

РОССИЙСКАЯ АКАДЕМИЯ НАУК
МИНИСТЕРСТВО НАУКИ И ТЕХНИЧЕСКОЙ ПОЛИТИКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ

На правах рукописи

ЛЯШЕВСКАЯ Ольга Николаевна

НЕСТАНДАРТНОЕ ЧИСЛОВОЕ ПОВЕДЕНИЕ
РУССКИХ СУЩЕСТВИТЕЛЬНЫХ

Специальность 05.13.17 - Теоретические основы информатики

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени
кандидата филологических наук

Москва 1999

Работа выполнена во Всероссийском институте научной и технической информации Российской АН и Министерства науки и технической политики РФ

Научный руководитель: доктор филологических наук
ПАДУЧЕВА Елена Викторовна

Официальные оппоненты:
доктор филологических наук
КРЫСИН Леонид Петрович,
кандидат филологических наук
ТАТЕВОСОВ Сергей Георгиевич

Ведущая организация: Институт проблем передачи информации
Российской АН

Защита состоится " ____ " _____ 199 ____ г. в ____ часов на заседании диссертационного совета Д.003.02.01 во Всероссийском институте научной и технической информации Российской АН и Министерства науки и технической политики РФ (125219, Москва, ул. Усиевича, 20а).

С диссертацией можно ознакомиться в библиотеке ВИНТИ.

Автореферат разослан " ____ " _____ 199 ____ г.

Ученый секретарь
диссертационного совета

М.А.Каменская

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Функционирование всемирной информационной сети интернет и других современных информационных систем делает весьма актуальной проблему представления и поиска информации о предметах реального мира, об их количестве, размере и других параметрах. Лингвистический компонент ИС, обслуживающий переработку и экспертизу больших массивов текстов и информационный поиск, а также интерфейс с пользователем на естественном языке, должен учитывать тот факт, что за поверхностным языковым выражением может стоять семантическое содержание, отличное от канона. В этой связи особый интерес вызывают семантические категории ЕЯ и, в частности, категория числа имен существительных. В русском языке форма числа существительных помимо передачи стандартного смысла:

единственное число (SG) – ‘один’

множественное число (PL) – ‘больше одного, несколько’ –

может передавать нестандартный смысл. Например, у существительного *транспорт* форма SG используется для выражения значения ‘совокупность’ (машин); у существительного *сани* форма PL передает оба смысла: ‘один’ ~ ‘несколько’ (предметов); форма PL *супруги* может, помимо прочего, выражать смысл ‘пара человек: супруг + супруга’; форма PL *масла* выражает значение ‘сорта’ (масла).

В грамматической литературе дефектные имена существительные имеют богатую традицию описания (А.А.Потебня, В.В.Виноградов, Д.И.Арбатский, А.А.Зализняк, Е.В.Падучева, И.А.Мельчук, А.Вежбицкая, А.И.Сумкина, Л.К.Чельцова, Е.В.Красильникова и др.) Однако обычно подобные случаи отклонения от стандарта задаются списками имен, а это явно недостаточно для задач лингвистического компонента ИС. Во-первых, даже самый полный список не обладает предсказательной силой, т. е. невозможно предвидеть, каково будет числовое поведение в тексте названия нового объекта. Во-вторых, списки позволяют выделить лишь самые грубые случаи – как правило, имена с неполной числовой парадигмой (sg.tt. *беднота*, pl.tt. *сани*) – но неудобны для фиксации особенностей употребления в том или ином контексте имен с полной парадигмой, прежде всего, в случае синонимии числовых форм (ср. *По краю / краям*

фонтана стояли статуи). Сложившаяся практика описания числового поведения в словарях – система помет вроде "ед. и мн.", "чаще ед." – также не позволяет построить необходимые правила.

Тем самым, обнаруживается необходимость разработки особой зоны числовых форм в статьях словаря предметной лексики. **Целью работы** являлось исследование семантики русской предметной лексики и выявление факторов, влияющих на ее числовое поведение.

Задачи исследования. Для достижения поставленной цели необходимо было решить ряд конкретных задач:

- на основании существующих описаний выделить таксономические классы имен с общими особенностями строения числовой парадигмы;
- семантически мотивировать наблюдаемые типы дефектности;
- оценить выделенные семантические признаки с точки зрения их необходимости и достаточности;
- выявить типы контекстов, влияющих на выбор числовой формы;
- проследить частные особенности числового поведения у слов каждого класса и выявить дополнительные признаки, объясняющие расхождение в числовом поведении;
- определить, какие типы информации, релевантной для построения правил числового оформления, должна содержать словарная статья предметного имени;
- разработать алгоритм определения значения числовой формы, исходя из информации, содержащейся в словарной статье данного имени и типа контекста, в котором употребляется имя.

Материал. В работе использовались материалы базы данных предметных имен "Лексикограф", разрабатываемой в ВИНТИ. Исследование числовых употреблений проводилось на основе выборки текстов научно-технического характера, а также художественной литературы. Наряду с этим в работе использованы сконструированные примеры. Ряд примеров заимствован из словарей русского языка (МАС, Словарь Ожегова и др.). Кроме того, учитывались приведенные в словарях толкования лексических единиц. Список введенных в рассмотрение имен существительных насчитывает порядка единиц.

Методы исследования. В основу работы были положены не только дескриптивный, но и объяснительный метод. Последний предполагает, что особенности оформления и употребления слов естественного языка могут быть мотивированы их семантикой. В работе был использован контекстуальный анализ, активно использовался языковой эксперимент, к рассмотрению был привлечен отрицательный языковой материал.

Следующие особенности работы определяют ее **научную новизну**:

- трактовка числа как лексически обусловленной грамматической категории, т. е. такой, значения которой определяются не только значением категориального признака, но и лексическим значением имени;
- построение эффективной модели числового поведения предметной лексики русского языка;
- разработка фрагмента словаря предметных имен, который может быть интегрирован в систему автоматического анализа в качестве модуля, позволяющего производить диагностику значений числовых форм в тексте.

