ и КВИ должен быть рассмотрен полный спектр цепочек произвольной длины с $F \geq F_r$ в подборке песнопений, представляющих конкретный глас;
- 3) формировать словари ВИ и КВИ следует на основе *беспометных* текстов. Тогда словари будут иметь формат «беспометное знамя – нота», единственно пригодный для решения задачи дешифровки в общем случае (беспометная нотация). Указанные двознаменники содержат пометы, которые следует предварительно устранить. Это равносильно агрегированию исходного знаменного алфавита, при котором в одну группу объединяются и обозначаются одним символом все пометные и беспометные варианты конкретного знамени. Так, в одну группу будут объединены  (стрела мрачная крыжевая с оттяжкой), она же с указательной пометой «борзо» () и «тихая» ()⁶, имеющие разные распевы. Предполагается, что в словарях ВИ и КВИ «агрегированная стрела» будет встречаться в разных контекстах, определяющих тот или иной вариант ее распева;
- 4) возможны различные варианты определения КВИ. В используемом авторами приближении цепочка знамен трактуется как КВИ, если одна из ее интерпретаций доминирует над остальными по частоте, а именно: выполняется соотношение $F_{\text{dom}}/F > 1/2$, где F — частота встречаемости цепочки в песнопениях гласа; F_{dom} — максимальная из частот встречаемости ее интерпретаций. Доминирующая интерпретация трактуется в дальнейшем как значение КВИ. Важно отметить, что понятия ВИ и КВИ имеют относительный характер: они зависят от объема и состава исходной (обучающей) подборки. При увеличении ее объема или изменении жанрового состава отдельные ВИ могут перейти в категорию КВИ, а некоторые КВИ перестают быть таковыми («исчезает» доминирующая интерпретация). Эффекты такого рода можно проследить при переходе от отдельных Октоихов (так строились словари, описанные в [8, 9]) к их совокупности (в данной работе словари формируются на основе всех трех Октоихов).

Построенные по двознаменным Октоихам словари ВИ и КВИ, представленные в формате «беспометное знамя – нота», служат основой для нотолинейной реконструкции контрольных *беспометных* песнопений. В качестве контрольного материала авторы использовали пометный двознаменник «Праздники» (РНБ, г. С.-Петербург, Кирилло-Белозерское собрание, шифр 797/1054) и фрагмент двознаменного Ирмология из собрания В. Ф. Одоевского (М., РГБ, Ф210. №18). Пометы в обоих случаях предварительно удалялись, т. е. на вход алгоритма дешифровки подавались беспометные знаменные песнопения, а нотолинейная компонента указанных двознаменников использовалась лишь для сравнения реконструированного распева с реальным.

Схема дешифровки фактически сводится к покрытию песнопения всевозможными цепочками знамен из словарей ВИ и КВИ и приписыванию найденным фрагментам их певческого значения, зафиксированного в словарях. При этом возможны отказы (не существует ни одного ВИ или КВИ, покрывающего данное знамя) и конфликты интересов (знамя входит в состав разных ВИ или КВИ, где ему приписываются неидентичные певческие значения). Более детально процесс дешифровки описан ниже.

4 Алгоритм дешифровки. Результаты апробации

Пусть $T = t_1 t_2 \cdots t_n \cdots t_N$ — представление песнопения одного из 8 гласов в виде последовательности знамен. Гласовая принадлежность песнопения обычно известна, хотя возможны исключения, и тогда это становится предметом отдельного рассмотрения. Покрытие песнопения цепочками знамен из словарей ВИ и КВИ, построенных для данного гласа, осуществляется итеративно по L , т. е. сначала используются ВИ и КВИ длины 1, затем длины 2 и т. д. При $L = 1$ произвольное знамя t_n песнопения, расположенное в n -й позиции, может быть покрыто единожды, если найдется идентичный ему ВИ или КВИ длины 1. В таком случае открываем список возможных интерпретаций знамени t_n из подсловаря инвариантов длины 1. При $L = 2$ знамя t_n может быть покрыто дважды, если среди ВИ (КВИ) длины 2 найдутся цепочки, совпадающие с биграммami $t_{n-1}t_n$ и $t_n t_{n+1}$ песнопения. В этом случае пополняем список возможных интерпретаций знамени t_n его значениями в составе указанных выше биграммных инвариантов. Знамя t_n при $L = 2$ может быть покрыто и единожды (вносим в список лишь одну интерпретацию) или не покрыто вовсе.

