

ДМИТРИЙ ВАДИМОВИЧ КУЗНЕЦОВ

магистрант математического факультета, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
kuznecov@cs.karelia.ru

АЛЕКСАНДР АЛЕКСАНДРОВИЧ ЛЕБЕДЕВ

магистрант филологического факультета, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
perevodchik88@yandex.ru

НИКОЛАЙ ДМИТРИЕВИЧ МОСКИН

кандидат технических наук, доцент кафедры информатики и математического обеспечения математического факультета, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
moskin@karelia.ru

АЛЕКСЕЙ ГЕННАДЬЕВИЧ ВАРФОЛОМЕЕВ

кандидат физико-математических наук, доцент кафедры теории вероятностей и анализа данных, декан математического факультета, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
avarf@petrsu.ru

ТЕОРЕТИКО-ГРАФОВЫЕ МОДЕЛИ С УПОРЯДОЧЕННОЙ ИЕРАРХИЧЕСКОЙ СТРУКТУРОЙ И ИХ ИСПОЛЬЗОВАНИЕ В АНАЛИЗЕ СИНТАКСИСА ПОЭТИЧЕСКИХ ТЕКСТОВ*

В статье рассматриваются теоретико-графовые модели с упорядоченной иерархической структурой и методы их сравнительного анализа. Сравнение моделей выполняется с помощью определения расстояния, основанного на известном алгоритме сравнения строк Вагнера – Фишера. При этом теоретико-графовая модель преобразуется в строковое представление, где каждой вершине ставится в соответствие ссылка на родительский элемент, уровень вложенности и тип связи с этой вершиной. Для определения степени схожести типов синтаксических связей на основе экспертных оценок филологов была построена матрица S размером 43×43 со значениями от 0 до 1. Апробация данных алгоритмов проводилась на моделях синтаксической структуры сложных предложений из творчества П. А. Вяземского и И. А. Бродского (118 текстов). В ходе исследования была обнаружена довольно заметная зависимость значений метрики от авторства предложений. Зависимость выражается сильнее при исследовании больших предложений с нетривиальной структурой.

Ключевые слова: теоретико-графовая модель, иерархическая структура, сравнение, синтаксис, текст, сложное предложение

В математике и информатике графы являются удобным механизмом представления информации. Однако во многих приложениях (например, при моделировании синтаксической структуры текстов) требуются более сложные теоретико-графовые модели, обладающие иерархической структурой. Известны иерархические графы и иерархические графовые модели, описание которых приводится в [6; 399–408], однако они не в полной мере подходят для анализа текстов. Поэтому возникает потребность в разработке новых моделей и методов их исследования.

Дадим рекурсивное определение *теоретико-графовой модели с упорядоченной иерархической структурой*: это пятерка $G = (V, W, E, \alpha, \beta)$, в которой $V = \{v_i\}_{i=1}^n$ – непустое конечное упорядоченное множество вершин; $W = \left\{ V_0 \cup \left\{ G_j \right\}_{j=1}^m \right\}$ – объединение подмножества вершин $V_0 \subseteq V$

и множества вложенных теоретико-графовых моделей $G_j = (V_j, W_j, E_j, \alpha_j, \beta_j)$, таких что V_j попарно не пересекаются и их объединение совпадает с $V \setminus V_0$; $E \subseteq W \times W$ – множество ребер; $\alpha: V \rightarrow L_V$ – функция, задающая метки вершинам; $\beta: E \rightarrow L_E$ – функция, задающая метки ребрам, где L_V и L_E – множества меток и атрибутов объектов и отношений, определенные в некоторой предметной области (номера элементов включаются в состав их меток). Таким образом, внутри теоретико-графовой модели с упорядоченными узлами существует иерархия вложенных друг в друга подструктур, причем ребра могут соединять как обычные узлы, так и вложенные подструктуры, находящиеся на данном уровне вложенности. На рис. 1 показан пример подобной модели, которая подробно описана в [10].

