

УДК 811.161.1`342:004

ФОНЕТИЧЕСКИЙ РЕЧЕВОЙ МОДУЛЬ КАК СПОСОБ ОБРАБОТКИ ЗВУКОВОЙ ИНФОРМАЦИИ

© 2013 г.

А.М. Грачев

Нижегородский госуниверситет им. Н.И. Лобачевского

grachol@yandex.ru

Поступила в редакцию 20.05.2013

Описывается речевая сенсорная модальность – фонетический, или речевой, модуль, предназначенный исключительно для обработки информации, содержащейся в звуках, идентифицированных как речь. При анализе использованы мнения и точки зрения наиболее авторитетных учёных-лингвистов. Отмечено, что теория фонологического восприятия речи должна лежать в основе экспериментального моделирования автоматического распознавания речи.

Ключевые слова: речевой модуль, произношение, идентификация, информативность, фонема.

Звуки речи образованы сигналами, имеющими различные физические свойства (т. е. разную интенсивность, частоту, сложность и продолжительность). Чтобы «раскрылась» содержащаяся в речевых звуках информация, они, как и все сложные символы, должны претерпеть некую специальную обработку. Восприятие речи требует от слушателя не только извлечения информации о наличии тех или иных звуков или об их местоположении, но и раскрытия смысла — семантического послания, закодированного в звуках. Из этого следует, что восприятие речи отличается от восприятия других форм аудиальной стимуляции и потому требует своего собственного, специального механизма обработки информации.

Это обстоятельство и явилось предпосылкой создания на биологической основе единой теории, согласно которой нервная система человека имеет специальный «процессор» — **речевую сенсорную модальность — фонетический, или речевой, модуль**, предназначенный исключительно для обработки информации, содержащейся в звуках, идентифицированных как речь. Одни и те же акустические сигналы обрабатываются и воспринимаются по-разному в зависимости от того, принимает ли их слушатель за речь или за информацию другого рода [1].

Помимо того что одни и те же речевые звуки могут при разных обстоятельствах восприниматься и как речь, и как неречевая стимуляция, при определенных условиях один и тот же физический стимул одновременно способен возбуждать обе сенсорные модальности — и неречевую (психоакустическую), и речевую (фонетическую). Подобное одновременное восприятие звуковых сигналов в качестве речи и в ка-

честве неречевой (акустической) стимуляции называется двойственным восприятием (*duplex perception*). Полагают, что двойственное восприятие отражает различия, существующие в психоакустическом и лингвистическом механизмах восприятия одного и того же стимула.

Приобретение навыков восприятия речи начинается в детстве и происходит постепенно. Хотя младенцы сначала воспринимают только физические или психоакустические свойства речи, они демонстрируют предрасположенность к различению звуков, соответствующих различным фонемам, и к научению тем конкретным фонемам, которые важны для коммуникации.

Овладевая языком, маленькие дети в определенный момент начинают принимать во внимание фонологическую (язык—звук) структуру и другие грамматические и синтаксические (связанные со структурой предложения) правила языка и по-разному воспринимать один и тот же акустический стимул. В результате классификация речевых звуков, основанная на физических (или психоакустических) свойствах последних и характерная для маленьких детей, постепенно сменяется классификацией, основанной на лингвистической релевантности звуков. Следовательно, развитие речи есть результат приобретения языковых навыков и попыток придать лингвистический смысл акустической стимуляции.

При этом слуховая система человека выделяет в речевом сигнале только фонетически полезные признаки, создающие базу для лингвистической интерпретации сообщения, своеобразные **акустические ключи** к звуковому распознаванию речи. Умение человека пользоваться этими акустическими ключами при восприятии

речи рассматривается как «фонологическое воспитание слуха» [2].

В науке о языке в этом случае принято говорить о **фонологичности речевого слуха**. Иными словами, мы слышим не какофонию звуков, источниками которых являются не связанные между собой различные объекты, а способны разобраться в них и идентифицировать различные звучащие объекты как отдельные события — мы перцептивно анализируем сложную аудиальную сцену и раскладываем ее на компоненты.

«Чтобы понять, в каком смысле речевые звуки являются фонемным кодом, необходимо прежде всего определить, какие свойства комплексного акустического сигнала лежат в основе восприятия звуковых типов — фонем. В связи с этим можно назвать акустические особенности, которые являются достаточными сигналами (своего рода подсказками) для восприятия почти всех сегментных звуковых типов — фонем» [3, с. 126].