При $L \geq 3$ на каждой итерации имеем не более чем L возможностей покрытия знамени t_n цепочками ВИ и КВИ длины L . Продолжаем пополнять список возможных интерпретаций знамени t_n вплоть до исчерпания словарей ВИ и КВИ для данного гласа. Из множества полученных певческих значений для t_n выбираем самое частое. При равенстве голосов фиксируем случайным образом любой вариант.

По итогам описанной процедуры каждое знамя в песнопении либо получает единственную интерпретацию, хотя и не всегда правильную, либо не покрывается ни одной из цепочек словарей ВИ и КВИ. Сравнение результатов реконструкции с реальным нотопевиным текстом позволяет оценить количество правильно интерпретированных знамен (n_+) и число знамен, трактовка которых отличается от истинной (n_-). Коэффициент покрытия песнопения цепочками из словарей ВИ и КВИ определяем тогда в виде $k = n_+/N$, где N — число знамен в песнопении ($0 \leq k \leq 1$). Коэффициент покрытия гласа, представленного группой контрольных песнопений, можно определить либо как среднее коэффициентов покрытия отдельных песнопений, либо в виде отношения $\bar{k} = \bar{n}_+/\bar{N}$, где \bar{n}_+ — суммарное число знамен, правильно интерпретированных во всех песнопениях гласа, а \bar{N} — суммарная длина песнопений. Показатель \bar{k} характеризует эффективность использования словарей ВИ и КВИ для дешифровки беспометных песнопений.

Для апробации подхода было проведено несколько экспериментов. В первом из них разделение данных, используемых для обучения и контроля, осуществлялось с помощью процедуры «скользящего контроля». Суть ее (применительно к конкретному гласу) в следующем:

- песнопения i -го гласа ($1 \leq i \leq 8$) из трех Октоихов объединяются в одну группу $T_i = T_{i1} T_{i2}, \dots$);
- удаляется песнопение T_{i1} , а по оставшимся строятся словари ВИ и КВИ;
- с помощью построенных словарей вычисляется коэффициент покрытия k_{i1} удаленного песнопения T_{i1} ;
- T_{i1} возвращается в исходную подборку, удаляется следующее песнопение T_{i2} , и процесс повторяется до исчерпания всех n_i песнопений i -й группы;
- вычисляется усредненный показатель покрываемости песнопений i -го гласа: $\bar{k}_i = (\sum_{l=1}^{n_i} k_{il})/n_i$.

Таблица 1 Результаты первого эксперимента

Гласы	\bar{k}
1	0,786
2	0,701
3	0,744
4	0,730
5	0,735
6	0,685
7	0,801
8	0,725

Таблица 2 Результаты второго эксперимента

Гласы	\bar{k}
1	0,690
2	0,518
3	0,726
4	0,652
5	0,603
6	0,546
7	0,602
8	0,543

Результаты эксперимента представлены в табл. 1. Из них следует, что в среднем по гласам правильно реконструируются порядка 70%–80% знамен. Однако эти оценки несколько завышены, поскольку поочередно удаляемые песнопения сохраняли какие-то связи с оставшимися, что при реконструкции повышало их шансы.

Более реальные данные получены в ходе второго эксперимента, когда для контроля использовался двознаменник «Праздники» из Кирилло-Белозерского собрания (другая школа). Его в основном составляют стихиры, но они значительно отличаются от стихир из Октоихов сложностью и развитостью распева, вариативностью, шириной динамического диапазона. Количество песнопений в гласах было сопоставимо с аналогичным показателем для Октоихов 619/647 и 618/644 за исключением гласов 3 и 7, слабо представленных в двознаменнике «Праздники». Использовался второй вариант коэффициента покрытия ($\bar{k} = \bar{n}_+ / \bar{N}$). Результаты этого эксперимента приведены в табл. 2.

