



А. А. Зализняк

Лингвистические задачи



А. А. Зализняк

Лингвистические задачи

*С предисловием В. А. Успенского
и статьёй А. Ч. Пиперски*

Третье издание, дополненное

Москва, МЦНМО
2018

УДК 81
ББК 74.200.58:81.2
323

Зализняк А. А.
323 Лингвистические задачи / С предисловием В. А. Успенского и статьёй А. Ч. Пиперски. — 3-е изд., дополн. — М.: МЦНМО, 2018. — 56 с.

ISBN 978-5-4439-1272-1

В брошюре публикуется классическая работа А. А. Зализняка, в которой впервые появились самодостаточные задачи по лингвистике. Статья перепечатана из сборника «Исследования по структурной типологии» (М.: Изд-во АН СССР, 1963. С. 137—159).

В предисловии В. А. Успенского подробно рассказывается об этой работе и жанре лингвистических задач. Статья А. Ч. Пиперски посвящена работам выдающегося лингвиста, академика РАН Андрея Анатольевича Зализняка (1935—2017).

ББК 74.200.58:81.2

Андрей Анатольевич Зализняк

ЛИНГВИСТИЧЕСКИЕ ЗАДАЧИ

Подписано в печать 26.03.2018 г. Формат 60×90/16.
Гарнитура ITC Charter. Печать офсетная. Печ. л. 3,5. Тираж 2000 экз.

Издательство Московского центра
непрерывного математического образования
119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241-08-04

Отпечатано с готовых диапозитивов в ООО «Принт Сервис Групп».
105187, Москва, Борисовская ул., 14

ISBN 978-5-4439-1272-1

© Зализняк А. А., 1963.
© МЦНМО, 2018.

ПРЕДИСЛОВИЕ

Начну с изъявления двух радостных чувств.

Это, во-первых, чувство глубокого удовлетворения от состоявшегося события, причём, на мой взгляд, — события значительного. Событие заключается в том, что данное уникальное сочинение А. А. Зализняка появилось, наконец, в виде отдельного издания. Лингвистические задачи Зализняка, разделённые на две группы и представленные в этой книжке, могут не только составить тему занятий кружков для школьников — кружков как по лингвистике, так и по математике, — но и доставить истинное наслаждение всем, кто готов получать удовольствие от интеллектуальных упражнений. До сего времени это сочинение Зализняка печаталось хотя и дважды, но только в составе сборников. Таковыми сборниками были:

Исследования по структурной типологии / Ин-т славяноведения АН СССР; Отв. ред. Т. Н. Молошная. — М.: Изд-во АН СССР, 1963. — С. 137—159;

Задачи лингвистических олимпиад. 1965—1975 / Ред.-сост. В. И. Беликов, Е. В. Муравенко, М. Е. Алексеев. — М.: МЦНМО, 2006. — С. 516—545.

Теперь, когда «Лингвистические задачи» появились в виде отдельной брошюры, можно не только её прочесть, но и найти текст на сайте МЦНМО, то есть Московского центра непрерывного математического образования, в разделе «Свободно распространяемые издания». Достаточно набрать <http://www.mcsme.ru/free-books/> в адресной строке браузера и далее читать на своём компьютере. В настоящей брошюре, кстати, исправлены те немногочисленные опечатки, которые присутствовали в указанных сборниках.

Во-вторых, я вспоминаю тот восторг, который охватил меня, когда полвека тому назад я познакомился с текстом ещё в рукописи.

Уже первая фраза первой задачи сражала наповал: «Для лиц, незнакомых с баскским языком»! А сама задача вызвала у меня оторопь формулировкой задания: выписывались, без переводов, 12 предложений на баскском языке, указывалось, что в одном из них допущена грамматическая ошибка, и требовалось эту ошибку найти и исправить. Тут какая-то чушь, подумалось мне. Ведь может случиться, что в этом таинственном баскском языке именно так и положено сказать. Как же можно такое опровергнуть? С тем бóльшим удивлением я обнаружил через некоторое время, что решил задачу: и ошибку нашёл, и исправление предложил. Более того, оказалось, что почти все, кто брался за эту задачу, успешно её решали.

Напрашивающееся объяснение таково: в подсознании человека содержатся некоторые представления о том, как может и как не может быть устроен человеческий язык. Эти представления формализуются с большим трудом — если вообще формализуются. Их реализация, демонстрируемая возможностью успешного решения сформулированной только что задачи, происходит на интуитивном уровне. На том же, до сих пор не формализованном уровне происходит, скажем, узнавание со спины человека, которого мы до того видели только спереди, а также наша способность продолжить числовую последовательность, заданную своим началом, — например, последовательность 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ...¹. Всё это наводит на непростые размышления о природе человеческого знания. Задач подобного типа — основанных на эксплуатации подсознания — в книге четыре. Они образуют первую группу. Вот, к примеру, третья задача этой группы. Выписаны шесть фраз на албанском языке с их переводами на древнееврейский; после чего выписываются две фразы на древнееврейском языке и предлагается перевести их на албанский.

Вторую группу составляют три задачи совершенно другого типа, объединённые единством замысла. По существу, в них читателю предлагается самостоятельно построить лингвистическую теорию, способную объяснить некоторые нерегулярности современного языка.

1 Это — хорошо известная математикам *последовательность Фибоначчи*, каждый член которой равен сумме двух предыдущих. Опыт показывает, что и лица, с нею не знакомые, в состоянии извлечь из выписанных начальных членов необходимую для продолжения закономерность.

Чтобы стало понятно, о чём идёт речь, приведу примеры. Вот не вызывающие вопросов переходы от существительных к образованным от них глаголам: *коса* — *косить*, *мера* — *мерить*, *гроза* — *грозить* (и тому подобные; читатель может продолжить). Отчего же тогда не *калека* — *калекить*, *рука* — *приручить*, *слуга* — *служить*, *нега* — *нежить*, а почему-то *калека* — *калечить*, *рука* — *приручить*, *слуга* — *служить*, *нега* — *нежить*? Потому, отвечает лингвист, что когда-то как раз и было *калекить*, *приручить*, *служить*, *нежить*, но в дальнейшем эти слова постепенно превратились соответственно в *калечить*, *приручить*, *служить*, *нежить*. Тогда встаёт следующий вопрос: каковы те события, происшедшие в процессе развития языка, которые вызвали такое превращение? Историческая лингвистика русского языка даёт на это такой ответ: согласные звуки [к] и [г], сопровождаемые гласными звуками определённого типа, к которому, в частности, принадлежит звук [и], повсюду в русском языке превращались в звуки [ч] и [ж]. Слово «повсюду» может вызвать у читателя недоумение: если превращение происходило повсюду, то слова «калека», «рука», «слуга» и «нега», взятые в форме родительного падежа единственного числа или в форме именительного падежа множественного числа, должны выглядеть как *калечи*, *ручи*, *служи*, *нежи*. Но ведь эти формы выглядят по-другому, а именно *калеки*, *руки*, *слуги*, *неги*. Как же так? Подумав, можно прийти к следующему заключению. Если бы первоначальными, древними формами были *руки*, *слуги* и так далее, то в силу правила превращения — перед звуком [и] — звуков [к] и [г] соответственно в звуки [ч] и [ж], сегодняшними формами существительных были бы *ручи*, *служи* и так далее. Но раз это не имеет места, то, значит, древние формы имели какой-то другой вид. И действительно, этот вид был таков: *калекы*, *руки*, *слуги*, *негы*. И лишь впоследствии — причём позже, чем совершился переход звуков [к] и [г] перед звуком [и] соответственно в звуки [ч] и [ж], — звук [ы] в этих формах в силу некоторого другого перехода превратился в звук [и].

Описанная процедура восстановления древних форм слов и установления законов развития, приведших от древних форм к современным, называется *внутренней реконструкцией*. (Бывает и так называемая *внешняя реконструкция*, но о ней нет необходимости здесь говорить.) В задачах второй группы читателю как раз и предлагается осуществить внутреннюю реконструкцию, только

не для русского, а для других языков — латышского и древнеиндийского.

Сборники задач по лингвистике публиковались и до того, как А. А. Зализняк обнародовал свои задачи. Некоторый список таких сборников приводит автор в сноске к первой же фразе своего сочинения. Для полноты добавим к этому списку два более поздних задачника:

В. А. Малаховский. Сборник задач и упражнений по курсу «Введение в языкознание». — М.: Учпедгиз, 1960;

В. Н. Клюева. Задачник по курсу «Введение в языкознание». — М.: Высшая школа, 1962.

Уже из самих названий этих сборников задач (как сборников, названных А. А. Зализняком, так и добавленных нами) видно, что они являются задачниками, цель которых — поупражняться в усвоении некоторого языковедческого курса. Таким образом, задачи, собранные в этих сборниках, предполагают некие знания, в соответствующих курсах полученные. Задачи Зализняка — совершенно иного рода. Они *самодостаточны* в том смысле, что не предполагают никаких лингвистических знаний и тем самым доступны для попытки решения любому желающему.

Поскольку раньше подобных задач не было, то оказалось, что для них нет общепринятого названия. Термин лингвистические задачи, кажущийся ныне совершенно естественным, в то время далеко не всеми воспринимался таковым. Характерна реакция одной лингвистки, прослышавшей, что вышел сборник «Исследования по структурной типологии» со статьёй «Лингвистические задачи». Принадлежа к кругу хороших знакомых автора статьи, она, с деликатно смягченным неодобрением, сказала кому-то из того же круга: «Это что же, выходит, Андрей решил уже выступить как академик Виноградов?» Академик Виктор Владимирович Виноградов, чьё имя носит сейчас Институт русского языка Академии наук, пребывал на вершине лингвистического Олимпа того времени, и только ему и равным ему принадлежало неписанное право писать программные статьи о задачах лингвистики, а единственный смысл, который она смогла приписать странному и никогда не встречавшемуся словосочетанию лингвистические задачи, — это ‘задачи лингвистики’.

Таким образом, Зализняк открыл новый жанр — *жанр самодостаточных лингвистических задач*. С одной стороны, задачи

этого жанра дают прекрасный материал для исследования мыслительной деятельности человека, о чём уже говорилось выше, с другой — сыграли поистине историческую роль в деле подготовки лингвистов. Дело в том, что именно опубликование в 1963 году этих задач сделало возможным лингвистические олимпиады школьников. В 1965 году в Москве состоялась Первая традиционная московская лингвистическая олимпиада ²; XLII Московская традиционная олимпиада по лингвистике состоялась в конце 2011 года. Традиционные московские олимпиады стали образцом и побудительной причиной для Международных олимпиад по лингвистике, которые теперь проводятся ежегодно, начиная с 2003 года. Первая Международная олимпиада состоялась в Болгарии в 2003 году, а далее они последовательно проводились в России, в Нидерландах, в Эстонии, в России, в Болгарии, в Польше, в Швеции, в США, в Словении. XI Международная олимпиада по лингвистике намечена на июль 2013 года в Великобритании. Конечно, в процессе проведения лингвистических олимпиад в их задачах не могли не появиться новые идеи, но все они так или иначе были развитием тех идей самодостаточности, которые были заложены в пионерской работе А. А. Зализняка. Без появления этой работы олимпиадное движение в области лингвистики не могло бы родиться.

Подводя итоги, можно сказать, что лингвистические задачи Зализняка имеют выдающееся научное, педагогическое и эстетическое значение.

Enjoy!

май 2013 г.

В. А. Успенский

2 Тогда она называлась по-другому: «Первая традиционная московская олимпиада по языковедению и математике».

АНДРЕЙ АНАТОЛЬЕВИЧ ЗАЛИЗНЯК
(1935—2017)

24 декабря 2017 года в возрасте 82 лет ушёл из жизни великий российский лингвист, академик РАН Андрей Анатольевич Зализняк. На его счету — множество выдающихся открытий. Он изучал древненовгородские берестяные грамоты и описал грамматику древненовгородского диалекта; он убедительно доказал, что «Слово о полку Игореве» действительно было написано около 1200 года, а не является подделкой XVIII века, как многие считали раньше; он создал словарь русского словоизменения, который и по сей день лежит в основе компьютерных программ, работающих с русским языком; он развил и доработал определение падежа, данное А. Н. Колмогоровым, — и многое другое. Строгость и доказательность работ Андрея Анатольевича была и остаётся идеалом для всех его учеников — а таковыми считают себя почти все лингвисты Москвы и не только.

Отличительная черта научной деятельности Зализняка — интерес к математике и точным методам в языкознании. В 1950-е годы он учился на французском отделении филологического факультета Московского университета — казалось бы, что может быть более гуманитарным и далёким от точных наук? Однако он стал ходить к В. А. Успенскому на занятия по математике для филологов и оказался в числе лучших студентов. «Он меня абсолютно потряс. Он задавал какие-то вопросы, чрезвычайно глубокие, но совершенно перпендикулярные к вопросам всех остальных и к тому, чего я мог ожидать. Какой-то такой был поворот в его вопросах, словно он с другой стороны на всё смотрел. Сразу стало ясно, что это гений», — вспоминает Успенский¹. Вскоре Зализняк стал одним из руководителей семинара по математической лингвистике на механико-математическом факультете

1 Борьба за лингвистику. Беседа с математиком Владимиром Успенским. Часть 1.
<http://polit.ru/article/2009/07/09/linguist/>

МГУ. О некоторых достижениях Зализняка, связанных с формализацией лингвистического знания, мне бы и хотелось рассказать поподробнее.

Русское словоизменение

В 1957—1958 гг. Зализняк был на стажировке в Париже. Там он преподавал французским студентам русский язык и обнаружил, что существующих описаний русского склонения и спряжения для иностранцев недостаточно: они слишком расплывчатые. Уже в 1961 году Зализняк издал «Краткий русско-французский учебный словарь», приложив к нему абсолютно формальный 150-страничный очерк русского словоизменения. В 1965 году из этого очерка выросла кандидатская диссертация; работа оказалась настолько новаторской, что её сразу засчитали за докторскую. В 1967 году Зализняк издал книгу «Русское именное словоизменение», а в 1977 году — «Грамматический словарь русского языка».

В этом словаре при каждом слове стоят непривычные пометы: *дуб* м 1с, *порог* м 3а, *фонарь* м 2в. Ясно, что «м» обозначает мужской род, а зачем всё остальное? Ведь мы со школьных лет и так знаем, что *дуб*, *порог* и *фонарь* — это II склонение.

Однако для полного описания русского языка такого знания недостаточно. Чтобы убедиться в этом, образуем родительный падеж единственного числа и именительный падеж множественного числа:

нет дуба — есть дубы́,
нет порога — есть порóги,
нет фонаря́ — есть фонари́.

