

УДК 81–114.2

UDC 81–114.2

10.00.00 Филологические науки

Philology

**ИЕРАРХИЯ КЛЮЧЕВЫХ СЛОВ В СТРУКТУРЕ ПОИСКОВОГО ЗАПРОСА****HIERARCHY OF KEYWORDS IN A SEARCH QUERY**

Потехина Марина Сергеевна  
аспирантка  
SPIN-код: 5778-4170  
*Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, Россия*  
[marina.s.sh@bk.ru](mailto:marina.s.sh@bk.ru)

Potehina Marina Sergeevna  
Postgraduate student  
RSCI SPIN-code: 5778-4170  
*Saratov State University named after N.G. Chernyshevsky, Saratov, Russia*  
[marina.s.sh@bk.ru](mailto:marina.s.sh@bk.ru)

Данная статья написана в русле современных лингвопрагматических исследований и посвящена анализу ключевых слов с точки зрения степени их семантической значимости для репрезентации прагматической установки продуцента текста. Исследование проведено на материале текстов поисковых запросов. Выбор данной разновидности интернет-коммуникации обусловлен, в первую очередь, тем, что поисковый запрос представляет собой текст, полностью состоящий из ключевых слов. В качестве основных методов исследования использованы экспериментальный метод и методы семантического и коммуникативного анализа текста. В статье поставлен вопрос об иерархии ключевых слов в структуре поискового запроса, предложены критерии классификации ключевых слов, разработана методика анализа поискового запроса в коммуникативном аспекте. Подчеркнуто, что членение поискового запроса на более и менее значимые компоненты обусловлено коммуникативным намерением и прагматической установкой продуцента запроса. Особое внимание уделено сложностям, возникающим при анализе запросов, в которых интенция продуцента выражена недостаточно четко. Результаты исследования важны для понимания лингвистической сущности ключевого слова, принципов формирования поисковых запросов и могут быть использованы в ходе анализа ключевых слов на материале разных типов текстов

This article was written in the mainstream of modern Lingua-pragmatic research and is dedicated to the analysis of keywords from the point of view of the degree of their semantic significance for representation of intensions of the producer text. The study was made on a material of texts of a search queries. The choice of this type of Internet communication is due, first of all, to the fact that the search query is a text consisting entirely of keywords. The main methods studies, were continuous sampling of the material, linguistic experiment and semantic and communicative analysis. The article raises the question of the hierarchy of keywords in the structure of the search query. The author suggests criteria for classifying keywords, develops a technique for analyzing the search query in a communicative aspect. The division of the search query into more and less significant components is due to the communicative and pragmatic intention of the author of the text. Particular attention is paid to the difficulties arising in the analysis of requests in which the producer's intention is not clearly expressed. The results of the research are important for understanding the linguistic essence of the keyword, the principles of the formation of search queries and can be used in the analysis of keywords in different types of texts

Ключевые слова: КЛЮЧЕВЫЕ СЛОВА, ПОИСКОВЫЙ ЗАПРОС, СЕМАНТИКА, ПОРЯДОК СЛОВ, ПРОДУЦЕНТ ЗАПРОСА

Keywords: KEYWORDS, SEARCH QUERY, SEMANTIC, ORDER OF WORDS, SEARCH QUERY PRODUCER

Doi: 10.21515/1990-4665-128-066

Изучение ключевых слов как наиболее значимых для интерпретации информации лексических единиц [5] – актуальная сфера современной лингвистики, поскольку данное понятие активно используется не только в филологической теории, но и в ежедневной речевой практике основной

части современного общества. В отечественной лингвистике ключевые слова анализируются на материале отдельных текстов [1], текстов конкретного автора [4], текстов определенного исторического периода [6, 11], выбранной сферы [9] и т.д. Особый интерес для исследования представляет функционирование ключевых слов в различных типах текстов современного интернет-дискурса. Материалом для данной статьи послужили тексты поисковых запросов системы Яндекс, по данным wordstat.yandex.ru.

