

Санкт-Петербургский государственный университет

Филологический факультет

Кафедра математической лингвистики

Направление: «Лингвистика»

Образовательная программа: «Прикладная и экспериментальная  
лингвистика»

Профиль: «Компьютерная лингвистика и интеллектуальные технологии»

**Выявление информации о конфликтных ситуациях в новостных  
текстах**

**(на базе семантического описания RussNet)**

Выпускная квалификационная работа  
соискателя на степень магистра филологии

Лукина Никиты Сергеевича

Научный руководитель: к. ф. н., доц. Азарова

Ирина Владимировна

Санкт-Петербург

2016

## Оглавление

Введение.....	3
Глава 1. Теоретические аспекты выявления конфликтных ситуаций.....	7
1.1 Семантическое поле и лексико-семантическая группа слов.....	7
1.1.1 Семантическое поле.....	7
1.1.2 Лексико-семантические группы слов.....	11
1.2 Понятие валентности в лингвистике.....	14
1.2.1 Структурно-семантическое описание валентностей.....	14
1.2.2 Структурно-семантические типы валентностей.....	17
1.2.3 Практическое применение теории валентностей семантических ролей.....	22
1.3 Фреймовый способ описания семантики.....	24
1.3.1 Понятие фрейма.....	24
1.3.2 Фрейм как способ представления знаний.....	31
1.4 Электронные тезаурусы как способ описания семантических отношений.....	33
1.4.1 WordNet.....	33
1.4.2 RussNet – электронный тезаурус русского языка.....	35
1.5 Выводы.....	37
Глава 2. Состав и структура фреймов конфликтных микро- и макроситуаций.....	39
2.1 Исследуемый материал.....	39
2.2 Лексика и рамки валентностей.....	40
2.3 Фреймовое описание конфликтных ситуаций.....	43
2.3.1 Структура фрейма микроситуации.....	44
2.3.2 Структура фрейма макроситуации.....	47
2.4 Выводы.....	51
Заключение.....	54

Список использованной литературы.....	56
Приложение.....	60

## **Введение**

Представленная работа производится на базе описания Russnet – электронного тезауруса русского языка, перенявшего большую часть методологии построения лексико-семантической базы данных у WordNet-проектов, представляющими собой сеть, элементами которой являются синонимические ряды (синсеты) и слова, входящие в них. Аналогичный проект RussNet, начало создания которого приходится на 1999 год, поставил перед собой цель охватить всю базовую общеупотребительную лексику. Наполнение словаря ведется методом отбора и анализа лексико-семантических полей, в состав которых входят высокочастотные лексем, которые предположительно формируют ядерную часть лексико-семантических полей.

Целью данной работы является выявление структуры лексико-семантического поля глаголов, описывающих конфликтные ситуации. Для этого были поставлены следующие задачи:

1. Выявить основные параметры лексико-семантического поля (в том числе при помощи описания теории валентностей и теории фреймов)
2. Собрать корпус с последующим выделением наиболее частотных лексем, относящихся к конфликтной ситуации, а также создать фрейм с целью обеспечить наиболее полное исчисление лексем, соответствующих структуре поля.

Корпус, на основе которого было проведено представленное исследование, собран на материале Википедии. В него вошло около 30 обширных статей, посвященных описанию вооруженных действий в 2000-2015 годах, например, таких, как: Афгано-пакистанский пограничный конфликт, конфликт в Южной Осетии, конфликт на Украине, гражданская война в Ливии, Ливии и другие. Общий объем корпуса составляет около 60 тысяч словоупотреблений. Несмотря на свой объем корпус репрезентативен как с точки зрения представленности в нем конфликтных ситуаций (в нем подробно описаны все виды боевых действий и связанные с ними элементы

(место, время и т.д.), интересующие нас в данном исследовании), так и с точки зрения лексики.

При помощи Томита-парсера были отобраны наиболее часто встречающиеся глаголы, при этом частотными считались те глаголы, которые встретились в корпусе более 15 раз. Из них были отобраны глаголы, имеющие явное отношение с точки зрения семантики к военным действиям, а также подобные глагольные конструкции с глаголами «быть» и «стать». Полученные глаголы (20 единиц) были разделены на те, которые составляют «ядро» семантического поля и «периферию». Для глаголов, относящихся к ядру, были подстроены актантные модели и были отмечены особенности каждой из конструкций. Кроме того, эти глаголы были объединены в синсеты с указанием их рамок валентностей, значением слов и примерами употребления для дальнейшей интеграции в RussNet.

Основная мысль, заложенная во фреймовом описании конфликтной ситуации, заключается в существовании конфликтных макроситуаций и микроситуаций, которые можно описать при помощи фреймов.

В нашем случае конфликтной ситуацией считается ситуация, которая обладает тремя составляющими конфликта: биполярностью, наличием сторон конфликта, наличием действий.

Микроситуацией является мельчайшее действие, в нашем случае – конфликтного характера. Микроситуация опирается на текст, она вербоцентрична. Фрейм микроситуации состоит из слотов «действие», «субъект», «объект» и «место» с указанием их лексического выражения и грамматического значения.

Макроситуацией же считается сумма связанных между собой микроситуаций, описывающая конфликт в конкретном тексте. Как правило, при наличии нескольких микроситуаций связь между ними представляют собой логические операции (импликация или конъюнкция). Конфликтная ситуация также опирается как на текст, так и на контекст, поэтому структура фрейма конфликтной макроситуации более сложна. Она состоит из слотов

«Общая характеристика», в которой прописывается точка зрения на конфликт, аспекты и участники конфликтной ситуации, «Микроситуации», слот, ссылающийся на фреймы микроситуаций, а также слот «Характеристика действия», в которой указывается связь между микроситуациями, характеристика локализации конфликта и особенности времени.

Таким образом, представленная работа состоит из трех частей:

1. В первой главе описаны теоретические аспекты выявления конфликтных ситуаций, а именно:

- описание теории семантического поля и лексико-семантических групп
- описание теории валентностей и семантических ролей
- описание теории фреймов
- описание структур WordNet-тезаурусов и RussNet'а в частности

2. Во второй главе произведено исследование состава и структуры фреймов конфликтных ситуаций, а именно:

- описан материал, на основе которого проведено представленное исследование
- произведено исследование лексики и рамок валентностей
- спроектированы фреймы конфликтных микро- и макроситуаций

3. В приложении представлены синсеты, включающие в себя глаголы, относящиеся к «ядру» исследуемого семантического поля, описанные в терминах RussNet, а также приведены примеры фреймов конфликтных микроситуаций с каждым из этих глаголов с примерами.

Методы, используемые в работе:

- Теория семантических полей (с целью описать лексико-семантического поле конфликтной ситуации)
- Теория валентностей и семантических ролей (с целью определить основные параметры конфликтной ситуации с их грамматическим значением).
- Теория фреймов (с целью построения фрейма конфликтной ситуации)

Новизна представленной работы заключается в построении фрейма, описывающего конфликтную ситуацию, так как подобных исследований в рамках методики RussNet еще не было.

Актуальность же исследования заключается в том, что в настоящее время получен грант РГНФ на интеграцию данных RussNet в YARN (Yet Another RussNet, проект Уральского Государственного Университета им. Горького, направленный на создание большого открытого тезауруса русского языка с использованием краудсорсинга), поэтому возникает задача проверить имеющиеся данные по областям RussNet, пополнить их максимально и поместить синсеты и семантические отношения в YARN.

# **Глава 1. Теоретические аспекты выявления конфликтных ситуаций**

## **1.1 Семантическое поле и лексико-семантическая группа слов**

### **1.1.1 Семантическое поле**

Идеи и принципы семантического анализа языка, которые впоследствии были объединены под общим понятием метода семантического поля, складывались постепенно и восходят к концу XIX – началу XX вв. В числе тех, кто занимался этой проблемой, были такие известные ученые, как А.А. Потебня, Р. Мейер, М.М. Покровский и др.

З.К. Тарланов, ссылаясь на академика М.М. Покровского, отмечает (Тарланов, 1995, с. 58), что тот, пытаясь найти систематизирующее начало в содержательной организации языка, писал в своей работе «Семасиологические исследования в области древних языков» следующее: "Слова и их значения живут не отдельной друг от друга жизнью, но соединяются в нашей душе, независимо от нашего сознания, в различные группы, причем основанием для группировки служит сходство или прямая противоположность по основному значению. Понятно уже а priori, что такие слова имеют сходные или параллельные семасиологические изменения и в своей истории влияют одно на другое; понятно также, что эти слова употребляются в сходных синтаксических сочетаниях».

Рихард Мейер в работе 1910 года «Bedeutungssysteme, Zeitschrift für vergleichende Sprachforschung» выделяет три темы семантических систем (классов):

1. Естественные (названия деревьев, животных, частей тела и пр.)
2. Искусственные (названия воинских чинов, составные части механизмов и пр.)
3. Полуискусственные (терминология рыбаков и охотников, названия этических понятий и пр.)



З.К. Тарланов, опираясь на работы Трира и Покровского, пишет (Тарланов, 1995), что семантическая система (семантический класс) и есть упорядоченность определенного числа выражений с той или иной точки зрения. "Точки зрения", с которыми "упорядочиваются" соответствующие слова, называются дифференцирующим фактором. Такими факторами могут быть, к примеру, значение целенаправленности, свойственное глаголу *ersteigen* и отсутствующее у глагола *steigen*, значение равномерной повторяемости, свойственное глаголу *marschieren* и отсутствующее у глагола *gehen* и т.д. С учетом этого задача семантического анализа сводится к тому, чтобы, во-первых, установить принадлежность слова к той или иной семантической системе и, во-вторых, выявить системообразующий, дифференцирующий фактор этой системы.

Общий вывод этих и подобных подходов к анализу и описанию лексической семантики заключается в постепенном утверждении системного взгляда на значение слов. Впоследствии этот взгляд нашел своё выражение в методе семантического поля.

Принципы метода семантического поля были сформулированы в 30-х годах XX в., его основоположником считается немецкий ученый Йост Трир. З.К. Тарланов, ссылаясь на Трира, выводит некоторые важнейшие постулаты, легшие в основу метода семантического поля Трира (Тарланов, 1995):

1. Вслед за Ф. де Соссюром Трир исходит из того, что язык определенного периода – это устойчивая и относительно замкнутая система, в которой слова наделены смыслами не в изолированном виде, а постольку, поскольку ими наделены и другие слова, смежные с первыми.
2. Общая система языка складывается из двух соотносительных друг с другом типов полей: а) понятийных полей, подразделяемых на элементарные единицы – понятия, и б) словесных полей, также подразделяемых на элементарные единицы – слова.
3. Единицы словесных полей полностью покрывают соответствующие понятийные поля, создавая своеобразную мозаику.

4. Семантические поля связаны между собой по принципу иерархичности подчинения (более широкие и более узкие). С течением времени семантические поля меняют свою структуру, тем самым изменяется лексическая система языка в целом.
5. Вслед за В. Гумбольдтом язык трактуется не как отражение объективной действительности, а как мировоззрение, характеризующееся самодовлеющей ценностью и по-своему расчленяющее действительность.

О.С. Ахманова определяет семантическое поле следующим образом: «Семантическое поле — это компактный, внутренне спаянный отрезок словаря, элементы которого взаимно ограничивают друг друга и, подобно мозаике, покрывают данную «понятийную сферу». Семантическое поле — уникальная монолитная структура, управляемая своими собственными законами. В ней реализуется, раскрывается «картина мира» и «иерархия ценностей», специфическая как для разных языков, так и для одного и того же языка в разные периоды его развития. Задача исследователя заключается, следовательно, в том, чтобы определить специфические для данного языка распределение, связь и взаимообусловленность значений» (Ахманова, 1957, с. 79).

Само описание лексической системы языка в свете сказанного предполагает прежде всего разбиение ее на лексические макросистемы, обслуживающие те или иные понятийные сферы, например, семантическое поле радости, семантическое поле родства, семантическое поле быта, семантическое поле обучения, семантическое поле цвета, семантическое поле одежды и украшений и т.д. Лексика, сосредоточенная вокруг каждого из семантических полей, образует отдельную микросистему. К установлению конечного числа таких микросистем и взаимных отношений между ними в языке определенного периода и сводится главная задача исследователя, отмечает З.К. Тарланов (Тарланов, 1995). При этом исследователь идет от общих понятий к конкретным языковым фактам.