Теоретическая значимость данного исследования состоит, во-первых, в анализе дефектности числовой парадигмы имени и в выявлении типов дефектности на множестве всех предметных имен. Введено понятие нестандартного числового поведения, позволяющее привлечь к рассмотрению имена с полной парадигмой, у которых, тем не менее, отмечается нетривиальное значение числовых форм в тексте.

Во-вторых, выявлены семантические группы и те семантические признаки, которые "отвечают" за тот или иной тип нестандартного числового поведения.

В-третьих, проведенные исследования позволяют по-новому взглянуть на грамматическую категорию числа имен существительных. Субстантивное число оказывается гораздо более лексикализированным, чем это традиционно представляется в лингвистических описаниях.

В-четвертых, обоснована необходимость значительного пополнения числовой зоны словарных статей – как в целях полнейшего описания языка, так и в целях автоматического анализа текста. В работе детализирован формат словарного представления релевантной с числовой точки зрения семантической информации.

В-пятых, сформулированы правила, предсказывающие числовое поведение предметных имен в русском языке, и на их основе построен алгоритм определения значений числовых форм в тексте.

Практическая ценность работы. Результаты диссертационного исследования могут найти применение при разработке лингвистического обеспечения современных информационных, информационно-поисковых и экспертных систем с интерфейсом на естественном языке. Полученные результаты призваны повысить точность поиска количественной информации в ИС.

Кроме того, настоящая работа может иметь применение в лексикографии русского языка – при создании словаря предметных имен, позволяющего прогнозировать, исходя из толкования, особенности употребления и интерпретации грамматических форм существительных в тексте, а также особенности сочетаемости лексем.

Методика проведенных исследований может быть распространена на анализ числового поведения предикатных имен существительных.

Апробация работы. Основные результаты исследований были отражены в выступлениях на международном семинаре "Диалог'96" (Пушино, 1996); конференции молодых преподавателей кафедры русского языка для иностранных учащихся филологического факультета МГУ (Москва, 1997); конференции "Язык и пространство" Института языкознания РАН (Москва - Дубна, 1997); на Первой традиционной школе по лингвистической типологии для студентов, аспирантов и молодых научных сотрудников (Москва, 1998).

Структура работы. Диссертация состоит из введения, 3-х глав, заключения, указателя рассматриваемых существительных, библиографии (около 150 работ) и приложения.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается выбор темы, определяются цели и задачи исследования, показана научная новизна, теоретическая и практическая ценность работы.

В первой главе диссертации "**Теоретические принципы описания числового поведения**" дается обзор литературы, ставится задача, излагаются основные принципы описания, уточняется терминология.

Ключевое слово метаязыка описания – **дефектность** парадигмы по числу. Различаются понятия формальной (морфологической) и функциональной (семантической) дефектности. При формальной дефектности¹ в парадигме существительного отсутствуют грамматические формы с окончаниями единственного или же множественного числа:

<i>singulare tantum</i>			<i>plurale tantum</i>		
	SG	PL		SG	PL
И.	бельё	–	И.	–	сани
Р.	белья	–	Р.	–	саней
Д.	белью	–	Д.	–	саням
В.	= И.	–	В.	–	= И.
Т.	бельём	–	Т.	–	санями
П.	белье	–	П.	–	санях

Семантическая дефектность² означает, что семантика данного имени 'X' не допускает сочетания смыслов 'один X' и 'несколько X'. Из вышеприведенных имен *белье* семантически дефектно, а у существительного *сани* оба значения (ср. *одни сани* ~ *несколько саней*) передаются омонимичными формами с окончаниями множественного числа (PL/PL)³. Существительные с семантически дефектной числовой парадигмой иначе называются **несчетными**. Классические примеры несчетных имен – это вещественные имена (*вода, песок, сливки*) и имена совокупностей (или собирательные имена в широком понимании этого термина: *аристократия, бельё, низы общества, руины*).

Свойство счетности имен вытекает из исчисляемости обозначаемых объектов – понятия, относящегося к сфере как лингвистики, так и когнитивной науки. Исчисляемость обусловлена не столько материальными свойствами объекта, сколько тем, как человек воспринимает этот объект – дискретно или не дискретно. Как было показано А. Вежбицкой,

¹ Формальное определение дефектности см.: А.А. Залязник, Русское именное словоизменение. М., 1967.

² Ср. понятие функциональной неполноты парадигмы глагола в работе: Ю.С. Маслов. Очерки по аспектологии. Л., 1984.

³ Здесь и далее используется формула употребления числовых форм имени со значением 'X', где в числителе сообщается, какая форма передает смысл 'один X', в знаменателе – какая форма обозначает 'несколько X'. Знак # обозначает, что семантика 'X' не сочетается со смыслами 'несколько' или 'один'.

дискретные объекты в специальных случаях могут восприниматься как континуум-субстанция (англ. *Put some **table** in the salad* – букв. "Положите немного **стола** в салат"). Однако в принципе языковая картина мира задает обычный способ восприятия объекта, на основании чего любому существительному может быть приписан признак COUNT / UNCOUNT. Если понимать семантику широко, то и сведения о материальных свойствах обозначаемого, и сведения о его восприятии следует включать в толкование имени, и следовательно, по Вежицкой, исчисляемость становится выводима из семантики слова.

В работе обсуждаются дополнительные аспекты понятия исчисляемости:

- потенциальная счетность, отмечаемая у уникальных имен типа *луна*;
- ситуативная несчетность, характеризующая названия объектов типа *горизонт*. Такие объекты определяются только в конкретной ситуации (в данном случае – это ситуация, когда наблюдатель видит местность вокруг себя, и *горизонтов* не может быть больше одного);
- счетность, не допускающая обозначения 'один X', которая наблюдается в переносных употреблениях существительных: *требуются новые рабочие **руки*** ('несколько человек'), *отряд в сто **сабель*** (т. е. человек, ср. [#]*одна сабля*).

Морфологическая дефектность парадигмы не влияет на счетность: существительные, не имеющие по формальным причинам форм PL, допускают употребление с числительными от 2 до 4, и не сочетаются только с "большими" числительными и словами *много*, *несколько*, требующими формы родительного множественного.