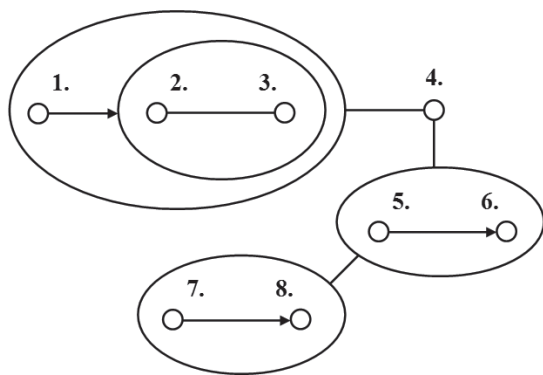


Рис. 1. Пример иерархической теоретико-графовой модели синтаксической структуры предложения «Их гений мужествен, как гений вод твоих...» [3; 83]

В данной работе рассматриваются модели синтаксической структуры многочленных сложных предложений из творчества П. А. Вяземского и И. А. Бродского. Выбор творчества этих поэтов не случаен: именно интеллектуальность поэзии, ее глубина стали отличительной чертой их творчества; стремление к выстраиванию причинно-следственных, глубинных связей, нехарактерное для лирических произведений в целом, отразилось в творчестве этих авторов. Значительную роль в оформлении интеллектуальности и сверхсодержательности поэзии П. А. Вяземского и И. А. Бродского сыграли сложные предложения.

ТЕОРЕТИКО-ГРАФОВАЯ МОДЕЛЬ СИНТАКСИЧЕСКОЙ СТРУКТУРЫ СЛОЖНОГО ПРЕДЛОЖЕНИЯ

Сложное предложение в современной лингвистике является объектом пристального научного интереса. Одна из немаловажных проблем, актуальных для данной сферы исследования, – типологизация сложных предложений, описание их структуры и составляющих. Исследователи-грамматисты отмечают, что «трудности, с которыми сталкивается синтаксическая наука при решении проблем типологии сложного предложения, начинаются уже на первых этапах классификации» [8]. В связи с этим ощущается потребность в создании некоего механизма, позволяющего размечать сложные предложения с точки зрения их структуры, что облегчило бы последующий анализ, установление взаимосвязей между частями предложения, а также определение роли того или иного сложного предложения в тексте произведения или в целом в творчестве писателя. Сложные предложения в стихотворных произведениях могут рассматриваться, с одной стороны, как показатель той эпохи, к которой принадлежит творчество поэта (очевидно, что со временем синтаксическая структура текстов меняется), а с другой стороны, – как проявление идиостиля писателя [7].

В современной теории синтаксиса широко используются так называемые деревья зависимостей и деревья составляющих [4]. Одним из первых их применил А. М. Пешковский [11] для изображения сочинительных и подчинительных связей между словами. Суть деревьев зависимостей сводится к следующему: считается заданным анализ предложения, устанавливающий подчинительные связи между словами согласно правилам традиционной грамматики и грамматики зависимостей. Результаты анализа представляются в виде графов, вершины которых соответствуют словам, а дуги соединяют их в соответствии с синтаксическими связями. В терминах деревьев зависимостей можно выражать многие стилистические характеристики текстов [12].

Деревья зависимостей обычно используются в описаниях языков со свободным порядком слов (например, русского). Для описания языков с фиксированным порядком слов преимущественно используется второй тип графов – деревья составляющих [4]. При этом в предложении выделяются группы слов, функционирующие как отдельные синтаксические единицы, – составляющие. Система составляющих – это множество отрезков предложения, которое обладает тем свойством, что каждые два входящих в него отрезка либо не пересекаются, либо один из них содержится в другом. При графическом изображении система составляющих тоже приобретает вид дерева.

В качестве модели представления структуры сложного предложения нами была выбрана вертикальная структурная схема. По филологической традиции, вертикальные структурные схемы используются в рамках выполнения синтаксического разбора многочленных сложных предложений. С ее помощью можно наблюдать расположение предикативных частей, средства связи между ними, а также направление зависимостей между частями сложного предложения. Вертикальная схема удобна в силу того, что не теряется наглядность направления связей (может быть продемонстрирован уровень глубины соотношения блоков внутри многокомпонентных сложных предложений). Вертикальная схема является оптимальным вариантом для представления тех сложных предложений, которые состоят из четырех и более компонентов. В некоторых случаях в качестве компонента сложного предложения может выступать блок, который объединяет несколько тесно связанных между собой частей. Внутри такого блока при более подробном анализе также можно выстроить систему связей (рис. 2). Обозначим сочинительные связи с помощью прямых линий, подчинительные связи – с помощью стрелок, ведущих от главной части к придаточной.