Показатели покрываемости заметно снижаются, что объясняется существенными различиями в обучающем и контрольном материале.

В третьем эксперименте контрольный материал был представлен 22 песнопениями первого гласа Ирмология. Коэффициент покрываемости этих песнопений инвариантами первого гласа, полученными на материале трех Октоихов, составил 0,687.

5 Обсуждение результатов

1. Словари ВИ и КВИ, построенные на основе трех Октоихов, достаточно представительны (в сумме порядка полутора тысяч цепочек знамен разной длины в каждом гласе). Они в состоянии обеспечить дешифруемость контрольного материала, близкую в среднем к 60%–70% даже при отсутствии помет у знамен. С определенной осторожностью подход может быть применен и к беспометным песнопениям XVI в. У авторов имеется минимальный опыт такого рода, однако достоверность реконструкции требует дополнительного исследования из-за отсутствия двознаменников, датированных указанным периодом.
2. Анализ покрываемости отдельных песнопений выявил редко встречающиеся аномалии — песнопения с коэффициентами покрытия порядка 0,2 и ниже. Предварительный анализ показал, что причинами могут быть: *неверно указанная гласовая принадлежность* песнопения (оно плохо дешифруется по словарю «собственного» гласа, но гораздо лучше по словарю другого); *неоговариваемые звуковысотные переносы* отдельных фрагментов мелодии; *наличие в песнопении большого количества лиц и фит* — структурных единиц, относящихся к категории мелизматических украшений мелодии, когда

	+	+	+	?	?	?	+
поз. 73:				()
	<i>d2</i>	<i>e2</i>	<i>d1</i>	<i>d4H4</i>	<i>c4d4</i>	<i>e2</i>	<i>e2</i>
	<i>d2(7)</i>	<i>e2(6)</i>	<i>d1(4)</i>	<i>c4H4(2)</i>	<i>e4f4(1)</i>	<i>g2(1)</i>	<i>e2(1)</i>
					<i>c4d4(1)</i>		
поз. 80:	()			
	<i>d4H4</i>	<i>c4d4</i>	<i>e2</i>	<i>c2</i>	<i>H2</i>	<i>A1</i>	
		<i>e4f4(1)</i>	<i>g2(1)</i>	<i>e2(2)</i>	<i>e2(2)</i>	<i>d1(2)</i>	
			<i>e2(1)</i>			<i>e1(2)</i>	
поз. 86:							
	<i>e2</i>	<i>d2</i>	<i>e4f4</i>	<i>g2</i>	<i>f2</i>	<i>e1</i>	
	<i>e2(4)</i>	<i>d2(4)</i>	<i>e4f4(3)</i>	<i>g2(5)</i>	<i>f2(4)</i>	<i>e1(2)</i>	

Рис. 2 Пример дешифровки фрагмента песнопения

на один слог текста приходится много звуков мелодии. Разводы лиц и фит «рядовым» знаменем не всегда стандартны. Их можно сравнить с идиомами в естественном языке. Индикатором лиц и фит в распеве является стихотворная компонента параллельных текстов.

3. Ошибочно трактуемые знамена часто располагаются между попевочными структурами. Среди таких «связующих» знамен много высокочастотных: , , , и пр. Порой они образуют тандемные повторы. Это удобный способ разнесения основных структурных единиц текста (в данном случае попевок) на нужное расстояние, используемый в разных языковых системах. Сложность дешифровки тандемов в том, что тандемному повтору на знаменном уровне не всегда соответствует повтор на нотолинейном уровне (и наоборот). Приведем реальный пример (рис. 2) дешифровки фрагмента песнопения первого гласа из двознаменника «Праздники» (лист 78 оборотный). Фрагмент начинается с 73-й позиции.