На этапе языковой реконструкции сообщения слушающий восстанавливает цепочку фонем, которая была реализована в речевом сигнале. Отдельные звуки, слоги и псевдослова образуют наиболее удобный материал для определения акустических ключей, которые используются в этом процессе. Главным критерием при оценке фонетической полезности тех или иных акустических характеристик речевого сигнала является их вклад в идентификацию и различение звуковых единиц языка. Поэтому вопрос о наборе акустических ключей, которые использует человек для опознания фонемного состава речевого сообщения, требует исследования того, как носители данного языка классифицируют и различают те или иные речевые образцы. В то же время общие свойства целевых артикуляций основных классов звуков, закономерности коартикуляции и формирования речевого сигнала, а также особенности обработки звуковых сигналов человеческим слухом вносят в звуковое восприятие речи определенные универсальные черты, позволяя выделить главные типы слуховых признаков фонем.

Перцептивные эксперименты, направленные на поиск таких ключей, проводятся в двух направлениях. В одних тестах исследуется способность человека производить слуховое членение речевого сигнала, т. е. обнаруживать резкие изменения в его акустических свойствах и делить сигнал на отрезки, которые таких изменений не содержат. В других — изучается умение слушателей субъективно оценивать (измерять) акустические характеристики выделенных отрезков и соотносить эти оценки с фонемными единицами языка и их фонологическими признаками.

Результаты перцептивных тестов на материале синтезированных речеподобных стимулов подтвердили, что слух человека очень чувствителен к резким изменениям акустических параметров сигнала во времени. Например, если на протяжении звучания сигнала, подобного гласному, в какой-то момент времени происходит скачок амплитуды звукового давления, сигнал может восприниматься по-разному: как гласный, как двойной гласный или же как последовательность гласных и сонорных согласных (Г, ГГ, СГ, ГС). Конкретная фонетическая интерпретация зависит от величины амплитудного скачка, его скорости и положения относительно границ воспринимаемого сигнала. Так, для слогов типа СГ стабильность восприятия слоговой структуры и опознание ее отдельных элементов существенно выше в тех случаях, когда первый звук является шумным, а не сонорным согласным [2, с. 68]. Это говорит о различной чувствительности слуха к акустическим изменениям разного типа: на границе шумного согласного и гласного происходит не только более резкий скачок амплитуды, чем в слогах с сонорными, но и смена источника звука.

Трудности членения речевого сигнала на вокальных участках, соответствующих сочетаниям гласных с сонорными согласными, хорошо известны фонетистам. Эксперименты показывают, что при опознании слогов типа СГ наибольшее количество ошибок, связанных с опознанием слоговой структуры, приходится на слабые безударные слоги, которые часто воспринимаются как единый гласный [2, с. 70]. В этом случае фонетически полезными являются два типа акустических ключей: *перепад интенсивности* и *изменение спектра в низкочастотной области*, где на участке носового согласного появляется сильная форманта назализации и ослабляются ротовые форманты. Поэтому для сочетаний гласных с носовыми сонантами, особенно типа ГСГ, вероятность слухового членения выше, чем для сочетаний с такими сонорными, которые характеризуются плавным ходом формант (типа русских [л]).

«Исследователи, занимающиеся изучением слухового членения, считают, что обнаружение акустических изменений в сигнале приводит к образованию граничных маркеров (меток), которые используются слушающим для считывания и оценки спектральной информации в их окрестности. Предполагается также, что слух вырабатывает разные маркеры для положительных и отрицательных приращений амплитуды сигнала. В первом случае образуются метки начал акустических отрезков, во втором — метки концов. В частности, экспериментально по-

казано, что слуховая система с большой надежностью формирует метки начал и концов гласных в сигналах, соответствующих последовательностям гласных и шумных согласных. Образование таких маркеров является основой для измерения длительности отрезков сигнала, выделяемых слухом» [4, с. 236–237].

Перцептивная сегментация не совпадает с членением на звуки (фрагменты фонемной размерности) в лингвистическом понимании. Некоторые отрезки, выделяемые слухом, могут рассматриваться как линейные компоненты звуковых единиц. Например, на участках реализации глухих согласных в позиции между гласными обычно выделяются такие отрезки, как пауза и шум взрыва. В других случаях один перцептивный отрезок может соответствовать целой последовательности звуков. В то же время слуховое членение явно коррелирует с артикуляцией и резкими изменениями в работе источников звука и перестройками в передаточной функции речевого тракта. Излишняя детализация такой акустически ориентированной сегментации может быть устранена в процессе дальнейшей фонологической интерпретации, а недостаточная – едва ли.

