

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ЭКСПЕРТНЫЕ СИСТЕМЫ (Вычислительные системы)

1997 год

Выпуск 160

УДК 519.237.8; 519.764

СИНОНИМИЧНЫЕ ЗАМЕНЫ В ЗНАМЕННЫХ ПЕСНОПЕНИЯХ¹

И.В.Бахмутова, В.Д.Гусев, Т.Н.Титкова

В в е д е н и е

Данная работа продолжает цикл исследований, связанных с трудноформализуемой проблемой перевода древнерусских знаменных песнопений в современную нотолинейную форму. Объектом исследования являются *двознаменники* — сборники песнопений, записанных параллельно в знаменной и нотолинейной форме. Проводимый нами компьютерный анализ двознаменников преследует своей целью изучение контекстных особенностей знаменного распева и получение дешифровочных таблиц как для отдельных знамен, так и для более крупных структурных единиц (типа попевок). Дошедшие до нас весьма немногочисленные руководства по дешифровке (азбуки и кокизники), составленные древнерусскими распевщиками, как правило, неполны, порою противоречивы и не содержат количественных оценок проявления многозначности соответствия "знамя-нота", что не позволяет выявить звуковысотные инварианты в каждом гласе.

Общее введение в проблему содержится в [1]. Из работ авторов, имеющих непосредственное отношение к теме данной статьи,

¹Работа выполнена при поддержке РГНФ (проект № 96-04-08258) и РФФИ (проект № 96-06-80576)

отметим [2÷5]. В частности, в [2,3] описаны частотно-позиционные алгоритмы выделения структурных единиц непосредственно из текстов знаменных песнопений, не содержащих параллельного нотного текста. Отмечены трудности, возникающие при идентификации низкочастотных структурных единиц. Для их преодоления зачастую приходится использовать рассуждения "по аналогии".

В [4,5] намечена общая схема количественного исследования многозначности прямого ("знамя-нота") и обратного ("нота-знамя") соответствий. Выделены группы близких в функциональном отношении знамен, которые в определенном контексте могут заменять друг друга вследствие пересеклемости множеств своих нотолинейных эквивалентов. Такого рода замены будем для краткости называть синонимичными, хотя речь идет лишь о частичном совпадении интерпретации знамен. В этом существенное отличие от синонимии в естественном языке, которая обычно подразумевает совпадение по основному значению при сохранении различий в оттенках.

В [4] рассматривались лишь синонимичные замены типа $1 \rightarrow 1$ (одно знамя заменяется другим с той же нотолинейной интерпретацией), а также замены типа $2 \rightarrow 2$, $3 \rightarrow 3$, составленные из замен типа $1 \rightarrow 1$. В данной работе рассматривается более широкий класс замен, допускающий возможность замены одного сложного знамени двумя или даже тремя простыми (замены типа $1 \rightarrow 2$, $1 \rightarrow 3$), а также замены типа $2 \rightarrow 2$, $2 \rightarrow 3$ и т.п., непредставимые в виде конкатенации замен типа $1 \rightarrow 1$.

Целью работы является выявление и сопоставление основных типов синонимичных замен по различным гласам на материале певческой книги "Октоих", представленной в форме двознаменника. Результаты могут найти применение: а) для выявления семантики сложных знамен путем интерпретации их через более простые; б) для уточнения классификации попевочных структур, представленных в нестандартной форме; в) для выявления некорректностей в существующей классификации. Отметим, что существенно расширился по сравнению с [4] не только класс анализируемых замен, но и состав гласов: к гласам 1 и 5 добавились еще четыре: 2, 3, 4 и 6.

1. Обоснование и реализация подхода

Выявление синонимичных замен — один из способов анализа *вариативности* эволюционно развивающихся языковых систем. Он широко используется в исследованиях по молекулярной эволюции [6,7], в задачах информационного поиска [8,9] и апробирован авторами данной работы на музыкальных (нотоподобных) текстах (мелодиях песен [10]). С его помощью оценивается скорость эволюции языковых систем, строятся матрицы сходства между элементами алфавита, используемые для поиска близких фрагментов текста, выявляются информативные признаки для задач классификации.

Из многочисленных критериев естественной лексической синонимии (см., например, [9]) для наших целей больше всего подходят следующие три: 1) функциональная общность; 2) полная или частичная взаимозаменяемость (в контексте); 3) общность звучания. В соответствии с третьим критерием из песнопений каждого гласа отбираем всевозможные цепочки знамен (в общем случае *разной* длины) с *одинаковой нотоподобной* интерпретацией. Отбрасываем цепочки одинаковой длины, не содержащие различающихся знамен (это "чистые" повторы без замен). В соответствии со вторым критерием дальнейшему рассмотрению подлежат максимально длинные цепочки знамен, обеспечивающие *совпадение контекстов* в окрестности участка, где имела место замена. Критерий 1 в неявной форме задает ограничение на минимальную длину контекста: она должна быть достаточной для того, чтобы спрогнозировать функциональное назначение цепочки (начало строки, архетип, подвод и т.п.) и избежать тем самым сопоставления случайных цепочек с *разной функциональной нагрузкой*. Заметим, что иногда о функциональной близости цепочек можно судить уже по одноэлементному совпадающему контексту (например, наличие "параклита" ζ в начале цепочек).

Существует ряд дополнительных признаков, учет которых может способствовать проведению более четкой границы между синонимичными заменами и ложными, обусловленными случайными факторами. К ним относятся: 1) частота цепочек, содер-

жащих взаимозаменяемые блоки (чем она выше, тем менее вероятен случайный характер замены); 2) число разных контекстов, в которых встречается данная замена (чем оно больше, тем вероятнее неслучайность замены); 3) число различных нотолинейных эквивалентов у многозначных знамен, для которых наблюдается взаимозаменяемость этих знамен (чем оно больше, тем вероятнее, что мы имеем дело с синонимичной заменой).

