

Б. В. Орехов
Институт языкознания РАН,
Национальный исследовательский университет
«Высшая школа экономики»

Моделирование
терминологического тезауруса
работ Р. Г. Назирова
о мифологии и истории
фольклорных сюжетов

В 2011 году Л. А. Каракуц-Бородина писала: «Ближайшей задачей анализа языковой личности Ромэна Гафановича Назирова видится нам описание его тезауруса в ключевых концептах; предположительно — Человек, Истина, Красота, Театр, однако максимально достоверный ответ на этот вопрос можно будет дать лишь на основе машинной обработки максимально полого корпуса его текстов»¹. Удовлетворительного по объёму корпуса текстов Р. Г. Назирова в электронном виде мы не получим ещё долго (возможно, никогда). Но уже сейчас нам доступны достаточно существенные, тематически автономные и внутренне законченные фрагменты назировского текста², наиболее привлекательным среди которых я бы счёл фольклорно-мифологический цикл.

В отличие от статей и книг, посвящённых Достоевскому, материал, на котором основаны эти работы, более универсален и многообразен, позволяет автору продемонстрировать свою эрудицию не только в координатах русской литературы, но и выйти на масштабные обобщения историко-культурного плана. Кроме того, чётче осмыслен контекст этих трудов: Назиров указывал Веселовского в качестве одного из своих ориентиров в науке³ (и интуитивно понятно, почему) в то время как формирование Назирова-литературоведа — это сложная в том числе и методологическая проблема, факторы и механизмы этого процесса не ясны и требуют специального осмысления.

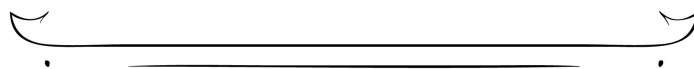
Литературовед Р. Г. Назиров в печати попробовал себя в новом качестве историка архаических сюжетов в 1980 году, выпустив статью «Наглядная дипломатия» и предметное иносказание в фольклоре и литературе (к постановке вопроса)» в сборнике «Фольклор народов РСФСР». «С этого времени интерес Назирова к фольклору, мифологии и их связям с литературой и культурой не ослабевает, становясь своего рода продолжением его литературоведческих штудий. В течение двенадцати лет каждый год учёный публикует по статье, каждая из которых представляет реконструкцию истории некоторого мотива, чаще всего восходящего к мифоритуальным истокам»⁴. Всего в круге этой тематики Назировым опубликовано 16 статей. Уже после смерти учёного увидели свет другие его труды, тематически смыкающиеся с названными. Прежде всего, это монография «Становление мифов и их историческая жизнь» (а также подготовительные материалы к ней), восемь

¹ Каракуц-Бородина Л. А. Человек эпохи пробуждения // Назиров Р. Г. Избранные газетные рецензии. Уфа, 2011. С. 73

² О системе назировского текста см. Орехов Б. В., Шаулов С. С. Архив Р. Г. Назирова как система // Назировский архив. 2013. № 2. С. 118—126.

³ См. Назировский архив. 2013. № 1. С. 139.

⁴ Орехов Б. В., Шаулов С. С. Сумма мифологии. Предисловие // Становление мифов и их историческая жизнь. Уфа, 2014. С. 4



найденных в рукописях статей, и наброски неосуществлённой монографии об истории сюжетов, в которую все эти тексты и должны были войти на правах отдельных параграфов (см. № 25, 27, 31, 34—36, 38, 40, 42, 43, 46, 50, 56, 58, 74, 75, 82, 86—89, 91, 97, 112, 120, 123 в сводной библиографии трудов Р. Г. Назирова в настоящем номере).

В силу относительной завершенности этого корпуса можно обратиться к нему как к репрезентанту определённой научной идеологии, проявляющей себя прежде всего в языке и системе понятий. Для научного текста система понятий концентрируется в терминах и особенностях их употребления. Попытку изучения терминологического тезауруса Назирова я и предпринимаю в настоящем исследовании. Прежде всего, здесь будут продемонстрированы некоторые количественные и околоквантитативные методы. Скорее, можно говорить даже не о законченном исследовании, а о первом пробном подходе к многоаспектной проблеме.