С другой стороны, на иных началах строится «семантическое поле» Вальтера Порцига. Г.С. Щур отмечает (Щур, 1974), что в отличие от абстрактно-концептуального метода Й. Трира, метод В. Порцига за исходную точку принимает сами языковые факты. Для Порцига принципиальное значение имеет анализ отношений между тремя классами слов: глаголами, именами существительными и прилагательными. При этом самыми важными оказываются слова, способные выражать признаки и выполнять предикативную функцию: это глаголы и прилагательные. Они и семантически конкретнее существительных. В силу этого ядром "элементарного семантического поля", согласно В. Порцигу, могут быть только глагол и прилагательное. Задача, которая решается с помощью метода В. Порцига, – это выявление того, каким образом отдельные языковые элементы включаются в семантическое поле, ядро которого составляют те или иные глаголы или прилагательные. Так, например, глаголы *идти*, *хватать*, *петь* предполагают соответственно существительные *ноги*, *руки*, *голос*; прилагательные *белокурые*, *карие* в русском языке предполагают соответственно *волосы*, *глаза*. Таким образом можно включать все слова, обозначающие предметы и признаки, в том числе и производные, в элементарные семантические поля, характерные для данного языка. Это так называемые синтаксические поля, основывающиеся на валентных свойствах слов, в отличие от парадигматических полей Трира (Щур, 1974).

И.М. Кобозева в работе «Лингвистическая семантика» определяет семантическое поле как совокупность языковых единиц, объединенных общностью содержания и отражающих понятийное, предметное или функциональное сходство обозначаемых явлений, а также выделяет его характерные основные свойства (Кобозева И.М., 2000, с. 99):

- 1) наличие семантических отношений (корреляций) между составляющими его словами;
- 2) системный характер этих отношений;
- 3) взаимозависимость и взаимоопределяемость лексических единиц;

- 4) относительная автономность поля;
- 5) непрерывность обозначения его смыслового пространства;
- 6) взаимосвязь семантических полей в пределах всей лексической системы (всего словаря).

С точки зрения функционального подхода о функционально-семантических полях писал А.В. Бондарко в работе «Теория функциональной грамматики: Введение. Аспектуальность. Временная локализованность. Таксис» (Бондарко, 1987).

Стоит отметить, что метод семантических полей применим, в частности, к исследованию терминологической лексики. М.А. Чигашева в работе «Исследование терминологической лексики методом семантического поля» описывает с этой позиции группировку слов, обозначающую воинские звания (Чигашева, 2004).

К методу семантических полей близко примыкает также широко используемый в классической (традиционной) лексикологии метод тематического описания лексики.

### **1.1.2 Лексико-семантические группы слов**

О лексико-семантических группах (ЛСГ) написано очень много работ на материале разных языков. Однако выделение групп проходило как обычный рабочий прием и не квалифицировалось в качестве самостоятельного метода. Серьезная попытка придать ему статус метода была предпринята Ф.П. Филиным в середине 50-х годов XX века. Подвергнув краткому критическому анализу теорию семантических полей Йоста Трира, которую он расценивает как разновидность семантического структурализма, слишком абстрактную и потому оторвавшуюся от "реальной языковой почвы" (Филин, 1982), Ф.П. Филин противопоставляет ей теорию лексико-семантических групп, а также дает определение лексико-семантической

группе и описывает способ её выделения следующим образом (Филин, 1982, с. 231):

1. Лексико-семантические группы слов представляют собой объединения двух, нескольких или многих слов по их лексическим значениям. Под лексическим значением подразумевается при этом предметное содержание слова, т.е. соотнесенность слова с миром вещей, явлений и т.д.

2. Лексико-семантические группы слов складываются исторически и не являются раз и навсегда данными.

3. Лексико-семантическая группа близка к тематической группе лексики, но вместе с тем существенно отличается от нее. Выделение лексико-семантических групп сближается с тематической классификацией лексики тем, что а) и то и другое имеют свою "тему"; б) лексико-семантическая группа слов входит в соответствующее тематическое объединение слов, в) тематическая соотнесенность – один из важнейших, но не единственный признак лексико-семантической группы.

Различия между лексико-семантической группой и тематической группой заключаются в том, что, во-первых, в лексико-семантических группах слов находят свое выражение действующие в данном языке закономерности и правила развития значений слов, в то время как тематические группы слов характеризуют уровень познавательной деятельности народа: что ему известно и отразилось в его языке и что неизвестно, и, следовательно, во-вторых, тематические группы слов – это объединения слов на основе классификации обозначаемых ими предметов и явлений, в то время как в лексико-семантических группах реализуются собственно языковые закономерности.

4. Классифицировать слова по тематическим группам можно с разной степенью дифференциации и с разными целями. Например, можно выделить

группы лексем, являющиеся наименованиями предметов и явлений, относящихся к темам "Домашние животные", "Дикие животные", "Строительство и строительные процессы", "Растения", "Небо и небесные тела", "Домашняя утварь", "Сезонный крестьянский труд", "Школа", "Театр" и т.д. Таким образом, в принципе, можно расписать всю лексику языка и затем сделать соответствующие выводы в зависимости от целей исследования. Детализация тематических групп логически и последовательно ведет к группам лексико-семантическим. При таком последовательном переходе от одного плана к другому по принципу детализации отчетливее проявляется и различие между ними. Так, например, названия частей тела человека объединяются в одну тематическую группу, однако между ними может и не быть каких бы то ни было семантических связей, например: *спина* и *печень*, *голова* и *кожа*, *зуб* и *локоть*; *руки* и *ноги*, *глаза* и *уши*.

В тех случаях, когда тематические группы слишком ограничены, они могут совпадать с лексико-семантическими группами, тем самым обычно существующие между ними принципиальные различия нейтрализуются. Ф.П. Филин считает (Филин, 1982), что в отличие от слов тематических групп слова лексико-семантических групп характеризуется наличием между ними родовидовых соотношений, например: *руки* и *ноги* – *конечности*. Однако и это не является определяющим. Существенно то, что отношения между словами в тематических группах строятся только на внешних отношениях между понятиями. С другой стороны, лексико-семантические группы слов представляют собою внутреннее специфическое явление языка, обусловленное ходом его исторического развития. Ярким примером внутренней обусловленности компонентов лексико-семантических групп, согласно Ф.П. Филину, являются синонимы и антонимы.

Стоит также добавить, что А.Е. Бертельс в работе «Разделы словаря, семантические поля и тематические группы слов», ссылаясь на Ю.Н. Караулова, отмечает, что авторы в настоящее время зачастую произвольно оперируют терминами «семантическое поле», «лексико-семантическая

группа», «тематическая группа», «синонимический ряд», «ономасиологическая группа» и т.д. На основании того, что центром системы (макро- или микро-) является слово, а сама система предстает как сеть связей разных значений одного слова и связей между словами, он условно выделяет два вида групп слов, значительно отличающихся одна от другой (Бертельс, 1982, с. 54): «семантическое поле» — группу слов, которые в содержательном отношении определяют значения друг друга (поле «радости», «знания») и «тематическая группа» — слова, обозначающие чем-то сходные предметы действительности («названия птиц, растений»). Все эти группы занимают положение между единичными словами и совокупностью словаря. Они являются частью целого и сходны со словами в том, что вступают в сочетания, образуя единицы более высокого уровня, а также сходны со словарем в том, что могут быть разложены на меньшие единицы.

В нашей работе, исследуя семантическое поле конфликта, мы будем опираться на точку зрения В. Порцига, так как, с одной стороны, мы будем описывать глаголы и глагольные группы, являющиеся ядром нашего поля и образующие семантически маркированные лексико-семантические группы, и, с другой стороны, метод выделения семантических полей Порцига основывается на валентных свойствах слов.

## **1.2 Понятие валентности в лингвистике**

### **1.2.1 Структурно-семантическое описание валентностей**

Расцвет структурной лингвистики, пришедшийся на середину XX века, связан в том числе с развитием научно-технического прогресса. Возникновение ЭВМ привело к тому, что появилась возможность быстрой и эффективной обработки информации, а также решения таких прикладных задач, как машинный перевод, автоматическое реферирование текстов, поиск информации и т.д. Рассматривая внутренние причины, следует отметить, что исторически структурная лингвистика возникла намного раньше, а её

основоположником принято считать Фердинанда де Соссюра. Структурная лингвистика являла собой реакцию на описательную грамматику. Как отмечает Ю.Д. Апресян (Апресян, 1966), не отрицая заслуг описательной грамматики, накопившей бесценный фактический материал и опыт исследования связей между языковыми категориями, структурная грамматика старалась выработать общий план для используемых понятий. Отсутствие единой, последовательно построенной системы привело к критическому замечанию А. Мейе о том, что в лингвистике имеется столько же разных лингвистик, сколько лингвистов, её изучающих.

Отталкиваясь от учения Соссюра, структурная лингвистика со временем охватила все пласты языка с одной стороны, и все подходы к его изучению с другой. Возникновение классических школ структурной лингвистики (Пражской, Копенгагенской и Американской) связано лишь с различным акцентированием проблем, выделенных Соссюром. Кроме того, в начале XX века наметилась, а в дальнейшем развилась тенденция совмещать в один подходы к изучению различных дисциплин.

В 1960-е годы возникла Парижская семиотическая школа, основанная А.Ж. Греймасом. «Структурная семантика» Греймаса, первая фундаментальная работа школы, опиралась на лингвистический структурализм Соссюра, взгляды Копенгагенской и Пражской школ, а также структуралистский подход в антропологии К. Леви-Стросса и, в особой степени, фольклористике В.Я. Проппа. Семантику, как отмечает Греймас (Греймас, 2004), структуралистские школы, разрабатывавшие фонологию и лингвистическую теорию для обновления грамматических исследований, обходили стороной сознательно, так как вопрос о том, поддается ли семантика структурному анализу и может ли она рассматриваться как отдельная лингвистическая дисциплина, долгое время был открытым.

Так или иначе, для нас важно то, что в работе «Структурная семантика» Греймасом была выведена универсальная нарративная структура текста. Основанная на идеях Проппа, почерпнутых из «Морфологии волшебной

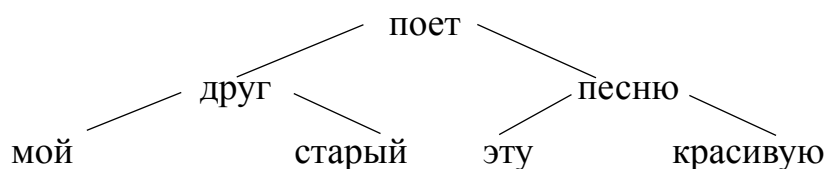


сказки», так называемая актантная схема включала в себя шесть актантных ролей, связанных между собой в бинарные оппозиции: отправитель/получатель, объект/субъект, помощник/оппонент.

В дальнейшем эта теория была развита в теорию семантических валентностей. В работе «Введение в общий синтаксис» Тарланов пишет (Тарланов, 1995), что семантической валентностью лексемы L называется любая (несвязанная) переменная X, входящая в толкование (описание значения) L. При этом всякая лексема, имеющая один или несколько партиципантов, называется предикатом. Иными словами, партиципантами глагола *наполнять* являются X (наполнять кого/что) и Y (наполнять чем). О том же говорит Богуславский (Богуславский, 1996, с. 26): «Слово имеет столько валентностей, сколько участников ситуации необходимо упомянуть, чтобы столковать его исчерпывающим и неизбыточным образом».

С позиций структурализма к синтаксису также подошел Л. Теньер в работе «Основы структурного синтаксиса», посмертно опубликованная в 1959 году. В ней Теньер последовательно различает статический и динамический синтаксис. Статический синтаксис, как пишет Теньер, является системой, охватывающей парадигмы слов, которые знает говорящий до момента формирования высказывания. Динамический синтаксис, в свою очередь, представляет собой реализацию элементом статического синтаксиса в процессе формирования речи. При этом Теньер говорит скорее о лексической составляющей, которая есть в сознании у говорящего, нежели о каких-либо моделях предложений, появившихся в научном поле зрения гораздо позже.

