

Лингвистические исследования на АРСО (обзор)

Надеина Т.М.

В вышедших ранее номерах журнала «Речевые технологии» были опубликованы статьи Н.Г. Загоруйко (Речевые технологии. 2008. №3. С. 81–96), В.В. Люблинской и Э.И. Столяровой (Речевые технологии. 2010. №2. С. 19–44), в которых рассказывалось о работе Всесоюзной школы-семинара «Автоматическое распознавание слуховых образов» (АРСО 1965–1992 гг.). В этих статьях анализировались основные итоги работы математиков, инженеров, физиологов в области звучащей речи и речевых технологий в СССР.

Данная статья посвящена лингвистическим исследованиям, выполненным в русле проблем распознавания и синтеза речи, а также пограничных научных проблем. Задача статьи заключается, прежде всего, в том, чтобы обозначить полный спектр проблем, которые находились в фокусе внимания лингвистов того времени.

Сборники материалов Всесоюзной школы-семинара «Автоматическое распознавание слуховых образов» АРСО в настоящее время стали библиографической редкостью. К счастью, нам удалось получить доступ к части из них¹, и этого оказалось достаточно, чтобы охарактеризовать весьма широкий спектр научных проблем, над которыми работали лингвисты, представлявшие коллективы Московского государственного университета им. М.В. Ломоносова, Санкт-Петербургского государственного университета (тогда Ленинградского), Московского государственного лингвистического университета (тогда Института иностранных языков им. М. Тореза), университетов Минска, Киева, Одессы, Еревана и других городов бывшего СССР.

К числу концептуальных докладов, определивших основные направления лингвистических исследований в области АРСО, следует, прежде всего, отнести доклад Л.Р. Зиндера «Фонетические основы автоматического распознавания речи», в котором обосновывается идея о том, что процесс распознавания речи требует привлечения **высших уровней языка** и всего языкового опыта человека. Человек прогнозирует соответствующее языковое явление, используя при этом связи, существующие между всеми ярусами языка. Лёгкость, с которой воспринимается родная речь, объясняется исключительной избыточностью языка, создаваемой смыслом и вероятностной структурой высказывания. Только этим объясняется тот факт, что при огромном разнообразии и вариативности акустических характеристик речи (в частности, и фонем) осуществляется её успешное и надёжное распознавание человеком. [Зиндер, 1971. С.33–34]

В.Б. Касевич в докладе «О критериях адекватности модели распознавания речи» подчёркивает, что построение адекватной модели распознавания речи требует учёта существенных свойств человеческого **восприятия**. По его мнению, именно дальнейшее изучение восприятия речи сможет помочь в создании функциональных аналогов соответствующих естественных механизмов. [Касевич, 1974. С.18–20]

В докладе В.С. Файна «Речевой процесс и общая проблема изменчивости» высказывается мысль о том, что для достижения стопроцентной надёжности распознающая система

¹ См. список литературы к статье.

должна быть снабжена способностью **понимания смысла**, и этим должно определяться генеральное направление речевых исследований. [Файн, 1974. С.24–26]

Тем самым в уже начале 70-х годов прошлого века были намечены основные направления лингвистических исследований в области автоматического распознавания и синтеза речи.

I. Уровни языка

Большая часть лингвистических исследований была посвящена единицам **фонетико-фонологического уровня языка** и нацелена на изучение сегментных характеристик согласных и гласных, явлений коартикуляции, слога, просодических характеристик и интонации.

В рамках **сегментных характеристик согласных** изучались:

- способы и особенности реализации консонантных сочетаний различных типов [Богомазов, 1974]; [Богомазов, Пауфошима, Шеворошкин, 1974];
- особенности акустической реализации мягких и твёрдых согласных в русском языке (на основании сравнительного анализа коартикуляции) [Кривнова, Кузнецова, 1974];
- акустическая реализация ассимилятивных явлений в сочетаниях согласных в русском и английском языках [Карневская и др., 1980 акустические признаки, различающие аспирированные смычные от неаспирированных [Хачатрян, 1980];
- конститутивные акустические признаки сонантов и шумных согласных [Касаткин, Пауфошима, 1982];
- собственная напряжённость согласных звуков (на материале украинского, немецкого и французского языков) [Кулина и др., 1984];
- индивидуальные признаки аффрикат [Николаенко, 1984].

