

УДК 519.766.44

ВЫЯВЛЕНИЕ И АНАЛИЗ СХОДНЫХ ФРАГМЕНТОВ В МУЗЫКАЛЬНЫХ  
ПРОИЗВЕДЕНИЯХ

И.В.Бахмутова, В.Д.Гусев, Р.Х.Зарипов, Т.Н.Титкова

## В в е д е н и е

В основу выявления сходных фрагментов может быть положено понятие повтора, трактуемое в данной работе весьма широко. Повторы являются важными структурными элементами символьных последовательностей различной природы — текстов на естественном языке, первичных структур нуклеотидных молекул, текстов программ. Интуитивно ясно, что повторы играют существенную роль и в формировании музыкальных произведений. Это инициировало ряд исследований кибернетического [1] и математического [2,3] плана.

Идеальным (или совершенным) повтором назовем два идентичных участка (слова) текста, расположенных произвольным образом относительно друг друга. Если какое-то слово встречается в тексте  $F$  раз, различные его вхождения, рассматриваемые попарно, образуют  $F(F-1)/2$  совершенных повторов. Неидеальным (или несовершенным) повтором будем называть пару участков текста, близких в определенном смысле (например, в смысле хэмингова или редакционного [4] расстояния). Описание последовательностей в терминах повторов обеспечивает широкие возможности для решения разнообразных задач классификации текстов (по жанру, стилю, времени создания и т.д.).

Объектом исследований в данной работе выступают тексты музыкальных произведений — мелодии песен. Сочиняя мелодию, композитор не разделяет этот сложный психологический процесс на отдельные операции. Мелодия часто появляется в его сознании сразу, в относительно завершенной форме. Композитор не всегда осознает, что при сочинении музыки он использует не только общеизвестные закономерности и правила композиции, но и специфические, присущие

только ему, его стилю. Многие композиторы осознанно или интуитивно используют в своих произведениях в качестве основы или темы известные мелодии народных песен или профессиональных сочинений. Композитор не всегда сообщает о факте заимствования темы для своих сочинений, поскольку иногда он может об этом не подозревать.

Целью данной работы является развитие подходящего аппарата для выявления и классификации сходных музыкальных фрагментов произвольной длины и апробация его на достаточно больших массивах музыкальных текстов. Основу подхода составляют, с одной стороны, понятия интонации и интонационно-метрической структуры как носителя образа сходных интонаций [5], с другой стороны, — понятия частотного и совместного частотного спектра [6] музыкальных текстов как аналога интонационной памяти композитора.

Развиваемый аппарат предоставляет интересные возможности для изучения психологии музыкального творчества. Наряду с этим рассматриваются и многочисленные практические аспекты использования предлагаемой методики. Главным из них является создание и пополнение машинного банка интонационных структур, играющего роль интонационной памяти композитора, как реального (конкретного), так и условного (типичного), аккумулирующего наиболее характерные черты данного стиля, жанра, временной эпохи. Подобный банк может служить информационной базой для экспертизы тех или иных музыкальных произведений, а также основой для создания новых мелодий и автоматической количественной оценки их "благозвучности".

## §1. Интонация и интонационно-метрическая структура

Интонация — элементарное музыкальное высказывание, обладающее относительно самостоятельным выразительным значением (семантическая единица в музыке). Интонация обычно состоит из 2-3 и более звуков в одноголосии. Членение музыкального потока на интонации и определение их смысла обусловлено как объективными факторами, так и субъективными, в том числе музыкально-слуховым воспитанием и опытом слушателя.

Аналогом интонации в естественном языке является морфема — элементарная смыслонесущая единица языка. В отличие от интонации элементы субъективизма сказываются в определении морфемы в значительно меньшей степени. Несомненный интерес представляют попытки автоматического вычлениения морфем из слитного текста, основанные

на отождествлении элементарных повторяющихся участков текста с отдельными морфемами [7, 14]. Подобный подход к интонации в рамках излагаемой ниже методики, по-видимому, позволил бы в значительной степени избавиться от элементов субъективизма в ее определении.

Интонация является важной классификационной характеристикой. У каждого композитора можно выделить типовые интонации, вариантами которых являются все остальные. Совокупность таких интонаций образует интонационный словарь композитора. Аналогично можно говорить об интонационном словаре какого-либо жанра, об интонационном словаре нации, эпохи. Каждый из этих словарей находится в постоянном развитии. Выявление закономерностей этого развития представляет большой интерес и немислимо без использования средств вычислительной техники.

Изучению интонации посвящено много работ, но они, в основном, носят качественный характер. Из количественных характеристик интонации можно отметить высотную, интервальную, длительностную, метрическую, ритмическую. В работах по количественному анализу рассматривается, как правило, лишь одна звуковысотная сторона интонации, т.е. последовательность интервалов между соседними звуками. В данной работе интонация рассматривается сразу с двух характеризующих ее сторон: звуковысотной (интервальной) и метрической, оказывающей, как показано ниже (см. гл. 3, § 4), сильное влияние на звуковысотную.

Аргументировать такой подход можно экспериментами по слуховому восприятию интонаций с одинаковой звуковысотной структурой, но разными метром, ритмом, тональностью и т.д. Как правило, человек способен на слух уловить сходство между двумя интонациями, если кроме звуковысотной структуры у них совпадает еще и метрическая, т.е. соотношенность отдельных звуков с сильными и слабыми долями такта. Роль остальных варьирующих факторов проявляется слабее.

Итак, пусть  $z_1, z_2, \dots, z_l$  — интонация, состоящая из последовательности  $l$  нот. Интервально-метрической структурой  $l$ -го порядка ((IS) $_l$ -структурой) назовем последовательность пар  $I_k S_k$ ,  $k = 1, 2, \dots, l-1$ , где  $I_k$  — количество ступеней между высотами  $k$ -го и  $(k+1)$ -го звуков интонации, а  $S_k$  — метрическая характеристика, дающая представление об относительной силе этих звуков. Формально  $I_k = W_{k+1} \ominus W_k$ , где  $W_k$  — высота  $k$ -го звука, а  $\ominus$  — символ специальной операции вычитания;

$$S_k = \begin{cases} \text{sign}(p_{k+1} - p_k) & \text{при } p_{k+1} \neq p_k, \\ -1 & \text{в противном случае,} \end{cases}$$

где  $p_k$  - относительная сила доли такта, соответствующей  $k$ -му звуку.

Способ получения  $(IS)_1$ -структуры из нотной записи проиллюстрируем на примере мелодии романса "Мой костер в тумане светит" (рис. 1)

The figure shows a musical staff with a treble clef, a key signature of one sharp (F#), and a 3/4 time signature. The melody consists of six notes: G4, A4, B4, A4, G4, F#4. Below the staff, the notes are numbered 1 through 6. To the right of the staff, the relative strength of each note is indicated by a vertical line with a horizontal bar at the top, labeled  $P_k$ . Below this, the notes are numbered 1, 4, 2, 5, 3, 6. To the left of the staff, the interval structure  $I_k$  is given as 1-, I-, 2+, 2+, 2-. Below that, the relative strength  $P_k$  is given as 2, 3, 1, 3, 1, 3. At the bottom, the sign  $S_k$  is given as +, -, +, -, +. To the right of the staff, the text reads: "Относительная сила восьмых нот в такте (размер 3/4)".

Рис. 1

При переходе к горизонтальной записи принимаем следующее соглашение: сначала указываем значение интервала (т.е. цифру), затем его знак, а после него знак стопы. Соответствующая  $(IS)_6$ -структура тогда имеет вид: (I-+ I--2++2+- 2-+).

Из определения  $(IS)_1$ -структуры следует, что одно и то же ее значение может отвечать некоторому множеству различных 1-нотных интонаций. Они могут отличаться ритмическими фигурами, перемещением мелодического рисунка вверх или вниз, но при сохранении интервально-метрических отношений воспринимаются на слух как сходные между собой. Если в тексте мелодии присутствует несколько таких интонаций, то при вышеописанном переходе от нотной записи мелодии к ее  $IS$ -представлению они будут образовывать идеальные повторы. Ниже кратко описаны способы представления и обнаружения повторов в произвольных (в общем случае) символьных последовательностях и дана семантическая интерпретация предлагаемых конструкций применительно к задачам анализа музыкальных текстов.

## §2. Частотные спектры музыкальных текстов

Будем называть музыкальным текстом интервально-метрическое представление анализируемой мелодии, а 1-граммой - любую связанную подпоследовательность этого текста, состоящую из 1 символов. Очевидно, что каждую 1-грамму можно интерпретировать как  $(IS)_{1+1}$ -структуру.

Всякий текст

$$T = z_1 z_2 z_3 z_4 z_5 \dots z_N \quad (1)$$

длины  $N$  содержит ровно  $N-1+1$  1-грамм ( $1 \leq N$ ), выделяемых последовательными сдвигами вдоль текста рамки шириной в 1 символ (случай  $1 = 3$  иллюстрирован подчеркиванием в (1)). Поскольку среди них могут быть повторяющиеся, количество разных 1-грамм  $M_1 \leq N-1+1$ . Частотная характеристика порядка 1 текста  $T$  есть совокупность элементов  $\Phi_1(T) = \{\phi_{1_1}, \phi_{1_2}, \dots, \phi_{1_{M_1}}\}$ , где каждый элемент  $\phi_{1_i}$  ( $1 \leq i \leq M_1$ ) есть пара:  $\langle i$ -я 1-грамма, частота ее встречаемости в тексте  $\rangle$ . Полный частотный спектр текста  $T$  - совокупность частотных характеристик

$$\Phi(T) = \{\phi_1(T), \phi_2(T), \dots, \phi_{1_{\max}}(T), \phi_{1_{\max}+1}(T), \phi_{1_{\max}+2}(T)\}, \quad (2)$$

где  $1_{\max}$  - максимальное значение  $1$ , при котором в тексте присутствуют повторяющиеся 1-граммы.

Каждая последующая характеристика в (2) содержит все большую информацию о тексте. По характеристике  $\phi_{1_{\max}+2}(T)$  текст уже может быть восстановлен однозначно [6], и в этом смысле мы и говорим о "полном" спектре. На практике чаще всего используется редуцированный спектр ( $1 \leq 1_{\max}$ ); более того, единичные, т.е. однократно встречающиеся 1-граммы, в него не включают (они легко могут быть выявлены по тексту, если пометить в нем все кратные 1-граммы).

Из форм представления частотных характеристик отметим две: представление с упорядочением 1-грамм по убыванию частоты встречаемости (оно удобно для выявления наиболее типичных 1-грамм для данной выборки музыкальных текстов) и представление с упорядочением 1-грамм  $((IS)_{1+1}$ -структур) по величине интервалов между соседними звуками (оно удобно для быстрого поиска нужной 1-граммы).

Понятие полного частотного спектра удобно использовать, когда мы хотим в целом охарактеризовать какое-то отдельное произведение либо группу родственных в каком-то смысле произведений, например, произведений одного автора или одного жанра. Если же мы хотим как-то дифференцировать отдельные произведения анализируемой выборки по степени их сходства или различия, удобно воспользоваться понятием совместного частотного спектра.