Алгоритм отыскания цепочек знамен произвольной длины с одинаковой нотолинейной интерпретацией выглядит следующим образом.

Для каждого из гласов:

1) Объединяем песнопения гласа в один текст $T = T_1 \vee \vee T_2 \vee \dots \vee T_k$, где T_i — i -е песнопение гласа, $1 \leq i \leq k$, k — число песнопений в гласе, \vee — разделитель между песнопениями. Нотолинейную составляющую текста T представляем блоками переменной длины с разделителем " + " между ними. Каждый блок — это цепочка нот, интерпретирующая одно знамя. Для анализируемого двознаменника число нот в интерпретирующей цепочке колеблется от 1 до 6.

Дальнейшая схема обработки повторяет с некоторыми модификациями алгоритм отыскания повторяющихся отрезков текста "АППОРТ-2", основанный на хешировании и описанный в [11].

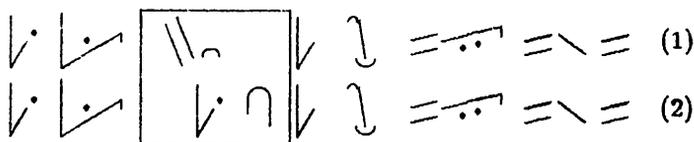
2) Осуществляем итерации по нотолинейной составляющей текста T . Каждая итерация связана с просмотром текста скользящим окном, содержащим ровно l блоков ($l = 1, 2, \dots, l^*$). Окно на каждом шаге сдвигается на один блок за исключением границ песнопений (разделитель \vee не должен попадать в окно). Содержимое окна (в виде конкатенации нотных цепочек от отдельных блоков) служит основой для формирования хеш-кода. Одинаковыми хеш-кодами будут обладать окна, состоящие, возможно, из разного числа блоков, но представимые одной и той же последовательностью нотных знаков (разделители типа " + " в формировании кода не участвуют).

В ходе итераций по l в участках памяти, нумеруемых одинаковыми хеш-кодами, накапливаются цепочки знамен (возможно, разной длины) с одинаковой нотолинейной интерпретацией. Ситуации, когда разным нотолинейным цепочкам соответствует

одинаковый хеш-код (так называемые наложения), устраняются с помощью стандартной списковой схемы [11]. Итерации продолжаются до тех пор, пока списки цепочек знамен с одинаковой нотолинейной интерпретацией продолжают пополняться. Значение l^* , определяющее число итераций, удовлетворяет условию: $l^* \geq l_{max}$, где l_{max} — длина максимального повтора в знаменной составляющей текста T .

3) По окончании итераций позиции хеш-поля оказываются заполненными цепочками знамен с одинаковыми нотолинейными эквивалентами. Для дальнейшего анализа оставляем группы таких цепочек, которые при одинаковой нотолинейной интерпретации отличаются числом знамен либо (при одинаковом числе знамен) способом расстановки разделителя " * " в нотолинейной интерпретации (последнее означает, что одна цепочка знамен не может быть переведена в другую путем замен типа $1 \rightarrow 1$). Рассматривая попарно такие цепочки, локализуем замены как участки, на которых происходит локальная рассинхронизация в расстановке разделителей " * " в нотолинейных интерпретациях цепочек.

ПРИМЕР 1. Цепочки знамен (1) и (2) из гласа 4



каждая с частотой встречаемости $F = 2$ являются попевками — производными от архетипа "грунка" и имеют при разной длине одинаковые нотолинейные интерпретации

$$e4d4 * e2* \quad \boxed{d4c4H2*} \quad e2 * f4e4 * d4e4f2 * e1 * d1, \quad (1')$$

$$e4d4 * e2* \quad \boxed{d4c4 * H2*} \quad e2 * f4e4 * d4e4f2 * e1 * d1, \quad (2')$$

с точностью до расстановки разделителя " * " в (1') и (2'). Здесь и далее звуки первой октавы обозначены заглавными буквами (G — "соль", A — "ля", H — "си"), а звуки второй — прописными (c — "до", d — "ре", e — "ми", f — "фа", g — "соль", a — "ля", b — "си-бемоль"). Стоящие рядом с буквами справа цифры обозначают длительности нот: 1 — целая, 2 — половинная, 4 — четвертная, 8 — восьмая, 2. — половинная с точкой и т.д.

Нетрудно видеть, что рассинхронизация в расстановке разделителя " * " в (1') и (2') имеет место в пределах взятого в рамку участка. Здесь и локализована замена. Анализ соответствующего участка в (1) и (2) показывает, что интерпретируемое тремя нотами $d4c4H2$ знамя $\parallel \cap$ ("сложитие с запятой") заменяется

парой знамен $\downarrow \cdot \cap$, интерпретируемых соответственно как $d4c4$ и $H2$, в результате чего цепочки (1) и (2), имея различия по знаменному составу, не отличаются своими нотолинейными представлениями. Сам факт синонимичной замены, т.е. эквивалентности $\parallel \cap$ и цепочки $\downarrow \cdot \cap$ в указанном выше контексте, будем обозначать знаком " \leftrightarrow " ($\parallel \cap \leftrightarrow \downarrow \cdot \cap$). Такого рода замены могут происходить, когда распеваемые на один и тот же мотив участки стихотворного текста отличаются числом слогов.

