

ТЕКСТ КАК ИНФОРМАЦИОННЫЙ ПАКЕТ

В статье рассматривается форма текста какместилище информации, необходимой для сохранения или трансляции в социуме. Исходя из целостных представлений о тексте, рассматриваются возможные пути его анализа как информационного пакета.

Естественный язык для тех, кто им пользуется, всегда современен (особенно в его устных репрезентациях). В процессе совместной эволюции человека и его языков складывались и совершенствовались механизмы фиксации значимой для социума информации в виде текстов, позволяющие трансляцию ее во времени и пространстве.

Осмысленная речь всегда организована по принятым носителями языка правилам и способам, изучаемым лингвистикой. При этом всегда (пусть и неосознаваемо) учитываются соотношения эффективности воздействия сообщения с надежностью его трансляции и ретрансляции в социуме и культуре. То есть в самом процессе порождения текста заложена и эволюционно отработана многими поколениями носителей естественных языков коммуникативно-прагматическая целесообразность бытия текста как информационного пакета, упаковки для ценной с точки зрения человека информации. При этом важен целостный подход к тексту как становящейся отдельности – паттерну организации: «Паттерн организации любой системы, живой или неживой, – это конфигурация взаимоотношений между компонентами системы, определяющая существенные характеристики этой системы» [1, с. 176].

Форма текста – это эволюционно-природный конструкт, необходимый для надсинтаксической достройки системы лингвистических уровней и включения текста как многомерного объекта в систему единиц языка. «Упаковка – это процесс отражения информации в той или иной структурной форме и результат (продукт) отражения в виде этой формы» [2, с. 300]. Форма текста рассматривается как вид абстрактной «упаковки» информации, репрезентант целостного гештальта – образа текста. Функции формы заключаются в отделении, изоляции некоторого лингвоментального и перцептивного квазипространства, воспринимаемого и/или интерпретируемого нами как границы данного текста, от иных лингвоментальных пространств.

При весьма широком разбросе мнений и истолкований понятия «информация» от него все же сложно отказаться. Поэтому мы придерживаемся мнения Д.С. Чернавского, А.М. Хазена и др.: «Информация есть запомненный

выбор одного варианта из нескольких возможных и равноправных» [3, с. 13; 4]. «Память» текста формируют сохраняющиеся в нем компоненты и подсистемы. Д.С. Чернавский различает следующие типы информации: микроинформацию, которая тут же «забывается», и макроинформацию, которая «запоминается» системой, воспроизводится в ней [3, с. 13].

Представляется соблазнительным дать количественную оценку информационной емкости всего текста и его частей. Согласно классической теории информации [5, с. 667] количество информации в тексте равно:

$$I = -N \sum_{i=1}^M p_i \log_2 p_i,$$

где N – число символов/букв в тексте, p_i – частота встречаемости i -ой буквы в русском языке, $M = 32$ – число букв в русском алфавите. При «скачивании» текста из Интернета мы будем платить как раз за этот объем информации. Полагая для простоты частоту встречаемости букв одинаковой, т. е. $p_i = 1/32$, получим количество информации в тексте $5N$. В соответствии с работой [6, с. 22-24] это лишь «информационная тара» текста. А вербальной полезной для нас информации в нем может и не быть. Количество же ценной информации зависит не только от качеств текста, но и от реципиента, его целей и установок. Оно может быть нулевым, положительным или даже отрицательным (дезинформация) [3]. Подсчитать это количество пока не представляется возможным. Поэтому необходимо искать иные пути оценивания информационной составляющей текста.

Поскольку уникальный текст вписан в повторяющиеся, стандартные формы организации, то важен вопрос о форме текста как некоемместилище информации и всех структур и процессов речевой деятельности человека. Он может быть решен путем моделирования внутритекстовых пространств, например семантических, просодических, эмотивных и других. Интерес представляет также взаимодействие и взаимное сопряжение собственно языкового субстрата, распределенного в пространстве-времени текста (в его границах от абсолютного начала до абсолютного конца), и различных

составляющих структуры. Разумеется, можно говорить лишь о достаточно грубых моделях текста как сложного целостного объекта. Координацию составляющих можно осуществить с помощью методик позиционного анализа и синтеза наблюдений.