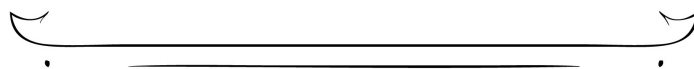
В качестве материала я воспользовался всеми названными текстами. Полный их объём составил 168 474 слова. Это довольно маленький корпус, как уже говорилось ранее, он обладает другими достоинствами — на концептуальном уровне.

Я построил сеть совместной встречаемости ключевых терминов статей. Эта методика хорошо вписывается в активно развивающееся направление применения сетей (*complex networks*) для описания функционирования лексики, в том числе отдельных авторов. Последнее позволяет осуществить выход в стилометрию⁵.

Исследование было проведено следующим образом. Все входящие в корпус тексты были очищены от чисел, небуквенных символов, лемматизированы (то есть все слова приведены к своим словарным формам), а также из текстов были удалены первые 50 самых частотных лексем русского языка (список был заимствован из Частотного словаря О. Ляшевской и С. Шарова). Затем для всего объёма текстов была посчитана совместная встречаемость слов в двух различных окнах: 3 и 5 слов до и после искомого. Это значит, что предложение «От этого сюжета о превращении статуи в живое человеческое существо принципиально отличаются рассказы о статуях, живущих скрытой жизнью» было представлено в следующем виде: «сюжет превращение статуя живой человеческий существо принципиально отличаться рассказ статуя жить скрытый жизнь». При этом в окне 3 сочетаемость слова «принципиально» со словом «статуя» имеет величину 1, а в окне 5 та же сочетаемость имеет величину 2.

Представление сети всех лексем из упомянутых текстов не слишком информативно, поэтому мы сосредоточились на центральных понятийных точках, которые имеют ключевое значение для всего концептуального плана работ о мифологии и фольклоре. Исследователя прежде всего интересовали такие термины

⁵ Amancio D.R. A complex network approach to stylometry // PLoS ONE. 10(8): e0136076; Antiqueira L., Pardo T., Nunes M., Oliveira O. N. Some issues on complex networks for author characterization // *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*. 2007. 11 (36). P. 51—58.



как миф, мифология, сюжет, жанр, происхождение, поэтика, мотив, фабула, герой, нарратив. Построенный в ходе исследования граф смог бы продемонстрировать текстовое, а, следовательно, и концептуальное соседство тех слов, которые оказываются в одном контексте с искомым. Иными словами, статьи и монография Назирова были представлены в виде графа, в котором узлами являются слова, а ребра отражают их совместную встречаемость в тексте.

Для начала следует ограничиться только теми терминами, которые употребляются в одном контексте во всем корпусе не менее 50 раз, то есть ребро в графе появляется между узлами только тогда, когда во всех анализируемых текстах Назирова слова-узлы оказались рядом 50 и более раз. Рассмотрим окно в 5 слов (см. рис. 1).

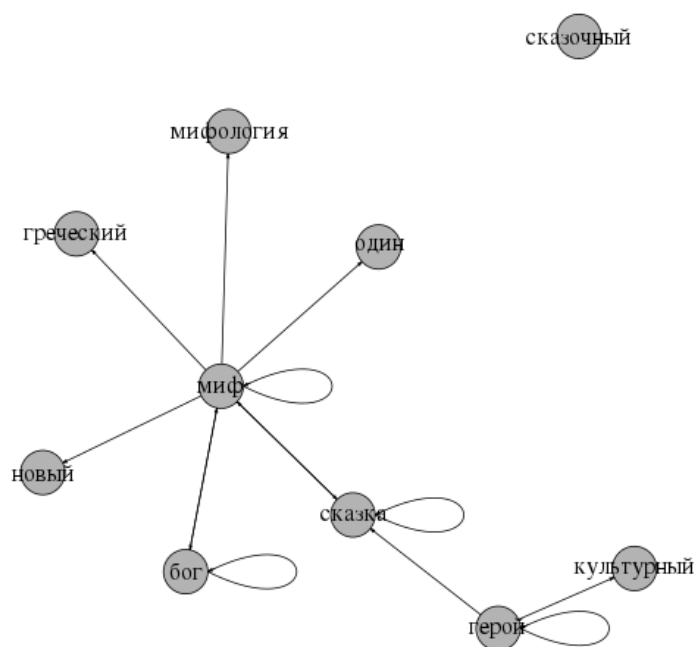
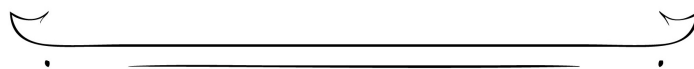


