

Тема 3: СВЕРТЫВАНИЕ ИНФОРМАЦИИ

Свертывание, информационное свертывание, развертывание, семантическое свертывание в языке, опущение, совмещение, замещение, текст, знак, значение, информация, информация социальная, информативность, содержательность синтаксическая структура текста, сложное синтаксическое целое, автосемантическое предложение, синсемантические предложение, микротекст, цельный микротекст, коннекторы, коммуникативная структура текста, синтаксическая структура предложения, текстовые субъекты, текстовые предикаты, содержательный аспект.

3.1. Понятие “свертывание информации”.

3.2. Виды свертывания.

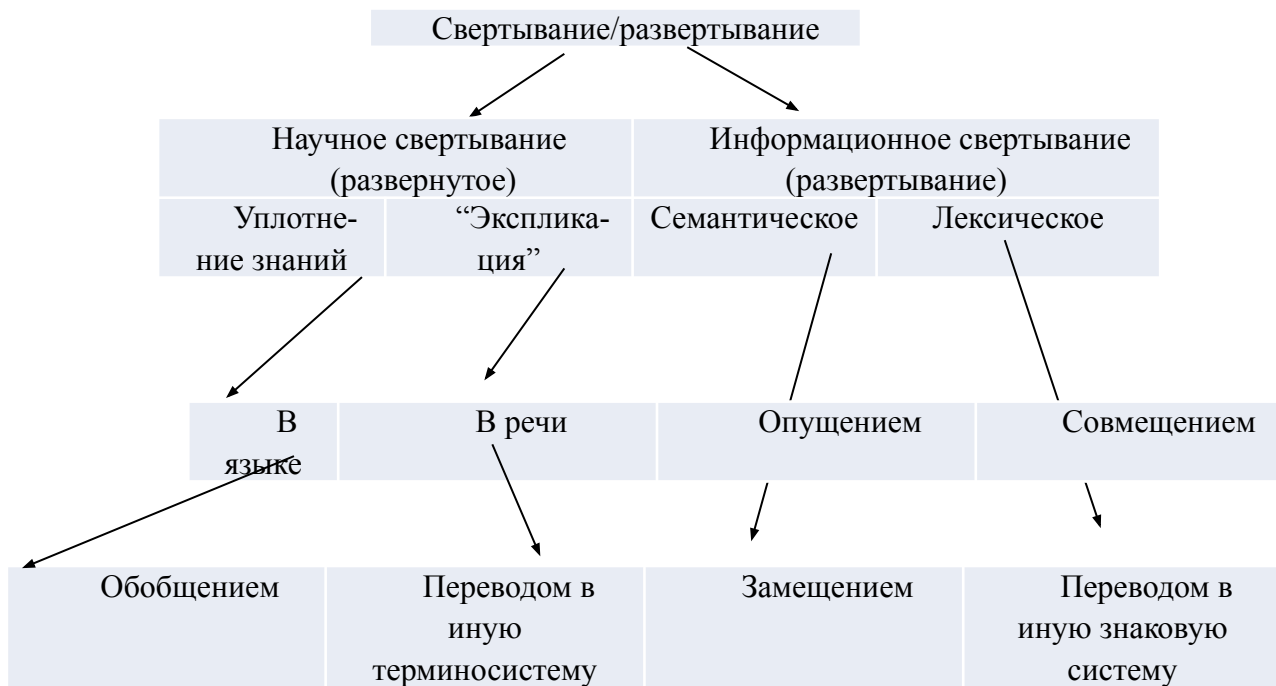
3.3. Понятие и структура текста.

3.1. Понятие “свертывание информации”

Свертывание – это совокупность операций аналитико-синтетической переработки информации, преследующих цель выразить содержание исходного текста в более краткой форме при сохранении или некотором уменьшении его информативности в производном тексте.

Развертывание – это увеличение физического объема текста за счет внесения в него необходимых дополнений, уточнений, пояснений, комментариев, обеспечивающих наилучшее его восприятие, понимание. Развертывание повышает информативность текста.

