

*Е. Г. Соколова, И. С. Кононенко*  
*E. G. Sokolova, I. S. Kononenko*

**КОРПУС ТЕКСТОВ КАК ОСНОВА МЕТОДОВ  
МНОГОФАКТОРНОГО АНАЛИЗА  
(ГЛАГОЛ «РАЗЛИЧИТЬ – РАЗЛИЧАТЬ»)<sup>1</sup>**

**TEXT CORPUS AS A BASE FOR  
MULTI-FACTOR ANALYSIS  
(VERB «RAZLICHT' – RAZLICHAТ'»)**

**Аннотация.** В статье изложены результаты корпусного исследования семантики и употребления русской глагольной пары *различить-различать* с помощью метода дискурсивных контекстов. Под дискурсивным контекстом понимается сочетание в дискурсе текста сущностных и грамматических элементов, которое может быть представлено конкретным глаголом.

**Abstract.** The method of Discourse Contexts introduced earlier is applied to describe the semantics and use of Russian verb pair *razlichit' – razlichat'* (distinguish). Discourse contexts are presented as combinations of essential elements (idea of discrimination and entities involved such as subject, entities to be compared, and features of discrimination) with grammar constituents. The analysis of the material of Russian National Corpus gives six types of discourse contexts described and exemplified in the paper.

## **1. Введение**

В статье приведены результаты корпусного исследования лексической семантики глагола *различить(-ать)* как продолжение разработки метода Дискурсивных Контекстов (ДК)<sup>2</sup>. Особенность метода ДК состоит в опоре на сущности, участвующие в

---

<sup>1</sup> Работа выполнена при финансовой поддержке Президиума СО РАН (Интеграционный проект № 15/10).

<sup>2</sup> *Соколова Е.Г., Кононенко И.С.* Какие ситуации обозначаются русскими глаголами «отличить – отличать» // Труды Международной конференции Диалог 2013 (в печати).

обозначаемой ситуации, а не в описании «фрейма», представляющего «значение» глагола. Для глагола *различить(-ать)* в толковых словарях<sup>3</sup> выделяются значения, которые а) не увязаны между собой (общее) и в то же время б) не учитывают специфику ситуаций, лежащих за выделенными значениями (различия).

Метод ДК, основанный на ручной обработке корпусных данных, решает эти проблемы. ДК – это сочетание сущностных и грамматических элементов, образующее определенный контекст, в котором может быть использован тот или другой конкретный глагол. По методике ДК, конкретный токен глагола в тексте, в нашем случае, *различить(-ать)*, ассоциирован с Имманентной Ситуацией – ИС[размежевание], которая задает структуру ситуации, Сущностной Составляющей СС[различить(-ать)], которая задает общее значение глагола, Грамматической Составляющей ГС[различить(-ать)], которая описывает грамматические позиции участвующих сущностей и, наконец, ДКі[различить(-ать)] – часто используемые контексты.

Мы рассматриваем современное употребление глагола только в активной форме. Все примеры взяты из НКРЯ.

## **2. Факторы ДК для глагола «различить – различать»**

В следующем разделе рассматривается только СС[различить(-ать)], ИС[размежевание] определена ранее (см. сноску 2).

### **2.1. Сущностная ситуация**

В результате ассоциации абстрактных элементов ИС[размежевание] с конкретными типами сущностей, участвующих в контекстах глагола формируется СС[различить(-ать)], для которой элементы ИС определим следующим образом:

– I – «ментальная диверсификация», суть которой состоит в ориентации в исходном множестве / пространстве М путем распознавания в нем сущности(ей) Е по признаку F;

---

<sup>3</sup> Малый академический словарь русского языка (МАС) <http://mas-dict.narod.ru/>; словарь Дмитриева <http://enc-dic.com/dmytriev/>.

– М – исходное множество / пространство, которое представлено как а) диффузное пространство (однородный фон), например, «картина поля» в примере (3), или б) дискретное множество М, состоящее из n однородных сущностей, например, «звезды» в (2); или М – имплицитное множество  $M_i$ , образуемое вновь введенными в дискурс  $i$  понятиями, источником которых является субъект S;

– S – субъект явный (1), (5) и др. или имплицитный (во втором случае в предложении может иметься ИСТОЧНИК в виде документа или раздела науки, представляющий мнение / знание интеллектуальной среды S (23);

– E – узнанные объекты в исходном М или  $E_i$  – введенные объекты, образующие имплицитное множество  $M_i$ .  $E_i$  диверсифицируются именами при актуализации, например, «вера» и «церковь» в (22) и свойствами  $F_i$ , выделенными, как правило, по «экспертному» параметру, см.  $F_1$  и  $F_2$  в (22). Сущности E не противопоставляются друг другу или нераспознанной части М, например, в (2) это все элементы М – звезды – с установленными характеристиками F – «внешнего вида и возраста».

