

Т. М. Николаева

ОПЫТ АЛГОРИТМИЧЕСКОЙ МОРФОЛОГИИ РУССКОГО ЯЗЫКА

(К проблеме построения «грамматики распознавания»)

§ 1. В настоящее время можно говорить о двух принципиально различных типах описания языка — о «статическом» и «динамическом». Первое описание, дескриптивная грамматика, представляет собой перечень всех явлений, специфичных для данного языка и образующих систему этого языка. К такому типу относится большинство грамматик.

Понятию цели, возникшему в последние годы в языкознании и все более и более становящемуся осознанным при всяком лингвистическом описании, во многом обязано возникновение нового типа грамматик — «динамического типа». Эти грамматики описывают язык в его движении — от языковой модели к тексту и от текста к языковой модели. Исследование первого пути, «порождение текста», вызывает широкий интерес среди лингвистов, создающих различные типы конструктивной грамматики (работы Н. Хомского, отдельные работы советских языковедов)¹.

Второй тип «динамической грамматики» — описание движения от текста к модели — можно назвать «грамматикой распознавания».

Обе последних грамматики являются алгоритмическими грамматиками в том понимании алгоритма, при котором для некоторого класса элементов при заданных условиях при помощи однозначно определенной последовательности операций получается запись решения².

Настоящая работа посвящена попытке описания языка (русского) в его движении от текста к модели, проблеме формального распознавания грамматических категорий по данным текста, т. е. построению некоторой грамматики «анализа» для русского языка, как будет показано далее, не совпадающей с грамматикой порождения — грамматикой синтеза. Задача такой грамматики — описание языка с «точки зрения декодирования»³.

Методы построения грамматики анализа отличаются от обычных методов описания текста, предлагаемых дескриптивной лингвистикой. Известные методы дескриптивного анализа текстов (Э. Харрис и др.) предполагают в качестве цели извлечение модели неопisanного языка

¹ См.: «Тезисы докладов на конференции по обработке информации, машинному переводу и автоматическому чтению текста». М., 1961.

² См.: А. Н. Колмогоров и В. А. Успенский. К определению алгоритма. «Успехи мат. наук», т. XIII, вып. 4, 1958.

³ Об описании языка с точки зрения «кодирования» и «декодирования» в целях машинного перевода см.: V. N. Y n g v e. Sentence-for-sentence translation. «Mechanical translation», v. 2, № 2, 1955.

по данным текста и предлагают соответствующую методику этого извлечения.

При построении грамматики анализа статическая модель языка считается заданной, внимание же переносится на «количество шагов», необходимое для того, чтобы каждая единица текста могла быть переписана в виде совокупности определенных грамматических признаков.

Как подобный «грамматический анализ» текста производится человеком, в настоящее время не исследовано.

§ 2. Автором предлагается классификация различных словоформ по их формальной «распознаваемости» в тексте. Поскольку речь идет о русском языке, для которого характерно широкое присоединение флексий к большому числу знаменательных слов, значительное место в описании будет занимать характеристика русских флексий как грамматических показателей.

В целом работа состоит из следующих разделов: 1) характеристика русских флексий по способу распознаваемости их грамматического значения; 2) описание того, каким способом выводятся по разным данным морфологические категории в русском языке.

§ 3. В русском языке все флексии, присоединяемые к изменяемым словам, можно разделить на два класса: а) флексии, формально выраженные, б) флексии нулевые.

Иными словами, формы изменяемых слов могут быть с флексиями и без флексий.

Для русского языка (письменной нормы) нами выделено 67 ненулевых флексий для всех частей речи ⁴.

Основной задачей анализа считалось исследование того, как определить отношение единицы текста к минимальным единицам грамматики данного языка, или, как их называют, граммемам ⁵. Для русского языка выделяются следующие грамматические единицы, распознаваемые в тексте.

Для существительного: число (единственное и множественное) и падеж (в принятой системе шести падежей), для глагола — время (настоящее, прошедшее или будущее), лицо (1-е, 2-е или 3-е), число (единственное или множественное), наклонение (изъявительное, сослагательное или повелительное), залог (действительный или страдательный), глагольная форма (или, по терминологии А. И. Смирницкого, репрезентация) — личная форма, инфинитив, причастие или деепричастие. Для прилагательного приняты характерные граммемы рода, числа, падежа и степени сравнения ⁶. Анализу подвергались лишь те единицы текста, в которых грамматические единицы выражены морфологическим путем. Неморфологический способ выражения граммем ⁷ сознательно оставлен в стороне.

Опыт подобной классификации русских флексий по их распознаваемости недавно предложен М. Пакаком, построившим следующую таблицу русских окончаний ⁸.

⁴ В данном случае под флексиями в целях удобства понимались и формообразующие суффиксы. Список всех выделенных автором флексий (окончаний) приводится в табл. 1. В ряде случаев деление на флексию и основу произвольно. Например, слово *ручей* может иметь флексии: *-й*, *-я*, *-ю*, тогда слово будет иметь две основы: *руче-*, *ручь-*. В другом случае можно считать флексиями *-ей*, *-ья*, тогда слово будет основным (*руч-*). В данной работе был избран первый путь.