Видно, что в окончаниях этих слов появляются разные буквы (в одном — *а* и *ы*, в другом — *а* и *и*, в третьем — *я* и *и*), да и с ударением всё непросто: у слова *дуб* оно падает то на окончание, то на основу, у слова *порог* оно в обеих формах на основе, а у слова *фонарь* — на окончании. Именно эту информацию и несут в себе пометы.

Цифры обозначают, на что оканчивается основа и как будут выглядеть окончания:

1 — основа на твёрдый согласный (кроме заднеязычных *к*, *г* и *х*, шипящих *ш* и *ж* и свистящего *ц*);

2 — основа на мягкий согласный (кроме *ч* и *й*);

3 — основа на *к*, *г* или *х*; и т. д.

Таких типов у существительных насчитывается девять: 0 — несклоняемые слова (например, *кенгуру*), 1—7 — варианты традиционных I и II склонений (например, *акула, поезд, ведро*) и 8 — традиционное III склонение (например, *крепость*).

Буква обозначает схему ударения:

a — всегда на основе;

b — всегда на окончании;

c — подвижное ударение: в единственном числе на основе, а во множественном — на окончании; и т. д.

Всего у существительных насчитывается восемь схем, включающих в себя по крайней мере 10 слов. Больше 90 % существительных в «Грамматическом словаре» Зализняка относятся к схеме *a* — но это не значит, что остальными схемами можно пренебречь. Дело в том, что редкие схемы включают в себя очень частотные слова: например, многие названия частей тела. Посклоняйте слово *рука* и посмотрите, когда ударение будет на основе, а когда на окончании: распределение там очень прихотливое. И вот если считать не по словарю, а по текстам, то на все схемы, кроме *a*, придётся более четверти всех существительных (табл. 1).

Таблица 1

Схема ударения	Примеры	Число существительных в словаре	Доля в словаре (%)	Доля в текстах (%)
<i>a</i>	<i>карта, порог</i>	44933	91,62	73,40
<i>b</i>	<i>фонарь, ступня</i>	3117	6,36	6,85
<i>c</i>	<i>дуб, море</i>	408	0,83	8,75
<i>d</i>	<i>изба, вино</i>	336	0,69	3,30
<i>e</i>	<i>ухо, доля</i>	185	0,38	5,04
<i>f</i>	<i>губа, червь</i>	35	0,07	0,38
<i>d'</i>	<i>спина, душа</i>	11	0,02	0,91
<i>f'</i>	<i>рука, гора</i>	15	0,03	1,37
Всего		49040	100,00	100,00

Кстати, в этой таблице пары слов в каждой ячейке упорядочены так же, как и в самом «Грамматическом словаре» — в инверсионном (обратном) порядке, то есть по алфавиту при чтении справа налево, а не слева направо. Если объединить эти 16 слов в один список и упорядочить его таким образом, получим: *изба, губа, рука, спина, гора, карта, душа, дуб, порог, море, вино, ухо,*

червь, фонарь, доля, ступня (как и при обычной алфавитной сортировке, если последние буквы одинаковы, смотрим на предпоследние и т. д.).

Зализняк не зря выбрал такой порядок. Благодаря этому в словаре оказываются рядом слова с одинаковой правой частью, то есть те, которые изменяются похоже: глаголы на *-нуть*, существительные на *-а* и т. п. В нашем примере видно: подряд расположились 7 слов традиционного I склонения, потом 7 слов II склонения и потом ещё 2 слова I склонения. Сейчас цель такой сортировки кажется надуманной: можно же взять оцифрованную версию словаря и за доли секунды выбрать все слова на *-нуть*, — но в 1977 году такой возможности не было. Однако с распространением компьютеров обнаружилось новое применение «Грамматического словаря»: именно на нём основаны программы, которые работают с русской морфологией. Без словаря гораздо хуже функционировали бы современные поисковые системы и голосовые помощники. Например, введя в поиск *Пушкин стихи*, вы получаете страницы, на которых написано *в стихах Пушкина*. Это не просто слова, которые отличаются на 1—2 буквы (иначе находились бы и страницы со словами *пушки* и *стихия*, а такого не происходит), а именно формы слов, построенные по моделям Зализняка. А голосовой помощник без словаря не построил бы родительный падеж от слова *фонарь*, а если бы он озвучивал текст, то не узнал бы, что надо читать *фонаря́*, а не *фона́ря*.

Легко понять, как на основе словаря Зализняка работает морфологический синтез. Если мы строим формы от слова *фонарь*, надо взять его помету (2b), выбрать из окончаний типа 2 то, которое выражает нужную комбинацию падежа и числа: например, *-я* для род. ед., — и узнать во вводной части словаря, что основа слов типа 2 получается отбрасыванием конечного *ь*. В результате выходит: *фонарь*- + *-я* = *фонаря*. Чуть сложнее со словами, у которых несколько основ: так, слово *огурец* имеет помету 5*b, где * обозначает наличие беглой гласной (т. е. основ будет две: *огурец*- и *огурц*-). Опять-таки из вводной части словаря узнаём, что окончание род. ед. в типе 5 имеет вид *-а*; если есть знак *, то берётся основа без гласной в случае, если окончание не входит в множество {-∅ (пустая строка); -й; -ь; -ью}. Получаем *огурц*- + *-а* = *огурца*.

Противоположная задача — опираясь на «Грамматический словарь», понять по строке символов, является ли она формой какого-либо русского слова, и если да, то какие это могут быть

формы и от каких слов — с вычислительной точки зрения решается сложнее. Требуется не только хранить все возможные наборы окончаний и помету для каждого слова, но и заранее построить все возможные основы и записать, в какой форме какая из основ используется. Далее можно предложить такой алгоритм: разбиваем слово на две части всеми возможными способами и смотрим, является ли первая часть основой каких-либо слов. Если да, проверяем, бывает ли вторая часть окончанием у этих слов при такой основе, и если да, то выписываем те слова и формы, в которых это возможно. Проанализируем в качестве примера слово *паром*, для которого получается три разбора (табл. 2).

Таблица 2

Часть 1 (основа)	Часть 2 (окончание)	Возможная начальная форма	Есть ли у этого слова нужное окончание?	Результат
<i>паром-</i>	-∅	ПАРОМ	да, им. ед. и вин. ед.	ПАРОМ, им. ед. ПАРОМ, вин. ед.
<i>паро-</i>	-м	—	—	
<i>пар-</i>	-ом	ПАРИТЬ	—	
		ПАРИТЬ	—	
		ПАРА	—	
		ПАР	да, твор. ед.	ПАР, твор. ед.
<i>па-</i>	-ром	ПА	—	
<i>п-</i>	-аром	—		
<i>∅-</i>	-паром	—		

Аркадий Волож, руководитель «Яндекса», вспоминает, что в начале 1990-х его другу и коллеге Илье Сегаловичу на основе словаря Зализняка удалось создать морфологический анализатор, который работал настолько эффективно, что умещался в оперативную память тогдашнего персонального компьютера — 640 килобайт.

Можно, конечно, сказать: зачем все эти сложные обозначения, схемы ударения, чередования основ и т. д.? Давайте просто для каждого слова выпишем все формы с ударением, и готово. Но это очень неэкономно: на каждое существительное придётся по 12 форм, а на некоторые глаголы — больше 100 (считая четыре причастия во всех родах, числах и падежах). А «Грамматический словарь» — не такая уж и толстая книга, хотя и имеет большой фор-

мат: в 4-м издании описание русского словоизменения занимает 118 страниц, а за ними следуют 586 страниц собственно словаря.

Стремясь описать язык оптимальным образом, Зализняк ставил и решал в том числе и математические задачи. К одной из них подводит пример, с которого мы начали. Напомним, что слово *фонарь* имеет схему ударения *b*, то есть ударение всегда на окончании. Прежде чем читать дальше, подумайте: нет ли здесь некоторого противоречия, которое надо разрешить?

Зададимся странным вопросом: куда падает ударение в форме *фонарь* — на основу или на окончание? На этот вопрос есть три варианта ответа:

1) на основу: ведь ясно, что гласный *a* входит в основу, а окончание здесь нулевое;

2) мы не знаем куда: основа и окончание здесь не соревнуются за ударение, поскольку окончание нулевое;

3) на окончание, но введём дополнительное правило: если ударение попало на нулевое окончание (и получается, что ему негде прозвучать), оно переносится на последний слог основы.

Обозначим ударение на основе через +, ударение на окончании через —, а «не знаем» через 0. Примем вариант 1 — самый, как кажется, естественный — и изобразим схему ударения во всех числах и падежах² для нескольких русских существительных (табл. 3). Нулей в полученной таблице не будет, поскольку этот вариант не предполагает ответа «не знаем».

Таблица 3

	ед.					мн.					Примеры	
	И.	Р.	Д.	В.	Т.	П.	И.	Р.	Д.	Т.		П.
I	+	+	+	+	+	+	+	+	+	+	+	юноша, карта, тигр, порог ступня фонарь вратарь сапог
II	—	—	—	—	—	—	—	—	—	—		
III	+	—	—	+	—	—	—	—	—	—		
IV	+	—	—	+	—	—	—	—	—	—		
V	+	—	—	+	—	—	—	+	—	—		

Видно, что вариант 1 не очень экономный: нам приходится вводить разные схемы ударения для неодушевлённого слова

2 За исключением винительного падежа множественного числа, поскольку он всегда совпадает либо с родительным, либо с именительным в зависимости от одушевлённости (вин. *вижу столы* как им. *стоят столы*, вин. *вижу кошек* как род. *глаза кошек*).

фонарь и одушевлённого слова *вратарь* (у них различается вин. ед.: *вижу фона́рь*, но *вижу вратаря́*), не говоря уже про очень похожее слово *сапог*, для которого понадобится схема V из-за того, что у него род. мн. имеет нулевое окончание: *нет сапо́г*. Всего таких схем, как посчитал Зализняк, окажется 22.

Ещё менее экономен вариант 2, который даёт нам таблицу с нулями (табл. 4). Если использовать его, схем выйдут целых 39. Слова *карта*, *тигр* и *порог* получают не ту же схему, что *юноша* (строка II не изменится, а в строках III—V вместо плюсов будут нули).

Таблица 4

	ед.						мн.					Примеры
	И.	Р.	Д.	В.	Т.	П.	И.	Р.	Д.	Т.	П.	
I	+	+	+	+	+	+	+	+	+	+	+	<i>юноша</i>
I'	+	+	+	+	+	+	+	0	+	+	+	<i>карта</i>
I''	0	+	+	+	+	+	+	+	+	+	+	<i>тигр</i>
I'''	0	+	+	0	+	+	+	+	+	+	+	<i>порог</i>

Таблица 5

	ед.						мн.					Примеры
	И.	Р.	Д.	В.	Т.	П.	И.	Р.	Д.	Т.	П.	
a	+	+	+	+	+	+	+	+	+	+	+	<i>юноша, карта, тигр, порог</i>
b	—	—	—	—	—	—	—	—	—	—	—	<i>ступня, фонарь, вратарь, сапог</i>

А самым экономным оказывается вариант 3 (табл. 5): если использовать его, выяснится, что *юноша*, *карта*, *тигр* и *порог* — это одна схема; *ступня*, *фонарь*, *вратарь* и *сапог* — ещё одна схема (и да, как это ни странно, в слове *порог* ударение на основе, а в слове *фонарь* — на окончании). Всего для русского языка потребуется 11 таких схем. Выше были продемонстрированы только 8 из них, обозначенные как a, b, c, d, e, f, d' и f', поскольку остальные представлены совсем небольшим числом слов.

Зализняк не просто получил это решение интуитивно, но и доказал строго математически, что оно оптимально, если мы хотим минимизировать число строк в таблице. Вот как выглядит задача, которую он перед собой поставил:

Имеется матрица из m строк и n столбцов, элементами которой являются знаки +, — и 0. Заменить каждый из имеющихся

в матрице нулей плюсом или минусом так, чтобы в полученной матрице без нулей (назовём ее результирующей для данной матрицы) число s различных строк было минимальным.

Читателей, заинтересовавшихся полным решением, отсылаем к книге А. А. Зализняка «Русское именное словоизменение» (М.: Наука, 1967; переиздание: М.: Языки славянской культуры, 2002). Заметим, что для решения этой задачи автор сформулировал и доказал три теоремы; это нечасто встречается в лингвистических работах — но, оказывается, встречается!

Лингвистические задачи

В 1963 году Зализняк опубликовал статью «Лингвистические задачи», которая и составляет основную часть этой брошюры. В статье он продемонстрировал, что можно анализировать факты неизвестных языков, опираясь на строгую логику в сочетании с самыми базовыми представлениями о том, как устроены тексты на человеческом языке. Так, мы вправе ожидать, что порядок слов в языке не случаен, а подчиняется каким-то закономерностям: например, определение всегда стоит после определяемого слова или всегда перед ним, связанные друг с другом слова располагаются рядом, а не раскиданы в разные концы предложения и т. п. И тогда применение логических операций позволит нам понять, что есть что в непонятном тексте. Именно на этом основываются задачи 1—4, приводимые в статье.

Задачи Зализняка самодостаточны, то есть никаких специальных знаний не нужно: например, в задаче 4 даются 14 фраз на баскском с переводами на венгерский, а дальше читателю предлагается самому перевести три фразы с венгерского на баскский. Автор специально подчёркивает, что это задача для тех, кто не знает ни того ни другого языка и, таким образом, даже не понимает, что эти фразы значат. Тем не менее, после тщательного анализа примеров становится ясно, что задача является разрешимой.

Другой тип задач, предложенных Зализняком, — это задачи на внутреннюю реконструкцию (№№ 5—7 в его статье). Это, конечно, не было так непривычно для лингвистов того времени, как перевод с неизвестного языка на неизвестный язык, но Зализняк в своих задачах довёл процедуру внутренней реконструкции до очень высокого уровня строгости. Видимое разнообразие форм

того или иного языка сводится к глубинным представлениям, в которых одному значению соответствует одна форма и из которых реально наблюдаемая картина получается при помощи упорядоченных формул перехода. С конца 1960-х такой подход почти на 20 лет стал доминирующим в фонологии (это область лингвистики, которая изучает, как люди обрабатывают звуки в своём сознании, и стремится строго это описать); и в наши дни он продолжает использоваться наряду с другими моделями.

Самодостаточные лингвистические задачи впервые были предложены школьникам на Первой традиционной олимпиаде по языковедению и математике, которая прошла в 1965 году. Некоторые смеялись: «Что это такое — первая традиционная?». А она действительно стала традиционной, как того и хотел В. А. Успенский, включивший это слово в название. В 2017/18 учебном году олимпиада прошла в 48-й раз и была посвящена памяти Зализняка — создателя первых самодостаточных задач на стыке языка и логики. А с 2003 года проводится и Международная олимпиада по лингвистике, в которой участвуют уже три десятка стран: от Швеции до Бразилии.