– F – признак (в случае  $E_i$ , см. (22)) или параметр (см. (2) и (3)) диверсификации.

Элементы ИС и СС обнаруживаются в контекстах из НКРЯ:

- (1) *Иными словами, кошки(S) тонко различают звуки(E=M). Так же хорошо они различают запахи(E=M), особенно представляющие для них интерес.*
- (2) *Астрономам(S), научившимся различать звезды(E=M) по внешнему виду и возрасту(F), стало легче разбираться в том, как протекает жизнь звезд от рождения к смерти*
- (3) *Внешние объекты искажают картину поля(M), и рыба чувствует искажения(F), что позволяет ей(S) обнаруживать и различать объекты(E), то есть осуществлять локационную функцию.*

## 2.2. Грамматическая составляющая

ГС – это инвентарь грамматических позиций, используемых данным глаголом в активной форме. Включает позиции, реализующие элементы СС и дополнительные позиции, описывающие стандартные обстоятельства процесса. ГС[различить(-ать)] в соответствии с корпусными данными, имеет следующие позиции:

А. прямое дополнение. Реализует Е, и только Е, представленную в виде:

– «Х-Асс», см. (1) – (3);

– пропозиционального актанта в форме дополнительного придаточного предложения:

(4) *Скоро бес(S) различил, что рядом с ним летит не кто иной, как его охранник и мучитель – двойник шамана(E).*

или косвенного вопроса:

(5) *Обычно Роберт(S) различал по этой стукотне(F), кто из друзей взял за работу(E).*

Наличие сентенциальных дополнений позволяет характеризовать глагол *различить(-ать)* как ментальный типа «*знать*»<sup>4</sup>.

В (1) и (2) прямое дополнение, представленное существительным во множественном числе, обозначает одновременно Е и М.

Б. «по X-Dat» – всегда реализует F, см. (2). Но обратное неверно.

В. «X-Instr» – Ожидаемая реализация в этой позиции сущности СРЕДСТВО как инструмента не подтверждается корпусными данными. Вместо этого реализуется сущность ОРГАН – неотъемлемая физическая или ментальная часть S (*ухо, мышление, глаз, пальцы, профессиональный взгляд*), которая является результатом расщепления S:

---

<sup>4</sup> *Падучева Е.В.* Прямая и косвенная диатеза ментального глагола: корпусное исследование // Труды международной конференции «Корпусная лингвистика – 2008». – СПб, 2008. С. 303–317.

- (6) *Ведуны(S) могли третьим глазом(ОРГАН) различать его облик(Е).*
- (7) *Тогда маленькая Милена(S) все время путала луковицу с головкой чеснока, не различая маленькими пальчиками(ОРГАН) чесночных ребрышек(Е)...*

Г. «с помощью X-Gen», «используя X-Acc» – реализует СРЕДСТВО или МЕТОД:

- (8) *Его(Е) можно было различить лишь с помощью лупы(СРЕДСТВО).*
- (9) *Биологи(S) научились различать легальную(Е1) и браконьерскую икру(Е2), используя генетические маркеры(МЕТОД).*

Д. «в/на W-Loc», «среди W-Gen» – регулярно реализует М в современном языке с помощью общелокативных предлогов «в» и «на», а также предлога «среди», «между»:

- (10) *Среди декоративных томатов(М) различают растения низкорослые( $E1 \cap F1$ ) и высокорослые( $E2 \cap F2$ ).*

Ж. «X-Nom» – подлежащее. Синтаксическая функция позволяющая реализовать и при этом коммуникативно выделить разные элементы СС:

– S – живые существа (1), (4), (5) и др., также интеллектуальные устройства:

- (11) *Тормоз-накат «инерционного» типа Knott(S) различает режим торможения(Е1) и движения задним ходом(Е2) .*

– ОРГАН как неотъемлемая часть S, при этом S может быть не выражен или выражен родительным падежом в составе той же позиции:

- (12) *Зоркий взгляд(ОРГАН) птицы(S) различает табуника(Е1), молодого мужчину в седле, с винтовкой за плечами.*

– ИСТОЧНИК представляет документ или область науки:

- (13) *Фитохимия(ИСТОЧНИК) уже различает в этой белковой массе(М) два тела(Е1-2): одно, – растворимое в теплом (вернее в горячем) спирте, другое, – нерастворимое в нем.*

- F, МЕТОД или СРЕДСТВО встречаются в позиции подлежащего модального глагола (*позволяет*) (16) или в составе

выделительной конструкции (14) или является антецедентом относительного местоимения (15), хозяином причастного оборота, но примера непосредственно в позиции подлежащего глагола *различать* нам не встретилось, хотя интроспективно это кажется возможным.

(14) *Пожалуй, единственное, что различает корабли(Е), – это их назначение(Ф).*

(15) *Для каждого вида мы на удивление легко нашли тот самый ключевой признак(Ф), который позволял легко различать спектрограммы звуков самцов и самок(Е).*

(16) *Это свидетельствует о неадекватности стандартов и в то же время показывает, что MAPREC-тест(МЕТОД) позволяет различать пробы(Е), различающиеся по содержанию 472-С-последовательностей всего на 0,2%.*

### **3. Дискурсивные контексты**

#### **3.1. Диверсификация в признаковом пространстве**

М – пространство признаков F, в котором S, биологическая сущность или интеллектуальное устройство, распознает совокупности свойств, образующие сущности E. I – выделение и диверсификация сущностей как интеллектуальный или жизненно важный процесс (а не просто как восприятие). Признаки F, по которым выделяется объект E, часто указываются вне актантажной структуры предиката – в другой клаузе (3) или в составе именной группы E:

(17) *«Вы(S) различаете с первого звука их голоса(Е) с их модуляциями, затверженным интонационным регистром(Ф)».*

**ДК1** – идентификация: E актуализована в сознании S в виде образа (и имени), и S знает (думает), что E присутствует в M (5). ДК1 имеет вариант:

**ДК1.1** – комбинаторная идентификация: S знает имена объектов E, данных ему в непосредственном восприятии, но не знает, какое имя соответствует какому объекту. В этом случае

процесс I сводится к отождествлению конкретного объекта с конкретным образом (именем), а E выражено косвенным вопросом:

(18) *«С этим у меня вообще проблема: когда я читаю чужие романы или смотрю «толстые» фильмы, то лишь к концу начинаю различать, кого из персонажей как зовут».*

Синонимы: *узнавать, отличать, опознать, распознать, идентифицировать.*

**ДК2** – ориентация: E не актуализована, т.е. S не обладает образом E, а имеет способность обнаружить в M сущности E в процессе I по типовым свойствам (3). Ориентация имеет вариант:

**ДК2.1** – выделение нюанса: M – это совокупность свойств F уже узнанного, выделенного объекта E, т.е. E=M как признаковое пространство. «Распознается» «дополнительная» E'(E) – некоторый индивидуальный (ожидаемый или сиюминутный) нюанс:

(19) *Голос его звучит мечтательно, но вместе с тем я(S) различаю и металлические нотки(E').*

Синоним: *улавливать.*

### **3.2. Диверсификация в множестве однородных сущностей**

M – множество абстрактных или предметных однородных сущностей. В зависимости от M и E выделяется несколько ДК.

**ДК3**: M – понятия.  $E_i = M$ . В толковом словаре Дмитриева значение глагола для этого ДК определено только для несовершенного вида<sup>5</sup> как «обладание способностью распознавать отличия, характеристики каких-либо объектов, явлений и т. п.». Ментальность в определении выражена как «обладание способностью». Отличия между  $E_i$ , элементами множества M, существенные, что подтверждается примером:

(20) *В 9 лет мальчик(S) едва различал буквы( $E_i=M$ ),*

---

<sup>5</sup> Значение, определенное только для одного, несов., вида, заставляет нас рассматривать *различить* и *различать* как два глагола, хотя более адекватным кажется объяснение, что они – точная видовая пара, но при совпадении в дискурсе множеств E и M возникает ДК, в котором глагол приобретает семантику отношения, которое, в свою очередь, исключает результативность или совершенный вид.

