

ПРОБЛЕМЫ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ТЕКСТОВ

Д. В. Березюк

Минск, МГЛУ

ТЕКСТ ОТЗЫВА КАК РЕЗУЛЬТАТ ОТРАЖЕНИЯ РЕПУТАЦИИ РЕКЛАМНОГО АГЕНТСТВА

Рассматривается лингвистический аспект формирования репутации рекламного агентства на материале англоязычных отзывов специализированного интернет-сайта. Особенностью отобранных текстов является то, что представленная в них информация содержит оценочную и эмоциональную составляющие. В ходе анализа были выделены группы лексических единиц, положенных в основу формирования оценочных суждений пользователей о его репутации. Был сформирован словарь положительной и отрицательной оценочной лексики, список аспектных категорий и ключевых слов предметной области «Рекламное агентство», а также дополнительные списки слов для более полного анализа степени оценочности текста.

Согласно толковому словарю И. С. Ожегова под репутацией понимается «приобретаемая кем-нибудь или чем-нибудь общественная оценка; общее мнение о качествах, достоинствах и недостатках кого-нибудь или чего-нибудь» [1]. В Энциклопедическом словаре В. Ф. Брокгауза и И. А. Ефрона слово *репутация* (от франц. *réputation*) трактуется как «создавшееся общее мнение о качестве, достоинствах и недостатках кого-либо, чего-либо» [2]. Значение слова *репутация* в Словаре бизнес-терминов определяется следующим образом: «это установившееся общественное мнение о качествах, достоинствах и недостатках определенного лица или компании» [3]. Таким образом, ключевыми единицами приведенных выше определений являются словосочетания *общественная оценка*, *общее мнение* о ком-либо или чем-либо. С развитием Интернета огромное количество мнений ее пользователей о различных объектах, явлениях, событиях размещаются на специализированных сайтах отзывов, форумах, в блогах, гостевых книгах, социальных сетях и т.п. Поэтому тщательный анализ и обобщение мнений большого числа людей о персонале рекламного агентства, его услугах, качестве и сроках выполнения рекламных продуктов позволяет сделать вывод о его репутации.

При рассмотрении мнения (оценки) в традиционной лингвистике особое внимание уделяется исследованию семантических и прагматических особенностей оценочных значений. Оценка имеет объективный характер как факт языка, т.е. владение словами с оценочной семантикой присуще всем носителям языка. Речевой уровень оценки связан, главным образом, с ее функциональными характеристиками. Согласно функциональному подходу оценка играет чрезвычайно важную роль при непосредственной коммуникации, в речи, где затрагиваются интересы участников общения.

Материалом исследования послужили несколько десятков текстов англоязычных отзывов об услугах рекламных агентств, взятых со специализированного сайта <http://www.clutch.co>. Особенностью отобранных текстов является то, что представленная в них информация носит, преимущественно, субъективный характер, т.е. содержит оценочную и эмоциональную составляющие. Субъективная информация неоднородна, может быть как избыточной, так и недостаточной. Основываясь на личном опыте, пользователи отмечают как положительные, так и отрицательные стороны заданного объекта и его аспектов, т.е. выражают к ним личное отношение (мнение).

В ходе анализа отобранного для исследования материала были выделены совокупности лексических единиц, позволяющих сформировать оценочные суждения пользователей о самом рекламном агентстве, его персонале, качестве оказываемых им услуг и т.п., т.е. о его репутации. Был сформирован словарь оценочной лексики, список аспектных категорий и ключевых слов предметной области «Рекламное агентство и оказываемые им услуги», а также дополнительные списки слов для дальнейшего анализа степени оценочности текста. Выделение и классификация оценочной лексики проводились с опорой на знание английского языка. Полученные результаты фрагментарно представлены в табл. 1–4.

Словарь оценочной лексики включает словоформы и состоит из двух частей: перечня положительных оценочных слов и перечня отрицательных оценочных слов. Каждая единица лексики размечена и наделена определенным семантическим весом от +3 до -3 (см. табл. 1).

Т а б л и ц а 1

Фрагмент словаря оценочной лексики, составленного на основе анализа текстов англоязычных отзывов о рекламных агентствах и оказываемых ими услугах

Положительная оценочная лексика	Вес	Отрицательная оценочная лексика	Вес
above-average	1	angry	-2
acceptable	1	annoyed	-2
adjustable	1	annoying	-2
advantage	1	anxious/anxiously	-2
advantageous/advantageously	1	arrogance	-2
affordable/affordably	1	arrogant	-2
amaze/amazed/amazes	1	ashamed	-3
amazing	2	austere	-2
...

В ходе анализа англоязычных отзывов о работе рекламных агентств были определены слова, обозначающие основной объект (*advertising agency*), аспектные категории (*advertisement, advertisement campaign, advertisement product, price, product placement, service*) и их ключевые слова (например, *brand, trade mark, logo, staff* и т.д.) (см. табл. 2). Выделение аспектных категорий и связанных с ними слов помогает исключить те лексические единицы, которые не влияют на оценку конкретной услуги.

