

В сети терминов М. М. Бахтина: теория графов о диалоге, карнавале и хронотопе

Б. В. Орехов

НИУ «Высшая школа экономики»

Изучение терминологической системы Бахтина уже имеет свою научную историю. Кажется, самым заметным её эпизодом стало издание «Бахтинского тезауруса»¹, которое в свою очередь породило оживлённый обмен репликами в периодике². Названными публикациями библиография проблемы, разумеется, не исчерпывается³. В этой заметке нам хотелось бы обратить внимание филологического сообщества и - специально — бахтинистики на имеющиеся в распоряжении современной науки методы решения некоторых задач, связанных с анализом текста и, в частности, терминологии.

Среди этого круга методов особенно перспективными нам представляются приложения к терминологии теории графов: см. наши публикации на эту тему⁴. Этот инструментарий позволяет работать с научными моделями, в которых ключевыми элементами являются некоторые сущности, представляемые как узлы (или вершины) графа (сети), а принципиальным свойством моделей является учёт связей между этими сущностями, фиксируемых в модели как рёбра между узлами. Поскольку мы имеем дело с универсальным принципом моделирования, и узлом, и ребром может выступать эквивалент любого реального объекта. В нашем случае мы попытаемся построить такую модель словоупотребления М. М. Бахтина, в которой узлами выступают частотные полнозначные лексемы в его текстах, а рёбра будут отражать частотную совместную встречаемость слов в бахтинских произведениях.

¹Бахтинский тезаурус: материалы и исследования: сборник статей / ред. Н. Д. Тмарченко, С. Н. Бройтман, А. Садецкий. М.: РГГУ, 1997.

²Зенкин С. Н. Испытание тезаурусом (Заметки о теории, 10) // Новое литературное обозрение. 2005. № 72. С. 327–335; Бройтман С. Н. Испытание Бахтиным // Новый филологический вестник. 2006. № 1 (2). С. 199–205; Зенкин С. Н. Обсудим разногласия // Новый филологический вестник. 2006. № 1 (2). С. 206–210; Тмарченко Н. Д. Легко ли пройти такое испытание? // Новый филологический вестник. 2006. № 1 (2). С. 211–215.

³См., например, Киржаева В. П., Осовский О. Е. О двух терминах М. М. Бахтина в контексте истории отечественного литературоведения XX века // Филология и культура. 2016. № 1 (43). С. 223–228.

⁴Орехов Б. В. Что может рассказать теория графов о терминологической системе О. М. Фрейденаберг // Вестник РГГУ. Серия «История. Филология. Культурология. Востоковедение». 2017. № 4. С. 89–98; Орехов Б. В. Гуманитарная терминология как сеть: теория графов о закономерностях научного стиля // Критика и семиотика. 2016. № 2. С. 94–101; Орехов Б. Моделирование терминологического тезауруса работ Р. Г. Назирова о мифологии и истории фольклорных сюжетов // Назировский архив. 2015. № 2. С. 118–131

Построение модели во многом повторяет путь, описанный нами в предыдущей работе на эту тему⁵. Исходным материалом послужил текст, опубликованный в собрании сочинений М. М. Бахтина в 7-ми томах. Этот корпус представляет собой массив в 911366 слов, из которого были изъяты неполнозначные лексические единицы. Финальный объём анализируемого текста после удаления служебных слов составил 542878 словоупотреблений. Словоформы этого корпуса автоматически приведены к своей начальной форме, и на таком образом обработанном материале построена сетевая модель. Узлами сети стали употреблённые М. М. Бахтиным лексемы, а связь (ребро) между ними регистрировалось в том случае, если эти лексические единицы встречались в его текстах на расстоянии 5 слов друг от друга не менее пяти раз. Узлы без рёбер в сеть не включались⁶.

Как показали предыдущие исследования, узлами сети при таком моделировании становятся прежде всего важнейшие для системы автора термины. Но этот подход позволяет включить в модель не все интересующие исследователей единицы: Бахтин мог говорить (и мы знаем, что говорил) о чём-то в форме заметок, конспективно, эпизодически упоминая интересующие его понятия. Таким образом, метод предполагает количественный порог, отсекающий некоторые термины, но не по принципу частотности их употребления, а по принципу интенсивности их взаимодействия друг с другом. Иными словами, термины представлены именно как система, то есть приоритетным становится их взаимное функционирование в тексте.