Структурные особенности поискового запроса как сложного вербального комплекса, полностью состоящего из ключевых слов [7], и возможность его членения на отдельные значимые компоненты позволяет поставить вопрос об иерархии ключевых слов в пределах поискового запроса. Под иерархией ключевых слов в данном случае понимается степень семантической значимости каждого отдельного ключевого слова для репрезентации содержащейся в запросе информации. Возможно ли выделение наиболее значимых компонентов (условно назовем их детерминирующими ключевыми словами или ДКС) в структуре поискового запроса и на каком основании? Ответ на этот вопрос важен для понимания принципов построения поискового запроса, выбора ключевых слов, а также сущности ключевого слова как особой позиции лексической единицы, позволяющей репрезентовать коммуникативное намерение и прагматическую установку продуцента запроса.

Предположим, что выделение ДКС в структуре поискового запроса возможно, а основанием для такого выделения служат следующие критерии: большая (по сравнению с другими ключевыми словами) семантическая значимость для продуцента запроса и порядок слов.

Использование однокомпонентных поисковых запросов приводит к тому, что прагматическая установка продуцента трактуется расширенно. На практике об этом свидетельствует большое количество результатов выдачи, т.е. сайтов, содержащих указанное ключевое слово. При этом поиск

информации, действительно актуальной для продуцента запроса, осложняется. Поэтому, как правило, используются двух- или многокомпонентные запросы, что позволяет конкретизировать/уточнить прагматическую установку и сузить сферу поиска. Следовательно, логично предположить, что некоторые компоненты запроса более актуальны для продуцента, а остальные являются менее значимыми или дополнительными. Наиболее значимые ключевые слова или ДКС представляют собой семантическое ядро поискового запроса. ДКС можно также определить как тему поискового запроса, поскольку они позволяют ответить на вопрос: о чем данный поисковый запрос. Следующие по значимости компоненты конкретизируют аспект ДКС, актуальный для продуцента запроса (условно определяются как переходная зона между ядром и периферией запроса). Наименее значимые компоненты содержат уточнение/дополнение содержания запроса (периферийная зона).

Представляя поисковый запрос в виде сложной иерархической структуры, необходимо отметить такой признак ключевого слова, как текстогенность (выделен Т.В. Шмелевой [10]). В данном случае под текстогенностью понимается способность ключевого слова порождать новые поисковые запросы. Поскольку все компоненты поискового запроса являются ключевыми словами, текстогенностью обладает каждый из них. Однако важно выяснить, какое именно слово породило данный конкретный поисковый запрос, т.е. вокруг какого слова он строится, поскольку это слово обладает наибольшей степенью семантической значимости. Для ответа на этот вопрос необходим анализ прагматической установки, конечной цели поиска и образа искомой информации.

Следует разграничить реальную потребность продуцента, обусловившую создание поискового запроса, и его прагматическую установку, репрезентованную данным запросом. Если реальная потребность связана с каким-либо фактом действительности, то прагматической установкой яв-

ляется поиск информации о данном факте. Таким образом, прагматической установкой, выраженной запросом *купить квартиру*, является не покупка квартиры, а поиск информации о покупке квартиры. Конечной целью поиска является получение искомой информации, а не собственно покупка. Объект поиска, соответственно, не квартира, а источник информации, сайт, на котором представлены объявления, связанные с продажей квартир. А образом искомой информации является объявление о продаже квартиры, а не сама квартира. Подобное разграничение важно для членения поискового запроса на более и менее семантически значимые ключевые слова.