Рис. 1. Граф, отражающий совместную встречаемость терминов в текстах Назирова в окне 5 (случаев совместной встречаемости не менее 50)

Визуализация хорошо показывает центральное понятие среди рассматриваемых, это понятие мифа, который связан прежде всего с понятиями бога (этот термин в графе имеет т. н. «петлю», то есть зафиксирована его встречаемость с самим собой), сказки и мифологии в целом. Любопытны попавшие в эту модель прилагательные: «новый» и «греческий». Несмотря на широту эрудиции ученый всё-таки чаще, чем можно было бы ожидать, обращается к иллюстрации своих идей с помощью



греческого материала.

Своё собственное ответвление создает понятие «сказки»: именно с ней оказывается связан термин «герой», который, разумеется, в текстах такого рода тематики может быть только «культурным».

Чуть более разветвленную картину демонстрирует граф, построенный на наборе терминов, вес ребра между которыми составляет не менее 40 случаев совместного употребления (окно вокруг искомого слова — всё те же 5 слов слева и справа), см. рис. 2.

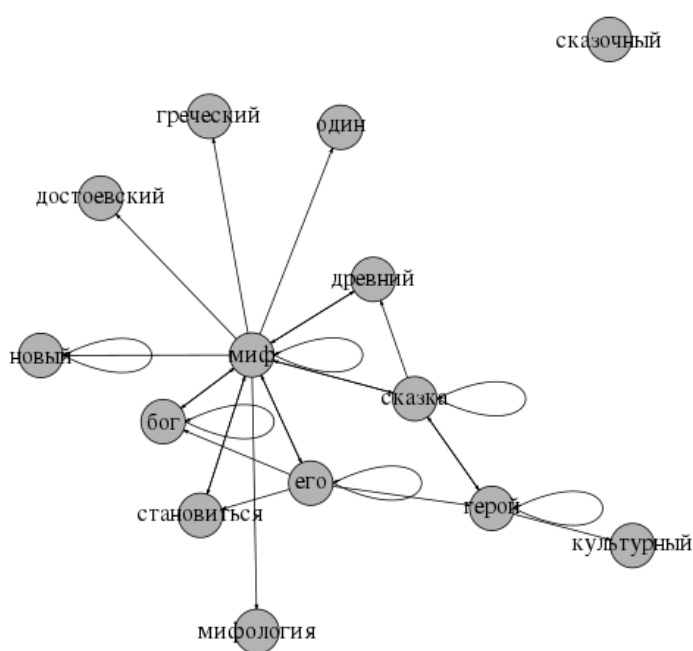


Рис. 2. Граф, отражающий совместную встречаемость терминов в текстах Назирова в окне 5 (случаев совместной встречаемости не менее 40)

Здесь на граф попали новые термины (разумеется, все те, которые мы уже видели на предыдущем рисунке тоже присутствуют): «становится», «Достоевский», «древний». Примечательно, что последний термин оказывается общим для «мифа» и «сказки».

Гораздо более сложную систему представляет собой граф, вбирающий слова, для которых случаев совместной встречаемости находится не менее 30 (см. рис. 3)

