

В первой синтагме это определение будет двусмысленным. Однако известно, что в рамочной конструкции со значением места глагол невозможен, следовательно альтернатива $\frac{S}{V}$ решается в пользу S. В окончательном виде первая синтагма выглядит следующим образом: S-ды SSS.

Во второй синтагме определение также двусмысленно. Однако известно, что в усилительной конструкции возможно только A. Кроме того, известно, что в синтагме может быть только один глагол, следовательно альтернатива в субсинтагме $S \frac{S}{V} V$ -ды должна быть разрешена в пользу S. В окончательном виде вторая синтагма выглядит следующим образом: SSV -ды SVcA -ды.

В третьей синтагме снова имеет место альтернатива. Однако суффикс -ды показывает, что эту альтернативу следует разрешить в пользу A. В окончательном виде третья синтагма выглядит следующим образом: PS-ды S Ad Vm Adv.

3. На третьей ступени выделяются определительные субсинтагмы. Приведенный пример не совсем удачен для того, чтобы показать, почему определительные субсинтагмы следует выделять после определения принадлежности слов к тем или иным частям речи. В этом примере определительные субсинтагмы находятся в левой части синтагм, поэтому их левые границы совпадают с левыми границами синтагм. Однако если бы они находились в правой части синтагмы, их левой границей послужил бы глагол, т. е. не формальный, а знаменательный элемент, который сам по себе нуждается в определении.

После выделения субсинтагм получаем синтагмы, в которых полностью отсутствуют служебные слова, и все грамматические отношения между словами определяются только их порядком.

祖國的	цзуго-ды	синтагма (1)
科建設事業中在社會主義	цзай шэжуйчжун цзянь- шэ шисэ чжун	синтагма (3)
學工作相負的	кэнькэ гуицзо дан фу-ды	синтагма (8)
任務是繁重的	жэньу ши фаньчжун-ды	синтагма (10)
程度還將一年比一年增加。	чэнду хай цзян и-нянь би-и-нянь цзэньцзя	синтагма (8)

4. По изложенным выше правилам осуществляется перевод отдельных синтагм: 1) «родины»; 2) «в деле социалистического строительства»; 3) «за которые несет ответственность научная работа»; 4) «задачи разнообразны»; 5) «этого разнообразия»; 6) «степень еще будет увеличиваться из года в год». На этой же ступени осуществляется и согласование между центральными словами синтагм.

5. Соединение переведенных синтагм проводится в обратном порядке. В первую очередь переводы определительных субсинтагм присоединяются к тем существительным, к которым относятся: 1 + 2 «в деле социалистического строительства родины»; 3 + 4 «задачи, за которые несет ответственность научная работа»; 5 + 6 «степень разнообразия еще будет увеличиваться из года в год».

Перед нами перевод синтагм, которые получились в результате расчленения предложения на первой ступени перевода. Затем переводы этих синтагм располагаются в том же порядке, как и в китайском предложении. При этом возможны некоторые отклонения, регулируемые специально введенными правилами.

В результате получается перевод: «Задачи, за которые несет ответственность научная работа в деле социалистического строительства родины, весьма разнообразны, степень разнообразия еще будет увеличиваться из года в год».

М. В. Софронов

О ЛИНГВИСТИЧЕСКОЙ ВЕРОЯТНОСТИ

Знать язык — это значит хранить в памяти слова и грамматические модели этого языка, определяемые его грамматическим строем. Слова хранятся в памяти в виде звуковых образов, а у грамотных людей — и в виде зрительных. Последние либо соотносимы со звуковыми образами, если мы имеем дело с языками, пользующимися буквенным (или, как его часто называют, звуковым) письмом, либо не соотносимы с ними, если мы имеем дело с языками, пользующимися иероглифами. Звуковой облик отдельных слов состоит из некоторого числа повторяющихся в разнообразных сочетаниях элементов, именуемых звуками речи или фонемами, составляющими в каждом языке свою особую систему.

Так как отдельные слова могут состоять из различного — иногда довольно большого — числа фонем, а сочетаемость фонем хотя и ограничена в каждом языке¹, но все же достаточно велика, то язык обладает возможностью образовать из нескольких десятков фонем сотни тысяч слов. То же можно сказать о письменной форме языка и о системе букв, используемой в ней.