Проанализируем следующий поисковый запрос:

*роман идиот скачать бесплатно в формате txt*

В данном примере выделяются следующие ключевые слова: *роман идиот, скачать, бесплатно, в формате txt*. Прагматической установкой, репрезентованной данным поисковым запросом, является поиск источника (сайта), на котором представлен текст романа Ф.М. Достоевского «Идиот» с возможностью его скачать бесплатно в определенном формате. Цель определяется как получение текста романа с помощью скачивания на конкретных условиях: бесплатно и в формате txt. Образом искомой информации является именно текст романа, затем процесс скачивания, ключевые компоненты *бесплатно* и *в формате txt* являются условиями, на которых будет осуществляться скачивание романа. Таким образом, структура приведенного поискового запроса с точки зрения семантической значимости составляющих его компонентов может быть представлена следующим образом:

1) ядро (наиболее значимый семантический компонент или ДКС):  
*роман идиот*;

2) переходная зона (менее значимый семантический компонент):  
*скачать*;

3) периферийная зона (наименее значимые семантические компоненты): *бесплатно* и *в формате txt*.

С этой точки зрения, легче всего выделяются наименее значимые (периферийные) компоненты. Например, в запросах

- 1) *скачать полное собрание сочинений пушкина в формате txt*;
- 2) *купить квартиру без посредников*;
- 3) *фильмы 2016 смотреть онлайн бесплатно*;
- 4) *танки играть онлайн без регистрации*;

компоненты *в формате txt*; *без посредников*; *2016*, *онлайн*, *бесплатно*; *онлайн*, *без регистрации* менее значимы, поскольку не репрезентуют прагматическую установку, конечную цель поиска, его объект и образ искомой информации, а уточняют и дополняют их. Кроме того, в приведенных примерах указанные ключевые слова не определяются как тема поискового запроса и не являются текстогенными, т.е. не они порождают данные поисковые запросы. Следовательно, слова *в формате txt*; *без посредников*; *2016*, *онлайн*, *бесплатно*; *онлайн*, *без регистрации* относятся к периферийной зоне поискового запроса.

Сложнее определить границы между семантическим ядром поискового запроса и переходной зоной. В анализируемых запросах компоненты *скачать*, *полное собрание сочинений пушкина*; *купить*, *квартиру*; *фильмы*, *смотреть*; *танки*, *играть* более значимы, по сравнению с периферийными, для репрезентации прагматической установки продуцента запроса, конечной цели поиска, объекта и образа искомой информации. Каждый из них может быть определен как текстогенный, т.е. тот, вокруг которого строится поисковый запрос. При этом остается открытым вопрос о том, какие из перечисленных компонентов являются семантическим ядром запроса, а какие – его конкретизацией.

Если попытаться выделить ДКС среди более значимых компонентов запроса на основании степени семантической конкретности, то в запросах

*купить квартиру без посредников* и *фильмы 2016 смотреть онлайн бесплатно* ДКС будут являться слова *купить* и *смотреть*, поскольку соотносятся с более конкретными фактами действительности: покупка и просмотр. В то время как компоненты *квартиру* и *фильмы* соотносятся с большим количеством фактов действительности, поскольку производитель запроса подразумевает под ними множество разных квартир и разных фильмов. В запросах *скачать полное собрание сочинений пушкина в формате txt* и *танки играть онлайн без регистрации* компоненты *скачать, полное собрание сочинений пушкина* и *танки, играть* обладают одинаковой степенью семантической конкретности, поскольку каждый из них соотносится в сознании производителя запроса с одним фактом действительности.

Таким образом, семантический критерий оказывается недостаточным для выделения ДКС в структуре поискового запроса. Необходимо также учитывать более формальный критерий, которым является порядок слов в поисковом запросе.

Под порядком слов в поисковом запросе понимается расположение компонентов запроса по отношению друг к другу.

О роли порядка слов в организации словосочетания, предложения, высказывания писали много и часто [2, 3, 8]. О.Б. Сиротина рассматривает различия функций порядка слов в письменной и разговорной речи. В письменной речи порядок слов выполняет следующие функции:

- 1) коммуникативную (выражение актуализации высказывания, степени коммуникативной значимости слова);
- 2) грамматическую (выражение синтаксической организации словосочетания, предикативной конструкции, высказывания);
- 3) стилистическую (выражение экспрессивности и стилевой окраски высказывания) [8].

В разговорной речи при помощи порядка слов выражается, в первую очередь, степень коммуникативной значимости компонента высказывания: наиболее значимые слова располагаются в начале высказывания.