Процесс свертывания можно представить в следующем виде:



Уровни свертывания текста

Текст

Информация

Смысл

Значение

В основе процедуры информационного свертывания лежит психолингвистический механизм обратной связи, который в многоуровневой цепочке текстопорождения (“смысл – текст”) обеспечивает постоянную корреляцию между реальным результатом речевой деятельности и смысловым образом, порожденным на этапе мотивации этой деятельности. Этот механизм определяет необходимый и достаточный набор лексических средств и синтаксических конструкций, который использует отправитель информации, чтобы обеспечить оптимальные условия коммуникации.

3.2. Виды свертывания

Семантическое свертывание в языке – это стихийный процесс терминообразования, в ходе которого возникают понятия более краткие, чем их синонимичные предшественники, например, *компьютер* и ЭВМ. Семантическое свертывание терминов в языке предстает в виде логического произведения некоторых атомарных значений – сем, семантических множителей. Так, к примеру, в качестве сем лексемы *трансформатор* могут выступать слова *устройство, изменение, напряжение*.

Семантическое свертывание в речи:

а) с изменением информативности путем обобщения (или, наоборот, деления) понятий и целых смысловых отрезков или опущения (или, наоборот, внесения) некоторых деталей, подробностей, уточнений;

б) с изменением информативности путем перевода в иную систему понятий.

Пример изменения информативности путем обобщения.

1. Свертывание.

Давняя проблема соотношения биологического и социального в человеке, которую намеревались окончательно решить в XIX веке при помощи методов точных наук, так и не была решена.

Проблема соотношения биологического и социального в человеке в XIX веке не была решена.

Другими примерами высокого уровня обобщения больших фрагментов текстов могут служить заголовки и подзаголовки в документах, фразы в аннотациях, тезисы к докладам и пр.

Развертывание

Социология – это наука об обществе; социология – это наука, изучающая структуры общества, их элементы и условия существования, а также социальные процессы, протекающие в этих структурах; социология – это наука об обществе как социальной системе в целом, функционировании и развитии этой системы через ее составные элементы: личности, социальные общности, институты.

Пример изменения информативности путем перевода в иную систему понятий:

- Бедность – это не только минимальный доход, но особый образ и стиль жизни.*
- Бедность – это экономическое и социокультурное состояние людей, имеющих минимальное количество ликвидных ценностей и ограниченный доступ к социальным благам.*

При лексическом свертывании производится преобразование знаковой формы сообщения при сохранении его информативности. В отдельных случаях в результате лексического свертывания информативность может даже возрасти за счет повышения читабельности текста (путем опущения неинформативных, паразитных слов, упрощения синтаксической конструкции и пр.).

Различают три способа лексического свертывания:

- *опущение*, при котором то или иное место в предложении, предназначенное для определенных повторяющихся или не повторяющихся в тексте языковых единиц, остается пустым, но может быть заполнено этими единицами благодаря опоре на сохраняющиеся в предложении элементы (например: *В котором часу ты придешь? - В седьмом*);
- *совмещение*, при котором два или несколько предложений, некоторые элементы которых тождественны, накладываются друг на друга, образуя сокращенную конструкцию, где тождественный компонент употреблен только однажды, но сохраняет самостоятельные связи с нетождественными частями совмещенных предложений;
- *замещение*, при котором повторяющийся или неповторяющийся отрезок текста замещается другим, более кратким, с сохранением в последнем необходимого уровня смысла первого (например, замена местоимением слова или выражения, применение более кратких синонимичных вариантов).

К перечисленным видам лексического свертывания можно добавить также применение аббревиатур, условных сокращений, перевод в иную знаковую систему (употребление формул, графиков, кодов и т.п.). Возможно, в качестве одной из форм лексического свертывания можно рассматривать употребление метафор.

Например: Трудно даже перечислить все затронутые в книге вопросы: читатель как бы попадает на интеллектуальный пир, откуда он не уйдет голодным, хотя некоторые блюда ему, может быть, не придутся по вкусу (М. В. Арапов).