ПРИМЕР 2. Цепочки знамен (3) и (4) из глоса 5

$$\begin{aligned} \approx \cup \cap \cap \text{---} &= \boxed{\begin{array}{c} \cap \cap \text{---} \\ \downarrow \end{array}}, & (3) \\ \approx \cup \cap \cap \text{---} &= \boxed{\begin{array}{c} \cap \text{---} \\ \downarrow \end{array}}, & (4) \end{aligned}$$

с частотой встречаемости 2 и 1 соответственно имеют одинаковые нотолинейные интерпретации:

$$\begin{aligned} G4c2H4 * A1 * G1 * \boxed{A4H4c2 * H2} &, & (3') \\ G4c2H4 * A1 * G1 * \boxed{A4H4 * c2H2} &, & (4') \end{aligned}$$

отличающиеся лишь расстановкой разделителя " * " на правом конце (см. выделенные фрагменты из (3') и (4')). Нетрудно видеть, что здесь имеет место замена типа $2 \rightarrow 2$, несводимая к последовательности замен типа $1 \rightarrow 1$. Ее можно представить в виде

$\cap \cap \text{---} \downarrow \leftrightarrow \cap \text{---} \downarrow$. Заметим, что как и в примере 1 сравниваемые цепочки имеют одинаковую функциональную нагрузку: их левая часть — производная от архетипа "долинка", а правая, где имеет место замена, — начальный фрагмент следующей строки (или полевки).

2. Наиболее характерные синонимичные замены для разных гласов

В нижеследующей таблице представлены наиболее часто встречающиеся замены типа 1 → 2, полученные в соответствии с описанной выше методикой на материале гласов 1 ÷ 6 двознаменного "Октоиха" XVII века (РНБ им. Салтыкова-Щедрина, г. Санкт-Петербург, Соловецкое собрание, шифр 618/644). В столбце 2 указан вид замены, в столбце 3 — нотолинейные интерпретации данной замены в тех гласах, где она имеет место (знак "*" разделяет нотные цепочки, соответствующие каждому из знамен в правой части синонимичной замены). В столбце 5 даются преимущественно комментарии к контексту, в котором встречается данная замена. Упоминаемые при этом названия архетипов соответствуют классификации [12].

Отметим следующие моменты, связанные с таблицей.

1) В левой части соотношений, фиксирующих вид замены (см. столбец 2), стоят знамена, интерпретируемые цепочками из 2 ÷ 4 нотных знаков. Однако не все знамена такого типа там представлены. В частности, отсутствуют средне- и высокочастотные знамена, фиксирующие восходящее или нисходящее движение двумя четвертными долями (такие, как $\overset{\cdot}{\curvearrowright}$,

 и т.п.). Это можно объяснить двумя факторами. Во-первых, число знамен, интерпретируемых одной нотой четвертной длительности, мало во всех гласах (как видно из таблиц, представленных в [4,5], четвертные доли встречаются лишь среди множеств нотолинейных эквивалентов знамен \downarrow , $\downarrow\cdot$, \curvearrowright и составляют они не более 5% от их общего числа). Во-вторых, создатели песнопений в той или иной степени использовали, по-видимому, общезыковой принцип "минимизации усилий", сводящийся в данном конкретном случае к минимизации описания — замене двух знамен одним для представления наиболее распространенных мелодических ходов знаменного распева (на ступень вверх или вниз).

Для подтверждения последнего приведем лишь один пример. Комбинация *d4c4* (нисходящее движение) встречается в третьем гласе 134 раза. Из них 75 раз она представлена знаменем \downarrow° , 37 раз — \downarrow , 7 раз — $\downarrow\downarrow$ и т.д. И лишь один раз она фигурирует в виде *d4*c4*, т.е. представлена двумя знаменами: $\downarrow \cap$. Обратная комбинация *c4d4* (восходящее движение) встречается в том же гласе 84 раза, из них почти всегда (81 раз) представлена знаменем \cap° . И вновь лишь один раз эта комбинация фигурирует в виде *c4*d4*, т.е. представлена двумя знаменами: $\cap \downarrow^\circ$. Такого рода примеры характерны для всех гласов.

Если игнорировать в нотолинейном представлении знамен звуковысотную составляющую и ограничиться только ритмической, то все сказанное в данном пункте можно резюмировать следующим образом: в текстах песнопений практически не встречается синонимичных замен с ритмической структурой типа *44* ↔ *4*4*.

2) Продолжим анализ замен в терминах ритмических структур, что вполне оправдано, поскольку многие замены отличаются лишь секвентным переносом при сохранении интервальной структуры. Подавляющее число замен из таблицы имеет ритмическую структуру, относящуюся к одному из четырех типов:

а) *22* ↔ *2*2*; б) *442* ↔ *44*2*;

в) *244* ↔ *2*44*; г) *4444* ↔ *44*44*.

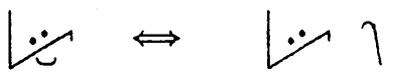
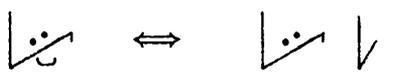
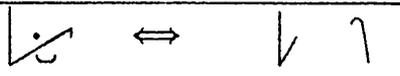
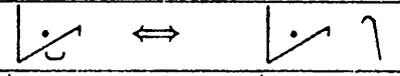
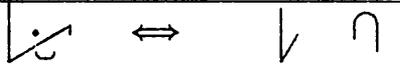
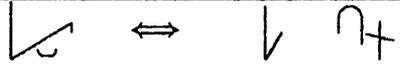
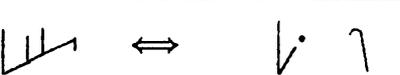
Структура "а" в основном представлена "крюками с подчашием"