Как и в разговорной речи, порядок слов в поисковом запросе является отражением речемыслительного акта, процесса формирования высказывания, соответствующего последовательности возникающих в сознании продуцента запроса понятий. И.И. Ковтунова для обозначения указанной особенности порядка слов использует термин «принцип ассоциативного присоединения» [3]. Следовательно, поисковый запрос можно представить в виде линейной структуры, отражающей расположение ключевых слов в порядке убывания семантической значимости: в начале – ДКС, затем – ключевые слова переходной зоны, в конце – ключевые слова периферийной зоны.

Сильной позицией, свидетельствующей о наибольшей коммуникативной значимости компонента, является начало поискового запроса.

Таким образом, в поисковых запросах

- 1) *фильмы смотреть онлайн бесплатно в хорошем качестве;*
- 2) *подарок своими руками на 23 февраля;*
- 3) *танки играть онлайн без регистрации;*

слова *фильмы, подарок, танки* являются ДКС, что подтверждается их расположением в сильной позиции — начале поискового запроса.

Однако порядок слов в поисковом запросе может определяться не только коммуникативным, но и грамматическим аспектом. В этом случае его функцией будет выражение синтаксической организации высказывания. Например, в запросах

- 1) *когда масленица в 2017 году;*
- 2) *как сделать скриншот экрана на компьютере;*
- 3) *где можно посмотреть фильм викинг;*

препозитивное расположение союзов *когда, как, где* не свидетельствует о более высокой степени их семантической значимости, а отражает синтаксическую организацию предикативной конструкции. Порядок слов в приведенных примерах является прямым (нейтральным) и выполняет грамматическую, а не коммуникативную функцию.

Прямой порядок слов отмечается также в запросах:

- 1) *купить квартиру;*
- 2) *сделать яндекс стартовой страницей;*
- 3) *сдать это в саратове;*
- 4) *смотреть фильмы онлайн.*

Расположение прямого дополнения в позиции после сказуемого в русском языке является нейтральным и не свидетельствует о степени семантической значимости компонентов.

Как и в живой разговорной речи, в поисковом запросе структура словосочетания, предикативной конструкции, высказывания может разрушаться. При этом порядок слов становится коммуникативно значимым. Например, в поисковых запросах

- 1) *квартиры купить;*
- 2) *фильмы смотреть;*
- 3) *это сдать;*

нарушение прямого порядка слов (постпозиция прямого дополнения) свидетельствует о повышении степени семантической значимости компонентов *квартиры, фильмы, это*.

В запросах

- 1) *фильмы советские;*
- 2) *романы исторические;*
- 3) *цветы искусственные;*

постпозиция согласованного определения отражает непрямой порядок слов, в разговорной речи это свидетельствует об актуализации определе-



ния как более значимого компонента. В поисковом запросе наблюдается обратная ситуация: семантически более значимым оказывается первый компонент запроса. Следовательно, ДКС в приведенных примерах являются слова *фильмы, романы, цветы*. Компоненты *советские, исторические, искусственные* оказываются менее значимыми.

Приведем также примеры разрушения структуры предикативной конструкции, позволяющего подчеркнуть актуальность одного из ключевых слов:

- 1) *восстание декабристов в каком году;*
- 2) *кормушка как сделать;*
- 3) *это саратов где сдать.*

Следовательно, главным критерием для выделения ДКС в структуре поискового запроса при непрямом порядке слов является его расположение в начале запроса. Прямой порядок слов в коммуникативном аспекте оказывается нейтральным.

Таким образом, для членения поискового запроса на семантическое ядро (или ДКС) и менее значимые компоненты необходимо учитывать как семантический (степень семантической значимости), так и формальный (порядок слов) аспекты.