Свойственные тексту завершенность и отдельность лишь постулированы как его категории [7 и др.], их изучение представляется важной теоретической задачей. На этом пути представляют интерес процессы, приводящие сообщение (текст) к состояниям изоляции информации и структуры текста в его внешне обозначенных границах. Для этого необходимо видение текста как фазового пространства, в границах которого распределены и опредмечены в различные структуры все те процессы, которые привели завершенный и отдельный текст к такому состоянию. При этом мы пока что не можем даже приблизительно судить о полноте учета именно всех факторов сложной деятельности субъекта (автора данного текста). Поэтому любые наши суждения о тексте будут лишь частичными, заведомо неполными, а модели взаимодействия, факторы и параметры организации текста – огрубленными. И здесь возникает проблема выбора параметров описания сложного объекта, которая решается путем оценки существенности одних показателей и возможности пренебречь другими, второстепенными. Синергетика как раз и занимается анализом подобных ситуаций и построением моделей [8; 3 и др.].

Исследования структуры коммуникативно успешных и завершенных текстов позволили сформулировать некоторые симметрично-асимметричные закономерности их организации, позволяющие судить о наиболее предпочитаемых структурах устных и письменных текстов. Этот путь позволяет увидеть во всем многообразии конкретных текстов общее и всеобщее.

Поэтому изучение роли повторов в тексте, казалось бы, избыточного дублирования его компонентов, понятого как совокупность повторяющегося языкового материала в противовес всей совокупности различного, позволило установить устойчивые тенденции неслучайного взаимодействия двух конструктов относительно сильных и слабых позиций текста. А позиция в тексте суть репрезентант внутритекстового пространства-времени. Именно анализ естественной спонтанной речи диалектоносителей старшей возрастной группы показал важную роль симметричного в организации устного текста. Тексты, порожденные практически

неграмотными людьми, располагаются в диапазоне размеров от 4 до 150 словоформ. И лишь немногие из них совсем лишены разного рода повторов [9].

Дальнейшее изучение позиционной локализации в текстах устной и письменной речи всего повторяющегося в них языкового материала позволило вскрыть неслучайность повторений разного рода и сформулировать инвариант структурной организации для устной и письменной речи, организующий систему форм русского текста [9; 10; 11; 8 и др.]. Инвариантное в структуре текста принадлежит системе языка, это тот суперконструкт, который возник в языке эволюционно, он неосознаваем его носителями, но регулярно воспроизводится в речевой деятельности в виде вариантов с различной употребительностью [8, с. 281-289].

«Инвариант» структуры текста, по всей видимости, сформировался в процессе эволюции речевой деятельности как оптимальный вариант передачи информации между членами социума. Под оптимальной мы понимаем структуру текста, соответствующую надежному и адекватному восприятию информации реципиентом. Эта структура «подстроена» под емкость оперативной памяти человека и его речемыслительные способности / возможности.

Инвариант структуры текста является достаточно грубой моделью, лишенной излишних подробностей, что делает ее универсальной. Содержание инварианта характеризует наиболее вероятный сценарий возникновения и сменяемости симметричных и асимметричных структур в порядке их появления в тексте: что именно характерно и предсказуемо с различной вероятностью в тех или иных местах в русском тексте, то есть что характерно в началах, серединах, концах самых разнообразных текстов (устных и письменных, спонтанных и литературно обработанных).

Геометрические параметры инварианта, т. е. пропорции в пространстве-времени частей текста организованы на основе ряда Фибоначчи (по принципу золотого сечения) [8, с. 58-64]. Это ставит текст в ряд обычных природных объектов и подтверждает эволюционную природу его структурной организации.

В достаточно обширных выборках (изучено более 10 тысяч разнообразных текстов) наблюдаются повторяющиеся с различной вероятностью способы упаковки уникального содержания в формы, обладающие повторяющимися паттернами организации независимо от устной или письменной формы осуществле-

ния текста, от индивидуального авторства и т. п. Это позволяет рассматривать текст как информационный пакет (форма и скрытые ее правила организации и самоорганизации). Путем позиционного моделирования структуры текста выявлены 243 разнообразных способа спонтанной самоорганизации формы текста, которые позволяют зафиксировать сами формы в виде свернутых моделей и открывают возможность управлять параметрами формы текста. Это важно для разработки алгоритмов гармонической правки текста и компьютерных программ на их основе для редакционно-издательской, переводческой, рекламной и другой практики (см. подробнее: [8, с. 241-253]).

