

## НЕКОТОРЫЕ АСПЕКТЫ ПРОБЛЕМЫ ВОСПРИЯТИЯ РЕЧИ

In the article results of the analysis of the scientific researches devoted to a problem of speech perception, developed in the 60-70<sup>th</sup>, are shown. The main principles of modern representations about process of perception of speech are considered. The received data allow to track the tendencies in the manufacture of the experimental methods of research of speech perception process.

Успешность дальнейшего развития образования во многом зависит от наличия и качества соответствующих электронных ресурсов. Современные информационные технологии, успехи которых обусловлены достаточно высокими темпами программно-технического оснащения и модернизации, объективно вовлекают в сферу своего влияния и такую отрасль знаний, как педагогика, которая до недавнего времени в определенной степени отстаивала "чистоту" своего инструментария и теоретико-методологических основ. На рынке компьютерных продуктов с каждым годом возрастает число обучающих программ, электронных учебников и т. п. Однако мультимедийная книга как новый способ обучения часто разрабатывается в рамках традиционной и преобладающей в области создания электронных учебников методологии, когда осуществляется полный перенос текстов полиграфических изданий учебного или дидактического характера, написанных профессиональными предметниками, на CD-ROM. При этом фактор субъективного восприятия учащимися программной оболочки, которая, во-первых, была бы привычна и удобна для изучения данного конкретного предмета, а, во-вторых, наиболее полно использовала бы преимущества работы на компьютере, либо не принимается во внимание, либо не оказывает влияния на итоговый результат. Исходя из необходимости учета этих требований, электронное учебное издание следует разрабатывать таким образом, чтобы оно превосходило по объему включенной информации свой полиграфический прообраз, было удобно в использовании в качестве справочно-учебного издания, так как содержало бы обширную разветвленную структуру. При этом полностью сохраняется авторский взгляд на изложение материала, как иллюстративного, так и текстового, и производится значительное его наращивание. Дополнительный материал должен присутствовать в таком издании в специфически "мультимедийной" форме (иллюстрации — в виде анимации или видеосфрагментов; текстовый материал — в качестве звуковых дополнений, комментариев и справок), которая не нарушала бы "исходный" авторский вариант подачи материала.

Одно из важнейших направлений развития компьютерных технологий связано с использованием речевых технологий и интеллектуальных речевых интерфейсов.

Еще в 1983 г. в предисловии к известному монографическому изданию "Методы автоматического распознавания речи" один из американских специалистов У. Ли пророчески заявил: "... нет никаких сомнений в том, что речевое взаимодействие будет преобладающим направлением развития вычислительной техники".

Необходимость исследования восприятия речи подчеркивала и российский ученый Р. К. Потапов, утверждая, что психологические эксперименты, проведенные над человеческой памятью, показали, что речь играет более важную и существенную роль, чем зрение, применительно к памяти и мышлению [1].

Такая точка зрения расходится с более ранними исследованиями особенностей освоения нового материала через зрительный, слуховой, а также тактильный анализаторы. Под анализаторами мы понимаем своего рода окна, через которые мозг человека связывается с внешней средой. Через сенсорные системы мозг получает информацию об окружающем мире, а также о состоянии внутренних органов.

Роль анализаторов в высшей нервной деятельности больше всего изучалась на животных в физиологических лабораториях, возглавляемых И. П. Павловым. Образование условных рефлексов у собак, как показали П. С. Купалов и В. Гент (1928), происходит быстрее на звуковые раздражители, чем на световые и кожно-механические. Степень генерализации процессов в слуховом анализаторе выше, чем в зрительном (П. К. Анохин, 1941). Н. С. Поповой (1961, 1962) проведено сопоставление особенностей образования условных рефлексов в слуховом и зрительном анализаторах собак и низших обезьян. Ею установлено, что различие между анализаторами у собак выступает более отчетливо, чем у обезьян.

При изучении высшей нервной деятельности ребенка отмечалось, что в раннем возрасте скорость выработки условных рефлексов с разными анализаторами неодинакова. Но по мере того, как ребенок становится старше, различия становятся менее выраженными (Н. И. Касаткин, 1951; З. И. Барбашева, 1955; М. М. Кольцова, 1958).

У детей старшего возраста почти с одинаковой скоростью образуются динамические стереотипы на световые и звуковые раздражители (Е. Н. Дегтярь, 1962).