Таким образом, в процессе информационного свертывания не просто сокращается текст, а именно “сворачивается”, причем так, чтобы иметь возможность затем вновь его развернуть на основе сохраненных “смысловых вех”, “смысловых опорных пунктов”. Это особенно наглядно проявляется при составлении индивидуального конспекта (мы записываем обычно то, что позволяет нам впоследствии мысленно восстановить конспектируемый текст), но это также характерно и для других видов свертывания.

Свертывание можно классифицировать и по ряду других критериев, например, по степени формализации процесса – интуитивное, алгоритмизированное (анкетное в том числе), автоматизированное; по количеству свертываемых документов – отдельных или некоторого множества (к примеру, реферат и реферативный обзор); по глубине (степени) свертывания – библиографическое описание, реферат) и т.д.

Наиболее важным *основанием классификации* видов свертывания является *вид анализа* — *документографический, фактографический и концептографический.*

Виды свертывания и порождаемые на его основе документы приведены в таблице.

Классификация видов свертывания информации

Признак	Вид анализа документов		
	документографический	фактографический	концептографический
Вид документов в результате свертывания	Библиографическое описание	Фактографические справки	Рекомендательные аннотации
	Аннотации (справочные и реферативные)	Рефераты	“Развернутые” рефераты
		Экспресс-информация	Рецензия
	Библиографические обзоры	Реферативные обзоры	Аналитические обзоры
	Классификационные индексы	Сопоставительные таблицы	Критические библиографические обзоры и др.
	Предметные рубрики	Фрагменты	
Функциональное назначение документов	Идентификация документов и ориентация в документальном потоке	Ориентация в информационном потоке	Интерпретированная ориентация в документальном в информационном потоке

Понятие текста

Текст – это многоуровневая система, причем не только в семиотическом аспекте (знак-слово, знак-предложение, знак-сложное синтаксическое целое), но и с точки зрения других измерений ею синтаксической, коммуникативной, семантической, информативной структуры и др.

Синтаксическая структура текста

Текст с точки зрения его синтаксической структуры представляет собой сложную многоуровневую систему суперсинтаксических единиц, очищающихся друг от друга «масштабностью» своих микротем.

[В литературе эти суперсинтаксические единицы называются по-разному — “сложное синтаксическое целое”, “суперсинтаксическое единство”, “прозаическая строфа”, “сегмент”, “суперсегмент”, “субтекст” и др.]

Автосемантическое предложение

Автосемантическим является то предложение, которое, будучи изъятым, из контекста, не содержит в себе никаких эксплицитно, т. е. явно, выраженных связей с другими предложениями, оно всегда самостоятельно и имеет твердое начало (начинается с существительного-подлежащего или группы подлежащего).

Синсемантическое предложение

Синсемантические предложения подчинены контексту и вне его лишаются ситуативной определенности, т. е. несамостоятельны и имеют мягкие начала (соединительный союз или глагол-сказуемое, указательные, притяжательные или личные местоимения 3-го лица, неполное предложение, т. е. слова-заместители, слова-связки).

Особенность рассмотрения сложного синтаксического целого как объекта свертывания

Автосемантические предложения обычно выступают в качестве метаинформативных элементов сложного синтаксического целого, синсемантические – в качестве его информативных элементов. Поэтому, если стоит задача в процессе свертывания повысить уровень метаинформативности сложного синтаксического целого, из него последовательно удаляется определенное количество синсемантических предложений. При этом важно, чтобы в оставшихся синсемантических предложениях слова – заместители (типа *эти, такой, по добный, они* и т.д.) были заменены на соответствующие им смысловые элементы из удаляемых вышестоящих предложений. В противном случае свернутое сложное синтаксическое целое потеряет свое смысловое единство.

Основные текстообразующие признаки сложного синтаксического целого

- *Цельность*

Цельным называется такой текст, который воспринимается как осмысленное целесообразное единство. Цельный текст характеризуется иерархией смысловых предикатов. В этом смысле можно определить цельный текст как такой текст, который при переходе от одной последовательной ступени смысловой компрессии к другой, более “глубокой”, каждый раз сохраняет свое смысловое тождество, лишаясь маргинальных (т.е. пограничных) элементов.