В результате, в русском языке выделяются следующие типы дефектности числовой парадигмы:

1) несклоняемые существительные;		} sg.tt.
2) SG / (PL) – потенциально счетные типа <i>луна</i> ;		
3) SG / # – несчетные типа <i>бельё</i> ;		
4) PL / PL – счетные типа <i>сани</i>		} pl.tt.
5) PL / # – несчетные типа <i>низы</i> ;		
6) # / PL – счетные типа (<i>рабочие</i>) <i>руки</i> .		

Отдельную теоретическую проблему составляет область определения дефектности парадигмы. В традиционной грамматике свойство быть *singulare* или *plurale tantum* приписывается имени существительному на

пространстве всех значений (т. е. слову как целому). Однако как указывалось еще в работах Л.К. Чельцовой и П.А. Соболевой, более корректно определять полноту парадигмы для каждого значения имени отдельно. Например, считать, что существительное *сапоги* в значении 'пара' имеет дефектную парадигму PL/PL, а в значении 'разрозненные предметы обуви' обладает полной парадигмой SG/PL. В ряде случаев и этого бывает недостаточно. В работе А.К. Поливановой было показано, что есть классы слов, которые, в принципе, имеют полную парадигму, но различаются числовым оформлением в ряде контекстов. Так, в контекстах, требующих родового употребления, у существительных *яблоко*, *помидор* выбирается форма PL, а у существительных *свекла*, *репа* – форма SG (*Любите ли вы **яблоки** / **помидоры**? ~ Любите ли вы **свеклу** / **репу**?*). Разница может выявляться не только при родовом, но и при конкретном употреблении. Например, формы SG и PL имени *арбуз* в принципе могут обозначать одну и ту же ситуацию (*У них вчера на десерт давали **арбузы** / **арбуз*** [в смысле 'некоторое количество']), а числовые формы имени *помидор* таким свойством не обладают (*Они вчера ели **помидоры** ≠ Они вчера ели **помидор***).

В то же время приписывание признака полноты парадигмы одному значению слова осложнено тем, что частные значения могут быть выделены по-разному, особенно у таких семантически богатых слов, как *место*, *земля*, *вода* и др. Очевидно, что приписывать в словаре характеристики парадигмы каждому употреблению слова – задача и вовсе нереальная. С учетом этого в настоящей работе было принято решение несколько укрупнить область определения дефектности парадигмы, усложнив при этом формат описания числового поведения.

За единицу описания принимается существительное во всех значениях, в которых обозначаемые объекты принадлежат одной денотативной области (например, числовое поведение слова *земля* в значениях 'планета', 'суша' и 'страна, государство' рассматривается в комплексе (пространственные значения); отдельно рассматривается числовое поведение слова *земля* в вещественном значении (*песок с **землей***). Общая формула числового поведения многозначных существительных представляется в виде набора формул SG/PL, SG/# и т. д., каждая из которых соотнесена со своим лексическим значением. В

случае необходимости общая формула сопровождается комментарием относительно частных особенностей контекстных употреблений форм числа.

Задача диссертационного исследования состоит в том, чтобы не просто приписать каждому существительному формулу парадигмы, но **предсказать** форму имени с данной семантикой в релевантном наборе контекстов.

Во второй главе "**Числовое поведение русских предметных существительных**" выделяются основные таксономические классы предметных имен: "отдельные объекты" (*девочка, стол*), "сложные объекты" (*брюки, очки*), "пары" (*руки, сапоги*), "наборы" (*клавиши, карты*), "множества" (*букет, лес*), "вещества и материалы" (*песок, линолеум*), "совокупности" (*белье, мусор, руины*). Общая стратегия числового оформления состоит в том, что имена отдельных объектов и множеств счетны и имеют полночисловую парадигму SG/PL. Имена пар, наборов и сложных объектов также счетны, но имеют формально дефектную парадигму PL/PL. Имена совокупностей и веществ-материалов несчетны и, как правило, имеют дефектную парадигму SG/#.

Таксономические классы предметной лексики могут пересекаться, и слова, относимые к двум классам, наследуют свойства обоих. Так, имя *булыжник* относится к классам "отдельный объект" и "вещество / материал" (*улица вымощена булыжниками / булыжником*); имя *шнурок* входит в классы "отдельный объект" и "сложный объект" (*завязать шнурок / шнурки на левом ботинке*). Названия парных частей тела наследуют числовое поведение класса "пары" и класса "части", релевантного для выбора числовых форм в дистрибутивном контексте (см. главу III).

Вторая глава состоит из двух частей. Вначале предлагается описание числового поведения наименований пар и наборов, сложных объектов, веществ и материалов, совокупностей. Затем рассматривается нестандартное числовое поведение имен, объединенных по признаку общей предметной области: это имена пространств, еды и растений, не получившие достаточного освещения в литературе по числу.

ПАРЫ и НАБОРЫ – это, в первую очередь, парные и вообще неединственные части – части тела (*руки, глаза, пальцы*) и части артефактов

(*губы* плоскогубцев, *линзы* очков, *клавиши* пишущей машинки), – а также предметы, функционально связанные с этими частями (*сапоги*, *носки*, *шины*). Наряду с "парным PL" или "PL набора" у существительных – наименований пар и наборов есть омонимичная форма PL, обозначающая тривиальное множество разрозненных объектов: *пожать **руки** всем присутствующим*, *ремонтировать **сапоги*** и т. д. (в этом значении слова имеют парадигму SG/PL и относятся к классу "отдельный объект").

Особый случай представляют названия людей. Необходимое условие импликации значения формы PL 'пара' – наличие в семантике имен указания на ситуацию, в которой участвуют два контрагента, например, *дуэлянты* (2 участника дуэли), *собеседники* (пара, участвующая в разговоре). Пара – это не обязательно "два X, обозначенных единственным числом", ср. имена *супруги*, *молодожены*, *родители* и др. Имя в форме SG может иметь валентность контрагента, которая насыщается внутрилексемно в форме PL в зависимости от контекста (ср. *собеседники удалились* – обычно пара участвующих в разговоре и *мои **собеседники** удалились*). Интересно, что слова типа *собеседник* и *партнер* в форме PL могут совмещать значение парности (набора) и тривиального множества: в принципе, каждую сторону могут представлять несколько людей.