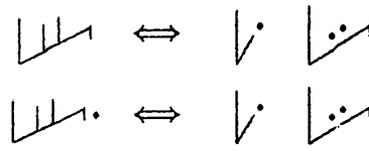
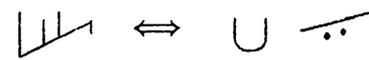
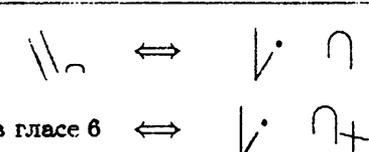
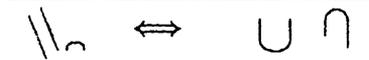
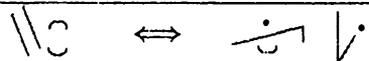
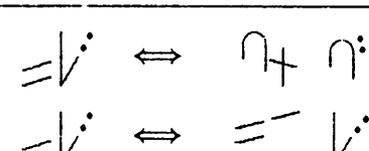
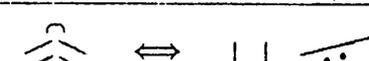
($\downarrow\downarrow$, $\downarrow\downarrow^\circ$, $\downarrow\downarrow^\circ$) и "голубчиком тихим" (\cap°); часто встречающаяся структура "б" — модификациями "стрелы простой"

и "громной" (\uparrow° , \uparrow° , $\uparrow\uparrow^\circ$, $\uparrow\uparrow^\circ$ и т.п.),

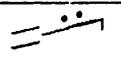
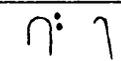
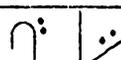
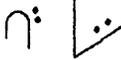
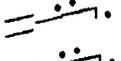
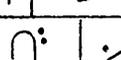
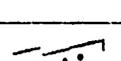
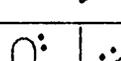
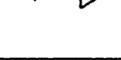
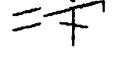
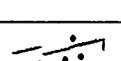
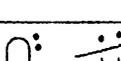
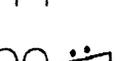
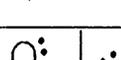
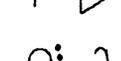
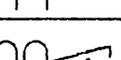
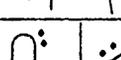
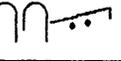
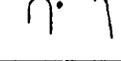
"сложития с запятой" ($\downarrow\downarrow\cap$) и т.д.; относительно редко встре-

Список наиболее распространенных синонимичных замен
типа 1 → 2 для гласов 1 ÷ 6

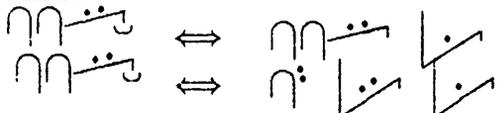
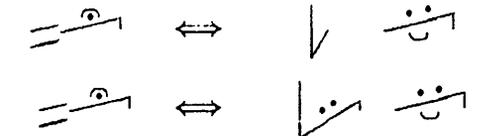
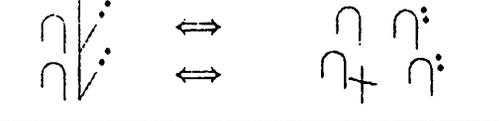
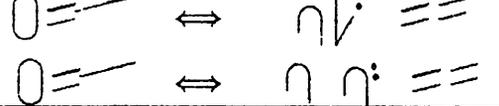
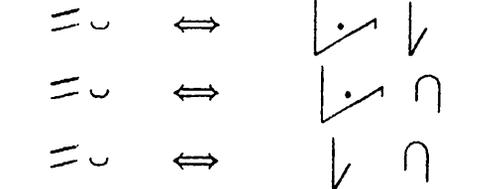
№	Вид замены	Нотолинейная интерпретация	Представлена в гласах:	Комментарии
1	2	3	4	5
1		$g2 * f2$ $f2 * e2$ $d2 * c2$	1,4 1,4 3	Фигурирует в архетипе "кокiza"
2		$g2 * f2$ $f2 * e2$ $a2 * g2$ $e2 * d2$	1,2,3,6 1,4 4 5	В гласе 3 может иметь форму  Фигурирует в начале строк и подводах
3		$f2 * e2$ $a2 * g2$ $g2 * f2$ $c2 * H2$ $e2 * d2$ $d2 * c2$	1,4 2 4 5 5,6 5	Встречается в начале строк и подводах к архетипу "грунка" (гласы 1, 4)
4		$f2 * e2$	2,6	Фигурирует в архетипе "кокiza"
5		$c2 * H2$	5	—— // ——
6		$e2 * d2$	2,4	Фигурирует в подводах и началах строк
7		$H2 * A2$ $e2 * d2$	3 5	—— // ——
8		$d2 * c2$ $A2 * G2$	4 5	—— // ——
9		$f4e4 * f2$ $c4H4 * c2$ $e4d4 * e2$	1 3 4	Встречается в подводах к архетипу "долинка" (глас 1) и в начале строк

1	2	3	4	5
10		$e4d4 * f2$ $g4f4 * g2$ $d4c4 * d2$ $e4d4 * g2$ $f4e4 * g2$	1 1 3,4,5 2 4	Фигурирует преимущественно в начальных фрагментах внутренних разделов песнопений
11		$d4e4 * d4c4$	2,6	Редкая замена в подводе у "мережи"
12		$f4e4 * d2$ $c4H4 * A2$ $d4c4 * H2$ $e4d4 * c2$	1,2,4 1,5,6 4 3,6	Встречается в подводах к "долинке" (глас 5), "грунке" (глас 4) и т.п.
13		$f4e4 * d2$ $c4H4 * A2$	4 3	— // —
14		$e4d4 * c4H4$	1,6	Замена, проясняющая семантику редкого знака
15		$d2 * e4f4$ $A2 * H4c4$ $c2 * d4e4$ $e2 * f4g4$	1 5 6 1	Слабый каданс с переходом на следующую структуру
16		$d4e4 * d2$ $d4e4 * d4c4$ $A4H4 * A4G4$	1 2,3,4,6 3	Замена, размывающая грань между модификациями "грунки" и "колеса"
17		$g4a4 * g4f4$ $f4g4 * f4e4$ $A4c4 * H4A4$ $d4f4 * e4d4$ $d4e4 * d4c4$	2 2,4,6 5 5 6	Часто встречающаяся замена в подводах к архетипам "кулизма", "подъем", "кокиза" и т.п.
18		$c4d4 * c4H4$ $g4a4 * g4f4$ $f4g4 * f4e4$ $d4e4 * d4c4$ $A4c4 * H4A4$	1,5 1 2,4,6 3,4 5	Встречается в довольно изменчивых попевках "срединного" типа: "хамила", "подъем", "срединка"

Продолжение таблицы.