Он определяется по той же схеме, что и (2), но уже применительно к двум (или более) текстам. Назовем совместной частотной характеристикой 1-го порядка текстов  $T_1$  и  $T_2$  совокупность элементов  $\Phi_1(T_1, T_2) = \{\phi_{11}(T_1, T_2), \phi_{12}(T_1, T_2), \dots, \phi_{1M_1}(T_1, T_2)\}$ , где  $M_1 = M_1(T_1, T_2)$  - количество различных 1-грамм, общих для обоих текстов ( $0 \leq M_1(T_1, T_2) \leq \min(M_1(T_1), M_1(T_2))$ ), а каждый элемент  $\phi_{1i}(T_1, T_2)$  ( $1 \leq i \leq M_1(T_1, T_2)$ ) есть тройка:  $\langle i$ -я 1-грамма, частота ее встречаемости в  $T_1$ , частота встречаемости в  $T_2 \rangle$ . Совместный частотный спектр текстов  $T_1$  и  $T_2$  есть совокупность совместных частотных характеристик

$$\Phi(T_1, T_2) = \{\phi_1(T_1, T_2), \phi_2(T_1, T_2), \dots, \phi_L(T_1, T_2)\}, \quad (3)$$

где  $L$  - максимальное значение  $l$ , при котором в текстах  $T_1$  и  $T_2$  еще есть общие 1-граммы (т.е.  $M_l(T_1, T_2) \neq 0$ , а  $M_{L+1}(T_1, T_2) = 0$ ).

Аналогично (3) может быть определен и частотный спектр для  $m > 2$  текстов -  $\Phi(T_1, T_2, \dots, T_m)$  - за единственным исключением. Слишком сильным, как правило, является требование о включении в него лишь 1-грамм, общих для всех  $m$  текстов. Целесообразнее определять совместную частотную характеристику текстов как совокупность 1-грамм, общих хотя бы для пары текстов. Именно так понимается совместный частотный спектр в данной работе.

На 1-граммном уровне совместный частотный спектр есть аналог операции пересечения двух текстов. Если же нас интересуют отличия между текстами, следует вычислить  $\Phi(T_1)$ ,  $\Phi(T_2)$  и удалить из каждого 1-граммы, вошедшие в  $\Phi(T_1, T_2)$ . Подобная схема действий может быть положена в основу классификации текстов по стилю, жанру, авторству и т.п.

Наряду с совершенными повторами, фигурирующими в характеристиках (2), (3), интерес могут представлять и различные классы несовершенных повторов, в первую очередь класс (1, k)-повторов. Под (1, k)-повтором ( $1 \leq k < l$ ) будем понимать произвольную пару 1-грамм тек -

ста, отличающихся друг от друга по  $k$  позициям. Совокупность всевозможных  $(1, k)$ -повторов текста  $T$  образует  $\Phi_{1, k}(T)$ -характеристику текста. По аналогии с (3) для пары текстов можно ввести понятие  $\Phi_{1, k}(T_1, T_2)$ -характеристики, подразумевая, что 1-граммы, образующие каждый  $(1, k)$ -повтор, берутся из разных текстов.

Несовершенные повторы указанного типа возникают в результате вариаций определенных интонаций путем замен отдельных элементов  $(IS)_1$ -структуры другими. Анализ  $\Phi_{1, k}(T)$ -характеристик текста позволяет типизировать характер вариаций, осуществляемых целенаправленно, осознанно. Анализ  $\Phi_{1, k}(T_1, T_2)$ -характеристик дает примеры неосознанных заимствований с вариациями.

Более широкий класс несовершенных повторов составляют участки текста, близкие в смысле редакционного расстояния [4]. Под редакционным расстоянием двух последовательностей понимается минимальное число редакционных операций, переводящих одну последовательность в другую. Множество редакционных операций включает в себя операции замены одного символа другим, вставки нового символа или устранения имеющегося. Возможными (но не единственными) характеристиками несовершенного повтора указанного вида могут служить два параметра:  $l$  - длина минимального (из двух рассматриваемых) участка и  $\epsilon$  - значение редакционного расстояния. При фиксированных  $l$  и  $\epsilon$  длина второго участка, образующего несовершенный повтор, может колебаться в пределах от  $l$  (участки отличаются одними заменами) до  $l + \epsilon$  (участки отличаются одними вставками).

Имея в виду указанные параметры, можно аналогично тому, как это делалось выше, ввести понятия  $\Phi_{l, \epsilon}(T)$ -характеристики одного текста и  $\Phi_{l, \epsilon}(T_1, T_2)$ -характеристики двух текстов (последняя легко обобщается на случай  $n > 2$  текстов).  $\Phi_{l, \epsilon}(T)$  - и  $\Phi_{l, \epsilon}(T_1, T_2)$  - характеристики дают примеры варьирования интонаций путем замен, вставок и устранения отдельных элементов интонационной структуры.

Получение характеристик (2) и (3) реализовано в рамках пакета прикладных программ "СИМВОЛ" [8]. Отметим, что если не оптимизировать процедуру вычисления (2) и (3), а осуществлять прямой перебор путем сопоставления каждой пары 1-грамм из числа тех, что присутствуют в тексте, то вычисление всего лишь одной характеристики 1-го порядка потребует затрат  $O(N^2 \cdot l)$ , где  $N$  - длина текста, а  $l$  - размер фиксируемых повторов. Таким образом, уже для объемов текста  $N \sim 10^4$  символов и выше и значений  $l \sim 20+50$  (а

в обработанных текстах встречались повторы и большей длины) такой подход оказывается неперспективным даже при использовании достаточно мощных ЭВМ.

Для преодоления этих трудностей потребовалась разработка специальных итеративных методов, основанных на идеях хеширования [16] и использовании по ходу 1-й итерации (т.е. при вычислении  $\Phi_1(T)$ ) информации, полученной на (1-1)-й итерации. На содержательном уровне идея хеширования состоит в построении достаточно простого в вычислительном плане отображения множества 1-грамм текста на множество адресов тех участков оперативной памяти, где хранится информация о каждой 1-грамме (число вхождений ее в текст, места вхождения и т.д.). Однократный просмотр текста позволяет рассортировать все 1-граммы по своим участкам, т.е. частотная характеристика 1-го порядка строится со средней трудоемкостью  $O(N \cdot 1)$ . Учет информации о предыдущей итерации обеспечивает более медленный (по сравнению с линейным) рост общей трудоемкости (т.е. трудоемкости вычисления всего спектра) в зависимости от числа итераций.

Алгоритмы вычисления совместного частотного спектра  $\Phi(T_1, T_2, \dots, T_m)$  и спектра несовершенных повторов  $\Phi_{1,k}(T_1, T_2, \dots, T_m)$  для  $m > 2$  текстов реализованы впервые в рамках данной работы. В обоих алгоритмах предусмотрена возможность фиксации информации о местах вхождения 1-грамм, образующих повторы. Прототипом первого алгоритма послужил алгоритм, описанный в [16], прототипом второго – алгоритм, представленный в данном сборнике (см. стр. 107). Важную роль при отыскании несовершенных повторов играет учет специфических свойств используемой метрики и порядок сравнения 1-грамм в процессе перебора.

Отметим также, что в ППП "СИМВОЛ" включены некоторые средства классификации последовательностей, в частности, процедуры вычисления редакционного расстояния и различных мер близости, алгоритмы кластеризации и выявления грамматических единиц и отношений между ними по ограниченной выборке. Простейшие процедуры реверсирования текста (записи его в обратном порядке) и перекодировки элементов алфавита обеспечивают обнаружение закономерностей, отличных от перечисленных, но, в принципе, также сводящихся к повторам, например, симметрий (т.е. закономерностей вида  $\dots z_1 z_2 \dots \dots z_{1-1} z_1; z_1 z_{1-1} \dots z_2 z_1 \dots$ , где  $z_i (1 \leq i \leq l)$  – произвольный элемент алфавита.



Перекодировкой можно отразить и факт изменения длительностей (замены четверти на две восьмушки (  $\text{♩} = \text{♪♪}$  ), восьмушки на две шестнадцатых (  $\text{♩} = \text{♩̣̣}$  ) и т.д.), что приведет к увеличению длины мелодии, так как к каждому интервалу будет справа приписан интервал "прима", означающий повторение звука. Эта перекодировка обеспечивает возможность обнаружения некоторых новых закономерностей (инвариантов по отношению к изменению ритмического рисунка). Примером может служить песня "Молодежная" И. Дунаевского, которая по мелодическому рисунку является копией песни "По Дону гуляет", но отличается по ритмическому рисунку.

В целом перекодировку следует рассматривать как весьма универсальное средство выявления инвариантной структуры, замаскированной вариациями. Под каждый способ вариации следует подбирать свой способ перекодировки. В достаточно простых случаях перекодировка носит "бесконтекстный" характер, т.е. каждый элемент текста заменяется одной и той же кодовой комбинацией вне зависимости от окружения. Более сложные случаи могут потребовать применения "контекстно-зависимой" перекодировки.

### §3. Апробация методики

Экспериментальный материал включал в себя 4 сборника: русских народных песен (в дальнейшем - сборник № 1 [9], 106 мелодий), песен советских композиторов (сборник № 2 [10], 216 мелодий), песен о войне (сборник № 3 [11], 58 мелодий) и песен Шуберта (сборник № 4 [12], 122 мелодии). Совместный объем закодированного и обработанного материала составил примерно  $5 \cdot 10^4$  нот. Целью эксперимента являлось: а) выявление и классификация бытующих интонаций; б) анализ влияния метрической структуры на звуковысотную; в) иллюстрация классификационных возможностей предлагаемой системы описания; г) выявление и анализ неосознанных заимствований. Проведены следующие серии экспериментов:

1) для каждого множества мелодий был получен полный частотный спектр  $\Phi(t^{(i)})$ , где  $t^{(i)} = t_1 \sqcup t_2 \sqcup \dots \sqcup t_n$  - конкатенация всех мелодий  $i$ -го сборника (здесь  $n_i$  - число мелодий в  $i$ -м сборнике,  $i=1 \div 4$ , а " $\sqcup$ " - символ-разделитель, используемый для удаления из частотного спектра 1-грамм, попадающих на стыки мелодий);

2) для каждого множества мелодий был вычислен совместный частотный спектр  $\Phi(T_1, T_2, \dots, T_{m_i})$  с указанием мест вхождения обших 1-грамм (для  $l \geq 7$ ) в разные мелодии;

3) осуществлена перекрестная обработка сборников 2 и 4 (т.е. вычислен  $\Phi(T^{(2)}, T^{(4)})$ ), а также сборников 1, 2, 3 (т.е. вычислен  $\Phi(T^{(1)}, T^{(2)}, T^{(3)})$ ).

Ниже отмечены наиболее интересные, на наш взгляд, интонационно-метрические закономерности, характеризующие отдельные мелодии или классы мелодий. Не все закономерности сопровождаются содержательной трактовкой и даже там, где это имеет место, она не является бесспорной. Авторы отдают себе отчет в том, что по мере накопления материала будет уточняться как само понятие закономерности, так и трактовка обнаруживаемых закономерностей.

Иллюстрацию большинства закономерностей будем проводить на примере сборника №2 (песни советских композиторов). Для сокращения записи будут указываться лишь номера мелодий. В приложении I приведены названия соответствующих мелодий и фамилии композиторов.

#### §4. Анализ полных частотных спектров (эксперимент I)

I. Интегральные характеристики спектров. Важнейшими характеристиками спектров являются значения параметров  $M_1, E_1^1, F_{\max}^{(1)}$  ( $l = 1, 2, \dots, l_{\max}$ ) и  $l_{\max}$  (здесь  $M_1$  - число различных 1-грамм в  $\Phi_1(T)$ ,  $E_1^1$  - число единичных 1-грамм,  $F_{\max}^{(1)} = \max_{1 \leq i \leq M_1} (F_{i1})$ , где  $F_{i1}$  - частота встречаемости  $i$ -й 1-граммы из  $\Phi_1(T)$ ,  $l_{\max}$  - длина максимального повтора в тексте). Табл. I иллюстрирует динамику изменения этих параметров с ростом  $l$  для всех четырех текстов.

