

## ЛИТЕРАТУРА

1. Харрис, Р. Дж. Психология массовых коммуникаций / Р. Дж. Харрис ; под ред. Д. Гиппиус. – 4-е междунар. изд. – М. : Олма-Пресс, 2001. – 448 с.

The article deals with specific linguistic features of the news about road accidents. Morphological, syntactical, lexical and structural characteristics of such texts are under consideration.

**А. В. Скрябина**

Минск, МГЛУ

### ПОИСКОВЫЙ ОБРАЗ АНГЛОЯЗЫЧНЫХ НОВОСТНЫХ ТЕКСТОВ

Рассматривается создание поискового образа англоязычных публицистических текстов на основе выделения ключевых или опорных слов.

В области анализа публицистических текстов ведутся разносторонние исследования, так как в текстах данной предметной области затрагиваются вопросы весьма широкой тематики – любые актуальные вопросы современности, представляющие интерес для общества: политические, экономические, моральные, философские, вопросы культуры, воспитания, повседневного быта.

На данный момент существует огромное количество публицистических текстов информационных и аналитических жанров, чтобы ознакомиться с их содержанием и смыслом, человеку требуется немало времени и усилий. В связи с этим появилась необходимость автоматической обработки и анализа данных текстов. Системы анализа текста решают сегодня целый ряд задач, такие как формирование информационного портрета текста в терминах ключевых понятий, выявление смысловых связей между понятиями, автоматическое реферирование. Таким образом, создание систем, автоматизирующих обработку текста с целью выявления его основного содержания и структуры, является актуальной проблемой в наши дни.

В данной работе материалом исследования послужили англоязычные новостные публицистические тексты по теме «культура», взятые из интернет-ресурсов [www.bbc.com](http://www.bbc.com) и [www.theguardian.com](http://www.theguardian.com). Для определения основного содержания каждого из текстов мы использовали статистический метод. В основе этого метода лежит критерий выделения ключевых или опорных слов, которые пронизывают текст и влияют на его понимание. Затем полученные списки потенциальных опорных слов были дополнены словами схожей тематики. Последним этапом составления словаря опорных слов по тематикам стало указание словоформ для слов, входящих в базу данных.

Таким образом, в созданной базе данных поисковый образ, например, для тематики «Film» выглядит следующим образом (табл.):

## Поисковый образ для тематики «Film»

act	acts	acting	acted
actor	actors		
contractor	contractors		
director	directors		
drama			
film	films		
finance	financial	financed	
movie	movies		
play	plays	playing	played
star	stars	starring	starred
take	takes	took	taken

На основании разработанных поисковых образов по каждой тематике был составлен алгоритм автоматической идентификации тематики англоязычных публицистических текстов.

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. О. Третьяк**  
Минск, МГЛУ

### АВТОМАТИЧЕСКОЕ ОБЪЕДИНЕНИЕ ФРАНКОЯЗЫЧНЫХ НОВОСТНЫХ СООБЩЕНИЙ В ЕДИНЫЙ СЮЖЕТ

Формальная процедура автоматического объединения текстов новостей в единый сюжет базируется на решении двух взаимосвязанных задач: определении тематики текстов новостных сообщений на основе дескрипторного метода и автоматическом построении их информационных портретов с опорой на статистический метод. Считается, что несколько новостных сообщений относятся к определенному сюжету при условии сходства их информационных портретов. Последние формируются на основе главных опорных слов каждого текста и формальных синтаксических групп (субъекта, предиката, объекта, места действия, времени действия), выделенных из ключевых предложений новостных сообщений.

Для решения целого ряда аналитических задач возникает потребность оценить содержание текста, получить его информационный портрет, то есть статистически значимую совокупность информационных характеристик. В большинстве существующих систем такой портрет состоит из статистически значимых слов и выражений, сопровождающих упоминание конкретного объекта. Построение информационных портретов позволяет автоматически группировать тексты по сюжетам. В общем плане под сюжетом понимается последовательное