Зализняк сочинил для олимпиады 26 задач и много лет активно участвовал в обсуждении и подготовке задач других авторов. И даже в те годы, когда обилие прочих дел уже не позволяло ему ходить на регулярные заседания задачной комиссии, все сочиненные задачи полагалось «озализнячивать» — показывать Зализняку перед тем, как давать школьникам.

История русского ударения

В 1970-е годы Зализняк обратился к истории русского языка и впоследствии плодотворно занимался ею более 40 лет. Изучение схем ударения в современном русском языке привело его к тому, что он решил разобраться, откуда они взялись, и построил строгую формальную модель, которая позволяет по составу древнерусского слова определить, какое в нём было ударение.

Рассмотрим современные русские слова *были* и *рыли*. В них обоих ударение падает на первый слог — но и в этом случае, как и в примере про *фонарь*, всё сложнее, чем кажется. Добавим отрицание и получим: *не были* и *не рыли*. У слова *рыли* — своё собственное, «настоящее» ударение, которое оно никому не отдаёт. А вот у слова *были* ударение «ненастоящее»; скажем, что

у него как бы нет ударения, а просто автоматически усиливается первый слог; если же первым слогом становится частица *не*, она и получает это автоматическое усиление. Для нас это неожиданно: ведь ударение в *бы́ли* звучит так же, как в *ры́ли*, но Зализняк установил, что в древнерусском языке около 1000 года это различие было вполне живым. Имелось два типа ударения, различающихся по звучанию и совпавших только к XIV веку: автономное (самостоятельное, «настоящее») ударение, которое могло падать на любой слог слова, и автоматическое («ненастоящее»). Автоматическое ударение — это и есть то самое усиление на первом слоге. Условно будем обозначать автономное ударение в древнерусских словах привычным способом (*ры́ли*), а автоматическое — чёрточкой перед словом (*̀были*).

Возьмём три тройки существительных (возможно, с предлогом), которые до наших дней сохранили древнерусское ударение: *гру́ша — гру́шу — за гру́шу*, *женá — жену́ — за жену́* и *ногá — нóгу — за нóгу*. Вспоминается что-то знакомое: ударение всегда на основе, ударение всегда на окончании и подвижное ударение — это те же самые схемы, которые мы уже видели в работах Зализняка про современный русский язык. В своей книге «От праславянской акцентуации к русской» (1985) он наглядно объяснил, откуда и как эти схемы произошли.

Припишем каждой морфеме (корню, окончанию и т. п., а также предлогам, союзам и частицам) одну из трёх маркировок: ↓, → или —. Договоримся, что ударение определяется самой левой стрелкой. Если это ↓, то автономное («настоящее») ударение падает на отмеченный ею слог; если это →, то автономное («настоящее») ударение падает на слог правее неё; если же стрелок в слове нет, а есть только минусы, слово имеет автоматическое («ненастоящее») ударение. Для того чтобы описать все приведённые выше примеры, достаточно разметить морфемы так: *за —, груш ↓, жен →, ног —, а ↓, у —*. После этого получаем все реально наблюдаемые ударения:

<i>гру́ш-а</i>	<i>гру́ш-у</i>	<i>за гру́ш-у</i>
↓ ↓	↓ —	— ↓ —
<i>жен-á</i>	<i>жен-у́</i>	<i>за жен-у́</i>
→ ↓	→ —	— → —
<i>ног-á</i>	<i>̀ног-у</i>	<i>̀за ног-у</i>
— ↓	— —	— — —

Этот простой алгоритм (с некоторыми добавлениями, которые, разумеется, изложены в работах Зализняка) позволяет полностью описать древнерусское ударение. А чтобы потренироваться, сделайте небольшое упражнение.

Упражнение I

Даны пять глаголов в двух формах прошедшего времени: *могла́, могли́; мы́ла, мы́ли; жи́ла, жи́ли; вы́мыла, вы́мыли; вы́жила, вы́жили*. Известно, что место современного ударения во всех этих формах совпадает с древнерусским.

Задание. Для каждой морфемы, содержащей хотя бы один гласный (приставка *вы-*; корни *-мог-*, *-мы-*, *-жи-*; окончания *-а*, *-и*), укажите её древнерусскую маркировку (↓, → или –) так, чтобы получить правильное ударение во всех формах.

Примечание. Поскольку суффикс прошедшего времени *-л-* не содержит гласного, ему можно не приписывать маркировку или приписать маркировку –.

Как определить ударение по рукописи

Схема, представленная выше, описывает древнерусскую ситуацию, которая затем заметно изменилась. Мы и сами это замечаем: например, сейчас вполне можно сказать *за но́гу*. В своих работах Зализняк очень подробно описал, как изменялось древнерусское ударение. Для этого он проанализировал 89 старинных рукописей и печатных книг с проставленными ударениями. Ударения ставились в текстах далеко не всегда — но иногда получить информацию об ударении можно даже в случае, если ударение не проставлено.

Сделаем небольшой экскурс в историю русского языка. Наряду с «обычными» гласными, в древнерусском языке около 1000 года встречались два «особых» сверхкратких гласных: *ѣ* (как *ы*, но короче) и *ь* (как *и*, но короче). Например, слова *сто́лѣ*, *го́дѣ* и *мѣхѣ* ‘мох’ были двусложными (*сто-лѣ*, *го-дѣ*, *мѣ-хѣ*), слова *го́лубь*, *ну́жно* ‘нужно’, *по́льза*, *тѣ́мно* ‘темно’, *шьвьѣць* ‘швец’ — трехсложными (*го-лу-бѣ*, *ну-жѣ-но*, *по-ль-за*, *тѣ-мѣ-но*, *шь-вь-ѣць*). В XI—XIV веках с *ѣ* и *ь* происходило изменение, которое можно описать так³:

3 Эта закономерность была открыта ещё до Зализняка и часто называется законом Гавлика в честь чешского лингвиста Антонина Гавлика (1855—1925).

- 1) в конце слова гласные *ѣ* и *ь* выпадают⁴;
- 2) далее движемся по слову справа налево; если мы находим *ѣ* или *ь*, то смотрим на ближайший гласный справа от него:
 - а) если это выпавший *ѣ* или *ь*, то рассматриваемый на данном шаге *ѣ* или *ь* проясняется, т. е. превращается в *о* или *е* соответственно;
 - б) в любом другом случае (т. е. если справа стоит «обычный» гласный или прояснившийся *ѣ* или *ь*) *ѣ* и *ь* выпадают.

В табл. 6 продемонстрировано, как эти правила работают (каждую строку следует читать справа налево)⁵.

Таблица 6

Стало	3-й слог с конца	2-й слог с конца	1-й слог с конца	Было
<i>стол</i>		<i>сто</i>	<i>лъ</i> выпадает (1)	<i>столъ</i>
<i>год</i>		<i>го</i>	<i>дѣ</i> выпадает (1)	<i>годѣ</i>
<i>мох</i>		<i>мѣ</i> <i>ѣ</i> → <i>о</i> (2а)	<i>хѣ</i> выпадает (1)	<i>мѣхѣ</i>
<i>голубь</i>	<i>го</i>	<i>лу</i>	<i>бѣ</i> выпадает (1)	<i>голубѣ</i>
<i>нужно</i>	<i>ну</i>	<i>жѣ</i> выпадает (2б)	<i>но</i>	<i>нужно</i>
<i>польза</i>	<i>по</i>	<i>ль</i> выпадает (2б)	<i>за</i>	<i>польза</i>
<i>темно</i>	<i>ть</i> <i>ь</i> → <i>е</i> (2а)	<i>мь</i> выпадает (2б)	<i>но</i>	<i>тьменно</i>
<i>швец</i>	<i>шь</i> выпадает (2б)	<i>вь</i> <i>ь</i> → <i>е</i> (2а)	<i>ць</i>	<i>шьвьць</i>

Видно, что в результате этого процесса появилось много слов, которые оканчиваются на согласный: *мох*, первый слог в *поль-за* и т. п.

4 Буква *ѣ* продолжает сохраняться на письме до 1918 года, а *ь* во многих случаях и до сих пор, обозначая мягкость согласного.

5 Согласный, который стоял перед выпавшим *ь*, иногда сохраняет мягкость (*голубѣ*, *польза*), а иногда нет (*нужно*, *темно*, *швец*). Для нашего изложения этот вопрос несущественен.

Упражнение 2

Проверьте себя: какие слова (не обязательно в начальной форме) получились из древнерусских слов *днь, львъ, сѣнъ, пышено, лѣжьска, почѣсть, льстимъ, льсть, чьтъць, Смольньску?*

А теперь, после этого экскурса, перенесёмся в XIV век. Именно тогда была написана рукопись, которая благодаря Зализняку стала известна как ценный источник знаний об ударении — справочник для судей под названием «Мерило праведное» (это значит «Весы правосудия»). Знаков ударения в рукописи нет, но зато различаются два типа *о*: «закрытое *о*», которое обозначалось буквой *ω* («омега») и звучало как нечто среднее между *о* и *у*, и «открытое *о*», которое обозначалось буквой *o* («он») и звучало как нечто среднее между *о* и *а*. Закрытое *о* возникло из раннедревнерусского *o* (причём только из «обычного» *o*, а не из *ъ*) в трёх случаях:

- 1) в начальном слоге, который ранее оканчивался на гласный, а после выпадения *ъ* и *ь* стал оканчиваться на согласный: *гωдъ* (было двусложное *˘го-дъ*, которое превратилось в односложное *год*), *тωч-но* (было *тѣ-чь-но*); назовём такой слог перестроенным;
- 2) под автономным ударением: *зλω* (*зѣлѡ*), *поклωнъ* (*поклѡнъ*);
- 3) на два слога правее автономного ударения (по счёту слогов уже после выпадения *ъ* и *ь*): *обладаемωе* (*обладаѣмое*) *вωинствω* (*вѡинство*).

В остальных случаях в «Мериле» раннедревнерусское *o* даёт открытое *o*: *гoду*, *зoлото* (*˘гoду*, *˘зoлото* — автоматическое ударение, начальные слоги оканчиваются на гласный), *мoлодa* (*мoлодa* — оба *o* безударные, начальный слог оканчивается на гласный). Если *ъ* превратился в *o*, это всегда открытое *o*: *мохъ* (*мѣхъ*) — не помогает ни то, что начальный слог перестроенный, ни автоматическое ударение.

Зная всё это, как нам понять ударение по написанию в «Мериле праведном», при том что знаков ударения там нет? Например, мы видим в «Мериле» слово *теплωтою* (из раннедревнерусского *теплотою*). Могло ли там быть автоматическое ударение (*˘теплотою*)? Нет, потому что тогда во втором слоге не появилась бы *ω*. А могло ли там быть автономное ударение на 1-м, 3-м или 4-м слоге: *тѣплωтою*, *теплѡтѡю* или *теплωтѡю*? И опять-

таки нет, потому что не появилась бы ω в нужном месте: *тѣплотою* и *теплотою* дали бы *теплотою* (проверьте, понимаете ли вы, почему в этих двух случаях написание совпадёт), а *теплотою* вовсе не дало бы буквы ω . Значит, это слово имело ударение *теплотою*, непривычное для нас сегодня.

А какое было ударение в слове *старость* (из раннедревнерусского *старость*)? Точно определить не удаётся, но у нас есть основания сказать, что там могло быть автоматическое ударение ($\bar{\text{старость}}$) или автономное ударение на 1-м слоге (*ста́рость*)⁶, но только не автономное ударение на 2-м слоге: из *ста́рость* в «Мериле» получилось бы *ста́рѡсть*. В слове *золоту* (из *золоту*) могло быть автоматическое ударение ($\bar{\text{золоту}}$) или автономное ударение на 3-м слоге (*золоту́*), но не автономное ударение на 1-м или 2-м слоге (*зо́лоту* дало бы *зѡлоту*, а *золоту́* — *золѡту*). В слове *пѡпѣ* (из *попѣ*) может быть любое ударение: если там было $\bar{\text{попѣ}}$, то ω получится по правилу про перестроенный слог, а если было *по́пѣ*, то сразу по двум правилам — про перестроенный слог и про автономное ударение.

Зализняку было труднее, чем вам: ему пришлось выявить все эти правила, читая рукопись. Но и вы можете попробовать повторить заключительную часть его открытия и решить следующую задачу.

Задача

Пусть вам известно написание слова в «Мериле праведном» (XIV век) и его раннедревнерусское написание, отражающее произношение около 1000 года. Постройте алгоритм, который позволяет определить, каким могло, а каким не могло быть древнерусское ударение в этом слове. Иначе говоря, для слова, которое в «Мериле праведном» является *n*-сложным, вам надо сказать, могло ли в нём быть: автоматическое ударение; автономное ударение на 1-м, 2-м, ..., *n*-м слоге.

Чтобы вы могли решать эту задачу не чисто умозрительно, а на реальном материале, вот несколько слов в обоих видах: *истиненѣ* (*истиненѣ*), *можемѣ* (*можемѣ*), *живѡтъ* (*живѡтъ*), *подѡбалѡ* (*подѡбалѡ*), *свѡбоденѣ* (*свѡбодѣнѣ*), *свѡбоднѡму* (*свѡбодѣному*),

⁶ Разумеется, если для какого-то слова у нас получилось несколько вариантов ударения, это не значит, что все они реально встречались: скорее всего, использовался только один из них, но мы не можем по написанию надёжно определить, какой именно.

золото (золото), морьскую (морьскую), погребенѡ (погребено), свитокъ (сѡвитѡкъ), полезно (пользьно), золь (зѡль).

Указание. Помните, что ко времени написания «Мерила праведного» буквы ѡ и ѡ уже не читаются как гласные и о них при определении ударения думать не надо.

Подсказка 1. Ровно для пяти из данных слов вы точно можете определить древнерусское ударение (т. е. доказать, что все варианты ударения, кроме одного, невозможны).

Подсказка 2. Ровно про три из данных слов вы точно можете сказать, что место ударения в них не такое же, как в современном русском языке.

* * *

Вот уже почти десять лет я бережно храню один листок бумаги. В 2008/09 году Зализняк читал в МГУ имени М. В. Ломоносова курс по истории русского ударения и дал на дом задание — то, что я только что предложил вам как задачу. Проверив мою работу, он исправил одну опечатку и написал:

*Превосходно. Безукоризненно.
А.З.*

С тех пор, читая книги и статьи самого Андрея Анатольевича и слушая его лекции, я всегда думал и продолжаю думать, что эти слова в первую очередь характеризуют не моё домашнее задание, а его самого и всё то, что он делал. Превосходно. Безукоризненно.

Март 2018 г.