1	2	3	4	5	
19			$e4f4 + g2$	1,2,4,6	Часто встречается в подводах к архетипу "кулизма"
20	 	 	$e4f4 + g2$ $H4c4 + d2$ $e4f4 + g2$	1,2,4,6 3,5 3	Встречается в начальных фрагментах разделов песнопений
21	 	 	$H4c4 + d1$	3	Фигурирует в редких модификациях "подъема"
22			$H4c4 + d2$ $c4d4 + e2$	1,3 6	Встречается в началах строк
23			$d4e4 + f2$ $c4d4 + e2$ $H4c4 + d2$	4 4 5,6	— // —
24			$c2 + d2$ $c2 + e2$ $d2 + e2$	3 3 1,3	— // —
25			$e4f4 + g4f4$ $c4d4 + e4d4$	1,4 5	— // —
26	 	 	$f4g4 + a2$ $c4d4 + e2$ $d4e4 + f2$	1,2,4 3,5 3	Фигурирует в подводе к "кокизе" (глас 1) и в начале разделов песнопений
27	 	 	$f4g4 + a2$ $d4e4 + f2$ $A4H4 + c2$ $d4e4 + f2$ $c4d4 + e2$	1,4 2,3,4,6 5 2,3,4 5	Встречается в начале разделов и в подводах к "кокизе" (глас 1), "мереже" (гласы 2,3,4,6), "кулизме" (глас 3, вариант с "палкой")
28	 	 	$d4e4 + f1$ $c4d4 + e1$	2,4,6 5	см. п.21

Окончание таблицы.

1	2	3	4	5
29		$f4g4a2 * g2$ $f4g4 * a2 * g2$	4 4	Начало строки
30		$e2 * f4e4$ $d2 * e4d4$	1,2 3,5	Встречается в подводах к архетипам "кокiza" (гласы 1,3) и "долинка" (глас 5)
31		$d2 * e4f4$ $A2 * H4c4$ $A2 * c4d4$	2,4,6 3,5,6 5	Фигурирует в архетипе "подъем" (гласы 2,3,4)
32		$d2e4f4 * g1$ $d2e4f4 * g1$	2,6 2,6	Замена, размывающая грань между архетипами "подъем" и "дуда"
33		$e2 * f2$ $d2 * e2$	4 2,4	Замена в архетипе "подъем"
34		$d2 * e2$	3	Новый знак, похожий по смыслу на 
35		$d2 * c2$	3	Замена в архетипе "кулизма"
36		$c2 * H4c4$	6	Замена в архетипе "грунка"
37		$e2 * d2$ $e2 * d2$ $H2 * A2$	2,4,6 4,6 5	Встречается в началах разделов песнопений

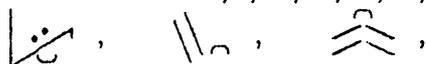
чающаяся структура "в" — "запятой с переводкой" ($\cap \dot{V}$) и "статьей с переводкой" ($\equiv \dot{V}$); структура "г" — многофункциональными знаменами типа $\hat{\equiv}$.

Отсутствуют замены с ритмической структурой вида: *244* \leftrightarrow *24 * 4*; и *442* \leftrightarrow *4 * 42*. Это объясняется тем, что очень ограниченное число знамен интерпретируется цепочками из двух нот разной длительности (типа 4 * 2 или 2 * 4), о чем свидетельствуют статистики, приведенные в [4,5].

3) Наличие синонимичных замен с ритмическими структурами типа "а" \div "г" означает, что в данном случае принцип "минимизации описания" уже не работает или работает лишь частично, иначе для представления какой-либо последовательности звуков (цепочки нот) в знаменной форме всегда бы использовалось знамя из левой части структуры. Поскольку такого не происходит, это означает, что имеются дополнительные факторы, влияющие на выбор знамен для представления той или иной последовательности звуков, например, фактор стихотворного текста (число слогов, ударность и т.п.).

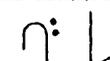
Анализ реальных песнопений показывает, что не всегда такого рода факторы лежат на поверхности, о чем свидетельствует следующий пример. В гласе 6 цепочка *e4d4e2* встречается 19 раз, из них 18 раз она представлена одним знаменем ||| ("два в челну"). Двухзнаменное же представление этой цепочки в форме *e4d4 * e2* встречается всего один раз ($\text{||} \dot{V}$), т.е. работает, по-видимому, принцип минимизации описания. С другой стороны, отличающаяся лишь секвентным переносом на 1 ступень цепочка *d4c4d2* из гласа 5 встречается 2 раза и оба раза представлена тем же знаменем ||| , однако двухзнаменный ее вариант *d4c4 * d2* встречается 29 раз в виде комбинации $\dot{V} \dot{V}$, т.е. ни о какой минимизации описания речь уже идти не может.

4) Важной особенностью замен, представленных в таблице, является встречаемость каждого типа замены, как правило, более чем в одном гласе. Это подтверждает неслучайный характер замен. Наиболее часто встречающимися заменами, представленными в разных контекстах и в разных нотолинейных эквивалентах во всех 6 гласах, являются замены № 2, 3, 12, 18, 20, 27, связанные со знаменами

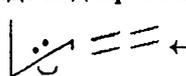
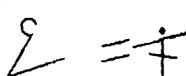
. Некоторые замены носят гласоспецифичный характер (например, № 5, 6, 11, 32 и т.п.).