⁵ K. Pike. Grammatic theory. «General linguistics», v. II, № 2, 1957; K. Pike. On tagmèmes, néés gramèmes. «International Journal of American Linguistics», v. XXIV, № 4, 1958.

⁶ Другие части речи автор не анализировал.

⁷ Согласно терминологии И. А. Мельчука, предложившего исчерпывающую классификацию способов выражения языковых значений.

⁸ M. P a c a k. Rieviev of noun, adjectives and verb suffixes. Seminar of Georgetown University, 1958.

Флексия	Существительное	Прилагательное	Глагол
Однбукв. многозначн.			ø
Однбукв. однозначн.	ø		
Двухбукв. многозначн.			
Двухбукв. однозначн.			
Трехбукв. многозначн.	ø		
Трехбукв. однозначн.			

Знак ø показывает, что для данной части речи таких окончаний нет, пустые клетки в его таблице заполняются соответствующими окончаниями! Однако такая классификация только констатирует наличие многозначных и однозначных флексий в русском языке и ничего не говорит о степени многозначности каждой флексии и о различии в количестве шагов, требуемых для ее распознавания.

§ 4. По степени распознаваемости грамматических категорий, показателями которых они являются, все русские флексии делятся на четыре типа.

1. Категория распознается сразу, без каких бы то ни было сведений об основе (например, слово с флексией *-ью* может быть только творительным падежом единственного числа существительного III склонения).

2. Для распознавания морфологических категорий необходимо не только знать флексию, но и иметь сведения о части речи (например, если о слове с флексией *-а* известно, что это прилагательное, то можно сказать, не требуя никаких дополнительных сведений, что это единственное число, краткая форма, женский род).

3. Для распознавания категории необходимо не только знание части речи, но и более подробные сведения о слове (например, род, одушевленность, тип словоизменения — для существительного⁹; вид, спряжение — для глагола; тип присоединения флексий — для прилагательного и т. п.)¹⁰.

4. Для распознаваемости формы необходимы не только сведения о флексии и о слове, но и анализ контекста. Например, для распознавания грамматической характеристики слова *мыши* необходим анализ структуры того предложения, в которое входит это слово.

Итак, для распознавания различных форм необходимы три вида знаний: а) сведения о флексии, б) сведения о флексии + сведения об основе, в) сведения о флексии + сведения об основе + сведения о контексте.

Составим таблицу, характеризующую все флексии по отношению к выделенным типам (анализу подвергались флексии письменного русского языка).

⁹ Более подробно, чем принятые три типа склонения.

¹⁰ Речь идет о всей совокупности морфологических признаков. Например, по флексии *-ого* можно указать, что это слово не женского рода, единственного числа (если речь идет о прилагательном), но о падеже этого слова по флексии делать выводы нельзя. Поэтому данная флексия считается распознаваемой лишь по контексту.

№	Оконча- ние	Тип				№	Окончание	Тип			
		1	2	3	4			1	2	3	4
1	-а		+	+	+	30	-ло				+
2	-е	+				31	-ое				+
3	-е		+		+	32	-ов			+	+
4	-и		+	+	+	33	-ой		+	+	+
5	-й		+	+	+	34	-ом	+			
6	-л				+	35	-от				+
7	-м			+		36	-ст			+	
8	-о		+		+	37	-ти	+			
9	-у		+	+		38	-ут			+	
10	-ы		+	+	+	39	-ую	+			
11	-ь		+	+	+	40	-чь	+			
12	-ю		+	+		41	-ши	+			
13	-я		+	+	+	42	-ые				+
14	-ам	+				43	-ый				+
15	-ат			+		44	-ым		+		+
16	-ах	+				45	-ых				+
17	-ая	+				46	-ью	+			
18	-ее				+	47	-юю	+			
19	-ев			+	+	48	-ют			+	
20	-ей		+		+	49	-ям	+			
21	-ем		+	+		50	-ят			+	
22	-ет			+		51	-ях	+			
23	-ие				+	52	-яя	+			
24	-ий				+	53	-ами	+			
25	-им		+		+	54	-вши	+			
26	-ит			+		55	-его				+
27	-их				+	56	-ему	+			
28	-ла				+	57	-ете-			+	
29	-ли				+	58	-ешь			+	

№	Окончание	Тип				№	Окончание	Тип			
		1	2	3	4			1	2	3	4
59	-ими	+			+	64	-ому	+			
60	-йте			+	+	65	-ыми	+			
61	-ишь			+		66	-ьте	+			
62	-йте	+				67	-ями	+			
63	-ого				+						

* Окончания расположены по количеству букв, а внутри этого деления — по алфавиту.

Как видно из этой таблицы, не все окончания распределяются по намеченным четырем типам, а некоторые из них представляют собой смешанные типы (плюсы в нескольких графах). Например, окончание *-и* имеет плюсы во второй, третьей и четвертой графе. Это значит, что при знании того, что данное слово — глагол или прилагательное (т. е. знании части речи — тип 2) слово получает нужную грамматическую характеристику. Если это глагол, то слово получает признак повелительного наклонения, если прилагательное — то это краткая форма, множественное число. Если же это слово существительное, то необходимы более подробные сведения о слове (тип 3). Если это слово I склонения, мужского рода и одушевленное, то данная форма есть именительный падеж множественного числа. Но для полного анализа недостаточно и этих сведений. Если это слово III склонения, мы не сможем определить его форму без контекстного анализа (тип 4).

