

260 сокращений, которые встречались в большем числе исследованных словарей сокращений. Ниже приводятся первые 10 включенных в создаваемый словарь сокращений.

A-bus

галоўная шына (магістраль) у мікрапрацэсары

AC = alternating current

пераменны ток

ACC = ACCUMULATOR

аккумулятар

ADA

Ада; мова праграмавання ADA

ADP = automatic data processing

аўтаматычная апрацоўка даных

Advanced Power Management (APM)

удасканаленае кіраванне сілкаваннем

AGP = accelerated graphics port

паскораны графічны порт

AIFF = audio interchange file format

фармат файлаў абмену аўдыёданымі; фармат файлаў лічбавага гуку

AIO = asynchronous input-output

асінхронны ўвод-вывад

Algol

Алгол (ад ALGOrithmic Language); мова праграмавання ALGOL

Н. А. Метлицкая

**ЭКСПЕРИМЕНТ ПО АВТОМАТИЧЕСКОМУ СОЗДАНИЮ
АНГЛОЯЗЫЧНЫХ РЕКЛАМНЫХ ТЕКСТОВ
ПО КОСМЕТИКЕ И ПАРФЮМЕРИИ**

Главной задачей разрабатываемой компьютерной системы является автоматическое порождение англоязычных рекламных текстов по косметике и парфюмерии трех предметных областей: губная помада, тушь для ресниц, шампунь. Для создания данной системы было разработано необходимое лингвистическое и алгоритмическое обеспечение, наделяющее компьютерную систему необходимыми знаниями о структуре и содержании порождаемого текста, а также лингвистическими знаниями, которые позволяют синтезировать правильный на всех языковых уровнях (синтаксическом, лексическом, морфологическом, семантическом) текст.

Лингвистическое обеспечение системы представлено лингвистической базой знаний, которая включает следующие компоненты:

1) главные опорные слова (ГОС) и второстепенные опорные слова (ВОС) исследованных текстов, отражающих их основное статическое содержание;

2) электронный словарь лексических единиц, содержащий 555 словарных статей с указанием морфологических и семантических сведений о каждой лексической единице;

3) семантико-синтаксические формулы исследованных рекламных текстов на формальном языке СЕМСИНТ, на основании которых происходит построение синтаксической структуры предложений, а также их лексическое наполнение с опорой на электронный словарь.

Далее был разработан алгоритм порождения рекламных текстов, который определяет порядок следования операций, выполняемых при планировании содержания текста, а также при выборе лексических и грамматических средств, представленных в лингвистической базе знаний, и их оформление в связный текст. В самом общем виде порождение текста в соответствии с данным алгоритмом предполагает последовательное решение следующих задач:

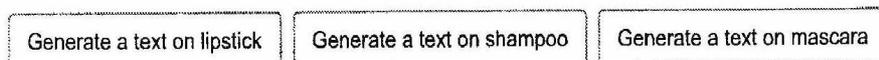
1) определение семантико-синтаксического содержания порождаемого текста (чтение семантико-синтаксической формулы текста);

2) лексическое заполнение семантико-синтаксической формулы текста из электронного словаря в соответствии с семантическими кодами слов;

3) морфологическое оформление каждой лексической единицы в соответствии с указанной в формуле грамматической информацией и грамматическими данными, содержащимися о ней в словарной статье электронного словаря.

Программа для порождения англоязычных рекламных текстов по косметике и парфюмерии реализована на языке программирования JavaScript.

Пример работы программы представлен на рисунке 1.



Dry scalp care shampoo

Clinically proven to enliven scalp natural moisture.

With rosemary oil.

Forget about sensitive, dry scalp by enlivening your scalp moisture.

Dry scalp care shampoo with rosemary oil brings dual-action thickening formula that targets dryness and brings you 100% healthy hair.

This lift shampoo moisturizes your dry scalp and leaves your hair perfectly chic.

Dual-action complex appeases dry scalp and itch, while giving extra conditioners for your hair.

Improves scalp in 2 weeks.

Пример работы программы автоматического порождения рекламных текстов по косметике и парфюмерии

Структура окна программы очень проста для пользователя и представляет собой три «кнопки», каждая из которых отвечает за создание рекламного текста определенной предметной области: *lipstick* ‘губная помада’, *shampoo* ‘шампунь’, *mascara* ‘тушь для ресниц’. При каждом очередном нажатии на определенную кнопку программа синтезирует новый текст по соответствующей тематике. Для текстов каждой предметной области база знаний системы включает по 10 вариантов семантико-

синтаксических формул, из которых при каждом нажатии кнопки датчиком случайных чисел выбирается одна. Лексическое наполнение выбранной формулы также определяется датчиком случайных чисел, но с учетом семантических кодов слов. Таким образом, очевидным преимуществом созданной системы можно считать очень большое разнообразие порождаемых текстов. Ведь синтезированные тексты отличаются друг от друга не только синтаксической структурой входящих в них предложений, но и словами, которые их образуют.

