

УДК 81'42(045)

М. В. Белла

ГИПЕРТЕКСТ И ЕГО СОСТАВЛЯЮЩИЕ

Компьютерная лингвистика, как направление прикладной лингвистики, наиболее активно начала разрабатываться в 60 – 70 гг. XX века и в первую очередь подразумевала использование статистических методов в языкознании. Сегодня компьютерная лингвистика – это одно из важных направлений в конструктивном описании языка. Этим и определяется интерес к проблемам компьютерного моделирования функционирования языка и сферам его применения, возникший в лингвистике в последние десятилетия. Одно из прикладных направлений компьютерной лингвистики – гипертекстовые технологии представления текста. Под гипертекстовой технологией понимается формирование, поддержка, наращивание и просмотр организованного в виде сети текста на компьютерной основе.

Актуальность данного исследования обусловлена многогранностью такого явления как гипертекст, которое характеризуется нелинейностью, децентрализованностью и максимальной приближенностью к процессу порождения высказывания. Идея гипертекста легла в основу многих современных научных исследований (А. Н. Баранов, М. М. Субботин, Е. В. Зыкова, О. В. Дедова, Р. К. Потапова и др.), что, прежде всего, связано с успешным применением гипертекстовых систем во многих исследовательских системах, в том числе и лингвистических. Таким образом, возникает один из важных вопросов компьютерных исследований – вопрос разработки теории гипертекста, ее реализации в гипертекстовых технологиях, а также возможности их использования.

Цель нашей статьи – выявление специфики электронного гипертекста как объекта лингвистического исследования. В круг задач входит рассмотрение гиперссылки как средства создания единого гипертекстового пространства, изучение возможности использования наработок теории гипертекста в лингвистических исследовательских программах.

Термин гипертекст ввел Тед Нельсон, определив гипертекст как соединение текста на естественном языке с создаваемой компьютером возможностью интерактивного создания внутри него новых ветвей или динамичной организации нелинейного текста, который уже не может быть напечатан обычным образом на обычной странице [1]. Гипертексту в его нынешнем «компьютерном» понимании предшествует нелинейный текст «ручного» составления. Одним из прообразов гипертекста были карточки со взаимными ссылками. Удобство таких карточек состояло в разбиении больших записей на несколько более мелких, которые в

дальнейшем можно было легко реорганизовать с учетом новой информации [2]. Иной формой представления «ручного» гипертекста являются словари и энциклопедии, в которых статьи или определения содержат ссылки друг на друга. Следуя по этим ссылкам, можно получить более полную информацию. В современных научных дискуссиях все чаще звучат вопросы, связанные со способами организации электронного гипертекстового пространства и возможностями навигации внутри него, обсуждаются технологии построения гипертекстов, описываются области применения гипертекстовых систем [2; 3].

Основными структурными элементами гипертекста являются узел и ссылка. Узел гипертекста – это относительно самостоятельный фрагмент текста, который обладает связью либо связями с другими фрагментами текста. Узлы в гипертексте не регламентируются в отношении видовой принадлежности (в гипертекст могут включаться разнородные элементы: аудио, видео, анимация и др.), масштаба, количества ссылок и степени «открытости», включённости в общий гипертекст (узел может содержать ссылки как на один, так и на неограниченное количество фрагментов текста). Ссылка в гипертексте – это элемент текста, который устанавливает связь между различными узлами. Характер ссылок может быть различен в отношении видовой принадлежности, масштаба, связности с основным текстом узла. Функция ссылки является наиболее важной функцией гипертекста. Она дает возможность использовать связи информационной сети, в виде которой представлен гипертекст как объект.

Гипертекст можно рассматривать в качестве своеобразной открытой базы данных, которая организуется в виде свободно наращиваемой и изменяемой сети, узлы которой (линейные тексты) соединяются самим пользователем. От обычной базы данных гипертекст отличается прежде всего тем, что в нем отсутствуют заданные ограничения на характер связей (как, например, в иерархических структурах). Последовательно соединенные связями узлы образует цепь. Расстояние между узлами, которое соответствует близости или неблизости их содержания, равно минимальному количеству промежуточных узлов. В качестве узла могут выступать слово, словосочетание, предложение, абзац, параграф, документ, собрание документов, относящихся к одной теме, отдельные сообщения, мультимедиа и т. п. Характер связей между узлами может быть различным. Переход может осуществляться между текстами, между текстом и комментарием к нему, между разными редакциями текста, между текстом и его возможными продолжениями, между текстами, пересекающимися по содержанию, отвечающими или возражающими друг другу, и т. д.