В связи с представлением данных имен в словаре встает вопрос, можно ли из толкования вывести таксономический класс существительного – ведь схожие с перечисленными выше парами и наборами объекты могут представляться в языке и как множества (ср. *супруги* ~ *чета*). Ответ на этот вопрос неоднозначен. Пожалуй, только семантический признак "быть частью" более или менее однозначно соотносит имя с классами пар и наборов.

При сопоставлении таких имен, как *инициалы* (пара) и *вензель* (множество) можно проследить некоторое семантическое наращение у имени множества (\approx 'сочетание... в виде вязи'), но оно не настолько единообразно и регулярно, чтобы его можно было формализовать для всех имен множеств. Укажем только, что обычно у имен множеств идея "состоять из" уходит на задний план, причем иногда настолько, что количество денотатов становится нерелевантным, так, *пригоршня* – это одна или две ладони, сложенные ковшом, а имя пары различает число

ладоней: *горсть / горсти*; *занавес* – это одно полотнище или две половины, а имя *занавеска / занавески* строго различает число полотнищ.

Некоторые наименования пар и наборов обнаруживают частные особенности употреблений числовых форм в тексте. У имен *ключи*, *поручни*, *шторы* и нек. др. форма PL не различает значений ‘один X’ – ‘пара (набор) X-ов’: *Я забыл **ключи** от квартиры* [может быть и связка ключей, и один ключ]; *Он перевесился через **поручни***; *На окне висят тяжелые французские **шторы*** [могут состоять и из одного полотнища]. У некоторых названий парных частей (в особенности частей тела) форма SG сообщает информацию о паре частей: *У него маленькая **рука***; *Туфли ей как раз по **ноге***; *Рубашка с коротким **рукавом***; *Тут воды по **колену***.

Объективно сложно устроенные, состоящие из нескольких частей предметы могут концептуализироваться и как СЛОЖНЫЙ ОБЪЕКТ (PL/PL), и как обычный; ср. синонимы *плечики* (PL/PL) и *вешалка* (SG/PL), где форма PL подчеркивает составность объекта. Таким образом, принадлежность к таксономическому классу в достаточной степени лексикализована, хотя в русском языке отмечается последовательное числовое оформление pl.tt. в группах наименований брюк и трусов, а также у инструментов типа *клещи*. На основании этого выделяются семантические признаки ‘быть раздвоенным’, ‘иметь заметные и функционально значимые части’, которые способствуют концептуализации объекта как сложного.

Различного рода веревки (ремни, цепи и т. п.) также представляются как сложный предмет (пара), если их концы находятся рядом (видны как пара) и выполняют какую-либо значимую функцию, например, привязаны к рукам и не дают им двигаться (*цепи*), находятся в руках у прыгающего (*прыгалки*) или правящего лошадью (*поводья*), завязаны в узел (*шнурки*). Следует оговориться, что такая концептуализация все же лексикализована: ср. имена *прыгалки* и *скакалка*.

Кроме того, таксономический класс "сложный объект" должен быть формально приписан в словаре тем существительным, у которых подобное представление объекта ничем не мотивировано или объясняется этимологически (*мостики*, *гусли* и т. п.). В ряде случаев числовая форма существительного объясняется не способом концептуализации, а тем, что она строится по аналогии с формами близких слов. Так, имена-гипонимы

слова *сани*: *салазки, дровни, розвальни и др.* – выстраиваются в ряд с парадигмой PL/PL; ср. также имена *часы* и *ходики*.

Некоторые имена могут представлять один и тот же объект и как простой, и как сложносоставный, от чего проистекает так называемая синонимия числовых форм. В работе исследуются различия в употреблении форм SG и PL слов *мозг / мозги, дверь / двери, сеть / сети* и др.

На названиях ВЕЩЕСТВ И МАТЕРИАЛОВ можно продемонстрировать, как внутри таксономического класса числовое оформление зависит от действия множества факторов и разнится от слова к слову. Кратко обозначим самые характерные особенности.

Вещества и материалы как правило обозначаются неисчисляемыми существительными с парадигмой SG/#. Такие существительные в форме родительного падежа SG обычно сочетаются с числительными, но обозначают не само вещество, а его порцию (*два пива*). Формы множественного числа этих существительных обозначают (а) виды, типы, сорта веществ/материалов: *чаи, лаки, чугуны, нефти, миткали*⁴; (б) скопления больших количеств вещества/материала или бесконечные пространства, заполненные ими: *воды, льды, снега, пески*; (в) вещество, сходное с веществом, обозначаемым существительным sg. tt., или его особый подвид: *пары, газы, воды, грязи, осадки*.

В качестве исключения вещества и материалы могут обозначаться неисчисляемыми существительными *pluralia tantum*: *сливки, чернила, белила, духи*. Среди них выделяется компактная группа имен, отличающихся от прочих как значением ‘отбросы или остатки вещества в результате какого-нибудь процесса’, так и морфологическим строением: *высевки, выгребки, отбросы, подонки, опилки*.

Некоторые обозначения веществ - *pluralia tantum* соотносятся с одноименными полночисловыми или *singularia tantum* существительными в другом значении: душистое вещество – его запах (*благоухания – благоухание, благовония – благовоние*); вещество как результат процесса – сам процесс (*испарения – испарение*); разнородные примеры других типов отношений: *капли – капля, нечистоты – нечистота* и проч.

⁴ Следует учитывать, что помимо семантических факторов, в языке действуют морфологические и акцентуационные ограничения на выбор числовых форм: так, существительное *пиво* не имеет формы PL, хотя она была бы приемлема для выражения смысла ‘сорта пива’.