Грамматические модели, к которым относятся не только модели предложений, словосочетаний и словоизменения, но и словообразовательные модели, не существуют в речи отдельно от слов; они реализуются в тех или иных формах слов (т. е. в разнообразии их звуковых обликов), в сочетаниях с так называемыми «служебными» словами (предлогами, союзами и т. д.) или же в том или ином расположении слов. Вместе с тем грамматические модели абстрагируются в языке от слов и существуют в виде правил как бы независимо от них. Соответственно они и в памяти хранятся в таком абстрагированном, «схематическом» виде.

Восприятие или понимание² речи заключается в том, что воспринятые звуковые сигналы в устной речи или же зрительные сигналы в письменной речи отождествляются с хранящимися в памяти звуковыми или зрительными образами слов³. Чем богаче такими образами память слушающего, тем быстрее и легче достигает он понимания речи. Активный запас слов отдельного человека, т. е. число употребляемых им слов, относительно невелик: он едва ли превышает (по крайней мере у представителей высокообразованных языков) 10% всего словарного состава данного языка. Пассивный запас слов отдельного человека, т. е. число слов, значение которых ему известно, значительно больше активного, но и этот запас слов, несомненно, далеко не достигает всего числа слов языка⁴.

Следует отметить, что, благодаря наличию в памяти словообразовательных моделей, число «понятных» слов больше числа хранящихся в памяти. Так, значение впервые услышанного слова *холодильник* выясняется благодаря суффиксу *-ильник* (модель — *кипятильник*). Приблизительное значение слова *вертолет* узнается по модели *самолет* и т. п. Некоторое число слов, неизвестных слушающему, следовательно, может быть «расшифровано» им. Однако в речи встречаются и такие слова, которые остаются на первых порах непонятными. Развитие словарного запаса отдельного индивидуума обязательно проходит этап превращения непонятных слов в понятные.

Все сказанное позволяет различать восприятие и понимание. Первое заключается в о п о з н а в а н и и з в у к о в о г о о б л и к а с л о в а; второе — в связи этого облика со значением. Восприятие, как мы видели, предшествует пониманию.

Возможность восприятия отдельных слов (но отнюдь не речи в целом) без понимания их обусловлена тем, что в языке существует, а значит и хранится в памяти слушающего абстрагированная от слов система звуковых элементов языка — фонем. Через них и происходит опознавание звукового облика слова.

На различении восприятия и понимания основан давно применяемый в телефонии метод слоговой артикуляции, использующий для испытания линий связи бессмысленные звукосочетания, возможность восприятия которых не подлежит никакому сомнению, так как она доказана практикой. Вместе с тем, несмотря на то, что здесь исключено понимание, восприятие звуков (фонем), составляющих такие звукосочетания, существенно отличается от восприятия любых других звуков в природе⁵. Объясняется это тем, что состав фонем языка, не являющихся самостоятельными носителями смысловых значений, в конечном счете определяется смысловыми отношениями, существующими в данном языке. Поэтому можно сказать, что хотя восприятие звуков речи, фонем и их сочетаний возможно и без понимания, оно все же связано с ним.

Понимание заметным образом влияет на восприятие, облегчая его, делая его более доступным для слушающего. Эта способность понимания давать дополнительные сведения о передаваемом (будь то предложение, слово, морфема) вместе с существующими в языке вероятностными ограничениями и определяет то, что в теории информации на-

¹ В русском языке, например, глухие согласные не могут в пределах одного слова сочетаться со звонкими, твердые согласные не могут стоять перед гласным *и* и т. д.

² Ниже будет показано, что эти два понятия не совпадают, хотя и теснейшим образом связаны друг с другом.

³ Поскольку в рассматриваемом в настоящей работе аспекте между письменной и устной формой языка существует полная аналогия, в дальнейшем речь будет идти только о последней форме.

⁴ Относительно активного запаса слов в литературе имеются кое-какие данные; они касаются числа слов, встречающихся в произведениях таких писателей, как Шекспир, Пушкин и т. п. Что же касается пассивного запаса, то не существует даже надежной методики для его учета.

⁵ Л. Р. З и н д е р, Специфические особенности восприятия звуков речи, сб. «Восприятие звуковых сигналов в различных акустических условиях», М., 1956, стр. 69.

зывают «избыточностью»¹. Последняя, как пишет М. П. Долуханов, «показывает относительное количество „лишней“ информации, которая определяется структурой языка и известна уже на приемном конце из статистических данных»².

