

## ФИЗИЧЕСКИЕ И СИНТАКСИЧЕСКИЕ ТЕКСТОВЫЕ СТРУКТУРЫ В АСПЕКТЕ ПРОСТРАНСТВЕННОЙ МОДЕЛИ ТЕКСТА

**В работе с позиции системнодеятельностного подхода обосновывается пространственная модель текста и в рамках данной модели доказывается отсутствие изоморфизма между организацией разнопредметных текстовых пространств.**

С позиции системнодеятельностного подхода текст, как любой объект, имеет многообразные предметные области (проекции), выделяемые в конкретных деятельности с объектом. Каждая из предметных областей (назовем их *пространствами текста* или, что то же самое, *текстовыми пространствами*) может быть представлена в виде структуры. Таким образом, в деятельности над текстом (и / или с текстом) мы в первую очередь вычлняем текстовое пространство, которое затем представляем как систему со своими компонентами и структурой. Поскольку список предметных проекций текста не может быть закрытым, возникает теоретическая проблема, можно ли создать такую модель текста, в рамках которой осуществлялся бы синтез структур разнообразных текстовых пространств? Положительный ответ на этот вопрос мы связываем с вводимым в исследовательское пространство общей теории текста понятием *формы текста*.

Форма текста представляет собой способ организации и саму организацию языкового субстрата в целостном пространстве речемыслительной деятельности. *Если текст как объект-феномен имеет своими предметами текстовые пространства, то форма текста – структуры данных пространств. Структура – это проекция формы на предметную область.* Форма обеспечивает необходимую общность структур предметов одного и того же объекта. Если структура – производная исследовательской деятельности, конструкт, то форма, как и текст в целом, – феномен (см. рис. 1). Полагаем, что форма текста в нашей

интерпретации является новым лингвистическим объектом<sup>1</sup>.

Признание текста полипредметным и, соответственно, полиструктурным, языковым объектом вызывает ряд новых проблем: в каком отношении находятся текстовые пространства по отношению друг к другу? Какой метод может быть применен для осуществления полипредметного (полиструктурного) синтеза?

Для того чтобы ответить на первый вопрос, необходимо реконструировать деятельностные схемы, производимые всякий раз при изучении того или иного текстового пространства. Такая реконструкция выявит, что пространства текста (а значит, и их структуры) могут находиться по отношению друг к другу в двух планах: 1) уровне и 2) независимом. Так, традиционная знаково-уровневая модель текста включает в себя ряд тексто-

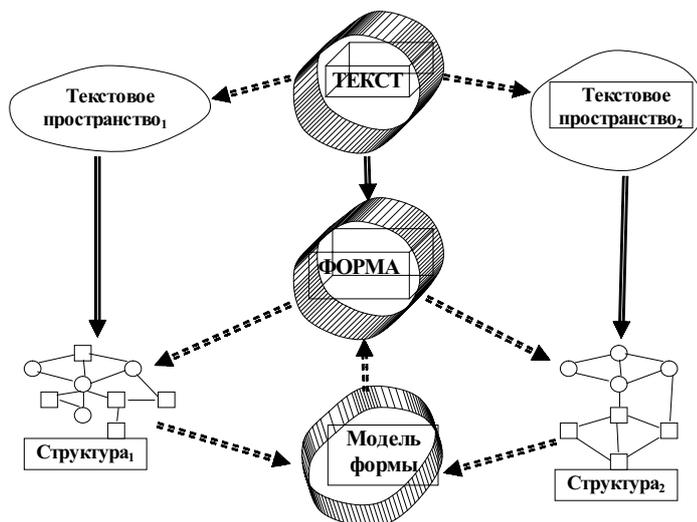


Рисунок 1. Взаимосвязь основных конструктов теории формообразования текста

<sup>1</sup> «Новым лингвистическим объектом я называю такое представление фактов, языковых данных, которое генерирует новую информацию о языке» [3, с. 140].

вых пространств, находящихся (в целом) в отношениях иерархической упорядоченности: фонетический, словообразовательный, лексический, морфологический и синтаксический уровни. Вместе с тем далеко не все текстовые пространства могут быть представлены в виде уровневой модели, например просодическое, эмоциональное и семантическое пространства текста (подробнее см.: [1]).