И. А. Кулаком было опубликовано несколько работ [2, 3, 4], в которых изучались особенности освоения нового материала через зрительный, слуховой и тактильный анализаторы. Проанализировав отдельные исследования с точки зрения изучения роли различных нервных образований головного мозга в освоении нового материала, ученый пришел к выводу, что в процессах формирования сложных систем связей через изученные анализаторы существенной разницы нет. Анализаторы не вносят особых изменений не только в продуктивность освоения, но и в процесс формирования систем временных связей. Через любой из указанных анализаторов в мозгу человека фиксируется одинаковый объем поступающей информации.

Аналогичные исследования провела Л. А. Гуринович [5] на людях пожилого возраста (50–75 лет). Формирование сложных систем через три анализатора в этом случае идет с различной скоростью в начале исследования (у большинства испытуемых освоение материала быстрее шло через зрительный анализатор), а под влиянием тренировки различия в продуктивности освоения нового материала через разные анализаторы сглаживаются.

Сосредоточим внимание на исследованиях речевого сигнала. Голос и речь являются основными и сугубо специфическими средствами коммуникации в человеческом обществе. Звуковой язык, обусловивший неограниченное развитие социального опыта и тем самым человеческого общества, является первичной и основной формой общения между людьми. Все другие средства (письменная речь, речь глухонемых, разные коды и т. д.) носят по отношению к устной речи вторичный характер и возникли на ее основе.

Голос и речь человека имеют непосредственную связь с внутренней жизнью личности. Поэтому голос человека не может не отличаться богатством оттенков, многими ступенями изменчивости, обилием структурных элементов. Индивидуальные особенности разговорного голоса сделали его особо интересным объектом для изучения в работах А. С. Рылова [6]. Именно с наличием индивидуальных особенностей он связывает разнообразие информации, имеющейся в реальном речевом процессе, а также особенности ее восприятия.

В первую очередь следует отметить, что всякая речевая активность содержит конкретную смысловую информацию, роль которой трудно переоценить в процессе коммуникативных актов между людьми. Однако речевая деятельность не сводится просто к механическому использованию средств языка для передачи какого-либо сообщения. Исходным моментом для речевого общения является мотив высказывания или потребность в передаче какого-либо

сообщения. Затем путем внутреннего программирования общий замысел высказывания в соответствии с имеющимися у данного лица физиологическими, психическими, социальными и ситуационными стереотипами кодируется в форме предметно-образного языкового кода, который и передается посредством речевого сигнала. Поэтому речевой сигнал как материальный носитель передаваемого смысла содержит комплексную информацию, а именно информацию о семантике высказывания и информацию о свойствах личности. Последнюю В. И. Галунов в статье “Речь, эмоции и личность” [7] называет дополнительной по отношению к основной (семантической). К дополнительной информации он относит:

- индивидуальные особенности голоса и речи;
- информацию, характеризующую эмоциональное состояние говорящего;
- обликовые (личностные) характеристики человека (социальный статус, возраст, пол и некоторые другие особенности).

Таким образом, особенности восприятия речи определяются во многом характеристиками голоса и особенностями речи как неотъемлемыми свойствами личности, заложенными в свойствах акустического сигнала и особенностях использования языковых средств для формирования устного текста. При этом акустические особенности голоса и речи, ее моторика обусловлены физиологическими характеристиками артикуляционного аппарата, а также врожденными и приобретенными навыками, общей психической и социальной активностью человека.

Как смысловая, так и дополнительная виды информации, заключенные в речевом сигнале, даются в едином комплексе. Поэтому вряд ли можно найти такие характеристики, которые бы, например, всецело относились к индивидуальности голоса и не участвовали бы в выражении семантики или при оценке эмоционального состояния говорящего.

То, что речевой сигнал может быть очень избыточным, утверждает Ю. А. Косарев [8]. Это действительно так, если говорить только о передаче семантики речи. Это явление явилось следствием экспериментов Дж. Фланегана [9] по оценке скорости восприятия и передачи информации человеком. Он сделал вывод, что человек не способен обрабатывать информацию со скоростью, превышающей 50 бит/с. Такой пропускной способности канала вполне достаточно для восприятия семантики речи, передаваемой с помощью 42 фонем ( $2^6$ ) русского (или любого другого) языка.