5) Замен типа $1 \rightarrow 3$ существенно меньше, чем типа $1 \rightarrow 2$. Причиной является ограничение сверху на длину интерпретирующей цепочки ($l \leq 6$) и невысокая частота встречаемости сложной интерпретируемых знамен. Обычно замены типа $1 \rightarrow 3$ не носят самостоятельного характера и возникают в результате замены типа $1 \rightarrow 2$, где одно из знамен в правой части — сложное и, в свою очередь, допускает замену типа $1 \rightarrow 2$ (см. для иллюстрации № 29 и 32 из таблицы).

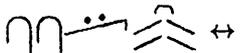
6) Замен типа $2 \rightarrow 2$, $2 \rightarrow 3$, непредставимых в виде конкатенации замен типа $1 \rightarrow 1$ или $1 \rightarrow 2$, достаточно много (особенно первых). В них фигурируют в основном те же знамена, что и в таблице. Некоторые комбинации замен устойчиво повторяются в разных гласах, о чем свидетельствуют приводимые примеры.

а)  \leftrightarrow  \leftrightarrow $\left\{ \begin{array}{l} \overline{d4e4} \overline{f2} \overline{e2} \quad (\text{глас 4}), \\ \overline{A4H4} \overline{c2} \overline{H2} \quad (\text{глас 5}). \end{array} \right.$

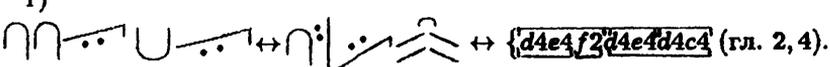
Здесь нотолинейные эквиваленты отдельных знамен из левой части замены указаны в виде надчеркиваний, а из правой части — в виде подчеркиваний.

б)  \leftrightarrow  \leftrightarrow $\left\{ \begin{array}{l} \overline{d2c2} \overline{e1} \quad (\text{глас 3}) \\ \overline{g2f2} \overline{a1} \quad (\text{глас 4}). \end{array} \right.$

в)

 \leftrightarrow  \leftrightarrow $\left\{ \overline{c4d4e2} \overline{d4e4d2} \quad (\text{гл. 3, 5}). \right.$

Замены типа $3 \rightarrow 3$ часто представимы в виде конкатенации замен типа $1 \rightarrow 2$ и $2 \rightarrow 1$:

г)


Замены еще более длинных блоков (3 → 4, 4 → 4 и т.п.) возможны, но всегда представимы в виде конкатенации более простых замен (1 → 1, 1 → 2, 2 → 1 ...). В общем случае разница длин цепочек знамен, образующих синонимичную замену, не превышает 2.

3. Возможности практического использования

3.1. *Выявление "семантики" сложных знамен.* Сложные знамена обычно образуются путем соединения простых знамен (например, ) ("дерблица"), ) и (или) использования (иногда весьма нестандартным образом) вспомогательных знаков — точек, подчаший, крыжей и т.п. (, , ). В первом случае, как правило, интерпретация сложного знамени не является конкатенацией интерпретаций составляющих его простых знамен. Так, "дерблица" () обычно интерпретируется четырьмя четвертными нотами, образующими восходящее движение, тогда как составляющие ее графический образ элементарные знамена интерпретируются (каждое) одной нотой половинной длительности.

Аналогичная ситуация возникает при нестандартном использовании вспомогательных знаков. Добавление точки сверху над "стрелой поводной" () вместо ) должно, казалось бы, привести лишь к секвентному переносу, однако меняется и длина, и ритмическая структура интерпретирующей цепочки.

Уяснению семантики сложных знамен может способствовать выявление синонимичных замен с их участием. При этом сложное знамя выражается через более простые, интерпретация которых нам известна. Приведем соответствующие примеры.

ПРИМЕР 3. Одной из интерпретаций "крюка светлого с подчашием" ($\downarrow \cdot \cdot \nearrow$) является цепочка из двух "крюков" — светлого и мрачного ($\downarrow \cdot \cdot \nearrow \downarrow \cdot \cdot \nearrow$) (см. №2 из таблицы). Поскольку каждый из них интерпретируется одной нотой половинной длительности и "крюк светлый" поется выше "крюка мрачного", заключаем, что добавление "подчашия" к "крюку светлому" увеличивает на одну ноту длину интерпретирующей цепочки и реализует нисходящее движение.

ПРИМЕР 4. Знамя $\downarrow \cdot \cdot \searrow$ ("сложитие с запятой") заменяется во многих контекстах цепочкой $\downarrow \cdot \cdot \searrow \cap$ (см. № 12 из таблицы). Поскольку "стопа с точкой" обычно подразумевает нисходящее движение двумя четвертными, а по поводу "запятой" в толкованиях азбук фигурирует фраза "взять низко", можно предполагать, что при добавлении \cap к $\downarrow \cdot \cdot \searrow$ нисходящее движение будет продолжено и завершится половинной нотой (доминирующая длительность у "палки"). Это и подтверждается данными из таблицы. В рассмотренном примере интерпретация составного знамени $\downarrow \cdot \cdot \searrow$ является конкатенацией интерпретаций составляющих его знамен.

ПРИМЕР 5. Попытаемся спрогнозировать, к чему приведет появление "подчашия" у рассмотренного в предыдущем примере знака. По аналогии с примером 3 можно предположить, что длина интерпретирующей цепочки увеличится на одну ноту, будет продолжено нисходящее движение и уменьшится длительность третьей ноты, поскольку она перестает быть последней в комбинации (в противном случае возникает немотивированное ритмическое замедление посреди нисходящего движения). Анализ синонимичной замены для $\downarrow \cdot \cdot \searrow$ (№ 14 в таблице) подтверждает приведенное рассуждение.