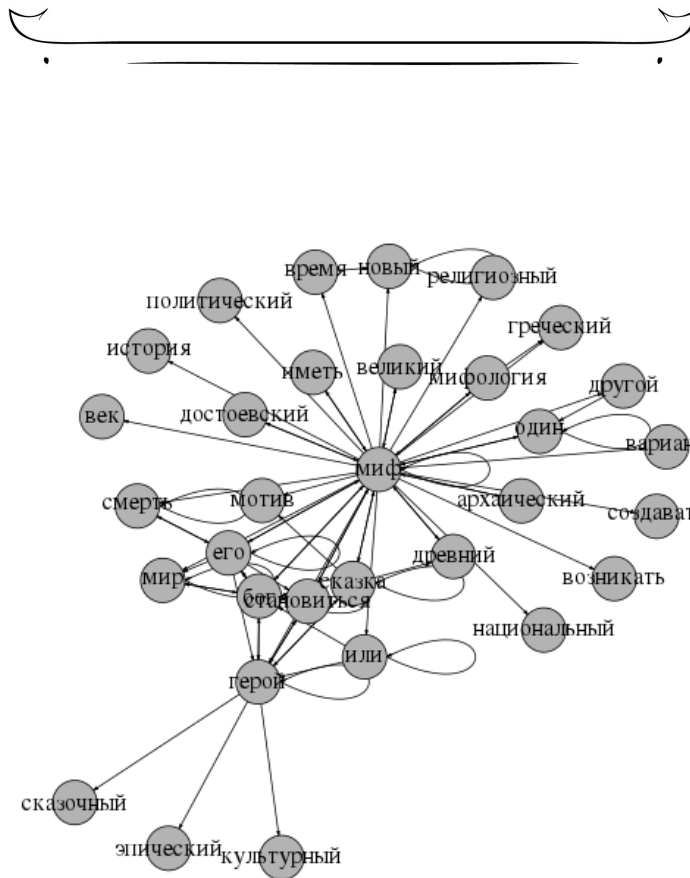
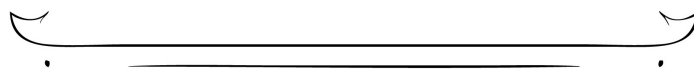


Рис. 3. Граф, отражающий совместную встречаемость терминов в текстах Назирова в окне 5 (случаев совместной встречаемости не менее 30)

Хорошо видно, как термины обрастают важными определениями. Так, центральное понятие для анализируемого корпуса, миф, в концепции Назирова предстает в таких аспектах как «политический», «великий» (сказывается характерная для Назирова эмоциональность подачи исторического материала), «религиозный», «архаический», «национальный», а герой может быть не только «культурным», но и «сказочным» и «эпическим».

Тот же миф при более пристальном рассмотрении соседствует с такими словами как «век», «возникать», «история», так или иначе отражающие диахронический аспект темы. Благодаря этому видно, что Назиров сосредотачивается на времени жизни мифа, а не пишет его вневременную «теорию».

Любопытны и такие понятия, которые оказываются связаны внутри графа сразу с несколькими узлами. Так, греческий дискурс востребован и в контексте мифа, и в контексте мифологии в целом. Важное в этом контексте понятие бога сопряжено и с понятием «героя», и с понятием «мира». Впервые встретившийся нам «мотив» оказывается связан как с мифом, так и со сказкой, а последняя связана с определением «древний», ровно как и сам миф. Кроме «мотива» граф показывает нам и ещё один специфичный термин фольклористики: «вариант».



В то же время граф с окном в 3 слова даже для лимита в 30 совместной встречаемости предельно аскетичен (см. рис. 4).

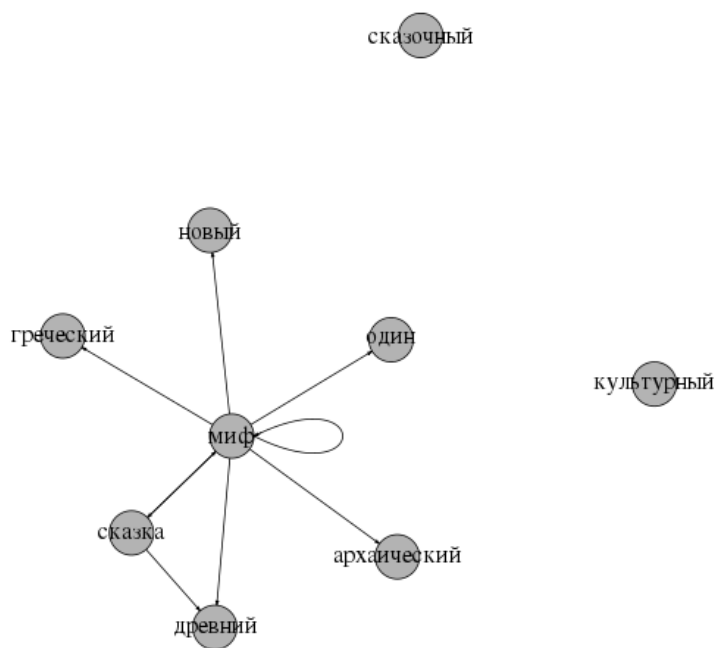


Рис. 4. Граф, отражающий совместную встречаемость терминов в текстах Назирова в окне 3 (случаев совместной встречаемости не менее 30)