Здесь каждая строка двознаменного текста — это последовательность знамен, а ниже — их певческие значения. Каждое знамя помечено сверху одним из трех символов: «-» означает, что знамя не покрыто ни одним ВИ или КВИ; «+» означает, что знамя интерпретировано правильно; в третьей строке его певческое значение повторено с указанием (в скобках) числа проголосовавших за него ВИ и КВИ; «?» говорит о том, что знамя интерпретировано неверно, в этом случае в третьей строке приводится ошибочная трактовка, одна или несколько (строки 4, 5 и т. д.), если за них подано одинаковое число голосов. В строках 4, 5 и т. д. может быть представлено и истинное значение (см., например, знамена 77, 82), но механизм случайного выбора, упрощенный до предела, сработал не в их пользу (выбирался первый по списку кандидат среди равных по числу голосов).

Пример не слишком богат попевочными структурами: их три, и все они, по классификации А. Н. Кручининой [10], являются разновидностями «кокиз», кадансовая структура которых представлена позициями 73–75, 83–85, 89–91. Наибольший же интерес представляет фрагмент между первым и вторым вхождением «кокиз» (поз. 76–83), содержащий тандемный повтор из 4 знамен (выделен скобками). Нетрудно видеть,

что на нотопрофильном уровне повтор искажается при переходе от «крюка светлого» (𐌆) к стопице (𐌇). Практически весь указанный фрагмент характеризуется пониженной дешифруемостью, хотя отличия от реального текста (вторая строка) касаются, в основном, лишь звуковысотной привязки (интервально-ритмическая структура не нарушается). Более того, интерпретация «стопицы с очком» (𐌇̇, поз. 76) в тексте двознаменника (скачок вниз на две ступени) представляется менее предпочтительной, чем стандартный ход на одну ступень в реконструированном тексте.

Приведенный пример показывает, что тандемные повторы претендуют на роль самостоятельных структурных единиц знаменного распева и нуждаются в отдельном изучении. Первый шаг в этом направлении сделан в [11]. Следует отметить также, что проявления тандемной повторности в знаменном распеве зависят от жанровой специфики. Так, например, биграмма 𐌆̇ 𐌇̇ часто используется в тандемном варианте (𐌆̇ 𐌇̇ 𐌆̇ 𐌇̇ ~ $e2*d1*e2*d1$) в песнопениях гласа 4 («Праздники»), а в том же гласе всех Октоихов встречается всего 4 раза, не образует тандемных вхождений и имеет доминирующую интерпретацию $c2*N1$, отличную от $e2*d1$. Как следствие, данная цепочка в песнопениях гласа 4 («Праздники») интерпретируется неверно. Из этого следует, что кроме настройки на конкретный глас желательна настройка и на конкретный жанр.

4. Созданные для дешифровки электронные словари ВИ и КВИ можно рассматривать как *систему описания гласов*, каждый из которых характеризуется достаточно представительным набором песнопений. Эти словари могут быть использованы не только для дешифровки песнопений, но и для решения вспомогательных задач, таких как определение гласовой принадлежности песнопения, обнаружение ошибок кодирования, выявление функционирующих в гласе структурных единиц и устойчивых их комбинаций, а также для *количественной и качественной характеристики системы осмогласия* в целом. В последнем случае речь идет о выявлении сходства и различий между гласами.

Применительно к задаче дешифровки нас больше интересует сходство гласов, а именно: наличие достаточно длинных общих для разных гласов цепочек знамен, одинаково интерпретируемых в разных гласах — своего рода *межгласовых инвариантов* (МИ). Степень дешифруемости таких цепочек в разных гласах с помощью словарей ВИ и КВИ может сильно отличаться, поскольку она зависит от частоты встречаемости цепочки в конкретном гласе и ее вариативности. В одном гласе она может попасть в словари ВИ и КВИ и иметь хорошие показатели по дешифруемости, а в другом — наоборот. В таком случае возникает возможность использования взаимосвязей между гласами на уровне МИ для улучшения показателей дешифруемости одного гласа за счет другого. Фактически речь идет о дешифровке *по прецедентам* (порой однократно встречающимся). Удобным инструментом для выявления межгласовых связей на уровне МИ является предложенный авторами аппарат сложностных разложений [12], в основе которого применительно к рассматриваемому случаю лежит представление песнопений одного гласа в виде конкатенации цепочек знамен (или нот) из песнопений другого гласа.