### **Литература**

1. Аверьянов Л.Я. Контент-анализ [Текст]: учеб. пособие / Л. Я. Аверьянов. М., 2009. С. 456.
2. Золотова Г.А. Труды В.В. Виноградова и проблемы текста // Вестник Моск. ун-та. Сер. 9. Филология, 1995. № 4. С.84-98.
3. Ковтунова И.И. Современный русский язык [Текст]: Порядок слов и актуальное членение предложения / И. И. Ковтунова. – М.: Просвещение, 1976. – 239 с.
4. Кухаренко В. А. Интерпретация текста [Текст] / В. А. Кухаренко. – М.: Просвещение, 1988. – С. 192.
5. Матвеева Т. В. Ключевые слова [Текст] / Т. В. Матвеева // Матвеева Т.В. Учебный словарь : русский язык, культура речи, стилистика, риторика. – М., 2003. – С. 103.
6. Романенко А.П. Советская словесная культура: образ ратора [Руко-пись]: дис. ...докт. филол. наук / А. П. Романенко. Саратов, 2001. С. 355.
7. Сахарный Л.В. Тексты-примитивы и закономерности их порождения //

Человеческий фактор в языке: язык и порождение речи. – М.: Наука, 1991. – С. 221-237.

8. Сиротинина О. Б. Порядок слов // Сиротинина О. Б. Лекции по синтаксису русского языка. М., 2003. С. 124-132.

9. Фомина З.Е. Слова-хронофакты в языке политических текстов [Текст] / З. Е. Фомина // Язык и эмоции : сб. науч. тр. Волгоград, 1995. С.207–324.

10. Шмелева Т. В. Ключевые слова текущего момента [Текст] / Т. В. Шмелева // Collegium. – 1993. – № 1. – С. 33–41.

11. Яшин В.Н. Архетипические ключевые слова отечественной политической речи (советский и постсоветский периоды) / В.Н. Яшин. Saarbrucken, 2011. С. 257.

### References

1. Aver'janov L.Ja. Kontent-analiz [Tekst]: ucheb. posobie / L. Ja. Aver'janov. M., 2009. S. 456.

2. Zolotova G.A. Trudy V.V. Vinogradova i problemy teksta // Vestnik Mosk. un-ta. Ser. 9. Filologija, 1995. № 4. S.84-98.

3. Kovtunova I.I. Sovremennyj russkij jazyk [Tekst]: Porjadok slov i aktual'noe chlenenie predlozhenija / I. I. Kovtunova. – М.: Prosveshhenie, 1976. – 239 s.

4. Kuharenko V. A. Interpretacija teksta [Tekst] / V. A. Kuharenko. – М.: Prosveshhenie, 1988. – S. 192.

5. Matveeva T. V. Kljuchevye slova [Tekst] / T. V. Matveeva // Matveeva T.V. Uchebnyj slovar' : russkij jazyk, kul'tura rechi, stilistika, ritorika. – М., 2003. – S. 103.

6. Romanenko A.P. Sovetskaja slovesnaja kul'tura: obraz ritora [Ruko-pis']: dis. ...dokt. filol. nauk / A. P. Romanenko. Saratov, 2001. S. 355.

7. Saharnyj L.V. Teksty-primitivy i zakonomernosti ih porozhdenija // Chelovecheskij faktor v jazyke: jazyk i porozhdenie rechi. – М.: Nauka, 1991. – S. 221-237.

8. Sirotinina O. B. Porjadok slov // Sirotinina O. B. Lekcii po sintaksisu russkogo jazyka. М., 2003. S. 124-132.

9. Fomina Z.E. Slova-hronofakty v jazyke politicheskikh tekstov [Tekst] / Z. E. Fomina // Jazyk i jemocii : sb. nauch. tr. Volgograd, 1995. S.207–324.

10. Shmeleva T. V. Kljuchevye slova tekushhego momenta [Tekst] / T. V. Shmeleva // Collegium. – 1993. – № 1. – S. 33–41.

11. Jashin V.N. Arhetipicheskie kljuchevye slova otechestvennoj politicheskoy rechi (sovetskij i postsovetskij periody) / V.N. Jashin. Saarbrucken, 2011. S. 25.