Это говорит о том, что в целом одинаковые слова неохотно попадают в один контекст у Назирова, и совместная встречаемость именно отражённых на рисунке терминов предельно нагружена концептуально, то есть их представленность на графике значима, и к этому следует относиться серьёзно. Показательно здесь одновременное функционирование сразу двух синонимичных определений: «древний» и «архаический». Очевидно, вопросы происхождения и исторических истоков интересуют Назирова чуть больше, чем современная политическая проекция мифологических представлений. Хотя при этом мы наблюдаем и другое прилагательное временного ряда: «новый». Их оппозиция, по всей видимости, и отражает главную семантическую ось концепции Назирова.

Следующий шаг в расширении модели терминологического тезауруса мы сделаем, сбросив лимит до 25 совместных словоупотреблений. Получившаяся картина отражена на рис. 5.

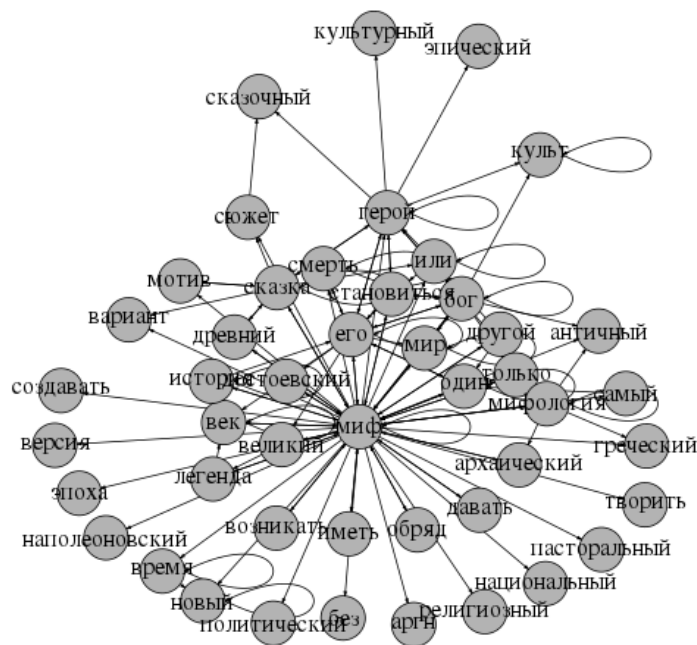
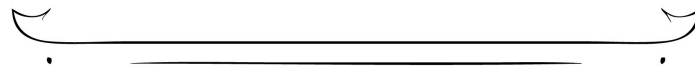
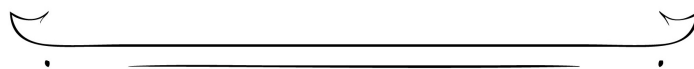


Рис. 5. Граф, отражающий совместную встречаемость терминов в текстах Назирова в окне 5 (случаев совместной встречаемости не менее 25).

Здесь проявляют себя новые для нас и весьма важные определения мифа: «наполеоновский», «пасторальный». Модель вбирает в себя и больше собственно литературоведческих терминов и терминов фольклористики: «легенда» (оказывается связана ребром с «веком»), «сюжет» (рёбра ведут к «мифу», «сказке» и «сказочному»), «культ» (связан с «героем» и «богом»). Значимо и появление ещё одного временно́го слова: «эпоха».

Отметим, что все представленные графы не случайно имеют звездообразную структуру: в центре модели всегда стоит слово «миф», а другие термины и понятия встраиваются в неё постольку поскольку связаны именно с ним. Разумеется, нужно учитывать, что самым объёмным текстом в анализируемом корпусе выступает монография, посвященная именно мифологической проблематике. В то же время эпицентровость понятия мифа и для других смежных трудов учёного также следует признать справедливой.