- *Тематичность*

Тематичность текста характеризуется единством предмета высказывания. Во всяком тексте, если он относительно закончен и последователен, высказана одна основная мысль, один тезис, одно положение. Все остальное подводит к этой мысли, развивает ее, аргументирует, разрабатывает.

- *Связность*

Одним из важнейших текстообразующих факторов общепризнанно является связность, без которой текст как функционально завершенное целое вообще немыслим. Различают *содержательный, логический* и *композиционный* аспекты связности

Содержательный аспект связности

Содержательный аспект связности проявляется через употребление лексико-грамматических средств связи самостоятельных предложений и абзацев: лексические повторы, указательные и личные (3-го лица) местоимения, наречия с причинно-следственным, пространственным и временным значениями.

Логический аспект связности

Логический аспект связности проявляется через выражение логических отношений между предложениями, абзацами и т.д. посредством функционально-синтаксических средств — вводные слова типа *так, таким образом, следовательно, итак, кроме того, наконец* и др.; наречия и наречные слова *сюда, отсюда, тогда, здесь, оттого, потому, затем, теперь, сейчас, сначала, сперва, после, выше, ниже* и др.; союзы *и, а, но, да, а и, однако, однако же* и др.

Композиционный аспект связности

Композиционный аспект связности выражается с помощью так называемых конструкций связи: *рассмотрим, заметим, что; важно подчеркнуть что; можно предположить, что; следует иметь в виду, что; понятно, что; как будет показано ниже; согласно сказанному заключается в следующем* и др., относящихся также к функционально-синтаксическим средствам.

Пример:

“Я пошел в кино. Кино на Остоженке. Остоженка – одна из самых старых улиц Москвы. Москва – центр всех железнодорожных путей страны”.

С другой стороны можно создавать вполне связные тексты без специальных сигналов связи.

Структуры связного текста

- В случае *цепочной структуры текста* все предложения оказываются связанными между собой так, что каждое данное предложение связано с предыдущим предложением, либо с одним предыдущим и одним последующим. При этом наблюдаются простые цепочки, когда каждое предложение связано с предыдущим с помощью лишь одного средства связи (либо повтора, либо коннектора и т. п.), и сложные, что встречаются чаще, когда предложения присоединяются друг к другу различными “скрепами”.

- В случае *рамочной структуры текста* сколь угодно много предложений текста образуют относительно законченное целое, состоящее из трех частей, которая по тем или иным причинам требует раскрытия в последующем контексте, средней (конкретизирующей) части, раскрывающей начальную (вводную) часть, и заключительной (обобщающей) части, – она подводит, итог образует своеобразную рамку.

Средства внутритекстовой связности

- 1) лексические и семантические повторы,
- 2) соотношенность временных форм глаголов,
- 3) анафорические и логико-смысловые коннекторы.

Виды повторов

– *основные:*

- а) когда каждое предложение включает некоторую словоформу одного и того же имени существительного,
- б) когда предложения включают синонимичные имена существительные (строительство – постройка),
- в) когда предложения включают имена существительные, находящиеся в парадигматических отношениях типа подчинения (врач – кардиолог),
- г) когда последующие предложения включают прилагательные или глаголы с той же основой, что и существительные первого предложения (опора – опорный);

– *расширенные*:

а) когда предложения включают имена существительные, являющиеся синонимами в рамках данного контекста (контекстуальными синонимами),

б) когда в последующих предложениях имя существительное заменено на личное, притяжательное или указательное местоимение.

Разновидности коннекторов

Класс коннекторов	Примеры коннекторов	Примеры употребления коннекторов
<i>Анафорические коннекторы:</i>		
Анафорическая связь с указанием на предмет схожий с данным	Аналогичный, данный, указанный, упомянутый, этот, такой,	Совмещение элементов разнородных знаковых систем создает специфические проблемы. Аналогичные проблемы возникают и в связи с применением товарных знаков.
Анафорическая связь с указанием на предмет, отличный от данного	Другой, иной, прочий	Одним из средств стала система препринтов. Другой путь заключается в сокращении объема первичных публикаций.