Вещества и материалы могут также обозначаться неисчисляемыми существительными, имеющими синонимичные формы SG и PL, одна из которых употребляется реже, будучи либо маркирована оценочно или стилистически (*харчи* и простореч. *харч*; *слюна* и разг. *слюни*), либо употребляясь в устойчивых словосочетаниях (*чаи гонять*, *неприкосновенный запас*).

Числовая форма вещественных имен зависит не только от семантики обозначаемого, но и от формального устройства значений слова: если слово многозначно, то его числовое оформление в данном значении может быть связано с числовым оформлением имени в мотивирующем значении. Так, для обозначения "веществ особого вида" (*газы, воды, соли*) в языке используются формы PL, не задействованные у данных слов в их основном значении "вещества". Помимо этого, числовое поведение может зависеть от выбранной говорящим точки зрения на мир – наивной или научной: например, одна и та же субстанция может называться именем в разных формах (*водяной пар* ~ *водяные пары*).

Имена веществ дифференцируются от обозначений "классов веществ" (*щелочи, антисептики, ингредиенты*) – имен, которые обладают полной парадигмой и в форме SG могут обозначать само вещество.

Весьма сходно с числовым поведением вещественных имен поведение названий С О В О К У П Н О С Т Е Й. Этот класс образуют существительные с суффиксами собирательности (*крестьянство, агентура, лапник, листва, белье* (-j-), *ветошь* (-0-) и т. п.), непроизводные имена (*одежда, мебель, промтовары* и др.) и переносные обозначения (*серебро* 'изделия из серебра', *спальня* 'мебель для спальни', *импорт* и т. п.). Имена совокупностей по-особому обозначают предметы внешнего мира. Это хорошо видно в их противопоставлении именам множеств (*лес, букет* и т. п.). Последние обозначают несколько лиц или предметов, объединенных какой-либо функцией, и характеризуются конечным количеством элементов и единством их местонахождения. Совокупность отражает представление о категориальном членении внешнего мира. Такая категория содержит открытое количество элементов и задается уникальным набором признаков (ср. *сословие* ~ *купечество*).

Так же, как и вещества, совокупности подходят под понятие "mass terms" (О. Есперсен), т. е. любое количество X можно назвать X, ср. не-считные *вода, мебель* и *стол* ≠ *столы*. В своей массе имена совокупнос-

тей являются sg. tt., и лишь считанное их количество относится к pl.tt. Мы не можем предложить никакой семантической мотивировки различию в числовом оформлении, ср. имена, относящиеся к одной денотативной области: *пожитки* и *утварь*, *низы* и *беднота*. Основная причина употребления в форме PL – морфологическая или формально-семантическая (перенос значения особого вида).

Как правило, имена совокупностей соотносимы с названиями отдельных предметов и могут быть заменены соответствующими формами PL: *Бюрократия не любит инициативных людей* = *Бюрократы не любят инициативных людей*. Однако, как доказывается в работе, отношения между этими именами квазисинонимичны; ср. *листья вишни* и **листва вишни*; *Американские профессора удивились, какое у нас современное оборудование* и *Американская профессура удивилась, какое у нас современное оборудование*.

В тексте наименования совокупностей могут иметь конкретную референцию (*Молодежь пустилась в пляс*; *Вся мебель в зале покрылась пылью*) и даже обозначать один предмет (*В транспорте нечем дышать*), хотя последнее свойство имеют далеко не все имена совокупностей.

Имена типа *инструмент*, *платье* имеют два значения: ‘отдельный предмет’ и ‘совокупность’. В отличие от имен типа *транспорт*, семантическая деривация сопряжена здесь с существенным сдвигом значения и изменением сочетаемостных свойств. *Инструмент 2* (совокупность) уже по значению, чем *инструмент 1*, а именно, обозначает только технические инструменты (ср. **пропал весь музыкальный инструмент*). *Платье 1* – это отдельный предмет одежды, который носят женщины, *платье 2* обозначает не обобщенную категорию *платьев 1*, а совокупность надетой на людей одежды, без различия пола владельца (ср. *советники в партикулярном платье*).

Кроме того, во второй главе прослеживается числовое поведение трех групп имен, объединенных по тематическому признаку: пространства, еда и растения. Показано, что состав парадигмы зависит от принадлежности имени к одному из вышеуказанных классов, а также от некоторых других признаков, которые будут упомянуты далее.

Обозначения ПРОСТРАНСТВА образуют четыре больших группы: точки, линии, территории и трехмерные пространства.

Точки (*центр, конец*) и линии (*меридиан, дорога*) исчисляемы, за исключением особого вида объектов, координаты которых определяются исключительно в конкретной ситуации наблюдения (например, *створ, горизонт*); территории (двухмерные пространства) могут быть как счетными (*поля, страны*), так и несчетными (*целина, чужбина*); трехмерные пространства (*комната, бассейн*) исчисляемы.

Особое внимание в разделе уделено объектам со сложным числовым поведением. Так, у имен *граница, край, берег* выделяются разные употребления в зависимости от геометрического вида линии. *Граница* и *край*, мыслимые как отрезок (*польско-белорусская граница, край листа*), имеют стандартное числовое оформление. Если же они мыслятся как замкнутая линия вокруг некоторого пространства, то ведут себя и как обозначения единого предмета, и как обозначения сложного предмета (*Граница Москвы протянулась на 109 км = Границы Москвы протянулись на 109 км*). Аналогично существительное *берег* оформляется по-разному в зависимости от объекта, границы которого он определяет (*берег ~ берега реки* и *берег = берега моря*). Употребление числовых форм имен *граница, край* чувствительно к таким семантическим характеристикам контекста, как точечная / неточечная локализация ситуации, ср. *Польша не намерена открывать границу = границы* vs. *Мы подъезжаем к границе / *границам с Польшей* [описывается точечная локализация субъекта]; *постучать по краю / краям стакана* vs. *щелкнуть по краю / *краям стакана* [однократное действие – точечная локализация]. Следовательно, у вышеуказанных имен взаимозамена форм допускается далеко не во всяком контексте, т. к. семантические характеристики имени могут входить в противоречие с семантикой контекста.