Для любого языка звуковые противопоставления, связанные со способом образования, базируются на таких акустических ключах, которые обладают максимальной устойчивостью и надежностью обнаружения. Это касается прежде всего противопоставления шумных согласных и вокальных звуков, к которым относятся гласные и сонорные согласные, а также противопоставлений по способу образования для шумных согласных. Важная акустическая информация о признаке содержится в частотных характеристиках акустических изменений, на основе которых производится слуховое восприятие.

Экспериментально установлено, что метки слухового членения «привязаны» к определенным частотным областям. Например, метка начала гласного распространяется обычно на область низких частот, а метка, соответствующая началу шума, появляется именно в том частотном диапазоне, где он обнаруживается, т. е. в области средних или высоких частот. На вокальный характер отрезка указывают и такие свойства, как периодичность и наличие формантной структуры. Человеческий слух отчетливо фиксирует присутствие этих свойств в звуковом сигнале.

Противопоставление шумных согласных по способу образования базируется на ключах иной природы. Для восприятия согласных этих классов важна последовательность определен-

ных акустических отрезков и их соотношение по длительности. Так, например, взрывным и аффрикатам в позиции между гласными соответствует последовательность, состоящая из паузы (глухой или звонкой) и отрезка шума. Если длительность паузы больше или равна длительности шума, последовательность воспринимается как взрывной согласный. Опознание аффрикат основано на обратном соотношении длительностей отрезков. Появление же паузы (даже очень краткой) перед длительным шумом искажает восприятие фрикативных согласных. В начале слова (фразы) акустическую паузу глухих взрывных и аффрикат выделить невозможно. В этой позиции «работают» другие ключи: скорость нарастания энергии и характер ее изменения на начальном участке шума, а также длительность шумового отрезка.

Акустические ключи, используемые для идентификации гласных, содержатся, прежде всего, в формантной картине их собственных отрезков. Акустический анализ шумных согласных показывает, что от места образования зависят как физические свойства шумового отрезка согласного, так и характеристики переходов на граничных участках соседних гласных. В настоящее время большой популярностью пользуется гипотеза, согласно которой наиболее важные ключи для распознавания места образования шумных согласных (взрывных и фрикативных) содержатся в окрестности начала последующего гласного.

Диффузный шум без выраженной области усиления и увеличение интенсивности сигнала в области высоких частот (выше последующего гласного) на границе с гласным соотносится с восприятием губных согласных. Широкополосный шум с выраженной областью усиления в области последующего гласного (или выше) и уменьшение интенсивности в этой области при переходе к гласному соответствуют восприятию зубных и альвеолярных согласных. Узкополосный шум (ширина меньше критической полосы слуха) с выраженной областью усиления в области последующего гласного и мало изменяющаяся интенсивность в этой области при переходе к гласному приводят к восприятию постальвеолярных согласных. Велярным согласным соответствуют характеристики, аналогичные постальвеолярным, однако область усиления шума ниже вблизи последующего гласного. В качестве дополнительного признака может выступать интенсивность шума: переднеязычные при прочих равных условиях громче остальных согласных.

Для восприятия согласных, различающихся местом образования, важны и формантные

ключи, привязанные к граничным меткам соседних гласных. При отсутствии взрыва у взрывного согласного (такое бывает в естественной речи) они становятся единственными признаками, сигнализирующими противопоставления согласных по месту артикуляции.

В целом обращает на себя внимание разная картина соответствия между слуховыми и лингвистическими свойствами согласных и гласных.

Акустические ключи для согласных распределяются по разным отрезкам сигнала (взрывные согласные в положении между гласными могут иметь ключи в окрестности границ четырех отрезков. Они локализируются в основном в области границ перцептивных отрезков, и по крайней мере некоторые из них не зависят от звукового окружения. Для фонологических признаков согласных характерна множественность ключей. Иначе говоря, действует схема «много ключей—один признак», а набор ключей может зависеть от контекста. Когда ключей много, возникает вопрос о весе отдельных ключей для восприятия звукового противопоставления, а также о необходимом или допустимом балансе количественных отношений между ними.

Акустические ключи для фонологических признаков гласных находятся на отрезке самого гласного. Для них не характерна соотносительность с границами отрезка, а контекстная зависимость, напротив, является типичной чертой. Самих ключей немного, и почти все они связаны с картиной гласного [4, с. 242–246].

В настоящее время неизвестны механизмы сбора и объединения информации, передаваемой акустическими ключами, в фонологические признаки и фонемные единицы языка. Большинство исследователей считают, что такие механизмы действуют за пределами слуховой системы человека.