Разновидности коннекторов

<i>Логико-смысловые:</i>		
Результат, итог	В итоге, в результате, итак, наконец, таким образом	Обычно центры анализа выписывают и просматривают весьма ограниченное число профильных изданий. Таким образом, центры анализа зависят от референтных служб.
Дополнения, сопутствующие обстоятельства	Более того, вдобавок, вместе с тем, кроме того, при этом, также	Адресность набора предполагает, что выбирается профильный журнал. Вдобавок следует отметить, что не нужно различать теоретические, научные и узкоотраслевые журналы.
Перифразы	Другими словами, иначе говоря, иными словами	Наука и техника существовала и развивалась и без таких информационных центров. Иначе говоря, следует уточнить, оправданы ли расходы на содержание центров.
Пример, общее – конкретное	В частности, например, примером, так	Ранее публикации издавались, в основном, на нескольких иностранных языках. Например, в 1909 году достаточно было знать три языка, чтобы читать 92% всей литературы по химии.
Импликация, причинно-следственные связи	Отсюда, следовательно, поэтому, стало быть	Связующим звеном между потребителем и системой служит узкий специалист. Поэтому общение между специалистом и потребителем информации осуществляется на профессиональном уровне.
Возражения, оговорки	Впрочем,	Научная информация обладает свой

Коммуникативная структура текста

- В отличие от синтаксической (формально-грамматической) структуры предложения, включающей несколько членов предложения, коммуникативная структура опирается на смысловое членение высказывания. Коммуникативная структура всегда двучленна. Она состоит из субъекта – **темы (Т)**, т.е. о чем сообщается в данном предложении (известное, данное), и предиката – **ремы (Р)**, т.е. что сообщается в данной теме (новое). Однако принцип смыслового членения характерен не только для предложения (высказывания), но и для текста в целом и различных его частей.

Текст в целом можно рассматривать как выражение суждения, субъект которого сформулирован в заглавии, а предикат – весь остальной текст. Этот принцип позволяет делить целый текст на смысловые фрагменты (единицы), которые в свою очередь представляют собой текстовое суждение более низкого уровня иерархии, включающие свои текстовые субъекты (темы) и текстовые предикаты (ремы). Следует, однако, иметь в виду, что предмет текстового субъекта не совпадает с предметом логического субъекта; не совпадают, соответственно, по содержанию текстовой и логической предикаты, поскольку логические субъекты и предикаты представляют собой понятия, а текстовые субъект и предикат – логические суждения (предложения, сегменты и т.д.). Для коммуникативной структуры научного текста характерны в основном четыре вида отношений типа $T - P$ (или четыре типа коммуникативных цепочек) в рамках того или иного сложного синтаксического целого.

Первый вид – *простая линейная корреляция*

T и P

В этом случае каждая рема предыдущего предложения является темой для ремы следующего предложения. Такой вид контекста иногда называют последовательным строем:

T1 – P1

T2 (=P1) – P2

T3 (=P2) – 3 и т.д.

Пример: Человеческая культура неотделима от традиций. Традиции, существующие десятки и сотни тысяч лет, могут превратиться в социальный институт. Институт – не помещающаяся в рамки традиции массовая привычка что-либо делать определенным образом.

Второй вид – последовательная корреляция T и P с постоянной темой (связь типа “нанизывание”)

В этом случае при одной повторяющейся теме имеется некоторое количество рем:

T1 – P1

T1 – P2

T1 – P3

Пример: Ценности – социально одобряемые и разделяемые большинством людей представления о том, что такое добро, справедливость, патриотизм, дружба и т.п. Ценности не подвергаются сомнению. Ценности служат эталонам, идеалам для всех людей.