К сожалению, сеть, построенная на таких основаниях, игнорирует временные различия в употреблении терминов, что не мешает построить несколько графов, каждый из которых отражал бы только тексты определённого периода, что, однако, выходит за рамки предлагаемого исследования. С течением времени автор может менять понятийное наполнение используемых единиц, смещаются акценты, совместная встречаемость слов зависит от тематики текстов. При «отвлечённом чтении» исходных текстов все эти нюансы неизбежно теряются. Однако общую картину терминологической системы эта модель всё же отражает. Граф улавливает реальные отношения терминов в тексте, а именно в конкретном текстовом словоупотреблении овеществляется потенциальная семантика. На этой идее основан, в частности, востребованный в прикладных лингвистических разработках дистрибутивный анализ, то есть метод исследования, который основан на изучении окружения единиц текста⁷.

⁵См.: Орехов Б. В. Гуманитарная терминология как сеть // Критика и семиотика. 2016. № 2. С. 94–101.

⁶Исключение составили так называемые «петли», то есть случаи, когда некоторый термин может встретиться 50 и более раз рядом с самим собой.

⁷Harris Z. Distributional structure // Word. 1954. N 10 (23). P. 146–162 и мн. др. О компьютерных перспективах см. Schutze H. Dimensions of meaning // Proceedings of Supercomputing'92. 1992. P. 787–796 и др.

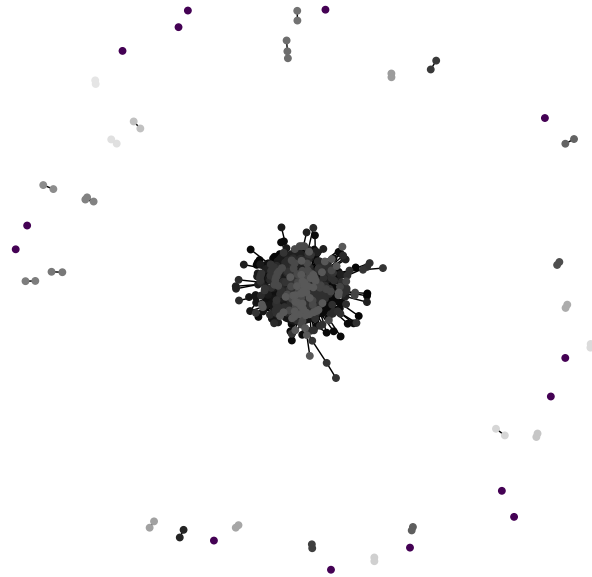


Рис. 1: Терминологическая сеть М. М. Бахтина

Описанный подход позволил построить терминологический граф⁸ М. М. Бахтина. В сети оказалось 710 узлов. Её структура видна на рис. 1: сеть имеет плотное центральное ядро, внутри которого мы наблюдаем высокую связность, а также несколько обособленных от этого ядра небольших «островов» с невысоким числом (от одного до трёх) устойчиво соединённых вершин. Такая структура должна интерпретироваться как наличие в языке Бахтина готового понятийного набора, много раз воспроизведённого текстуально как целостная лексическая система.

Рассмотрим сначала автономные кластеры, не входящие в центральное ядро. С одной стороны, это такие сочетания, как «князь» и «Мышкин»,

⁸Строго говоря, получившийся в результате граф не именно терминологический, а граф словоупотреблений М. М. Бахтина. В то же время несомненно, что в качестве его узлов (и что важнее — в качестве конструктивно наиболее значимых его узлов) большей частью мы находим именно те лексические единицы, которые несомненно должны быть признаны терминами. Статус остальных слов, попавших в граф, по всей видимости, должен быть определён как окказиональные термины, имеющие такое положение именно в системе М. М. Бахтина.