Синтаксические связи находятся в иерархии: вышестоящие элементы называются управляющими, а нижестоящие — подчиненными (Теньер, 1988). Нижестоящий элемент не может зависеть более чем от одного управляющего элемента, но не наоборот. По Теньеру схема зависимостей в предложении выглядит следующим образом:



Идея Теньера заключается в том, что подобная схема зависимостей, тем не менее, является многомерной, так как каждый управляющий элемент может иметь несколько подчиненных.

С другой же стороны речевая цепочка линейна и одномерна, так как разворачивается во времени, а значит одновременность каких-либо элементов в ней невозможна, зато в ней есть последовательность.

Таким образом, отмечает Теньер (Теньер, 1988), в основании всего структурного синтаксиса лежит соотношение между структурным порядком и порядком линейным.

Слово, по мысли Теньера, с зависимыми словами формирует так называемый узел или, иначе, ядро. Определение ядра носит частеречный характер: узел может быть глагольным, субстантивным, адъективным и адвербиальным. Во главе глагольного предложения всегда находится глагольное ядро, в подчинении которого находятся «сирконстанты» и «актанты». Число актантов, способных присоединяться к глаголу, Теньер называл валентностью.

### 1.2.2 Структурно-семантические типы валентностей

Теньер применял термин «валентность» исключительно к глаголам. К примеру, трехвалентным считается глагол «читать», так как при нем возможно возникновение 1) субъекта действия (*Тракторина читает*), 2) предмета действия (*читает Капитал*) и 3) объекта действия, то есть того, на кого действие направлено (*читает Паие*). При этом необходимо отметить, что, согласно Теньеру, совершенно необязательно все валентности какого-

либо глагола должны быть заняты соответствующими актантами. Некоторые валентности могут свободными.

Существуют безвалентные глаголы, которые не могут иметь актантов. Такие глаголы мы можем встретить в безличных предложениях (*Светает. Стемнело.*), а их значение чаще всего представляется как процесс или как состояние.

Далее идут одновалентные глаголы, известные под названием «непереходные»: *спать, падать, течь* и многие другие. Чаще всего, отмечает Теньер (Теньер, 1988), это глаголы как действия, так и состояния.

Двухактантные глаголы – это переходные глаголы. В предложении «*Колхозник косит траву*» действие как бы переходит от колхозника к траве.

Для трехактантных глаголов нет специального термина в традиционной грамматике, но они, с точки зрения Теньера, безоговорочно могут разделяться на два типа: глаголы речи (например, глагол *сказать*) и глаголы «давания» (например, глагол *передать*). Следует отметить, что подобное утверждение справедливо для французского языка. Для русского языка справедливо также утверждать о наличии глаголов, с валентностью более, чем три (например, глагол *купить* – *кто покупает, что покупает, у кого покупает и за сколько покупает*).

О типах валентности более подробно писал Соломон Давидович Кацнельсон в статье «К понятию типов валентности». С его точки зрения (Кацнельсон, 1987) валентность можно определить как заключающуюся в лексическом значении слова синтаксическую потенцию, то есть способность присоединить к себе другое категориально вполне определенное полнозначное слово. При этом валентность являет собой нечто большее, чем просто способность вступать в синтаксические отношения. Если само значение слова указывает на необходимость его восполнения словами различных типов, таким образом валентностью обладают не все полнозначные слова, а только обладающие неполнотой высказывания.

При этом следует различать валентность формальную и содержательную (Кацнельсон, 1987). В случае формальной валентности наблюдается связь с определенной словоформой и обусловленность элементами синтетической морфологии. Содержательная же валентность подразумевает под собой зависимость исключительно от значения слова.

Содержательна валентность одинакова во всех языках (приблизительно одного уровня развития), формальная же находится в зависимости от морфологического строя, а значит, что иногда содержательная и формальная валентности могут совпадать.

С точки зрения Кацнельсона все предикаты делятся на событийные и предикаты, выделяющие объекты. Первые сводятся к выделению события как такового. Последние же делятся на предикаты-свойства и предикаты-отношения. Предикаты-свойства одновалентны, выражают такие события, которые совершаются в объекте. Иными словами, это предикаты непереходные: *падать, лежать, прыгать* и т.д. Предикаты-отношения выражают события, которые затрагивают несколько объектов, при этом у каждого объекта есть своя роль. Это предикаты переходные: событие *сжечь* предполагает огонь и объект, который сжигается (*Пожар сжег дома* или *Дома сожгло пожаром*).

Кажется важным отметить посыл Кацнельсона, заключающийся в том, что «место» при глаголе следует четко отграничить от обстоятельственной характеристики глагольного действия, так как обстоятельственная характеристика возможна при любом глаголе. Место и время сопутствуют каждому действию и шире – предикативному слову.

Про семантическую валентность также подробно пишет И.В. Кобозева в книге «Лингвистическая семантика» (Кобозева, 2000). Она пишет, что у некоторого слова L есть семантическая валентность X, если слово L описывает ситуацию, в которой есть обязательный участник, выполняющий роль X. Значения слов, связанные с L и обозначающие обязательных участников ситуации, называются семантическими актантами этого слова.

Семантические актанты заполняют валентности слова L. Актанты и валентности при этом имеют разную логическую природу, потому их не стоит смешивать. Если семантический актант слова L является собой некоторое значение, которое отлично от значения слова L, то валентность подразумевает под собой постоянное внутреннее свойство самого слова L, обусловленное его значением. В семантической валентности слова содержатся роли участников ситуации – агенс, пациенс, инструмент, место и т.д.

Говоря о семантических отношениях, здесь будет уместным вспомнить схожее понятие в другой теории языка – *глубинный падеж* в падежной грамматике Ч. Филлмора (Филлмор, 1981) «Глубинных падежей», то есть описание семантических отношений, находящих формальное выражение в языке, в падежной грамматике Филлмора шесть:

1. Агентив – падеж одушевленного субъекта действия
2. Инструменталис – падеж неодушевленной силы или предмета, включенного в действие или состояние, которое глагол называет в качестве причины.
3. Датив – падеж одушевленного существа, которое затрагивается состоянием или действием, называемым глаголом.
4. Фактитив – падеж предмета или существа, которое возникает в результате действия, называемого глаголом.
5. Локатив – местоположение или пространственная ориентация действия или состояния, называемого глаголом.
6. Объектив – семантически нейтральный падеж; что-либо, что может быть обозначено существительным, роль которого определяется интерпретацией самого глагола.

Другие ученые по-разному выделяли количество и качество семантических отношений: у Ю.Д. Апресяна их двадцать пять (Апресян, 1995), у Н.Н. Леонтьевой их количество доходит до ста.

Кобозева также указывает (Кобозева, 2000) на то, что семантические валентности и актанты следует последовательно отличать от синтаксических

актантов и валентностей: первые характеризуют план содержания речевого отрезка, а вторые – план выражения.

Синтаксические актанты слова – это возможные при данном слове «сильноуправляемые» зависимые члены: подлежащее и дополнение. При этом, как уже указывалось ранее, нередко семантическому актанту (например, инструмента) не соответствует никакого синтаксического актанта (например, в предложении *Колхозник ударил председателя*). В таком случае справедливо говорить о синтаксической факультативности соответствующей валентности.

Суммируя вышесказанное, Кобозева указывает на главную проблему определения семантических валентностей слова – проблема правильного проведения границы между валентностями, связывающими значение слова с его актантами, и более слабым типом семантических зависимостей, соединяющих значение слова (или более сложного языкового выражения) с его сирконстантами. Для разрешения этого вопроса Кобозева рекомендует придерживаться следующих критериев (Кобозева, 2000).

1. Обязательность соответствующих участников (аспектов) для обозначаемой словом ситуации говорит о том, что они являются семантическими актантами. Время и место в таком случае принято считать сирконстантами, так как они присущи любому действию вообще.

2. Семантических валентностей у слова бывает от одной до трех, реже четыре и более.

3. Для семантических актантов характерна идиоматичность морфологического выражения: оно зависит не только от содержания соответствующих валентностей, но и от лексемы, которой принадлежит актант. Так, участник ситуации с одной и той же семантической ролью адресата при глаголе *сообщать* выражается именем в дательном падеже, а при глаголе *извещать* — в винительном. Это явный показатель того, что мы имеем здесь дело с семантическим актантом соответствующих лексем.

Стоит обратить внимание еще на один факт. Несмотря на то, что, к примеру, предложения *Колхозник подружился с председателем* и *Колхозник и председатель подружались* являются синонимичными, при добавлении некоторого обстоятельственного значения они проявляют сильные различия: *После долгих уговоров колхозник подружился с председателем* и *После долгих уговоров колхозник и председатель подружались*.

Подобную закономерность можно объяснить в терминах семантических ролей. Тестелец отмечает (Тестелец, 2001), что у одного предикатного слова не может быть двух актантов с одной и той же семантической ролью, что выглядит очевидным. Так или иначе, в результате совместной работы Филлмора, Апресяна, Грубера, Богданова и других лингвистов был разработан примерный универсальный перечень семантических ролей. В этот перечень обычно включают такие роли, как:

1. Агенса
2. Пациенса
3. Бенефактив (реципиент)
4. Экспериенцер
5. Стимул
6. Инструмент
7. Адресат
8. Источник
9. Цель

### **1.2.3 Практическое применение теории валентностей семантических ролей**

В настоящее время в свет выходит большое количество работ прикладного характера, использующие вышеизложенные идеи в разрешении проблем компьютерной лингвистики.

Хотелось бы особенно выделить работу под названием *Exploiting semantic role resources for preposition disambiguation* за авторством Tom O’Hara и Januse Wiebe. Эта работа посвящена разрешению семантической

неоднозначности предлогов при помощи семантических ролей. Семантические роли, используемые для решения задачи, прописаны в аннотированном корпусе Penn Treebank и базе знаний Factotum. Разрешение неоднозначности, как утверждают авторы (O’Hara T., Wiebe J., 2008), опирается на отношения, определенные по существующим фразам, а также при помощи гиперонимов в WordNet как коллокаций.

Авторы использовали 2 разных типа ресурсов семантических ролей, делая акцент на аннотированный корпус, при этом в Penn Treebank отображены лишь основные роли, в то время как FrameNet включает в себя большое количество ролей, характерных для аргументов из специфичных областей. Проводя различные эксперименты, авторы статьи установили (O’Hara T., Wiebe J., 2008), что наилучший результат разрешения предложной неоднозначности показал метод, совмещающий в себе как использование семантических ролей, так и словесных и гиперонимических коллокаций.

Еще одна работа, напрямую связанная с семантическими ролями – «The importance of syntactic parsing and inference in semantic role labeling». В работе авторы представляют метод автоматического выделения и классификации семантических ролей, основанный на машинном обучении с линейной процедурой логического решения (Punyakanok V., 2008). В рамках этого метода авторы исследуют роль синтаксической информации в классификации семантических ролей.

Также представляется интересной совместная работа Lluís Marquez и Xavier Carreras под названием «Semantic role labeling: an introduction to the special issue». Эта обзорная работа совмещает в себе результаты различных экспериментов по автоматическому определению семантических ролей, использующих совершенно разные методы (Marquez L., Carreras X., 2008): некоторые из них опираются исключительно на статистические модели, другие используют такие базы, как FrameNet, ProbBank и VerbNet, третьи же делают упор на синтаксис. Авторы утверждают, что у каждого из методов есть не только свои исключительные достоинства, но и не менее



исключительные недостатки. Помимо этого, необходимо также решить такие методологические вопросы, как способы проведения исследования и, что особенно важно, оценка результатов. Авторы призывают универсализировать измерения, так как это поможет более точно составить качественный анализ.

Отдельно хотелось бы выделить статью Е.В. Мартынюка под названием «От валентности к фреймовой семантике», описывающей взаимосвязь понятий валентности и фрейма, а также как эти понятия развивались. О развитии понятия «валентность» у разных ученых было подробно описано выше. Стоит лишь отметить, что Мартынюк, ссылаясь на книгу «Теория функционально-коммуникативного синтаксиса» М.В. Всеволодова, рассматривает валентность на трех уровнях (Мартынюк, 2014, с. 170):

1. Логический уровень валентности универсален (определяется событиями и предметами реального мира).
2. Семантический уровень определяется значением слова.
3. Синтаксический уровень определяет: 1) облигаторное / факультативное заполнение позиции актантов, 2) их количество, 3) синтактико-морфологические характеристики.