Надо полагать, что вывод Фланегана имеет прямую связь с другим известным фактом, установленным П. Линдсеем и Д. Норманном [10] в результате психоакустических и других экспериментов, согласно которым человек в отличие от

животных обладает кратковременной (оперативной) памятью емкостью приблизительно  $7 \pm 2$  единиц, предъявляемых для запоминания. Очень упрощенно функционирование этой кратковременной памяти может быть представлено так. Любой предъявляемый материал подвергается обработке в сенсорной системе и интерпретируется на разных уровнях механизма распознавания образов. Затем опознанный образ предъявленного материала вводится в одну из "пустых ячеек" кратковременной памяти. При этом если количество "ячеек" ограничено, например, цифрой 7, то при введении восьмой единицы одна из семи заполненных ячеек изменит свое содержание. Из-за малости объема такой оперативной памяти у человека возникают ограничения на структуру и степень сложности процессов мышления, а также восприятия информации, поскольку ему необходимо контролировать продвижение процесса решения задачи и получение промежуточных результатов. В принципе возможно предсказать, когда человек окажется неспособным решить ту или иную задачу просто потому, что он не в состоянии уследить за всеми событиями, происходящими в данный момент.

Более детально феномен наличия оперативной памяти у человека и ее объем изучались Л. А. Чистович и ее коллегами [11]. При этом в качестве единиц, запоминаемых в оперативной памяти, использовались слоги как единицы речи. В этих психоакустических экспериментах слоги предъявлялись изолированно, с паузами между ними. Было сделано предположение о том, что в случае сильно искаженных фраз человеку требуется значительное время для перехода от запомненной последовательности фонетических элементов к решению о смысле сообщения. Если это так, то, учитывая ограниченный объем оперативной памяти, предъявление длинных фраз должно приводить не к улучшению разборчивости, как это имеет место, а наоборот, к ее ухудшению за счет пропуска части фразы.

Повышение разборчивости с увеличением длины фразы должно наблюдаться до тех пор, пока длина фразы не превысит объем оперативной памяти на запоминание последовательности слогов. Искажения создавались путем пропуска речевого сигнала через полосовой фильтр (906–1146 Гц). Зависимость фразовой и словесной разборчивости от числа слогов во фразе показывает, что разборчивость несколько возрастает в пределах до 7 слогов, затем резко падает. Это падение обусловлено преимущественно пропуском слов.

Аналогичные результаты были получены в экспериментах Хурмэ по имитации речевого сигнала без понимания его смысла, когда фразы предъявлялись аудиторам на финском языке, при этом аудиторы не знали финского языка.

В экспериментах Д. М. Лисенко и Н. А. Федоровой [12] объем оперативной памяти человека исследовался на восприятии ритмического рисунка последовательности слогов (tatata...) в количестве от 5 до 13.

Таким образом, все приведенные выше данные достаточно хорошо согласуются между собой и позволяют считать, что объем оперативной памяти на восприятие фонетических элементов действительно составляет примерно семь единиц. Эксперименты Лисенко и Федоровой показали, что, кроме восприятия семантики, закодированной в виде фонетического состава высказывания, человек может переключить свой пятидесятибитный канал и оперативную память на восприятие любой другой дополнительной информации.

Полная модель восприятия речи должна обеспечивать преобразование речевого сигнала в описание смысла сообщения. Одно из фундаментальных положений науки о языке заключается в том, что переход к смыслу сообщения возможен лишь после того, как речевой сигнал преобразован в последовательность дискретных элементов. Для понимания смысла речевого сообщения необходима фонемная и просодическая информация, в последнюю входит все воспринимаемое в речевом сигнале, помимо последовательности фонем.

Выше было отмечено, что речевой сигнал несет в себе многозначную информацию, выполняя коммуникативные функции между индивидуумами, а также заключает в себе информацию, всесторонне характеризующую самого индивидуума.

Анализируя причины, обуславливающие индивидуальность устной речи, в первую очередь нужно назвать зависимость ее характеристик от особенностей строения и функционирования речевого аппарата личности. Но не только анатомо-физиологические особенности влияют на параметры речи, важная роль здесь принадлежит социальному опыту человека, его общению с внешним миром.

Факторы, формирующие индивидуальность устной речи, можно представить следующим образом [13]:

- социальные факторы;
- анатомо-физиологические факторы;
- психические факторы.