Критика знаково-урвневой модели текста звучала еще в 1980-х годах: «В последнее время в лингвистике широко распространилось убеждение, что наряду с традиционно выделяемыми фонологическим, морфологическим, синтаксическим и лексико-семантическими структурными уровнями существуют также и другие, которые не могут быть просто включены в иерархию «урвней языковой системы» в качестве дополнительных урвней, но, напротив, лежат «вне» этой системы урвней и иным, по-видимому, более непосредственным образом соотносятся с моментами экстралингвистической действительности» [10, с. 55-56]. В последние годы критика урвневой модели языка и текста особенно усиливается (одна из последних ярких книг в этой области принадлежит Б.М. Гаспарову [2]).

В качестве альтернативы предлагается модель текста, которую мы обозначили как *пространственная* модель. Она является более общей, по отношению к ней урвневая модель – частное проявление. Собственно, только в рамках пространственной модели можно осуществлять полипредметный синтез текста, реконструировать форму текста. Для этого необходимо на первом этапе так структурировать разнопредметные текстовые проекции (пространства), чтобы выявленные структуры обладали качеством сопоставимости, были бы построены на основе применения одних и тех же методов и принципов. Базовыми методами выступают метод позиционного анализа текста и метод контуров текста<sup>2</sup>, основополагающим принципом является квантитативный характер создаваемой модели.

Суть метода позиционного анализа состоит в позиционировании интересующих иссле-

дователя языковых единиц в линейном ряду всех единиц текста. Поскольку текст имеет границы (начало и конец), начало принимается за «0», а конец за «1» независимо от размера текста. Это позволяет сопоставлять как тексты разных объемов, так и разные аспекты (текстовые пространства) одного текста. За единицу счета принимается графическое слово (слово в тексте между пробелами). Чтобы локализовать какой-либо элемент на отрезке (0; 1), нужно разделить порядковый номер слова на общее количество слов в тексте. Позиционная локализация компонентов текста – один из основных *независимых* параметров, описывающих динамику становления текста как целостности. Метод контуров предполагает единообразное построение двух и более контуров текста, каждый из которых отражает сукцессивную развертку структуры того или иного текстового пространства, и синхронизацию контуров на одном позиционном отрезке (0, 1).

*Предметом* работы является сопоставление двух текстовых пространств и их структур, выделенных в рамках пространственной модели. В отличие от ранее проводимых сопоставлений текстовых пространств (семантического, просодического, эмоционального) в центре внимания данной статьи будут находиться текстовые пространства, структурируемые на основании *общего* объекта – предложения в тексте. В одном случае предложение будет изучаться со стороны своей физической протяженности в тексте, в другом – в аспекте синтаксического описания (синтаксической модели).

Исследование ритма размера предложений в тексте (*физическое пространство*) проводится с опорой на метод позиционного анализа, позволяющий локализовать каждое предложение в границах всего текста (0, 1). В итоге текст разбивается на ряд отрезков, соответствующих длине каждого предложения в тексте в порядке их следования от начала к концу. В основе метода лежит *положение о необходимости рассмотрения размеров предложений только в составе целого текста*. Ра-

<sup>2</sup> Под *контуром текста* понимается графическая фиксация распределения вероятности появления языковых единиц или значений интенсивности (а также любых других показателей) тех или иных процессов, производимая поступательно от абсолютного начала текста к его абсолютному концу.

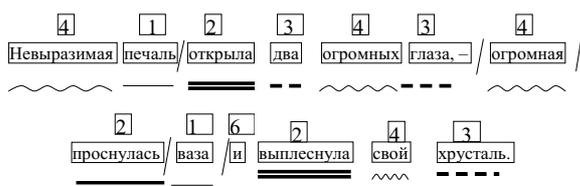
боты, в которых качественные параметры размеров предложений устанавливаются на основании иных факторов, например короткие до 10 слов, средние до 30 слов и т. д., не могут выйти на уровень исследования *текста*. Говорить о коротких, средних и больших размерах предложения можно только относительно того *целого*, в котором эти предложения *содержатся*. В этом смысле априорные шкалы, структуры и стандарты при исследовании ритмов размеров предложений в тексте становятся ненадежными. Границы же размеров всех обозначенных типов предложений могут быть установлены статистически на основании определения среднего размера предложения ( $x_{cp.}$ ) в тексте и среднеквадратичного отклонения ( $y$ ):  $x_{cp.} \pm y (2y)$ . Таким образом, для каждого типа предложения существует определенный коридор значений, в рамках которого предложение в тексте сохраняет свою качественную определенность (оставаясь кратким, средним, длинным и т. д.). В то же самое время количественная граница полагает качественный переход – «предел, за которым изменение количества влечет за собой изменение качества и наоборот» [9, с. 360].