В общем виде комбинации из этих четырех типов в русском языке распределяются следующим образом:

Комбинация	Тип			
	1	2	3	4
1	+			
2		+		
3			+	
4				+
5		+	+	+
6		+		+
7			+	+
8		+	+	

§ 5. Покажем, как распределяются по этим комбинациям все флексии русского языка (перечислим все 8 групп).

Г р у п п а 1.

№	Окончание	Распознаваемая грамматическая характеристика
1	-ть	Неопределенная форма глагола
2	-ти	То же
3	-чь	»
4	-в	Глагол, деепричастие прош. вр.
5	вши	То же
6	-ши	»
7	-йте	Глагол, повел. накл., мн. ч.
8	-ьте	То же
9	-ами	Сущ., тв. п., мн. ч.
10	-ями	То же
11	-ом	Сущ., тв. п., ед. ч.
12	-ью	То же
13	-ам	Сущ., дат. п., мн. ч.
14	-ям	То же
15	-ах	Сущ., предл. п., мн. ч.
16	-ях	То же
17	-ая	Прил., ж. р., ед. ч., им. п.
18	-ья	То же
19	-ому	Прил., м. р., ед. ч., дат. п.
20	-ему	То же
21	-ую	Прил., ж. р., ед. ч., вин. п.
22	-юю	То же
23	-ыми	Прил., мн. ч., тв. п.
24	-ими	То же

Г р у п п а 2. В чистом виде в русском языке не представлена.

Г р у п п а 3 (нужны сведения о слове):

№	Окончание	При каком условии	Что означает
1	-м	Сов. вид	1-е лицо, ед. ч., изъявит. накл., буд. вр.
		Несов. вид	» » » » » » наст. вр.
2	-ишь	Сов. вид	2-е лицо, ед. ч., изъявит. накл., буд. вр.
		Несов. вид	» » » » » » наст. вр.
3	-ешь	Сов. вид	» » » » » » буд. вр.
		Несов. вид	» » » » » » наст. вр.
4	-шь	Сов. вид	» » » » » » буд. вр.
		Несов. вид	» » » » » » наст. вр.
5	-ет	Сов. вид	3-е лицо, ед. ч., изъявит. накл., буд. вр.
		Несов. вид	» » » » » » наст. вр.
6	-ит	Сов. вид	» » » » » » буд. вр.
		Несов. вид	» » » » » » наст. вр.
7	-ст	Сов. вид	» » » » » » буд. вр.
		Несов. вид	» » » » » » наст. вр.
8	-ете	Сов. вид	2-е лицо, мн. ч., изъявит. накл., буд. вр.
		Несов. вид	» » » » » » наст. вр.
9	-ут	Сов. вид	3-е лицо, мн. ч., изъявит. накл., буд. вр.
		Несов. вид	» » » » » » наст. вр.
10	-ют	Сов. вид	» » » » » »
		Несов. вид	» » » » » »
11	-ат	Сов. вид	3-е лицо, мн. ч., изъявит. накл., буд. вр.
		Несов. вид	» » » » » » наст. вр.
12	-ят	Сов. вид	» » » » » » буд. вр.
		Несов. вид	» » » » » » наст. вр.

Г р у п п а 4 (нужны сведения о контексте):

№	Окончание	Что может означать
1	-ый	Им. или вин. п. ед. ч. м. р. прил.
2	-ий	То же
3	-ые	Им. или вин. п. мн. ч. прил.
4	-ие	То же
5	-ых	Род., вин. или предл. п. мн. ч. прил.
6	-ой	Им. или вин. п. ед. ч. м. р. прил.
7	-ое	Им. или вин. п. ед. ч. ср. р. прил.
8	-ее	Им. или вин. п. ед. ч. ср. р. прил., сравн. ст.
9	-ого	Род. или вин. п. ед. ч. м. р. или вин. п. ед. ч. ср. р. прил.
10	-его	То же
11	-их	Род., вин. или предл. п. мн. ч. прил.
12	-ла	Прош. вр. изъявит. накл. или сослаг. накл.
13	-ло	То же
14	-л	»
15	-ли	»

Г р у п п а 5 (нужны сведения о части речи + сведения о слове + анализ контекста). К этой группе относятся следующие окончания: 1) -ь, 2) -й, 3) -и, 4) -ы, 5) -а, 6) -я, 7) -е.

Покажем для каждой флексии, какие категории она может означать, по следующей общей модели:

Окончание	При каком условии	Что означает		
-ь	Глагол	Повелительное наклонение		
	Прилагательное	Краткая форма, единственное число, мужской род		
	Существительное	Женский род	III скл.	Им. или вин. п. ед. ч.
			I скл.	Одуш. Род. или вин. п. мн. ч. Неодуш. Род. п. мн. ч.
	Мужской род	Одуш. Неодуш.	Им. п. ед. ч. Им. или вин. п. ед. ч.	
-й	Глагол	Повелительное наклонение		
	Существительное	Женский род	Неодуш. Одущ.	Род. или вин. п. мн. ч. Род. п. мн. ч.
		Средний род	Род. п. мн. ч.	
	Мужской род	Неодуш. Одущ.	Им. или вин. п. ед. ч. Им. п. ед. ч.	

