

2. *Захаров, В. П.* Информационные системы (документальный поиск) / В. П. Захаров. – СПб., 2002. – 188 с.

The article analyses the procedure of creating a search image for English publicist texts, which is based on the algorithm of identifying key or prop words.

А. Л. Холод
Минск, МГЛУ

ЛИНГВИСТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ СИСТЕМЫ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ СРЕДСТВ ВЫРАЖЕНИЯ СТРАТЕГИИ «ФЛЕЙМ» В ВИРТУАЛЬНОЙ КОММУНИКАЦИИ

В статье рассматриваются десять основных классов лексических единиц, однозначно указывающих на наличие в испаноязычных твитах деструктивной стратегии «флейм». Отмечается, что деструктивный эффект высказывания может быть усилен путем использования особых идеографических знаков (эмодзи), передающих в виртуальной коммуникации определенную эмоцию или эмоциональное состояние пользователя. В работе также приводятся второстепенные маркеры реализации флейма, которые проявляют свой деструктивный потенциал в случае, если их количество в твите превышает две единицы. Выявленные в ходе исследования маркеры составили лингвистическую базу данных системы автоматического обнаружения и блокировки твитов с флеймом.

Периодическое отклонение от норм и правил кооперативного общения в Интернете, а также агрессивное речевое поведение его пользователей явно указывает на деструктивный характер виртуальной коммуникации, в рамках которой можно выделить различные деструктивные стратегии, и в частности флейм. Флейм представляет собой эмоциональное обсуждение определенной темы, сопровождающееся нарушением принципов конструктивной дискуссии. Флейм относится к запрещенным сетевым коммуникативным стратегиям, однако особенности идентификации пользователей превращают контроль за соблюдением правил сетевого общения в довольно сложную задачу. Модераторы блогов, форумов, социальных сетей и других сетевых средств общения блокируют сообщения агрессивных пользователей вручную. В то же время выявленные в полном объеме вербальные и невербальные средства реализации данной стратегии могут быть положены в основу создания системы автоматического обнаружения и блокировки сообщений с флеймом.

Материалом исследования послужили 278 испаноязычных аккаунтов социальной сети Твиттер (общим объемом более 250 000 твитов), из которых детально было проанализировано 5000 твитов и выделено 720 твитов, содержащих признаки деструктивной стратегии «флейм». Общение на этих аккаунтах характеризуется повышенной эмоциональностью и экспрессивностью, а также категоричностью высказываемых точек зрения. На первом этапе исследования отобранных твитов были определены основные классы вербальных и невербальных маркеров, однозначно указывающих на наличие деструктивной стратегии, важных для разработки процедуры автоматического

определения средств ее реализации. К ним относятся табуированная и обценная лексика (однокомпонентная, двухкомпонентная, трехкомпонентная и многокомпонентная); пейоративно-оценочная лексика; слова, называющие понятия, связанные с интимной жизнью человека; слова, называющие понятия, связанные с сексуальной ориентацией и сексизмом; слова, называющие животных; англицизмы с негативной коннотацией; эмодзи. Помимо однозначных маркеров флейма, также были выявлены второстепенные единицы, чей негативный потенциал проявляется в комплексном употреблении. Среди них можно выделить личные и вопросительные местоимения; слова-интенсификаторы; некоторые пунктуационные знаки и графические средства.

На втором этапе исследования из отобранного массива испаноязычных твитов были выделены конкретные лексические единицы, вошедшие в состав указанных выше классов. Выделение маркеров флейма проводилось с опорой на знание испанского языка, а также с учетом ряда учебных и электронных источников. Внутри некоторых классов лексические единицы были разделены по группам в зависимости от числа входящих в них компонентов. Как показало исследование, важную роль среди всех маркеров играют лексико-семантические индикаторы флейма. При анализе отобранного массива твитов особое внимание уделялось таким маркерам деструктивной стратегии, как лексика с эмоционально-экспрессивной стилистической окраской, грубо-просторечная и табуированная лексика, основной функцией которых является стимулирование у речевого партнера негативных эмоций. Фрагментарно представленные в табл. 1–5 лексические единицы указывают на явное наличие флейма.

Т а б л и ц а 1

Фрагмент списка однокомпонентной табуированной и обценной лексики

№	Лексическая единица
1.	chucha
2.	idiota
3.	imbecil
4.	gilipollas
...	...

Т а б л и ц а 2

Фрагмент списка двухкомпонентной табуированной и обценной лексики

№	Лексическая единица
1.	diciendo pendejadas
2.	gringa comemierda
3.	lame huevos
4.	retraso mental
...	...

Т а б л и ц а 3

Фрагмент списка трехкомпонентной табуированной и обценной лексики

№	Лексическая единица
1.	egocéntrica de mierda
2.	hijo de puta
3.	infórm a cagar
4.	ojala te mueras
...	...

Т а б л и ц а 4

Фрагмент списка многокомпонентной табуированной и обценной лексики

№	Лексическая единица
1.	de la familia de tordo
2.	mal de la cabeza
3.	no me interesa tu opinión
4.	te pongan a comer mierda
...	...

Т а б л и ц а 5

Фрагмент списка слов, называющих понятия,
связанные с интимной жизнью человека

№	Лексическая единица
1.	chinga tu madre
2.	coge por el culo
3.	haz mamada
4.	haz una cubana
...	...