*Синтаксическое пространство текста* строится на основе «традиционного» структурного синтаксического описания, выбор которого обусловлен большими возможностями для формализации, относительной простотой анализа и меньшим исследовательским произволом по сравнению с другими вариантами описания (коммуникативным, семантическим, трансформационным и др.) в случае обращения к *корпусам* текстов. При этом используется предельно обобщенный вид структурного описания синтаксиса, что оправдано объектом исследования, которым является не предложение, а текст. Организацию синтаксического пространства текста мы стремимся представить в виде *конечной последовательности переходящих друг в друга синтаксических групп (синтаксических звеньев текста), список которых обозрим и статистически упорядочен*. Переходы от одного синтаксического звена к другому, вероятно, подчиняются статистическим законам / закономерностям самоорганизации текста. Задача поиска таких закономерностей предполагает проведение исследования на текстовых

*выборках*, что вынуждает нас использовать в качестве единиц анализа *предельно обобщенные формализованные* синтаксические модели. Для обозначения *компонентов* структурных моделей нами применяется цифровая символика: цифрой «1» обозначается подлежащее; «2» – сказуемое; «3» – дополнение; «4» – определение; «5» – обстоятельство; «6» – остальное. Таким образом, для выполнения поставленных задач мы будем использовать цифровую комбинаторику в качестве основы для получения *репертуара элементарных структур*.

На первом этапе построения синтаксического пространства текста возникает задача сегментации числовой цепи текста, поскольку сочетание даже из 20 чисел (текст из 20 слов) имеет вероятность повторения, близкую к нулю. Сегментация цепи может проводиться по двум основаниям: 1) принадлежности рассматриваемых членов предложения одному и тому же предложению и 2) отнесенности того или иного члена предложения к группе: а) подлежащего, б) группе сказуемого. Слова, не являющиеся членами предложения, определяются в отдельную группу. Таким образом, в основе сегментации синтаксической цепочки текста лежит естественное его дробление на предложения и в рамках каждого предложения членение на звенья, каждое из которых состоит из элементов, связанных друг с другом подчинительной или сочинительной связью. Компоненты предложения, не входящие в выделяемые таким образом синтаксические звенья, обособляются в отдельные группы. В русском тексте компоненты, относящиеся к группам подлежащего или сказуемого, могут чередоваться в процессе *поступательной развертки* предложения. Тем самым возникает дополнительное членение синтаксических групп, образование более дробных звеньев. Рассмотрим сказанное на примере одного из предложений, входящего в стихотворение «Невыразимая печаль...» О.Э. Мандельштама:

*Невыразимая печаль  
Открыла два огромных глаза, –  
Огромная проснулась ваза  
И выплеснула свой хрусталь.*



Знаком « / » обозначена сегментация предложения, проведенная в соответствие с описанными принципами членения. Таким образом, вся синтаксическая цепочка предложения состоит из ряда звеньев и может быть сведена к следующей формуле: 452-41-2343-4-1-6243. Обращает на себя внимание растущая дискретность, связанная с инверсией сказуемого *проснулась*, которое «вклинивается» в состав группы подлежащего. Это приводит к тому, что вместо двух нормативных компонентов (звеньев): *огромная ваза / проснулась и выплеснула свой хрусталь* – возникает четыре звена: *огромная / проснулась / ваза / и выплеснула свой хрусталь*.