Человек не смог бы достаточно экономно и естественно хорошо воспринимать речь, если бы каждая фонема была закодирована в виде отдельной звуковой единицы.

Специализация полушарий мозга при восприятии звучащей речи обуславливает наличие двух основных компонентов речевого слуха — **фонематического** и **интонационного**. Несмотря на определенный параллелизм в механизмах фонематического и интонационного слуха (интонационные средства языка имеют ту же материальную природу, что и фонемы — волновые колебания определенной частоты, спектра, длительности и интенсивности), компоненты речевого слуха являются относительно самостоятельными образованиями.

«Фонематический слух — это способность слушающего различать фонемы, или смысло-различительные звуки данного языка, на кото-

рых основан звуковой анализ отдельных элементов речи (звуков, слогов, слов), т. е. способность к анализу и синтезу речевых звуков» [5, с. 54].

Способность различать звуки речи и идентифицировать их с фонемами языка обеспечивается наличием в долговременной памяти человека специального запоминающего устройства — «фонемной решетки» [6, с.145]). Для каждой фонемы в этом устройстве имеется эталон, точнее «эталонная зона», допускающая отклонения в определенных пределах. В «эталонной зоне» представлены информативные (фонологически значимые) признаки фонемы. При восприятии звука фиксируется наличие или отсутствие определенных информативных признаков (просеивание их через «фонологическое сито»), и звук относится к той или иной фонеме.

Сформированный фонематический слух предполагает умение: 1) слышать отдельные звуки в слове, разделять слова на звуки и составлять из заданных звуков слова; 2) различать на слух все звуки языка, относящиеся к разным фонемам, не смешивая их друг с другом; 3) соотносить услышанный звук с определенной фонемой языка; 4) перекодировать физические звуки в фонемы; 5) мгновенно вычленять из речевого потока осмысленные звуковые комплексы, сличать их, дифференцировать, идентифицировать; 6) образовывать акустико-артикуляционные образцы предъявляемых сигналов с опорой только на слуховое ощущение; 7) удерживать предъявляемые сигналы (звуки) в кратковременной памяти; 8) внутренне проговаривать предъявленный звуковой образец; 9) соотносить предъявленные сигналы с определенным смысловым значением.

Помимо фонематического, в речевом слухе выделяют интонационный. «При восприятии на слух фраза как одна из единиц восприятия понимается в результате распознавания информативных признаков. Интонация — наиболее информативный признак, т. к. она обладает воспринимаемыми качествами, благодаря которым аудитор может сегментировать речь на синтаксические блоки, понять связь частей фразы и раскрыть содержание». [7, с. 33]. Под интонационным слухом понимается «способность аудитора воспринимать на слух интонационную структуру фразы и правильно соотносить ее с интонационным инвариантом языка» [8, с. 178]. Определенные «интонационные компоненты специфичны для каждого языка так же, как специфическими являются фонематические признаки. С помощью интонации передается большой объем информации, в том числе эмоциональное содержание высказывания, и, конечно, отношение самого субъекта к тому, о чем именно он говорит.

Интонационный слух вырабатывает одновременно с выполнением указанных функций индикаторные признаки (эталон), которые необходимы для успешной деятельности кратковременной памяти. При помощи интонационного слуха слушающий может делить речевой поток на смысловые куски, устанавливая связь между этими кусками и, наконец, понимать содержание прослушанного материала.

Сформированный интонационный слух предполагает умение: 1) различать физические свойства и воспринимаемые качества интонации; 2) соотносить интонацию со смысловым содержанием речи; 3) различать коммуникативные типы предложений; 4) дифференцировать интонационные особенности языка; 5) соотносить интонацию с контекстом и ситуацией; 6) распределять внимание между содержанием и формой.

Интонационные характеристики в каждом языке организованы в систему функциональных единиц, которые распознаются и декодируются посредством перцептивных представлений. Эти представления сами по себе являются результатом воздействия физических стимулов, которые достигают уха слушающего. В обычных условиях именно изменения высоты основного тона способствуют восприятию слушающим данного тона или интонации. В речи это относительная высота. Но не только изменения высоты основного тона способствуют распознаванию интонационных структур и тонов. Этому способствуют также изменения громкости, длительности, а также тембра.