Можно выделить следующие типы связей между объектами узлов гипертекста: отношение «*вид – род*», отношение «*род – вид*»,

отношение «целое – часть», отношение «часть – целое», отношение «предмет – процесс», отношение «процесс – предмет». Отношения задают потенциальные возможности навигации по гипертексту. Отношения могут быть однонаправленными, которые позволяют двигаться только в одну сторону, или двунаправленными, которые, соответственно, позволяют двигаться пользователю в обе стороны.

Цепочка узлов, по которым проходит пользователь (читатель) при просмотре элементов гипертекста, образует путь, или маршрут. В зависимости от объекта, на который ссылаются гиперссылки, можно выделить гиперссылки на web-страницу; гиперссылки на часть документа внутри текущей web-страницы; гиперссылки на мультимедийный ресурс Интернета (музыка, фото, видео); гиперссылки на документ FTP-сервера (протокол передачи файлов); гиперссылки на электронную почту; гиперссылки на сценарий на языке Java Script.

Создание гипертекста состоит, прежде всего, в формировании системы переходов от узла к узлу (системы ссылок). В зависимости от типа гипертекстовой системы такая система может задаваться как разработчиками, так и пользователем в процессе работы с гипертекстом. Основными параметрами создания гипертекстового пространства являются его нелинейность, гибкая система переходов между элементами гипертекста, способность всех элементов быть доступными в равной степени и координация нескольких дополняющих друг друга структурных элементов, связанных по содержанию [4].

Гипертекстовая технология реализуется в конкретной гипертекстовой системе, состоящей из двух частей: гипертекста и гипертекстовой оболочки. Основными функциями гипертекстовой оболочки являются поддержка ссылочных связей; создание, редактирование и наращивание гипертекста; просмотр; прямой доступ; выделение виртуальных структур.

Функция создания, редактирования и наращивания гипертекста принципиально отличает технологию гипертекста от технологии баз данных, в которых концептуальная схема данных заранее задана. Она позволяет вводить новые узлы, редактировать содержание узлов, устанавливать связи между узлами. В зависимости от специализации гипертекстовые системы могут обладать различным набором вышеперечисленных функций. Однако, если гиперсеть имеет сложную, разветвленную структуру, возникает проблема ориентации пользователя, т. е. определения, в каком месте сети в данный момент он находится.

Проблема ориентации присутствует и при работе с традиционным линейным текстом большого объема, но в этом случае пользователь имеет только два направления поиска – выше или ниже. Гипертекст предлагает гораздо больше возможностей в выборе направлений движения, поэтому работать с ним сложнее. Кроме того, важную роль здесь играют виды гиперссылок.

Анкорные ссылки являются более привлекательными (анкор –

«видимый» пользователем текст ссылки, при нажатии на который осуществляется переход), т.к. в самом анкоре уже заложены ключевые слова, которые заинтересовывают и заставляют пользователя нажать на ссылку. Безанкорные ссылки («тут», «здесь», «скачать», «смотреть», «по этой ссылке», «читайте здесь», «на этом сайте», «читать далее», <http://site.ua/> и др.) пользуются меньшей популярностью по кликам у пользователей. Переходы по ним осуществляются за счет околоссылочного текста, выполняющего роль анкера, который может стоять как перед, так и после ссылки.

Гипертекстовые технологии широко используются в различных прикладных системах: в издательских системах для создания документов большого объема со свойствами гипертекста (т. е. с системой ссылок); в системах управления документами; в системах подготовки электронных документов, позволяющих составлять гипертекстовые документы с возможностью осуществления навигации и др. [5].

Одним из наиболее перспективных направлений развития гипертекстовых систем является технология гипермедиа. Это соединение технологии гипертекста и технологии мультимедиа: в узлах гипертекста представляется разнородная, но семантически связанная информация – текст, рисунки, таблицы, графики, фотографии, видео, звук и т.п.

Другое важное направление развития гипертекстовых технологий – аналитическая обработка информации, которая подразумевает развитие методов автоматического установления смысловых связей как внутри предметной области, так и вне ее, что позволяет пользователю отбирать релевантные для него смысловые связи (например, смысловое упорядочивание документов). Это позволит решить одну из основных проблем пользователей современного Интернета – эффективный поиск информации.

