

Н. Г. Швец
Минск, МГЛУ

ТЕЗАУРУСНЫЙ МЕТОД ОПИСАНИЯ ИЛЛЮСТРАЦИИ РЕКЛАМНОГО ОБЪЯВЛЕНИЯ

Исследование механизмов порождения и восприятия смысла креолизованных текстов (в том числе и печатных рекламных объявлений), в которых в рамках единого сообщения сплетены вербальные и невербальные компоненты, вызывает большой интерес. Вследствие постоянного увеличения объема семиотически неоднородных массивов текстов современное общество испытывает острую потребность в их автоматизированной обработке. Важной задачей разрабатываемых технологий является поиск оперативных способов формального представления такой информации. В статье предлагается использование тезаурусного способа представления содержания изображения, представляющего собой описание иллюстрации в виде некоторого набора слов и отношений между этими словами в рамках некоторой предметной области. Преимуществом данного метода является то, что он позволяет учесть как доминирующие, так и второстепенные черты изображения.

При создании компьютерной системы, моделирующей понимание текста и изображения рекламного объявления человеком, возникает проблема поиска способа эффективной формализации изображения. Необходимо придать иллюстрации «неизобразительный» вид, ибо соответствующие алгоритмы приспособлены для переработки лишь неких символических описаний.

Как известно, одним из методов формального представления знаний о некотором фрагменте реальной действительности является тезаурусный способ представления информации о той или иной предметной области. В нашем исследовании мы предлагаем использовать этот подход для описания изображения.

Анализ 1000 рекламных объявлений на тему «Парфюмерия и косметика» показал, что в 99 % случаев на иллюстрации обязательно присутствует изображение человека. Поэтому в целях нашего исследования были выявлены наиболее употребительные существительные, используемые при описании внешнего облика человека и установлены системные отношения между ними.

Особенности лексического состава описания определяются, прежде всего, структурой самого описания, которое служит для детального изображения состояний действительности с подробным перечислением ряда одновременно существующих объектов и их признаков. Функции объектов выполняют слова, обозначающие основные компоненты внешнего описания, которыми являются:

- 1) тело человека и его части (соматическая лексика);
- 2) мимика, поза, жесты (кинетическая лексика);
- 3) предметы одежды (вестиальная лексика).

Рассмотрим структурные отношения в группе соматической лексики. Для ее систематизации подходит отношение меронимии. *Меронимия* – это способ ассоциативной сетевой организации понятий на основе отношения ЦЕЛОЕ–ЧАСТЬ. Это отношение обладает следующими свойствами:

- 1) оно транзитивно: ЦЕЛОЕ (ЦЕЛОЕ (X)) = ЦЕЛОЕ (X);
- 2) дескрипторы-части наследуют отношение АССОЦИАЦИЯ: АСЦ (ЦЕЛОЕ (X)) = АСЦ (X);
- 3) отношение ЧАСТЬ наследуется видовыми дескрипторами: ЧАСТЬ (ВЫШЕ (X)) = ЧАСТЬ (X);
- 4) отношение ВЫШЕ – НИЖЕ поглощается отношением ЦЕЛОЕ–ЧАСТЬ: ЦЕЛОЕ (ВЫШЕ (X)) = ЦЕЛОЕ (X), ВЫШЕ (ЦЕЛОЕ (X)) = ЦЕЛОЕ (X).

