

УДК 81: 004.55

ГИПЕРТЕКСТОВЫЕ ТЕХНОЛОГИИ ПРЕДСТАВЛЕНИЯ ЛЕКСИКИ

Е. Ю. Ченик

В статье рассматривается гипертекст как способ представления лексики в компьютерной лингвистике. Приведен ряд параметров, с использованием которых организуется гипертекстовое пространство.

Ключевые слова: компьютерная лексикография, гипертекст, электронный словарь.

Компьютерная лингвистика сегодня — это одно из важных направлений в конструктивном описании языка. Этим и определяется тот повышенный интерес к проблемам компьютерного моделирования функционирования языка и сферам его применения, который возник в лингвистике в последнее десятилетие. В данной статье рассматривается такой способ представления лексики в компьютерной лингвистике, как гипертекст. Идея гипертекста лежит в основе многих научных исследований, что связано, прежде всего, с успешным применением гипертекстовых систем во многих лингвистических системах, и, в частности, в лексикографических. В связи с этим одним из важных вопросов компьютерных исследований является разработка теории гипертекста, ее реализации в гипертекстовых технологиях, а также возможности использования их в лексикографии. Необходимость решения этих вопросов и обуславливает **актуальность** предлагаемого исследования.

Теория нелинейной организации информационных единиц, которые эксплицируются текстом, получила свое развитие в конце XX века. В современных научных дискуссиях все чаще звучат вопросы, связанные со способами организации электронного гипертекстового пространства, возможностями навигации внутри него [1], [3], [5], [8], [9]; а также обсуждаются технологии построения гипертекстов, описываются области применения гипертекстовых систем [6], [4]. Эти ключевые проблемы и определяют **научную новизну** работы.

В современной литературе встречаются различные толкования термина *гипертекст* и *гипертекстового пространства*. Приведем некоторые из них.

Определение гипертекста, данное J. Conklin в 1987 году, является важным и определяющим по своей сути: “Механизм, заключающийся в возможности связать отрывки текста, переходить от одного к другому, называется гипертекстом, или нелинейным текстом” [11].

А. Н. Баранов рассматривает гипертекст с нескольких позиций: “С одной стороны, это особый способ представления, организации текста, с другой — новый вид текста, противопоставленный по многим своим свойствам обычному тексту, сформированному в гутенберговской традиции книгопечатания. И, наконец, это новый способ, инструмент и новая технология понимания текста” [1, с. 31 – 32].

Определение гипертекстовой технологии Ю. И. Шемакина убедительно доказывает положение о том, что гипертекст является, прежде всего, системой ассоциативных

связей между элементами текста: “Гипертекстовая технология основана на такой машинной организации текстового материала, при которой лексические единицы представлены не только линейной последовательностью, но и ассоциативными связями между ними, определенными знаниями о предметной области. [...] При достаточно обширном материале с большим количеством связей возникает весьма сложное гипертекстовое пространство в виде своеобразной семантической сети, многомерно отражающей предметную область” [8, с. 104].

На наш взгляд, важным является тезис В. Л. Эпштейна о том, что “Гипертекст можно определить как нелинейную документацию, документацию, которая ветвится и взаимосвязывается, позволяя читателю исследовать содержащуюся в ней информацию, в последовательности, которую он сам выбирает” [9].

Принимая позицию Р. К. Потаповой, под гипертекстом мы понимаем “соединение смысловой структуры, структуры внутренних связей некоего содержания и технической среды, технических средств, дающих человеку возможность осваивать структуру смысловых связей, осуществлять переходы между взаимосвязанными элементами” [6, с. 149].

Теория гипертекста имеет свою историю, в разработку которой вносили свои идеи как лингвисты, так и программисты. Гипертексту в его нынешнем “компьютерном” понимании предшествует нелинейный текст “ручного” составления [2]. Одним из вариантов использования такого вида гипертекста являлись карточки, которые нумеровались и снабжались взаимными ссылками. Удобство таких карточек заключалось в разбиении больших записей на несколько более мелких, которые можно легко реорганизовать с учетом новой информации.