Окончание	При каком условии	Что означает			
-и	Глагол	Повелительное наклонение			
	Прилагательное	Краткая форма, множественное число			
	Существительное	Женский род	I скл.	Род. п. ед. ч. или им. п. мн. ч.	
			III скл. и на -и ^я	Род., дат., предл. п. ед. ч., им. или вин. п. мн. ч.	
		Средний род	Предл. п. ед. ч.		
Мужской род		На -и ^й Одуш. Неодуш.	Предл. п. ед. ч. или им. п. мн. ч. Им. п. мн. ч. Им. или вин. п. мн. ч.		
-а -я	Глагол	Деепричастие несовершенного вида			
	Прилагательное	Краткая форма, единственное число, женский род			
	Существительное	Женский род	Им. п. ед. ч.		
		Средний род	Род. п. ед. ч., им. или вин. п. мн. ч.		
		Одуш.	на -а *	Род. или вин. п. ед. ч., им. п. мн. ч.	
			не на -а	Род. или вин. п. ед. ч.	
		Мужской род	на -а	Род. п. ед. ч., им. или вин. п. мн. ч.	
не на -а			Род. п. ед. ч.		

* Этот признак получают существительные мужского рода I склонения, имеющие окончание -а в именительном падеже множественного числа.

Окончание	При каком условии	Что означает		
-е	Прилагательное	Наречие, краткая форма или сравнительная степень		
	Существительное	Женский род	Дат. или предл. п. ед. ч.	
		Средний род	на -е	Им., вин. или предл. п. ед. ч.
-ы	Прилагательное	Множественное число, краткая форма		
		Женский род	Одуш.	Род. п. ед. ч. или им. п. мн. ч.
	Неодуш.		Род. п. ед. ч., им. или вин. п. мн. ч.	
	Существительное	Мужской род	Одуш.	Им. п. мн. ч.
			Неодуш.	Им. или вин. п. мн. ч.

Г р у п п а 6. Сведения о части речи + сведения о контексте. К этой группе относятся следующие окончания: 1) *-ой*, 2) *-им*, 3) *-о*, 4) *-ей*, 5) *-ым*.

Выписываем окончания этой группы, анализируя их по тому же образцу, что и окончания группы 5:

Окончание	При каком условии	Что означает		
-ей	Прилагательное	Им. или вин. п. ед. ч. м. р., род., дат., тв. или пр. п. ед. ч. ж. р.		
		I скл.	Тв. п. ед. ч.	
	Существительное	III скл.	Неодуш.	Род. п. мн. ч.
			Одуш.	Род. или вин. п. мн. ч.
-ой	Существительное	Тв. п. ед. ч.		
	Прилагательное	Им. или вин. п. ед. ч. м. р. Род., дат., тв. или пр. п. ед. ч. ж. р.		

Окончание	При каком условии	Что означает
-им	Глагол	1-е лицо мн. ч.
	Прилагательное	Дат. п. мн. ч. или тв. ед. ч. м. или ср. р.
-о	Прилагательное	Наречие, краткая форма ср. рода или безлично-предикативное слово
	Существительное	Им. или вин. п. ед. ч. ср. рода
-ым	Существительное	Тв. п. ед. ч. м. р. или дат. п. мн. ч.
	Прилагательное	То же

Г р у п п а 7 (нужны сведения о слове + сведения о контексте)¹¹.
В эту группу входят следующие окончания: 1) *-ите*, 2) *-ов (-ев)*.

Расписываем окончания по той же модели:

Окончание	При каком условии	Что означает
<i>-ите</i>	II спряжение I спряжение	2-е лицо ед. ч. или пов. накл. Пов. накл.
<i>-ов (-ев)</i>	Одуш. Неодуш.	Род. или вин. п. мн. ч. Род. п. мн. ч.

Г р у п п а 8 (сведения о части речи + сведения о слове). В эту группу входят окончания: 1) *-у (-ю)*, 2) *-ем*.

Окончание	При каком условии	Что означает	
<i>-у (-ю)</i>	Глагол	1-е лицо ед. ч.	
	Существительное	Женский род	Вин. п. ед. ч.
		Мужской род	Дат. п. ед. ч.

¹¹ Часть речи, к которой принадлежит основа, может для данной группы определяться по самому окончанию (*-ите* — только глагол, *-ов* — только существительное).

Окончание	При каком условии	Что означает	
-ем	Существительное	Тв. п. ед. ч.	
	Прилагательное	Предл. п. ед. ч. м. р.	
	Глагол	Сов. вид	Буд. вр., 1-е лицо мн. ч.
		Несов. вид	Наст. вр., 1-е лицо мн. ч.

Из анализа всех восьми групп окончаний видно, что в одни и те же группы попадают флексии, характеризующие различные части речи.