Стилистика решает вопрос о повторах в языке и тексте с установкой на богатство и разнообразие кода. Повторы при этом часто оцениваются как нежелательные, излишние, неоправданные (кем? чем? почему?). Подобная точка зрения на повторы и их оценка как погрешностей стиля или приемов организации речевого произведения, безусловно, важна. Но анализ бытия повторов в пространстве целого текста лежит в другой плоскости и относится не к онтологии самого субстрата, а либо к его осознанному в свете языковой нормы, либо неосознаваемому отбору и существованию, что может оцениваться лишь в свете бытия текста как такового. Здесь представляется перспективной аналогия с ближним и дальним порядком в структуре целого. То, что существенно в тексте на неболь-

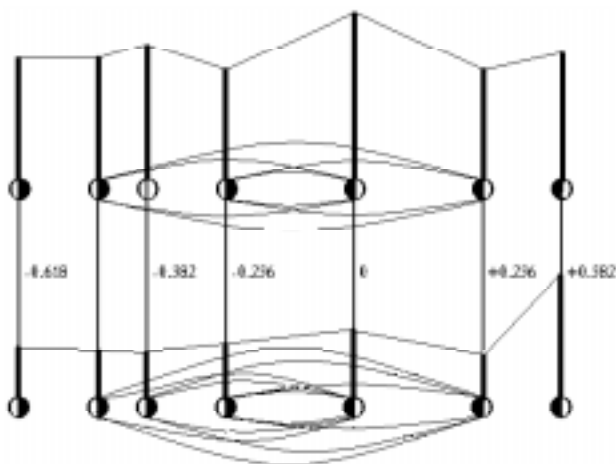


Рисунок 1. Соотношение инварианта структуры текста с процессами его циклической самоорганизации (вверху – инвариант устной речи, внизу – художественная проза) [8, с. 133].

Абсолютное начало текста – координата $-0,618$; абсолютный конец – координата $+0,382$; координата «0» – гармонический центр текста (пропорция золотого сечения). Жирные вертикальные линии указывают плотность размещения в указанных позициях элементов симметрии (повторов разного рода).

ших дистанциях, но «забывается» системой далее, то есть не повторяется больше, является фактом ближнего порядка, действует эпизодически. Повторяющееся же неоднократно образует в протяженности целого некоторую линейную конфигурацию, репрезентирует запомненный выбор данной системы и является носителем существенной для текста информации.

Текст является весьма сложной иерархической системой, создаваемой человеком для передачи и/или сохранения определенной информации. В качестве элементов этой системы можно рассматривать словоформы, которые могут служить единицами измерения протяженности (объема) текста в пространстве. Слова объединены в предложения, структурирующие весь текст на части / подсистемы. Так, еще Аристотель различал в тексте «начало», «середины» и «конец». У филологов (и обычных читателей) бытует представление о неповторимости, индивидуальности каждого текста. И это верно, если иметь в виду его содержание, смысл и их конкретную языковую репрезентацию в отдельном тексте. Именно такой подход принят в современном литературоведении, изучающем текст как уникальный и неповторимый. С такой точкой зрения трудно согласиться с позиций теоретической филологии, поскольку наука занимается поиском закономерного, стабильно воспроизводимого в природе и социуме. Уникальность объекта интересует искусствоведение, эстетику, но литературоведение пока слабо освоило подход к тексту с эстетических позиций, где также выявлены законы формообразования, композиции, восприятия и воздействия произведения на человека.

Поэтому текст как объект теории языка должен рассматриваться с позиций известной триады Гегеля: всеобщее – общее – отдельное. И здесь необходимым является выход за пределы каждого отдельного текста, обращение к тем его свойствам и качествам как завершеного и отдельного речевого произведения, присущим каждому и любому тексту вне зависимости от авторства и способа его осуществления (устный или письменный). Иными словами, текст в общей теории рассматривается в кругу других текстов.

Еще более высокий уровень абстракции необходим при вписывании каждого и любого текста как цельного и отдельного объекта в лингвоментальную и лингвопрагматическую деятельность человека и социума (текст в деятельности человека и социума, текст в кругу других объектов окружающей действительности).