Вместо параметра  $M_1$  в этой таблице указан параметр  $M_1 - E_1^1$ . Он характеризует число кратных 1-грамм в тексте. Зависимость  $M_1 - E_1^1$  от  $l$  представляет собой унимодальную кривую, максимум которой приходится на значения  $l=4+5$  (см. подчеркнутые значения).

Этому факту можно придать следующую трактовку. Если воспринимать интонацию как элементарную семантическую единицу в музыке и положить в основу ее автоматического выделения из музыкального текста свойство повторяемости (по аналогии с морфемой в естественном языке, см. §I), то максимум выразительных средств музыкального языка приходится на диапазон значений  $l = 4+5$  (5+6 нот), кото-

Т а б л и ц а I

Интегральные характеристики спектров (значения  $\times 1 < l_{\max}$  опущены; значения  $l_{\max}$  обведены кружками)

Т е к с т ы												
l	$T^{(1)}(N_1=4758)$			$T^{(2)}(N_2=16905)$			$T^{(3)}(N_3=5126)$			$T^{(4)}(N_4=15455)$		
	$M_1 - E_1^1$	$E_1^1$	$R_{\max}^1$	$M_1 - E_1^1$	$E_1^1$	$R_{\max}^1$	$M_1 - E_1^1$	$E_1^1$	$R_{\max}^1$	$M_1 - E_1^1$	$E_1^1$	$R_{\max}^1$
1	32	I	890	33	I	3163	33	I	1000	o	I	2355
2	294	I22	397	450	I06	I20I	306	I3I	34I	404	III	694
3	694	8I6	I35	I635	I343	37I	745	770	I24	I573	II44	2I3
4	<u>8I8</u>	I932	57	257I	46I8	I60	<u>933</u>	I874	57	263I	3955	94
5	673	2982	I5	<u>26I4</u>	8648	44	837	2900	I6	<u>2748</u>	72I3	29
6	489	3648	7	2054	II794	I3	684	3532	IO	2320	96I6	I6
7	357	399I	4	I497	I3609	6	559	389I	7	I947	IO870	IO
8	277	4I79	4	II07	I4576	4	473	4I02	7	I7I5	II54I	7
9	2I2	43I9	3	88I	I5077	4	408	4250	6	I552	II979	7
(27)	I	4730	2									
(50)				I	I6853	2	I	5076	2			
(II4)										I	I5340	2

рому соответствует максимальное разнообразие повторяющихся l-грамм. Значения  $l = 5 (\pm I)$  можно, по-видимому, считать оценкой средней длины интонации. Слишком короткие интонации не могут претендовать на роль сформировавшейся семантической единицы. Слишком длинные интонации уже не являются элементарными и состоят из более коротких.

Значения  $M_1 (l=1, 2, \dots, l_{\max})$  могут быть получены для каждого текста из табл. I попарным сложением элементов двух первых столбцов, содержащих сведения о параметрах  $M_1 - E_1^1$  и  $E_1^1$ . Величина  $M_1$  характеризует мощность исходного словаря. Поскольку максимальные

значения интервалов в анализированных мелодиях не превышали 9, максимальная мощность исходного словаря была ограничена величиной 40 ( $10 \times 2 \times 2$ ). В достаточно длинных текстах присутствуют почти все элементы словаря ( $M_1^{(i)} \geq 33$  для  $i = 1-4$ ), причем в мелодиях Шу - берта этот показатель достигает максимального значения  $M_1^{(4)} = 37$ ). В каждой же мелодии в отдельности коэффициент использования словаря существенно ниже (в среднем 12-15 элементов).

При  $i = 2$  даже на длинных текстах ( $N \sim 10^4$  нот и выше) реализуется меньше половины потенциально возможных значений словаря биграмм. Из двух компонент интервально-метрической структуры (I и S) существенную роль в ограничении мощности реально используемого словаря играет метрическая структура (см. ниже п.3).

2. Классификационные возможности системы описания можно проиллюстрировать уже на примере частотных характеристик первого порядка (табл.2). Общая закономерность, наблюдаемая для всех четырех массивов мелодий, заключается в том, что по мере возрастания величин интервалов частота использования их падает. Это не удивительно, поскольку человеческому голосу гораздо удобнее "пропевать" последовательно расположенные по высоте звуки, чем воспроизводить скачки. Наиболее распространенным интервалом, как правило, является секунда ( $I = \pm 1$ ), далее следует прима ( $I = 0$ ). Суммарная частота их встречаемости в любом из анализированных текстов превышает 60%, и это является характерной особенностью песенной мелодики в целом.

Текст  $T^{(4)}$  (песни Шуберта) отличается от первых трех текстов тем, что самым частым интервалом в нем является прима ( $\approx 25\%$ ). Соответствующий показатель для песен советских композиторов составляет лишь 14,6%. В связи с этим интересно отметить, что анализ распределения интервалов для 12 вокальных партий из оперы Л.Ян - чека "Средство Макропулоса", проведенный М.Штедром [5], показал, что интервал чистой прима ( $I = 0$ ) составляет 43,5% от общего числа интервалов. Высокий процент использования интервала чистой прима является, по-видимому, характерной особенностью оперной вокальной мелодики. Таким образом, песни Шуберта по своему характеру гораздо ближе к оперной вокальной мелодике, чем современные песни. Частично это объясняется наличием в их составе арий и баллад, а частично отражает различие в интонационных словарях двух рассматриваемых эпох.

Информация о более тонких по сравнению с жанровыми различиях (например, стилевых) содержится в характеристиках более высокого порядка. Тем не менее, некоторые стилевые особенности просматриваются и по характеристикам первого порядка. К примеру, относительная частота структуры 7+- (скачок на октаву вверх, попадающий на

Т а б л и ц а 2

Частотные характеристики первого порядка массивов  $T^{(i)}$   
( $i = 1-4$ , частоты  $F_{j1} \geq 2$ )

Т е к с т ы								
$T^{(1)} (N_1=4758)$		$T^{(2)} (N_2=16905)$		$T^{(3)} (N_3=5126)$		$T^{(4)} (N_4=15455)$		
j	I-гр.	$F_{j1}^{(1)}$	I-гр.	$F_{j1}^{(2)}$	I-гр.	$F_{j1}^{(3)}$	I-гр.	$F_{j1}^{(4)}$
I	2		3		4		5	
1	I→	890	I→	3163	I→	1000	0→	2355
2	I→	708	I→	2211	I→	642	I→	2300
3	0→	498	I→	1746	I→	539	0→	1544
4	I→	406	I→	1589	0→	528	I→	1510
5	I→	380	0→	1548	I→	445	I→	1507
6	2→	276	2→	964	0→	342	I→	1349
7	2→	242	0→	887	2→	250	2→	840
8	0→	201	2→	677	2→	199	2→	622
9	2→	153	3→	567	2→	180	2→	593
10	2→	140	2→	563	3→	126	3→	431
11	3→	136	3→	375	3→	114	3→	402
12	4→	91	2→	289	2→	87	2→	401
13	3→	83	4→	272	5→	86	3→	289
14	3→	70	3→	235	4→	80	4→	173
15	5→	58	3→	233	4→	60	4→	150
16	3→	55	5→	222	3→	60	3→	136
17	4→	54	4→	202	3→	43	5→	119
18	4→	40	4→	167	4→	43	4→	114
19	4→	38	4→	137	4→	42	5→	106

I	2		3		4		5	
20	5++	22	5++	120	5++	38	5++	89
21	7+-	22	5+-	105	5+-	32	4-	89
22	7++	22	7+-	84	6+-	29	7+-	61
23	5+-	20	6+-	75	7+-	20	7++	53
24	6+-	11	7++	58	6++	19	5-	42
25	7+-	9	6++	48	7++	16	6+-	18
26	6+-	8	7+-	45	7+-	11	6++	16
27	6++	6	6+-	35	6-	9	7+-	15
28	5-	5	6-	26	5-	9	6-	13
29	8++	4	7-	19	7-	8	6+	11
30	7-	2	5-	17	6+	7	7-	8
31	8+-	2	8++	4	-	-	9++	6
32	-	-	9+-	3	-	-	9+-	4
33	-	-	8+-	2	-	-	9+-	4
34	-	-	-	-	-	-	8++	3
35	-	-	-	-	-	-	8+-	2
36	-	-	-	-	-	-	8+-	2

сильную долю) в песнях Шуберта в 5-6 раз ниже, чем во всех других массивах. По-видимому, такое различие можно рассматривать как эмоциональную характеристику песен, подчеркивающую широту и удаль, присущую русской песне.

3. Влияние метрической структуры на интервальную подчас прослеживается довольно отчетливо, но иногда носит весьма неожиданный, непредсказуемый характер. Можно отметить две четкие закономерности. Анализ частотных характеристик первого порядка показывает, что понижение высоты звука чаще сопровождается переходом от ударного звука к безударному (понижению высоты соответствует "понижение ударности"). К примеру, для текста  $T^{(2)}$  эта закономерность выполняется для всех значений интервала от  $I = (I-)$  до  $I = (7-)$ .

Другая закономерность отчетливо проявляется при анализе частотных характеристик второго порядка, т.е. интонационных структур вида  $I_1 S_1 I_2 S_2$ . Если проследить за изменением одной лишь метрической структуры, т.е. пары  $S_1 S_2$ , видно, что комбинация "-" встречается гораздо реже остальных (при любых значениях интервала). Для текста  $T^{(2)}$ , к примеру, относительная частота встречаемости комбинации (++) составляет 15,8%, (+-) - 41,2%,

(+) - 40,4% и (-) - 2,6%. Отмеченная закономерность вносит существенный вклад в ограничение словаря биграмм.

Менее предсказуемые случаи иллюстрируют два фрагмента из  $\Phi_2(T^{(2)})$ , где упорядочение произведено по значениям интервала (табл.3).

Первый фрагмент (строки 1÷6 табл.3) содержит пары с фиксированным значением интервала  $I_1 = (2+)$  и непрерывно изменяющимися значениями интервала  $I_2$ . Второй фрагмент (строки (7÷10) построен аналогичным образом, но уже для  $I_1 = (1+)$ . Естественно было бы

Т а б л и ц а 3

№	$I_1 I_2$	$S_1 S_2$			
		++	+-	-+	--
1	2+3+	0	45	5	1
2	2+2+	36	135	61	4
3	2+1+	17	62	35	3
4	2+0+	14	29	52	0
5	2+1-	35	172	277	8
6	2+2-	29	2	38	0
7	1+2+	3	78	87	8
8	1+1+	118	810	283	21
9	1+0+	7	12	262	24
10	1+1-	106	188	396	23

ожидать довольно плавного (хотя и не унимодального) изменения частот по столбцам и отсутствия резких скачков по строкам, за исключением частот, соответствующих значению  $S_1 S_2 = (--)$ . Вместо этого наблюдаются резкие скачки по первым двум столбцам (особенно во втором фрагменте) и слабо предсказуемые изменения значений по отдельным строкам (№ 1,6,7,9), сильно отличающиеся от тех процентов, что приведены выше. Указанные аномалии свидетельствуют о том, что

попытки машинного синтеза мелодий с использованием одной лишь последовательности высот, никак не связанной с метроритмом, малоперспективны.