«Настасья» и «Филипповна», «Фёдор» и «Павлович», «дон» и «Кихот» то есть устойчивые двусловные обозначения персонажей. Сюда же примыкает кластер «преступление» и «наказание», очевидным образом указывающий на название романа Ф. М. Достоевского. В модели эти имена оказались изолированы от прочей понятийной системы, то есть остались концептуально нейтральными. С другой стороны, в этом обособленном ряду мы видим кластеры, образованные устойчивыми сочетаниями, репрезентирующими научный стиль М. М. Бахтина, но состоящими из однозначных слов и поэтому не исключёнными в процессе предобработки: «подлежать», «никакой», «сомнение» («не подлежит никакому сомнению»); «служить», «цель» («служить цели»); «вполне», «понятно» («вполне понятно»). Список этих фразеологизованных единиц интересен с точки зрения изучения стиля Бахтина, но не его понятийной системы. В этом же ряду должны рассматриваться словосочетания «удельный вес» и «исходный пункт», также частотные в текстах Бахтина, а лексемы, из которых они состоят, также составляют автономные группы узлов сети.

Наконец, третью группу обособленных вершин графа составляют сочетания действительно концептуально значимых единиц: «индийский» и «чудо», «гуманитарный» и «наука», «добро» и «зло», «исповедь» и «самоотчёт», «бранный» и «хвалебный», «эстетически» и «значимый». Их выделение из ядра сети отчётливо иллюстрирует судьба кластера «индийский» и «чудо», возникшего, разумеется, благодаря пятой главе «Творчества Франсуа Рабле и народной культуры средневековья и Ренессанса», где индийским чудесам как источнику гротескно-телесных образов посвящён целый параграф. Однако несмотря на наличие такого цельного понятия в бахтинской системе, текстуально она остаётся самостоятельной и не сопрягается с другими частями концептуальной схемы. Действительно, параграф об индийских чудесах в композиции книги стоит особняком, как важный, но обособленный элемент историко-литературных рассуждений.

Особенное место занимают самостоятельные узлы сети, связанные только с самими собой, то есть встречающиеся в текстах достаточное число раз в соседстве со своими повторениями, но не в текстуальной близости ни с какими другими словами, составившими вершины терминологического графа: «память», «истина», «бог», «спор», «сцена», «порог», «сон», «новелла», «случайный», «мистерия», «толпа», «любовь», «миф». Эти изолированные в бахтинской системе понятия также появляются из самодостаточных текстовых фрагментов, посвящённых, главным образом, одному историко-литературному или философскому сюжету, репрезентированному этой лексемой, как, например, несколько абзацев о мистерии в третьей главе «Творчества Франсуа Рабле».

Обычно сетевой анализ (как часто называют практическое приложение теории графов) позволяет получить два типа информации о построенной модели: во-первых, мы можем установить наиболее значимые узлы сети; во-вторых, мы можем установить, каким образом узлы группируются внутри сети, исходя из структуры их связей.

Рассмотрим эти сведения последовательно.

Как наиболее значимые узлы сети осмысляются такие, которые обладают наибольшим значением по шкале центральности и нагрузки. Это количественные, вычисляемые характеристики вершин графа, которые зависят от числа и качества рёбер в модели. Принято считать, что центральность узла — это его способность быть «несущим» для всей конструкции (то есть его удаление приведёт к разрушению системы), а нагрузка — это способность участвовать в связях между разными элементами системы (то есть его удаление приведёт к худшей «проводимости» между частями графа). Вычислив центральность и нагрузку узлов в терминологической системе Бахтина, мы узнаем, какие понятия в ней являются наиболее важными, организующими её внутреннюю связность.

В таблице 1 можно увидеть, что список наиболее центральных и наиболее нагруженных узлов практически совпадает. Естественно, что эти узлы принадлежат к связному ядру графа, состоящему из 648 вершин (см. рис. 2).