Третий вид – предложения объединены параллельной связью

- Здесь разрабатывается (развертывается) не рема, а тема. последующие темы являются производными от тем других тем, а не рем:
- Т
- 1————— Т —————1 ...
- T1–P1 T2–P2 T3–P3
- Пример: Т. Парсонс (США) рассматривает три системы. Социальная система включает совокупность институтов и структур, обеспечивающих социальные требования и организацию совместной жизни людей. Система личности состоит из способов организации мотивированного действия. Система культуры направляет взаимодействие людей, чтобы примирить личность и общество, поддержать порядок.

Четвертый вид – смесь 1-го и 2-го видов (смешанный тип).

- T1 – P1 или T1 – P1 или T1
 -
 - T2 – P2 T2 – P2 P1 P2 P3 P4
 -
 -
 - T2 – P3 T3 – P3 и т.д. T2 – P5
 -
- P6 и т.д.

Пример текста, имеющего структуру, изображенную последним графом:

К социально-демографической подструктуре (T1) относятся различия по полу (P1), возрасту (P2), образованию (P3), семейному положению (P4). Семейное положение (T2) зависит от ряда признаков (P5), характерных для данной социальной структуры (P6).

Аспектная структура текста

Семантический анализ текста (документа) предполагает разбиение его на отдельные содержательные аспекты, набор которых в достаточной степени универсален для однотипных документов. Под содержательным аспектом понимают контекст, характеризующий тот или иной предмет рассмотрения в документе с определенной точки зрения: его развития, назначения, принципа действия, условий применения и т.д., либо логику изложения содержания документа (постановка проблемы, целевая установка, описание предлагаемого варианта решения и т. д.). В практике реферирования такие перечни аспектов часто называют “сетками”. Можно выделить два типа содержательных аспектов: первый отражает структуру (логику) вида документа, например статьи, другой – структуру (логику) той или иной области знания или группы объектов.

В качестве примера характерной сетки первого типа можно привести следующий перечень аспектов:

1. Известный вариант решения (ИВР).
2. Недостатки и оценки известного варианта решения (ОцИВР).
3. Постановка проблемы (целевая установка) (ПП, ЦУ).
4. Описание предлагаемого варианта решения (ОПВР).
5. Оценка предлагаемого варианта решения (ОцПВР).
6. Особенности предлагаемого варианта решения (ОсПВР).
7. Экспериментальная проверка (ЭП).
8. Эффективность (экономическая, социальная) (Э).
9. Область применения (ОП).
10. Выводы (В).
11. Результаты (Р).
12. Рекомендации (Рек.).

Подобная структура документа отражает, как правило, саму логику исследования: вначале изучается состояние вопроса – что уже сделано, достигнуто в данной области, чем достигнутое не удовлетворяет требованиям нового этапа практики, т.е. в чем суть рассогласования между существующим и требуемым. Этот этап разработки находит отражение в аспектах ИВР и ОцИВР. Затем на основе анализа существующей ситуации формулируется ведущая гипотеза о способах и средствах устранения имеющегося рассогласования. Этот этап отражается в аспектах ПП, ЦУ. Далее следует сама разработка (эксперименты, расчеты и пр.) и осмысление полученных результатов с анализом их достоинств, ограничений, особенностей и возможностей дальнейшего использования. Эти этапы находят отражение, соответственно, в аспектах ОцПВР, ОсПВР, В, Рек и др.

Аспектная структура сеток второго типа

Аспектная структура сеток второго типа определяется характером описываемого в документе объекта – описывается ли методика в целом, отдельный метод, материал, конструкция и т. д.

Для документов – статей, посвященных, например, изложению технологического процесса, аспектная сетка будет иметь вид:

1. Сущность технологического процесса.
2. Последовательность отдельных технологических операций.
3. Объект, подвергавшийся обработке.
4. Применяемое оборудование и инструмент.
5. Результаты процесса, качество и свойства полученных изделий.
6. Область применения процесса.

Сетки, отражающие логику того или иного вида документа (первого типа), имеют значение прежде всего в практике реферирования (аннотирования) и, как правило, носят универсальный характер для многих отраслей знания. Сетки второго типа также используются для описания структуры вторичного документа, однако имеют большее значение для фрагментирования документов, выделения из них содержательно самостоятельных типовых фрагментов текста.