Числовое поведение существительного *след* отражает взаимодействие в его семантике двух концептов – отпечатка, который оставляет на поверхности ступня, и пути, который оставляет за собой движущийся объект. Таким образом, контекстная синонимия числовых форм мотивируется представлением объекта в виде множества точек или же единым объектом (о цепочке следов можно сказать и так: *На контрольной полосе был виден след сапога* – и так: *были видны следы сапога*). Концепты "от-

печаток" и "путь" в имени *след* переплетаются так тесно, что форма PL употребляется тогда, когда путь нельзя представить как последовательность отпечатков (ср. *Следы самолета терялись в безоблачном небе*).

У ряда названий трехмерных пространств также наблюдается известный эффект синонимии числовых форм, связанный с двойственностью восприятия пространства. Сочетание нескольких пространств может образовывать сложный объект: *внутренние **комнаты** дома, чердаки, подвалы* и др.).

Кроме того, в русском языке сохранились следы того, что в XIX веке пространственные обозначения тяготели к форме PL: *антресоли, покои, хоромы* и т. п. У подобных имен дефектность парадигмы в настоящее время не имеет семантической мотивации.

В заключении раздела рассматриваются имена, у которых значение пространства является производным. Во-первых, это метафорические обозначения пространства (*пяточок, лоскут земли, окно в облаках*), у которых сохраняются числовые характеристики имени в основном значении. Во-вторых, это отглагольные обозначения пространства и места, счетность которых зависит от счетности мотивирующих предикатов (*спуск – спуски*). В-третьих, это пространства, покрытые веществом (*пески, снега, воды, почвы*), пространства, покрытые злаками (*овсы, ржи*) и пространства, образуемые объектами (*горы, острова, огороды, облака, тропики; в **стенах** университета, стоять в **дверях***) – у этих имен пространственное значение закреплено за формой PL. В принципе, в контексте локативных предлогов любое предметное имя может обозначать пространство или место. Однако в данном случае объект, скрываемый за существительным в форме PL, – это пространство особого вида, что доказывают новые нетривиальные сочетаемостные свойства производного имени.

Круг имен существительных – наименований Е Д Ы весьма неоднороден по своему составу, так как выделяется по функциональному, а не по денотативному признаку. Сюда попадают названия как жидкостей (т. е. классических веществ – *сок, бульон*), так и таких сущностей, которые близки к прототипу предмета (например, *яблоко* или *арбуз*, твердые и обладающие четкой формой). Поэтому вполне оправданно, что

наименования еды различаются по исчисляемости, а значит, и по числовому поведению.

Исчисляемость неожиданных продуктов зависит, в основном, от четырех семантических признаков⁵: формы, размера, способа употребления и гомогенности. Отсутствие в толковании представления о форме объекта делает несчетными такие имена, как *мясо* и *шоколад*. Признак размера отделяет средние, исчисляемые объекты (*огурцы*, *яблоки*, по размеру соответствующие ладони) от маленьких, неисчисляемых (*фасоль*, *горох*) и больших, также неисчисляемых (*капуста*, *салат*). Способ употребления важен для продуктов среднего размера и противопоставляет *огурцы* и *яблоки морковке* и *свекле*, которые могут исчисляться, но все же тяготеют к классу неисчисляемых. Этот же признак регулирует распределение числовых форм у существительного *перец*, ср. *есть маринованный **перец*** [в резаном виде] и *фаршированные **перцы*** [плоды целиком]. Наконец, признак размера объясняет языковую логику, стоящую за неисчисляемостью больших объектов: их надо резать на мелкие части.

Последний признак – гомогенность – вытекает из свойства человека не различать на глаз отдельных частиц в массе и релевантен для мелких объектов; например, он предопределяет осмысление *винограда* как совокупности плодов – в отличие от исчисляемых *маслин*; совокупное осмысление *изюма* – в отличие от *орехов* и т. д. Надо сказать, что гомогенность слабо вытекает из реальных характеристик частиц и скорее может быть использован как семантический ярлык при формулировке правил. Впрочем, признак гомогенности позволяет провести параллель с исчисляемостью других вещественных объектов, ср. триаду *идти по **гравию*** (UNCOUNT) ~ *по **плитам*** (COUNT) ~ *по **булыжникам*** (COUNT) / *по **булыжнику*** (UNCOUNT).

Названия готовых блюд представляют замечательный пример лексикализации числовой формы, которая связана с привычной ситуацией потребления продукта. Так, говорят: *Я ем **котлеты** / **пельмени** / **блинчики*** [‘неопределенное количество’, PL], но *Я ем **шницель*** [SG]. Если *котлеты*, *фрикадельки* и *блинчики* обычно едят в количестве нескольких штук, то *шницель* – это, как правило, большой кусок мяса, который один

⁵ Выделены А. Вежбицкой: Wierzbicka A. The semantics of grammar. Amsterdam, 1988. Ch. X.

лежит на тарелке. В том же контексте употребление формы SG имен *плов* и *гуляш* говорит о том, что эти блюда концептуализуются в языке не как множество кусочков мяса, а как субстанция – рис или подливка, в которых пребывают эти кусочки.

Наименования жидкостей и сыпучих веществ допускают регулярную полисемию "вещество" – "сорта вещества", сопровождающуюся дополнением парадигмы имени формами множественного числа (*чаи, вина*). Готовые блюда жидко-сыпучего типа, кроме того, отличает очень интересный тип полисемии, когда форма PL обозначает 'блюда, которые готовят в разное время' (*По воскресеньям мама варит бульоны*).

Имена РАСТЕНИЙ по признаку счетности целесообразно разделить на четыре класса: деревья, кустарники, цветы и травы. Счетность деревьев и цветов и несчетность кустарников и трав выглядит вполне логично. Действительно, деревья и цветы представляют более выделенные в поле зрения человека объекты, нежели кусты и травы, первые – в силу своих размеров, наличия одного крупного ствола, вторые – в силу больших ярких цветов и своей роли в бытовой культуре.