К характеризующим лексическим единицам, описывающим внешние признаки, обычно относятся слова, указывающие цвет, форму, толщину, величину. В зависимости от типа характеризуемого объекта набор таких признаков меняется. Можно выделить следующие признаки соматемы «Лицо»:

- 1) цвет и окраска кожного покрова (*загорелое, бледное, землистое* и др.);
- 2) форма (*круглое, овальное, треугольное* и др.);
- 3) величина (*большое, крупное, маленькое* и др.);
- 4) толщина (*худое, сытое, тонкое* и др.);
- 5) характер кожного покрова (*гладкое, холменое, морщинистое* и др.);
- 6) характерная деталь (*длинноносое, курносое, скуластое, узколобое, большеглазое* и др.):
 - 6.1) величина и форма носа;
 - 6.2) форма скул;
 - 6.3) величина и форма губ;
 - 6.4) величина и форма лба;
 - 6.5) величина, форма и цвет глаз;
 - 6.6) наличие или отсутствие волосяного покрова на лице и голове;
- 7) состояние здоровья и возраста (*здоровое, усталое, молодое, свежее* и др.);

- 8) признак внешней выраженности интеллектуальных способностей (*умное, мудрое, простое, глупое* и др.);
- 9) общее впечатление:
 - 9.1) положительные оценки (*красивое, миловидное, приятное* и др.);
 - 9.2) отрицательные оценки (*неинтересное, невзрачное, непривлекательное* и др.);
- 10) доминирующая черта характера, особенность поведения (*благородное, честное, открытое, важное, самодовольное* и др.);
- 11) выражение черт характера и свойств личности (*волевое, властное, строгое, гордое, скромное, доброе, кроткое, хитрое, угодливое, насмешливое* и др.);
- 12) выражение доминирующих психоэмоциональных состояний (*живое, бодрое, радостное, счастливое, сердитое, равнодушное* и др.);
- 13) представление о социальном статусе личности (*благородное, интеллигентное* и др.).

Признаки других соматем представлены в табл. 1.

Т а б л и ц а 1

Признаки соматем

Соматема	Признак	Пример
Волосы	цвет форма длина густота мягкость толщина влажность блеск сила укладка здоровье тяжесть чистота общая положительная оценка общая отрицательная оценка	<i>русые прямые короткие редкие жесткие тонкие мокрые блестящие слабые растрепанные посеченные легкие грязные восхитительные некрасивые</i>
Глаза	цвет форма величина блеск возраст здоровье эмоции доброта общая положительная оценка общая отрицательная оценка образ из сферы флоры образ из сферы фауны образ на основе экзотизмов	<i>голубые круглые маленькие блестящие молодые воспаленные восторженные злые притягательные невзрачные ореховые змеиные цыганские</i>
Губы	цвет форма толщина величина физическое состояние влажность	<i>красные оттопыренные пухлые большие свежие сухие</i>

Зубы	цвет форма величина расположение физическое состояние	<i>белые ровные мелкие редкие жемчужные</i>
Кожа	цвет толщина пигментация возраст особое состояние влажность тип эластичность	<i>белая тонкая смуглая молодая гусиная сухая комбинированная дряблая</i>
Нос	форма величина цвет характер кожного покрова длина образ на основе сходства по форме с носами животных и птиц	<i>курносый большой красный блестящий короткий орлиный</i>
Ресницы	цвет величина длина густота	<i>черные большие длинные редкие</i>
Тело (фигура)	величина толщина тяжесть сложение гибкость длина нагота цвет чистота здоровье возраст сила	<i>щуплое худое легкое стройное гибкое короткое полубнаженное розовое грязное больное молодое дюжее</i>
Женщина	возраст осанка рост телосложение сила нагота чистота здоровье ухоженность интеллект психоэмоциональное состояние характер социальный статус общая положительная оценка общая отрицательная оценка	<i>молодая грациозная высокая стройная сильная голая чистая здоровая ухоженная умная веселая властная интеллигентная привлекательная несимпатичная</i>

Очень часто на иллюстрации рекламного объявления изображается предмет рекламы (флакон духов, баночка крема и др.). Для его описания используются такие признаки, как форма, цвет, содержимое, название, фирма-изготовитель, а для представления связей между дескрипторами – такие отношения, как ЦЕЛОЕ–ЧАСТЬ, АССОЦИАЦИЯ, ОБЪЕКТ–ХАРАКТЕРНОЕ СВОЙСТВО и др.