Другая форма представления “ручного” гипертекста — словари и энциклопедии, в которых статьи или определения содержат ссылки друг на друга. Следуя по этим ссылкам, читатель может получить более полную информацию.

Документы, где есть внутренние перекрестные ссылки и отсылки к другим документам, существуют уже много столетий. Таковы, например, сочинения Аристотеля и печатные издания Библии. Так, в современных изданиях Библии текст идет в сопровождении так называемых “параллельных мест”, ссылок, вынесенных на поля.

Таким образом, ценность гипертекста, заключающаяся в возможности предоставления дополнительной информации без отвлечения от “главной” темы повествования, была замечена не нашими современниками.

Термин *гипертекст* был введен в обращение Тедом Нельсоном (Ted Nelson) в 1965 г. для описания документов (например, представляемых компьютером), которые выражают нелинейную структуру идей, в противоположность линейной структуре традиционных книг, фильмов и речи. Более поздний термин “*гипермедиа*” близок к нему по смыслу, но он подчеркивает наличие в *гипертексте* нетекстовых компонентов, таких как анимация, записанный звук и видео [7].

Под “гипертекстом” Нельсон понимал не последовательную запись:

“Обычно процесс письма осуществляется последовательно по следующим двум причинам. Во-первых, потому, что он является производным от речи [...], которая не может не быть последовательной (так как у нас для этого только один канал), и, во-вторых, потому, что книги неудобно читать иначе как последовательно. Однако мыс-

ли образуют структуры, которые не являются последовательными — они связаны многими возможными переходами” [12].

По нашему мнению, гипертекстовое пространство организуется с использованием следующих параметров:

- 1) нелинейность;
- 2) взаимосвязь между элементами гипертекста, маршрут движения по которым выбирает сам пользователь;
- 3) способность всех элементов быть доступными в равной степени;
- 4) координация нескольких дополняющих друг друга структурных элементов, связанных по содержанию.

Рассмотрим признаки гипертекстового пространства на примере фрагмента разрабатываемого нами “Электронного словаря политической лексики”.

Нелинейность — наиболее характерная черта гипертекста — в словаре реализуется в разветвленной системе гиперссылок, которая может быть проиллюстрирована следующей схемой:

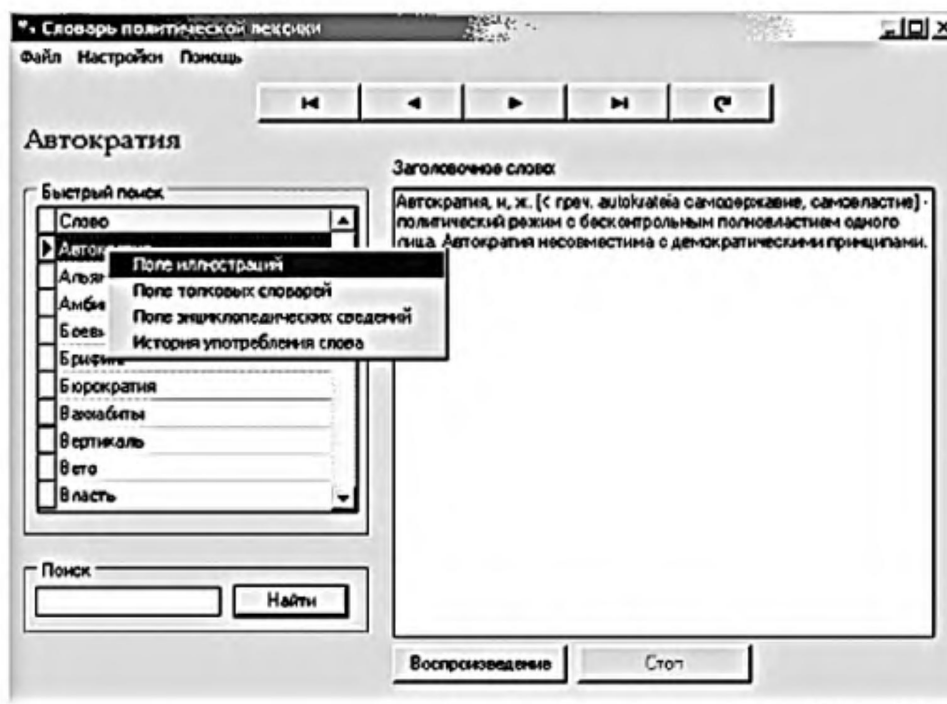
Нелинейность смысловых фрагментов обуславливает их *взаимосвязь*, навигационный путь, который пользователь может определять с помощью гиперссылок, являющихся названиями полей словаря и элементами его меню.