4. Классификация длинных повторяющихся 1-грамм. Наиболее частые 1-граммы в диапазоне значений  $2 \leq l \leq 6$  образуются, как правило, в результате конкатенации наиболее высокочастотных однограмм (с учетом тех ограничений, которые накладывает метрическая структура). В табл.4 приведены значения первых семи самых частых 1-грамм ( $2 \leq l \leq 7$ ) для текста  $T^{(2)}$ . Поскольку для данного текста при  $l = 1$  три первых места в частотном упорядочении занимают интонационные структуры со значением  $I = (I^\pm)$ , то они же представлены и в наиболее частотных интонационных структурах более высокого порядка. Эти интонации встречаются во многих произведениях разных авторов, поэтому их естественно назвать "бытующими".

Наиболее частые 1-граммы массива мелодий  $\Gamma^{(z)}$  ( $N_z = 16905$ )

1	№	1-грамма	$F_{11}$	1	№	1-грамма	$F_{11}$
2	1	I+I---	1201	5	1	I-I-I+I-I---	44
	2	I++I+-	810		2	I++I+-I-I+I-I---	31
	3	I-I-I+-	794		3	I+-I-I+I-I-I-I---	28
	4	I-I-I+-	539		4	I-I-I-I-I-I-I-I-	27
	5	I+-I+-	396		5	I-I-I-I-I-I-I-I+	24
	6	0++0+-	380		6	I-I-I-I-I-I-I-I-	23
	7	I-I-I+-	359		7	I-I-I-I-I-I-I-I+	23
3	1	I-I-I-I-I+-	371	6	1	I-I-I-I-I-I-I-I-I-	13
	2	I-I-I-I-I-	335		2	I-I-I-I-I-I-I-I-I-	13
	3	I-I-I-I-I-	218		3	I-I-I-I-I-I-I-I-I+	11
	4	I-I-I-I-I+-	167		4	I-I-I-I-I-I-I-I-I+	11
	5	I-I-I-I-I+-	157		5	0++I-I-I-I-I-I-I-I-	11
	6	I-I-I-I-I+-	147		6	I-I-I-I-I-I-I-I-I-	10
	7	I-I-I-I-I+-	145		7	I-I-I-I-I-I-I-I-I+	10
4	1	I-I-I-I-I+-	160	7	1	I-I-I-I-I-I-I-I-I-	6
	2	I-I-I-I-I-I-I-I-	90		2	I-I-I-I-I-I-I-I-I-	6
	3	I-I-I-I-I-I-I-I+	74		3	I-I-I-I-I-I-I-I-I-	6
	4	I-I-I-I-I-I-I-I-	73		4	I-I-I-I-I-I-I-I-I-	5
	5	I-I-I-I-I-I-I-I-	62		5	I-I-I-I-I-I-I-I-I-	5
	6	I-I-I-I-I-I-I-I-	60		6	I-I-I-I-I-I-I-I-I-	5
	7	0++I-I-I-I-I-I-I-	53		7	0++0+-0++I-I-I-I-I-I-	5



С увеличением  $l$  ( $l > 6$ ) в составе наиболее частых интонационных структур начинают появляться среднечастотные, а затем и малочастотные однограммы (соответствующие тенденции уже просматриваются для значения  $l = 7$  в табл.4). Интонации становятся более оригинальными, частота встречаемости их резко падает, и для характеристики факта использования одной и той же интонационной структуры разными композиторами уже больше подходит термин "заимствование" (неважно какое, осознанное или нет). В случае появления характерной интонационной структуры в разных мелодиях одного композитора будем говорить о "самозаимствовании".

Наконец, особого рассмотрения заслуживает случай повторения характерных интонационных структур в одной мелодии. Это основной источник повторяющихся интонационных структур, обнаруживаемых при анализе любого массива музыкальных мелодий. К примеру, число различных кратных 7-грамм, входящих в частотную характеристику 7-го порядка текста  $T^{(2)}$ , равно 1497, т.е. в терминах определения (2)  $M_7(T^{(2)}) - P_7^1(T^{(2)}) = 1497$ , где  $T^{(2)} = T_1 \sqcup T_2 \sqcup \dots \sqcup T_{m_2}$ ,  $m_2$  - число мелодий в массиве  $T^{(2)}$ . Число же различных кратных 7-грамм, каждая из которых является общей хотя бы для пары различных мелодий, равно 368, т.е. в терминах определения (3)  $M_7(T_1, T_2, \dots, T_{m_2}) = 368$ . Поскольку 1-граммы второго типа составляют подмножество 1-грамм первого типа, можно заключить, что свыше 1000 разновидностей кратных 1-грамм, т.е. большая их часть, повторяются (каждая) лишь в пределах какой-то одной мелодии.

В среднем на каждую мелодию массива  $T^{(2)}$  приходится порядка 5 внутренних 7-граммных повторов. Многие из них сцеплены друг с другом и образуют более длинные (иногда очень длинные) повторы (см. значения  $l_{max}$  в табл.1). Таким образом, наличие достаточно длинных повторяющихся IS-структур внутри отдельно взятой мелодии следует считать скорее правилом, чем исключением. Вот, к примеру, как выглядит в терминах повторов сжатая запись мелодии русской народной песни "Всю-то я вселенную проехал" ( $N = 43$ ):

$$T = X(1\rightarrow)Y(2\rightarrow+1\rightarrow)Z(4\rightarrow)X(1\rightarrow)Y(2\rightarrow+3\rightarrow)Z, \quad (4)$$

где  $X = I\rightarrow I\rightarrow I\rightarrow I\rightarrow 7\rightarrow+0\rightarrow I\rightarrow I\rightarrow I\rightarrow I\rightarrow+$ ,  $Y = 3\rightarrow I\rightarrow 3\rightarrow+2\rightarrow+$ ,  $Z = 2\rightarrow+I\rightarrow+2\rightarrow I\rightarrow$  - повторы длины 10, 4 и 4 соответственно (скобками выделены незатронутые перекодировкой элементы мелодии). Заметим, что фрагменты  $X(I\rightarrow)Y$  и  $X(I\rightarrow)Y$  отличаются лишь одним элементом, стоящим в круглых скобках, т.е. образуют  $(1,k)$ -повтор

со значениями параметров  $l = I5$ ,  $k = I$ . Изменение значения метрической структуры при сохранении значения интервала - типичный прием вариации интонации. Отметим также, что наиболее длинный повторяющийся фрагмент X содержит более мелкие внутренние повторы, т.е. в свою очередь может быть представлен в виде:

$$X = W(I-7++0+-) W(I++),$$

где  $W = I \rightarrow I \rightarrow I \rightarrow$ . Наличие внутренних повторов в составе более длинных повторяющихся фрагментов ("вложенность повторов") - характерная структурная особенность песенных мелодий. Довольно часто внутренние повторы следуют непосредственно друг за другом, т.е. образуют периодичности. Фрагменты, содержащие внутри себя периодичности, обычно являются наиболее высокочастотными (см., например, 6-грамму #I в табл. 4).

Примерами "бесповторных" мелодий (точнее, мелодий с очень короткими и малочисленными повторами) являются русские народные песни "По диким степям Забайкалья", "Уж ты, поле мое" [9], песня М. Табащикова на слова И. Френкеля "Давай закурим" [II] и ряд других. Каждая из них содержит не более одного-двух биграммных повторов ( $l_{\max} = 2$ ). Все мелодии такого типа (из сборников [9-II]) довольно короткие ( $N \sim 30$  символов).

Повторы внутри каждой мелодии можно в свою очередь разделить на бытующие и характерные (оригинальные), определяющие своеобразие данной мелодии, позволяющие легко отличить ее от других. Большой интерес представляет задача автоматического выделения из текста мелодии оригинальных интервально-метрических структур. Основой для этого может служить совокупность частотных характеристик  $\phi_1(T), \phi_2(T), \dots, \phi_{l_{\max}}(T)$  мелодии T.

Еще при анализе полных частотных спектров текстов на естественном языке [13] было замечено, что частотные деревья переходов от 1-грамм к их всевозможным (1+1)-, (1+2)- и т.д. расширениям содержат информацию, полезную для решения задачи автоматического выделения из слитного нерасчлененного текста самостоятельных смысловых единиц (например, слов). Особый интерес в этом смысле представляют неразветвляющиеся фрагменты дерева, соответствующие 1-граммам, сохраняющим свою частоту при расширении на один, два и более символов. Факт сохранения частоты при удлинении 1-граммы может интерпретироваться в том смысле, что исходная 1-грамма всегда встречается в фиксированном контексте, т.е. не может рассматри-

ваться как семантически самостоятельная единица. Когда же в процессе удлинения частота 1-граммы резко изменится, это означает, что формирование семантической единицы закончилось. Она встречается в разных контекстах и является тем связующим звеном, которое их объединяет. К примеру, набор словоформ {столы, стола, столом, столовая} может быть представлен в виде дерева следующим образом:

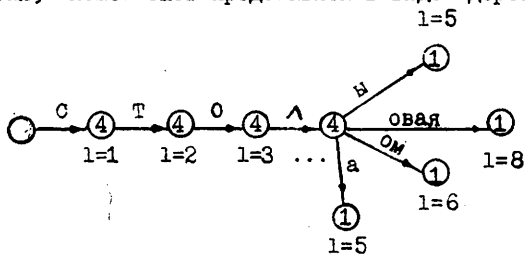


Рис. 2

(рис.2). Здесь пустой узел - корень дерева, каждая ветвь (от корня до листа) соответствует одной из словоформ, в непустых узлах стоят частоты соответствующих 1-грамм. До  $l = 4$  значения частот не меняются ( $F = 4$ ), далее все 1-граммы вырождаются в единичные. В соответствии с вышесказанным 1-граммы "с", "ст", "сто", имеющие частоту 4, еще не могут рассматриваться как семантически самостоятельные единицы, но 4-грамма "стол" - корень рассматривавшихся словоформ - уже претендует на роль таковой.

Приведенные рассуждения легко переносятся и на музыкальные тексты. Проиллюстрируем это на примере песни Соловьева-Седого "Давно мы дома не были". Выпишем необходимые нам фрагменты частотного спектра.

- $l = 1$ ; 1-гр. F      1-гр. F      1-гр. F      1-гр. F      1-гр. F  
           ↓        ↓            ↓        ↓            ↓        ↓            ↓        ↓  
           (I--)-26; (I+-)-24; (2--)-5; (2--)-5; (4+-)-5;  
           (5+-)-4; (I+-)-3; (2+-)-2; (3+-)-2; (0+-)-2;
- $l = 2$ ; (I--4+-)-5; (2--5+-)-5; (5+-I+-)-4; (4+-I--)-3;
- $l = 3$ ; (I+-I--4+-)-5; (5+-I+-I--)-4; (2--2--5+-)-4;  
           (2--5+-I+-)-4; (I--4+-I--)-3;
- $l = 4$ ; (I--I+-I--4+-)-5; (2--5+-I+-I--)-4;  
           (2--2--5+-I+-)-4; (I+-I--4+-I--)-3;
- $l = 5$ ; (I+-I--I+-I--4+-)-5; (2--2--5+-I+-I--)-4;  
           (2+-2--2--5+-I+-)-2;

$l = 6; (I \rightarrow I \rightarrow I \rightarrow I \rightarrow 4++I \rightarrow) - 3; (I \rightarrow I \rightarrow I \rightarrow I \rightarrow I \rightarrow 4++) - 2;$   
 $(2 \rightarrow 2 \rightarrow 5++I \rightarrow I \rightarrow 0++) - 2;$   
 $\vdots$   
 $l = 26; (2 \rightarrow 2 \rightarrow 5++I \rightarrow I \rightarrow I \rightarrow 2++2 \rightarrow 2 \rightarrow 5++I \rightarrow I \rightarrow 0++I \rightarrow I \rightarrow 3+-I \rightarrow$   
 $I \rightarrow I \rightarrow I \rightarrow 4++I \rightarrow I \rightarrow I \rightarrow I \rightarrow) - 2.$