Нетрудно увидеть, что избыточность — это в общем то же, что и «психологический фактор» или «фактор догадки», о котором раньше говорили в телефонии. Стремясь получить данные о «чистом восприятии» при артикуляционных испытаниях, старались устранить действие этого фактора, для чего и пользовались бессмысленными звуко-сочетаниями. Теория информации, следовательно, не открыла самого явления, но она, во-первых, позволила уточнить его сущность и, во-вторых, нашла для него единицу измерения.

Совершенно очевидно, что измерять действие психологического фактора невозможно, так как этот фактор — понятие субъективное; «догадливость» того или иного индивидуума зависит от его образованности, от его психического состояния, от степени утомленности и т. п. Однако в о з м о ж н о с т ь догадки определяется объективно закономерностями, присущими данному языку.

Количество избыточной информации, облегчающей догадку и тем самым понимание и восприятие речи, определяется степенью вероятности соответствующего явления в данном языке. Вероятностные ограничения, характеризующие язык, удобно назвать лингвистической вероятностью. Этим понятием охватываются различные явления: вероятность появления в речи отдельных элементов (слов, фонем); вероятность появления в речи тех или иных грамматических моделей; вероятность появления тех или иных звуко-сочетаний, тех или иных словосочетаний; иными словами, вероятность следования таких-то звуков за такими-то, таких-то слов за такими-то. Все это выявляется в той или иной степени в так называемой статистической структуре языка.

Вероятность словосочетаний подчиняется двум различным закономерностям; соответственно этому можно говорить о лексической и грамматической вероятности словосочетаний. Под первой следует понимать сочетаемость слов, зависящую от их смысловых значений. Зависимость эта, строго говоря, не лингвистическая; она определяется объективной действительностью, условиями жизни общества и является поэтому очень изменчивой как территориально, так и во времени. Например, словосочетание *сидеть на крыше* в условиях жизни в Европейской части СССР неизмеримо менее вероятно, чем словосочетание *сидеть в саду*; в Средней же Азии вероятность первого значительно возрастает. Равным образом словосочетание *стеклянная сковородка*, почти невероятное лет 20 назад, стало сейчас широко употребительным³. Таким образом, лексическая вероятность — это закономерность чисто статистическая, не зависящая от грамматического строя языка.

Грамматическая вероятность может быть двух видов. Первый вид — это вероятность появления после данного слова та кой-то части речи; второй вид — это вероятность появления после данной грамматической формы та кой-то грамматической формы. Первый вид имеет особенно большое значение в языках с твердым порядком слов (как, например, английский), в которых действует, таким образом, чисто грамматическая закономерность, — определенное грамматическое правило. Он играет известную роль также и в языках с более свободным порядком слов (как, например, русский). Так, сочетание *холодная зима*, в котором определяемое (имя существительное) непосредственно следует за определением, более вероятно, чем сочетание *холодная пыliche зима*, где определяемое и определение разделены третьим словом (наречием). Вероятность в этом случае определяется не только грамматически, но и статистически.

Второй вид грамматической вероятности имеет особенно большое значение в языках с развитой флективной системой (как, например, русской), в которых правила согласования и управления определяют появление соответствующей грамматической формы. В случае согласования действует чисто грамматическая закономерность: за словом *холодная* может последовать только им. падеж ед. числа слова женского рода. В случае управления может проявиться и статистическая закономерность, если управление колеблющееся; за словами *не читал*, как правило, будет следовать род. падеж (*не читал книги*), но не исключен и вин. падеж (*не читал книгу*).

¹ См.: М. П. Долуханов, Введение в теорию передачи информации по электрическим каналам связи, М., 1955; А. А. Харкевич, Очерки общей теории связи, М., 1955; С. Голдман, Теория информации, перевод с англ., М., 1957; G. A. Miller, Language and communication, New York, 1951.

² М. П. Долуханов, указ. соч., стр. 30. Термин «избыточность» следует понимать условно, так как при наличии помех может оказаться невозможным получить даже минимум информации, если нет избыточности.

³ Сказанное заставляет сделать вывод, что лексическую вероятность следует определять для более или менее ограниченного отрезка времени; сочетаемость слов, наблюдаемая в произведениях писателей XIX в., ничего не говорит о лексической вероятности в русском языке наших дней. То же относится, разумеется, и к вероятности появления в речи тех или иных слов, а также отдельных фонем.