Тем не менее, традиционное понимание валентности в настоящее время ограничено вербоцентричной концепцией Л. Теньера, так как на этом уровне легче соотнести между собой слово и предложение.

Другой попыткой, как отмечает Мартынюк (Мартынюк, 2014), описать взаимодействие элементов высказывания на семантическом уровне была предпринята в рамках генеративной лингвистики. Ч. Филлмор, отказавшись от термина «актант», а точнее от его морфологической интерпретации, использовал термины «падеж» и «падежная рамка», называя так падежное окружение глагола. В работе Филлмора «Дело о падеже» термин «рамка» – «frame» – «фрейм» еще не обладал самостоятельностью, но с развитием лингвистики, а в частности – психолингвистики, в 1974 году М. Минский был предложен психолингвистический термин «фрейм», обозначающий

«структуру знаний для представления стереотипной ситуации» (Минский, 1979, с.24).

### **1.3 Фреймовый способ описания семантики**

#### **1.3.1 Понятие фрейма**

В настоящее время понятия валентности и фрейма размежевывают, но между ними прослеживается четкая историческая связь. Термин «валентность» в настоящее время относят к вербоцентрической концепции предложения, в то время как «фрейм» является частью семантики и, что немаловажно, синтаксиса.

Как было отмечено в предыдущей главе, впервые термин «фрейм» был употреблен Ч. Филлмором в работе «Дело о падеже». Впрочем, тогда еще термин не обладал еще самостоятельностью и, как отмечает Мартынюк (Мартынюк, 2014), был скорее произвольно использован для описания употребления глагола, а точнее, глагольного управления, только в новых терминах.

В 1974 году Марвин Минский предложил термин «фрейм» как психолингвистическое понятие. Используя локальную и общую теории зрительного восприятия, Минский высказал идею о том (Минский, 1979), что все процессы нашего мышления основаны на многочисленных структурах, хранящихся в нашем сознании, основанных на знаниях и культурных условиях – фреймах. Зрительное восприятие, говорит Минский, крайне сложный и длительный процесс, но нам не нужно много времени, чтобы, зайдя в комнату, сориентироваться в ней. Мы как будто уже знаем, что в ней будет (Минский, 1979, с.15). Опираясь также на теории гештальт-психологов, в работе «Фреймы для представления знаний» Минский старается раскрыть суть процессов, лежащих в основе соотнесения сенсорных данных либо с общими понятиями, либо с их частями. «В теории фреймов упор делается на многочисленность взаимодействий между сенсорными данными и

образованной в процессе обучения громадной сетью символической информации» (Минский, 1979, с.15).

В конечном счете Минский ставил перед собой задачу создания модели, которая способна в формализованном виде представить те аспекты знаний, которые связаны с человеческим восприятием, в частности, с пониманием текстов. Также, что немаловажно, одной из задач было представление этой модели в памяти компьютера.

По Минскому фреймом является структура данных для представления стереотипной ситуации, с каждой из которых ассоциирована информация разных видов. «Одна ее часть указывает, каким образом следует использовать данный фрейм, другая - что предположительно может повлечь за собой его выполнение, третья - что следует предпринять, если эти ожидания не подтвердятся» (Минский, 1979, с.5).

Фрейм можно представлять себе в виде сети, состоящей из узлов и связей между ними. "Верхние уровни" фрейма четко определены, на более низких уровнях имеется много особых вершин-терминалов или "ячеек", которые должны быть заполнены теми или иными данными.

Каждым терминалом могут устанавливаться условия, которым должны удовлетворять его задания. Простые условия определяются маркерами, например, в виде требования, чтобы заданием терминала был какой-либо субъект, или предмет подходящих размеров, или указатель на субфрейм определенного типа.

Здесь следует отметить, что в приложении к русскому изданию работы «Фреймы для представления знаний» Ф.М. Кулаков приводит (Минский, 1979) примеры фрейма-визуального образа и, что наиболее важно для нас, фрейма-сценария.

Фрейм-визуальный образ, пример которого основан на изображении куба, заключается в процедуре построению системы фреймов, соответствующих перемещению точки наблюдения вокруг куба, так как каждая позиция наблюдения куба отличается друг от друга (появление новых

граней и исчезновение или перемещение предыдущих), при этом каждая новая позиция представляет собой отдельный фрейм.

Фрейм-сценарий же строится на более сложных принципах, ибо представляет собой систему невизуальных фреймов. Так или иначе, как утверждает Кулаков, любые знания о мире можно представить себе в виде некоторых общих универсальных фрагментов, которым соответствуют свои системы фреймов и из которых можно формировать новые системы фреймов, соответствующие некоторым новым представлениям.

Минский сам утверждает, что знания о мире следует строить в виде фреймов-сценариев. Подобный фрейм, описывающий какое-либо событие, должен включать в себя необходимые узлы фрейма. Например, фрейм защиты диссертации студентом будет иметь следующие заполненные данными слоты:

Одежа: строгая, классическая, праздничная.

Подарок: цветы,

и так далее. По существу, отмечает кулаков, фрейм-сценарий можно отождествить с собранием вопросов, необходимых задать относительно некоторой гипотетической ситуации, и способов ответа на них.

Что наиболее важно для нас, здесь крайне применима теория семантических ролей, так как наиболее важные вопросы (к примеру, относительно ситуации защиты диссертации) можно представить в виде следующей таблицы:

Вопрос	Ответ	Семантическая роль
Кто осуществляет действие?	Студент	Агент
Какова цель действия?	Успешная защита ВКР	Намерение
Каковы последствия?	Получение ученой	Эффект

	степени	
На кого это действие влияет?	Комиссия	Получатель
Каким образом оно произведено?	Публичное выступление, презентация	Инструмент

В зависимости от ситуации вопросы, конечно же, могут быть совершенно разные, но по мере того, как на каждый вопрос дается предварительный ответ, из памяти могут вызываться новые фреймы, соответствующие ситуациям, которые возникают в результате ответов на вопросы. Вопросы - терминалы этих новых фреймов становятся в свою очередь активными.

У каждого сценария имеется ряд исполнителей ролей. При этом немаловажно то, что он может иметь различные интерпретации, отражающие точку зрения того или иного исполнителя (к примеру, в нашей ситуации защиты диссертации этот фрейм может быть описан как с точки зрения студента, так и с точки зрения комиссии, и несмотря на то, что все находятся в рамках одной ситуации, описание будет кардинальным образом отличаться). Именно поэтому каждую ситуацию можно рассматривать как систему фреймов.

Здесь можно привести пример сценария защиты диссертации с точки зрения защищающегося студента:

Сценарий: защита выпускной квалификационной работы

Роли: студент, комиссия, научный руководитель, рецензент, другие студенты.

Цель: успешная защита ВКР

*Сцена 1. Вход*

- Войти в аудиторию
- Отдать деньги коллегам за цветы
- Найти свободное место

- Сесть

#### *Сцена 2. Защита чужих работ*

- Выслушать вступительное слово комиссии
- Выслушать выступления коллег
- Дождаться своей очереди
- Выйти, когда назовут твою фамилию

#### *Сцена 3. Защита своей работы*

- Поприветствовать комиссию и коллег
- Выступить с защитой
- Выслушать научного руководителя и рецензента
- Ответить на вопросы
- Сесть на своё место

#### *Сцена 4. Окончание защиты выпускных работ*

- Выслушать остальных докладчиков
- Выйти из аудитории
- Ждать, пока комиссия совещается
- Вернуться в аудиторию
- Выслушать итоги защиты и заключительное слово комиссии
- Покинуть аудиторию

Кажется очевидным, что у каждого фрейма есть название и набор слотов с уникальными названиями. В то же время слот может ссылаться на другой фрейм (например, «выступить с защитой» вполне может являться отдельным фреймом со своей структурой), образуя тем самым сеть фреймов.

Конечно же, в каждом сценарии средства выполнения действий, а порядок нарушаться (отклонение от сценария, возникновения препятствия и ошибка). Таким образом, сценарий – это не столько цепочка событий, сколько причинно-следственная цепочка действий, которая способна разветвляться на множество путей.

Возвращаясь к Ч. Филлмору, нужно отметить, что тот опирался на популярные в то время принципы когнитивной лингвистики: концепты нельзя изучать в изолированном виде, так как он всегда включен в какую-

либо более крупную область знаний. Поэтому фрейм считается структурой когнитивного уровня, совмещающая в себе область когнитивного с языковым.

Фрейм в когнитивной лингвистике – это оперативная единица памяти, включающая в себя стереотипные ситуации и сценарии. С другой стороны, как отмечает В.А. Маслова в книге «Введение в когнитивную лингвистику» (Маслова, 2011), фрейм (наряду с картиной, гештальтом, сценарием) является одним из типов структур представления знаний, описания того или иного концепта.

На лексическом уровне фреймовое понимание значения слова позволяет нам понять такое, на первый взгляд, странное предложение, как, к примеру, *магнитофон поднял ему настроение*. Даже при помощи толкового словаря «марсианин» Эрика Бёрна не смог бы понять смысл этого предложения, но мы, люди, включенные в культуру, понимаем, что магнитофон может выступать, например, в качестве покупки или подарка, приносящего удовлетворение, радость.

С точки зрения синтаксиса теория фреймов позволяет уйти от логического подхода к членам предложения и обратиться к семантической стороне слов. Мартынюк приводит следующий пример (Мартынюк, 2014, с.171): «В предложении *Ветер унес шляпу* формально *ветер* является подлежащим, т. к. этой единице приписывается действие, названное предикатом. Если рассуждать в терминах актантов и обращаться к семантике, то чтобы быть Агенсом лексема *ветер* не вполне удовлетворяет требованиям – ветер не обладает волей, чтобы совершать целенаправленное действие, и именно теория фреймов помогает разрешить дилемму. Во фрейме глагола *унести* обязательно «закодирован» тот, кто совершает действие, и даже если лексема *ветер* не обладала этим свойством сама по себе, но в позиции рядом с глаголом *унести* она приобретает это свойство».

Для нас более подходит определение понятия «фрейм» самого Чарльза Филлмора, которое он дает в рамках разработки FrameNet: «Фрейм – это схематическая репрезентация ситуации, включающая различных участников,

предметы и другие концептуальные роли, каждая из которых является элементом фрейма. Семантический аргумент предикативного слова соответствует элементам фрейма, связанного с этим словом» (Fillmore, 1976, с. 22).

### 1.3.2 Фрейм как способ представления знаний

FrameNet – один из самых известных проектов, основанных на теории фреймовой семантики. Электронный ресурс был основан в 1997 г. в Беркли, в International Computer Science Institute Чарльзом Филлмором, который, как уже было сказано, разрабатывал свою теорию с учетом её прикладного использования в информационных технологиях.

База данных FrameNet насчитывает около 1200 семантических фреймов, 13000 лексических единиц (включающих в себя пары слово-значение; полисемичные слова представлены несколькими лексическими единицами) и около 190000 примеров предложений. FrameNet стал проектом, повлиявшим как на лингвистику в целом, так и на обработку естественного языка в частности, что привело к задаче автоматического определения семантических ролей, работы по которым были описаны выше.

Элементы фрейма обеспечивают дополнительную информацию о семантической структуре предложения. Так каждый фрейм имеет некоторое число «ядерных» и «периферийных» фреймовых элементов, которые можно рассматривать как семантические роли. Элементы «ядра» имеют ведущую роль в семантическом значении фрейма, в то время как периферийные элементы, как правило, имеют описательный характер (например, время, место, способ и т.д.). Например:

1) единственным ядерным элементом фрейма *Being born* является *Child*, а периферийными фреймовыми элементами являются *Time*, *Place*, *Relatives* и т.д.



2) Ядерные фреймовые элементы фрейма *Commerce\_goods-transfer* включают в себя *Seller, Buyer, Goods*, в то время как на периферии находятся такие элементы, как *Place, Purpose* и т.д.

Лексические единицы представлены леммами с указанием их частеречной принадлежности, которые вызывают тот или иной фрейм. Другими словами, когда лексическая единица определена в предложении, она может быть соотнесена со своим отдельным фреймом. Существует множество лексических единиц, соотнесенных с одним фреймом, и множество фреймов, соотнесенных с одним словом (как правило, обладающим несколькими значениями). Например, лексические единицы, которые вызывают фрейм *Complaining* включают в себя слова «жаловаться», «сетовать», «ворчать» и другие.