¹Во мне идет с ожесточеньем
 Борьба враждебных двух начал:
²Ум не скудеет размысленьем,
³Но воли нет, ⁴но дух упал.

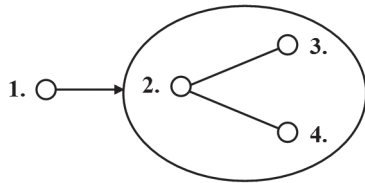


Рис. 2. Теоретико-графовая модель предложения «Во мне идет с ожесточеньем...» [3; 411]

Таким образом, теоретико-графовая модель представляет собой иерархическую структуру с несколькими уровнями вложенности. При этом типы связей делятся на три класса в зависимости от вида союза и характера отношений между частями сложного предложения:

- А. Сложносочиненные предложения (ССП);
- В. Сложноподчиненные предложения (СПП);
- С. Бессоюзные предложения (БСП).

В свою очередь, классы делятся на подклассы типов связей (18 – для ССП, 14 – для СПП и 11 – для БСП). Рассмотрим пример из творчества И. Бродского [2; 80–81], теоретико-графовая модель которого изображена на рис. 3 (слева вверху):

¹Ибо, ²когда расстанутся двое,
³то, перед тем как открыты ворота,
⁴каждый берет у другого что-то
 в память о том, ⁴как их век был прожит:
⁵тело – незримость; ⁶душа, быть может,
 зренья и слух.

Здесь выделяются следующие типы связей (пятая связь соединяет две вложенные подструктуры):

- 1→2: СПП с придаточным времени (BF);
- 1→3: СПП с придаточным времени (BF);
- 3→4: СПП с придаточным изъяснительным (BC);
- 5–6: БСП со значением перечисления (CA);
- [1–4]–[5–6]: БСП со значением пояснения (CD).

ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ ДЛЯ СРАВНИТЕЛЬНОГО АНАЛИЗА ТЕОРЕТИКО-ГРАФОВЫХ МОДЕЛЕЙ

Для сравнения и классификации теоретико-графовых моделей можно применить несколько подходов [9]. Одним из них является вычисление числовых характеристик, отражающих структурные особенности графов. В простейшем случае это параметры размерности – число вершин и ребер графа. К более сложным параметрам можно отнести параметр глубины (число уровней в дереве), параметр ширины (наибольшее

число узлов на уровне), коэффициент асимметрии (логарифм отношения количества узлов, расположенных слева от корня, к количеству узлов, расположенных справа от корня), веточный параметр (отношение числа концевых узлов к числу уровней глубины), параметр вложенности (максимальная степень вложенности структур друг в друга) и др.

Второй подход основан на введении метрики на множестве графов. Здесь задается мера, которая позволяет оценить, насколько те или иные структуры «похожи» друг на друга. В зарубежной литературе данное направление получило название «*graph matching*» [13]. К подобным способам оценки сходства графов относится мера на основе операций редактирования (*graph edit distance*). Эта мера является расширением известного правила сравнения строк Вагнера – Фишера [14]. Здесь строится минимальная последовательность операций редактирования (как правило, это добавление, удаление или переименование вершин и ребер), которая преобразует одну теоретико-графовую модель в другую. При этом на практике часто одни операции являются более значимыми по сравнению с другими, поэтому каждой операции ставится в соответствие ее положительный вес $\gamma: \Sigma \rightarrow R^+$, где Σ – множество операций редактирования. Расстояние вычисляется как минимальная по сумме весов последовательность операций редактирования, преобразующая одну модель в другую.