В рамках **сегментных характеристик гласных** внимание исследователей было обращено на:

- спектральные характеристики гласных фонем азербайджанского языка в сопоставлении с русским [Тахи-заде, 1974];
- формантные характеристики русских гласных в потоке речи [Бондарко, Щербаква, 1978];
- фонетические характеристики стационарных гласных [Кузнецов, 1978];
- компоненты дифференциальных признаков фонем [Джапаридзе, 1978];
- динамика основных фонемных тембров гласных, содержащихся в спектральных описаниях слогов типа Г, типа ГСГ [Чабан, 1980];
- редукция гласных (на материале немецкого языка) [Стериополо, 1978; 1980];
- фонемный состав современного русского языка [Грибанов, 1982];
- вариативность формантных характеристик гласных в речи [Стериополо, 1982; 1989];
- варьирование спектральных характеристик ударных гласных в условиях просодического контура фразы (на материале английского языка) [Штакина, 1984];
- вокалическая структура слова и морфемы в русском языке [Богданова и др., 1984].

Интерес вызывали также проблемы **слога и слогоделения**, в числе которых рассматривались:

- асимметричность структуры слога [Таранец, 1980];
- генезис слога и слогоделение [Таранец, 1982];
- артикуляторные характеристики слога (на материале немецкого языка) [Кулина, 1982];
- модели физиологической структуры открытого ударного слога типа CV в современном немецком языке [Кулина, 1984];
- слогоделение с позиций морфологического анализа [Князева и др., 1986].

Интерес к явлениям **коартикуляции** повысился в начале 80-х годов XX века, когда исследованию подвергались следующие проблемы:

- коартикуляционные и позиционные эффекты в динамических спектрах речевого сигнала [Поспелов и др., 1984];
- формантные признаки коартикуляции для многоязычного синтеза [Карневская и др., 1984];
- комплексная модель коартикуляции [Башкина и др., 1986];
- взаимное влияние фонем на параметры друг друга (на материале эстонского языка) [Кюннап, 1989];
- заударные суффиксально-флексийные комплексы русского языка [Захаров и др., 1989].

Супрасегментные фонетические исследования охватывали широкий круг **просодических явлений**, начиная от просодии слога и слова, заканчивая фразовой просодией и интонацией.

На уровне звуков и слогов изучались:

- различительные признаки слоговой просодии (на материале немецкого языка) [Прокопова, 1974];
- взаимодействие просодических характеристик высказывания в пределах ядерного слога [Бровченко, Волошин, 1982];
- взаимодействие просодии и коартикуляции (на материале эстонского языка) [Лийв, 1982];
- длительность гласного как результат взаимодействия двух факторов: ритмической структуры слова и его положения в синтагме [Златоустова, Кузнецов, Отт, 1984];
- воздействие супрасегментных факторов на длительность ударных гласных в синтагме [Кривнова, 1984];
- временная организация слога и собственная длительность гласного [Кузнецов, 1989];
- временная организация слога: собственная длительность согласного и её позиционные модификации [Кузнецов, 1991];
- «собственный тон» русских гласных [Поспелов, 1991].

Традиционным объектом изучения являлись также **типы ритмических структур** (ритмика фонетического слова) и их модификации в слитной речи:

- временная организация ритмической структуры и её реализация в речи [Грязнова, 1980];
- ритмика фонетического слова в разных видах русской прозаической речи [Кривнова, 1982];
- ритмические структуры русской речи [Хитина, 1984];
- типология временных характеристик акцентно-ритмической структуры слова в синтагме [Белая, 1984];
- «модели» структурной организации слов и фонетических слов [Хитина, 1986];
- взаимообусловленные изменения сегментной длительности и ритмических структур в слитной речи [Потапова, 1986];

— ритмическая организации синтагмы в русской устной речи [Прокофьева, 1991].

Довольно популярными были также исследования параметров **словесного ударения**, в которых изучались:

- акустические корреляты фразовых ударений в высказываниях различных стилей (на материале русского языка) [Лавенкова, Панова, 1974];
- тональные характеристики и процесс распознавания ударения во фразе [Бровченко, 1974];
- физические корреляты словесного ударения и возможности их использования в системе автоматической сегментации слитной речи [Лийв, 1980];
- степени ударности слов в односинтагменной фразе [Надеина, 1982];
- связь акустических и перцептивных параметров ударения в немецком языке при помощи аппарата статистики [Евтихова, 1984];
- артикуляторная природа русского ударения [Скалозуб, 1986].