Фреймы также соотнесены с примерами предложений, в которых фреймовые элементы маркированы. К примеру, предложение *She was born about AD 460* соотнесено по глаголу с фреймом *Being\_born*, в котором *She* помечено как фреймовый элемент *Child*, а *about AD 460* как элемент *Time*.

FrameNet также отображает отношения между различными фреймами, такие как *Inheritance, Subframe, Using, Perspective On, Inchoative Of, Causative Of, Precedes*, а также *See also*, соединяющий фреймы, имеющие некоторое сходство, но требующие тщательного различения.

FrameNet доказал свою полезность в числе компьютерных приложений, так как компьютером требуется дополнительное знание в случае, когда нужно распознать, что «Иван продал машину Марии» и «Мария купила машину у Джона» описывают одну и ту же ситуацию. Также FrameNet используется при разрешении таких задач, как перефразирование, извлечение информации, поиск ответа на вопрос, присвоение семантических ролей и т.д. Первая автоматическая система для присвоения семантических ролей (SRL), основанная на FrameNet, была разработана в 2002 году Daniel Gildea и Daniel Jurafsky.

Фреймово-семантический парсинг является также одним из перспективных направлений в последнем десятилетии. В работе *Frame-semantic parsing* описывается метод, совмещающий в себе как работу с данными FrameNet, так и статистические модели для разрешения семантической неоднозначности (Das, 2014). На FrameNet (а также BabelNet) опирались авторы статьи *Any-language frame-semantic parsing* (Anders, 2015), работа которых заключалась в аннотировании корпусов текстов, собранных на базе сайтов Wikipedia и Twitter, причем процент ошибок, согласно их результатам, относительно невелик.

Для российской науки представляется значимой работа О.В. Фурера «Фреймовый способ описания лексической семантики», в которой автор описывает цели создания метаязыка ALADDIN (*action lexemes and their derivatives describing invention*). Этот формализованный метаязык, как утверждает автор (Фурер, 2009), был создан для описания семантики лексических единиц, базирующийся на категориях теории фреймов и сценариев, с целью выявления основной реляционной структуры значений, которое необходимо для установления системного характера рассматриваемого семантического поля.

## **1.4 Электронные тезаурусы как способ описания семантических отношений**

### **1.4.1 WordNet**

WordNet – это электронный тезаурус английского языка, группирующий слова в наборы синонимов, называемых синсетамии. База данных начала разрабатываться в 1985 году Лабораторией Когнитивных Наук Принстонского Университета под руководством профессора психологии Дж. А. Миллера и продолжает пополняться до сих пор.

В настоящее время база насчитывает порядка 160 тысяч слов, собранных в 118 тысяч синсетов. WordNet включает в себя такие лексические

категории, как существительные, глаголы, прилагательные и наречия. Таким образом, слова одной части речи, представляющие собой синонимы в широком значении, объединяются в синсеты. При этом в случае, если слово появляется в нескольких синсетах, его считают многозначным.

Все слова и синсеты соединены между собой при помощи семантических отношений, например, таких, как:

- гиперонимия (Слово *зверь* – гипероним к слову *собака*)
- гипонимия (Слово *чихуахуа* – гипоним к слову *собака*)
- меронимия (Слово *окно* – мероним к слову *здание*)
- холонимия (Слово *здание* – холоним к слову *окно*)
- антонимия (Слово *жар* – антоним к слову *холод*)

и других.

При запросе к тезаурусу в виде какого-либо слова выводится следующая информация о нем:

- Принадлежность к части речи
- Список синсетов, в которые входит это слово
- Примеры употребления, определение значения слова, его парадигматические связи.

Также стоит отметить выделение семантических областей для каждой части речи. Для существительного их 25 (*время, действие, животные, знание, лицо, местоположение* и др.), для глаголов их 38 (*событие, состояние, действие, путь, место* и др.)

Проект EuroWordNet, построенный на базе американского WordNet, объединил в себе несколько систем для таких европейских языков, как французский, немецкий, итальянский, испанский, голландский, чешский и эстонский. При этом следует отметить, что EuroWordNet отличался от своего американского предшественника, с одной стороны, опорой на лексикографические данные (а не на психолингвистические, как при построении WordNet), а с другой добавлением отношений между разными частями речи.

Проекты наподобие WordNet представлены на различных языках: французском (WOLF), китайском (CWN), финском (FinnWordNet), немецком (GermaNet), русском (RussNet) и других. В силу особенностей языков, а также из-за различий в подходах и методиках многие тезаурусы отличны от WordNet в некоторых деталях. С целью объединения всех WordNet проектов, опирающихся на различные языки, в 2001 году была создана Всемирная ассоциация WordNet.

С 2013 по настоящее время Уральским Государственным Университетом под руководством Павла Браславского разрабатывается проект YARN (Yet Another RussNet), целью которого является создание большого открытого электронного тезауруса русского языка. Работа ведется с привлечением краудсорсинга. Основная идея проекта — эксперимент по комбинированию традиционных принципов создания ворднетов и вики-подхода к наполнению и редактированию лингвистических ресурсов.

#### **1.4.2 RussNet – электронный тезаурус русского языка**

RussNet – это проект, разрабатывающийся на кафедре математической лингвистики СПбГУ.

Следует назвать основные черты, которые унаследовал RussNet от Принстонского WordNet и EuroWordNet (Азарова, Митрофанова, Синопальникова, 2002):

1. Сохранена общая установка на отображение лексической системы языка в целом (а не только узкоспециальной, терминологической лексики), причем это должен быть, с одной стороны, словарь базовой лексики, активно употребляемой на протяжении длительного периода развития языка, с другой стороны, словарь активной лексики, широко употребляемой в газетно-публицистическом жанре в последние десятилетия.

2. Основными частями речи являются существительные, глаголы, прилагательные и наречия.

3. Основными структурными единицами являются слово и синонимический ряд (синсет).

4. Слова, связанные между собой отношением синонимии, образуют базовые единицы словаря — синсеты.

5. На синсетах заданы синтагматические и парадигматические отношения гипонимии/гиперонимии, тропонимии, меронимии/холонимии, отношения каузации, пресуппозиции, лексического вывода и т.д., а также для некоторых слов – отношение антонимии и деривационные отношения.

Корпус, на котором основывается RussNet, представляет собой корпус современных текстов, включающий в себя период с середины 80-х по начало нулевых годов. При этом в нем особенно представлен жанр газетных текстов (40%) как жанр, наиболее быстро откликающийся на изменения в языке, достаточно экспрессивный и вариативный; большая доля (30%) научно-популярных текстов как экспрессивно нейтральных и описывающих реалии не только обыденной жизни, но и других сфер; небольшая часть (20%) отрывков из художественной литературы, причем важным является отсутствие произведений, взятых целиком, а также небольшая часть (10%) текстов законов, договоров, инструкций и проч., обеспечивающая конструкциями современных клише делового употребления слов (Азарова, 2002).

Корпус текстов состоит из 21 миллиона словоупотреблений. На начальном этапе были отобраны наиболее встречающиеся слова, из которых существительных около 1000, глаголов 500, прилагательных 200, наречий 100. Кроме того, практически все частотные слова русского языка являются многозначными.

Отдельно следует обратить внимание на понятие синсета. Синсет (Азарова, Митрофанова, Синопальникова, 2002) объединяет в себе все лексико-семантические варианты слов, которые выражают данное понятие и связаны между собой отношением синонимии. К примеру, в проекте YARN

(открытый тезаурус отмечено (<https://russianword.net/words/206>), что слово *рука* входит в следующие синсеты:

- рука, покровительство, протекция, заручка, патронирование, патронаж
- длань, кисть, ладонь, ладоши, ладошка, ладошки, лапа, пятерня, рука, грабка
- рука, связи, блат, знакомство
- бок, грань, край, рука, сторона, фланг

и другие. Слова в синсетах RussNet объединены синонимией, а также дополняется словарным определением и примерами употребления слов в контексте.

RussNet имеет крайне широкое практическое применение. К примеру, в статье «Представление устойчивых лексических сочетаний в компьютерном тезаурусе RussNet» в рамках задачи по выделению устойчивых сочетаний используются рамки валентностей, причем в RussNet различается 2 типа валентностей: активная (для признаков слов – глаголов, прилагательных и их дериватов) и пассивная (для зависимых слов – наречий и существительных). Активная валентность, как отмечается в статье Азарова, Синопальникова, Смрж), наиболее информативна; активная рамка валентностей некоторого признакового слова задает грамматические и семантические параметры его зависимых слов.

## 1.5 Выводы

В приведенной главе были рассмотрены теоретические положения, необходимые для описания метод выявления конфликтных ситуаций. Отдельно были рассмотрены теории семантических полей и лексико-семантических групп, что необходимо нам для раскрытия семантического поля конфликтной ситуации как таковой. В нашей работе, исследуя семантическое поле конфликта, мы будем опираться на точку зрения В. Порцига, так как, с одной стороны, мы будем описывать глаголы и

глагольные группы, являющиеся ядром нашего поля и образующие семантически маркированные лексико-семантические группы, и, с другой стороны, метод выделения семантических полей Порцига основывается на валентных свойствах слов.

Далее была рассмотрена теория валентностей. В представленной работе она является одной из основных теорий, так как, опираясь на классическое понимание рамок валентностей (основанной на вербоцентрической концепции Теньера) мы выделяем в тексте основные глаголы и их окружение, при этом на передний план выходят их семантические роли. Мы придерживаемся точки зрения Кобозевой на разграничение синтаксических актантаов: для нас наиболее важными является обязательность наличия тех или иных участников ситуации, их «сильноуправляемость», выражающая в синтаксисе в ролях подлежащего и дополнения (например, субъекта и объекта). Кроме этого, для нас будет важна хронотопическая характеристика действия.

Теория фреймов, активно используемая в современных исследованиях, начало которой было положено Ч. Филлмором и М. Минским, также является для нас важной опорой, так как описание конфликтной ситуации с позиции фреймов и сценариев позволяет нам разложить конфликт на составные части. Опираясь на идеи Минского, во второй главе мы будем выстраивать прототип фрейма-сценария, связанного с теорией семантических ролей, в котором будут указаны действующие лица ситуации. Кроме того, для нас также будет важна идея Минского о точке зрения, что означает, что каждый сценарий может представлять собой целую систему фреймов.

Наконец, описание таких проектов, как WordNet, EuroWordNet и Russnet позволяет нам проиллюстрировать актуальность представленного исследования, так как, опираясь на методологию RussNet, наш подход может быть применен в прикладной направленности (в частности, при интеграции данных из RussNet в YARN).

## **Глава 2. Состав и структура фреймов конфликтных микро- и макроситуаций**

### **2.1 Исследуемый материал**

Первоначально нами был собран корпус текстов, основанных на статьях новостных сайтов (таких, как [lenta.ru](http://lenta.ru), [grani.ru](http://grani.ru), [gazeta.ru](http://gazeta.ru), [russian.rt.com](http://russian.rt.com)), посвященных вооруженным конфликтам с 2010 по 2014 год. Корпус насчитывал около 100 тысяч словоупотреблений. В процессе работы выяснилось, что особенностью подобных новостных текстов является то, что в них довольно редко описывается непосредственно ход вооруженных действий. Напротив, по теме какого-либо конфликта наиболее часто встречаемыми материалами являлись те или иные социальные, политические или экономические проблемы, сопряженные с самим конфликтом. Исследование метода по выявлению конфликтных ситуаций в подобном корпусе представляло собой неразрешимую задачу, а работа не принесла никаких результатов вследствие низкой репрезентативности корпуса. Поэтому было принято решение изменить подход к созданию корпуса.

Корпус, на основе которого было проведено представленное исследование, собран на материале Википедии. В него вошло около 30 обширных статей, посвященных описанию вооруженных действий в 2000-2015 годах, например, таких, как: Афгано-пакистанский пограничный конфликт, конфликт в Южной Осетии, конфликт на Украине, гражданская война в Либерии, Ливии и другие. Общий объем корпуса составляет около 60 тысяч словоупотреблений. Несмотря на свой объем корпус крайне репрезентативен как с точки зрения представленности в нем конфликтных ситуаций (в нем подробно описаны все виды боевых действий и связанные с ними элементы (место, время и т.д.), интересующие нас в данном исследовании), так и с точки зрения лексики.