А. Ч. Пиперски

ОТВЕТЫ К УПРАЖНЕНИЯМ И РЕШЕНИЕ ЗАДАЧИ

Упражнение 1. *вы-* ↓, *-мог-* →, *-мы-* ↓, *-жи-* —, *-а* ↓, *-и* —.

Упражнение 2. *день, лев, сон, пшено, ложка, почесть, льстим, леть, чтец, Смоленску.*

Задача. Шаг 1 (классификация омег). Разделим все буквы ω на два класса: свободная ω и несвободная ω . Назовём несвободными такие ω , которые стоят в 1-м перестроенном слоге, и свободными — все остальные.

Шаг 2 (первичная генерация вариантов). Если в записи «Мерила» есть хотя бы одна свободная ω , слово может иметь только автономное ударение на слоге с самой левой свободной ω или на два слога левее (если такой слог есть). Если в слове нет ни одной свободной ω , слово может иметь любое ударение (автоматическое или автономное на любом слоге)

Шаг 3 (ограничение возможностей). Если в записи «Мерила» есть хотя бы одна буква o на месте раннедревнерусского o , для каждой такой буквы из возможностей, полученных в шаге 2, следует исключить автономное ударение на этом слоге и на два слога левее (если такой слог есть).

Рассмотрим несколько примеров:

Написание «Мерила»	Шаг 1 (свободные ω подчёркнуты)	Шаг 2 (отмечены все варианты)	Шаг 3 (отмечены все варианты)	Число вариантов
<i>истиненъ</i>	<i>истиненъ</i>	<i>̀истиненъ</i>	<i>̀истиненъ</i>	4
<i>золото</i>	<i>золото</i>	<i>̀золото</i>	<i>̀золото</i>	1
<i>погребенω</i>	<i>погребен<u>ω</u></i>	<i>погрѐбенω</i>	<i>погрѐбенω</i>	2
<i>подωбалω</i>	<i>под<u>ω</u>бал<u>ω</u></i>	<i>подѡбалω</i>	<i>подѡбалω</i>	1
<i>свободенъ</i>	<i>свободенъ</i>	<i>̀свободенъ</i>	<i>̀свободенъ</i>	2
<i>свитокъ</i>	<i>свитокъ</i>	<i>̀свитокъ</i>	<i>̀свитокъ</i> ⁷	3

Однозначно определяется ударение в словах *мѡжемъ, живѡтъ, подѡбалω, свободнѡму* и *̀золото*. В словах *подѡбалω* и *свободнѡму* ударение отличается от современного (*подѡбало* и *свободнѡму*); в слове *̀свободенъ* мы не можем однозначно определить ударение, но оно в любом случае отлично от современного *свободен*⁸.

⁷ Так как здесь o из ω , исключить ударение на o нельзя.

⁸ Из других источников известно, что на самом деле было ударение *̀свободенъ*.

ЛИНГВИСТИЧЕСКИЕ ЗАДАЧИ

Важным средством обучения основным положениям и методам языкознания могут служить специально составленные задачи¹. В существующих сборниках в качестве материала для задач в большинстве случаев используются факты родного языка учащихся или наиболее известных европейских языков. Такие задачи, безусловно, полезны, но, к сожалению, они часто страдают тем недостатком, что в них трудно отделить собственно лингвистическое задание (не требующее ничего, кроме понимания основных лингвистических положений) от проверки знания конкретных фактов рассматриваемого языка. Наилучший (хотя отнюдь не единственный) способ избавиться от этого второго элемента задания, не имеющего прямого отношения к общему языкознанию, состоит в том, чтобы составлять задачи на материале языков, незнакомых учащемуся². Разумеется, составлять такие задачи труднее, поскольку все существенные для решения конкретные факты должны быть так или иначе представлены в исходных данных задачи, зато от учащегося в этом случае требуется только представление о свойствах языка вообще.

Ниже предлагается серия лингвистических задач, рассчитанных на читателей, незнакомых с рассматриваемыми языками.

Первую группу составляют задачи на грамматический анализ текста на незнакомом языке. Читатель должен изучить формальные особенности строения текста на незнакомом языке и на основании своего анализа выполнить контрольное задание (найти ошибку в тексте или перевести контрольные фразы), которое

1 Укажем наиболее интересные сборники задач: *И. А. Бодуэн де Куртене*. Сборник задач по «Введению в языковедение» по преимуществу применительно к русскому языку. СПб, 1912; *Л. Р. Зиндер*. Сборник задач по общему языкознанию. Л., 1957; *H. A. Gleason*. Workbook in descriptive linguistics. N. Y., 1955; *W. P. Lehmann*. Exercises to accompany «Historical Linguistics». N. Y., 1962.

2 По этому принципу составлены почти все задачи в указанной выше книге Г. Глисона.

должно показать, насколько правильно ему удалось выявить эти закономерности.

Задачи этого типа представляются нам интересными прежде всего потому, что для их решения недостаточно чисто формальных логических операций: для того, чтобы приступить к этим операциям, человек обязательно должен исходить из некоторого общего представления о строении всякого языкового текста. Если в результате он получает решение, верное с точки зрения реального языка, это следует рассматривать как косвенное подтверждение его исходных представлений (принцип «черного ящика»).

Дело, однако, в том, что исходное представление о тех или иных свойствах языка обычно не формулируется явно, а остается на интуитивном уровне. Более того, человек может полагать, что он и не использовал при решении ничего, кроме логических операций. Что это не так, легко показать, например, на задаче 1, где требуется найти грамматическую ошибку в тексте: если бы мы не знали, что перед нами фразы реального языка, и могли рассматривать текст как произвольную последовательность знаков или как фразы на искусственном языке, который может быть устроен как угодно, то, разумеется, о поиске ошибки не могло бы быть и речи.

Несомненно, что выявление и формализация интуитивных посылок такого рода, являющихся элементами правильного осознания человеком своего языка, очень важны для языкознания. Очевидно также, что ограниченный материал искусственной задачи создает благоприятные условия для выявления таких посылок. Поэтому составление задач данного типа можно рассматривать и как способ экспериментального изучения языковой интуиции человека.

Вторую группу составляют задачи на внутреннюю реконструкцию. В каждой из них приводится небольшая часть системы словоизменения (несколько глаголов в нескольких грамматических формах). Эта подсистема с внешней стороны отличается малой степенью регулярности: одно и то же грамматическое значение выражается многими способами, один и тот же глагол имеет несколько вариантов корня и т. д. Требуется построить систему, обладающую существенно большей степенью регулярности, которую можно было бы рассматривать как первоначальный вид заданной системы, подвергшийся впоследствии серии «фонети-

ческих изменений» (цепь формул перехода, выражающих эти изменения, также должна быть указана в решении).

В каждой задаче этой группы задание ставится строго формально, без учета каких бы то ни было фонетических соображений. Единственным критерием для оценки решения является простота полученной «квазиалгебраической» конструкции, выражаемая длиной цепи формул перехода. Тем не менее — и в этом, как нам представляется, заключена одна из важных общелингвистических закономерностей — самые простые «квазиалгебраические» решения оказываются обычно наиболее точными отражениями диахронического развития. Как известно, именно такая, «алгебраическая» постановка задачи привела Ф. де Соссюра к одному из наиболее замечательных открытий в области индоевропейской фонетики. (Факты, проанализированные Ф. де Соссюром, составляют материал последних задач этой группы.)

Автор приносит искреннюю благодарность В. А. Успенскому и Е. В. Падучевой за конструктивную критику настоящей работы. Автор благодарит А. А. Санчеса за помощь в составлении арабского текста и Г. Э. Влэдуца — за проверку венгерского текста. Автор глубоко признателен также М. И. Белецкому, А. Б. Долгопольскому, А. Л. Крылову, М. М. Ланглебен, Р. А. Минлосу, Т. М. Николаевой, И. И. Ревзину, В. М. Тихомирову, Б. А. Успенскому, Г. С. Цейтину, Т. В. Цивьян, Ю. А. Шихановичу, С. М. Шур и всем другим, любезно представившим ему свои решения публикуемых ниже задач.

Группа I

ЗАДАЧИ НА ГРАММАТИЧЕСКИЙ АНАЛИЗ

Общие указания

1. Все задачи этой группы направлены на выявление грамматических закономерностей в тексте на незнакомом языке. В задаче 1 требуется найти и исправить грамматически неправильную фразу в одноязычном тексте. В задачах 2—4 требуется на основе анализа двуязычного текста перевести контрольные фразы на незнакомый язык (с русского языка в задаче 2, с другого незнакомо языка — в задачах 3 и 4).

Читатель должен иметь в виду, что в этих задачах от предлагаемого им решения требуется не то, чтобы оно соответствовало

какому-то числу закономерностей, которые он обнаружил, а то, чтобы оно было правильно с точки зрения данного реального языка, т. е. чтобы оно соответствовало *всем* грамматическим закономерностям этого языка. Как показывает опыт, во всех задачах такое решение может быть получено; следует лишь остерегаться поспешных решений.

2. Обращаем внимание на то, что всякая буква с надстрочным или подстрочным знаком есть *особая* буква, отличная от соответствующей простой³. Самостоятельными буквами являются также знаки ' и ' в арабском тексте (задача 2) (ими обозначаются определенные арабские согласные).

3. В задачах 1, 3 и 4 читателю может быть предложено следующее *дополнительное задание*: построить наиболее вероятную гипотезу о грамматической структуре приведенных в задаче фраз (т. е. предложить разбор этих фраз по членам предложения, распределить слова по частям речи, определить роль отдельных морфем в многоморфемных словах). В задаче 2 лицам, изучающим методы лингвистического описания, может быть предложено следующее дополнительное задание: составить полное описание языка данного текста (в частности, рекомендуется: составить полный список морфем, указав правило распределения алломорфов там, где они есть; выявить морфологические группы слов; установить, какие грамматические категории имеет каждая группа и какие противопоставления эти категории включают; сформулировать правила порождения слов из морфем и предложений из слов).

Задача I («gizona»)

Предназначается для лиц, незнакомых с баскским языком.

Исходные данные

Дан текст из 12 фраз на незнакомом языке (баскском). Известно, что одна из фраз грамматически неправильна из-за ошибки в одном слове (в более строгой форме: из-за того, что в одном случае одна последовательность букв между пробелами заменена некоторой иной последовательностью букв).

3 Исключение составляет знак ударения, но в задачах этой группы он не встречается; знак ' в венгерской графике (задача 4) не является знаком ударения.

Текст

1. Gizona joaten da.
2. Gizonak zaldia ikusten du.
3. Astoa atzo joaten zan.
4. Gizonak atzo joaten ziran.
5. Astoak zaldia atzo ikusten zuen.
6. Zaldia gizona ikusten du.
7. Zakurrak joaten dira.
8. Gizonak zakurra atzo ikusten zuen.
9. Zakurrak astoak ikusten ditu.
10. Zaldia gizonak atzo ikusten zituen.
11. Zakurra atzo joaten zan.
12. Gizonak astoak atzo ikusten zituen.

Задание

Найти грамматически неправильную фразу и сделать ее грамматически правильной, изменив (или заменив) в ней только одно слово.

Задача 2 («samīrun») ⁴

Предназначается для лиц, незнакомых с арабским языком.

Исходные данные

Дан текст из 12 фраз на незнакомом языке (арабском) с *пословным* переводом на известный язык (русский).

ПОЯСНЕНИЯ К ФОРМЕ ЗАПИСИ. Арабский текст дан в латинской транслитерации. Под каждым словом арабского текста записана соответствующая ему часть буквального русского перевода (которая может состоять из одного или нескольких слов); начало этой части всегда находится точно под началом переводимого арабского слова. Арабские показатели определенности и неопределенности (соответствующие артиклям западноевропейских языков) условно переданы в переводе словами *этот* (опре-

⁴ Эта и следующая задачи были предложены осенью 1960 года участникам семинара по математической лингвистике на механико-математическом факультете МГУ под руководством А. А. Маркова, В. А. Успенского и автора настоящей работы.

деленность) и *один* или *некоторый* (неопределенность); эти слова выделены курсивом. Так, например, в записи «эта *эта* ...» первое слово соответствует арабскому указательному местоимению, а второе — показателю определенности. Слова, необходимые для ясности перевода, но не имеющие прямого соответствия в арабском тексте, даны в скобках. Таким образом, курсивом даны слова, которые в обычном (не буквальном) переводе были бы опущены, в скобках же — слова, которые были бы вставлены.

Текст

1. yas'alu samīrun sāmī'ahu:
спрашивает *один* рассказчик своего слушателя
2. 'ata'lamu ḳiṣṣata ṣabīyin waḡinnīyatin?
знаешь ли историю *одного* юноши и *одной* волшебницы
3. yaḳūlu ssāmī'u lissamīri:
говорит *этот* слушатель *этому* рассказчику
4. mā 'a'lamuhā, yā samīrī, wa'asma'uka.
не знаю ее о мой рассказчик и слушаю тебя
5. faḳaṣṣa ssamīru hādihi ḳiṣṣata:
и рассказал *этот* рассказчик эту *эту* историю
6. ḳālat ḡinnīyatun liṣabīyin
сказала *одна* волшебница *одному* юноше
7. 'as'aluka: 'ataryabu ṣafḳatan?
спрашиваю тебя желаешь ли *одну* сделку
8. fasa'alahā ṣṣabīyu: mā hādihi ṣṣafḳatu?
и спросил ее *этот* юноша какова эта *эта* сделка
9. faḳālat liṣṣabīyi: tamliku mi'ata
и сказала (она) *этому* юноше будешь иметь сто
- 'alfi dirhamin, fa'aḳluka yaḏhabu.
тысяч *некоторых* дирхемов а ум твой пропадет
10. faḏaḥika waḳāla lilḡinnīyati: lā, mā
и засмеялся (он) и сказал *той* волшебнице нет не
- 'aryabu ṣafḳataki.
желаю твоей сделки

11. fasa'alathu lġinnīyatu: lima taqūlu
и спросила его *эта* волшебница почему говоришь
hakaḍā walima taḍḥaku?
так и почему смеешься

12. faḳāla: yaḍhabu dirhamun ba'da
и сказал (он) пропадет *один* дирхем за
dirhamin fayadhabu mālī ḡamī'an
одним дирхемом и пропадет мое богатство целиком
li'adami wuġūdi l'aḳli fa'amliku
из-за отсутствия наличия *этого* ума и буду иметь
ḥumḳī faḳaṭ
мою глупость только

Задание

Перевести с русского языка на арабский следующие две фразы:

1. Рассказала эта *эта* волшебница историю своей сделки, и засмеялся ее слушатель.
2. Знает рассказчик твой, о мой юноша, тысячу *некоторых* историй, а не имеет (и) *одного* дирхема.