На восприятие речевой информации оказывают влияние такие параметры устной речи, как:

- соотношение звучащих отрезков и пауз, т. е. длительность синтагм и плавность интонации. Интонация характеризуется плавностью в тех случаях, когда физиологические паузы в речи (необходимые для вдоха воздуха) совпадают с лингвистически обусловленными, служащими для разграничения синтаксических единиц друг от друга;

— сила голоса, параметр речи, производный от интенсивности выдыхания;

— высота голоса — зависит от длины, толщины и натяжения голосовых связок (от частоты образующихся звуковых колебаний). Вариация основной частоты голоса — основного тона — в потоке речи имеет большое значение для выражения интонации предложения, ударения, а также эмоционального отношения говорящего к предмету разговора;

— тембр голоса, зависящий от числа и выраженности обертонов (или добавочных тонов) в данном звуке. Возникновение обертонов связано с тем, что голосовые связки колеблются не только всей своей длиной, воспроизводя основной тон, но и отдельными частями.

Образование звуков речи (артикуляция) происходит в результате деятельности артикуляционного аппарата, в состав которого входят активные и пассивные органы произношения. Активные органы произношения (нижняя челюсть, губы, язык, мягкое небо), изменяя в процессе речи свою форму и положение, образуют совместно с пассивными (твердое небо, задняя стенка глотки, зубы, полость носа) специфическую конфигурацию речевого тракта, обуславливающую качество речевого звука, его акустические особенности.

Нарушение артикуляции можно свести к двум основным явлениям: нарушение произношения звуков и нарушение скорости артикуляции, выражающейся в патологически измененном темпе речи.

Роль параметров устной речи в восприятии информации в настоящее время исследована слабо, поэтому нуждается в разработке.

Таким образом, автоматически выделяя акустические параметры речевых сигналов, подробно изучая вышеописанные индивидуальные характеристики голоса и речи, совокупность значений которых в информационных системах задается множеством вариантов, можно решать различные задачи влияния этих параметров на восприятие речи.

В заключение хотелось бы отметить, что, несмотря на все сделанное в области звукового восприятия речи, в сегодняшних условиях стремительного увеличения потока информации особое внимание следует уделить достижению

точности, эффективности и психофизиологического комфорта восприятия информации при заданных речевых параметрах.

## Литература

1. Потапова Р. К. Речь: коммуникация, информация, кибернетика: учеб. пос. Изд. 3-е, стер.— М.: Едиториал УРСС, 2003.— 568 с.

2. Кулак И. А. Формирование сложных систем временных связей у человека. — Мн.: Изд-во АН БССР, 1962. — 230 с.

3. Кулак И. А. Психофизиологические и физиологические функции организма человека и системы НОТ. — Мн.: Беларусь, 1974 — 256 с.

4. Кулак И. А. Психофизиологические принципы обучения: Функциональные возможности головного мозга в восприятии и переработке информации. — Мн.: Изд-во БГУ, 1981. — 287 с.

5. Гуринович Л. А. Функционирование анализаторов при формировании сложных систем временных связей у людей пожилого возраста // Изв. АН БССР. Сер. биол. наук, 1966. — № 3. — С. 93–97.

6. Рылов А. С. Анализ речи в распознающих системах. — Мн.: Бестпринт, 2003. — 264 с.

7. Галунов В. И. Речь, эмоции и личность: проблемы и перспективы // Речь, эмоции и личность: Мат-лы и сообщ. Всесоюз. симп. — Л., 1978. — С. 3–21.

8. Косарев Ю. А. Естественная форма диалога с ЭВМ. — Л.: Машиностроение, 1989. — 143 с.

9. Фланаган Д. А. Анализ, синтез и восприятие речи: пер. с англ. / Под ред. А. А. Пирогова. — М.: Связь, 1968. — 396 с.

10. Линдсей П., Норманн Д. Переработка информации у человека: пер. с англ. / Под ред. А. Л. Лурия. — М.: Мир, 1974. — 550 с.

11. Физиология речи. Восприятие речи человеком / Л. А. Чистович и др. — Л.: Наука, 1976. — 388 с.

12. Лисенко Д. М., Федорова Н. А. К вопросу о временных ограничениях слуховой сегментации // Автоматическое распознавание слуховых образов: Тез. докл. 8-го Всесоюзн. семинара. — Львов, 1974. — Ч. 2. — С. 49–51.

13. Идентификация лиц по фонограммам русской речи: Пособие для экспертов / Н. Ф. Попов и др. — М.: ВЧ 34435, 1996. — 102 с.