Таблица 1: Значение центральности и нагрузки для наиболее значимых узлов сети

Узлы с наибольшей центральностью	Значение центральности	Узлы с наибольшей нагрузкой	Значение нагрузки
образ	0.320	образ	0.228
роман	0.286	роман	0.167
слово	0.252	слово	0.139
мир	0.203	мир	0.079
форма	0.194	форма	0.071
герой	0.181	герой	0.064
язык	0.155	язык	0.053
рабле	0.119	рабле	0.030
отношение	0.102	достоевский	0.026
достоевский	0.101	новый	0.024

Итак, термины «образ», «роман», «слово», «мир», «форма», «герой» и «язык» (в таком порядке) находятся в центре понятийной системы Бахтина, что в практическом смысле означает, что они будут обнаруживаться в ближайшем текстуальном соседстве от любого ключевого для учёного термина. Иными словами, практически никакое концептуально значимое и достаточно пространное рассуждение не сможет обойтись без использования единиц этого ряда. Число соседей в графе для узла «образ» — 227, для узла «роман» — 203, для «слова» — 179. Со столькими идейно насыщенными элементами эти лексемы регулярно взаимодействуют в текстах.

Сетевая модель не сообщает нам ничего о семантике этих единиц (наверняка во множестве конкретных контекстов реализуют себя разные оттенки значения), зато может подсказать точки входа для размышлений о кон-

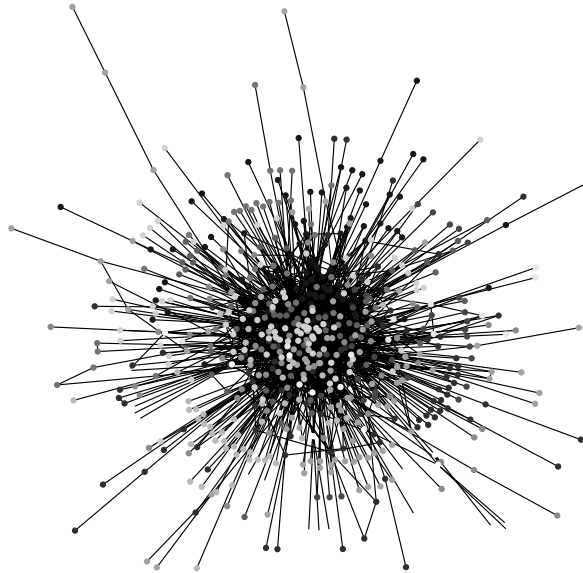


Рис. 2: Ядро понятийной сети М. М. Бахтина

центральном наполнении терминов из «подвала» этого списка. Так, термин «гротеск» имеет общие рёбра только с терминами «романтический» и «образ».

Кажущийся конструктивно важным элементом терминологической системы Бахтина «хронотоп» (этот эффект проявляет себя в том числе и благодаря названию нашего журнала) на самом деле обладает минимумом связей (в терминах сетевого анализа число связей узла называется степенью) внутри графа (см. рис. 3). Это, разумеется, не будет удивлять, если вспомнить, что к обсуждаемому понятию Бахтин обратился только в поздней работе, а не в центральных для его творчества книгах о Достоевском и Рабле.

Разработанные в рамках сетевого анализа автоматические алгоритмы позволяют распознавать внутри графа более тесно связанные друг с другом группы узлов, которые называют кластерами или сообществами. Таких алгоритмов существует много, и качество их работы зависит от структуры

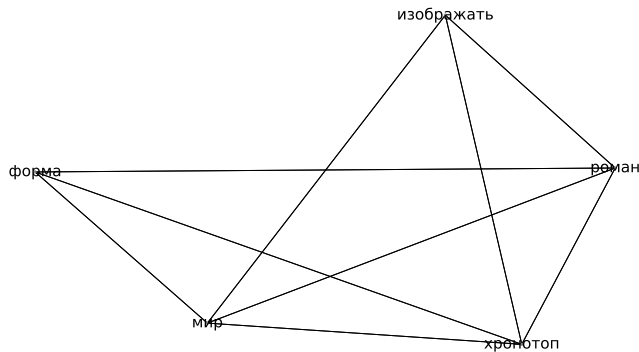


Рис. 3: Подграф ближайших соседей узла «хронотоп»

графа. Например, алгоритм, работающий с нагрузкой рёбер⁹ делит 710 узлов на 455 сообществ, в большинстве которых единственным участником остаётся всего один термин, а существенная часть ядра сети оказывается «слеплена» в один большой кластер из 139 лексем, который по-прежнему остаётся для исследователя неструктурированным.