Неизменный интерес лингвистов был обращён к изучению **интонации и её компонентов**. Анализу подвергались следующие аспекты:

- зависимость между параметрами контура основного тона и лингвистическими единицами [Хаавель, 1974];
- тональные контуры различных типов высказываний [Бровченко и др., 1986];
- суммарная энергия и взаимодействие её компонентов в выражении просьбы и приказа в украинском языке [Войцеховский, Камбуленко, 1982];
- обобщённый интонационный контур фразы как форма изображения речевого сигнала [Бровченко, Волошин, 1984];
- обобщённый тональный контур разновидностей приказа в русском языке [Колымба, Войцеховский, 1984];
- просодические структуры с восходящее-нисходящей конфигурацией в терминальном тоне [Дубовский, 1984];
- акустическая характеристика видов вопроса в покутско-буковинском говоре [Бузинская, 1984];
- модификации временных характеристик фразы при изменении темпа [Метлюк и др., 1986];
- временная организация речи: предельные длительности ударных гласных в русском языке при нормальном (среднем) темпе [Кузнецов, 1986];
- роль компонентов интонации в выделении ключевых информационных сегментов звучащего текста [Саенко, 1986];
- препаузальное продление гласных в связном тексте [Кривнова, 1991].

Среди работ данного направления можно выделить такие, которые были непосредственно ориентированы на задачи распознавания и синтеза речи. В них были представлены следующие темы:

- алгоритмы поиска интоном русского языка [Блохина и др., 1971];
- модели просодических контуров [Белявский, Светозарова, 1978];
- модели мелодических контуров [Башкина и др., 1978];
- модель временной организации речи: регрессионный анализ динамики гласных [Кузнецов, Жаромский, 1982];
- статистический анализ частотного контура коммуникативных сигналов [Златоустова, Краснова, Козленко, 1984].

Отдельной, и весьма важной, лингвистической задачей в рамках изучения интонации и просодии являлся поиск параметров **макросегментации речи** на синтагмы и более крупные интонационно-смысловые единицы, а также **паузирование** как одно из средств макросегментации. Анализировались следующие проблемы:

- акустические характеристики различных видов пауз в слитной речи [Блохина, 1978];
- проблема членения связной речи в некоторых фонетических стилях (на материале русского языка) [Барышникова, Гайдучик, 1974];
- ритмообразующая функция частоты основного тона и паузации в организации звучащего текста [Златоустова, Долански, 1978];
- акустические характеристики синтагматического членения в армянском языке [Тохманян, 1980];
- роль интонации в выражении межсинтагменных и межфразовых связей в устном тексте [Блохина, 1980];
- связь сегментации звучащего текста с формальными синтаксическими единицами и их акустическими коррелятами [Абрамова и др., 1982];
- темповые характеристики и их применение для членения слитной речи [Выгонная, Якушев, 1982];
- закономерности временной организации фразы и проблема разграничения слов в потоке речи [Грязнова, 1982];
- сегментация звучащего текста с опорой на просодию [Потапова, Камышная, 1982];
- фонологические критерии автоматического членения слитной речи [Яковлев, 1984];
- роль длительности физических пауз в перцептивной сегментации звучащей речи [Блохина и др., 1984];
- роль модуляций частоты основного тона в макросегментации слитной речи [Блохина, 1986].

Исследования, посвящённые единицам **морфологического уровня**, были представлены не столь часто. Их характерная особенность – преимущественная направленность на решение конкретных задач автоматического распознавания и синтеза речи, поэтому соответствующие единицы языка анализировались, прежде всего, в фонетическом аспекте. В них обсуждались следующие проблемы:

- вопросы использования морфологической информации в целях распознавания непрерывного речевого сигнала [Асиновский и др., 1978];
- статистические данные о фонетической организации морфем в русском языке [Бондарко и др., 1982];
- синтаксическая индексация морфем как средство построения минимальной семантической модели при автоматическом распознавании [Асиновский и др., 1982];
- фонетические характеристики единиц морфемного словаря распознавания [Алексеева и др., 1984];
- фонетическая организация морфем (на статистическом материале суффиксов) [Саломатина, Юдина, 1989];
- фонематический состав частотных корневых морфем [Джолдасбекова, 1991].