Цепочек, которые можно отнести к категории МИ, достаточно много. Приведем лишь один пример, иллюстрирующий возможность дешифровки по прецедентам. При сравнении гласов 1 и 4 Октоиха 619/647 выявляется общая цепочка знамен из разных песнопений (глас 1: «Слава и ныне. Богородичен», поз. 45 и глас 4: «На стиховне стихира», поз. 63):

$$\begin{array}{cccccc} \text{𐌆̇} & \text{𐌇̇} & \text{𐌆̇} & \text{𐌇̇} & \text{𐌆̇} & \text{𐌇̇} \\ f4e4d2 & f4g4 & f8e8d4 & f4g2f4 & e1 & d1. \end{array}$$

В соответствии с [10] — это одна из разновидностей «кулизм». С помощью словарей ВИ и КВИ, построенных для данного Октоиха, цепочка из гласа 1 дешифруется правильно, а из гласа 4 — с ошибкой в интерпретации второго и третьего знамени. Эта ошибка может быть скорректирована на основе информации о наличии указанного выше МИ в обучающей подборке.

6 Заключение

Проблема нотолинейной реконструкции (дешифровки) древнерусских церковных песнопений, представленных в знаменной форме записи, является одной из наиболее известных и актуальных в музыкальной медиевистике. Трудности ее решения связаны с многозначностью соответствия «знамя–нота». Авторы развивают новый компьютерно-ориентированный подход к ее решению, основанный на использовании двознаменников конца XVII — начала XVIII в., записанных в виде трех *параллельных* (синхронизованных между собой) *текстов*: знаменного, нотолинейного и стихотворного (старославянского). Двознаменники нужны для того, чтобы извлечь из них информацию о наиболее устойчивых (наименее вариативных) структурных единицах знаменного распева, названных авторами инвариантами и квазиинвариантами. Последние представлены в единственно пригодном для дешифровки формате «беспомятое знамя–нота».

На достаточно объемном обучающем материале построены электронные словари инвариантов и квазиинвариантов для каждого из гласов знаменного распева. Разработан алгоритм дешифровки знаменной нотации с использованием указанных словарей. Эксперименты на независимом контрольном материале показали, что даже без апелляции к структурной организации знаменного распева (резерв для дальнейшего развития подхода) удается обеспечить в среднем 60%–70%-ную дешифруемость по гласам.

Разбор ошибок говорит о том, что часть из них — это ошибки кодирования (авторские) и погрешности в самой певческой книге. Другая часть ошибок носит характер допустимого варьирования. Случаев радикального искажения нотолинейной структуры не отмечено.

Развитие подхода мыслится в направлении привлечения информации о внутренней структуре песнопений (попевки, лица, фиты, тандемные повторы). Важной может оказаться информация о взаимосвязях между гласами (приведены некоторые примеры и соображения на указанную тему). И, наконец, ощутимую пользу может принести настройка на отдельные классы песнопений (стихиры, ирмосы и др.).

К основным достоинствам подхода можно отнести опору на *двознаменники* периода наивысшего расцвета знаменного пения, ориентацию на *беспомятую нотацию* и экспериментально подтвержденную применимость к песнопениям разного жанра.

Литература

- [1] *Бражников М. В.* Древнерусская теория музыки. — Л.: Музыка, 1972. 423 с.
- [2] *Кутузов Б. П.* Русское знаменное пение. — 2-е изд. — М., 2008. 304 с.
- [3] *Бахмутова И. В., Гусев В. Д., Титкова Т. Н.* О функциях указательных помет (на материале двознаменника XVIII века) // Сибирский музыкальный альманах. — Новосибирск: Изд-во НГК, 2002. С. 81–92.
- [4] *Бахмутова И. В., Гусев В. Д., Титкова Т. Н.* Факторы, влияющие на точность нотолинейной реконструкции пометных знаменных песнопений // Сибирский музыкальный альманах. — Новосибирск: Изд-во НГК, 2004. С. 51–59.
- [5] *Металлов В. М.* Осмогласие знаменного распева (сборник нотолинейных попевок). — М., 1899, 50 с.