Такая классификация русских флексий может быть сравнима с речетом, через которое постепенно «просеиваются» все окончания.

§ 6. Формы изменяемых слов, имеющие нулевую флексию, разбиваются на три простых типа: 1) форма распознается при знании части речи, к которой принадлежит данное слово; 2) форма распознается при более подробных сведениях о слове; 3) форма распознается при анализе контекста, т. е. того предложения, в которое входит данное слово.

Все формы, имеющие нулевую флексию, следующим образом могут быть представлены в одной таблице:

Примеры форм	Тип 1	Тип 2	Тип 3
<i>греб, нес</i>	+		
<i>рук, сел красен брат</i>		+	
<i>дом солдат сапог жен</i>			+

Типы форм, представленные этими примерами, могут иметь следующие грамматические значения:

- 1) *греб* — прошедшее время глагола, мужской род;
- 2) *рук, сел* — родительный падеж множественного числа (если известна часть речи, род и отношение к категории одушевленности);
- 3) *красен* — краткая форма прилагательного (если известно, что оно непряжательное);
- 4) *брат* — именительный падеж единственного числа (если известна часть речи, род и одушевленность);
- 5) *дом* — именительный или винительный падеж единственного числа (если известна часть речи, род, неодушевленность);
- 6) *солдат* — именительный падеж единственного числа, родительный или винительный падеж множественного числа;
- 7) *сапог* — именительный или винительный падеж единственного числа или родительный падеж множественного числа;
- 8) *жен* — родительный или винительный падеж множественного числа.

По сравнению с анализом словоформ, имеющих ненулевую флексию, очевидно, что в случае с нулевой флексией первая группа флексий отпадает, происходит как бы «сдвиг» на следующий уровень, так как сама по себе нулевая флексия не свидетельствует о грамматических признаках.

§ 7. Все формы с нулевой и ненулевой флексией, требующие контекстного анализа, могут быть сгруппированы по частям речи.

Классификация этих форм будет представлять собой следующий раздел алгоритмической грамматики.

Все возможные типы совпадения форм существительного систематизированы в виде таблицы (плюсами отмечены категории, возможные для данной формы). При этом в одну группу попадают слова с разными флексиями.

Совпадение форм существительного:

№	Примеры форм	Единственное число						Множественное число						
		Им.	Род.	Дат.	Вин.	Тв.	Предл.	Им.	Род.	Дат.	Вин.	Тв.	Предл.	
1	<i>село, пень, строй, дом</i>	+			+									
2	<i>попугаев, бойцов, змей, жен</i>								+		+			
3	<i>дома, рога, болота</i>		+					+			+			
4	<i>учителя</i>		+		+			+						
5	<i>слухи, столы</i>							+			+			
6	<i>автора, героя</i>		+		+									
7	<i>книги, стены</i>		+					+			+			
8	<i>жены</i>		+					+						
9	<i>планетарии</i>							+	+			+		
10	<i>гении</i>							+	+					
11	<i>мышь</i>		+	+				+	+					
12	<i>печи</i>		+	+				+	+			+		
13	<i>сапог</i>	+			+				+					
14	<i>солдат</i>	+							+		+			
15	<i>стене, лесу</i>			+				+						
16	<i>поле</i>	+			+			+						

Из этой таблицы очевидно совпадение падежных различий для некоторых групп существительных. В парадигмах этих существительных не представлены полностью все падежные противопоставления¹².

Большое значение подобному совпадению форм — синкретизму па-

¹² Такие парадигмы русского являются объектом изучения в ряде групп, занимающихся машинным переводом с русского языка. Так, в Гарвардской группе (США) для каждого существительного определяется, имеет ли оно полную парадигму или «усеченную» (reduced paradigm).

дежей — придает Л. Ельмслев¹³, описание совпадений падежей в связи с трехмерным соотношением их противопоставления и в связи с фонетическим обликом основы и падежных окончаний дано Р. Якобсоном¹⁴. Однако представляется случайным наличие противопоставления для форм *автомобиля*, *автомобилю*, *автомобиле* и отсутствие противопоставления для форм *печи*, *печи*, *печи*. Само разнообразие «сокращенных парадигм» может быть свидетельством случайности (не в историко-фонетическом смысле) процессов, приведших к синкретизму.

§ 8. Эти 16 типов можно свести к меньшему числу укрупненных типов, объединенных по бинарному противопоставлению.

1. Именительный /винительный +*x*, где *x* — некий третий падеж (*x* — может быть равен нулю). Например:

дома = им. пад. ед. ч. + вин. пад. мн. ч. + род. пад. ед. ч. (*x* — род. пад. ед. ч.);

сапог = им. пад. ед. ч. + вин. пад. ед. ч. + род. пад. мн. ч. (*x* — род. пад. мн. ч.);

планетарии = им. пад. мн. ч. + вин. пад. мн. ч. + предл. пад. ед. ч. (*x* — предл. пад. ед. ч.).

2. Родительный /винительный¹⁵ (+*x*).

3. Родительный падеж единственного числа /именительный падеж множественного числа.

4. Предложный падеж единственного числа /именительный падеж множественного числа.

5. Дательный падеж единственного числа /предложный падеж единственного числа.