Эти фрагменты отобраны по следующему принципу. Из частотной характеристики первого порядка (она приведена полностью, за исключением трех единичных однограмм) выбираем наименее типичные элементы (обычно это элементы с максимальным по модулю значением интервала). В данном случае таковыми являются  $4++ (F=5)$  и  $5+- (F=4)$ . Далее прослеживаем всевозможные расширения этих однограмм в обе стороны. Однограмма  $(4++)$  расширяется влево на 4 символа с сохранением частоты, затем (при  $l = 6$ ) наблюдается скачок частоты (до  $F = 2$ ). Пятиграмму  $A = (I \rightarrow I \rightarrow I \rightarrow I \rightarrow 4++)$  естественно считать определяющей для данной мелодии. Однограмма  $(5+-)$  расширяется влево и вправо на два символа с сохранением частоты, затем частота изменяется. Естественно и пятиграмму  $B = (2 \rightarrow 2 \rightarrow 5++I \rightarrow I \rightarrow)$  считать определяющей для данной мелодии. Обе они входят в состав самой длинной повторяющейся 1-граммы в данном тексте (см.  $l = 26$ ). По аналогии с (4) ее можно представить в сжатом виде как

$$C = B(1 \rightarrow 1 \rightarrow 2++)B(0++1 \rightarrow 1 \rightarrow 3+-)A(1 \rightarrow 1 \rightarrow 1 \rightarrow 1++) \quad (5)$$

Бытующие интонации в данной мелодии представлены комбинациями наиболее высокочастотных однограмм ( $I \rightarrow$  и  $I \rightarrow$ ). Среди них следует выделить 4-грамму  $D = (I \rightarrow I \rightarrow I \rightarrow I \rightarrow)$ . Соответствующий ей фрагмент дерева представлен на рис. 3.

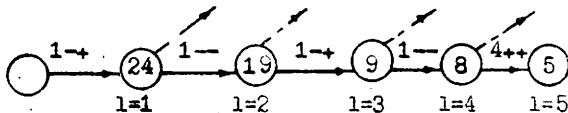


Рис. 3

При переходе от  $l = 3$  к  $l = 4$  частота изменяется всего на 1, что позволяет выделить  $D$  в качестве самостоятельной единицы. Фрагмент  $A$  можно тогда рассматривать как составной:  $A = D(4++)$ . На-

конец, еще одна бытующая интонация ( $E=(I-I+I+-)$ ) выделяется при левостороннем расширении однограммы ( $I+-$ ) (см.рис.4).

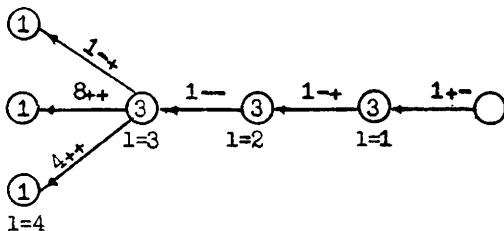


Рис. 4

Всю мелодию можно записать в терминах выделенных интонационных структур следующим образом:

$$T = C(1-4++)C(1+-8++)EAEED(3++1-)\Delta(2-2-+7+-) . \quad (6)$$

Заметим, что в соответствии с изложенной схемой выделения характерных 1-грамм единичные однограммы не могут входить в состав характерных. Места вхождения их в текст разбивают мелодию на отдельные фрагменты, которые ограничивают сверху размер характерных 1-грамм. В представлениях типа (4) и (6) единичные однограммы всегда фигурируют внутри скобок: в (4) - это ( $4+-$ ), в (6) - ( $3++$ ), ( $3++$ ) и ( $7+-$ ).

#### §5. Анализ совместных частотных спектров (эксперименты 2,3)

1. Интегральные характеристики совместных частотных спектров каждого из сборников  $T^{(i)}$  ( $i=1 \div 4$ ) приведены в табл.5. Здесь  $m_1(T_1, T_2, \dots, T_{m_1})$  - число различных 1-грамм, встретившихся как - дая как минимум в двух мелодиях  $i$ -го сборника,  $F_{\max}^{(i)}(T_1, T_2, \dots, T_{m_1})$  - максимум (по всем 1-граммам из  $\Phi_1(T_1, T_2, \dots, T_{m_1})$ ) частот встречаемости общих 1-грамм. Параметр  $L(T_1, T_2, \dots, T_{m_1})^1$  - максимум длин 1-грамм, общих для мелодий  $i$ -го сборника, - извлекается из последней непустой строки, характеризующей  $i$ -й сборник в табл.5. Все значения в табл.5 приведены для  $i \geq 6$ . Этот диапазон значений представляет наибольший интерес в плане анализа заимствований.

Анализ табл.5 показывает, что одинаковые и достаточно длинные интонационные структуры довольно часто встречаются в разных

Т а б л и ц а 5

## Интегральные характеристики совместных частотных спектров

1	Т е к с т ы							
	$T^{(1)}(N_1=4758)$		$T^{(2)}(N_2=16905)$		$T^{(3)}(N_3=5126)$		$T^{(4)}(N_4=15455)$	
	$M_1$	$F_{\max}^{(1)}$	$M_1$	$F_{\max}^{(1)}$	$M_1$	$F_{\max}^{(1)}$	$M_1$	$F_{\max}^{(1)}$
6	I25	7	877	I3	I20	8	504	I6
7	39	5	368	6	39	6	I59	I0
8	I4	3	II4	4	I3	4	45	7
9	7	3	33	4	I	2	I0	7
I0	3	3	I2	3	-	-	5	6
II	-	-	3	2	-	-	2	5
I2	-	-	2	2	-	-	I	4
I3	-	-	I	2	-	-	-	-

мелодиях. При  $F_{11} > 2$  преобладают бытующие интонации, при  $F_{11} = 2$  и  $1 > 7$  — характерные, которые могут быть отнесены к категории заммствований.

Среди бытующих интонаций можно выделить крупные классы гаммаобразных и речитативных интонаций.

Гаммаобразные интонации характеризуются достаточно длинными последовательностями положительных либо отрицательных единичных интервалов. Примером могут служить интонационно-метрические структуры 7-го порядка (I—I—II—I—II—II) и (I—II—II—II—II—II), встретившиеся каждая в 10 различных мелодиях сборника [I0], структура (I—II—II—II—II—II), вошедшая в 6 мелодий, и т.д.

Речитативные интонации характеризуются преобладанием интервалов с нулевыми значениями. Особенно характерны такие интонации для мелодий Шуберта. Так, IS — структура 7-го порядка (0+0+0+0+0+0+0++) встретилась (причем неоднократно) в 5 мелодиях сборника [I2] и в 3 мелодиях сборника [II]; структура (0++0+0+0+0+0+I+) встретилась 5 раз в [I2] и 4 раза в [I0].

В приложении 2 приведены все связи (типа общих IS-структур 9-го порядка и выше), существующие между мелодиями сборника [I0].

1=13

Да здрав-ству-ет соз-дан-ный во-лей на-ро-дов е-ди-ный  
И де-рет-ся Ча-пай пар-ти-зан-ский ге-рой  
За сво-бод

№ I

№ 51

1=10

нам вру-чил пу-тев-ки ком-со-моль-ский ко-ми-тет  
судь-ба за-бро-сит нас да-ле-ко, пус-кай !

№ I29

№ I91

1=115

Пус-кай по-ет о нас стра-на и звон-кой пес-не-ю  
то-го цве - ту ког-да яб-ло-ня (цветет)

№ 200

1=9

ой-мом пре-крас-ну-ю пе-сню сла  
си-ма-я за-ря мо-я ве-чер-ня я

№ 6

№ I99

1=10

Стра-на вста-ет со сла-во-ю на встре-чу дня

№ III

"С горки на гору ходила" (частушка)

бо-лят. Чер-но-бро-во-го лю-би-ла Чер-но

Рис. 5. Примеры заимствований на материале сборников [9] и [10].

С помощью приложения I идентифицируются названия мелодий, содержащих одинаковые структуры, и фамилии композиторов. Нотная запись наиболее интересных заимствований приведена на рис.5. Отметим, что граф связей мелодий содержит относительно небольшое число изолированных подграфов. Большая же часть мелодий объединяется в один связный подграф, для любой пары вершин которого может быть найден соединяющий их кратчайший путь.

Явной корреляции между длиной мелодии и числом входящих в нее интонационных структур, принадлежащих совместному частотному спектру, не обнаружено. Существуют длинные мелодии, не содержащие заимствований (например, №126 из [10]), и, наоборот, короткие мелодии с относительно большим числом заимствований (№33 из [10]).

Распределение мест вхождений общих 1-грамм по длине мелодии в целом по всем мелодиям сборника [10] близко к равномерному. На рис.6 приведены соответствующие гистограммы для общих 1S-структур

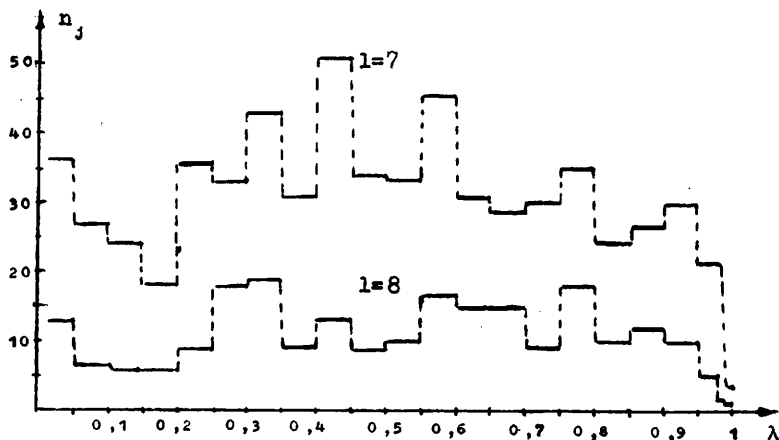


Рис. 6

8-го и 9-го порядков ( $l = 7$  и  $8$  соответственно). Место вхождения в мелодию  $T_k$  длиной  $N_k$  ( $1 \leq k \leq 216$ ) 1-граммы, начинающейся в  $i$ -й позиции, характеризуется значением параметра  $\lambda = i/(N_k - 1 + 1)$ , откладываемым по оси абсцисс. Возможный диапазон изменения значений  $\lambda$  ( $0 < \lambda \leq 1$ ) разбит на интервалы шириной  $\Delta = 0,05$  и  $s_j$ -м интервалом ( $1 \leq j \leq 20$ ) связано значение  $n_j$  - суммарное (по всем мелодиям сборника) количество общих 1-грамм со значением  $\lambda$ , попадающим в указанный интервал.



Анализ распределения мест вхождений общих 1-грамм ( $1 \geq 7$ ) в конкретные мелодии выявляет две основные схемы их расположения: наличие кластера в ограниченном районе (рис.7,а) либо равномерное размещение по длине мелодии (рис.7,б). Кластеризация характеризует, как правило, бытующие интонации.

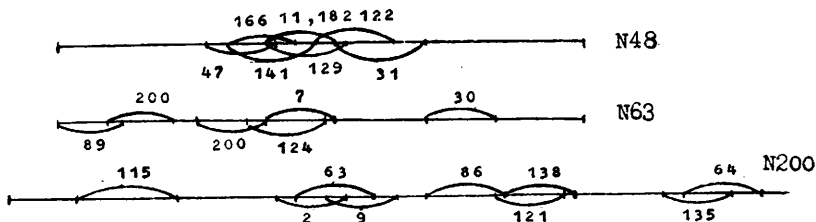


Рис. 7

2. Многократные заимствования. Нередки случаи, когда две мелодии имеют более чем один общий фрагмент достаточно большой длины ( $1 \geq 6$ ). Анализ многократных связей позволяет типизировать их следующим образом.