Несколько более эффективно действует алгоритм многоуровневой оптимизации модулярности¹⁰, он выделяет всего 47 кластеров, среди которых 9 крупных, то есть состоящих из нескольких десятков узлов.

К первому относятся термины, описывающие философию поступка: «сознание», «творческий», «ответственный», «конкретный», «этический», «созерцание», «ценностный», «определенный», «объект», «поступать», «возможный», «должный», «акт», «поступок», «эстетический», «деятельность», «событие», «познавательный», «участник», «взаимодействие» и др.

Здесь концентрируются термины, выражающие этическую концепцию Бахтина.

Второй включает, главным образом, лингвистическую терминологию, показывая, тем не менее, с какими общегуманитарными терминами она наиболее тесно соотносится: «переводить», «голос», «общение», «бытовой», «действительность», «чужой», «собственный», «смысл», «риторический», «слышать», «речевой», «мениппея», «мышление», «реплика», «латинский», «сократический», «стиль», «разговорный», «звучать», «персонаж», «жанр», «русский», «произведение», «диалог», «речь», «контекст», «говорить», «поэтический», «лингвистический», «разноречие», «французский», «вульгар-

⁹Описание алгоритма см. в Newman M. E. J., Girvan M. Finding and evaluating community structure in networks // Physical Review E. 2004. 69. P. 026113.

¹⁰Blondel V. D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks // Journal of Statistical Mechanics Theory and Experiment. 2008 (10).

ный», «предложение», «субъект», «язык», «высказывание», «драматический», «говорящий», «диалогический», «диалект», «монологический», «устный» и др.

Состав кластера хорошо согласуется с лингвистической концепцией Бахтина и даже термин «жанр», традиционно используемый литературоведением, закономерно оказывается в этом ряду благодаря идее «речевого жанра». Обращает на себя внимание, что именно в этом сообществе оказались термины «мениппея» и «персонаж». Первая попала сюда благодаря тесной связи с понятием «жанра», что обращает специальное внимание на множественную категориальную отнесённость этого понятия для Бахтина: жанр одновременно является для него и термином исторической поэтики, и термином лингвистики. Что касается «персонажа», то единственным его соседом в сети является как раз лингвистический термин «речь», что, конечно, отражает и их тесную концептуальную связь со всей идейной конструкцией Бахтина: персонаж воплощается в тексте в первую очередь через речь.

Занятно, что термины «субъект» и «объект» попали в разные кластеры (во второй и первый соответственно), хотя многие другие «парные» термины часто представлены вместе (см. «добро» и «зло» выше). Собственно, в сети они являются соседями (см. рис. 4), но при этом в соответствии с решением алгоритма тяготеют к разным группам вершин.

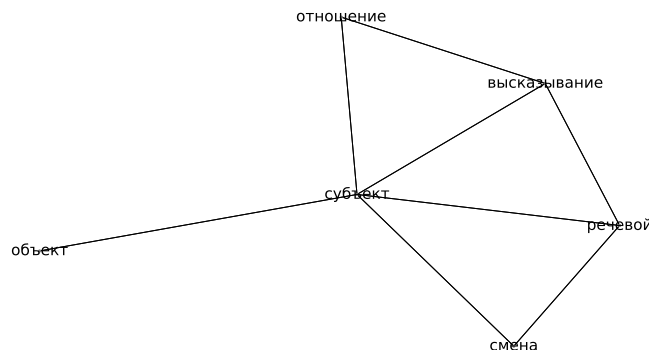


Рис. 4: Подграф ближайших соседей узла «субъект»

Третий кластер объединяет термины, описывающие механизмы карнавализации: «экспрессивный», «текст», «инстанция», «авторитарный», «тема», «площадной», «писание», «прозаический», «стилистика», «праздник», «диалогически», «яркий», «слияние», «хвала», «смысловой», «интонация», «стилизация», «глупец», «поэзия», «эстетика», «священный», «слово», «эмоционально», «шут», «пародийный», «романс», «фамильярный», «дурак», «осел», «брань», «зона», «понятие», «полемика» и др.