В данном исследовании был рассмотрен алгоритм нахождения расстояния, который является модификацией алгоритма Вагнера – Фишера нахождения расстояния между строками. Для представления графа в виде строки каждой вершине (простому предложению в составе сложного в порядке следования) ставятся в соответствие следующие параметры (рис. 3):

1. Тип связи между этим предложением и родительским (в силу древоподобной структуры такая связь всегда одна или ее нет для корневых элементов);

2. Ссылка на порядковый номер родительского элемента;

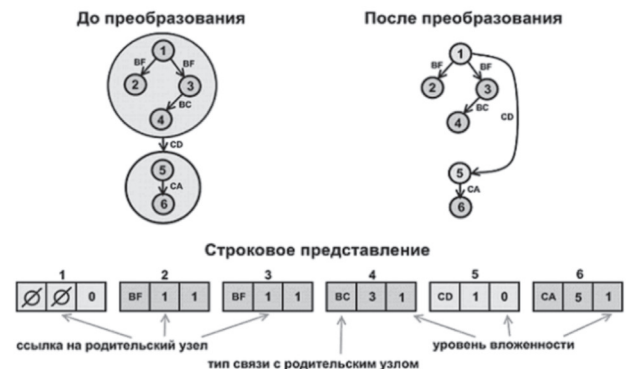


Рис. 3. Строковое представление теоретико-графовой модели

3. Уровень вложенности (равный нулю для внешних узлов структуры и увеличивающийся на единицу для всех вложенных подструктур). Вложенная подструктура на последующем уровне вложенности ведет себя как обычный узел, то есть может иметь входящие и исходящие связи; в то же время она всегда имеет корень, у которого отсутствует непосредственный родительский узел.

Таким образом, вычисление стоимости операций проводилось не только в зависимости от меток узлов, но и при помощи функций от таких параметров, как уровень вложенности и связь с другими узлами. Заметим, что операции редактирования, изменяющие уровень вложенности узлов, являются недопустимыми, что позволяет вычислять расстояния для всех уровней вложенности за одну итерацию алгоритма, без рекурсии.

Заметим, что определенное таким образом расстояние удовлетворяет условиям метрики [14]:

1. $d(G_1, G_2) \geq 0$ (неотрицательность): выполняется, поскольку сумма положительных весов является положительным числом;

2. $d(G_1, G_2) = 0 \Leftrightarrow G_1 = G_2$;

3. $d(G_1, G_2) = d(G_2, G_1)$ (симметричность): обратная последовательность операций редактирования, преобразующая G_2 в G_1 , совпадает с исходной и имеет тот же суммарный вес;

4. $d(G_1, G_2) \leq d(G_1, G_3) + d(G_3, G_2)$ (неравенство треугольника): расстояние $d(G_1, G_2)$ определяется как минимальная по стоимости последовательность операций, преобразующая структуру G_1 в структуру G_2 . Значит, любая последовательность, преобразующая структуру G_1 в G_3 , а затем G_3 в G_2 (то есть с заданным промежуточным состоянием), по определению имеет не меньшую стоимость, чем $d(G_1, G_2)$.

Алгоритм Вагнера – Фишера заключается в последовательном заполнении матрицы $m \times n$ промежуточными значениями расстояний, вычисляемых по формуле:

$$D(i, j) = \begin{cases} \max(i, j) & i = 0 \text{ или } j = 0 \\ \min(D(i, j-1)+1, D(i-1, j)+1, \\ D(i-1, j-1)+\gamma(S_1[i], S_2[j])); & i = 1, \dots, m, j = 1, \dots, n \end{cases}$$

где $\gamma(S_1[i], S_2[j])$ – функция стоимости замены элемента $S_1[i]$ на $S_2[j]$, и равна нулю, если эти элементы идентичны, m и n – количество простых элементов в G_1 и G_2 соответственно. В матрице шаг по i символизирует удаление элемента из первой строки, по j – вставку в первую строку, а шаг по обоим индексам одновременно символизирует замену элемента или отсутствие изменений, если элементы равны.

Для определения степени схожести типов синтаксических связей совместно с экспертами-филологами была построена матрица C разме-

ром 43×43 со значениями от 0 до 1. Вопросы, связанные с синкретизмом сложных предложений, промежуточными видами между типами и видами сложных предложений, входят в сферу интереса филологов [1], [5]. В рамках данного исследования невозможно провести полный и обстоятельный анализ трансформации между разными типами связей с привлечением большого числа примеров и контекстов, однако вместе с тем в основу таблицы коэффициентов положены определенные выводы, связанные с анализом структуры различных сложных предложений. В связи с этим были рассмотрены два вопроса: 1) трансформация одной связи в другую в пределах одного класса; 2) трансформация для связей разных классов.