6. Комплексный тип: тип 3 + тип 5 (+*x*).

Внутри этих шести типов следующим образом располагаются 16 перечисленных выше неукрупненных групп:

Укрупн. тип	Какие мелкие типы входят	Основное противопоставление	Добавочное противопоставление
I	1	Им. и вин. п. ед. ч.	
	3	Им. и вин. п. мн. ч.	Род. п. ед. ч.
	5	Им. и вин. п. мн. ч.	
	7	Им. и вин. п. мн. ч.	Род. п. ед. ч.
	13	Им. и вин. п. ед. ч.	Род. п. мн. ч.
	9	Им. и вин. п. мн. ч.	Предл. п. ед. ч.
II	16	Им. и вин. п. ед. ч.	Предл. п. ед. ч.
	2	Род. и вин. п. мн. ч.	
	4	Род. и вин. п. ед. ч.	Им. п. мн. ч.
III	14	Род. и вин. п. мн. ч.	Им. п. ед. ч.
	8		
IV	10		
V	15		
VI	11		
	12		Вин. п. мн. ч.

¹³ L. Hjelmslev. La catégorie des cas. Etude des grammaire générale. Aarhus, 1935, стр. 60.

¹⁴ Р. О. Якобсон. Морфологические наблюдения над славянским склонением. «American contributions to the Fourth International Congress of slavists». S-Gravenhage, 1958.

¹⁵ В 1 и 2 типах число совпадает (т. е. может быть: им. п. мн. ч. и вин. п. мн. ч.; род. п. ед. ч. и вин. п. ед. ч.; им. п. ед. ч. и вин. п. ед. ч.; род. п. мн. ч. и вин. п. мн. ч.).

Как видно из таблицы, возможное добавочное противопоставление (x) всегда обозначает падеж, расходящийся в числе с двумя составными противопоставляемыми падежами, обязательно совпадающими в числе. Исключение представляет лишь легко распознаваемый предложный падеж единственного числа для форм типа *поле*. Поэтому, на наш взгляд, при классификации омонимичных форм русского существительного признак числа является релевантным признаком. Для распознавания формы существенную помощь оказывает именно знание того, какое число мы можем ожидать в данном случае. Для 16 мелких типов совпадения, приведенных выше, в девяти случаях характерно несовпадение числа при противопоставлении падежей. В связи с этим при анализе контекста разрешаемая дизъюнкция в некоторых случаях полностью снимается, если известно число сопровождающего прилагательного (*молодые жены*, но *молодой жены*).

Если учесть тот факт, что добавочное противопоставление (x) к выделенным укрупненным типам может в свою очередь входить в бинарное противопоставление с одним из основных для данного типа противопоставленных падежей, все выделенные противопоставления падежей в русском языке могут быть еще более точно описаны в виде следующей таблицы, где выделяется пять простых типов и шесть сложных, комбинирующихся из простых (названия падежей даны заглавными буквами):

Простой тип Комбинация						Возможные падежи	Примеры
	1	2	3	4	5		
1	+					И/В — число одно	<i>стол, поля</i>
2		+				Р/В — число одно	<i>брата, жен</i>
3			+			Р/И — числа разные	<i>жены</i>
4				+		П/И — числа разные	<i>гении</i>
5					+	Д/П — число одно	<i>жене, лесу</i>
6	+		+			И/В — одного числа, Р — другого	<i>дома, сапог</i>
7	+			+		И/В — мн. ч., П — ед. ч.	<i>планетарии</i>
8		+	+			Р/В — одного числа, И — другого	<i>учителя, солдат</i>
9			+		+	Р/Д/П — ед. ч., И — мн. ч.	<i>мышь</i>
10	+		+		+	Р/Д/П — ед. ч., И/В — мн. ч.	<i>печи</i>
11	+			+		И/В/П — ед. ч.	<i>море</i>

§ 9. Опишем контекстные сведения, необходимые для распознавания каждой формы для пяти простых типов (бинарные противопоставления). Каждая категория идентифицируется при знании одного из этих условий, однако не все предложенные условия в равной степени доста-

точные для различения форм. Порядок расположения этих условий в таблице не определяет степени их достоверности.

Категория	Тип простого противопоставления	Контекстные условия идентификации
Винительный падеж	1	<p>Если предложение, в которое входит эта форма, есть деепричастный оборот;</p> <p>если в предложении есть категория состояния;</p> <p>если в предложении есть глагол во 2-м или в 1-м лице любого числа;</p> <p>если число глагола в личной форме не совпадает с предполагаемым числом существительного;</p> <p>если перед данным словом есть предлог, требующий винительного падежа;</p> <p>если перед словом стоит переходный глагол;</p> <p>если в предложении есть глагол в личной форме с двойным управлением (один из падежей — винительный);</p> <p>если в предложении есть слово, о котором известно, что оно в именительном падеже</p>
Именительный падеж	1	<p>Если перед данным словом стоит тире и в предложении нет глагола в личной форме;</p> <p>если перед данной формой стоит слово <i>чем</i>;</p> <p>если сказуемое в предложении выражено существительным</p>
Родительный падеж	2	<p>Если перед данной формой стоит существительное;</p> <p>если перед данной формой стоит числительное;</p> <p>если перед данным словом стоит предлог, управляющий родительным падежом (с пропуском только прилагательных и наречий);</p> <p>если перед данным словом стоит отрицание;</p> <p>если перед данным словом стоит глагол, управляющий родительным падежом</p>
Винительный падеж	2	<p>Если данное слово не подходит под признаки родительного падежа (см. выше);</p> <p>если перед словом стоит предлог, требующий винительного падежа;</p> <p>если перед словом стоит глагол, требующий винительного падежа;</p> <p>если в предложении есть прилагательное в краткой форме и нет глагола в личной форме</p>