Симптоматично, что в этой группе тесно соседствуют термины, относящиеся к сфере текста (собственно «текст», а также «тема», «стилистика», «поэзия»), и относящиеся к сфере уличного карнавала («площадной», «праздник», «шут», «осел»). Органичное взаимопроникновение этих двух «реальностей» для историко-литературной оптики Бахтина вполне естественно.

Следующий кластер объединяет термины попавших в сферу внимания Бахтина мотивов, большинство из которых мы можем отнести к мифопоэтическим: «смерть», «рождение», «страх», «верх», «низ», «обновление», «мотив», «тело», «гротескный», «система», «романтический», «орган», «земля», «начало», «небо», «гротескно», «рот», «стихия», «пир», «встреча», «дорога», «женщина», «космический», «дух», «гигант», «питие», «сатирирование», «драма», «образ», «душа», «война», «игра», «еда», «половой», «мертвый», «подтирка», «пространство», «телесный», «гротеск» и др. Не трудно заметить, что эти лексемы востребованы Бахтиным для описания карнавала у Рабле, и дополняют таким образом предыдущую группу терминов с историко-литературной стороны.

Ещё одно сообщество терминов концентрирует бахтинскую теорию автора: «ценность», «авторский», «слушатель», «героиня», «герой», «точка», «зрение», «носитель», «рассказ», «мысль», «чувство», «автор», «биографический», «самосознание», «читатель», «интенция», «активность», «творец», «идеолог», «изображать», «судьба», «находиться», «созерцатель», «оценка», «рассказчик», «жизненный», «переживание», «позиция» и др. Ещё раз укажем на то, что «персонаж» и «герой» в этой системе не являются синонимами и выражают принципиально разные для Бахтина понятия.

Хорошо интерпретируется кластер, связанный с концепцией полифонического романа: «полифонический», «раскольников», «романный», «идея», «плутовский», «барокко», «комический», «странствование», «роман», «жанровый», «художественно», «достоевский», «ответный», «вопрос», «испытание», «философский», «проза», «художник», «материал», «эпопея», «словесный», «толстой», «теория», «реалистический», «изображение», «рыцарский», «видение», «функция», «идеологический», «эпос», «поэтика», «сюжетный», «личность», «творчество» и др.

Отдельную группу составляют термины бахтинской поэтики: «авантюрный», «зрелищный», «мироощущение», «выражать», «праздничный», «композиционный», «хронотоп», «сюжет», «карнавал», «форма», «раскрывать», «архитектонический», «сатира», «комик», «карнавальный», «характер», «атмосфера», «описание» и др. Несмотря на то, что с разными лексическими составляющими теории карнавала мы уже встречались в других кластерах, сам термин «карнавал» оказался именно в этой группе вместе с «хронотопом», «сюжетом» и «формой».

Два последних кластера имеют менее отчётливый облик. Здесь мы видим ещё одну грань репрезентации теории карнавала, на этот раз с высокой степенью историко-литературного обобщения: «литература», «средневековый», «стадия», «этап», «источник», «проблема», «искусство», «смех», «средневековье», «традиция», «смеховой», «готический», «сервантес», «ев-

ропейский», «возрождение», «ренессанс», «фольклорный», «канон», «шекспир», «античный», «книга», «фольклор», «эпоха», «пародия», «культура», «народный и др.

Наконец, последний кластер не демонстрирует очевидной концептуальной цельности. Кажется, что сгруппированные здесь термины обслуживают другие понятийные зоны бахтинских текстов по остаточному принципу: «объективный», «правда», «знать», «четкий», «картина», «настоящее», «прошлое», «будущее», «теоретический», «плоскость», «социальный», «аспект», «познание», «модель», «глаз», «официальный», «граница», «эпический», «дантовский», «власть», «мир», «абсолютный», «друг», «мировоззрение» и др.

В заключение сравним содержание сети со словником «Бахтинского тезауруса»¹¹. Какие термины совпадают в списке словника и среди вершин графа?