Т а б л и ц а 2

Фрагмент списка аспектных категорий и ключевых слов, составленного на основе текстов англоязычных отзывов о рекламных агентствах и оказываемых ими услугах

Объект	Аспектная категория	Ключевое слово
Advertising agency	common	atmosphere conditions quality ...
	advertisement	to inform to suggest ...
	advertisement campaign	mass-media Internet ...
	advertisement product	bill board banner ...
	location	place

Исследование отобранного массива текстов англоязычных отзывов показало, что их авторы применяют различные интенсификаторы, т.е. слова, увеличивающие или уменьшающие вес оценочного слова, например, *almost, pretty, very, too, so, rather, less*. Выявленные интенсификаторы были ранжированы по следующей шкале: низкий, средний, высокий. Знак @ перед списком слов означает глобальный поиск этого слова по списку. Таким образом, если слово проверяется по @списку, то оно автоматически сопоставляется со всеми включенными в список лексическими единицами до тех пор, пока не будет найдено соответствие (см. табл. 3).

Т а б л и ц а 3

Фрагмент списка интенсификаторов, составленного на основе анализа текстов англоязычных отзывов о рекламных агентствах и оказываемых ими услугах

№ п/п	@высокая степень усиления/уменьшения	№ п/п	@средняя степень усиления/уменьшения	№ п/п	@низкая степень усиления/уменьшения
1	absolutely	1	actually	1	almost
2	amazingly	2	admittedly	2	bit
3	awfully	3	enough	3	fairly
4	bitterly	4	even	4	little
4	certainly	5	frankly	5	nearly
5	completely	6	positively	6	poorly
6	considerably	7	pretty	7	quite
...

В результате анализа англоязычных отзывов о работе рекламных агентств также были выявлены слова-инверторы (см. табл. 4). Это местоимения, предлоги и частицы, меняющие направление оценочного веса слова на противоположное. Например, *pleasure* (+1) – *without pleasure* (-1); *attractive* (+1) – *less attractive* (-1), *special* (+1) – *nothing special* (-1), *bad* (-1) – *not bad* (+1), *lush* (+2) – *non-lush* (-2).

Т а б л и ц а 4

Список слов-инверторов, составленный на основе текстов англоязычных отзывов о рекламных агентствах и оказываемых ими услугах

№ п/п	Инвертор
1	least
2	less
3	never
4	no
5	non
6	not
7	nothing
8	without

Выделенные четыре списка лексических единиц могут составить лингвистическую базу данных системы автоматического анализа репутации рекламного агентства.

ЛИТЕРАТУРА

1. [Электронный ресурс]. – Режим доступа : <http://www.ozhegov.org/words/30637.shtml>. – Дата доступа : 12.05.2016.
2. [Электронный ресурс]. – Режим доступа : <http://www.vehi.net/brokgauz>. – Дата доступа : 12.05.2016.
3. [Электронный ресурс]. – Режим доступа : <http://www.klerk.ru/slovar/bus/term/63865/>. – Дата доступа : 15.05.2016.

The article deals with the linguistic aspect of advertising agency reputation formation. Four types of lexical units are under consideration as they can form certain opinion about the object of investigation.

П. А. Маракулина

Минск, МГЛУ

АВТОМАТИЧЕСКАЯ ЭКСПЕРТИЗА АНГЛОЯЗЫЧНОГО ХУДОЖЕСТВЕННОГО ТЕКСТА С ВЫРАЖЕНИЕМ КОНФЛИКТНОЙ СИТУАЦИИ

Рассматриваются основные принципы построения системы автоматической экспертизы англоязычного художественного текста с выражением конфликтной ситуации. В основу работы такой системы положен определенный набор языковых маркеров конфликтности. В полном объеме выделенные лексические единицы составили лингвистическое обеспечение экспертной системы, на основе которого была разработана формальная модель, реализованная на языке *JAVA*. В ходе выполнения компьютерного эксперимента программа не только констатировала наличие либо отсутствие в англоязычных текстах конфликтных речевых ситуаций, но и определяла степень ее выраженности (высокая, средняя, низкая степень).

Под речевым конфликтом понимается коммуникативная ситуация, в которой происходит столкновение двух сторон по поводу разногласия интересов, целей, взглядов и т.п., в результате чего одна из сторон (или обе стороны) предпринимают речевые действия, направленные против интересов партнера по коммуникации. Как разновидность речевого общения речевой конфликт характеризуется употреблением определенных вербальных и невербальных маркеров, сигнализирующих о разной степени выраженности конфликтной ситуации. Роль маркеров речевого конфликта выполняют обладающие конфликтными свойствами языковые единицы, относящиеся к лексическому и грамматическому уровням языка. Кроме того, речевой конфликт часто усиливается такими невербальными компонентами, как мимика, жесты, позы и различные действия коммуникантов.

Языковые маркеры конфликтности были положены в основу создания системы автоматической экспертизы англоязычного художественного текста с выражением конфликтной ситуации. Материалом исследования послужили