ПРИМЕЧАНИЕ. Система записи здесь та же, что в русском переводе арабского текста; следовательно, при переводе на арабский язык порядок слов должен быть сохранен, слово, стоящее в скобках, должно быть опущено, и т. д.

Задача 3 («tizë ri»)

Предназначается для лиц, незнакомых с албанским и древнееврейским языками.

Исходные данные

Дан текст из 6 фраз на незнакомом языке А (албанском) с переводом каждой фразы на незнакомый язык В (древнееврейский).

Албанский текст дан в обычной орфографии. Для древнееврейского текста дана латинская транслитерация консонантической записи (т. е. записи без обозначения гласных).

Текст

Язык А	Язык В	Язык А	Язык В
1. Mizë pi.	— yšth zbwb.	4. Mizat pinë.	— yštw hzbwbym.
2. Miza pinin.	— štw zbwbywm.	5. Miza pinë.	— yštw zbwbywm.
3. Mizë pinte.	— šth zbwb.	6. Miza pi.	— yšth hzbwb.

Задание

Перевести с языка В на язык А следующие две фразы:

1. šth hzbwb.
2. štw hzbwbym.

Задача 4 («dezu»)

Предназначается для лиц, незнакомых с баскским и венгерским языками.

Исходные данные

Дан текст из 14 фраз на незнакомом языке А (баскском) с переводом каждой фразы на незнакомый язык В (венгерский).

Текст

Язык А	Язык В
1. Agindutzen dezu.	— Ezt igéred.
2. Urdea billatu dezute.	— Kerestétek a sertést.
3. Billatu diozu.	— Ezt kerested neki.
4. Agindutzen diotezu.	— Ezt igéred nekik.
5. Agindutzen diozute.	— Ezt igéritek neki.
6. Agindu dizkiotezute.	— Ezeket igértétek nekik.
7. Arrek agindu dizu.	— Ezt igérte neked.
8. Arrek urdea billatutzen dizute.	— Keresi nektek a sertést.
9. Arrek agindutzen dizkizu.	— Ezeket igéri neked.
10. Aek labanak billatu dizkizute.	— Keresték neked a késeket.
11. Aek billatutzen diote.	— Ezt keresik neki.
12. Arrek urdeak agindu dizkio.	— Igérte neki a sertéseket.
13. Arrek billatutzen dizkiote.	— Ezeket keresi nekik.
14. Aek agindu dizkiote.	— Ezeket igérték neki.

Задание

Перевести с языка В на язык А следующие фразы:

1. Igéri neki a kést.
2. Keresed nekik a sertéseket.
3. Ezt ígérték neked.

Группа II

ЗАДАЧИ НА ВНУТРЕНнюю РЕКОНСТРУКЦИю

Общие указания

1. Все задачи этой группы имеют следующее строение.

Задается некоторая совокупность («система») словоформ. Она приводится в виде таблицы, в которой по горизонтали располагаются словоформы с одинаковым лексическим значением, а по вертикали — словоформы с одинаковым грамматическим значением.

Требуется построить другую систему словоформ, называемую «первообразной» (для данной системы), которая обладала бы следующими свойствами:

1) Она содержит столько же словоформ, сколько заданная система, и эти словоформы могут быть представлены в виде такой же таблицы (включающей те же лексические и грамматические значения). Таким образом, каждой словоформе заданной системы соответствует некоторая словоформа «первообразной» системы, имеющая то же лексическое и то же грамматическое значение.

2) Строение словоформ в этой системе удовлетворяет определенным требованиям, которые указываются отдельно в каждой задаче.

3) Существует такая «цепь формул перехода» (см. следующий пункт), которая, будучи применена к «первообразной» системе, преобразует ее в заданную. (В решении «цепь формул перехода» должна быть полностью приведена.)

Для записи словоформ «первообразной» системы можно использовать любой исходный алфавит (в частности, можно вводить в имеющийся алфавит любые дополнительные буквы).

2. Что понимается под «формулой перехода» и «цепью формул перехода»?

Рассмотрим вначале случай, когда место ударения не обозначается.

В этом случае под «формулой перехода» понимается правило следующего общего вида: последовательность букв (или букв и пробелов) α преобразуется в другую последовательность букв (или букв и пробелов) β ⁵; запись: $\alpha \rightarrow \beta$.

Примеры формул перехода: $o \rightarrow u$, $h \rightarrow \text{нуль}$, $ee \rightarrow i$, $d\# \rightarrow t\#$, $кы \rightarrow ки$, $нр \rightarrow ндр$, $est \rightarrow \acute{e}t$.

Число букв и пробелов в последовательности α мы будем называть «глубиной взаимодействия» в данной формуле перехода. Так, в первых двух примерах глубина взаимодействия 1, в последнем — 3, во всех остальных — 2.

Применить формулу перехода к системе словоформ — значит осуществить данное преобразование во всех словоформах, содержащих⁶ последовательность α . Если последовательность α входит в словоформу более одного раза, вначале преобразованию подвергается первое (самое левое) вхождение α ; затем, если в измененной словоформе все еще имеются вхождения α , преобразованию подвергается опять-таки самое левое из них, и т. д.

Например, применяя формулу перехода $ee \rightarrow i$ к системе словоформ *see*, *seee*, *seeree*, мы получим *si*, *sie* (не *sei!*), *sipi*.

ПРИМЕЧАНИЕ. Легко видеть, что «формула перехода» — это не что иное, как формальный аналог фонетического изменения; по понятным причинам в нашем изложении фигурируют буквы, а не звуки.

Несколько формул перехода, расположенных в определенном порядке, образуют «цепь формул перехода», например:

1. $ки \rightarrow чи$
2. $кы \rightarrow ки$

Цепь формул перехода применяется к системе словоформ так: вначале к системе применяется первая формула цепи, после че-

5 Последовательность β может быть также пустой, т. е. не содержать ни букв, ни пробелов. Заметим, что не следует смешивать пустую последовательность («нуль») и пробел: так, в последовательности *я здесь* между буквами *я* и *з* пробел, а в последовательности *язь* — нуль. Пробел обозначается в формулах перехода знаком #; нуль в соседстве с другими буквами не обозначается, а в изолированном виде обозначается записью «нуль».

6 Условимся считать, что пробелы в начале и в конце словоформы входят в нее. Это даст нам возможность говорить, что, например, словоформа *волк* содержит последовательности *#в* и *к#*.

го она уже ни в каких дальнейших операциях не используется; затем применяется вторая формула, и так далее до конца цепи ⁷.

Допустим, мы построили «первообразную» систему из трех словоформ: *волк, волкы, волчица*. Применяя к этой системе приведенную выше цепь формул перехода, получим: после 1-го перехода — *волк, волкы, волчица*; после 2-го (последнего) перехода — *волк, волки, волчица* ⁸.

Читателю могут быть рекомендованы следующие способы сокращения и обобщения записи.

Если часть последовательности α остается неизменной (т. е. не участвует в преобразовании, а только обуславливает его), можно использовать сокращенную запись. Пусть $\alpha = \mu\alpha'v$ и $\beta = \mu\beta'v$, где все буквы обозначают последовательности, причем α' или β' (но не обе сразу) может быть равно нулю. Тогда переход $\alpha \rightarrow \beta$ можно записать так: $(\mu+)\alpha'(+v) \rightarrow \beta'$. Таким сокращенным способом можно записать большинство формул, приведенных выше в качестве примеров: $\kappa(+u) \rightarrow \zeta$, $(\kappa+)_{\gamma} \rightarrow u$, $d(+\#) \rightarrow t$, $es(+t) \rightarrow \acute{e}$, $(n+)$ ноль $(+p) \rightarrow d$.

Две или несколько формул перехода, стоящих рядом в цепи и сходных по строению, при определенном условии (см. ниже) могут быть заменены одной так называемой обобщенной формулой. Обобщенная формула имеет вид $A \rightarrow B$ или $(M+)A'(+N) \rightarrow B'$, где каждая большая буква обозначает любую последовательность из некоторого списка (все такие списки должны быть тут же приведены); при этом должно быть указано, какой член списка B (или B') соответствует каждому члену списка A (или A').

Обобщенная формула применяется к системе словоформ так же, как обычная формула (см. выше), только в роли единичного объекта обработки выступает уже не вхождение последовательности α , а вхождение любой последовательности вида $A (= MA'N)$ (причем между вхождениями разных последовательностей вида A и разными вхождениями одной и той же последовательности не делается различия).

Примеры обобщенных формул:

$(K+)_{\gamma} \rightarrow u$, где $K = \kappa, \zeta$ или x [можно записать то же самое и без обобщающих символов:

$$\left(\begin{array}{c} \kappa \\ \zeta \\ x \end{array} \right) +_{\gamma} \rightarrow u];$$

7 Понятно, что тот же самый результат даст применение всей цепи формул перехода последовательно к каждой словоформе системы.

8 На приведенном примере видно, что порядок формул в цепи существен: при обратном порядке формул перехода мы получили бы другой конечный результат, а именно: *волк, волчи, волчица*.

$es (+C) \rightarrow \acute{e}$, где C — любая согласная⁹;

$(C+) V_1 (+V_2) \rightarrow$ нуль («гласная в зиянии исчезает»), где C — любая согласная, а V_1 и V_2 — любые гласные;

$\bar{V} (+C_1 C_2) \rightarrow V$ («гласная в закрытом слоге сокращается»), где C_1 и C_2 — любые согласные, \bar{V} — любая гласная со знаком долготы, V — та же гласная, что в левой части, но без знака долготы.

Условие, при котором обобщение допустимо, состоит в том, что применение обобщенной формулы к системе словоформ, для которой она предназначена, должно давать тот же результат, что и применение заменяемой ею цепи формул. Следует иметь в виду, что это требование нередко делает обобщение невозможным даже в простых на вид случаях. Ср., например,

$$\left. \begin{array}{l} \text{а) } 1. aw \rightarrow o \\ \quad 2. wa \rightarrow o \end{array} \right\} \begin{array}{l} \text{б) } aw \\ \quad wa \end{array} \rightarrow o$$

Можно ли заменить цепь а обобщенной формулой б? Оказывается, что если в системе словоформ к моменту применения этих формул имеется, например, словоформа aw , то такая замена недопустима: в случае а aw преобразуется в wo , а в случае б — в ow . Таким образом, при каждом обобщении формул необходимо проверить, соблюдено ли указанное основное требование.

В системе записи, где место ударения обозначается, действуют те же самые правила, но только к любой букве¹⁰, входящей в левую часть формулы, может быть присоединен показатель ударности (´) или безударности («безуд.»). Это значит, что преобразование происходит только при указанном положении ударения.

Примеры: $\acute{a} \rightarrow \bar{a}$ ¹¹;

$V_{\text{безуд.}} \rightarrow$ нуль, где V — любая гласная;

$(III+) o_{\text{безуд.}} \rightarrow e$, где III — $ш, ж, ч, щ$ или $ц$.

Если ни одна буква в левой части формулы не имеет показателей ударности или безударности, это значит, что преобразование не зависит от положения ударения.

3. Каждая задача этой группы допускает, вообще говоря, много решений, которые могут различаться как исходным набором

9 Здесь и далее под согласной (буквой) или гласной (буквой) подразумеваются члены определенных списков букв, которые должны быть предварительно заданы.

10 Практически, разумеется, только к гласной.

11 Ставить знак ударения и в правой части формулы нет необходимости, за исключением двусмысленных случаев вроде $\acute{a} \rightarrow \acute{e}i$ (ср. $\acute{a} \rightarrow ei$).

морфологических элементов, составляющих словоформы, так и цепями формул перехода, ведущими от «первообразной» системы к заданной. Эти решения не равноценны: лучшим считается то решение, при котором цепь формул перехода самая короткая.

Таким образом, после того, как получено некоторое решение, желательно проверить, не существует ли более короткого. При этом, чтобы легче было сравнивать разные цепи формул по длине, рекомендуется в каждой из них произвести все возможные обобщения.

4. Как и в задачах первой группы, следует иметь в виду, что буква с надстрочным или подстрочным знаком, кроме знака ударения, есть особая буква, отличная от соответствующей простой.

5. После того как задача решена, читателю, интересующемуся вопросами общей фонетики, может быть предложено (для любой из задач данной группы) следующее *дополнительное задание*: дать фонетическую интерпретацию найденного «квазиалгебраического» решения, согласующуюся с данными общей фонетики; если задача решена несколькими способами, — установить, какое из решений допускает наиболее приемлемую фонетическую интерпретацию.

ФОНЕТИЧЕСКИЕ УКАЗАНИЯ, необходимые для выполнения дополнительных заданий (для «трудных» букв латышского и древнеиндийского текстов указывается русская буква, передающая наиболее похожий звук).

Латышские буквы: *c* — *ц*, *š* — *ш*, *j* — *й*.

Буквы, используемые для транслитерации древнеиндийского письма: *c* — *ч*; *š* — *ш*; *y* — *й*; *j* — слитное *дж* (англ. *j*); *ç* — краткое *щ* (например, в словах *общность*, *женщина*); *ṭ*, *ḍ*, *ṇ* — разновидность *т*, *д*, *н* (сходны с английскими *t*, *d*, *n*); *ñ* — *нь*; *ṅ* — заднеязычное *н* (англ. *ng*); *h* — фрикативное *г* (например, в слове *aga!*)¹²; *r̥* — слогообразующее *р* (например, в словах *тембр*, *центр* при двусложном произношении).

Задача 5 («säkt»)

Предназначается для лиц, незнакомых с исторической грамматикой латышского языка.

12 После букв, обозначающих взрывные согласные, *h* не передает отдельной фонемы: пары букв типа *bh*, *dh*, *kh* обозначают единые придыхательные фонемы. В наших задачах, однако, этот факт не принимается во внимание и такие сочетания букв должны рассматриваться как обозначения пар фонем.

Исходные данные

Дано 9 латышских глаголов, каждый в семи грамматических формах (инфинитив и все лица единственного числа настоящего и прошедшего времени), итого 63 словоформы.

Инфинитив	Ед. ч. наст. врем.			Ед. ч. прош. врем.		
	1 л.	2 л.	3 л.	1 л.	2 л.	1 л.
sākt	sāku	sāc	sāk	sāku	sāki	sāka
nākt	nāku	nāc	nāk	nācu	nāci	nāca
jaukt	jaucu	jauc	jauc	jaucu	jauci	jauca
nest	nesu	nes	nes	nesu	nesi	nesa
vērst	vēršu	vērs	vērš	vērsu	vērsi	vērsa
mest	metu	met	met	metu	meti	meta
jaust	jaušu	jaut	jauš	jautu	jauti	jauta
sept	sepu	sep	sep	sepu	sepi	sepa
kāpt	kāpju	kāp	kāpj	kāpu	kāpi	kāpa

Задание

Построить для данной системы словоформ «первообразную» систему, обладающую следующими свойствами:

1. Каждая словоформа состоит из основы и окончания (окончание может быть нулевым). Основа состоит либо только из корня, либо из корня и суффикса; при этом суффикс во всех случаях один и тот же.