## 2.2 Лексика и рамки валентностей

При описании рамок валентностей, опираясь на вербоцентрическую концепцию Теньера, следующим этапом нашей работы было выявление наиболее частотных глаголов в корпусе.

Глаголы были отобраны при помощи Томита-парсера – разработки компании «Яндекс», являющейся инструментом для извлечения структурированных данных (фактов) из текста на естественном языке. Извлечение производится при помощи контекстно-свободных грамматик и словарей ключевых слов.

Частотными считались те глаголы, которые встретились в корпусе более 15 раз. Из них были отобраны глаголы, имеющие явное отношение с точки зрения семантики к военным действиям. Глаголы, имеющие наиболее абстрактную семантику были отброшены за исключением глаголов «быть» и «стать», которые были также обработаны при помощи Томита-парсера с целью извлечь наиболее встречаемые глагольные сочетания. Получившийся список отобранных глаголов и глагольных сочетаний выглядит следующим образом:

1. Нанести удар/авиаудар
2. Действовать
3. Пострадать
4. Атаковать
5. Отступить
6. Передать
7. Контролировать
8. Потерять
9. Занять
10. Захватить
11. Начаться
12. Погибнуть
13. Начать
14. Участвовать
15. Быть атакованным

- 16.Быть убитым
- 17.Быть примененным
- 18.Быть захваченным
- 19.Быть уничтоженным
- 20.Стать жертвой

Также была проведена работа по разделению получившегося списка на «ядро» семантического поля и его «периферию». Таким образом, «ядерными» глаголами оказались:

1. Атаковать
2. Потерять
3. Захватить
4. Нанести удар/авиаудар
5. Контролировать
6. Занять
7. Пострадать
8. Погибнуть
9. Отступить
- 10.Быть убитым
- 11.Быть уничтоженным

Основанием для подобного выделения служили частотность в корпусе с одной стороны и практически полное отсутствие полисемии с другой, что означает высокую используемость этих глаголов при описании интересующих нас вооруженных конфликтов.

После классификации отобранных слов было произведено построение их актантных моделей, а также были отмечены те или иные особенности каждой из конструкций. Для возможности дальнейшей классификации за условные обозначения основных единиц были приняты следующие символы:

V – опорный глагол

Суб – агенс, (тот, кто осуществляет действие)

Объ – пациенс (объект в широком смысле)

Кауз – факультативный элемент, обозначающий категорию причины

Мест – факультативный элемент, обозначающий место

Падежи: 1– именительный 2 – родительный 3 – дательный 4 – винительный

5 – творительный 6 – предложный

Список актантных моделей с различного рода пометами выглядит следующим образом:

Атаковать Суб1 V Обь4

Потерять Суб1 V Обь4

Захватить Суб1 V Обь4

Нанести удар Суб1 V по Обь3

Контролировать Суб1 V Обь4.

Занять Суб1 V Обь4

Пострадать Суб1 V от Кауз2;

Погибнуть Суб1 V в Кауз6 Кауз2

Отступить 1) Суб1 V из Мест2; 2) Суб1 V в Мест4 ; 3)Суб1 V из Мест2 в Мест4

Быть убитым Обь1 V Суб5

Быть уничтоженным Обь1 V Суб5

Быть атакованным Обь1 V Суб5

Быть захваченным Обь1 V Суб5

Для классификации получившегося списка требуются какие-либо признаки, объединяющие единицы в группы. С одной стороны, основываясь на грамматическом значении позиции субъекта, последние 4 конструкции, обозначающие пассивность действия, выделяются в отдельную группу.

С другой стороны, при исследовании полученных актантных моделей была выделена особенность, свойственная большинству конструкций и послужившая основой для некоторого их упорядочивания. Этой особенностью послужили 3 варианта семантической реализации позиции «объект». В этой позиции могли находиться лексемы и их сочетания, обозначающие:

- 1) людей и боевые единицы
- 2) территорию
- 3) как людей и боевые единицы, так и территорию

Таким образом, получилась следующая классификация:

- 1) Обь – люди и боевые единицы: *быть убитым, быть уничтоженным.*
- 2) Обь – территория: *нанести удар/авиаудар, контролировать, занять*
- 3) Обь – как люди и боевые единицы, так и территория: *атаковать, потерять, захватить, быть захваченным, быть атакованным.*

Из-за отсутствия позиции «объект» в некоторых актантных моделях, в отдельный ряд встали такие глаголы, как *пострадать, погибнуть, отступить.*

Также в Приложении 1 в таблицах 1 и 2 представлены опорные глаголы, объединенные в синонимические ряды – синсеты – по значению. Кроме того, в таблице приведены примеры употребления, а также лексическое значение, основанное на четырехтомном словаре русского языка, каждого из опорных глаголов с указанием номера значения.

### **2.3 Фреймовое описание конфликтных ситуаций**

Представленный прототип направлен на описание ситуаций военных конфликтов, но также может быть применен и к другим типам и формам конфликтов. Так или иначе, с представленных позиций возможно описание как военных операций в Сирии и конфликта в Фергюсоне, так и Куликовской битвы.

В нашем случае конфликтной ситуацией считается ситуация, которая обладает тремя составляющими конфликта: биполярностью, наличием сторон конфликта, наличием действий.

Основная мысль, заложенная в данном описании, заключается в существовании конфликтных макроситуаций и микроситуаций, которые можно описать при помощи фреймов.

Микроситуацией является мельчайшее действие, в нашем случае – конфликтного характера. Микроситуация опирается на текст, она вербоцентрична. Пример: в предложении *«Вооруженные боевики атаковали*

*КПП йеменской армии в Лахдже и убили 12 солдат в засаде у Лодара»* мы видим две микроситуации: 1) *боевики атаковали КПП в Лахдже* и 2) *боевики убили 12 солдат в засаде у Лодара*.

Макроситуацией же считается сумма связанных между собой микроситуаций, описывающая конфликт в конкретном тексте. Как правило, при наличии нескольких микроситуаций связь между ними представляют собой логические операции (импликация или конъюнкция). Конфликтная ситуация также опирается как на текст, так и на контекст. Пример: предложение «*Вооруженные боевики атаковали КПП йеменской армии в Лахдже и убили 12 солдат в засаде у Лодара*» описывает макроситуацию, в которой две микроситуации связаны между собой конъюнкцией, а сама конфликтная ситуация относится к афгано-пакистанскому пограничному конфликту.

### **2.3.1 Структура фрейма микроситуации**

Фрейм, описывающий конфликт на уровне микроситуации, состоит из следующих элементов:

Действие. Действие выражено глаголом или глагольным сочетанием, чаще всего имеющим лексическое значение, напрямую связанное с конфликтом. Например: *атаковать, обороняться, погибнуть, захватить* и др. Наиболее частотные глаголы, обозначающие действия конфликтной направленности, были отобраны на основе корпуса, собранного на материале Википедии, в который вошло около 20 обширных статей, посвященных описанию вооруженных действий в 2000-2015 годах: Афгано-пакистанский пограничный конфликт, конфликт в Южной Осетии, конфликт на Украине, гражданская война в Либерии, Ливии и другие. Общий объем корпуса составил ~60 тысяч словоупотреблений. Глаголы были разделены по значению на ядро и периферию. Ядро составили такие глаголы и глагольные группы, как *атаковать, потерять, захватить, нанести удар/авиаудар,*

*контролировать, занять, пострадать, погибнуть, отступить, быть убитым, быть уничтоженным.*

Субъект. Агенс обозначает субъекта действия. Чаще всего выражено существительным или именной группой в именительном или творительном падеже. Обозначает человека, группу людей, государство и т.д. Например: *Командующего повстанцами Мисраты Рамадана Аль-Свейхи предали и убили **представители племён Бени-Валида**. 7 боевиков были убиты **повстанцами**.*

Объект. Пациенс или объект в широком смысле слова. Пациенсом может выступать как человек, группа людей или страна, пострадавшая от каких либо действий, так и такие материальные объекты, как здания, города, техника и т.д. Так или иначе, пациенс – это то, на что направлено действие. Чаще всего выражено существительным или именной группой в винительном падеже. Пример: *Члены племени Сиаам атаковали **амазигов Налута**. 2 марта революционеры атаковали **нефтяной порт Брега**.*

Место. Факультативный элемент. Местом на уровне микроситуации является лексическое выражение локализации конфликта. Чаще всего выражается существительным или именной группой с предлогом в родительном, винительном или предложеном падеже. Пример: ***Из города Бенгази** начался поток беженцев **на восток**.*

Таким образом из текста можно вычленить базовые элементы конфликтной ситуации, опираясь на основные составляющие конфликтной ситуации: биполярность, наличие конфликтующих сторон и действия. Биполярность и конфликтующие стороны выражены отношением агенс-пациенс, действия – опорными глаголами с лексическим значением конфликта.

Примеры фреймового описания на уровне микроситуации:

*Фадл-Аллах Харун, командующий войсками ПНС в этом районе, заявил, что 12 сентября лоялисты атаковали гавань Рас-Лануфа.*

Элемент конфликта	Лексическое	Грамматическое
-------------------	-------------	----------------

	<b>выражение</b>	<b>значение</b>
Действие	<i>атаковали</i>	Гл., прош. вр.
Субъект	<i>лоялисты</i>	Сущ., им. пад.
Объект	<i>гавань Рас-Лануфа</i>	Им. группа, вин.пад.
Место	—	—

Из представленной таблицы видно, что опорным глаголом, обозначающим конфликтное действие, является *атаковали*. В данном случае он выражен глаголом в прошедшем времени.

Субъектом действия выступает существительное *лоялисты*, стоящее в именительном падеже, а объектом действия – *гавань Рас-Лануфа*, выраженная именной группой в винительном падеже. При этом слот «место» остается незаполненным.

*19 сентября появились сообщения, что войска ПНС, подошедшие с восточной стороны к Себхе, захватили аэропорт, крепость и район Манишия в южной части Себхи.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
Действие	<i>захватили</i>	Гл., прош. вр.
Субъект	<i>войска ПНС</i>	Им. группа, им. пад.
Объект-1	<i>аэропорт</i>	Сущ., вин.пад.
Объект-2	<i>крепость</i>	Сущ., вин. пад.
Объект-3	<i>район Манишия</i>	Им. группа, вин. пад.
Место	<i>в южной части Себхи</i>	Им.группа, предл. пад.

Из представленного фрейма видно, что опорным глаголом в данном случае выступает *захватили*. Субъектом действия являются *войска ПНС*, выраженные именной группой, стоящей в именительном падеже. Объектов в приведенном примере 3: *аэропорт* (выраженный существительным в винительном падеже), *крепость* (выраженный существительным в винительном), а также *район Манишия* (выраженный именной группой в

винительном падеже). Слот «место» занимает именная группа в южной части Себхи, стоящая в предложном падеже.

В Приложении 2 представленной работы вынесены примеры фреймов микроситуаций для каждого выделенного нами опорного глагола.

### 2.3.2 Структура фрейма макроситуации

Фрейм, описывающий конфликт на уровне макроситуации состоит из семи элементов, которые можно разбить на три группы:

- Общая характеристика
- Микроситуация (-и)
- Характеристика действия

Фрейм, описывающий макроситуацию, охватывает большое количество контекстуальной информации, не отраженной напрямую в каком-либо конкретном предложении, но воспроизводимой при помощи текста в целом.

#### **Общая характеристика**

1 . Точка зрения обозначает указание на то, с точки зрения какого участника КС ведется описание. Зависит от источника информации и контекста. Исходя из определения конфликтной ситуации, участники конфликтной ситуации являются противоборствующими сторонами, а значит справедливо их именовать как Агент и Контрагент. Таким образом, конфликтная ситуация в контексте может рассматриваться как с точки зрения агента, так и с точки зрения контрагента, так как зачастую в тексте описание конфликтной ситуации лишено объективности.