Уже приведенный пример свидетельствует о том, что наиболее частотные размеры элементарной структуры (синтаксического звена) включают в себя от одного до четырех компонентов. Проанализировав две выборки стихотворений (по 30 в каждой) В.Ф. Ходасевича и О.Э. Мандельштама, мы выделили в первой 1093, а во второй 628 элементарных структур. Результаты распределения встречаемости разнообъемных элементарных структур отражены в таблице 1.

Из таблицы 1 видно, что частотность появления каждой структуры в двух выборках не имеет заметных отличий, что позволяет говорить о единообразии распределения. 1-, 2- и 3-компонентные модели встречаются наиболее часто и в сумме «покрывают» 85 % употребления всех структур как в выборке стихов В. Ходасевича, так и О. Мандельштама. Наиболее частотной оказывается 1-компонентная структура, на нее приходится более половины всех встретившихся структур. Так как в основе каждой синтаксической структуры лежит принцип иерархизации – подчинения главному компоненту звена второстепенных (между которыми также могут быть установлены иерархические отношения), 1-компонентная структура и ее доминирование представляется весьма интересным.

Однокомпонентные структуры обладают еще одним важным свойством – они выражают максимальную величину синтаксической дискретности. Любая структура, состоящая более чем из одного компонента, уже не только дискретизирует текст, но и участвует в его континуализации. Действия дискретности и континуума – противоположные, но взаимосвязанные тенденции, в полной мере обнаруживающие себя в тексте. Но если «строевая» единица дискретности существует – это однокомпонентная структура, то «строевой» единицы континуума нет, т.к. такая структура должна вместить в себя все предложение и весь текст. Можно лишь говорить, что чем объемней структура, тем более проявляется в ней тенденция континуализации.

Понимая таким образом механизмы проявления дискретности и континуальности на уровне синтаксической организации текста (синтаксического текстового пространства), мы можем сопоставить дискретность / континуальность синтаксическую с физической. Как отмечалось, распределение дискретности / континуальности физического пространства текста устанавливается в ходе позиционного маркирования границ предложений в порядке их следования в тексте. Выделенным интервалам присваивается значение, равное величине размера предложения,

Таблица 1. Частотность употребления разнообъемных элементарных синтаксических структур

| Размер элементарной структуры | Стихотворения В. Ходасевича | Стихотворения О. Мандельштама |
|-------------------------------|-----------------------------|-------------------------------|
| 1-компонентная                | 617 (0,565)                 | 328 (0,522)                   |
| 2-компонентная                | 196 (0,179)                 | 121 (0,193)                   |
| 3-компонентная                | 119 (0,109)                 | 85 (0,135)                    |
| 4-компонентная                | 59 (0,054)                  | 39 (0,062)                    |
| 5-компонентная                | 27 (0,025)                  | 23 (0,037)                    |
| 6-компонентная                | 30 (0,027)                  | 16 (0,025)                    |
| больше 6 компонентов          | 45 (0,041)                  | 16 (0,025)                    |
| всего                         | 1093 (1)                    | 628 (1)                       |

*Примечание.* В таблице отражены два ряда значений: одно значение, выраженное в натуральных числах, характеризует количество встретившихся структур в каждой выборке, другое значение (в рациональных числах) отражает частотность появления каждой структуры. Для сопоставления выборок необходимо использовать значение частотности.

располагающегося в данном интервале. Затем все значения из абсолютных переводятся в относительные (максимальное значение равно «1» и принадлежит самому протяженному предложению текста). После этого вычисляются средние значения распределения дискретности / континуальности всей выборки.

Характеристики дискретности / континуальности синтаксического пространства текста выявляются на основе протяженности сукцессивной цепочки элементарных структур. После того, как синтаксическое пространство текста записывается в виде последовательности элементарных структур, с помощью метода позиционного анализа границы этих структур маркируются на позиционной оси (0, 1) подобно тому, как фиксируется локализация границ предложения физического пространства. Каждому сегменту отрезка (0, 1), которое занимает та или иная элементарная структура, присваивается значение, равное количеству графических слов в данной структуре. Затем значения из абсолютных переводятся в относительные и усредняются на основании данных всей выборки.

В результате за каждой координатой сукцессивной развертки текста закрепляются два значения дискретности / континуальности физического и синтаксического пространств. Полученные данные по выборкам стихотворений В. Ходасевича и О. Мандельштама можно представить графически (рисунки 2 и 3).