2. Каждая из семи грамматических форм имеет единое для девяти глаголов окончание.

3. У каждого глагола:

- корень имеет единый вид;
- основа инфинитива равна корню;
- основа всех форм настоящего времени едина, основа всех форм прошедшего времени — тоже, но основы разных времен могут и не совпадать.

ПРИМЕЧАНИЕ. Таким образом, в системе используется всего 17 элементов: 9 корней, 1 суффикс, 7 окончаний. Заметим, что некоторые из этих элементов могут быть омонимичны, т. е. совпадать по буквенному составу.

4. «Первообразная» система преобразуется в заданную некоторой цепью формул перехода (см. общие указания выше); глубина взаимодействия в каждой формуле перехода не должна превышать 2.

Задача 6 («yāmi»)

Предназначается для лиц, незнакомых с исторической грамматикой древнеиндийского языка.

Исходные данные

Дано семь древнеиндийских глаголов, каждый в шести грамматических формах (все лица единственного числа настоящего и прошедшего времени). Словоформы даны в латинской транслитерации без указания места ударения.

Ед. ч. наст. врем.			Ед. ч. прош. врем.		
1 л.	2 л.	3 л.	1 л.	2 л.	3 л.
yāmi	yāsi	yāti	ayām	ayāḥ	ayāt
vedmi	vetsi	vetti	avedam	avet	
rājmi	rākṣi	rāṣṭi	arājam	arāṭ	
bhajmi	bhakṣi	bhakti	abhajam	abhak	
vaçmi	vakṣi	vaṣṭi	avaçam	avaṭ	
vacmi	vakṣi	vakti	avacam	avak	
dveṣmi	dvekṣi	dveṣṭi	adveṣam	adveṭ	

Задание

Построить для данной системы словоформ «первообразную» систему, обладающую следующими свойствами:

1. Каждая словоформа состоит из корня, показателя лица и показателя времени; порядок этих элементов в разных словоформах может быть различным.
2. Показатель каждого из трех лиц един для всех восьми глаголов и для обоих времен.
3. Показатель каждого из двух времен един для всех восьми глаголов и для всех трех лиц.
4. Корень каждого глагола имеет единый вид.

ПРИМЕЧАНИЕ. Таким образом, в системе используется всего 12 элементов: 7 корней, 3 показателя лица и 2 показателя времени.

5. «Первообразная» система преобразуется в заданную некоторой цепью формул перехода (см. общие указания выше); глубина взаимодействия в каждой формуле перехода не должна превышать 3.

Задача 7 («yunákti») ¹³

Предназначается для лиц, незнакомых с исторической грамматикой древнеиндийского языка.

Исходные данные

Дано 12 древнеиндийских глаголов, каждый в 4 грамматических формах (3 л. единственного и множественного числа настоящего времени, причастие и инфинитив). Словоформы даны в латинской транслитерации с указанием места ударения (знак ´).

Наст. время		Причастие	Инфинитив
3 л. ед. ч.	3 л. мн. ч.		
yunákti	yuñjánti	yuktá	yóktum
bhinátti	bhindánti	bhinná	bhéttum
vṛṇákti	vṛñjánti	vṛktá	várktum
punáti	punánti	pūtá	pávitum
kṣiṇáti	kṣiṇánti	kṣiṇá	kṣáyitum
stṛṇáti	stṛṇánti	sṭirṇá	stáritum
rávati	rávanti	rutá	rótum
jáyati	jáyanti	jitá	jétum
bhárati	bháranti	bhṛtá	bhártum
bhávati	bhávanti	bhūtá	bhávitum
náyati	náyanti	nītá	náyitum
tárati	táranti	tīrṇá	táritum

Задание

Построить для данной системы словоформ «первообразную» систему, обладающую следующими свойствами:

1. Каждая словоформа может быть получена из корня по правилу, имеющему следующий общий вид:

¹³ Эта задача представляет собой переработанный и усложненный вариант задачи 4 из главы 6 книги: *W. P. Lehmann. Exercises to accompany «Historical Linguistics», N. Y., 1962.* В своей задаче У. Леман приводит факты, проанализированные Ф. де Соссюром в его знаменитом «*Mémoire sur le système primitif des voyelles dans les langues indo-européennes*», указывает решение де Соссюра и предлагает учащемуся проверить его и применить к другим аналогичным фактам. В настоящей задаче словоформы подобраны так, что становится возможным строго сформулировать проблему, впервые поставленную и разрешенную де Соссюром в этой книге. Таким образом, при решении этой задачи читатель должен самостоятельно «повторить» открытие де Соссюра.

а) К корню присоединяется такой-то аффикс или аффиксы. При этом для каждого аффикса должно быть указано его положение относительно корня: спереди (префикс), сзади (суффикс), внутри корня (инфикс); в последнем случае должно быть указано, между какими по счету буквами корня он вставляется. Для однородных аффиксов должен быть указан порядок присоединения.

б) Ударение падает на такую-то по порядку гласную букву¹⁴, такого-то морфологического элемента (корня или аффикса).

Каждое такое правило называется «способом образования».

2. В системе используются: два способа образования 3 л. ед. ч. наст. вр., два способа образования 3 л. мн. ч. наст. вр., два способа образования причастия и один способ образования инфинитива. Выбор способа образования одной формы не связан с выбором способа образования другой формы.

3. Корень каждого глагола имеет единый вид.

4. «Первообразная» система преобразуется в заданную некоторой цепью формул перехода (см. общие указания выше); глубина взаимодействия в каждой формуле перехода не должна превышать 3.

ДОПОЛНИТЕЛЬНОЕ УКАЗАНИЕ. Вопреки общему правилу, считать *n*, *ñ*, *ñ̃* и *ñ̄* одной и той же буквой (иначе говоря, заданный набор словоформ можно переписать, сняв все диакритические знаки при букве *n*)¹⁵.

Задача 7^a 16

ИСХОДНЫЕ ДАННЫЕ — те же, что в задаче 7, плюс еще один глагол в тех же четырех грамматических формах:

çṛṇóti çṛṇvánti çrutá çrótum

ЗАДАНИЕ — такое же, как в задаче 7.

14 Чтобы это правило было недвусмысленным, в решении должно быть указано, какие буквы исходного алфавита считаются гласными.

15 Это указание дано для того, чтобы несколько сократить работу по составлению формул перехода. Желающие могут, однако, отказаться от него и, таким образом, решать задачу в «необлегченном» варианте.

16 Эта и последующие задачи составлены так, что каждая из них включает в себя предшествующую (и, таким образом, задача 7^b включает в себя задачи 7, 7^a и 7^c). Мы рекомендуем, однако, решать эти задачи именно в той последовательности, в которой они даны. Это даст читателю возможность, переходя от более общих явлений к более частным, постепенно уточнять свои знания о предстории языка и тем самым как бы воспроизвести действительный ход лингвистического изучения.

Задача 7^б

ИСХОДНЫЕ ДАННЫЕ — те же, что в задаче 7^а, плюс еще три глагола в тех же четырех грамматических формах:

anákti	añjánti	aktá	ánktum
hánati	hánanti	hatá	hántum
jánati	jánanti	jātá	jánitum

ЗАДАНИЕ — такое же, как в задаче 7.

Задача 7^в

ИСХОДНЫЕ ДАННЫЕ — те же, что в задаче 7^б, плюс еще два глагола в тех же четырех грамматических формах:

svápati	svápanti	suptá	sváptum
jináti	jinánti	jītá	juátum

ЗАДАНИЕ — такое же, как в задаче 7.

РЕШЕНИЯ ЗАДАЧ

Публикуемые решения — не единственно возможные. Более того, они, быть может, и не наилучшие, и автор был бы благодарен читателям за сообщение ему более удачных решений. О приводимых ниже решениях можно сказать, таким образом, лишь то, что они наиболее точно соответствуют замыслу составителя задач.

В решениях учитывается только основное задание каждой задачи; дополнительные задания не рассматриваются.

РЕШЕНИЯ ЗАДАЧ НА ГРАММАТИЧЕСКИЙ АНАЛИЗ

Решение задачи I

ИСХОДНОЕ ПРЕДПОЛОЖЕНИЕ. Если будет обнаружено некоторое общее правило строения текста («грамматическая закономерность»), которое соблюдается в 11 фразах текста и нарушено только в одной фразе (причем, чтобы устранить нарушение, достаточно исправить в ней одно слово), то нарушение этого правила и есть искомая ошибка.

Поскольку смысл фраз нам неизвестен, мы можем искать только чисто формальные закономерности следующего общего вида:

при наличии (отсутствии, определенном взаимном расположении) во фразе некоторых слов или морфем в этой же фразе обязательно присутствуют (отсутствуют, располагаются в определенном порядке) некоторые другие слова или морфемы.

Рассмотрим строение фраз текста. Каждую фразу можно считать состоящей из следующих четырех «мест»:

I — занято одним или двумя из следующих восьми слов: *gizona(k)*, *astoa(k)*, *zaldia(k)*, *zakurra(k)*; при этом части *gizona-*, *astoa-*, *zaldia-*, *zakurra-* явно можно рассматривать как основы, а *-k* и, соответственно, нуль — как окончания.

II — Занято словом *atzo* или ничем.

III — Занято словом *joaten* или *ikusten*.

IV — Занято одним из восьми слов: *da*, *du*, *dira*, *ditu*, *zan*, *zuen*, *ziran*, *zituen*.

Здесь мы сделаем следующее предположение: слова, занимающие место I, принадлежат к одному и тому же синтаксическому классу; иначе говоря, замена одной основы на другую не нарушает грамматической правильности фразы. В тексте наиболее явными примерами такой взаимозаменяемости являются фразы 3 и 11, 2 и 6.

Теперь мы можем не учитывать выбор основы у слов этой группы и представить структуру фраз текста в виде следующей таблицы (порядок фраз изменен; основа слов группы I обозначена знаком ~; слова группы III обозначены сокращенно):

Фразы	«Место»			
	I	II	III	IV
1	~	—	j.	da
7	~k	—	j.	dira
3, 11	~	atzo	j.	zan
4	~k	atzo	j.	ziran
2, 6	~k ~	—	i.	du
9	~k ~k	—	i.	ditu
8	~k ~	atzo	i.	zuen
5	~k ~k	atzo	i.	zuen
10, 12	~k ~k	atzo	i.	zituen

При изучении этой таблицы мы замечаем ряд строгих зависимостей между разными членами одной и той же фразы:

$$1) \text{ I — III — IV: } \frac{\text{одно слово в I}}{\text{два слова в I}} = \frac{j.}{i.} = \frac{\text{гласная } a \text{ в IV}}{\text{гласная } u \text{ в IV}}.$$

$$2) \text{ II — IV: } \frac{\text{нет } atzo}{\text{есть } atzo} = \frac{\text{начальное } d \text{ в IV}}{\text{начальное } z \text{ и конечное } (e)n \text{ в IV}}.$$

ПРИМЕЧАНИЕ: конечное n выступает после a , конечное en — после u .

3) I — IV (соблюдается во всех фразах, кроме пятой):

$$\frac{\text{слово в I (если их два — второе) имеет окончание } -k}{\text{слово в I (если их два — второе) не имеет окончания } -k} = \\ = \frac{\text{слово в IV содержит } ir \text{ или } it}{\text{слово в IV не содержит } ir, it}.$$

ПРИМЕЧАНИЕ: ir выступает, если в I одно слово, it — если в I два слова.

Таким образом, искомое нарушение найдено: грамматически неправильна фраза 5. Чтобы ее исправить, надо выполнить 3-е правило соответствия, то есть либо а) исправить $zuen$ на $zítuen$, либо б) исправить $zaldíak$ на $zaldía$. В первом случае фраза 5 уподобится фразам 10 и 12, во втором случае — фразе 8.

С точки зрения баскского языка этот ответ действительно правилен.

Решение задач на перевод (2—4)

Предварительный этап решения, общий для задач 2—4

Назовем частичным алгоритмом перевода с языка В на язык А права перевода с В на А, применимые не к любым, а только к некоторым фразам языка В.

Допустим, что при решении задачи рассматриваемого типа нам удалось построить частичный алгоритм перевода с языка В на язык А¹⁷, обладающий следующими свойствами: 1) он применим ко всем контрольным фразам; 2) он применим ко всем фразам языка В из исходного двуязычного текста, и они переводятся с его помощью на язык А правильно¹⁸; 3) он не содержит правил, которые не применяются при переводе фраз из двуязычного текста; 4) каждой переводимой фразе языка В он ставит в соответствие только одну переводящую фразу языка А.

При решении задач 2—4 мы делаем следующее *основное предположение*: если получен алгоритм, обладающий всеми указанными свойствами, то контрольные фразы будут переведены с его

¹⁷ В задаче 2 в роли языка А выступает арабский, в роли В — русский.

¹⁸ При несоблюдении требования 4 (см. ниже) требование 2 должно быть ослаблено, а именно: хотя бы один из переводов каждой исходной фразы правилен.

помощью правильно (иначе говоря, для получения правильного перевода *достаточно* построить алгоритм описанного типа).

ПРИМЕЧАНИЕ. Легко понять, что это предположение не оправдывается, если в истинном переводе контрольных фраз на язык А встречаются морфологические или иные аномалии, не представленные в исходном тексте. Например, если в исходном тексте встречаются два ряда французских словоформ — *parlons, lisons, disons* и *parlez, lisez*, а для контрольного перевода нужна форма второго ряда от основы *dis-*, то наш алгоритм, разумеется, даст ошибочный ответ *disez* (вместо требуемой аномальной формы *dites*). Сформулированное выше основное предположение опирается, таким образом, на предположение о том, что в материале задач аномалий такого рода нет. В самом деле, пытаться угадать в таких случаях безнадежно, и поэтому, если мы вообще хотим надеяться на успех, мы должны исходить из предположения об отсутствии подобных аномалий.

Почему мы выдвигаем в качестве предположительной гарантии правильности перевода (при условии отсутствия аномалий) соблюдение *всех четырех* сформулированных выше требований к алгоритму? Дело в том, что никакие три из этих четырех требований заведомо не дают гарантии правильности ответа (даже при отсутствии аномалий).

Действительно, при несоблюдении первого требования мы вообще не получим перевода контрольных фраз.

Сняв второе требование, мы будем вынуждены признать удовлетворительным огромное множество алгоритмов, дающих любые фантастические переводы.