2 . Аспект конфликтной ситуации. Будет справедливым выделить 4 аспекта конфликтной ситуации: *Политический, Военный, Экономический* и *Социальный*. Следует отметить, что сама конфликтная ситуация практически всегда затрагивает сразу несколько типов, так как они сильно взаимосвязаны между собой. С другой стороны, новостной текст, как правило, может отражать один из аспектов конфликтной ситуации. Например, конфликт в



Донецкой области можно рассматривать как со стороны экономического аспекта (например, новостной текст о повышении налогов из-за войны), так и со стороны социального аспекта (например, текст о доставке гуманитарной помощи жителям).

3. Участники конфликтной ситуации. В данном слоте указываются стороны, принимающие участие в конфликте.

### **Микроситуации.**

Значение этого слота ссылается на значение одноименного фрейма, описанного ранее.

### **Характеристика действия**

1. Связь между микроситуациями. Зачастую связь между событиями в конфликтной ситуации можно выразить двумя логическими операциями: *импликацией* и *конъюнкцией*. Импликация подразумевает причинно-следственную связь между событиями, конъюнкция указывает на её отсутствие. Следует отметить, что импликация может зависеть от точки зрения. При этом можно считать объективность истинной импликативностью, а субъективность – ложной импликативностью.

Импликативность конфликтного действия может иметь 3 возможных варианта развития события:

- Разрешение КС
- Порождение другого Действия
  - а) Последующее действие совершается той же стороной
  - б) Последующее действие совершается противоположной стороной
  - в) Последующее действие совершается третьей стороной
- Неразрешение КС без последующего Действия

2. Характеристика локализации. Будет справедливым выделить 3 возможных характеристики локализации конфликта: *местный*, *региональный* и *международный*. Следует отметить, что зачастую, как и аспекты конфликтной ситуации, характеристики локализации конфликта

соприкасаются между собой и состоят в гипо-гиперонимических отношениях, в то время как в тексте может быть отражена лишь одна характеристика (например, описание Сталинградской битвы в контексте Второй мировой войны). В то же время иногда характеристика зависит от точки зрения. Например, конфликт на Донбассе с точки зрения пророссийских СМИ описывается как региональный (то есть как «гражданская война»), в то время как Украина и Запад считают его международным (так как с их точки зрения российские войска вторглись на территорию суверенного государства).

3 . Время. Описание длительности конфликтной ситуации сильно зависит от конкретной точки зрения. Понятие *долгосрочных* и *краткосрочных* конфликтных ситуаций могут отобразить такую объективную составляющую, как время, но из-за разнообразия конфликтных ситуаций и возможностей их описания, а значит и субъективном восприятии времени встает вопрос о соотносимости по времени тех или иных конфликтов. Исходя из этого кажется разумным ввести понятие *время неопределенное*, обозначающее субъективное начало в описании длительности конфликта и основанное на теории неопределенных множеств. Таким образом *время неопределенное* есть не долгосрочное или краткосрочное, но воспринимаемое субъективно с какой-либо точки зрения. Кроме того, кажется, что объективности в вопросе времени быть и не может, поэтому оставим его открытым.

На практике фрейм макроситуации выглядит следующим образом:

*В середине декабря федеральные войска начали артиллерийские обстрелы пригородов Грозного, а 19 декабря был нанесен первый бомбовый удар по центру города.*

<b>Общая характеристика</b>	
Точка зрения	РФ
Аспект КС	Военно-политический
Участники КС	Федеральные войска, войска ЧРИ
<b>Микроситуации</b>	

<u>Микроситуация-1</u>		
Элемент конфликта	Лексическое выражение	Грамматическое значение
Действие	<i>начали артиллерийские обстрелы</i>	Глагольная группа, прош. вр.
Субъект	<i>федеральные войска</i>	Именная группа, им. пад.
Объект	<i>пригородов Грозного</i>	Именная группа, род. пад.
<u>Микроситуация-2</u>		
Действие	<i>был нанесен удар</i>	Глагольная группа, прош. вр.
Субъект	—	—
Объект	<i>по центру города</i>	Именная группа, дат. пад.
<b>Характеристика действия</b>		
Связь	Имплицативность, последующее действие совершается той же стороной	
Характеристика локализации	Местный	
Время	Время неопределенное	

В приведенном примере можно обнаружить следующее. Точка зрения в данном предложении является контекстовой информацией, в данном случае её занимает *Российская Федерация*, так как в статье, из которой был взят этот отрывок, события описываются с проправительственных позиций. Аспект конфликтной ситуации также является по большей части контекстовой информацией, в нашем случае это пересечение военного и политического аспектов. Также и участники конфликтной ситуации находятся частично за пределами предложения: из предложения мы знаем одну сторону конфликта – *федеральные войска*, из текста, откуда был взят пример, мы знаем, что речь идет о Первой чеченской кампании.

В предложении имеются два предиката, а значит мы имеем дело с двумя микроситуациями, причем, исходя из семантики глагольных сочетаний, они обе имеют явно выраженный конфликтный характер. В первой микроситуации действием является глагольная группа *начали*

*артиллерийские обстрелы, субъектом – федеральные войска, а объектом выступает именная группа пригороов Грозного.*

Во второй микроситуации действием выступает глагольная группа *был нанесен удар*, слот «субъект» остался незаполненным (так как при описании микроситуации мы опираемся на данный текст, а не на контекст), а объектом выступает именная группа *по центру города*.

При характеристике самого конфликтного действия указано, что связь представляет собой импликацию, так как действия между собой взаимосвязаны, причем вариант импликации, когда действия совершаются одной и той же стороной. Конфликт имеет местный локальный характер, а время указано как *неопределенное*.

## **2.4 Выводы**

В представленной главе был описан способ выявления конфликтных сценариев при помощи вербоцентрической концепции Теньера и рамок валентностей с одной стороны, а также способ представить конфликтную ситуацию как сценарий в терминах теории фреймов.

В главе был рассмотрен используемый материал, а также обоснованность его выбора.

На основе материала были выделены опорные глаголы (встречающиеся в корпусе более 15 раз и имеющие отношение к конфликту с точки зрения семантики). Они, в свою очередь, были рассмотрены в рамках теории валентности, то есть были выделены основные семантические роли актантов, сопутствующие этим глаголам с указанием их грамматического значения.

Основную часть главы заняло фреймовое описание конфликтной ситуации. Было предложено подразделение на микроситуации и макроситуации.

Микроситуацией считается мельчайшее действие, в нашем случае – конфликтного характера. Микроситуация опирается на текст, она

вербоцентрична. Пример: в предложении *«Вооруженные боевики атаковали КПП йеменской армии в Лахдже и убили 12 солдат в засаде у Лодара»* мы видим две микроситуации: 1) *боевики атаковали КПП в Лахдже* и 2) *боевики убили 12 солдат в засаде у Лодара*.

Макроситуацией же считается сумма связанных между собой микроситуаций, описывающая конфликт в конкретном тексте. Как правило, при наличии нескольких микроситуаций связь между ними представляют собой логические операции (импликация или конъюнкция). Конфликтная ситуация также опирается как на текст, так и на контекст. Пример: предложение *«Вооруженные боевики атаковали КПП йеменской армии в Лахдже и убили 12 солдат в засаде у Лодара»* описывает макроситуацию, в которой две микроситуации связаны между собой конъюнкцией, а сама конфликтная ситуация относится к афгано-пакистанскому пограничному конфликту.

Также в главе были представлены примеры фреймов, описывающих микро- и макроситуацию, а в приложение были вынесены синсеты глаголов, описанные в терминах RussNet, а также примеры фреймов микроситуаций на каждый из выделенных нами глаголов.

## **Заключение**

В представленной работе был исследован способ выявления конфликтных ситуаций в новостных текстах посредством теории валентностей, семантических ролей и фреймового описания.

В ходе исследования были решены следующие задачи:

1. Были описаны основные параметры лексико-семантического поля конфликтной ситуации, указаны особенности рамок валентности и семантических ролей аргументов при глаголах, а также указаны принципы построения фреймов.
2. Был собран корпус, отобраны наиболее частотные лексемы и синтагмы, принадлежащие семантическому полю конфликтной ситуации, а также создан фрейм, описывающий конфликтную ситуацию.

Итогом работы можно считать полученные нами фреймы конфликтных микроситуаций и макроситуаций, а также синсеты отобранных нами слов, описанных в терминах RussNet.

При анализе материала при помощи Томита-парсера были отобраны наиболее частотные глаголы (встретившиеся в тексте более 15 раз) и классифицированы на семантические «ядро» и «периферию». В «ядро» вошли такие глаголы, как *атаковать*, *захватить*, *пострадать*, *погибнуть* и другие, а также такие глагольные сочетания, как *нанести удар*, *быть убитым* и *быть уничтоженным*. Основанием для подобного выделения служили частотность в корпусе с одной стороны и практически полное отсутствие полисемии с другой, что означает высокую используемость этих глаголов при описании интересующих нас вооруженных конфликтов.

После классификации была проделана работа по построению их актантных моделей с указанием семантических ролей и грамматических значений. Глаголы, объединенные в синсеты с указанием примеров употребления, рамки валентности и лексического значения, которые можно найти в приложении, были предоставлены RussNet.

Кроме того, была проведена работа по формированию прототипов микроситуаций и макроситуаций в терминах фреймового описания.

Микроситуацией считалось мельчайшее действие конфликтного характера. Микроситуация опирается на текст, она вербоцентрична. Пример: в предложении *«Вооруженные боевики атаковали КПП йеменской армии в Лахдже и убили 12 солдат в засаде у Лодара»* мы видим две микроситуации: 1) *боевики атаковали КПП в Лахдже* и 2) *боевики убили 12 солдат в засаде у Лодара*.

Макроситуацией же считалась сумма связанных между собой микроситуаций, описывающая конфликт в конкретном тексте. Как правило, при наличии нескольких микроситуаций связь между ними представляют собой логические операции (импликация или конъюнкция). Конфликтная ситуация также опирается как на текст, так и на контекст. Пример: предложение *«Вооруженные боевики атаковали КПП йеменской армии в Лахдже и убили 12 солдат в засаде у Лодара»* описывает макроситуацию, в которой две микроситуации связаны между собой конъюнкцией, а сама

конфликтная ситуация относится к афгано-пакистанскому пограничному конфликту.

В работе были подробно разобраны структуры фреймов, а также приведены примеры фреймового описания.

Фреймовое описание ситуаций в терминах RussNet представляется достаточно свежим направлением. В представленной работе было произведено описание ситуаций, носящих исключительно конфликтный характер, поэтому перспективность данного исследования видится в описании ситуаций других типов, а также, возможно, нахождение объединяющих их элементов с последующей классификацией.

### **Список использованной литературы**

1. Anders Johannsen. Any-language frame-semantic parsing. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, 2015. P. 2062-2066.
2. Benítez P. Framing Terminology: A Process-Oriented Approach. Meta: Translators' Journal, 2005. Vol. 50, №4. P. 40-64.
3. Braslavski P., Ustalov D., Mukhin M. A Spinning Wheel for YARN: User Interface for a Crowdsourced Thesaurus // Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics. Gothenburg, Sweden, 2014. P. 101-104.
4. Das D. Frame-Semantic Parsing. Computational Linguistics, 2014. Vol. 40, #1. P. 9-56.