ПРИМЕР 6. Для "стрелы поводной простой" и "продленной" в гласе 5 выявляются следующие синонимичные замены:



Трактовка "стрелы продленной" показывает, что в данном случае указание об увеличении длительности, обусловленное дополнительной точкой, распространяется на все нотные знаки в интерпретации (длительности удваиваются). Это не совсем традиционная трактовка, поскольку чаще всего в подобной ситуации увеличивается длительность последнего нотного знака (см., например, № 20 и 21 из таблицы).

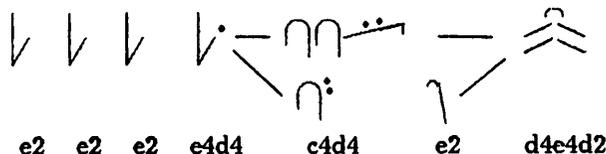
Заканчивая рассмотрение примеров на эту тему, заметим, что с определенной натяжкой синонимичные замены могут выявляться и без привлечения нотолинейного текста (т.е. без использования третьего критерия синонимии). Интерпретация сложных знамен через простые может быть в этом случае непосредственно использована для целей дешифровки.

3.2. Уточнение классификации попевок по аналогии. Известная нам классификация попевок, предложенная А.Н. Кручининой [12] на основе анализа наиболее известных кокизников — сборников попевок, составленных древнерусскими распевщиками — выделяет 24 семейства попевок. Каждое семейство характеризуется архетипом — цепочкой из трех знамен, определяющих наиболее устойчивую кадансовую часть попевки. Многообразие попевок в семействе задается варьированием архетипа и добавлением к нему различных предваряющих цепочек из нескольких знамен — так называемых подводов.

Попытки выделить в реальном песнопении все попевочные структуры с помощью данной классификации быстро убеждают в том, что она неполна.² Проявляется это в том, что обнаруживается структура явно попевочного характера, не вкладывающаяся в классификацию. Для ее отождествления приходится прибегать к рассуждениям по аналогии и отправной точкой здесь может служить поиск синонимичной в описанном выше смысле

²Это следствие неполноты исходных кокизников, а не результат недоработки автора [12].

ПРИМЕР 8. Цепочки



из гласа 5 имеют одинаковую нотолинейную интерпретацию и отличаются лишь заменой $\eta\eta\dot{\cdot}$ ↔ $\eta\dot{\cdot}\eta$ ↔ $c_4d_4 * e_2$. Суффиксная часть второй цепочки $\{\eta\dot{\cdot}\eta \cong\}$ идентифицируется как модификация архетипа "повертка". По аналогии и первую цепочку относим к этому семейству.

ПРИМЕР 9. Цепочки

$\{U \eta\dot{\cdot}\cong\}$ и $\{\eta \eta\dot{\cdot}\cong\}$ из гласа 4 имеют одинаковую нотолинейную интерпретацию с цепочкой $\eta\dot{\cdot}\eta\dot{\cdot}\cong$, которая в соответствии с [12] идентифицируется как "подъем". Две первые цепочки отличаются от последней синонимичными заменами типа $1 \rightarrow 1$ и $1 \rightarrow 2$:

$$U \leftrightarrow \eta \leftrightarrow \eta\dot{\cdot} \leftrightarrow f_4e_4 \text{ и } \eta\dot{\cdot} \leftrightarrow \eta \eta\dot{\cdot} \leftrightarrow d_2 * e_4f_4.$$

По аналогии относим их к семейству "подъем".

3.3. Выявление некорректностей в существующей классификации. В предыдущем разделе мы выявляли попевочные структуры, которые не вкладывались напрямую в существующую классификацию, и искали для них синонимичные аналоги среди известных структур. Возможна однако ситуация, когда обе рассматриваемые структуры однозначно идентифицируются и относятся в соответствии с [12] к разным семействам, однако имеют одинаковую нотолинейную интерпретацию и сводимы друг к другу с помощью синонимичных замен. Такого рода случаи мы относим к категории некорректностей в существующей классификации. Их трудно избежать, если отталкиваться при построении клас-

сификации не от текстов песнопений, а от уже существующих и, по-видимому, частично противоречивых классификаций.

Проиллюстрируем повторяющиеся во многих гласах случаи совпадения на мелодическом уровне попевочных структур, относящихся в соответствии с [12] к разным семействам.

ПРИМЕР 10. Дважды встречающиеся в гласе 2 цепочки

Diagram illustrating two chains of notes (цепочки) in the same vowel (гласе 2). The first chain consists of notes $d4c4$, $d4e4$, and $f2$. The second chain consists of notes $d4e4+$, $d4c4$, $d1$, and $c1$. The notes are represented by stems with flags and dots, and some have horizontal lines above them. The classification labels are placed below each note.

имеют одинаковые нотолинейные интерпретации, но относятся к разным семействам: первая — модификация "грунка", вторая — ("мережа") — производная от "колеса". Точно эти же пары с той же интерпретацией встречаются в гласах 4,6, с секвентным переносом — в гласе 3, с небольшими ритмическими отличиями — в гласе 1.

ПРИМЕР 11. Цепочки из гласа 6

Diagram illustrating two chains of notes (цепочки) in the same vowel (гласе 6). The first chain consists of notes $d4c4$, $d4e4+$, and $d4c4$. The second chain consists of notes $d2e4f4+$ and $g1$. The notes are represented by stems with flags and dots, and some have horizontal lines above them. The classification labels are placed below each note.