Способность всех элементов быть доступными в равной степени как один из важных организующих параметров реализуется в визуальном представлении каждого структурного элемента. Это показано на рисунке, иллюстрирующем интерфейс словаря (см. Рис.).



Четвертый параметр — *координация нескольких дополняющих друг друга структурных элементов, связанных по содержанию*, определяется в каждой лексеме. При этом фиксируются 5 полей, являющиеся структурными элементами словаря, например, в лексеме автократия (см. Рис) выделяются следующие поля:

- поле авторского толкования;
- поле иллюстраций;
- поле толковых словарей;
- поле энциклопедических сведений;
- поле истории слова.



Выводы

1. Одно из направлений современной лингвистики — это построение гипертекстовых систем, что является важным и актуальным не только для интерпретации текста, но и для лексикографического моделирования лексики.

2. Одной из перспектив современных гипертекстовых систем является развитие методов автоматического установления смысловых связей как внутри предметной области, так и вне ее, что позволяет пользователю отбирать релевантные для него смысловые связи.

3. 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.

Список литературы

1. Баранов А. Н. Введение в прикладную лингвистику: Учебное пособие. Изд. 2-е, исправленное. — М.: Едиториал УРСС, 2003. — 360 с.
2. Зефирова В. Л., Челноков В. М. Современные гипертекстовые системы // <http://www.computer-museum.ru/histsoft/newsyst.htm>. — 8.10.06
3. Купер И. Р. Гипертекст как форма коммуникации // <http://www.nir.ru/sj/sj/sj1-2-00kuper.html>. — 2.10.06.
4. Ованесбеков Л. Г. Технология построения гипертекстов. Диссертация на соискание ученой степени кандидата физико-математических наук. — М., 1993.
5. Орехов С. И. Гипертекстовый способ организации виртуальной реальности. // www.omsk.edu/article/vestnik-omgru-21.pdf. — 8.10.06.
6. Потапова Р. К. Новые информационные технологии и лингвистика: Учебное пособие. Изд. 4-е, стереотипное. — М.: КомКнига, 2005. — 368 с.
7. Толковый словарь по вычислительной технике/ Пер. с англ. — М.:Издательский отдел «Русская редакция» ТОО «Channel Trading Ltd», 1995. — 496 с.
8. Шемакин Ю. И. Начала компьютерной лингвистики: Учеб. пособие. — М.: Изд-во МГОУ, А/О «Росвузнаука», 1992, 116 с.
9. Эпштейн В. Л. Введение в гипертекст и гипертекстовые системы// <http://www.ru/publ/epstn.htm> — 8.10.06.
10. Bush V. As we may think // Atlantic monthly. 1945. V. 176. №1.
11. Conklin J. Hypertext: an introduction and survey // Computer. 1987. Vol.20. No 9.
12. Nelson, T. N. A file structure for the complex, the changing, and the indeterminate // in: ACM 20th National Conference — Proceedings. Cleveland, Ohio, 1965. — pp. 84 – 100.

Ченик О. Ю. Гіпертекстові технології представлення лексики.

У статті розглядається гіпертекст як спосіб представлення лексики в комп'ютерній лінгвістиці. Приведений ряд параметрів, з використанням яких організовується гіпертекстовий простір.

Ключові слова: комп'ютерна лексикографія, гіпертекст, електронний словник.

Chepik E. Y. Hypertext technologies of lexis representation.

In article the hypertext is considered as a way of representation of lexicon in computer linguistics. A number of parameters which use the hypertext space will be organized is given.

Key words: a computer lexicography, hypertext, the electronic dictionary.

Статья поступила в редакцию 30 октября 2006 г.