Для элементов одного уровня стоимость операции редактирования мы определяем по формуле:

$$(a, b) = 1 - C[a, b] \cdot l(a, b),$$

где $C[a, b]$ – степень схожести типов синтаксических связей a и b ; $l(a, b)$ – функция со значениями от 0 до 1, принимающая значение 1, если родительский элемент не меняется (не производится перенос элемента), и значение меньше 1, если родительский элемент a не равен родительскому элементу b . В этом случае мы можем либо определить значение как 0, либо сделать его зависимым от расстояния, на которое узел переносится. Для этого была выбрана функция следующего вида:

$$l(a, b) = \text{const} \cdot |link_a - link_b|,$$

где $|link_a - link_b|$ – модуль разницы между номерами родительских узлов, а const – константа от 0 до 1 (в нашем случае взята 0,5), характеризующая гибкость данной метрики относительно переносов узлов на небольшие расстояния.

В результате исследования была построена выборка из 61 предложения из творчества П. А. Вяземского, 57 предложений И. А. Бродского и их теоретико-графовых моделей. В результате работы алгоритма была построена матрица расстояний размером 118×118 . Эту матрицу можно использовать, например, для выявления возможной зависимости расстояний между структурами предложений от авторства этих предложений. На рис. 4 представлена диаграмма средних значений и отклонений расстояний: между всеми предложениями выборки (а), только между предложениями П. А. Вяземского (б), только между предложениями И. А. Бродского (с) и перекрестное расстояние между предложениями П. А. Вяземского и И. А. Бродского (д). Здесь средние значения представлены точками в центре отрезков, характеризующих среднее отклонение расстояний.

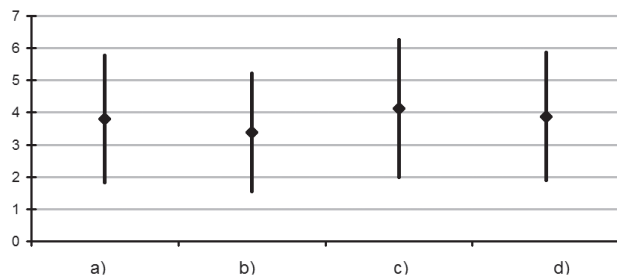


Рис. 4. Расстояния между всеми предложениями

Кроме того, было проведено исследование расстояния между предложениями с достаточно сложной структурой, которые содержали более чем 4 части. В результате получилась диаграмма, изображенная на рис. 5.

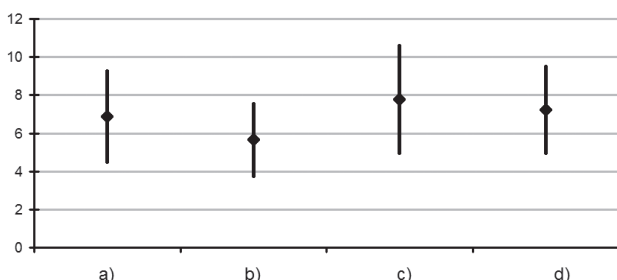


Рис. 5. Расстояния между предложениями с более чем 4 частями

Как видно, значения расстояний между предложениями в случаях (b) и (c) заметно отличаются друг от друга, причем это отличие проявляется более ярко для предложений с достаточно сложной структурой. Можно сделать предположение, что И. А. Бродский использовал более

широкий спектр синтаксических структур предложений, чем П. А. Вяземский, причем это разнообразие структур становится заметнее с увеличением степени сложности предложения. Для проверки этой гипотезы необходимо сравнить расстояния для других выборок предложений из стихотворений тех же авторов, кроме того, интерес представляет изучение других способов определения расстояний между моделями предложений.