Многие предлагаемые в тезаурусе термины состоят из нескольких лексем, а в сети единицей является одно слово. Если разбить составные термины словника на отдельные словоформы, то соответствии узлам графа находит 126 лексем: «авантюрный», «автор», «авторский», «акцент», «амбивалентность», «анализ», «биографический», «большой», «будущее», «верх», «веселый», «вещь», «видение», «внутренне», «внутренний», «возможность», «высказывание», «герой», «говорящий», «голос», «граница», «греческий», «гротеск», «гротескный», «действительность», «действие», «диалог», «доминанта», «драматический», «дурак», «жанр», «задание», «зона», «игра», «идея», «интенция», «интонация», «исповедь», «истина», «исторический», «канон», «карнавал», «карнавальный», «категория», «контакт», «концепция», «кругозор», «культура», «лазейка», «литература», «личность», «материал», «менишья», «мир», «мироощущение», «мистерия», «миф», «мотив», «начало», «необходимость», «низ», «образ», «объект», «оценка», «память», «пародия», «персонаж», «пир», «площадь», «плутовской», «позиция», «полифонический», «постановка», «поступок», «поэтика», «поэтический», «правда», «принцип», «произведение», «разноречие», «рассказ», «рассказчик», «реализм», «речь», «роль», «роман», «романтический», «рыцарский», «самосознание», «свобода», «серьезность», «система», «слово», «смех», «смысл», «событие», «содержание», «сознание», «сократический», «сон», «стилизация», «стилистика», «стиль», «структура», «сюжет», «текст», «тело», «тип», «тон», «точка», «установка», «философский», «фольклорный», «форма», «функция», «характер», «хронотоп», «целое», «ценность», «читатель», «чужой», «шут», «эпопея», «эстетика», «эстетический», «язык».

По всей видимости, эти лексические единицы охватывают понятийное ядро системы М. М. Бахтина как с точки зрения текстового воплощения, так и с рецептивной точки зрения.

Из обращающих на себя внимание терминов, присутствующих в словнике, но не попавших в сетевую модель, отметим следующие: «апперцептив-

¹¹Словник «Бахтинского Тезауруса» (проект) / Сост. Н. Д. Тамарченко // Бахтинский тезаурус. М.: РГГУ, 1997. С. 7–16.

ный», «внесюжетный», «время», «выбор», «диалогизирующий», «диахрония», «древнехристианский», «завершенность», «заданность», «заочный», «идеология», «инверсия», «канонизация», «карнавализация», «конструкция», «литературность», «материя», «мезальянс», «модус», «натурализм», «невоплощенный», «незавершенность», «обман», «опредмечивание», «переакцентуализация», «плут», «полифония», «просветительский», «профанация», «редуцированный», «ритуальный», «самокритика», «сентиментальность», «символика», «синхрония», «ситуация», «счастье», «тематика» «трехмерность».

Вместо терминов «плут», «полифония», «карнавализация», «идеология» в сети мы встречаем однокоренные. Значимо отсутствие в сети слова «время» при наличии термина «пространство».

Редакция «Бахтинского тезауруса» в 1997 году констатировала: «Итак, признаемся себе, что мы еще плохо понимаем М. М. Бахтина, не видим его как целое, не овладели его научным языком и системой понятий»¹². Хочется верить, что движение от непонимания к пониманию в этой сфере может быть поддержано таким мощным инструментом, как сетевой анализ. При определённых в начале исследования степенях свободы в отношении рёбер между узлами графа мы получили довольно общую модель, которая может быть усовершенствована и более тонко настроена. В частности, понижение порога для связи понятий введёт в систему другие, менее частотные термины, позволит установить их соседей (то, что в «Бахтинском тезаурусе» называется «семантическим гнездом») и определить их место при кластеризации вершин.

Но уже сейчас мы можем сказать, что центральными терминами бахтинской понятийной сферы являются «образ», «роман», «слово», «мир», «форма», «герой», «язык», а центральной концепцией, состоящей, в свою очередь из нескольких вполне оформленных автономных частей, — концепция карнавала в литературе Возрождения.

¹²Бахтинский тезаурус. С. 5.