Единицы **синтаксического** и **семантического уровней** исследовались в неразрывном единстве, а посвящённые им работы были ориентированы на решение задач распознавания речи. В этих работах рассматривались:

- проблемы описания смысловой структуры текста [Щекотова, 1978];
- семантико-синтаксическое строение частного вопроса [Носенко, 1980];
- семантико-синтаксическая модель и её применение [Белявский, Кузнецова, 1980];
- структура побудительных высказываний (команд), связанных с глаголами движения [Щекотова, 1980];

- возможные стратегии использования семантико-синтаксической информации при распознавании речевых сообщений [Данилов, 1980];
- формализация семантико-синтаксических ограничений в системах понимания речи [Чымбаев, 1982];
- вопросно-ответное единство в диалоговых системах [Носенко, 1982];
- принципы организации диалога в системах речевого управления [Щекотова, 1982];
- интонационно-смысловая структура ключевого слова в смысловой динамике текста [Дворжецкая, 1986].

II. Исследования процессов восприятия и понимания речи

В фокусе внимания лингвистов находились также механизмы восприятия различных языковых единиц, поскольку эти механизмы могли лечь в основу алгоритмов распознавания и синтеза речи.

Объектом изучения в перцептивных исследованиях являлись языковые единицы разных уровней, а также **общие закономерности восприятия их лингвистических характеристик**, в частности:

- лингвистические факторы, влияющие на восприятие связного текста [Штерн, 1982];
- роль сегментных и целостных признаков слова в процессах слухового и автоматического анализа [Данилов, 1982];
- реализация и восприятие наиболее частых словоформ в спонтанной речи [Александров, Гейльман, 1982];
- языковая специфика слуховой интерпретации речевых стимулов [Никифоров, 1984];
- специфика восприятия синтезированных слов [Штерн, 1989];
- дифференциальные признаки и восприятие речи (обзор) [Касевич, 1989];
- особенности восприятия гласноподобных стимулов с изменяющейся второй формантой разноязычными аудиториями [Кузнецов, 1991].

При изучении перцептивных коррелятов **словесного ударения** внимание уделялось следующим аспектам:

- роль изменения основного тона в восприятии словесного ударения в изолированных словах и фразах [Светозарова, Щербакова, 1971];
- влияние длительности согласного и гласного на восприятие ударения в русском языке [Венцов, 1974];
- зависимость восприятия русского ударения от консонантной структуры слова [Богомазов, Пауфошима, 1978];
- роль частоты основного тона и длительности гласных при восприятии места фразового ударения [Надеина и др., 1982];
- восприятие ударности гласного [А] в русском и белорусском языках [Поплавская, 1984];
- восприятие редуцированных гласных русского языка в ритмическом контексте [Захаров и др., 1991].

Для выявления закономерностей восприятия **интонации** изучались следующие вопросы:

- слуховое восприятие частотного контура фразы [Блохина, 1974];
- значимые воспринимаемые и акустические признаки при определении степени близости интонационной структуры фразы к естественному образцу [Барышникова и др., 1974];

- слуховая оценка временного контура ударных сегментов фразы [Потапова, 1974];
- восприятие темпа речи естественного речевого материала [Выгонная, 1980];
- восприятие ритмической организации фраз-команд [Хитина, 1984];
- восприятие ритмической структуры стиха [Красноперова, 1984];
- зависимость между просодической выделенностью и смысловой значимостью слов в тексте [Надеина, 1984];
- восприятие общего темпа синтагмы [Кривнова, 1986];
- длительность вокалического цикла и перцептивные пороги по темпу [Кривнова, 1989];
- роль просодии в механизмах восприятия речи [Белякова, Светозарова, 1991].

Довольно частым предметом исследования являлись закономерности восприятия речи в условиях **шумов и помех**:

- влияние контекста на восприятие слов и псевдослов в шумах [Синицына, 1982];
- связь механизма опознания и тяжести условий приёма речи [Штерн, 1984];
- восприятие ритмической организации фразы в шуме [Светозарова, 1982];
- восприятие в шуме слов разных степеней просодической выделенности [Надеина и др., 1984];
- вариативность восприятия речевых сегментов в затруднённых условиях [Прокофьева и др., 1986];
- трансформация фонетических признаков при восприятии речи в условиях помех [Кедрова, 1989];
- восприятие звуков речи в затруднённых условиях при использовании слуховой и визуальной информации [Захаров и др., 1989];
- условия адекватного восприятия речи на фоне однородной акустической информации [Надеина, 1991].