Если обратиться к рисунку 2, то можно увидеть, что распределение дискретности / континуальности имеет неслучайный характер. В случае физических и синтаксических структур четко выделяются области минимума (максимальная дискретность) и максимума размеров (максимальная континуальность). Синтаксический контур состоит из трех областей, граница между которыми пролегает по точкам максимальной дискретности: 0 – 0,42; 0,42 – 0,85; 0,85 – 1. Фактически эти области коррелируют с областями зоны начала (0 – 0,382), зоны гармонического центра (0,382 – 0,854), зоны конца (0,854 – 1), выделенными в ходе исследования процессов формообразования текста в аспекте распределения в нем элементов симметрии [4; 7]. Распределение дискретности / континуальности

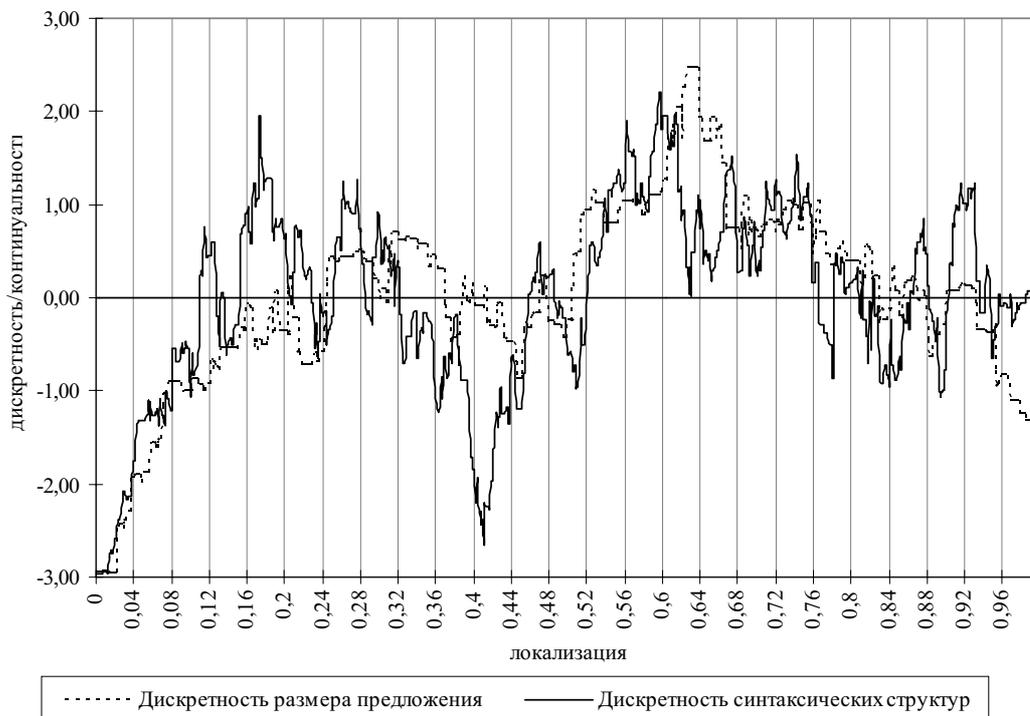


Рисунок 2. Дискретность / континуальность текста на уровне размера предложения и синтаксических структур (выборка из 30 стихотворений В. Ходасевича)

Примечание. Все значения центрированы, т.е. распределены относительно средних значений по выборке, принимаемых за «0».

физического пространства текста имеет сходную динамику (различие касается локализации зоны начала – (0 – 0,45)). Представляется интересным усиление континуальности синтаксического и физического текстовых пространств в области гармонического центра текста (0,618), где соответствующие показатели достигают максимальных значений, что является оптимальным с позиции формообразования текста [1, с. 36].