Смысл третьего требования можно показать на примере следующей задачи:

Язык А	Язык В
1. il marche	— he walks
2. il marchait	— he walked
3. il saute	— he jumps

Перевести на язык А фразу: *he jumped*.

Можно, например, предложить следующий частичный алгоритм перевода¹⁹: *he* → *il*, *walks* → *marche*, *walked* → *marchait*, *jumps* →

19 Строго говоря, данная запись не является алгоритмом, поскольку не указано, в каком порядке должны производиться действия и как из полученных элементов языка А составить фразу. Поскольку, однако, и то и другое очевидно, мы здесь и в дальнейшем ограничиваемся указанием соответствий, считая, что все прочие указания подразумеваются.

→ *saute, jumped* → *abc*²⁰ (запись $a \rightarrow b$ здесь и далее в решениях задач на перевод означает: «перевести элемент a элементом b »). Этот алгоритм удовлетворяет требованиям 1, 2, 4, но не удовлетворяет требованию 3. Контрольный перевод неправилен: *il abc*. Следовательно, соблюдение всех требований, кроме третьего, еще не обеспечивает правильного ответа.

В данном случае, чтобы соблюсти требование 3, нужно построить алгоритм не пословного, а поморфемного перевода, например, такой: *he* → *il*, *walk-* → *march-*, *jump-* → *saut-*, *-s* → *-e*, *-ed* → *-ait*. Этот алгоритм даст правильный перевод: *il sautait*.

Таким образом, важное следствие требования 3 состоит в том, что в качестве элементарных объектов перевода в алгоритме должны выступать достаточно мелкие единицы, а именно такие, на которые можно разложить как исходные, так и контрольные фразы. Так, ни в одной из наших задач в качестве элементарных объектов перевода не могут выступать целые фразы. В задаче 3, где каждая из словоформ, встречающихся в контрольных фразах, встречается также и в исходных фразах, возможен пословный перевод. В задачах 2 и 4 в контрольных фразах встречаются словоформы, отсутствующие в исходных фразах, и, следовательно, переводимой единицей может быть только морфема.

Смысл четвертого требования можно показать на примере следующей задачи:

	Язык А	Язык В
1.	<i>mon chat</i>	— <i>my cat</i>
2.	<i>mes chats</i>	— <i>my cats</i>
3.	<i>mon chien</i>	— <i>my dog</i>

Перевести на язык А фразу: *my dogs*.

Представим себе, что предложен следующий частичный алгоритм перевода: *cat* → *chat*, *dog* → *chien*, *-s* → *-s*, *my* → 1) *mon*, 2) *mes*. Этот алгоритм удовлетворяет всем требованиям, кроме четвертого. Контрольная фраза получит два перевода: 1) *mon chiens*, 2) *mes chiens*. В действительности правилен только один перевод (второй). Таким образом, соблюдение только первых трех требований не обеспечивает правильного ответа.

Итак, соблюдение всех четырех требований, как мы предполагаем (хотя и не можем этого доказать), является *достаточным*

²⁰ Взято в качестве образца произвольного слова.

для получения правильного ответа. А является ли хотя бы одно из этих требований *необходимым* для получения правильного ответа? Безусловно, нет. Дело в том, что правильный ответ можно вообще получить не с помощью алгоритма перевода, а, например, на основании разного рода разрозненных соображений ad hoc или даже на основании неосознанных интуитивных соображений. Если все же ответ достигается с помощью алгоритма, то, по-видимому, необходимо только соблюдение требования 1; прочие требования могут и не соблюдаться. Более того, существуют случаи, когда одновременное соблюдение всех четырех требований невозможно. Так, невозможно одновременно соблюсти требования 2 и 4, например, в задаче со следующими исходными данными:

Язык А Язык В

1. ты шел — you went
2. ты шла — you went

Заметим, что, решая такую задачу, мы, разумеется, пожертвуем требованием 4, но не требованием 2.

Из всего сказанного вытекает следующая программа действий при решении задач рассматриваемого типа. Прежде всего, мы пытаемся построить частичный алгоритм перевода, удовлетворяющий всем четырем сформулированным выше требованиям. Если это удалось, мы переводим с его помощью контрольные фразы и считаем задачу решенной. Если это не удалось, мы стараемся проверить, не являются ли в данном случае требования 2 и 4 несовместимыми. Если мы каким-либо образом убедились в том, что это действительно так, мы пытаемся построить частичный алгоритм перевода, удовлетворяющий хотя бы трем первым требованиям. В случае новой неудачи, по-видимому, можно и далее снимать поставленные вначале требования, однако наша уверенность в правильности ответа будет при этом каждый раз уменьшаться.

Во всех трех рассматриваемых ниже задачах имеет место оптимальный случай, то есть удается построить алгоритм, удовлетворяющий всем выдвинутым выше требованиям.

Решение задачи 2

Ход решения излагается сокращенно.

Единицей перевода, как указано выше, в данной задаче должна служить морфема. Основной способ установления соответ-

ствия между морфемами обоих текстов состоит в сопоставлении арабских слов, русские переводы которых имеют одинаковое лексическое значение при разном грамматическом или наоборот, одинаковое грамматическое значение при разном лексическом. В результате удастся получить частичный алгоритм перевода с русского языка на арабский, удовлетворяющий всем основным требованиям (см. выше, стр. 43).

Приведем основные правила передачи грамматических значений, входящие в этот алгоритм (лексические соответствия мы опускаем, поскольку они очевидны).

В арабских словах, соответствующих русским существительным, выражаются: 1) *падеж* — «именительный», т. е. падеж субъекта (показатель *-u*); «винительный», т. е. падеж прямого объекта (*-a*) или «родительный», т. е. падеж несогласованного определения (*-i*); последний выступает также после предлогов; 2) *степень определенности* (не выражается только в том случае, если слово имеет показатель обладателя или за ним следует определение в родительном падеже); различаются определенная форма (показатель *s*- перед *s*, *ṣ*- перед *ṣ*, *l*- в прочих случаях) и неопределенная форма (показатель *-n*, ставится после показателя падежа); 3) *лицо и род* (пол) *обладателя* (если таковой имеется): 1 л. *-ī*, 2 л. м. р. *-ka*, ж. р. *-ki*, 3 л. м. р. *-hu*, ж. р. *-hā*; показатели обладателя ставятся после показателя падежа; при этом *-ī* «поглощает» показатель падежа. Кроме того, в состав словоформы могут входить элементы: *li-* (по функции соответствует английскому предлогу *to*), *wa-* «и» (соединяет однородные члены предложения), *fa-* «и» (соединяет предложения). Никакого выражения грамматической категории числа обнаружить не удастся.

В арабских словах, соответствующих русским глаголам, выражаются: 1) *время* — настоящее-будущее или прошедшее; это различие выражается противопоставлением двух основ, имеющих одинаковый набор согласных, но разный набор гласных, например: *ḏḥaku, kūlu* — основы наст.-буд. времени, *ḏaḥika, kāla* — основы прошедшего времени²¹; 2) *лицо и род* (пол) *субъекта*: в наст.-буд. времени 1 л. *'a-*, 2 л. *ta-*, 3 л. *ya-*; в прош. времени 3 л. м. р. нуль, ж. р. *-t*; 3) *лицо и род* (пол) *объекта*, если он имеется (выражается так же, как лицо и род обладателя у имени). Кроме того,

21 Разделить словоформы типа *taḏḥaku, kālat* на основу и показатель субъекта можно и другими способами, например *ta-ḏḥak-u, t-aḏḥak-u, kāl-at*.

в состав словоформы могут входить элементы: 'а- «ли» (показатель вопроса), *wa-* «и», *fa-* «и» (см. выше).

Полученный алгоритм дает следующий перевод контрольных фраз:

1. *qaṣṣat hāḍihi lġinnīyatu kiṣṣata ṣafqatihā faḍaḥika sāmi'uhā.*

2. *ya'lamu samīruka, yā ṣabīyī, 'alfa kiṣṣatin wamā yamliku dirham.*

Этот перевод действительно правилен с точки зрения арабского языка.

Решение задачи 3

Эта задача может быть решена многими способами. Описанный ниже путь решения представляется нам одним из самых простых.

Как уже указано выше, эта задача, в отличие от задач 2 и 4, в принципе может быть решена с помощью алгоритма *пословного* перевода с В на А. Таким образом, здесь нет необходимости разбивать слова на морфемы (хотя это и нетрудно сделать).

Каждая пара переводящих друг друга фраз состоит из четырех слов. Обозначим эти слова по порядку в каждой паре: А I, А II, В I, В II.

Замечаем, что изменение В I, при неизменном В II (фразы 1 и 3, 2 и 5), отражается в языке А в виде изменения А II (при неизменном А I); с другой стороны, изменение В II, при неизменном В I (фразы 1 и 6, 4 и 5), отражается в виде изменения А I (при неизменном А II). Отсюда мы можем заключить, что слова фраз А и слова фраз В соответствуют друг другу (переводят друг друга) перекрестно: В I — А II, В II — А I.

Выпишем оба ряда соответствий (в направлении от В к А):

В I — А II	В II — А I
ṣth (3) → pinte	zbwb (1, 3) → mizē
ṣtw (2) → pinin	hzbwb (6)
yṣth (1, 6) → pi	zbwbym (2) } → miza
yṣtw (4, 5) → pinē	hzbwbym (4) → mizat

Эти ряды соответствий сами по себе можно рассматривать как частичный алгоритм перевода с В на А. Как нетрудно проверить, этот алгоритм удовлетворяет всем четырем основным требованиям (см. выше, стр. 43).

Применяем этот алгоритм к контрольным фразам:

1) *šth hzbwbw* → *miza pinte*; 2) *štw hzbwbwym* → *mizat pinin*.

Полученные переводы действительно правильны с точки зрения взятых реальных языков.

Решение задачи 4

1. Внутренний анализ каждого из двух одноязычных текстов.

Каждая фраза языка *A* состоит из следующих четырех «мест».

I — *arrek*, *aek* или нуль.

II — *urdea(k)*, *labanak* или нуль. Внутри этого «места» мы можем выделить два ряда морфем: IIa (основы) — *urdea-*, *labana-*, нуль; IIб (окончания) — *-k*, нуль

III — *agindu(tzen)* или *billatu(tzen)*. И здесь легко выделить два ряда морфем: IIIa — *agindu-*, *billatu-*, IIIб — *-tzen*, нуль.

IV — одно из 13 разных слов, начинающихся на *de-* или *di-*. Можно выделить следующие ряды: IVa — *de-*, *di-*; IVб — *-zki-*, нуль; IVв — *-o-*, нуль; IVг — нуль, *-zu*, *-te*, *-zute*, *-tezu*, *-tezute* (разбить ряд IVг на ряды нуль/*zu* и нуль/*te* нельзя из-за возможности разного порядка элементов *zu* и *te*).

Каждая фраза языка *B*²² состоит, вообще говоря, тоже из четырех «мест»: 1) *ezt*, *ezeke*t или нуль; 2) одно из слов с основой *igér-* или *keres-*; 3) нуль или одно из слов с основой *nek-*; 4) *a sertés(eke)t*, *a kés(eke)t* или нуль (слово *a* явно не составляет самостоятельного «места»). Легко заметить, однако, что «места» 1 и 4 тесно связаны между собой, а именно, если одно из них занято, то другое свободно (представлено нулем), и наоборот; кроме того, слова, выступающие на этих местах, имеют формальное сходство — они обнаруживают одно и то же противопоставление *-t/-eket*. Это позволяет нам рассматривать «места» 1 и 4 как единое «место».

Таким образом, мы выделяем во фразе языка *B* только три «места»:

I (= 1 и 4) — *ez(eke)t*, *a sertés(eke)t* или *a kés(eke)t*. Выделяются ряды: Ia — *ez-*, *a sertés-*, *a kés-*; Ib — *-t*, *-eket*.

II — одно из слов с основой на *igér-* или *keres-*. Выделяются ряды: IIa — *igér-*, *keres-*; IIб — *-i*, *-ik*, *-ed*, *-itek*, *-te*, *-ték*, *-ted*, *-tétek*. Восемь окончаний, выступающих в IIб, явно противопоставлены друг другу по трем бинарным признакам:

²² При анализе учитываются также контрольные фразы.

	Тип <i>i</i> (нет согласных, кроме начального <i>t</i> и конечного <i>k</i>)		Тип <i>ed</i> (есть согласная — <i>t</i> или <i>d</i> , — кроме начального <i>t</i> и конечного <i>k</i>)	
	не оканчивается на <i>k</i>	оканчивается на <i>k</i>	не оканчивается на <i>k</i>	оканчивается на <i>k</i>
Не начинается с <i>t</i> Начинается с <i>t</i>	- <i>i</i> - <i>te</i>	- <i>ik</i> - <i>ték</i>	- <i>ed</i> - <i>ted</i>	- <i>itek</i> - <i>tétek</i>

Можно предположить, что здесь в каждом окончании три морфемы: 1) *t* или нуль, 2) *i* или *ed*, 3) *k* или нуль, — но при их соединении происходят изменения (*ed + k* дает *itek*; *t + i* дает в середине слова *té*, в конце слова *te*). Однако подтверждение этому мы получим лишь в том случае, если окажется, что каждая из этих морфем имеет самостоятельное отражение в языке А.

III — нуль или одно из слов с основой *nek-*. Выделяются ряды: IIIа — нуль или *nek-*; IIIб (только для слов с *nek-*) — *-i*, *-ik*, *-ed*, *-tek*. Заметим, что полученные окончания явно сходны с первым рядом окончаний в IIб.

2. Сопоставление результатов отдельного анализа текстов.

Выписываем для каждого ряда морфем (Iа, Iб и т. д.) обоих языков распределение членов этого ряда по фразам, например: В IIIа — нуль (1, 2), *nek-* (остальные фразы). Сравниваем последовательно распределение членов в каждом ряду морфем языка В с распределением членов в рядах морфем языка А. (При рассмотрении рядов В IIб и IIIб учитываем также гипотезу о многоморфемном составе «окончаний».) Если мы обнаруживаем два ряда морфем (разных языков), имеющие одинаковое распределение членов, мы считаем, что эти ряды соответствуют друг другу, а входящие в них морфемы переводят друг друга. Например, для рассмотренного выше ряда В IIIа мы найдем соответствующий ему ряд А IVа с распределением: *de-* (1, 2), *di-* (остальные фразы). Отсюда правило перевода с В на А: нуль в В IIIа → *de-* в А IVа, *nek-* в В IIIа → *di-* в А IVа.

Действуя таким способом, мы обнаруживаем следующие соответствия.

В Iа — А IIа: *ez-* → нуль, *a sertés-* → *urdea-*, *a kés-* → *labana-*.