- [6] *Бахмутова И. В., Гусев В. Д., Титкова Т. Н.* Электронная азбука знаменного распева: Предварительная версия // Вычислительные системы, 2005. Вып. 174. С. 29–53.
- [7] *Бразнников М. В.* Лица и фиты знаменного распева. — Л.: Музыка, 1984. 302 с.
- [8] *Бахмутова И. В., Гусев В. Д., Титкова Т. Н.* Компьютерный поиск инвариантных структурных единиц знаменного распева // Проблемы музыкальной науки. Российский научный специализированный журнал, 2011. №1(8). С. 20–24.
- [9] *Бахмутова И. В., Гусев В. Д., Титкова Т. Н.* Выявление инвариантов и квазиинвариантов знаменного распева с помощью билингв типа «знамя-нота» // Мат-лы Всеросс. конф. ЗОНТ–2013, 2013. Т. 1. С. 27–35.
- [10] *Кручинина А. Н.* Попевка в русской музыкальной теории XVII века. Автореф. дисс. . . . канд. иск. н. — Л., 1979.
- [11] *Бахмутова И. В., Гусев В. Д., Мирошниченко Л. А., Титкова Т. Н.* Тандемные повторы в знаменных песнопениях // Вычислительные системы, 2005. Вып. 174. С. 13–28.
- [12] *Gusev V. D., Miroshnichenko L. A.* Complexity expansions in comparisons of character sequences // Pattern Recogn. Image Anal., 2014. Vol. 24. Iss. 4. P. 467–472. <http://link.springer.com/article/10.1134/S1054661814040063>.

Поступила в редакцию 15.06.2015

References

- [1] Brazhnikov, M. V. 1972. *Drevnerusskaya teoriya muzyki* [Ancient Russian theory of music]. Leningrad: Music. 423 p.
- [2] Kutuzov, B. P. 2008. *Russkoe znamennoe penie* [Russian znamennyi chant]. Moscow. 304 p.
- [3] Bakhmutova, I. V., V. D. Gusev, and T. N. Titkova. 2002. On functions of the indicative marks (in the material of dvoeznamenniks XVIII sc.). *Siberian Musical Almanac*. Novosibirsk: NGK. 81–92.
- [4] Bakhmutova, I. V., V. D. Gusev, and T. N. Titkova. 2004. Factors acting on accuracy of the noted reconstruction of Russian znamennyi marked hymnals. *Siberian Musical Almanac*. Novosibirsk: NGK. 51–59.
- [5] Metallov, V. M. 1899. *Osmoglasie of znamennyi chant* [Set of noted popevoks]. Moscow. 50 p.
- [6] Bakhmutova, I. V., V. D. Gusev, and T. N. Titkova. 2005. Electronic alphabeth of Russian chant: Previous version. *Computing Syst.* 174:29–53.
- [7] Brazhnikov, M. V. 1984. *Litsa and fitas of Russian znamennyi chant*. Leningrad: Music. 302 p.
- [8] Bakhmutova, I. V., V. D. Gusev, and T. N. Titkova. 2011. Computerized search for the invariant structural elements of znamennyi chant. *Music Scholarship* 1(8):20–24.
- [9] Bakhmutova, I. V., V. D. Gusev, and T. N. Titkova. 2013. Revelation of invariants and quasi-invariants of znamennyi chant using “neume-note” bilinguas. *Russian Conference KONT-2013 Proceedings*. Novosibirsk. 1:27–35.
- [10] Kruchinina, A. N. 1979. Popevka in Russian musical theory of XVII sc. Ph.D. Thesis. Leningrad.
- [11] Bakhmutova, I. V., V. D. Gusev, L. A. Miroshnichenko, and T. N. Titkova. 2005. Tandem repeats in the neume hymns. *Computing Systems* 174:13–28.
- [12] Gusev, V. D., and L. A. Miroshnichenko. 2014. Complexity expansions in comparisons of character sequences. *Pattern Recogn. Image Anal.* 24(4):467–472. Available at: <http://link.springer.com/article/10.1134/S1054661814040063> (accessed December 28, 2015).

Received June 15, 2015