Существительные *слива, вишня, черешня* образуют переходный класс между деревьями и кустарниками, так как в той или иной мере обладают признаками обоих классов (ср. куст *вишни, черешни, ?сливы; два ряда сливы/слив*). Имена *айва, хурма* имеют только формы SG (дефектность их парадигмы, по-видимому, объясняется слабой освоенностью этих существительных в русском языке), поэтому, несмотря на то, что эти существительные сочетаются с числительными от 1 до 4, одна и та же форма SG у них обозначает как одно растение, так и их неисчисляемое множество. Еще два заимствованных названия – *алыча, инжир* – вообще не допускают единичной интерпретации, хотя неисчисляемое существительное *инжир* является синонимом исчисляемого *смоковница-смоковницы*.

Вышеназванные существительные обладают еще одной интересной особенностью: их плоды (в отличие от плодов других деревьев) либо вообще не поддаются исчислению (*рябина, черемуха, инжир*), либо в некоторых употреблениях имеют значение неисчисляемой совокупности (COUNT *сливы, две вишни*, но UNCOUNT *покупать вишню, варенье из сливы*).

Заметим также, что у кустарников (существительные sg.tt.) плоды (ягоды) тоже являются неисчисляемыми объектами: *виноград, барбарис, смородина, земляника*. Такая последовательная аналогия между недискретной концептуализацией плодов и недискретной концептуализацией самого растения получает языковое выражение именно в числовом оформлении.

Цветы выделяются как класс дискретно вычлняемых человеком растений, прежде всего, по характерному внешнему виду (обычно это растение, состоящее из одного крупного и яркого цветка на ножкестебле). Существительное *ромашка* употребляется для обозначения двух видов достаточно похожих растений: лекарственная ромашка представляет собой кустик с мелкими цветками, а полевая ромашка является классическим представителем класса цветов. Напротив, у других растений есть два варианта наименования, один из которых – исчисляемое, а другой – неисчисляемое существительное, ср.: *подсолнух* и *подсолнечник, примула* и *первоцвет*.

Тем не менее, класс цветов включает в себя также некоторые растения с мелкими цветками: *незабудка, лютик, ландыш, подснежник, флокс*. По-видимому, более важным фактором, чем размеры цветка, является выполнение этими растениями эстетической функции (украшения жилища человека и т. п.).

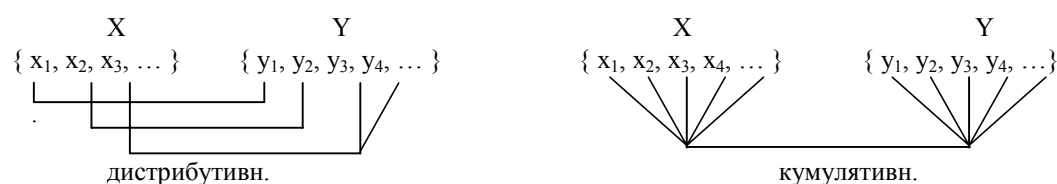
Существительное *герань* имеет две стратегии числового поведения: для обозначения исчисляемого (класс цветы) и неисчисляемого (класс кусты) названия комнатного растения, ср.: *на окне растут герани* – *растет герань*, однако конструкция *горшки с* + *Твор.* требует формы SG этого существительного, хотя, напротив, требует форму PL у названий *фикус, кактус (горшки с геранью ~ горшки с кактусами)*.

Пожалуй, наиболее необычный представитель класса цветов – *сирень* (sg.tt.). Среди названий цветов это существительное уникально по числовому поведению. В отличие от других цветов сирень растет на ветках больших кустов, и ее цветы представляют соцветие мелких цветков (ср. сочетания *цветы сирени* и *цветки сирени*). Форма ед. числа обозначает не отдельный цветок-соцветие, но только неисчисляемое множество.

Дополняет вторую главу экскурс "Исчисляемость и производность: числовое оформление производных существительных", в котором исследуется феномен неисчисляемости существительных с диминутивными суффиксами (*письмецо, чаёк*), других суффиксальных производных с оценочным значением (*парнюга, лесничко*), а также семантических производных (*аудитория* "люди в аудитории" и др.).

В третьей главе "Числовое поведение предметных имен в дистрибутивном контексте" описывается особенное употребление числовых форм, свойственное существительным некоторых классов.

Если в предложении идет речь о некоторой ситуации Р с двумя участниками X и Y, каждый из которых является множеством объектов, то между именами этих множеств устанавливаются дистрибутивные или кумулятивные отношения.



Дистрибутивность предполагает, что каждому члену x_1, x_2, \dots множества X соответствует один или несколько членов y_1, y_2, \dots множества Y.

В работе М.М. Бурас⁶ выделяются следующие признаки, предопределяющие выбор числовых форм:

	форма X	форма Y
1) X выступает с квантификаторами <i>каждый, всякий, любой</i> : <i>К каждому приемнику прилагается <u>трансформатор</u>.</i> <i>К каждому приемнику прилагаются <u>запасные антенны</u>.</i>	SG	SG - если только один объект PL - если мн-во
2) X и Y связаны одно-однозначным внеситуативным (постоянным) соответствием (обычно человек - часть тела (у каждого человека единственная)): <i>Казаки приподняли уже совсем понурившиеся было <u>головы</u>, и многие одобрительно кивнули <u>головой</u> (Гоголь).</i>	PL	SG=PL
3) X и Y связаны неодно-однозначным внеситуативным соответствием: <i>У остальных очков сломана <u>дужка</u>.</i> <i>У остальных очков сломаны <u>дужки</u>.</i>	PL	SG - если только один объект PL - если все

⁶ Бурас М.М. Выбор числового оформления русских существительных в дистрибутивном употреблении // ИРЯ АН СССР. Проблемная группа по экспериментальной и прикладной лингвистике. Предварительные публикации. Вып. 168. М., 1985.