Дальнейшее увеличение масштаба в исследовательской оптике (по крайней мере, если пользоваться только визуализацией как инструментом исследования тезауруса) следует признать неконструктивным: расширение модели создаёт трудности в её изучении, так как узлов становится слишком много, и теряется наглядность.



Обзор был бы неполным, если бы я не попытался сравнить полученные данные с аналогичными моделями, построенными на текстах других учёных. В качестве контрольных корпусов были выбраны тексты трёх авторов: А. Н. Веселовский (1838—1906) О. М. Фрейденберг (1890—1955) и Е. М. Мелетинский (1918—2005). Выбор не случаен: Веселовского сам Назиров назвал в числе своих ориентиров в науке, Фрейденберг известна своим ярким и неформальным научным языком, а Мелетинский больше всего «похож» на Назирова по времени творческой активности и сфере научных занятий (с «поколенческой» точки зрения, вероятно, больше подходили бы тексты В. Н. Топорова, но они и стилистически, и методологически сильнее отличаются от той традиции, которую в своих работах продолжает Назиров).

Веселовский представлен в корпусе текстами объёмом в 258 719 словоформ (то есть этот текстовый объём гораздо существеннее, чем у Назирова), тематически это работы так или иначе вращающиеся в орбите «Исторической поэтики». В корпус Фрейденберг (282 061 словоформ) вошли ее большие труды об образе и понятии, о евангельском осле и происхождении ряда литературных явлений античности. Коллекция текстов Мелетинского (292 117 словоформ) сформирована его книгами и статьями о происхождении словесного искусства, о мифологических представлениях и их текстуальном воплощении, о конкретных мифах и жанрах.

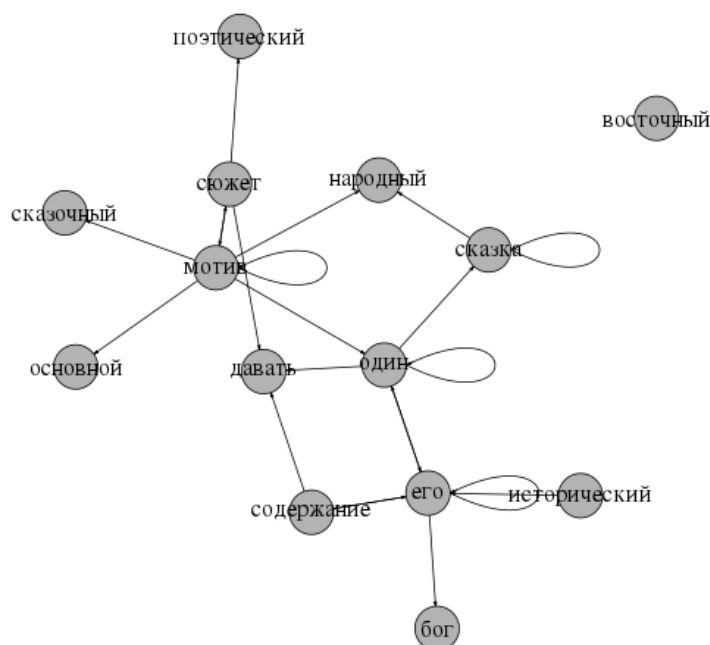
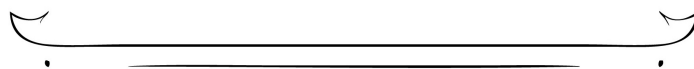


Рис. 6. Граф, отражающий совместную встречаемость терминов в текстах Веселовского в окне 3 (случаев совместной встречаемости не менее 15).



Несмотря на величину корпуса Веселовского, понятия в них не тяготеют к совместной встречаемости. Взаимодействие терминов мы видим на рис. 6, где несмотря на серьёзно пониженный порог попадания в граф, мы видим всего полторы дюжины терминов. Разумеется, центральный среди них — «мотив», образующий вокруг себя конструкцию остальных понятий. Второй по значимости — «сюжет», связанный как с «мотивом», так и с понятием «поэтического» и процессуальным «давать». Пересечения с терминологическим тезаурусом Назирова имеются («сказка», «сказочный», «исторический», «бог»), но все эти термины, очевидным образом, иначе участвуют в образовании общей системы, а «миф» и вовсе отсутствует.