Категория	Тип простого противопоставления	Контекстные условия идентификации
Родительный падеж	3	<p>Если число сопровождающего прилагательного совпадает с числом возможного родительного падежа;</p> <p>если слово подходит под признаки родительного падежа по типу № 2;</p> <p>если число глагола совпадает с числом возможного родительного;</p> <p>если в предложении есть категория состояния (глаголоид);</p> <p>если в предложении есть слово в именительном падеже</p>
Предложный падеж	4 и 5	Если перед словом стоит предлог, требующий предложного падежа
Дательный падеж	5	Если слово не стоит в предложном падеже
Именительный падеж	4	Если перед словом не стоит предлог

Как видно из приведенных сведений, для разрешения каждого бинарного противопоставления нужен особый анализ для каждого падежа. Например, определение винительного падежа для пары «винительный / именительный» и для пары «винительный / родительный» различно. В первом случае разрешается совпадение субъекта и объекта, во втором — двух объектов. Поэтому для разрешения первой дизъюнкции достаточно, например, иметь сведения о том, что в данном предложении уже есть слово в именительном падеже — подлежащее, или же, напротив, о том, что его не может быть (сказуемое выражено безлично-предикативным словом). Для разрешения второй дизъюнкции эти признаки не имеют значения. В некоторых случаях определить форму проще негативным, чем позитивным путем. Например, для определения именительного падежа при противопоставлении «именительный — предложный» более просто выяснить, что перед данным словом нет никакого предлога, чем перебирать все возможные случаи позиции именительного падежа.

§ 10. Формы прилагательного, требующие при распознавании анализа контекста, делятся на две группы: а) формы, совпадающие внутри одной парадигмы (*молодой жены, молодой жене, молодой женой. . .*), б) формы, совпадающие вне одной парадигмы (*хорошо сделано, это решение хорошо*).

Формы первого типа не анализировались, так как предполагалось, что распознать эти формы можно на основе сведений, характеризующих форму существительного, определяемого данным прилагательным.

Формы второго типа представляют собой следующие возможности совпадения.

1. *-o* — наречие, или категория состояния, или краткая форма.
 2. *-ee* — единственное число среднего рода или сравнительная степень.
 3. *-ей* — единственное число женского рода или сравнительная степень.
 4. *-e* — краткая форма или сравнительная степень, или наречие.
- Эти формы распознаются при следующих условиях:

Форма	Из какой возможности	При каком условии распознается
Наречие	4	Если в предложении есть глагол в личной форме; если данное предложение есть причастный или деепричастный оборот
Сравнительная степень	2 и 3	Если следующее слово <i>чем</i> ; если следующее слово — существительное в род. п.
Положительная степень	2	Если за словом стоит существительное ср. р. в им., вин. или предл. п. ед. ч.
Положительная степень	3	Если за данным словом следует существительное ж. р. ед. ч. не в им. п.
Краткая форма	1	Если в предложении нет сказуемого и подлежащего среднего рода

§ 11. Для глагола на основе анализа контекста распознаются 1) формы множественного числа повелительного наклонения — 2-го лица мн. ч. глаголов типа *положить* и 2) формы сослагательного наклонения — формы прошедшего времени.

Распознавание форм в этом случае происходит на основе следующих сведений:

Форма	Из какой возможности	При каком условии распознается
Настоящее время	1	Если перед глаголом стоит <i>Вы</i> *; если предложение начинается с союза
Повелительное наклонение	1	Если нет условий для определения 2-го лица
Сослагательное наклонение	2	Если в предложении есть частица <i>бы</i>
Прошедшее время, изъявительное наклонение	2	Если в предложении нет частицы <i>бы</i>

* Настоящее условие не имеет силы для художественной литературы.

§ 12. Рассмотрев различные уровни «распознавания» грамматических категорий по данным текста, можно как бы «перевернуть» вопрос и описать, какими различными способами может быть получена каждая грамматическая категория.

Основные категории русского существительного могут классифицироваться по распознаванию следующим способом:

Падеж, число	Распознается по флексии сразу	Распознается различными способами	Не распознается по флексии никогда
Им. ед.		+	
Род. ед.		+	
Дат. ед.		+	
Вин. ед.		+	
Тв. ед.		+	
Предл. ед.		+	
Им. мн.		+	
Род. мн.		+	
Дат. мн.	+		
Вин. мн.			+
Тв. мн.	+		
Предл. мн.	+		

§ 13. Для прилагательного распознаются следующие категории: род, число, падеж, степени сравнения. Покажем на нескольких таблицах, какими способами могут быть выведены эти категории.