ЗАКЛЮЧЕНИЕ

В данной статье рассмотрены теоретико-графовые модели с упорядоченной иерархической структурой и методы их сравнительного анализа. Апробация алгоритмов проводилась на моделях синтаксической структуры сложных предложений из творчества П. А. Вяземского и И. А. Бродского. В ходе исследования была обнаружена заметная зависимость значений расстояний между предложениями от их авторства. Зависимость выражается сильнее при сравнении больших предложений с нетривиальной структурой. Относительно большое внутреннее расстояние между предложениями И. А. Бродского означает большее разнообразие между структурами предложений и, вероятно, говорит о разнообразии и непостоянстве авторского стиля. Относительно малое расстояние между предложениями П. А. Вяземского, напротив, говорит о большей приверженности автора к определенному стилю. В дальнейшем данное исследование планируется продолжить с привлечением других выборок предложений, а также структурных параметров теоретико-графовых моделей.

* Работа выполнена при поддержке Программы стратегического развития ПетрГУ в рамках реализации комплекса мероприятий по развитию научно-исследовательской деятельности на 2012–2016 гг.

СПИСОК ЛИТЕРАТУРЫ

1. Бабайцева В. В. Переходные конструкции в синтаксисе. Воронеж: Центрально-Черноземное изд-во, 1967. 392 с.
2. Бродский И. А. Сочинения: В 4 т. СПб.: Пушкинский фонд, 1994. Т. 2. 479 с.
3. Вяземский П. А. Стихотворения. Л.: Сов. писатель, 1986. 544 с.
4. Гладкий А. В. Синтаксические структуры естественного языка. М.: Изд-во ЛКИ, 2007. 152 с.
5. Дружинина С. И. Синкретизм в системе сложноподчиненных предложений. Орел: Изд-во Орел ГАУ, 2008. 436 с.
6. Касьянов В. Н., Евстигнеев В. А. Графы в программировании: обработка, визуализация и применение. СПб.: БХВ-Петербург, 2003. 1104 с.
7. Коломийцева А. С. Полипредикативное сложное предложение-абзац как проявление идиостиля писателя и способ реализации текстовой категории «образ автора» // Историческая и социально-образовательная мысль. 2011. № 3. С. 108–111.
8. Мишланов В. А. Семантика и структура сложного предложения в сфере динамического синтаксиса. Пермь: Изд-во Пермского ун-та, 1996. 267 с.
9. Москвин Н. Д. Теоретико-графовые модели структуры фольклорных текстов, алгоритмы поиска закономерностей и их программная реализация: Дис. ... канд. техн. наук. Петрозаводск, 2006. 121 с.
10. Москвин Н. Д., Лебедев А. А., Варфоломеев А. Г. Представление моделей синтаксической структуры поэзии П. А. Вяземского с помощью технологии XML // Материалы IV междунар. науч. конф. «Информационные технологии и письменное наследие – EΓManuscript-2012» (Петрозаводск, 3–8 сентября 2012 г.). Петрозаводск; Ижевск, 2012. С. 174–178.
11. Пешковский А. М. Русский синтаксис в научном освещении. М.: Эдиториал УРСС, 2001. 450 с.
12. Сево И. П. Графическое представление синтаксических структур и стилистическая диагностика. Киев: Наукова думка, 1981. 192 с.
13. Bunke H. Graph matching: theoretical foundations, algorithms, and applications // Proc. Vision Interface. Montreal, 2000. P. 82–88.
14. Wagner R., Fischer M. The string-to-string correction problem // Journal of the Association for Computing Machinery. 1974. Vol. 21. № 1. P. 168–173.

Kuznetsov D. V., Petrozavodsk State University (Petrozavodsk, Russian Federation)
Lebedev A. A., Petrozavodsk State University (Petrozavodsk, Russian Federation)
Moskin N. D., Petrozavodsk State University (Petrozavodsk, Russian Federation)
Varfolomeyev A. G., Petrozavodsk State University (Petrozavodsk, Russian Federation)

GRAPH-THEORETIC MODELS WITH ORDERED HIERARCHICAL STRUCTURE AND THEIR USE IN ANALYSIS OF POETIC TEXTS' SYNTAX

The article examines graph-theoretic models with ordered hierarchical structures and methods of their comparative analysis. A comparison of the models is carried out by the distance definition, which is based on the well known Wagner-Fischer string comparison algorithm. Here, a graph-theoretic model is converted into a string, where each node is assigned a link to the parent element, the nesting level, and type of connection with the node (this distance satisfies all metric properties). To determine the degree of similarity between different types of syntactic relations on the basis of the expert assessments of philologists the matrix C sized 43×43 with values between 0 and 1 was built. Testing of these algorithms was performed on the models of syntactic structure of complex sentences taken from the works of P. A. Vyazemskiy and I. A. Brodskiy (118 texts). Noticeable dependence between metric values and authorship of the sentences was revealed. This dependence is expressed more strongly in the analysis of large sentences with non-trivial structure.