а) Мелодии связаны лишь одной 1-граммой, но она входит в каждую из них многократно. Например, русские народные песни "В темном лесе" и "На горе-то калина" [9] имеют общую 6-грамму ( $2++I--I--+I--0++0+-$ ), которая входит трижды в первую песню (позиции 25, 31, 37) и дважды во вторую (позиции 3, 12). Чаще всего такие связующие 1-граммы являются бытующими, и наличие их еще не определяет сходства мелодий в целом. В частности, в данном случае структурные представления обеих мелодий (аналогичные (4), (6)) сильно отличаются.

б) Мелодии связаны лишь одной 1-граммой, но в одну из них она входит многократно, а в другую - лишь один раз. Часто встречающейся разновидностью этой схемы является ситуация, когда (1+1)-грамма из одной мелодии представлена во второй двумя разнесенными 1-граммами, выделяемыми из (1+1)-граммы. Например, в мелодии № 12 из [10], начиная с 9-й позиции, расположена 7-грамма ( $0++3+-I++I+-0++I--I+-$ ), а с 18-й позиции - 7-грамма ( $3+-I++I+-0++I--I+-I--$ ), которые в мелодии № 21 представлены в виде 8-граммы ( $0++3+-I++I+-0++I--I+-I--$ ) - 30-я позиция. Наличие заимствований типа "б" также обычно не определяет сходства мелодий в целом, даже если заимствуется характерная (оригинальная) 1-грамма.

в) Мелодии связаны двумя или более независимыми 1-граммами. Если они бытующие и места их вхождения сильно отличаются, это, как правило, не приводит к ощущению сходства между мелодиями. Если связующие 1-граммы характерные, можно предположить, что одна мелодия сочинена под влиянием другой. Однако при слуховом восприятии сходство между мелодиями не всегда может быть уловлено, что происходит при изменении порядка следования характерных 1-грамм в обеих мелодиях или при сильном несоответствии их позиций (при правильном порядке). Например, мелодии №88 и №94 из [10] имеют общие 7-граммы А = (1--1--+1--3--+3--+1++1++) и В=(2--+5--+1--+1--1--+1--1++), а также 6-грамму В' = (2--+5--+1--+1--1--+1--), являющуюся подсловом слова В. Порядок же следования этих 1-грамм различен: ...А...В... ..В'... - в мелодии 88 и ...В...В'...А... - в мелодии 94.

При сохранении порядка и не слишком различающихся позициях общих характерных 1-грамм сходство между мелодиями воспринимается и на слух. Иллюстрацией могут служить мелодии №1 и №51 из [10]. Кроме общей 13-граммы, приведенной на рис. 5 и расположенной в близких позициях в обеих мелодиях (24 и 23 соответственно), имеется (7,2)-повтор (позиции 3 и 1 соответственно), общая 4-грамма (0++7+-1--+1++) - позиции 41 и 43, (7,1)-повтор (позиции 51 и 53). В целом начальные участки обеих мелодий длиной в 57 и 59 символов (при длинах самих мелодий, равных 97 и 62 символам соответственно) имеют общую подпоследовательность из 35 символов.

3. Самозаимствования, как это видно из приложения 2, довольно широко распространены в практике отдельных композиторов (Соловьев-Седой, Блантер, Листов, Захаров, Дунаевский). Примеры наиболее интересных самозаимствований приведены на рис.8.

Самозаимствования можно также делить на характерные и бытующие, многократные и однократные. Отметим, что далеко не всегда заимствование оригинальной интонационной структуры из популярной мелодии обеспечивает популярность новой мелодии. Еще в большей степени это относится к самозаимствованию бытующих интонаций, поэтому наличие их носит, скорее всего, неосознанный характер и отражает особенности стиля данного композитора.

При наличии многократных самозаимствований следует различать две ситуации: заимствование из разных мелодий и из одной и той же. Примером первой ситуации является песня Дунаевского "Эх, хорошо", в которой использованы четыре 7-граммных фрагмента из других произведений этого же автора, вошедших в [10]. В определенной

№ 157  
1=10  
под со-бст-веч-ну-ю му-зы-ку ша-га-ет (тракторист)

№ 159  
х ле-ба сто-ят вы - со-ки-е, а в серд-це

№ 118  
1=9  
(чисты) е у - па-ли на тра - ву

№ 133  
(соко)-лы и день и ночь па рят

Рис. 8

степени такого рода заимствования также характеризуют особенности стиля композитора, быть может, им не осознаваемые. Вторую ситуацию более естественно охарактеризовать как осознанное самозаимствование.

#### §6. Совместный частотный спектр песен советских композиторов и русских народных песен (эксперимент 3)

Существует мнение, что многие композиторы черпают темы для своих произведений в народном творчестве. В лучших произведениях композиторов - классиков часто ощущается заметное влияние народных мелодий. С этой точки зрения и представляла интерес совместная обработка сборников [9,10]. Аналогичная обработка сборников [10] и [12] дает возможность сопоставления интонационных словарей не только разных эпох, но и разных народов.

Результаты совместной обработки сборников [9] и [10] представлены в табл.6 в виде списка композиторов, упорядоченного по частоте встречаемости в их мелодиях интонационных структур, присущих народному песенному творчеству. Обозначим параметр, по ко-

тому ведется упорядочение, через  $\rho_1(t, T)$ , где  $t$  - выборка мелодий конкретного композитора, а  $T$  - выборка мелодий, с которыми сопоставляются мелодии данного композитора путем вычисления  $\Phi_1(t, T)$ . В качестве  $T$  может выступать любой из рассматривавшихся сборников, из которого исключены мелодии, вошедшие в  $t$ . Тогда  $\rho_1(t, T)$  определяется как отношение суммы длин всех фрагментов из  $t$ , представленных в  $\Phi_1(t, T)$ , к сумме длин мелодий из  $t$ . Участки мелодий, соответствующие наложению нескольких фрагментов, учитываются однократно. Иными словами,  $\rho_1(t, T)$  - относительная величина покрываемости текстов из  $t$  фрагментами из  $\Phi_1(t, T)$ .

В табл.6 упорядочение проведено по убыванию значений  $\rho_1(t, T^{(1)})$ . Для сопоставления приведены соответствующие тем же композиторам значения  $\rho_1(t, \bar{T}^{(2)})$ , где  $\bar{T}^{(2)}$  - текст  $T^{(2)}$  без вошедших в  $t$  мелодий. Поскольку абсолютные значения  $\rho_1(t, T^{(1)})$  и  $\rho_1(t, \bar{T}^{(2)})$  сильно отличаются из-за различия в объеме выборки, больший интерес представляет сопоставление рангов  $r_1(t, T^{(1)})$  и  $r_1(t, \bar{T}^{(2)})$  - порядковых номеров в упорядочении. Символ "↑" означает существенное повышение ранга при переходе от упорядочения по  $\rho_1(t, \bar{T}^{(2)})$  к упорядочению по  $\rho_1(t, T^{(1)})$ , символ "↓" - существенное понижение. Отсутствие пометки означает близость рангов по обоим упорядочениям.

Таким образом, наличие символа "↑" показывает, что в произведениях данного композитора в большей степени проявляется тенденция к использованию элементов (интонационных структур), присущих народному творчеству, чем элементов, находящихся "на слуху" нации в данный период. Соответственно, наличие символа "↓" означает, что данный композитор в большей степени склонен к использованию в своих произведениях "современных" интонационных структур, бытующих в музыкальной практике нынешнего дня. При трактовке абсолютных значений коэффициента  $\rho_1(t, \bar{T}^{(2)})$  следует учитывать, что вклад в него вносят как фрагменты, заимствованные (пусть, неосознанно) самим композитором, так и заимствуемые у него.

Общий вывод, который можно сделать по данному эксперименту, состоит в том, что в целом интонационные структуры, которые находятся "на слуху" нации в данный момент, используются композиторами чаще, чем те, что бытовали в прошлом. Некоторое представление о количественном различии дало вычисление совместного частотного спектра первых 135 мелодий сборника [10] с оставшимися, а так-

Т а б л и ц а 6

Список композиторов, представленных в сборнике [10],  
упорядоченный по степени встречаемости в их произведе-  
ниях интонационных структур из [9]

$r_1(t, T^{(1)})$	Композитор	Число мело- дий компо- зит.	$\rho_1(t, T^{(1)})$	$\rho_1(t, T^{(2)})$	$r_1(t, T^{(2)})$
1	Сорокин (†)	1	0,207	0,17	21
2	Акуленко (†)	1	0,189	0,095	32
3	Бабаджанян и Молчанов	1	0,186	0,355	3
4	Шостакович	4	0,136	0,353	4
5	Компанеец	3	0,112	0,242	8
6	Давиденко	2	0,104	0,222	12
7	Жарковский (†)	1	0,100	0,2	18
8	Хренников (†)	6	0,094	0,165	23
9	А.Александров	7	0,088	0,327	7
10	Новиков	17	0,078	0,221	13
11	Соловьев-Седой	21	0,076	0,219	14
12	Фрадкин (†)	5	0,076	0,079	33
13	Лепин (†)	1	0,072	0,258	7
14	Б.Александров( †)	1	0,071	0,337	5
15	Носов (†)	6	0,066	0,125	31
16	Блантер (†)	18	0,048	0,239	10
17	Туликов (†)	10	0,047	0,148	25
18	Мокроусов (†)	13	0,046	0,134	30
19	Макаров	11	0,040	0,205	17
20	Богословский( †)	5	0,038	0,240	9
21	Мурадели	8	0,036	0,183	19
22	Милютин (†)	6	0,034	0,214	15
23	Захаров	11	0,033	0,139	27
24	Дунаевский (†)	27	0,030	0,229	11
25	Покрасс	5	0,028	0,138	28
26	Листов	4	0,025	0,160	24
27	Кац	5	0,018	0,166	22

же этих же 135 мелодий с русскими народными песнями [9]. В обоих случаях добавлялись близкие по объему тексты. В первом случае из 135 мелодий 48 не оказались связанными с мелодиями второго текста на уровне общих 7-грамм; во втором случае таких мелодий было уже 79.

В целом же лишь 31 мелодия из 216, содержащихся в [10], не имеет связей длины 7 (и выше) с другими мелодиями этого сборника. Совместная обработка сборников [9] и [10] сокращает число мелодий до 28, т.е. очень незначительно. Совместная обработка [10] и [12] устраняет еще 5 мелодий из этого списка. В итоге лишь 23 мелодии из 216 не имеют связей ни с одной из мелодий сборников [9, 10, 12] (их номера в соответствии с приложением I: 17, 19, 23, 28, 37, 53, 77, 81, 98, 103, 107, 113, 126, 137, 155, 168, 170, 173, 177, 185, 197, 203, 205 - всего 15 композиторов из 33).

Подробный анализ этих мелодий не входит в цели данной статьи, но представляет несомненный интерес в плане высказываемого иногда тезиса о том, что композиторы-песенники не могут не использовать бытующие интонации, если хотят сочинить мелодию, которая воспринималась бы как благозвучная. Отметим в связи с этим, что среди перечисленных мелодий есть довольно популярные, например, такие как "Моя любимая" М.Блантера, "Смуглянка" А.Новикова, "Вечер на реке" В.Соловьева-Седого и ряд других. Их анализ, возможно, будет способствовать выявлению принципов, обеспечивающих создание достаточно популярных, но сохраняющих свою индивидуальность мелодий.