Количество избыточной информации, обеспечиваемой грамматической вероятностью, в общем очень велико, особенно в тех случаях, когда мы имеем дело с согласованием, а отчасти и с управлением. Каждое из согласуемых слов несет в себе информацию о грамматической форме второго из согласуемых слов, или, иными словами, — информацию о грамматической модели всего словосочетания. Достаточно услышать слово *дети*, чтобы знать, что сказуемое будет стоять в 3-м лице мн. числа, услышав определение *большого*, мы знаем, что определяемое является существительным мужского или среднего рода ед. числа. Если же перед этим словом стоит, например, слово *увидел*, то мы получаем еще больше сведений об определяемом, которое в данном случае почти обязательно будет существительным мужского рода (и притом обозначающим одушевленный предмет), стоящим в вин. надежде и выполняющим функцию дополнения. Следует заметить, что количество избыточной информации здесь будет тем большим, чем более всеобъемлющим является соответствующее грамматическое правило.

Количество избыточной информации, обеспечиваемой лексической вероятностью, гораздо более разнообразно, так как целиком зависит от статистической структуры языка. В словосочетании *красный флаг*, например, слово *красный* содержит в какой-то мере сведения о том, что за ним последует слово *флаг*. При этом количество избыточной информации, как уже указывалось выше, определяется вероятностью появления сочетания *красный флаг* по сравнению с вероятностями сочетаний — *красный цвет*, *красный мак* и т. п. Увеличение объема словосочетания влечет за собой увеличение количества избыточной информации. Так, например, в словосочетании *на балконе висел красный флаг* вероятность появления слова *флаг* после *красный*, разумеется, гораздо больше, чем в сочетании *красный флаг*, так как после слов *на балконе висел* невозможны сочетания — *красный цвет*, *красный пол*, *красный штаб*, *красный нос* и др. Соответственно с этим и количество избыточной информации, содержащейся в слове *красный*, в первом сочетании значительно больше, чем во втором.

Рассмотренные виды лингвистической вероятности связаны с пониманием речи, которое предполагает, что какие-то элементы речи (слова или словосочетания) правильно восприняты и поняты. Наряду с этим и само восприятие, которое, как было показано выше, сводится, в конечном счете, к восприятию отдельных звуковых единиц — фоном, облегчается лингвистической вероятностью. В данном случае речь идет о вероятности соответствующих звуковых сочетаний; вероятности того, что за данной ой фонемой последует такая-то. Эта вероятность, которую можно было бы назвать звуковой вероятностью, определяется двояко: во-первых, фонетическими правилами данного языка (в русском языке, например, за звонким согласным может следовать только гласный или звонкий же согласный), во-вторых, чисто статистически, как это имеет место и при лексической вероятности.

Так как звуковая вероятность не влияет непосредственно на понимание, то значение ее как источника избыточности не так очевидно, как значение лексической вероятности. Тем не менее звуковая вероятность, оказывающая существенное влияние на восприятие, также должна быть признана источником избыточной информации. Для иллюстрации этого положения можно указать на следующий факт. В результате специальных опытов, поставленных еще в 1949 г., обнаружено, что артикуляция согласных *ф*, *к*, *с* в звуковых сочетаниях *фс* и *ск* выше, чем артикуляция этих же согласных вне консонансов (т. е. сочетаний с другими согласными). Так, в одном из опытов с искажающим трактом артикуляция *с* вне консонансов была равна 66,6%, в сочетании с *ф* — 89,8%, в сочетании с *к* — 94,7%¹.

Расхождение между артикуляцией согласных в консонансах и вне консонансов объяснялось тогда тем общезвучающим положением, что акустическая характеристика отдельного согласного и того же согласного в составе консонанса должна быть различной. Однако оставалось непонятным, почему артикуляция согласного в консонансах оказывалась всегда выше, чем в других условиях. Последнее обстоятельство может быть истолковано, по-видимому, только с точки зрения теории информации, а именно в том смысле, что соседний согласный содержит в русской речи большее количество избыточной информации о данном согласном, чем соседний гласный. Что же касается консонансов *фс* и *ск*, то они по сравнению с другими консонансами обладают относительно высокой вероятностью. Этим можно объяснить столь резкое повышение артикуляции в них, наблюдавшееся в указанном опыте.