Также были выделены лексические единицы, которые в эмоционально-экспрессивном плане нейтральны, но обладают отрицательной, актуализируемой в ходе реализации высказывания коннотацией и, следовательно, высоким деструктивным потенциалом. Особенно часто для характеристики адресата во время спора используется пейоративно-оценочная лексика и зоосравнения. Зоосравнения и лексические единицы, которые называют понятия, связанные с сексуальной ориентацией или сексизмом, часто активизируют свой негативный потенциал, находясь в сочетании с соответствующей оценочной лексикой или формой 2-го лица единственного числа глагола *ser*. В то же время интернет-общение постоянно упрощается, например, путем употребления эллиптических предложений, что, в свою очередь, сокращает текст твита до одного слова, при этом усиливая его негативный смысл. Фрагменты списков подобных лексических единиц приведены в табл. 6–9.

Т а б л и ц а 6

Фрагмент списка лексических единиц, называющих животных

№	Лексическая единица
1.	burro
2.	cabrón (cabron)
3.	cerdo
4.	perra
...	...

Т а б л и ц а 7

Фрагмент списка лексических единиц, называющих понятия, связанные с сексуальной ориентацией и сексизмом

№	Лексическая единица
1.	feminista
2.	gay
3.	lesbiana
4.	machista
...	...

Т а б л и ц а 8

Фрагмент списка пейоративно-оценочной лексики

№	Лексическая единица
1.	hortera
2.	de mierda
3.	pendejo
4.	piojoso
...	...

Т а б л и ц а 9

Фрагмент списка форм глагола *ser* во 2-м лице единственном числе

№	Лексическая единица
1.	eres sos
2.	serás
3.	habrá sido
4.	serías
...	...

Использование английского языка является очень активным в интернет-коммуникации. Как и другие языки, этот язык богат лексическими единицами с негативной коннотацией. Употребление подобных единиц делает общение более

универсальным и в некоторых случаях позволяет усилить деструктивный эффект высказывания. В табл. 10 приведен фрагмент списка англицизмов, выявленных в ходе анализа испаноязычных твитов с маркерами флейма.

Т а б л и ц а 10

Фрагмент списка англицизмов с негативной коннотацией

№	Лексическая единица
1.	bitch
2.	burn in hell
3.	dickhead
4.	screw you
...	...

В ходе выполнения исследования в качестве экспликаторов конфликтной речевой ситуации были выделены идеографические знаки (эмодзи), указывающие на отрицательное невербальное поведение коммуникантов. В современной интернет-коммуникации эмодзи стали занимать особое место в процессе передачи эмоций и эмоциональных состояний. Эти графические единицы могут не только усилить деструктивный эффект высказывания, но и служить в качестве самостоятельного твита (например, эмодзи «средний палец»). В табл. 11 приведены эмодзи, которые являются однозначными маркерами флейма и выражают негативные эмоции, эмоциональное состояние и реакцию в отношении собеседника.

Т а б л и ц а 11

Список эмодзи, отражающих негативное эмоциональное состояние автора твита

№	Эмодзи
1.	
2.	
3.	
4.	

Помимо вербальных и невербальных единиц, которые явно указывают на наличие деструктивной стратегии «флейм», в текстах испаноязычных твитов были выделены те маркеры, которые проявляют свой деструктивный потенциал в комплексном употреблении. Под комплексным употреблением в работе понимается одновременное употребление как минимум трех подобных индикаторов. Второстепенные маркеры реализации флейма представлены в табл. 12, 13.

Список личных и вопросительных местоимений

№	Лексическая единица
1.	me
2.	qué (que)
3.	te
4.	tú (tu)
5.	yo

Личные и вопросительные местоимения позволяют персонализировать и/или окрасить сообщение в эмоционально-экспрессивном плане, что, в свою очередь, сигнализирует о появлении флейма.

Фрагмент списка слов-интенсификаторов

№	Лексическая единица
1.	absolutamente
2.	extremadamente
3.	imposiblemente
4.	increíblemente
...	...

Чаще всего, в роли интенсификаторов выступают наречия, прилагательные или написание всего сообщения либо его части прописными буквами. Такие средства позволяют пользователям наглядно подчеркнуть важность сообщения, его части или одного слова. Например: *tu vieja @UnaLindaDamita@unacomadreja LO SE POR ESO TE DIGO*. Пользователи Твиттера употребляют все виды интенсификаторов для увеличения негативного эффекта высказывания, выделения своей точки зрения, выражения своего эмоционального состояния и усиления приказного тона.

Среди пунктуационных знаков, выступающих в роли второстепенных маркеров флейма, можно отметить присутствие в тексте твита более одного вопросительного и/или восклицательного знака. Кроме того, из-за упрощения общения в Интернете пользователи могут, как случайно, так и целенаправленно (повышая уровень агрессии), сделать опечатку и вместо «!» использовать «1».

Выделенные вербальные и невербальные маркеры флейма в полном объеме составили лингвистическую базу данных системы автоматического определения средств реализации коммуникативной стратегии «флейм» в испаноязычных твитах.

The article deals with the verbal and non-verbal markers of flame commonly used as one of the virtual destructive strategies in twits. Ten classes of the main Spanish lexical and grammatical units, the emoji and two classes of the secondary units are under consideration as linguistic database elements of the system of automatic detection of means of flame strategy realization in virtual communication.