Результаты тезаурусного представления иллюстрации рекламного объявления представлены в табл. 2.

Т а б л и ц а 2

Тезаурусное представление иллюстрации рекламного объявления

Дескриптор (уровень тезаурусной иерархии)	Мероним (уровень тезаурусной иерархии)	Признак (уровень тезаурусной иерархии)	Действие (уровень тезаурусной иерархии) (ассоциация)
<i>женщина (1)</i>	<i>голова (2)</i> <i>руки (2)</i>		
<i>флакон (1)</i>	<i>шампунь (2)</i>	<i>зеленый (2)</i> <i>прямоугольный (2)</i>	
<i>фон (1)</i>	<i>фрукты (2)</i>		
<i>схема-пояснение (1)</i>	<i>волос (2)</i> <i>вещество (2)</i>		
<i>голова (2)</i>	<i>волосы (3)</i>		
<i>руки (2)</i>			<i>завязывать волосы в узел (3)</i> <i>(сила волос)</i>
<i>шампунь (2)</i>		<i>fructis (фруктус) (3)</i> <i>garnier paris (3)</i>	
<i>волос (2)</i>	<i>корень (3)</i>		
<i>вещество (2)</i>	<i>молекулы (3)</i>	<i>активное (3)</i> <i>зеленое (3)</i>	
<i>фрукты (2)</i>	<i>дольки (3)</i>		
<i>волосы (3)</i>		<i>блестящие (4)</i> <i>сильные (4)</i> <i>густые (4)</i> <i>шелковистые (4)</i> <i>длинные (4)</i> <i>гладкие (4)</i>	
<i>корень (3)</i>		<i>здоровый (4)</i> <i>сильный (4)</i>	
<i>молекулы (3)</i>		<i>зеленые (4)</i> <i>красные (4)</i>	<i>проникать в корень (4)</i> <i>(сила и блеск волос)</i>
<i>дольки (3)</i>		<i>зеленые (4)</i> <i>желтые (4)</i>	

Тезаурус является иерархической структурой, состоящей из дескрипторов различных уровней. Первый уровень тезауруса иллюстрации анализируемого РО (уровень тезауруса в табл. 2 обозначается цифрой в скобках) образуют следующие дескрипторы: *женщина*, *флакон*, *фон* и *схема-пояснение*.

Между дескриптором первого уровня *женщина* и дескрипторами второго уровня *голова* и *руки* устанавливаются системные отношения типа ЦЕЛОЕ (холоним) – ЧАСТЬ (мероним).

Дескриптор первого уровня *флакон* связан с дескриптором второго уровня *шампунь* отношением ЦЕЛОЕ–ЧАСТЬ, а с дескрипторами второго уровня *зеленый* и *прямоугольный* – отношением ОБЪЕКТ–ПРИЗНАК.

Между дескриптором первого уровня *схема-пояснение* и дескрипторами второго уровня *волос* и *вещество* устанавливаются отношения типа ЦЕЛОЕ–ЧАСТЬ.

Третий уровень тезауруса образуют следующие дескрипторы: *волосы, завязывать волосы в узел, сила волос, Fructis (Фруктис), Garnier Paris, корень, молекулы, активное, зеленое, дольки, волосы.*

Дескрипторами четвертого уровня тезауруса являются: *блестящие, сильные, густые, шелковистые, длинные, гладкие, здоровый, сильный, зеленые, красные, желтые, проникать в корень, сила и блеск волос.*

Таким образом, тезаурусный метод позволяет описать иллюстрацию в виде некоторого набора слов (дескрипторов) и отношений между этими лексическими единицами в рамках некоторой предметной области. Основным преимуществом данного подхода является то, что он позволяет учесть как доминирующие, так и второстепенные черты изображения.

This article discusses one of the possible ways to formally present an illustration of an advertisement. The use of the thesaurus method is proposed, which allows one to take into account both dominant and secondary features of the image.