### З а к л ю ч е н и е

Предложен метод анализа музыкальных произведений (мелодий песен), в основу которого положены, с одной стороны, понятия интонации и интонационно-метрической структуры как носителя сходных интонаций, с другой стороны, - понятия частотного и совместного частотного спектра музыкальных текстов как аналога интонационной памяти композитора. Метод имеет развитую программную поддержку в виде пакета прикладных программ "СИМВОЛ" и ориентирован на анализ больших массивов музыкальных данных. С его помощью могут решаться разнообразные задачи классификации музыкальных произведений (по жанру, стилю, времени создания), автоматического выделения наиболее существенных структурных элементов мелодии, сжатия музыкальных текстов и ряд других. Возможности метода проиллюстрированы на результатах обработки четырех сборников песен.

## Л и т е р а т у р а

1. ЗАРИПОВ Р.Х. Машинный поиск вариантов при моделировании творческого процесса. -М.: Наука, 1983. - 232 с.
2. KNUTH D.E. The Complexity of Songs.- Communications of the ACM, 1984, v.27, N 4, p.344-346.
3. Analysis and Compacting of Musical Texts /A.Bertoni, G.Hanus, G.Mauri, M.Torelli. - Journal of Cybernetics, 1978, N 8, p.257-272.
4. WAGNER R.A., FISCHER M.I. The string-to-string correction problem. - JACM, 1974, v.21, N 1, p.168-173.
5. ЗАРИПОВ Р.Х. Построение частотных словарей музыкальных интонаций для анализа и моделирования мелодий. -В кн.: Проблемы кибернетики. Вып. 41, 1984, с.207-252.
6. ГУСЕВ В.Д. Характеристики символьных последовательностей. -В кн.: Машинные методы обнаружения закономерностей (Вычислительные системы, вып. 88). Новосибирск, 1981, с.112-123.
7. СУХОТИН Б.В. Оптимизационные методы исследования языка. - Автореф. дис. на соиск. учен. степени доктора филолог.наук. М., 1979 (МГУ).
8. ГУСЕВ В.Д., КОСАРЕВ Ю.Г., ТИМОФЕЕВА М.К. и др. Пакет прикладных программ для анализа произвольных символьных последовательностей значительной длины (СИМВОЛ). -В кн.: Структурный анализ символьных последовательностей (Вычислительные системы, вып. 101). Новосибирск, 1984, с.3-21.
9. Русские народные песни. -М.: Музыка, 1985.
10. Русские песни. Выпуск 3. Песни советских композиторов. -Л.: Музгиз, 1965.
11. День победы. Песни великого подвига. -М.: Худ.литература. -Музыка, 1985.
12. ШУБЕРТ. Избранные песни, т.1,6. - Л.: Музыка, 1965.
13. ГУСЕВ В.Д., КОСАРЕВ Ю.Г., ТИТКОВА Т.Н. Методы поиска и анализ статистических закономерностей в символьных последовательностях. -В кн.: Машинные методы обнаружения закономерностей (Материалы Всесоюзного симпозиума). Новосибирск, 1976, с.75-84.
14. ТИМОФЕЕВА М.К. Индуктивная реконструкция грамматик флективных языков. -В кн.: Методы обнаружения закономерностей с помощью ЭВМ (Вычислительные системы, вып. 91). Новосибирск, 1981, с. 57-67.
15. КНУТ Д. Искусство программирования для ЭВМ, т.3. -М.: Мир, 1976. - 845 с.
16. ГУСЕВ В.Д., ТИТКОВА Т.Н. Алгоритмы вычисления совместного частотного спектра двух текстов. -В кн.: Структурный анализ символьных последовательностей (Вычислительные системы, вып. 101). Новосибирск, 1984, с.22-34.

Поступила в ред.-изд.отд.

12 декабря 1985 года

Приложение I

Русские песни. Выпуск 3. Песни советских композиторов  
Л., Музгиз, 1955

	Длина
1. Александров А. Гимн Сов.Союза	97
2. Мурадели М. Партия - наш рулевой	132
3. Новиков А. Есть такая партия	78
4. Новиков А. Партия, слушай, родная!	59
5. Александров А. Святое Ленинское знамя	86
6. Александров А. Кантата о Сталине	63
7. Дунаевский И. Песня о Родине	74
8. Александров Б. Да здравствует наша держава!	98
9. Новиков А. Родина моя	81
10. Новиков А. Россия	88
11. Компанеец З. Моя страна	66
12. Компанеец З. Всецел теперь сильны мы и богаты	74
13. Туликов С. Сов.Россия	51
14. Мокроусов Б. Песня о родной земле	80
15. Макаров В. Родимый край	58
16. Мурадели В. Гимн Москве	67
17. Новиков А. Слава Новой Москве!	95
18. Новиков А. Хороша ты, Москва	93
19. Маслов Ф. Колхозная песня о Москве	36
20. Туликов С. Город мира	102
21. Дм. и Дан.Покрасс.Москва майская	91
22. Ленин А. Здравствуй, Москва!	97
23. Дунаевский И. Моя Москва	78
24. Кац С. Здравствуй, столица!	90
25. Соловьев-Седой В. Наш город	75
26. Мокроусов Б. Песня о Сталинграде	79
27. Дунаевский И. Песня о Волге	69
28. Мокроусов Б. Песня о Волге	90
29. Туликов С. Волга	100
30. Макаров В. Солнечные плесы	77
31. Макаров В. Огоньки	66
32. Дунаевский И. Дорожная песня	117
33. Соловьев-Седой В. Играй, мой баян	59
34. Блантер М. Под звездами болгарскими	40



	Длина
35. Мокроусов Б. Лучше Родины на свете не найдешь!	47
36. Блантер М. Летят перелетные птицы	68
37. К.Листов Ходили мы походами	94
38. Шостакович Т. Тоска по родине	84
39. Хренников Т. Песня о Москве	114
40. Блантер М. Московский вальс	113
41. Блантер М. Московский марш	83
42. Носов Г. Песня о Ленинграде	90
43. Туликов С. Курский соловей	75
44. Соловьев-Седой В. У родного Иртыша	60
45. Давиденко А. Первая конная	86
46. Давиденко А. Конная Буденного	49
47. Александров А. Песня о Ворошилове	51
48. Александров А. Забайкальская	53
49. Дзержинский И. От края до края	54
50. Покрасс Дм. и Дан. Конармейская	45
51. Новиков А. Песня о Чапаеве	62
52. Н.Леви Чапаевская	32
53. Блантер М. Песня о Щорсе	28
54. Блантер М. Партизан Железняк	39
55. Дунаевский И. Песня о Каховке	119
56. Белый В. Орленок	62
57. Прицкер Д. Песня о баяне	76
58. Книшпер Л. Полшко-поле	34
59. Дзержинский И. Казачья песня	62
60. Кац С. Как у дуба старого	50
61. Покрасс Дм. и Дан. То не тучи - грозовые облака	43
62. Соловьев-Седой В. Казачья кавалерийская	45
63. Блантер М. В путь -дорожку дальнюю	53
64. Листов К. Песня о тачанке	62
65. Покрасс Дм. и Дан. Три танкиста	37
66. Блантер М. С нами поет вся страна	78
67. А.Александров Священная война	57
68. Компанеев З. В бой за Родину!	83
69. Белый В. Песня смелых	62
70. Блантер М. До свидания, города и хаты	43
71. Соловьев-Седой В. Комсомольская песня	117
72. Хренников Т. Песня артиллеристов	96

	длина
73. Новиков А. Самовары-самопалы	70
74. Милотин Ю. Морская гвардия	81
75. Захаров В. Ой, туманы мои	71
76. Кац С. Шумел сурово брянский лес	61
77. Фрадкин М. Песня о Днепре	54
78. Соловьев-Седой В. Песня о краснодонцах	31
79. Сорокин В. Матросов - наш однополчанин	87
80. Прицкер Д. Александр Чекалин	50
81. Мокроусов Б. Заветный камень	81
82. Александров А. Песня о Сов.Армии	94
83. Блантер М. Солнце скрылось за горой	42
84. Фрадкин М. Золотился закат	100
85. Дунаевский И. Ехал я из Берлина	76
86. Новиков А. Дороги	47
87. Макаров В. Широки поля под Сталинградом	46
88. Соловьев-Седой В. Азовская партизанская	74
89. Соловьев-Седой В. Гвардейская походная	81
90. Мокроусов Б. Мы люди большого полета	84
91. Шостакович Д. Родина слышит	63
92. Акуленко П. Песня молодых солдат	66
93. Соловьев-Седой В. Марш нахимовцев	79
94. Туликов С. На страже Родины	85
95. Новиков А. Гимн демократ. молодежи мира	125
96. Мурадели В. Гимн междунар. союза студентов	95
97. Белый В. В защиту мира	74
98. Туликов С. Мы за мир!	86
99. Шостакович Д. Песня мира	76
100. Захаров В. Наша сила в деле правом!	65
101. Мурадели В. Песня борцов за мир	67
102. Мурадели В. Москва-Пекин	103
103. Носов Р. За мир и свободу	78
104. Туликов С. Марш сов. молодежи	111
105. Левитин Ю. Марш молодых защитников мира	96
106. Кац С. Дай руку, товарищ далекий	88
107. Мурадели В. Песня молодежи	58
108. Мурадели В. Молодость, здравствуй!	97
109. Дунаевский И. Летите, голуби	73
110. Дунаевский И. Марш энтузиастов	115

	Длина
II1. Шостакович Д. Песня о встречном	57
II2. Новиков А. Мы славим труд	II3
II3. Макаров В. Как у Волги, у реки	49
II4. Дунаевский И. Марш трактористов	59
II5. Дунаевский И. Идем, идем, веселые подруги!	83
II6. Захаров В. Пройдут года	85
II7. Богословский А. Спят курганы темные	51
II8. Захаров В. Дороженька	52
II9. Захаров В. Вдоль деревни	55
I20. Дунаевский И. По реке, реке Кубани	97
I21. Макаров В. Ты в хлебах, земля!	82
I22. Мокроусов Б. Зацветает степь лесами	63
I23. Туликов С. Лес - богатый	67
I24. Новиков А. Хороша земля	98
I25. Хренников Т. Как в степи Кубанской	46
I26. Туликов С. Приезжали на побывку	I40
I27. Милютин Ю. Посреди полей России	79
I28. Туликов С. Цвети, наш край	91
I29. Мурадели В. Едем мы, друзья!	43
I30. Бабаджанян А. и Молчанов К. Комсомольская прощальная	II3
I31. Коваль М. Юность	78
I32. Макаров В. Золотые зори	60
I33. Захаров В. Зелеными просторами	71
I34. Дунаевский И. Марш веселых ребят	86
I35. Новиков А. Вперед, молодежь!	48
I36. Островский А. Комсомолы	71
I37. Дунаевский И. Наша молодость идет	59
I38. Дунаевский И. Молодежная	95
I39. Новиков А. Прибавить надо шаг!	I05
I40. Дунаевский И. Эх, хорошо!	II7
I41. Дунаевский И. Спой нам, ветер	I27
I42. Соловьев-Седой В. Студенческая попутная	I56
I43. Хренников Т. Студенческая песня	83
I44. Хренников Т. Поезд идет все быстрее	61
I45. Мокроусов Б. В добрый час!	82
I46. Милютин Ю. Провожают гармониста	50
I47. Милютин Ю. Ленинские горы	88