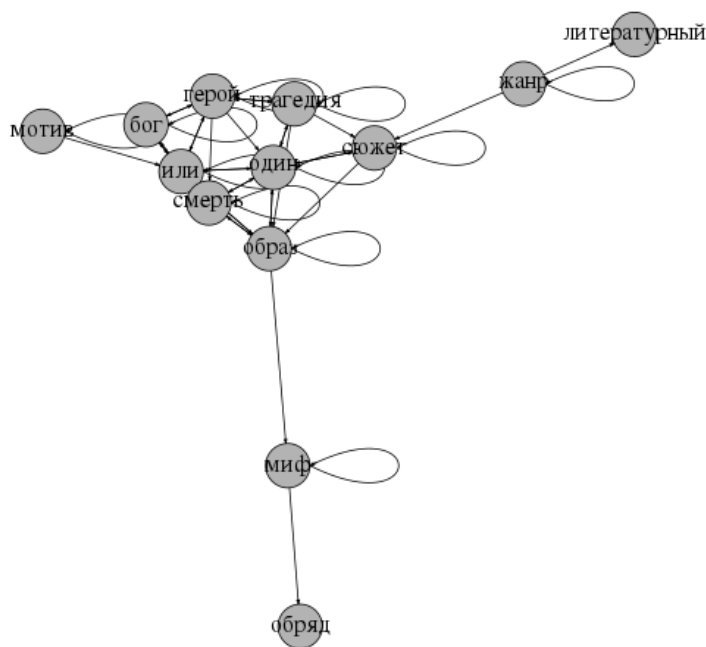
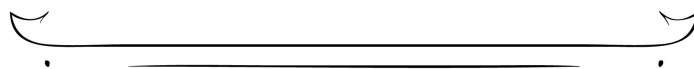


Рис. 7. Граф, отражающий совместную встречаемость терминов в текстах Фрейденберг в окне 5 (случаев совместной встречаемости не менее 50).

Совсем иную картину рисуют тексты Фрейденберг (рис. 7). Впервые мы сталкиваемся с ситуацией, когда у терминологической модели нет очевидного центра, ее структура не звездообразна, и в центральной части графа узлы густо переплетены рёбрами совместной встречаемости. «Герой», «бог», «трагедия», «образ», «смерть», «сюжет» имеют каждый по несколько соседей, и все они (то есть по-настоящему — никто) пытается претендовать на статус главного термина. «Мотив» (главный для Веселовского), «миф» (главный для Назирова) и «жанр» оказываются



периферийными понятиями, так как связаны с центральной кликой узлов всего одним ребром.

Ещё более запутанную для визуального анализа картину мы наблюдаем при окне в 5 слов и весе ребра не менее 40 (рис. 8).