Например, слушающий получает сообщение в двух реализациях: в одном случае он распознает высокое падение, в другом – нисходяще-восходящую интонацию. Естественно, его восприятие будет сильно детерминировано в обычных условиях коммуникации изменениями высоты основного тона, которые встречаются на протяжении всего высказывания. Однако они будут также подвергаться влиянию различий в изменениях громкости. Изменения длительности также имеют значение. Наконец, будет изменение и в тембре, по крайней мере в конечной гласной, которая может быть более закрытой или более дифтонгизированной. Восприятие слушающим, его лингвистическое решение в каждом отдельном случае является результатом всех этих перцептивных факторов, вместе взятых.

Восприятие тембра также способствует распознаванию ритмического рисунка практически во всех языках, так как существует почти универсальная тенденция для качества гласного, который будет более открытым или более центральным в более слабых слогах, чем в удар-

ных. Этот эффект очень распространен в русском языке, где существует постоянное качественное различие, когда один и тот же гласный встречается в ударном и безударном положениях.

Фонематический и интонационный слух, являясь компонентами одного явления, функционируют взаимосвязанно. А.А. Леонтьев видит сходство фонетического и интонационного слуха в том, что они «вырабатываются онтогенетически, т. е. усваиваются каждым человеком в ходе его развития. В обоих случаях человек имеет дело с единой природой воспринимаемых свойств раздражителей. Общность природы соответствующих психических функций позволяет человеку допустить известную аналогию в построении системы фонематического и интонационного слуха» [9, с. 105].

В целом отметим, что именно теория фонологического восприятия речи должна лежать в основе экспериментального моделирования автоматического распознавания речи, чтобы в результате можно было построить адекватные модели распознавания речи, имеющие определенные аналогии с естественным восприятием речи человеком.

На основе проведенного анализа можно сделать следующие выводы.

Восприятие речи основано не на фонетическом, а на фонематическом членении речи, при котором учитывается сигнификативное членение звучащей речи на смыслоразличительные элементы. В этом плане важно определить принципы классификации наблюдаемых в потоке речи единиц, которые позволили бы отделить значимые звуковые различия от незначимых.

Восприятие речи отличается от восприятия других форм аудиальной стимуляции и потому требует своего собственного, специального механизма обработки информации. Нервная система человека имеет специальный «процессор» – речевую сенсорную модальность – фонетический, или речевой, модуль, предназначенный исключительно для обработки информации, содержащейся в звуках, идентифицированных как речь. Таким образом, речевой слух фонологичен по своей природе.

Поэтому теория фонологического восприятия речи должна лежать в основе экспериментального моделирования автоматического распознавания речи.

Список литературы

1. Шиффман Х.Р. Ощущение и восприятие. СПб.: Питер, 2003. 928 с.
2. Бондарко Л. В. Фонетическое описание языка и фонологическое описание речи. Л.: ЛГУ, 1981. 200 с.

- 3 Златоустова Л. В., Потапова Р. К., Потапов В. В., Трунин-Донской В. Н. Общая и прикладная фонетика. 2-е изд., перераб. и доп. М.: Изд-во МГУ, 1997. 50 с.
4. Кодзасов С.В., Кривнова О.Ф. Общая фонетика: Учеб. пособие. М.: РГГУ, 2001. 592 с.
5. Фролова И.А. Речевой слух – основа правильного произношения // Актуальные проблемы обучения русскому языку как иностранному и дисциплинам специализации: Сб. науч. тр. / Под ред. В.Д. Виноградова. Н.Новгород, 1999. С. 54–57.
6. Жинкин Н.И. Механизмы речи. М.: Изд-во Акад. пед. наук СССР, 1958. 370 с.
7. Гез Н.И. К вопросу об обучении распознаванию речи при зрительном и слуховом восприятии // Иностранные языки в школе. 1962. № 2. С. 27– 37.
8. Сайдалиев С.О. О взаимосвязи фонематического и интонационного слуха при обучении восприятию русской речи на начальном этапе языкового вуза // Теоретические и экспериментальные исследования в области обучения иноязычной речевой деятельности: Сб. науч. тр. / Под ред. Ю.А. Дементьева. М., 1981. Вып. 180. С. 172–187.
9. Леонтьев А.А. Язык, речь, речевая деятельность. М.: Просвещение, 1969. 214 с.

SOUND DATA PROCESSING BY MEANS OF PHONETIC VOICE MODULE

A.M. Grachev

The subject of this paper is speech sensation modality – a phonetic or speech module intended for processing pieces of sound data identified as speech. The ideas and views of some of the most reputed linguists have been used in the course of analysis. The theory of phonologic speech reception is accepted as the basis for experimental modeling of computer-aided speech reception.

Keywords: voice module, pronunciation, identification, information capacity, phoneme.