4) X и Y связаны только в конкретной речевой ситуации: <i>Пожарные, надев <u>шлемы</u>, приступили к тушению огня</i> – одно-однозначн. соотв.; <i>Пожарные, размотав <u>брандспойты</u>, приступили к тушению огня</i> – неодно-однозначн. соотв.	PL	PL
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----	----

Таким образом, решающим семантическим критерием числового оформления (см. правила 2-4) является признак "постоянное vs. ситуативное соответствие между X и Y", который можно вычислить, исходя из лексического значения имен X и Y: если имеет место постоянное соответствие, то имя Y может быть употреблено в форме SG (но обозначать при этом реальное множество!), если же имеет место ситуативное соответствие, допускается только форма PL.

В области предметной лексики признак "постоянное соответствие" реализуется у имен класса "части". В частности, части тела входят в постоянное соотношение с целым (человеком), причем важна дихотомия "единственная ~ неединственная (парная) часть". Удивительно, но соотношения типа "отец - сын", заложенные в семантике имен родства, должны в соответствии с данным алгоритмом трактоваться как ситуативные (*В легких экипажах подъезжали к подъезду **девочки**_X и их матери_Y*).

Проведенный контекстуальный показал, что данные правила нуждаются в некотором уточнении.

Во-первых, данные правила следует распространить на те случаи, когда множество объектов Y названо именем совокупности или иным именем с дефектной парадигмой: форма PL имени Y должна быть сменена на SG и наоборот (ср. *Под ногами у **толпы** были рассыпаны яркие шарики*)

Во-вторых, следует учитывать, что в дистрибутивном употреблении наследуются нестандартные черты числового поведения, отмечаемые у отдельных лексем; ср. выбор числовой формы SG у имени *рукав* в контексте *рубашка с коротким **рукавом** ~ рубашки с коротким **рукавом***.

В-третьих, бывает важна синтаксическая структура предложения, если она делает связь между элементами X и Y опосредованной, ср. ***Продавцы**_X за **рыбным прилавком** как **окаменели** со своими **ножами**_Z в **руках**_Y* (М. Булгаков). [У каждого продавца задействована одна рука, тем не менее требуется форма PL, согласуемая с формой PL имени Z]

В-четвертых, некоторые сочетания и синтаксические конструкции оставляют выбор числовых форм без альтернативы. Например, в русском языке по возможности избегаются сочетания *правый / левый Y* во множественном числе. Ср. предложение: *У студентов уже затекла правая **рука**, а лектор все продолжал диктовать домашнее задание*. Равнозначность форм SG и PL здесь сомнительна: ??*У студентов уже затекли правые **руки**...*

Имена частей тела с локативными предлогами дают идиоматичные сочетания *за спиной, за плечами, под носом*, которые не трансформируются в дистрибутивном контексте (*За **спиной** у солдат оставался Сталинград – За **спиной** / за **спинами** у солдат висели винтовки*)

Наконец, в ряде конструкций, например, в конструкции *X с Y-ом* у формы SG отмечается употребление, близкое к родовому: Y может обозначать видовую особенность объектов X: *сигареты с **фильтром***.

В **приложении** приводится описание программы распознавания типа дистрибутивных отношений в тексте, а также исходный код программы. Программа написана на языке Visual C++, являющемся инструментом разработки многооконных приложений в среде Windows. Информационным ресурсом программы является семантический словарь предметных имен, построенный на основе базы данных системы "Лексикограф". Фрагмент словаря также приводится в приложении.

На основании проведенных исследований можно сделать следующие **краткие выводы**:

1. Детальный анализ употребления числовых форм в контексте показал, что круг имен с нестандартным числовым поведением значительно шире, чем это полагалось ранее.
2. Правила числового поведения в существенной степени предопределяются принадлежностью слова к таксономическому классу.
3. Принадлежность слова к таксономическому классу по большей части выводима из семантики. Например, классы "пары" и "наборы" соотносятся с семантическими признаками 'быть частью' и 'привычная ситуация функционирования предполагает участие пары или набора однородных объектов'.

4. Полнота числовой парадигмы существительного зависит от исчисляемости обозначаемого объекта. Исчисляемость в свою очередь выводима из семантики слова.

5. Числовая форма, выбираемая для выражения смысла 'один X', может зависеть:

а) от способа номинации объекта (номинация через составляющие элементы у сложных объектов; перенос значения по внешнему подобию, по схожести функций и т.п.);

б) от числового оформления синонимов или аналогов данного слова (выбор формы "по аналогии" вследствие принципа системоцентричности языка);

в) от морфологической структуры слова (формальные ограничения);

6. Язык может использовать незадействованные по прагматическим причинам числовые формы для выражения принципиально другого смысла (*вода ~ воды*).

7. Некоторые объекты внешнего мира допускают двоякую концептуализацию, в связи с чем называющие их существительные получают числовые характеристики двух таксономических классов (простой объект и сложный), – явление, известное как синонимия числовых форм или колебание числового поведения.

8. Контекстуально зависимые правила употребления числовых форм должны учитывать такие семантические признаки контекста, как точечная / неточечная локализация ситуации, дистрибутивные отношения и др., а также семантику синтаксических конструкций и сочетаний.

Основные положения диссертации отражены в следующих публикациях автора.

1. Число и пространство: формат описания числового поведения русских существительных // Нариньяни А.С. (ред.), Труды Международного семинара Диалог'96 по компьютерной лингвистике и ее приложениям, Москва, 1996, с.147-148.
2. Группа предметных имен-совокупностей и категоризация предметного мира // Пospelов Д.А. (ред.), КИИ-96. Тезисы докладов Конференции по искусственному интеллекту, Москва, 1996.
3. Исчисляемость и морфологическая производность (проблема грамматического оформления производных существительных) // Нариньяни А.С. (ред.), Труды Международного семинара Диалог'98 по компьютерной лингвистике и ее приложениям, т. 1, Казань, 1998, с.89-94.

4. Числовое оформление предметных существительных в дистрибутивном контексте // Научная и техническая информация. Сер. 2. Информационные процессы и системы, 1999, № 11 (в печати).
5. [Рец. на кн.:] L. Janda. A Geography of Case Semantics: The Czech Dative and the Russian instrumental. B.-N.Y., 1993 // Вопросы языкознания, 1998, № 1, с. 151-156.