Key words: graph-theoretic model, hierarchical structure, comparing, syntax, text, complex sentence

REFERENCES

1. Babaytseva V.V. *Perekhodnye konstruksii v sintaksise* [Transitional constructions in syntax]. Voronezh, The Central Black Earth Publ., 1967. 392 p.
2. Brodskiy I.A. *Sochineniya: V 4 t.* [Works: In 4 vol.]. St. Petersburg, Pushkin fund Publ., 1994. Vol. 2. 479 p.
3. Vyazemskiy P.A. *Stikhotvoreniya* [Poems]. Leningrad, Sovetskiy pisatel' Publ., 1986. 544 p.
4. Gladkiy A.V. *Sintaksicheskie struktury estestvennogo yazyka* [The syntactic structures of natural language]. Moscow, LKI Publ., 2007. 152 p.
5. Druzhinina S.I. *Sinkretizm v sisteme slozhnopodchinennykh predlozheniy* [Syncretism in the system of complex sentences]. Orel, Orel GAU Publ., 2008. 436 p.
6. Kasyanov V.N., Evstigneev V.A. *Grafy v programmirovanii: obrabotka, vizualizatsiya i primeneniye* [Graphs in programming: processing, visualization, and application]. St. Petersburg, BHV-Petersburg Publ., 2003. 1104 p.
7. Kolomyitseva A.S. Polypredicative compound sentence-paragraph as realization of writer's idyostyle and method of realization of textual category of "author's character" [Polipredikativnoe slozhnoe predlozhenie-abzats kak proyavlenie idostilya pisatelya i sposob realizatsii tekstovoy kategorii "obraz avtora"]. *Istoricheskaya i sotsial'no-obrazovatel'naya mysl'* [Historical and social-educational idea]. 2011. № 3. P. 108–111.
8. Mishlanov V.A. *Semantika i struktura slozhnogo predlozheniya v sfere dinamicheskogo sintaksisa* [Semantics and complex sentence structure in dynamic syntax]. Perm', Perm University Publ., 1996. 267 p.
9. Moskin N.D. *Teoretiko-grafovye modeli struktury folklornykh tekstov, algoritmy poiska zakonomernostey i ikh programmnaya realizatsiya. Diss. kand. tekhn. nauk* [Graph-theoretic models of the folklore texts' structure, algorithms for regularity searching and their program implementation. Cand. tech. sci. diss.]. Petrozavodsk, 2006. 121 p.
10. Moskin N.D., Lebedev A.A., Varfolomeyev A.G. Presentation of syntactic structure models of P.A. Vyazemsky poetry using XML technology [Predstavlenie modeley sintaksicheskoy struktury poezii P.A. Vyazemskogo s pomoshch'yu tekhnologii XML]. *Proceedings of the IV International conference «Informatsionnye tekhnologii i pismennoe nasledie»* [Information Technologies and Textual Heritage] – EI Manuscript-2012». (Petrozavodsk, 3–8 September 2012). Petrozavodsk; Izhevsk, 2012. P. 174–178.
11. Peshkovskiy A.M. *Russkiy sintaksis v nauchnom osveshchenii* [Russian syntax in scientific light]. Moscow, Editorial URSS Publ., 2001. 450 p.
12. Sevbo I.P. *Graficheskoe predstavlenie sintaksicheskikh struktur i stilisticheskaya diagnostika* [Graphic presentation of syntactic structures and stylistic diagnosis]. Kiev, Naukova Dumka Publ., 1981. 192 p.
13. Bunke H. Graph matching: theoretical foundations, algorithms, and applications // *Proc. Vision Interface*. Montreal, 2000. P. 82–88.
14. Wagner R., Fischer M. The string-to-string correction problem // *Journal of the Association for Computing Machinery*. 1974. Vol. 21. № 1. P. 168–173.

Поступила в редакцию 03.06.2013