Фрактальность (самоподобие) текста обычно рассматривается как существование структурного изоморфизма между уровнями текста (в рамках его «уровневой» модели). Однако, как таковой, структурный изоморфизм даже в рамках «уровневой» модели текста может и не наблюдаться. Об этом, на наш взгляд, пишет Ю.М. Лотман, вводя в филологическую терминологию категорию «дополнительности»: «...между отдельными структурными уровнями может возникать отношение дополнительности. Например, давно было отмечено, что в пределах определенных поэтических структур ослабление ограничений, наложенных на ритм, сопровождается усилением требований к рифме...» [6, с. 84]. Или более конкретная формулировка: «Можно сказать, что в отношении эквивален-

тности формальные и семантические элементы естественного языка, входя в поэтическую структуру, выступают как дополнительные множества: совпадение одних влечет за собой несовпадение других» [там же, с. 89].

Отсутствие структурного изоморфизма можно наблюдать на графике распределения дискретности / континуальности синтаксического и физического пространств текста стихотворной выборки О. Мандельштама (рисунок 3). Коэффициент корреляции (Пирсона) равен 0,26. Видно, что синтаксическая дискретность последовательно убывает (растет континуальность), затем, начиная с координаты 0,44, входит в стабильное состояние с небольшими флуктуациями. Физическое пространство в аспектах дискретности / континуума имеет более сложную пространственно-временную организацию. Заметим, что и на материале анализируемой выборки выделяется область начала текста (0 – 0,45); в гармоническом центре также отмечается максимальное действие континуума.

Для художественного произведения континуальность означает большее, чем просто синтаксическая усложненность, – *континуальность является, по нашему мнению, средством имитации континуальности реального*

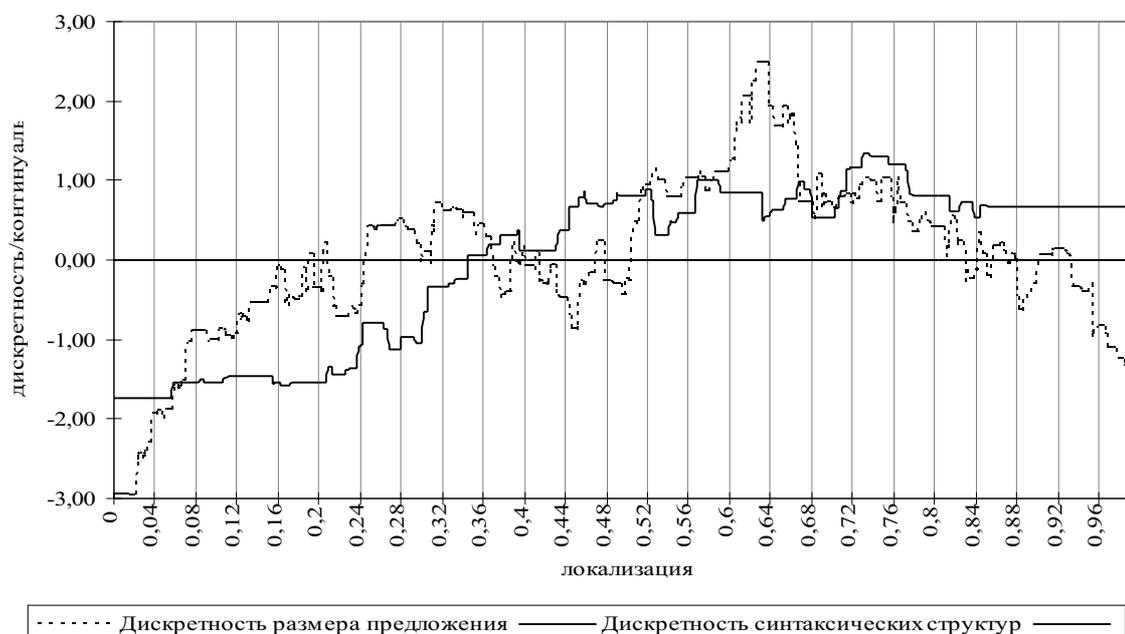


Рисунок 3. Дискретность / континуальность текста на уровне размера предложения и синтаксических структур (выборка из 30 стихотворений О. Мандельштама)

Примечание. Все значения центрированы, т.е. распределены относительно средних значений по выборке, принимаемых за «0».

мира. Действительность непрерывна<sup>3</sup>, и для писателя (поэта) создание универсума художественного текста возможно в первую очередь при условии моделирования непрерывности. Только двигаясь по континууму, можно создать иллюзию бесконечности ограниченного пространства<sup>4</sup>. В таком контексте дискретность / континуальность имеют художественные и коммуникативные функции.