Подобно тому, как это имеет место при лексической вероятности, увеличение объема звукового сочетания влечет за собой увеличение количества избыточной информации относительно каждого из сочетающихся звуков. Совершенно очевидно, например, что появление *р* после *ост* ожидается с гораздо большей вероятностью, чем после *с* и даже чем после *ст*, а следовательно, в *ост* содержится больше избыточной информации относительно *р*, чем в *ст*, а в *ст*, в свою очередь, больше, чем в *с*. Таким образом, можно различать звуковую вероятность «первого порядка», когда учитывается сочетание двух звуков, «второго порядка», когда учитывается сочетание трех звуков, и т. д.

¹ Л. Р. З и п д е р, Русские артикуляционные таблицы, «Труды [Военной краснознаменной академии связи им. С. М. Буденного]», сб. 29—30, Л., 1951, стр. 37—38.

На кафедре фонетики Ленинградского университета им. А. А. Жданова в 1956 г. была выполнена работа по определению звуковой вероятности «первого порядка» для русской речи. В качестве исходного материала по соображениям, изложенным выше, были взяты статьи из текущих газет, а также произведения современных советских писателей (Горбатова, Гайдара, Тендрякова, Паустовского, Пановой и др.). Общий объем материала — 88 538 фонем. После фонетического транскрибирования всех текстов был произведен подсчет того, сколько раз в них встретилось каждое из возможных в русском языке сочетаний, а затем вычислены вероятности каждого сочетания.

Звуковая вероятность внутри слова определяется частотностью употребления слов и морфологической структурой их; на стыке же слов она определяется сочетаемостью слов и синтаксическими правилами, т. е. лексической и грамматической вероятностью. Соответственно были произведены подсчеты и составлены две таблицы.

Кроме указанных видов звуковой вероятности, была определена вероятность следования друг за другом отдельных групп звуков внутри слова: гласных и согласных; шумных и сонантов; взрывных и щелевых; носовых и плавных. При этом оказалось, что вероятность появления гласного после согласного равна 0,7449, а согласного после согласного — 0,2551; вероятность же появления гласного после гласного равна 0,0017 а согласного после гласного — 0,9983.

В приводимых ниже таблицах представлена вероятность следования друг за другом разных групп согласных. В первой таблице за единицу принято сочетание «согласный + согласный», во второй — сочетание «согласный + любой звук (гласный или согласный)»¹.

Таблица 1

Последующий Предшествующий	<i>C</i>	<i>M</i>	<i>S</i>	<i>E</i>	<i>F</i>	<i>N</i>	<i>L</i>
<i>C</i>	1,0000	0,5695	0,4305	0,3385	0,2310	0,1800	0,2505
<i>M</i>	0,7850	0,4278	0,3572	0,2732	0,1546	0,1161	0,2411
<i>S</i>	0,2150	0,1417	0,0733	0,0653	0,0764	0,0639	0,0094
<i>E</i>	0,3303	0,0981	0,2322	0,0350	0,0631	0,0420	0,1902
<i>F</i>	0,4547	0,3297	0,1250	0,2382	0,0915	0,0741	0,0509
<i>N</i>	0,0878	0,0471	0,0407	0,0230	0,0241	0,0358	0,0049
<i>L</i>	0,1272	0,0946	0,0326	0,0423	0,0523	0,0281	0,0045

Таблица 2

Последующий Предшествующий	<i>M</i>	<i>S</i>	<i>E</i>	<i>F</i>	<i>N</i>	<i>L</i>
<i>C</i>	0,2551	0,1453	0,1098	0,0864	0,0589	0,0639
<i>M</i>	0,2003	0,1092	0,0911	0,0697	0,0395	0,0615
<i>S</i>	0,0548	0,0361	0,0187	0,0167	0,0194	0,0024
<i>E</i>	0,0842	0,0250	0,0592	0,0089	0,0161	0,0485
<i>F</i>	0,1161	0,0842	0,0319	0,0608	0,0234	0,0189
<i>N</i>	0,0224	0,0120	0,0104	0,0059	0,0061	0,0013
<i>L</i>	0,0324	0,0241	0,0083	0,0108	0,0133	0,0011

Л. Р. Зиндер

¹ Буквы в таблицах обозначают: *C* — согласный, *M* — шумный, *S* — сонант, *E* — взрывной, *F* — щелевой и аффрикату, *N* — носовой, *L* — плавный (включая *й*).