Р о д:

Определяется по окончанию		Окончания совпадают	
окончание	род	по роду	по числу
-ая, -яя, -ую, -юю, -а, -я	Женский	-ой, -ей	-ым, -им
-ый, -ий	Мужской	-ого, -его	
-ос, -ес, -о, -е	Средний	-ому, -ему -ом, -ем	

Степень сравнения:

	Распознается по окончанию	Распознается по контексту
Положительная	+	+
Сравнительная *		+
Превосходная	+	+

* Сравнительная степень распознается только при анализе контекста, так как она может быть выражена или аналитическим путем или синтетическим. В последнем случае все показатели сравнительной степени имеют аналогию среди показателей форм положительной степени прилагательного.

Число:

Распознается по окончанию		Распознается по контексту
окончание	число	
<i>-ая, -яя, -ое, -ея, -ую, -юю, -ый, -ой, -ий, -ей, -ого, -его, -ому, -ему, -ем, -ом, -а, -о, -е, -о</i>	Единственное	<i>-ым, -им</i>
<i>-ие, -ые, -и, -ы, -ими, -ыми, -их, -ых</i>	Множественное	

Падеж:

Падеж, число, род	Распознается по окончанию	Распознается по контексту	Падеж, число, род	Распознается по окончанию	Распознается по контексту
Им. ед. ж.	+		Тв. ед. м. — ср.		+
Род. ед. ж.		+	Предл. ед. м. — ср.	+	
Дат. ед. ж.		+	Им. ед. ср.		+
Вин. ед. ж.	+		Вин. ед. ср.		+
Тв. ед. ж.		+	Им. мн.		+
Предл. ед. ж.		+	Род. мн.		+
Им. ед. м.		+	Дат. мн.		+
Род. ед. м. — ср.		+	Вин. мн.		+
Дат. ед. м. — ср.	+		Тв. мн.	+	
Вин. ед. м.		+	Предл. мн.		+

§ 14. Основные грамматические характеристики глагола в русском языке могут быть распознаны следующими способами.

В р е м я:

	Распознается при помощи		
	окончания	сведений о слове	сведений о контексте
Настоящее	+	+	
Прошедшее	+		+
Будущее	+	+	+

Г л а г о л ь н а я р е п р е з е н т а ц и я (личная форма, инфинитив, причастие или деепричастие) распознается только на основе окончаний.

Ч и с л о распознается только на основе окончаний.

Н а к л о н е н и е распознается и при помощи окончаний (*дела-ет* — только изъявительное наклонение) и при помощи контекста (*сделал* — *сделал бы*).

Л и ц о распознается при помощи окончания, кроме 2-го лица множественного числа части глаголов II спряжения, распознаваемого при помощи контекста.

З а л о г распознается 1) при помощи окончания (*делает* — действительный залог, *сделан* — страдательный залог); 2) при помощи окончания + сведения о слове: *смеется* — действительный залог, так как глагол этого типа без *-ся* не употребляется; *пишется* — страдательный залог, так как глагол этого типа с *-ся* всегда страдательного залога; 3) при помощи контекста: *Это произведение читается с эстрады лучшими декламаторами* — страдательный залог. *Это произведение читается с трудом* — действительный залог.

* * *

Различие в «количестве шагов», требуемых при формальном распознавании каждой грамматической категории языка, было бы удобным образом сопоставимо при наличии некоторой общей меры грамматической информации — единицы, в которой могло быть описано количественно различие путей прохождения от единицы текста до единицы грамматики — граммы. Различие в количестве подобных единиц могло бы определять расхождение между построением и распознаванием словоформ (формы *жены* и *жену* однотипны для построения, но различны для распознавания). Такая мера грамматических сведений еще не выведена, что значительно затруднило предложенное описание.

1. Анализ различных уровней сложности при распознавании категорий в русском языке говорит о немотивированности требований, предъявляемых при анализе конкретных словоформ. При определении одной и той же категории (например, одного падежа) могут требоваться самые различные по количеству требуемой информации сведения (например, при распознавании дательного падежа единственного числа для формы типа *столу* и для формы типа *мышь*).

2. Описание языка по степени его формальной распознаваемости по категориям может служить основой для создания алгоритма грамматического анализа этого языка. При этом алгоритм должен быть построен с учетом особенностей данного языка и наиболее экономным образом, т. е. сведения, необходимые на каждом этапе анализа, должны основываться на сведениях, полученных на предыдущем этапе.

3. Характеристика языка по степени и уровням распознаваемости категорий в тексте может быть одним из критериев типологического сопоставления языков.

При этом основой типологического сопоставления является различие в степени «асимметричности» грамматик анализа и синтеза (грамматик распознавания и порождения) для каждого из сопоставляемых языков по сравнению с некоторым идеальным языком, в котором каждая флексия соответствует одной из возможных комбинаций грамматических категорий и, наоборот, каждой возможной комбинации грамматических категорий соответствует одна флексия. Заранее очевидно, что степень этого отклонения будет (наряду с другими показателями) в значительной степени определяться степенью и типами нейтрализации грамматических категорий, представленных в данном языке.