где речь идет о грамотности мальчика, а не его зрении. Примеры из словаря Дмитриева: *«различать породы собак, сорта вин, болезни».*

Синоним: *знать.*

**ДК4** – практическая диверсификация, М – эксплицитное.

**ДК4.1**: маркирование элементов E по актуализованному параметру F, индивидуализация. Значения параметра F индивидуальны для каждого  $E_i$ . F находится в фокусе внимания (2), выделяется синтаксической конструкцией (15).

Синоним: *отличать друг от друга.*

**ДК4.2**: разделение множества сущностей на группы в реально протекающем процессе I, практическая классификация. Актуализованы типы сущностей в группах  $E_i$ , F – не актуализован. Часто присутствует МЕТОД (9).

Синоним: *выделять, разделять.*

**ДК5** – теоретическая диверсификация. М – имплицитное. Это констатация различия с указанием конкретных значений некоторого параметра, очевидного или эксплицитного, например, в (10) – это «высота», в (21) – «форма»:

(21) *Различают(S) плоды смородиновидные( $E_1 \cap F_1$ ), сливовидные, грушевидные( $E_2 \cap F_2$ ), и другой формы( $E_i \cap F_i$ ).*

Если S вводит новые понятия или интерпретирует известные, параметр обычно не актуализован, а конкретные значения параметра диверсификации формулируются в смежной клаузе:

(22) *Я(S) все-таки очень сильно различаю веру( $E_1$ ) и церковь( $E_2$ ). и вера—она для меня вне морали( $F_1$ ), это ни хорошо, ни плохо; все зависит от того, как человек ее применяет. церковь же для меня безусловное зло( $F_2$ ).*

Часто классификации сопутствуют УСЛОВИЯ классификации. Они могут быть имплицитными (очевидными, стандартными) (10), (21) или эксплицитными, т.е. выражаться обстоятельствами, например, *«в соответствии с (документом)», «с точки зрения (аспект, параметр)», «при (макропроцесс)»:*

(23) При анализе качества программ(УСЛОВИЯ) целесообразно различать их дефекты и ошибки(E).



Констатация различия может исходить от конкретного S (22) или обобщенного лица (неопределенно-личная форма глагола), как в (21), но может исходить и от автора текста, который ссылается на ИСТОЧНИК. Последний в этом случае занимает освободившуюся позицию подлежащего глагола «различать»:

(24) *Закон 1641 г.(ИСТОЧНИК) различает в составе крестьянской крепости три исковые части(Е1-3): крестьянство(Е1), крестьянские животы(Е2) и крестьянское владение(Е3).*

У дериватов глагола с подавленным подлежащим ИСТОЧНИК выражается в позиции локатива: *у Иванова, в трудах Иванова.*

Синоним: *разделять(ся) (несов.!) на (22), разделять (23)*

### ***3.3. Диверсификация как результат работы распознающего устройства***

**ДК6.** М и F не выражены. S обладает распознающим ОРГАНОм, который либо назван (6), либо очевиден по S. Также может присутствовать СРЕДСТВО усиления разрешающей способности ОРГАНа (8) или преграда. Часто предложение включает оценочные слова (1) или условия наилучшего восприятия:

(25) *«Человек с хорошим зрением **различает** мелкие детали лучше всего в 25 см от глаз (это расстояние наилучшего зрения)».*

Синоним: *~распознавать* или конкретно – *видеть, слышать...*

## **4. Заключение**

Таким образом, выделено 6 дискурсивных контекстов, покрывающих значения глагола *различить(-ать)* в толковых словарях и основанных на единой системе сущностных элементов.

Только одно из двух значений, описанных в МАС, найдено методом ДК. Примеры на второе значение из А.С. Пушкина и И.С. Тургенева, по-видимому, просто устарели, их контексты в современном языке передаются глаголом «отличить» (строка 2

Табл. 1). Выделено 5 типов контекстов, в которых описаны актуальные «значения» глаголов «различить(-ать)».

Таблица 1. Сравнение значений толковых словарей и ДК для глагола *различить (-ать)*

	<b>различить</b>		<b>различать</b>
1	Распознать, обнаружить зрением или с помощью других внешних чувств;	ДК6	Распознать, обнаружить зрением или с помощью других внешних чувств;
2	Распознать какой-л. предмет, явление среди других; отличить.	–	Распознать какой-л. предмет, явление среди других; отличить.
3		ДК3	«обладание способностью распознавать отличия, характеристики каких-либо объектов, явлений и т. п.» (Дмитриев)
4	ДК1, ДК2, ДК4, ДК5		