Тем не менее, если большая часть порождаемых человеком текстов воспринимается как осмысленные сообщения, то данное обстоятельство можно не учитывать в анализе, а сосредоточиться на рассмотрении самих способов организации повторяющегося в каждом отдельном тексте и в тексте как инварианте. Вероятно, что на этом пути могут быть обнаружены тонкие подстройки особенностей структурирования дискретной информации в дискретную же структуру, могут выявиться различные функциональные связи, воплощающиеся в некоторых группах текстов или свойственные любым текстам. То есть изучение моделей организации симметричного в тексте позволяет отвлечься от конкретного и, на первый взгляд, пугающего многообразия отдельных текстов и перейти к анализу собственно структур и процессов, приводящих к их появлению. Таким образом, мы полагаем, что все наиболее важное для данной системы в ней уже находится, проявилось и органично встроилось в данную форму. В тексте такими компонентами, репрезентантами его «памяти» (макроинформации) являются элементы симметрии, то есть широко понятое повторяющееся, сложно переплетающееся с неповторяющимся (различным).

Словоцентризм имеющихся теорий текста заслоняет подлинный масштаб его макроструктурирования. Слово в данной системе является лишь структурной ячейкой синтаксической фактуры целого. Предложение также предстает лишь как дискретность, но в отличие от слова в предложениях явственно проступает их размер, а он изменчив в достаточно широком диапазоне. Но, тем не менее, естественно определились наиболее употребительные размеры и предложений, и текстов. Как реальность можно воспринимать и характер внутреннего соотношения размера предложения и текста, их инвариантную взаимоподстройку в рамках некоторой меры.

Мера и является интегративной характеристикой целостного объекта, поскольку даже самый длинный текст конечен в физическом времени и в лингвоментальном пространстве субъекта (наблюдателя). Различная протяженность внутритекстового пространства-времени служит его объективной характеристикой, позволяющей проводить измерения и фиксацию на оси протяженности разнообразных наблюдений. Текст является направленно развертывающейся в пространстве-времени структурой. Процесс направленного роста осуществляется от абсолютного начала текста к его абсолютному концу.

Размер конечного продукта речемыслительной деятельности (текста) является по сути единственно объективным и самым фундаментальным его параметром, характеризующим целостные качества информационного пакета. Поэтому текстосимметрия базируется на размерах целых текстов и их разнообразных составляющих. Это размеры композиционных зон текста, пропорциональные соотношения повторяющегося и различного в нем, отражающие характер подстройки целого и его частей для выполнения общих функций. Разумеется, исследования форматов текста должны быть продолжены, и на этом пути возможны весьма многообещающие результаты.

В свете изложенного продуктивной представляется идея для осуществления синтеза – внутритекстовое физическое пространство-время системы и вписанные в целостный континуум структуры и процессы. Форма текста рассматривается при этом как эволюционно-природный объект, вмещающий в своем пространстве-времени возможные в данных граничных условиях инвариантно и вариативно реализованные разнообразные и взаимодополнительные собственно структурные и информационные характеристики.

Список использованной литературы:

1. Капра Ф. Паутина жизни. Новое научное понимание живых систем. М.: ИД «София», 2003.
2. Карпов В.А. Язык как система. М.: Едиториал УРСС, 2003.
3. Чернавский Д.С. Синергетика и информация. Динамическая теория информации. М.: Едиториал УРСС, 2004.
4. Хазен А.М. Разум природы и разум человека. М., 2000.
5. Шеннон К., Бандвагон Е. Работы по теории информации и кибернетике. М.: ИЛ, 1963.
6. Корогодина В.И. Информация и феномен информации. Пушкино: АН СССР, 1991.
7. Гальперин И.Р. Текст как объект лингвистического исследования. М.: Наука, 1979.
8. Москальчук Г.Г. Структура текста как синергетический процесс. М.: Едиториал УРСС, 2003.
9. Москальчук Г.Г. Фразовый повтор в диалектной речи: Автореф. дис. ... канд. филол. наук. М., 1990.
10. Корбут А.Ю., Москальчук Г.Г. Инвариант структуры диалектного и художественного текста // Явление вариативности в языке. Кемерово, 1997. С. 182-188.
11. Корбут А.Ю. Повтор как средство структурной организации художественного прозаического текста (элементы симметрии): Автореф. дис. ... канд. филол. наук. М., 1995.