Анализ порожденных системой рекламных текстов на лексическом, морфологическом и синтаксическом уровнях свидетельствует о том, что благодаря разработанной лингвистической базе знаний в целом нам удалось достичь достаточно высокого лингвистического качества создаваемых системой текстов. Во-первых, синтезированные тексты представляют собой корректные семантически произведения, которые соответствуют заявленному пользователем требованию, т.е. порождению текста необходимой тематики (*губная помада, шампунь или тушь для ресниц*). Во-вторых, все порожденные тексты правильны с точки зрения грамматики (осуществляется согласование подлежащего и сказуемого в лице и числе, правильно употребляются необходимые предлоги и частицы, и т.п.).

Следует обратить внимание на то, что создаваемые тексты полностью соответствуют требованиям, предъявляемым к рекламному тексту как особому типу текстов, т.е. отражают его лингвистические особенности, которые обнаруживаются как в его своеобразной структурной организации, так и в содержании. А именно:

- создаваемые тексты краткие, ибо в рекламной коммуникации реализуется принцип «максимум информации при минимуме слов»;
- все тексты имеют характерную для рекламного текста структуру (включают название рекламируемого продукта, заголовок и основной рекламный текст);
- создаваемые тексты в значительной степени экспрессивны и эмоциональны, что достигается, прежде всего, благодаря наличию лексических единиц эмоционально-оценочного характера.
- в создаваемых текстах достаточно ярко представлен один из основных параметров рекламного дискурса – *диалогичность*. В массмедиальном дискурсе вовлечение адресата в диалог и установление контакта с ним осуществляется такими лексическими средствами, как *личные и притяжательные местоимения 1 и 2-го лица*.

Однако в ходе эксперимента в системе были выявлены и некоторые недочеты.

1. В некоторых случаях датчик случайных чисел выбирает одно и то же слово для рядом стоящих в предложении слов с одинаковым семантическим кодом. В результате в порожденных текстах можно встретить предложения следующего вида: *The caring amazing applicator brings a carefree, carefree coverage* ‘Изумительный заботливый аппликатор обеспечивает **беззаботное, беззаботное** нанесение’.

2. В процессе анализа обнаружены некоторые нарушения норм лексической сочетаемости. Например, не корректны с точки зрения сочетаемости выражения типа *colorful color* ‘цветной цвет’, *ideally magnificent* ‘идеально великолепный’, *comfortable temptation* ‘комфортное обольщение’, *fascinate ideally* ‘идеально изумлять’.

Полученные во время тестирования результаты позволяют наметить дальнейшие пути повышения качества работы системы, учитывающие выявленные ошибки и связанные, в первую очередь, с усовершенствованием лингвистической базы знаний. В частности, действенной нам видится корректировка семантической классификации лексических единиц словаря системы с учетом лексико-семантической и стилистической сочетаемости слов в связном отрезке речи.

В целом анализ текстов, полученных в ходе тестирования программы, показал, что разработанная система является эффективной, позволяет оперативно порождать рекламные сообщения по выбранной пользователем тематике.

В. Ф. Руцкая

К ВОПРОСУ О ТИПОЛОГИИ ТОПОНИМИЧЕСКИХ СИСТЕМ

Топонимика – важный раздел ономастики, исследующий топонимы – географические названия. Топонимы подвергаются анализу с точки зрения их этимологии, значения, структуры, функционирования в языке. В ходе анализа определяются их ареалы и этапы диахронического развития. Под топонимией понимается совокупность топонимов на определенном географическом пространстве. Топонимия содержит ценные данные для описания исторического развития языка, изучения лексикологии, диалектологии, лингвогеографии. Она помогает описать характер взаимодействия языков, типы языковых контактов, восстановить черты исторического прошлого народов. Поэтому топонимику справедливо называют «языком земли». Среди топонимов выделяют: ойконимы – названия населенных мест; аstownимы – названия городов; гидронимы – названия рек; дрионимы – названия лесов; оронимы – названия гор; урбанонимы – названия внутригородских объектов; годонимы – названия улиц; агоронимы – названия площадей; дромонимы – названия путей сообщения. Исходя из величины объектов топонимы подразделяются на макротопонимы и микротопонимы. Существуют также антропотопонимы – названия географических объектов, произведенных от личного имени. По составу топонимы подразделяются на однословные, словосочетания и топонимические фразеологизмы. Они могут иметь простую и сложную структуру. Топонимы могут объединять элементы нескольких языков, относящихся к различным историческим периодам, как, например, *London* – латино-кельтский комплекс (из *Londinium*). Они могут восходить к языкам-субстратам тех народов, которые жили на этой территории в прошлом.