имеют одинаковую нотолинейную интерпретацию, хотя принадлежат к разным семействам: первая — модификация "подъема", вторая — производная от "дуды". Аналогичная пара (с "подверткой" у "палки") встречается в гласе 2.

ПРИМЕР 12. Дважды встречающиеся в гласе 1 цепочки

Diagram illustrating two chains of notes (цепочки) in the same vowel (гласе 1). The first chain consists of notes $e2$, $e2$, $f4e4$, and $d2$. The second chain consists of notes $e2d4$ and $e4d2$. The notes are represented by stems with flags and dots, and some have horizontal lines above them. The classification labels are placed below each note.

переводятся друг в друга синонимичной заменой типа 2 → 2 (надчеркивание относится к интерпретации знамен первой цепочки, подчеркивание — ко второй). Три последних знамени в 1-й и 2-й цепочках соответственно — архетипы "хамилы" и "повертки" в чистом виде.

Перечень подобных примеров может быть легко расширен. В частности, еще большее число совпадений, но уже в пределах одного семейства, наблюдается за счет синонимии подводов (см. пример 1). И, наконец, немало случаев совпадения на мелодическом уровне цепочек попевочной структуры, ни одна из которых не вкладывается в [12]. Они требуют специального рассмотрения.

З а к л ю ч е н и е

Синонимичные в определенном контексте замены (а в более общем случае — блочные подстановки) являются важным элементом эволюционного процесса, сопровождающего функционирование языковых систем естественного типа. Важную роль в понимании эволюционного процесса имеют количественные оценки проявления вариативности в целом и синонимии в частности.

В данной работе предложена *схема выявления синонимичных замен* в древнерусских знаменных песнопениях, основанная на использовании *двознаменников* — текстов, написанных параллельно в знаменной и нотолинейной форме. Она удовлетворяет ряду критериев естественноразговорной синонимии и напоминает по смыслу процедуру определения синонимичных кодонов для разных аминокислот в гипотетической ситуации, когда генетический код неизвестен, но заданы для одной и той же аминокислотной последовательности ее всевозможные ДНК-варианты. С некоторой натяжкой разработанная схема может быть перенесена впоследствии на обычные знаменные тексты, не содержащие нотолинейной интерпретации.

Разработан алгоритм выявления синонимичных замен, основанный на технике хеширования. Алгоритм апробирован на шести из восьми гласов двознаменного "Октоиха" XVII века. Проведен анализ выявленных замен и блочных подстановок. Проиллюстрированы возможности использования полученных результатов для целей дешифровки, в частности, для *интерпретации*

сложных знамен путем замены их цепочками более простых знамен, для уточнения *классификации нестандартных попевочных структур*, а также для *выявления некорректностей* в существующей классификации.

Л и т е р а т у р а

1. БРАЖНИКОВ М.В. Древнерусская теория музыки. — Л.: Музыка, 1972. — 423 с.

2. Дешифровочный подход к анализу древнерусских песнопений /И.В.Бахмутова, В.Д.Гусев, Т.Н.Титкова, Б.А.Шиндин //Анализ последовательностей и таблиц данных. — Новосибирск, 1994. — вып. 150: Вычислительные системы. — С. 107-130.

3. Об одном подходе к проблеме дешифровки древнерусских песнопений в невменной записи /И.В.Бахмутова, В.Д.Гусев, Т.Н.Титкова, Б.А.Шиндин //Труды Сибирской конференции по прикладной и индустриальной математике, посвященной памяти Л.В.Канторовича. — Новосибирск, 1997. — Т.2. — С.1-10.

4. БАХМУТОВА И.В., ГУСЕВ В.Д., ТИТКОВА Т.Н. Компьютерный анализ древнерусских двознаменников: многозначность соответствий "знамя - нота" и "нота - знамя" //Искусственный интеллект и экспертные системы. — Новосибирск, 1996. — Вып. 157: Вычислительные системы. — С. 68-100.

5. БАХМУТОВА И.В., ГУСЕВ В.Д., ТИТКОВА Т.Н. Компьютерный анализ древнерусских двознаменников: алфавит знамен и гласоразличительные признаки //Искусственный интеллект и экспертные системы. — Новосибирск, 1997. — Вып. 160: Вычислительные системы. — С. 124-146.

6. Проблемы теории молекулярной эволюции /Под ред. Ратнера В.А. — Новосибирск: Наука, 1985. — 263с.

7. RATNER V.A., ZHARKIKH A.A., KOLCHANOV N.A. e.a. The Problems of the Theory of Molecular Evolution. — Berlin e.a.: Springer - Verlag, 1992.

8. МИХЕЕВ М.Ю. Алгоритм автоматического индексирования: разрешение синонимии при отождествлении новых слов и словосочетаний //ИТИ, сер 2. — 1988. — №2. — С. 10-17.

9. РЫЧКОВА Л.В. О критерии синонимичности лексических единиц, участвующих в фактографическом информационном поиске //НТИ, сер 2. – 1988.

10. БАХМУТОВА И.В., ГУСЕВ В.Д., ТИТКОВА Т.Н. Закономерности варьирования в текстах различной природы и методика их количественного исследования //Анализ текстов и сигналов. – Новосибирск, 1987. – Вып. 123: Вычислительные системы. – С. 25-49.

11. ГУСЕВ В.Д., КОСАРЕВ Ю.Г., ТИТКОВА Т.Н. О задаче поиска повторяющихся участков текста //Ассоциативное кодирование. – Новосибирск, 1975. – Вып. 62: Вычислительные системы. – С. 49-71.

12. КРУЧИНИНА А.Н. Попевка в русской музыкальной теории XVII века: Дисс. ... канд. искусствовед. наук: 17.00.02 – Л., 1979.

Поступила в редакцию
2 декабря 1997 года.