В Iб — А IVб и А IIб: *-eket* → *-zki* в А IVб и *-k* в А IIб (если только А IIа не нуль), *-t* → нуль в А IVб и А IIб.

В IIа — А IIIа: *igér-* → *agindu-*, *keres-* → *billatu-*.

В IIб имеет более сложные соответствия:

1) наличие или отсутствие начального t отражается в А Шб: есть $t \rightarrow$ нет $-tzen$, нет $t \rightarrow$ есть $-tzen$;

2) противопоставление типов i и ed и наличие или отсутствие конечного k отражается в А I: тип $ed \rightarrow$ нуль, тип $i \rightarrow$ не нуль, а именно: если есть конечное $k — aek$, если нет конечного $k — arrek$. Как отражается в языке А наличие или отсутствие k в окончаниях типа ed , остается пока неясным.

В Ша — А IVа: нуль $\rightarrow de-$, $nek-$ $\rightarrow di-$.

В Шб: противопоставление типов i и ed отражается в А IVв: тип $i \rightarrow o$, тип $ed \rightarrow$ нуль. Как отражается в языке А наличие или отсутствие конечного k , неясно.

Таким образом, на данном этапе решения остается невыясненным лишь то, как отражаются некоторые элементы рядов В Пб и В Шб и что отражает ряд А IVг. Естественно искать связь между самими этими рядами. Такая связь действительно есть: чтобы в этом убедиться, достаточно сравнить, например, фразы 4 и 5.

Выпишем из каждой пары фраз представленные в этой паре члены рядов В Пб, В Шб и А IVг. Например, для первой пары фраз получим тройку $-ed$, нуль, $-zu$, для шестой пары — тройку $-tétek$, $-ik$, $-tezute$ и т. д. Сравнивая эти тройки, замечаем следующие соответствия: 1) конечное k как члена В Пб, так и члена В Шб отражается в А IVг в виде элемента te ; 2) любой элемент типа ed (то есть $-ed$, $-ite-$, $-éte-$, $-te-$) как члена В Пб, так и члена В Шб отражается в А IVг в виде элемента zu . При этом элементы te и zu располагаются внутри члена А IVг в строго определенном порядке. Если обозначить через te_2 и zu_2 отражения элементов члена В Пб, а через te_3 и zu_3 — отражения элементов члена В Шб, то этот порядок можно записать так:

$$zu_3 — te_3 — zu_2 — te_2.$$

Итак, мы можем теперь уточнить составленные выше правила соответствия следующим образом.

Член В Пб:

1) наличие или отсутствие начального t отражается в А Шб (см. выше);

2) противопоставление типов i и ed отражается в А I и А IVг: тип $ed \rightarrow$ нуль в А I и zu_2 в А IVг, тип $i \rightarrow$ не нуль ($arrek$ или aek) в А I и отсутствие zu_2 в А IVг;

3) наличие или отсутствие конечного k отражается в А I и А IVг: есть $k \rightarrow aek$ (если не нуль, см. выше) в А I и te_2 в А IVг,

нет $k \rightarrow arrek$ (если не нуль, см. выше) в А I и отсутствие te_2 в А IVг.

Член В IIIб:

1) противопоставление типов i и ed отражается в А IVв и А IVг: тип $ed \rightarrow$ нуль в А IVв и zu_3 в А IVг, тип $i \rightarrow o$ в А IVв и отсутствие zu_3 в А IVг;

2) наличие или отсутствие конечного k отражается в А IVг: есть $k \rightarrow te_3$, нет $k \rightarrow$ отсутствие te_3 .

Теперь нам удалось найти в языке А отражение всех элементов фразы языка В; в то же время каждый элемент фразы языка А получил объяснение, т. е. мы выявили, что он отражает.

ЗАМЕЧАНИЕ. Из наших наблюдений следует, что первоначальное деление члена А IV было неполным и не совсем точным. Теперь мы могли бы выделить в нем следующие ряды морфем: 1) $de-$, $di-$, 2) zki , нуль, 3) zu_3 , o , нуль, 4) te_3 , нуль, 5) zu_2 , нуль, 6) te_2 , нуль. Ряд 1 отражает В IIIа, 2 — В Iб, 3 и 4 — В IIIб, 5 и 6 — В IIб.

Сформулированные выше правила соответствия можно рассматривать как частичный алгоритм перевода с В на А (в алгоритм должна также быть включена формула расположения элементов внутри А IVг). Этот алгоритм, как показывает проверка, удовлетворяет всем четырем основным требованиям (см. выше, стр. 43).

Применяя этот алгоритм к контрольным фразам, получаем следующие переводы:

1. Arrek labana agindutzen dio.
2. Urdeak billatutzen dizkiotezu.
3. Aek agindu dizute.

Полученные переводы действительно правильны с точки зрения взятых реальных языков.

РЕШЕНИЯ ЗАДАЧ НА ВНУТРЕНнюю РЕКОНСТРУКЦИЮ

ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ. 1. В приводимых ниже решениях сокращение записи и обобщение формул произведено не везде, где это возможно, а только там, где это действительно упрощает запись. В некоторых случаях (см., например, формулу 3 в решении задачи 5) обобщение сделано не по форме, описанной на стр. 34, а выражается только в том, что обобщаемые формулы получают единый номер. Во всех таких случаях имеется в виду тип обобщения $A \rightarrow B$, где в качестве значений А выступают левые части формул, а в качестве значений В — правые.

2. В большинстве решений в исходный алфавит вводятся дополнительные буквы. Разумеется, в качестве таковых могут быть использованы, вообще говоря, любые знаки. Выбор дополнительных букв в наших решениях определяется фонетическими ассоциациями. Выбор знака Н в решении задачи 7 соответствует современной практике сравнительного языкознания (Ф. де Соссюр использовал в своей работе знак A).

Решение задачи 5

«Первообразная» система словоформ

В исходный алфавит, помимо букв, встречающихся в заданной системе словоформ, включаем букву i_2 .

КОРНИ: $s\bar{a}k$ -, $n\bar{a}k$ -, $jauk$ -, nes -, $v\bar{e}rs$ -, met -, $jaut$ -, cep -, $k\bar{a}p$ -.

ОКОНЧАНИЯ: инфинитив $-t$; наст. время $-u$, $-i_2$, нуль; прош. время $-u$, $-i$, $-a$.

СУФФИКС: $-j$ -. Он используется в наст. времени глаголов $jaukt$, $v\bar{e}rst$, $jaust$, $k\bar{a}pt$ и в прош. времени глаголов $n\bar{a}kt$, $jaukt$.

Цепь формул перехода

- | | |
|---|---------------------------|
| 1. $j(+i_2) \rightarrow$ нуль | 4. $tt \rightarrow st$ |
| 2. $k(+i_2) \rightarrow c$ | 5. $i_2 \rightarrow$ нуль |
| 3. $kj \rightarrow c$ | |
| $\left. \begin{array}{l} sj \\ tj \end{array} \right\} \rightarrow \check{s}$ | |

Решение задачи 6

«Первообразная» система словоформ

В исходный алфавит, помимо букв, встречающихся в заданной системе словоформ, включаем букву j_2 . Буквы a , \bar{a} , e , i считаются гласными, все прочие — согласными. В формулах С обозначает любую согласную.

КОРНИ: $y\bar{a}$ -, ved -, $r\bar{a}j_2$ -, $bhaj$ -, vac -, $va\check{s}$ -, $dve\check{s}$ -.

ПОКАЗАТЕЛИ ЛИЦА (ставятся после корня): 1 л. $-m$ -, 2 л. $-s$ -, 3 л. $-t$ -.

ПОКАЗАТЕЛИ ВРЕМЕНИ: наст. время $-i$ (ставится после показателя лица), прош. время a - (ставится перед корнем).

Цепь формул перехода

1. $A \left(+ \begin{Bmatrix} s \\ t \end{Bmatrix} \right) \rightarrow B$,
 где A: соответственно, B:
 d — t
 с или j — k
 ç, ş или j₂ — t
3. $s(+\#) \rightarrow \text{h}$
 4. $\left. \begin{matrix} ks \\ ts \end{matrix} \right\} \rightarrow ks$
 tt → şt
 5. $j_2 \rightarrow j$
2. $(C+) \left\{ \begin{matrix} s \\ t \end{matrix} \right\} (+\#) \rightarrow \text{нуль}$
 (C+) m (+\#) → am

Решение задачи 7

«Первообразная» система словоформ

В исходный алфавит, помимо букв, встречающихся в заданной системе словоформ, включаем буквы *H* и \bar{j} . Буквы *a, ā, i, ī, u, ū, r, ṛ, e, o* считаются гласными, все прочие — согласными. В формулах *V* обозначает любую гласную, а *C, C₁, C₂* — любые согласные.

корни: *yuj-, bhid-, vṛj-; puH-, kṣiH-, strḡH-; ru-, ji-, bhṛ-; bhuH-, niH-, tṛH-*.

СПОСОБЫ ОБРАЗОВАНИЯ ФОРМ:

НАСТ. ВРЕМЯ:

Первый способ (первые шесть глаголов) — в 3 л. ед. ч. инфикс *-ná-*²³, вставляемый перед последней буквой корня, и суффикс (окончание) *-ti*; в 3 л. мн. ч. инфикс *-n-*, вставляемый перед последней буквой корня, и суффикс (окончание) *-ánti*;

Второй способ (вторые шесть глаголов) — в 3 л. ед. ч. суффикс (окончание) *-ati*, в 3 л. мн. ч. суффикс (окончание) *-anti*; ударение в обеих формах на корне²⁴.

ПРИЧАСТИЕ:

Первый способ (все глаголы, кроме четырех, — см. ниже) — суффикс *-tá*;

Второй способ (глаголы с корнями *bhid-, kṣiH-, strḡH-, tṛH-*) — суффикс *-ná*.

инфинитив: суффикс *-tum*, ударение на корне.

²³ Здесь и ниже знак ударения над служебным элементом показывает, что при данном способе образования ударение всегда падает на этот элемент.

²⁴ Здесь и ниже подразумевается: на единственной гласной корня.

Цепь формул перехода

1. $\acute{u} \rightarrow \acute{a}v$
 $\acute{i} \rightarrow \acute{a}y$
 $\acute{f} \rightarrow \acute{a}r$
2. $(C_1 +)H(+C_2) \rightarrow i$
3. $VH(+C) \rightarrow \bar{V}$,
где \bar{V} — та же гласная, что в левой части формулы, но только с добавлением знака долготы, если гласная в левой части этого знака не имеет.
4. $H \rightarrow$ нуль
5. $\bar{r} \rightarrow \bar{i}r$
6. $av(+C) \rightarrow o$
 $ay(+C) \rightarrow e$
7. $dn \rightarrow nn$
 $dt \rightarrow tt$
 $jt \rightarrow kt$

Решение задачи 7^а

Для решения этой задачи достаточно внести несколько дополнений в решение задачи 7.

В список корней добавляем корень ζru -. Глагол с этим корнем образует как настоящее время, так и причастие первым из описанных выше способов.

В цепь формул перехода добавляем следующие формулы:

$$2^a. (C_1 +) r (+C_2) \rightarrow \bar{r}$$

$$5^a. ua \rightarrow va$$

$$6^a. au \rightarrow o$$

Решение задачи 7^б

Для решения этой задачи достаточно внести несколько дополнений в решение задачи 7^а.

В исходный алфавит вводим букву η , которую объявляем гласной. В список корней добавляем корни ηj -, $h\eta$ -, $j\eta H$ -. Глагол с корнем ηj - образует настоящее время первым способом, глаголы с корнями $h\eta$ - и $j\eta H$ - — вторым способом; все три глагола образуют причастие первым способом.

В цепь формул перехода добавляем следующие формулы:

$$1^a. \acute{\eta} \rightarrow \acute{a}\eta$$

$$1^b. \eta \rightarrow a$$

Решение задачи 7^в

Эту задачу можно решить тем же способом, что и задачи 7^а и 7^б, — внося дополнения в решение предшествующей задачи. Однако в данном случае более короткую цепь формул перехода

дает другой путь, состоящий в существенной перестройке предшествующих решений.

«Первообразная» система словоформ

В исходный алфавит входят те же буквы, что в решении задачи 7, и, кроме того, буква a_2 (гласная).

КОРНИ: $yavj-$, $bhayd-$, $varj-$; $pavH-$, $kşayH-$, $starH-$; $rav-$, $jay-$, $bhar-$; $bhavH-$, $paуH-$, $tarH-$; $şrav-$; $anj-$, $han-$, $janH-$; $svap-$, $jyaH-$.

СПОСОБЫ ОБРАЗОВАНИЯ ФОРМ — те же, что в решении задачи 7, за исключением второго способа образования настоящего времени. Второй способ образования настоящего времени в данном случае таков: в 3 л. ед. ч. суффикс (окончание) $-a_2ti$, в 3 л. мн. ч. суффикс (окончание) $-a_2nti$; ударение в обеих формах на корне.

Глагол с корнем $jyaH-$ образует настоящее время первым способом, глагол с корнем $svap-$ — вторым способом; оба эти глагола образуют причастие первым способом.

Цепь формул перехода

- | | |
|--|--|
| 1. $a_{\text{безуд.}} \rightarrow \text{нуль}$ | 8. $VH (+C) \rightarrow \bar{V}$ (см. формулу 3 в решении задачи 7). |
| 2. $a_2 \rightarrow a$ | 9. $H \rightarrow \text{нуль}$ |
| 3. $\left. \begin{array}{l} \{C_1 \\ \# \} + \end{array} \right) v (+C_2) \rightarrow u$ | 10. $\bar{r} \rightarrow \bar{ir}$ |
| 4. $- „ - y - „ - \rightarrow i$ | 11. $av (+C) \rightarrow o$ |
| 5. $- „ - r - „ - \rightarrow r$ | $ay (+C) \rightarrow e$ |
| 6. $- „ - n - „ - \rightarrow a$ | 12. $dn \rightarrow nn$ |
| 7. $- „ - H - „ - \rightarrow i$ | $dt \rightarrow tt$ |
| | $jt \rightarrow kt$ |

ПРИМЕЧАНИЕ. Если не принимать во внимание дополнительное указание о буквах η , \bar{n} , \bar{n} (стр. 40), мы должны будем дополнить эту цепь следующими формулами:

13. $nj \rightarrow \bar{n}j$
 $nk \rightarrow \bar{n}k$
14. $(R+) n \rightarrow \bar{n}$, где $R - r, \bar{r}, \bar{s}$ или \bar{s} .

Аналогичные добавления потребуются и в трех предшествующих задачах.

25 Все переходы этой группы происходят в одинаковых условиях (между двумя согласными или между пробелом и согласной), но обобщить их нельзя, так как порядок обработки букв v, y, r, n, H существенен.