Существующие сегодня инструменты поиска документальной информации в Интернет не обладают достаточной точностью выдачи и возвращают слишком много документов, из которых лишь небольшая часть действительно соответствует запросу пользователя, поэтому очевидно, что актуальность этой проблемы будет возрастать, так как объем документальной информации в Интернет постоянно увеличивается.

Основная идея гипертекстовых технологий заключается в том, что поиск документальной информации осуществляется с учётом множества связей, существующих между документами, а значит, более эффективно, чем при традиционных методах поиска. При этом доступ к информации организуется не последовательным просмотром текста, как в обычных информационно-поисковых системах, а путём движения от фрагмента к фрагменту.

Еще одно перспективное направление, которое мы можем выделить в рамках развития гипертекстовых технологий, – развитие технологий организации информационных ресурсов, распределенных в

сетях различных типов (локальных, корпоративных, глобальных) и, прежде всего, Web-технологии.

Сегодня степень освоения гипертекста лингвистами-практиками еще невелика. К вопросам, нуждающимся в дальнейшей разработке, можно отнести вопросы соотношения текста и гипертекста, гипертекста и его единиц. Однако исследование гипертекста далеко не всегда воспринимается в качестве объекта деятельности лингвистов по изучению принципов организации и понимания текста. Тем не менее, успешная разработка программ-оболочек гипертекста, организация удобного взаимодействия компьютерной программы с пользователем невозможна без привлечения фундаментальных знаний о языковой системе, которую могут обеспечить филологи-специалисты по прикладной лингвистике.

Список использованной литературы

- 1. Баранов А. Н.** Введение в прикладную лингвистику : Учеб. пособие // А. Н. Баранов – М. : Едиториал УРСС, 2003. – 360 с.
- 2. Потапова Р. К.** Новые информационные технологии и лингвистика: Учебное пособие. Изд. 4-е, стереотипное // Р. К. Потапова – М. : КомКнига, 2005. – 368 с.
- 3. Шемакин Ю. И.** Начала компьютерной лингвистики: Учеб. пособие // Ю.И. Шемакин – М. : Изд-во МГОУ, А/О «Росвузнаука», 1992. – 116 с.
- 4. Какое** применение нашли гипертекстовые технологии в Интернет? // Market-journal: ИТ в управлении [Электронный ресурс]. – Режим доступа : <http://www.market-journal.com/itvupravlenii/49.html>
- 5. Эпштейн В. Л.** Введение в гипертекст и гипертекстовые системы [Электронный ресурс]. – Режим доступа : <http://www.ipu.ru/publ/epstn.htm>

Белла М. В. Гіпертекст та його складові

Статья посвящена прикладным аспектам лингвистической науки. Рассматривается гипертекст как объект лингвистического исследования и гипертекстовые технологии, которые непосредственно связаны с эксплуатацией и развитием глобальной сети Интернет. Выделены собственно лингвистические аспекты компьютерной лингвистики и поставлен вопрос о возможности использования наработок теории гипертекста в лингвистических исследовательских программах.

Ключові слова: гіпертекст, гіперпосилання, анкор, гіпермедіа.

Белла М. В. Гипертекст и его составляющие

Статья посвящена прикладным аспектам лингвистической науки. Рассматривается гипертекст как объект лингвистического исследования и гипертекстовые технологии, которые непосредственно связаны с эксплуатацией и развитием глобальной сети Интернет. Выделены собственно лингвистические аспекты компьютерной лингвистики и

поставлен вопрос о возможности использования наработок теории гипертекста в лингвистических исследовательских программах.

Ключевые слова: гипертекст, гиперссылка, анкор, гипермедиа.

Bella M. V. Hypertext and its Components

Article deals the applied aspects of linguistic science. It observes the hypertext as an object of linguistic research and hypertext technologies which directly related to using and development of the Internet. Allocated strictly linguistic aspects of computational linguistics and raised the question about the use of hypertext theory developments in linguistic research programs.

Key words: hypertext, hyperlink, anchor, hypermedia

Стаття надійшла до редакції 09.02.2013 р.

Прийнято до друку 28.02.2013 р.

Рецензент – к. філол. н., доц. Гусева О. І.