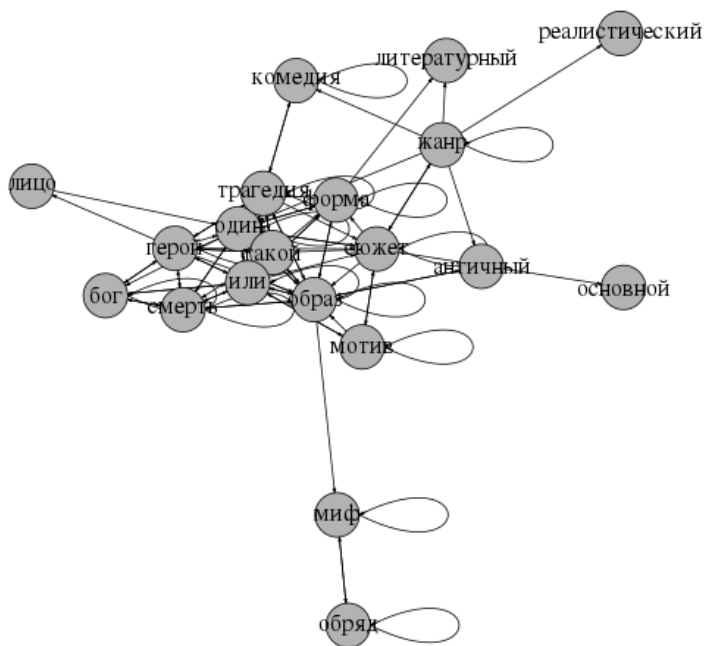


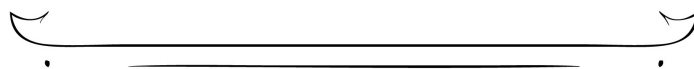
Рис. 8. Граф, отражающий совместную встречаемость терминов в текстах Фрейденберг в окне 5 (случаев совместной встречаемости не менее 40).

Ситуация с центральным понятием системы не проясняется, в центральную клику добавляются новые понятийные единицы («форма», «лицо»); «миф» по-прежнему образует только ответвление «образа» в то время как «мотив» теснее связывается с «образом» и «сюжетом».

Примечательно, что и у Назирова, и у Фрейденберг как центральное содержательное понятие мифа, проявляющее себя в терминологической системе, отчётливо выделяется «смерть».

Объём корпуса Мелетинского и связанность терминов в его текстах приводит к тому, что даже высокие пороги пропускают на граф солидное число узлов.

Центральным понятием терминологической системы Мелетинского является понятие «героя» (см. рис. 9). Именно с ним объединяются важный для Фрейденберг термин «образ», определения «сказочный» и «мифический». При этом «миф» играет



роль второстепенного центра.

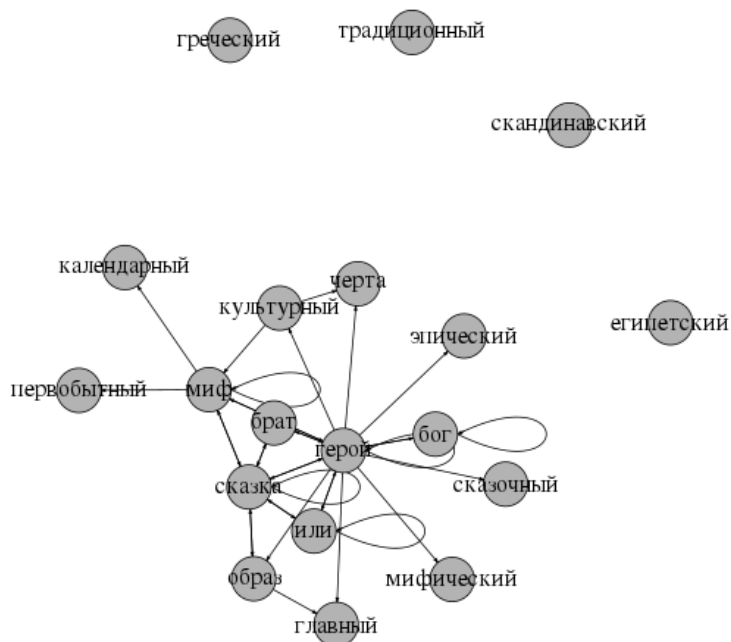


Рис. 9. Граф, отражающий совместную встречаемость терминов в текстах Мелетинского в окне 3 (случаев совместной встречаемости не менее 50).

Обращают на себя внимание повторяющиеся элементы терминологической системы Назирова соседи миф — сказка, миф — первобытный, герой — бог, герой — культурный. Это не слишком удивительно, известно, что Назиров во многом ориентировался на обобщения Мелетинского в сфере мифологии (хотя и не всегда был с ними согласен), часто цитировал их. Впрочем, он цитировал и Веселовского, и Фрейденберг. Очевидно, сходство между терминологическими системами Мелетинского и Назирова лежит на более глубоком уровне, чем простое внимание к работе старшего коллеги.

Ещё более явно тенденцию к образованию двух центральных структур графа вокруг понятия «героя» (главного для Мелетинского) и понятия «мифа» (важного для Мелетинского и центрального для Назирова) демонстрирует рис. 10. Хорошо видно, что у «мифа» на треть меньше соседей, чем у «героя».