Вхождение в художественную действительность начинается с максимальной дискретности материи текста, которая постепенно и целенаправленно преодолевается, на первый план выдвигается континуальность, которая достигает максимальных значений в области гармонического центра – наиболее оптимальной для восприятия. И то обстоятельство, что в изучаемых выборках в данном интервале находятся наиболее протяженные предложения и по преимуществу наиболее сложные элементарные синтаксические структуры, не является случайностью (аналогичное явление наблюдалось нами и при изучении «Стихотворений в прозе» И.С. Тургенева [1, с. 105-109]). В области гармонического центра сосредоточиваются структурные компоненты того или иного текстового пространства, наиболее значимые с точ-

ки зрения *цельнооформляющей* функции (континуальность и цельность с необходимостью взаимосвязаны). В зоне конца, напротив, наблюдается процесс усиления действия дискретности текстовой материи, что свидетельствует о подготовке воспринимающего сознания к выходу из художественной действительности.

В заключении отметим, что изучение структурного изоморфизма между уровнями текста в большинстве случаев ограничивается либо созданием / использованием «изоморфного» языка описания данных уровней (*фонема, морфема, лексема...*), либо сопоставлением *фрагментов* структур тех или иных уровней. Обычно в практике анализа текста структурный изоморфизм усматривается в тех случаях, когда одни и те же текстовые значения передаются единицами, относящимися к разным уровням текста. Модельное же изучение уровней текста и его структур с позиций а) *целостности предметных областей текста* (к которым относятся и текстовые уровни), б) *общности методов*, с помощью которых можно было бы единообразно подходить к изучению разных текстовых реальностей и в) *единства языка описания* – еще только зарождается.

#### Список использованной литературы:

1. Белоусов, К.И. Текст: пространство, время, темпоритм: монография / К.И. Белоусов. – Новосибирск: Сибирские огни, 2005. – 248 с.
2. Гаспаров, Б.М. Язык, память, образ. Лингвистика языкового существования / Б.М. Гаспаров. – М.: Новое литературное обозрение, 1996. – 352 с.
3. Караулов, Ю.Н. Вторичные размышления об эксперименте в языкознании / Ю.Н. Караулов // Теория языка, методы его исследования и преподавания. – М.: Наука, 1981. – С. 135-140.
4. Корбут, А.Ю. Текстосимметрия: Монография. / А.Ю. Корбут. – Иркутск: Изд-во Иркут. гос. пед. ун-та, 2004. – 200 с.
5. Лосев, А.Ф. В поисках построения общего языкознания как диалектической системы / А.Ф. Лосев // Теория и методология языкознания: Методы исследования языка. – М.: Наука, 1989. – С. 5-92.
6. Лотман, Ю.М. Об искусстве / Ю.М. Лотман. – СПб.: Искусство-СПб, 2000. – 704 с.
7. Москальчук, Г.Г. Структура текста как синергетический процесс / Г.Г. Москальчук. – М.: Едиториал УРСС, 2003. – 296 с.
8. Паршин, А.Н. Размышления над теоремой Геделя / А.Н. Паршин // Вопросы философии. – 2000. – № 6. – С. 92-109.
9. Философский энциклопедический словарь. – М.: Сов. энциклоп., 1983. – 840 с.
10. Хартунг, В. Деятельностный подход в лингвистике: результаты, границы, перспективы / В. Хартунг // Общение. Текст. Высказывание. – М.: Наука, 1989. – С. 41-55.

**Исследование выполнялось при финансовой поддержке Роснауки  
(грант Президента Российской Федерации для государственной поддержки  
молодых докторов наук МД-2549.2008.6).**

<sup>3</sup> Континуум есть «диалектическое слияние раздельной системы со сплошными и непрерывными переходами одного ее момента в другой» [5, с. 23].

<sup>4</sup> Как пишет А.Н. Паршин, «знаменитые парадоксы Зенона, такие как Ахиллес и черепаха, как раз и состояли в том, что если мы ограничиваемся рациональными числами, то движение невозможно <...> движение возможно только по континууму» [8, с. 104].