5. Fillmore C. Frame semantics and the nature of language // Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech, 1976. Vol. 280. P. 20-32.
6. M'arquez L., Carreras X. Semantic Role Labeling: An Introduction to the Special Issue. Computational Linguistics, 2008. Vol. 34, #2. P. 145-159.
7. O'Hara T. Exploiting Semantic Role Resources for Preposition Disambiguation. Computational Linguistics, 2008. Vol. 35, #2. P. 151-184.
8. Punyakanok V. The Importance of Syntactic Parsing and Inference in Semantic Role Labeling. Computational Linguistics, 2008. P. 257-287.
9. Азарова И.В. RussNet как база представления лексического блока в системах автоматизированного анализа текстов. СПб, 2015.
10. Азарова И.В. Выявление лексикализованных понятий в RussNet с использованием контекстной информации из корпуса текстов. [http://project.phil.spbu.ru/RussNet/papers/azarova\\_facultet\\_2004.doc](http://project.phil.spbu.ru/RussNet/papers/azarova_facultet_2004.doc) (Дата обращения: 15.05.2016)
11. Азарова И.В. Схемы управления в грамматике и рамки валентностей в RussNet. [http://project.phil.spbu.ru/RussNet/papers/azarova\\_facultet\\_2005.doc](http://project.phil.spbu.ru/RussNet/papers/azarova_facultet_2005.doc) (Дата обращения: 15.05.2016)
12. Азарова И.В., Митрофанова О.А., Синопальникова А.А. Разработка компьютерного тезауруса русского языка типа WordNet // Материалы конференции «Корпусная лингвистика и лингвистические базы данных» 5-7 марта 2002 г. СПб, 2002.
13. Азарова И.В., Синопальникова А.А., Смрж П. Представление устойчивых лексических сочетаний в компьютерном тезаурусе RussNet. <http://project.phil.spbu.ru/RussNet/papers/azarova-sinopalnikova-smrz.doc> (Дата обращения: 15.05.2016)
14. Азарова И.В., Синопальникова А.А., Яворская М.В. Принципы построения wordnet-тезауруса RussNet. // Компьютерная лингвистика и

- интеллектуальные технологии. Труды международной конференции Диалог'2004 (Верхневолжский, 2–7 июня 2004 г.) М., 2004. С. 542–547.
15. Апресян Ю.Д. Идеи и методы современной структурной лингвистики (краткий очерк). М.: Просвещение, 1966.
  16. Апресян Ю.Д. Избранные труды, том I. Лексическая семантика. М., 1995.
  17. Арутюнова Н.Д. Язык и мир человека. М.: Языки русской культуры, 1999.
  18. Ахманова О.С. Очерки по общей и русской лексикологии. М.: Государственное учебно-педагогическое издательство министерства просвещения РСФСР, 1957.
  19. Богуславский И.М. «Сфера действия лексических единиц», М.: Языки русской культуры, 1996.
  20. Бондарко А.В. Теория функциональной грамматики: Введение. Аспектуальность. Временная локализованность. Таксис. Л., 1987.
  21. Валгина Н.С. Активные процессы в современном русском языке. М.: Логос, 2003.
  22. Васильев Л.М. Современная лингвистическая семантика. М., 1990.
  23. Греймас А. Структурная семантика. М.: Академический проект, 2004.
  24. Добров А.В. Автоматическая рубрикация новостных сообщений средствами синтаксической семантики. СПб, 2014. Диссертация на соискание ученой степени к.ф.н. СПб, 2014.
  25. Засорина Л.Н. Введение в структурную лингвистику. М.: Высшая школа, 1974.
  26. Кацнельсон С.Д. К понятию типов валентности // Вопросы языкознания, 1987. №3. С. 20-32.
  27. Кобозева И.М. Лингвистическая семантика. М.: Эдиториал УСССР, 2000.
  28. Мартынюк Е.В. От валентности к фреймовой семантике // Вестник Кемеровского государственного университета, 2014, № 4 (60) Т. 2. С. 169-172.

29. Маслова В.А. Введение в когнитивную лингвистику. М.: Флинта, 2011.
30. Минский М. Фреймы для представления знаний. М., 1979.
31. Тарланов З.К. Методы и принципы лингвистического анализа. Петрозаводск: Издательство Петрозаводского Университета, 1995.
32. Теньер Л. Основы структурного синтаксиса. М.: Прогресс, 1988.
33. Тестелец Я.Г. Введение в общий синтаксис. М.: Российский Государственный Гуманитарный Университет, 2001.
34. Уфимцева А.А. Опыт изучения лексики как системы (на материале английского языка). М.: Изд-во Академии наук СССР, 1962.
35. Ушакова А.А. «Формализованное описание структуры лексико-семантического поля глаголов деятельности для компьютерного идеографического словаря типа WordNet (на материале русского языка). Диссертация на соискание ученой степени к.ф.н. СПб, 2007.
36. Фурер О. В. Фреймовый способ описания лексической семантики. Альманах современной науки и образования, № 8 (27). Самара, 2009. С. 178-180.
37. Чигашева М.А. Исследование терминологической лексики методом семантического поля // Вестник РУДН, №6. 2004.
38. Шаумян С.К. Структурная лингвистика. М.: Наука, 1965.
39. Щур Г.С. Теории поля в лингвистике. М.: Наука, 1974.

## Приложение

### Приложение 1

Таблица 1

Синсет	Значение	Примеры	Валентности
Захватить	4. Взять силой; овладеть.	Несколько сотен экстремистов захватили несколько городов в провинции Абьян.	Суб1 V Обь4
Занять	2. Овладеть какой-	По словам Абду	Суб1 V Обь4

	л. территорией, населенным пунктом и т. п.	Салема, революционеры заняли центр Зувары.	
Атаковать	1. Произвести (производить) атаку.	Войска Каддафи атаковали нефтяной комплекс возле Рас-Лануфа.	Суб1 V Обь4
Нанести удар	8. В сочетании с некоторыми существительным и означает: причинить, сделать то, что выражено существительным. <i>Нанести ножевую рану. Нанести урон.</i>	20 марта ВВС Франции, Великобритании и США утром нанесли удар по танковой колонне.	Суб1 V по Обь3

Таблица 2

Синсет	Значение	Примеры	Валентности
Погибнуть	Подвергнуться уничтожению, полному разрушению (обычно при катастрофе, бедствии и т. п.).	28 сентября Аль-Джазира сообщила, что в результате ракетного обстрела возле Бени-Валида погибли 11	Суб1 V в Кауз6 Кауз2

		бойцов Национальной армии	
Потерять	2. Оставаться без кого-, чего-л., переставать обладать кем-, чем-л.; лишаться, утрачивать.	Представитель ПНС заявил, что в бою они потеряли от 15 до 17 солдат.	Суб1 V Обь4
Быть убитым	УБИТЬ 1. Лишить жизни, умертвить.	По дороге Юнис и ещё два сопровождавших его офицера были убиты.	Обь1 V Суб5
Быть уничтоженным	УНИЧТОЖИТЬ 1. Прекратить существование кого-, чего-л.; истребить.	Атаке войск Каддафи подвергся аэропорт Триполи, в котором, как сообщают источники революционеров, ракетным огнём были уничтожены как минимум 4 самолёта.	Обь1 V Суб5
Пострадать	2. <i>от чего, из-за чего и без доп.</i> Терпеть ущерб,	Пять человек пострадали от взрыва в	Суб1 V от Кауз2

	урон.	Стамбуле	
Быть захваченным	<b>ЗАХВАТИТЬ</b> 4. Взять силой; овладеть.	Также повстанцами были захвачены районы Сабамия и Эль-Гиза.	Объ1 V Суб5
Быть атакованным	<b>АТАКОВАТЬ</b> 1. Произвести (производить) атаку.	В районе города З а в и я 32 - й бригадой спецназначения Х а м и с б ы л о атаковано подразделение английского спецназа SAS.	Объ1 V Суб5

## Приложение 2

Примеры фреймов микроситуаций

1. Атаковать

*2 марта революционеры атаковали нефтяной порт Брега.*

Элемент конфликта	Лексическое выражение	Грамматическое значение
Действие	<i>атаковали</i>	Гл., прош. вр.
Субъект	<i>революционеры</i>	Сущ., им. пад.
Объект	<i>Нефтяной порт Брега</i>	Им. группа, вин.пад.
Место	—	—

2. Потерять

*23 декабря власти Южного Судана потеряли контроль над городом Бенгиу.*

Элемент конфликта	Лексическое	Грамматическое
-------------------	-------------	----------------

	<b>выражение</b>	<b>значение</b>
Действие	<i>потеряли контроль</i>	Глагольная группа, прош. вр.
Субъект	<i>власти Южного Судана</i>	Именная группа, им. пад.
Объект	<i>над городом Бенгиу</i>	Именная группа, творительный падеж без предлога
Место	—	—

### 3. Захватить

*По данным Росбалта, российские миротворцы и подразделения захватили в зоне конфликта более 100 единиц бронетехники, в том числе 65 танков.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
Действие	<i>захватили</i>	Гл., прош. вр.
Субъект	<i>российские миротворцы</i>	Именная группа, им. пад.
Объект-1	<i>более 100 единиц бронетехники</i>	Именная группа, винительный падеж без предлога
Объект-2	<i>65 танков</i>	Именная группа, винительный падеж без предлога
Место	<i>в зоне конфликта</i>	Именная группа, предложный падеж с предлогом <i>в</i>

### 4. Нанести удар

*26 мая несколько вертолётчиков также нанесли удары по зданию завода «Точмаш».*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
--------------------------	------------------------------	--------------------------------



Действие	<i>нанесли удары</i>	Глагольная группа, прош. вр.
Субъект	<i>несколько вертолётов</i>	Именная группа, им. пад.
Объект	<i>по зданию завода «Точмаш».</i>	Именная группа, дательный падеж с предлогом
Место	—	—

### 5. Контролировать

*В тот момент они всё ещё контролировали большую часть Фаллуджи и Аль-Кармаха, а также части городов Хадита, Джурф-Аль-Сахар, Ана, Эль-Кайм, Абу-Грейб и несколько более мелких поселений мухафазы Анбар.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
Действие	<i>контролировали</i>	Гл., прош. вр.
Субъект	<i>они</i>	Местоимение, им. пад.
Объект-1	<i>большую часть Фаллуджи и Аль-Кармаха</i>	Именная группа, винительный падеж без предлога
Объект-2	<i>части городов Хадита, Джурф-Аль-Сахар, Ана, Эль-Кайм, Абу-Грейб</i>	Именная группа, винительный падеж без предлога
Объект-3	<i>несколько поселений мухафазы Анбар</i>	Именная группа, винительный падеж без предлога
Место	—	—

### 6. Занять

*Развивая наступление, они 2 мая подошли к Яфрану с севера и заняли город.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
--------------------------	------------------------------	--------------------------------

Действие	<i>заняли</i>	Гл., прош. вр.
Субъект	<i>они</i>	Местоимение, им. пад.
Объект	<i>город</i>	Сущ., винительный падеж без предлога
Место	—	—

## 7. Пострадать

*В ходе столкновений пострадали 38 человек.*

Элемент конфликта	Лексическое выражение	Грамматическое значение
Действие	<i>пострадали</i>	Гл., прош. вр.
Субъект	<i>38 человек</i>	Именная группа, им. пад.
Объект	—	—
Место	—	—

## 8. Погибнуть

*28 сентября Аль-Джазира сообщила, что из-за ракетного обстрела возле Бени-Валида погибли 11 бойцов Национальной армии, в том числе и один из старших командующих — Дау эль-Шахин эль-Джадак, на машину которого пришёлся удар.*

Элемент конфликта	Лексическое выражение	Грамматическое значение
Действие	<i>погибли</i>	Гл., прош. вр.
Субъект-1	<i>11 бойцов Национальной армии</i>	Именная группа, им. пад.
Субъект-2	<i>один из старших командующих Дау эль- Шахин эль-Джадак</i>	Именная группа, им. пад.
Объект	—	—
Место	<i>возле Бени-Валида</i>	Именная группа, родительный падеж с предлогом <i>возле</i>

### 9. Отступить

*Как сообщается, приблизительно 70 солдат Каддафи отступили в направлении Себхи.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
Действие	<i>отступили</i>	Гл., прош. вр.
Субъект	<i>70 солдат</i>	Именная группа, им. пад.
Объект	—	—
Место	<i>в направлении Себхи.</i>	Именная группа, винительный падеж с предлогом <i>в</i>

### 10. Быть убитым

*По дороге Юнис и ещё два сопровождавших его офицера были убиты.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
Действие	<i>были убиты</i>	Глагольная группа, прош. вр., пассивная конструкция
Субъект	—	—
Объект-1	<i>Юнис</i>	Сущ., им. пад.
Объект-2	<i>два офицера</i>	Именная группа, им. пад.
Место	<i>По дороге</i>	Сущ., дательный падеж с предлогом <i>по</i>

### 11. Быть уничтоженным

*В результате столкновения силовиками были уничтожены, по данным повстанцев, 4 танка и 1 БТР.*

<b>Элемент конфликта</b>	<b>Лексическое выражение</b>	<b>Грамматическое значение</b>
Действие	<i>были уничтожены</i>	Глагольная группа,

		прош. вр., пассивная конструкция
Субъект	<i>силовиками</i>	Именная группа, творительный падеж
Объект-1	<i>4 танка</i>	Именная группа, им. пад.
Объект-2	<i>1 БТР</i>	Именная группа, им. пад.
Место	—	—