Несмотря на многообразие научных документов (множество типов и видов научной литературы, отраслей знания), различие в языке и стиле их авторов, в структуре научных документов наиболее часто встречаются следующие аспекты содержания:

Семантический блок *Аспекты содержания*

1. Вводная часть Актуальность проблемы

Известный вариант решения Недостатки известного варианта решения
Целевая установка

2. Основная часть Предлагаемый вариант решения или предмет рассмотрения
Особенность предлагаемого варианта решения
Место исследования Область применения (назначения) Методы исследования
Оборудование или технические средства реализации

Примеры: Наглядное представление информации

Математический аппарат Экспериментальная проверка

Преимущества предлагаемого варианта решения

3. Заключение Результаты

• Выводы

• Рекомендации

Информативная структура текста

- С точки зрения информативной структуры, в терминах теории информации текст представляет собой последовательность сигналов (знаков), отражающих содержание какого-либо объекта и несущих не только новую информацию, но и некоторое количество информации известной (в сумме – полезную информацию), а также и неизбежные помехи.
- Новая информация – это целенаправленно создаваемая часть текста, которая, по мнению автора (*коммуниканта*), способна внести изменения в структуру как общественного, так и индивидуального тезаурусов. Это содержательная часть текста.

ШУМ

- Составной частью новой информации является шум, который в то же, время относится и к сфере помех (но не входит в сферу полезной информации).
- Шум – это дезинформация, вольно или невольно закравшаяся в текст: ошибки в расчетах и обобщениях, предвзятые, необоснованные выводы и рекомендации, недостаточно корректно трактуемые результаты экспериментов и др. Шум – самая опасная категория помех, поскольку снижает уровень содержательности информации.
- Искажения выступают в виде различных стилистических ошибок, нечетких формулировок и пр.

Избыточная информация

- К избыточной информации относятся уже известные сведения, повторения, излишние подробности. Избыточная информация разделяется на *релевантную* и *нерелевантную*, причем граница между ними весьма подвижна и зависит от степени развитости тезауруса человека и уровня свертывания ситуации в сообщении.
- Наряду с избыточной нерелевантной информацией в документе не только допустима, но и обязательна избыточная релевантная информация, обеспечивающая понимание и восприятие новой информации. Известно, что ничто новое не может быть воспринято без введения его в контекст известного. Задача заключается в том, чтобы известное было изложено на уровне свертывания, соответствующем назначению документа. Таким образом, избыточная информация может относиться не к сфере помех, а к полезной информации. Полезностью обладают также и метаинформативные элементы текста, ориентирующие в структуре документа.

Функционально-смысловая структура

- Традиционно различают три типа речи (текста): описание, повествование, рассуждение (иногда добавляют определение).
- *Текст-описание* характеризуется перечислением признаков описываемого объекта или его частей. Здесь построение текста определяется структурой и признаками данного объекта.
- Для *текста-повествования* характерно сообщение о развивающихся действиях, событиях, состояниях. Здесь материал излагается часто в хронологической последовательности, отражая процессуальность жизни.

В *текстах-рассуждениях* содержатся причинно-следственные (выводные) конструкции, авторские оценки, модальные слова. Здесь тексты отражают, как правило, ход процесса мышления. Подобная функционально-смысловая типология текстов применима лишь к сравнительно небольшим фрагментам текста, так как более крупные текстовые образования содержат в себе нередко все три типа речи.

В рамках различных функциональных стилей (*научного, официально-делового, публицистического*) типы изложения (речи) в общем универсальны и, попадая в любой из стилей, сохраняют присущие им как типам изложения лингвистические черты. Однако в пределах конкретного функционального стиля (в нашем случае – научного) доля разных способов изложения меняется в зависимости от вида науки. Так, в экспериментальных науках и в работах, описывающих эксперимент, чаще используется повествовательный тип речи. В логико-теоретических науках (философии, например), а также в работах теоретического характера (независимо от принадлежности к той или иной науке) чаще используется такой тип речи, как рассуждение. Для наук описательных характерен, естественно, описательный тип речи.