	Длина
148. Дунаевский И. Школьный вальс	103
149. Богословский Н. Где ты, утро раннее ?	71
150. Соловьев-Седой В. Грустная песенка	50
151. Дунаевский И. Физкультурный марш	131
152. Дунаевский И. Спортивный марш	115
153. Блантер М. Марш физкультурников	78
154. Дунаевский И. Крали беззаветно отчизну свою	119
155. Фрадкин М. Подмосковная песня	60
156. Фрадкин М. Вернулся я на родину	90
157. Соловьев-Седой В. Поет гармонь за Вологдой	95
158. Макаров В. Хорошо весной на Волге	76
159. Соловьев-Седой В. Моя родная сторона	69
160. Макаров В. Хороши вы, Жигули !	56
161. Мокроусов В. Моей ночью	80
162. Новиков А. Хороши колхозные покосы	81
163. Блантер М. Ишеница золотая	74
164. Блантер М. Колыбельная	71
165. Блантер М. Катюша	58
166. Носов Г. Парень кудрявый	76
167. Захаров В. Шел со службы пограничник	74
168. Милютин Ю. Чайка	51
169. Пушкин С. Лейся, песня, на просторе	44
170. Дм. и Дан. Покрасс Прощание	30
171. Богословский Н. Я на подвиг тебя провожала	102
172. Богословский Н. Любимый город	71
173. Соловьев-Седой В. Вечер на рейде	104
174. Жарковский Е. Прощайте, скалистые горы	70
175. Блантер М. В лесу прифронтовом	169
176. Соловьев-Седой В. Соловьи	117
177. Блантер М. Моя любимая	29
178. Богословский Н. Лизавета	71
179. Листов К. В землянке	71
180. Дунаевский И. Далеко, далеко за снегами	121
181. Кац С. Сирень цветет	157
182. Новиков А. Вася-Василек	114
183. Соловьев-Седой В. Как за Камой, за рекой	106
184. Соловьев-Седой В. На солнечной поляночке	93
185. Новиков А. Смуглянка	110
186. Захаров В. Чем же, чем я виновата	46

	Длина
187. Соловьев-Седой В. Где же вы теперь, друзья-однополчане?	83
188. Носов Г. Далеко-далеко	103
189. Носов Г. За горами за Карпатскими	107
190. Носов Г. Плясовая	121
191. Соловьев-Седой В. Пора в путь-дорогу	131
192. Захаров В. И кто его знает	62
193. Захаров В. Провожанье	67
194. Дунаевский И. Ох ты, сердце	63
195. Дунаевский И. Если Волга разольется	132
196. Дунаевский И. Дорогой широкой	53
197. Нушков В. Тайга золотая	59
198. Хренников Т. Несня Глаши	120
199. Соловьев-Седой В. Улынькь моя, хорошая	48
200. Блантер М. Лучше нету того цвету	94
201. Мокроусов В. Хороши весной в саду цветочки	65
202. Милютин Ю. Сирень-черемуха	67
203. Фрадкин М. Ходит по полю девчонка	65
204. Дунаевский И. Под луной золотой	47
205. Соловьев-Седой В. На лодке	98
206. Макаров В. Рос на опушке рощи клен	73
207. Листов К. Пой, моя хорошая !	49
208. Макаров В. Мерцают звезды ранние	100
209. Мокроусов Б. Сормовская лирическая	45
210. Дунаевский И. Ой, цветет калина	59
211. Дунаевский И. Каким ты был	59
212. Блантер М. На горе, белым-бела	54
213. Будашкин Н. За дальнею околицей	89
214. Мокроусов Б. Одинокая гармонь	56
215. Захаров В. Ходят двое	57
216. Мокроусов Б. Застольная	70

Приложение 2. Список повторяющихся интервально-метрических структур девятого (и более высокого) порядков для массива песен советских композиторов (см. примечания к списку и приложение 1)

№ мелодии	№ мелодий, имеющих общую 1-грамму с мелодией, указан. в 1 столбце	2	3	час-тота		длина
				1-граммы	1-граммы	
1		2	3	4	5	
1	51, 209		(51) 5+-I--I+-2+-2+3+-I--I+-4--0++3+- (209) 4--5+-I--I+-1+-2+-2+3+-	2	2	13
4	13		(13) I++I+-2++I--I++I+-I++I+-	2	2	8
6	6, 67, 84, 199		(8) I++I+-I--I--I--I++I++I+- (67) I--I--I--I--I--I--I--I--2++ (84) I--I--I--I--I--I--I--I-- (199) I--I--2++I+-I+-2+-I--I--I--	2	2	8
7	41		(41) I++I+-0++0+-I--I--I++I+-	2	2	8
8	6, 96		(96) I++I+-2+-2++2+-I--I--I--	3	3	8
12	148, 191		(148) 0++I--I--I--I--I--I-- (191) 3--3+-I++I++0++I--I--I--	2	2	8
13	4			2	2	8
14	120, 194, 214		(120) I++I+-I++I+-I--I++I+-0++ (194) I--I++I+-I--I++I+-0++I-- (214) I+-0++I+-I--I++I+-I--	2	2	8
16	64, 67, 68		(64) 2+-2--4+I--I--I--I--I-- (67) I--2--2--4+I--I--I--I-- (68) I--I--2++I--I--I--I--I--	2	2	8
18	140		(140) 0++I+-0++I++I+-I--I--	2	2	8
22	41, 112		(41) 4+-0++0+-0+I+-I--I--I-- (112) I--I--2--3--0++0+-3+3+-	2	2	8

1	2	3	4	5
31	48, 55, 176	(48) 5+-I+-I+-I+-2-+3+-I-+3-- (55) I+-I+-I+-I+-I+-I+-I+-I+- (176) 2-+I+-I+-I+-2-+I--I+-	2 2 3	8 9 8
32	38, 175, 162	(38) 0++5-0++I+-I+-I+-I+-I-- (175) I+-I+-I+-I+-I+-I+-3+- (182) 2++I+-I+-I+-I+-I+-I--	2 2 2 2	8 8 8 8
33	210	(210) I+-I--I+-I+-I--I+-3+-I+-	2	8
35	161	(161) I+-I--2++I--I+-I--2++I--	3	8
38	32, 55, 72	(55) I+-I+-I+-I+-I--2-+I+-I+- (72) 2++2-I+-I--I+-I+-2-+0++	2 2	8 8
40	85	(85) I+-0+0+-I+-I+-I+-2--2--	2	8
41	22			
42	74	(74) I+-I--I+-I--3++I--I--I--I+-	2	9
45	162, 172	(152) 0++0++I+-I--I--I--I--I+- (172) I+-I+-I--3-+4+-I+-I+-I+-	2 2	8 8
47	183	(183) I--4-0+0+-0-+0++0+-I+-	2	8
48	31, 141	(141) I+-I+-I+-I+-I--I--I+-I+-	3	9
49	121	(121) I+-2-+I+-I+-I--2++I--I+-	2	8
51	I, 108, 141	(108) I+-I+3+-I+-I+-I--2-+I+-I+- (141) I+-I+-0++0+3+-I+-I+-I+-	2 2	9 8
52	165	(165) I--I+-I+-I--0++I--I+-I+-4-+5+-	3	10
54	104, 171	(104, 171) 2--0++I+-I+-I+-I--3+-	3	8
55	31, 38, 135	(135) I+-I+-I+-I+-I+-2-+0++0+-	2	8
59	121	(121) I+-I+-I+-2+-I+-I--I+-I+-	2	8

I	2	3	4	5
62	195	(195) I--I--I--2++3--2++I--I--+	2	8
63	124, <u>200</u>	(124) I++I++0++I++I++0++0++I--+	2	8
		(200) I--I--I--I--0++0++I++I--	2	8
64	16, <u>207</u>	(207) I--I--I--I--I--0++7++I--I--	2	9
67	6, 16, 72	(72) 2+-I--I--I--I--I--I--I--+	2	8
68	16, 124	(124) 2--I--I--I--I--I--2++I--I--+	2	8
71	108, 159	(108, 159) 0++0+-0++0+-0++I--I++I--	3	8
72	38, 67			
74	42, 110, 140	(110) I--I--I--I--I--I--I--I--0++	2	8
		(140) I--I--I--I--I--2--I--I++I--	2	8
76	83	(83) I--I--I--I--I--2++2++2+-	2	8
79	115	(115) I--I--0++I--I++I--I--I--	3	8
83	76, 141	(141) 2++2+-I++I--I--I--I++I--I--+	3	9
84	6			
85	40, 216	(216) I--I--I--2--2--I++I--I--I--I--+	2	10
86	200	(200) 0+-0++I--I--I--2--4++I--	3	8
87	142	(142) I++I++0++0++I++I--I--I--I--+	2	10
93	144	(144) I--I++I--I++I--I++I--I++	2	8
96	8, 130, 141	(130, 141) I++I--I--I--I--I--I--I--I--+	4	9
100	51			
104	(54, 171)			
108	51, (71, 159)			
110	74			



I	2	3	4	5
112	130	(130) I++I+-0++I--I+-I+-I+-I+-I++I+-	2	10
115	79, 200	(200) 0++5+-I+-I+-I+-I--0++0+-I++I+-	2	10
118	133	(133) 0++3+-2++0+-I+-I+-I--I++I+-	2	9
120	14			
121	49, 59			
124	63, 68			
128	192	(192) I++I--I+-I+-I+-I+-I--I++	2	8
129	191	(191) 2+-2++2+-0++I--I+-I+-I+-I++	2	10
130	96, 112, 141, 191	(191) I++I+-I+-I+-0++I--I+-I--	2	8
132	151	(151) I+-I--I+-I--4++I+-I+-I--	2	8
133	116			
134	138	(138) I++I++I+-0++0+-0++0+-I--	2	8
135	55			
138	134, 140, 199, 212, 213	(140) I+-I+-I--I+-I--2+-I+-I++ (199) I--I+-I--2+-I+-I--2+-I-- (212) I+-0++0+-0++0+-I+-I+-I--0++ (213) I--2+-I+-I++2+-I+-0+-0++I--	2 2 2 3	8 8 10 9
140	18, 74, 138			
141	48, 51, 83			
142	87, 183	(183) 0+-I--2+-I--0+-I--I+-I--I--	3	9
144	93			
148	12			

	2	3	4	5
149	162, 202	(162) I--I--I--I--3--I+-I++I+- (202) I++I+-2++I--2--2--I--I--	2	8
151	122, 157	(157) I--I--I--I--5++I+-I--I--	2	8
152	45			
157	151, 159	(159) I+-I+-I--I--2+-7--2+-2++2+-I--	3	10
163	(71, 103), <u>157</u>			
164	<u>35</u>			
165	<u>52</u>			
171	54, 104			
172	45			
175	32			
176	31			
182	32, 149, 165	(195) I--I--I--I--2++3--2++I--	2	8
183	47, 142			
191	12, <u>129</u> , 130, 195	(195) I+-3--2+-I--I--I--I--2++	2	8
204	214	(214) I--I--I++I+-I++I+-I--0++	2	8
192	128			
194	14			
195	62, 182, 191			
199	6, 135			
200	<u>63</u> , 86, 115			
202	<u>132</u>			
204	214			
207	<u>64</u>			

I	2	3	4	5
209	I			
210	33			
212	22, 138			
213	138			
214	14, 204			
216	86			

- Примечания: 1. Мелодии, отсутствующие в графе "1", не имеют связей с другими мелодиями на уровне 8-грамм и выше.
2. Подчеркивание мелодий во втором столбце фиксирует факт самозамыкания: мелодия, указанная в той же строке в первом столбце, написана тем же композитором, что и подчеркнутая.
3. Пропуск в графе "3" означает, что 1-грамма приведена ранее для мелодии с меньшим порядковым номером.
4. Если значение частоты 1-граммы (графа 4) превышает число мелодий, в которых она встретилась, это указывает на многократность вхождения ее в отдельные мелодии.