Закономерности **смыслового восприятия (понимания)** исследовались в следующих аспектах:

- значение пауз для понимания речи [Гайдучик, 1980];
- оценка сохранности смысла в искажённых текстах [Штерн, 1980];
- теория понимания речевого сообщения [Городецкий, 1982];
- механизм смысловой интерпретации воспринимаемых слов [Ушакова, 1982];
- роль ключевых и фонетически выделенных слов в распознавании связного текста [Кукольщикова и др., 1986];
- опорные элементы смыслового восприятия речевого сообщения [Павлова, 1986].

Отдельную группу образуют исследования, выполненные лингвистами (зачастую совместно с инженерами и математиками), которые имели ярко выраженную практическую направленность и в той или иной мере использовались в действующих системах автоматического распознавания и синтеза речи.

III. Автоматическое распознавание речи

В лингвистических исследованиях затрагивались **общие проблемы автоматического распознавания речи**:

- фонетические аспекты автоматического распознавания разговорной речи [Гейльман, Светозарова, 1980];
- иерархическая структура языковой способности, проявляющаяся в речевом и мыслительном поведении человека, и разработки устройства речевого ввода и вывода [Винарская, 1984];
- теоретические и прикладные аспекты распознавания спонтанной речи [Гейльман и др., 1989].

Многие работы были ориентированы на создание **лингвистических баз данных**:

- принципы построения фонетически представительной обучающей выборки [Гейльман и др., 1971];
- пути минимизации словаря распознавания [Штерн, 1974];
- статистические характеристики русской разговорной речи [Гейльман, 1978];
- способ формирования словарных эталонов [Кузнецова, 1982];
- словообразовательный словарь и его использование для автоматического распознавания речи [Бондарко и др., 1982];
- методики получения языкового материала для разработки диалоговых систем [Кузнецова, Галашевская, 1984];
- принципы организации словарного блока и блока лингвистической смежности в системах распознавания [Кузнецова, Галашевская, 1984];
- структура словаря для диалоговых систем с речевым вводом-выводом [Носенко, 1984];
- возможности различного представления единиц с одинаковыми структурно-лингвистическими свойствами в распознающих системах [Русакова, 1984];
- статистические характеристики слов, начинающихся с согласной [Саломатина, Юдина, 1986];
- фонетическая база данных как основа для эталонного описания русской речи [Бондарко и др., 1989];
- слоговая статистика русской устной речи [Кузнецов, Прокофьева, 1989];
- структура словарного блока распознающей системы [Кузнецова, 1991];
- фонетически представительный русский материал в целях совершенствования систем синтеза и распознавания речи [Богданова и др., 1991].

Важной задачей являлась также разработка **транскрипции и автоматических транскрипторов**:

- анатомо-физиологическая общefonетическая транскрипционная подсистема согласных [Симонян, 1980];
- теоретические и прикладные аспекты автоматической транскрипции [Бондарко и др., 1986];
- автоматический транскриптор в системах распознавания, синтеза и обучения [Богданова и др., 1989].

Наиболее многочисленную подгруппу в данной категории исследований составляют работы, в которых предлагались лингвистически ориентированные **алгоритмы и методики распознавания (экспертные системы)**:

- алгоритм синтаксического анализа устной речи [Асиновский и др., 1980];
- структура информационной системы, основанной на смысловой интерпретации устных и письменных запросов [Демьянков, Кузнецов, 1980];
- оптимальный коллектив линейных алгоритмов распознавания ударных слогов [Бровченко и др., 1982];
- методика чтения «слепых» сонограмм [Зиновьева и др., 1982];
- структура лингвистического процессора для диалога «Человек — ЭВМ» [Родичева, 1984];
- принципы моделирования диалога «Человек — ЭВМ» [Носенко, 1984];
- параметры эмоциональной речи на устройстве выделения речевых признаков [Нушикян, Трунин-Донской, 1984];
- определение опорных звуковых характеристик при анализе «слепых» сонограмм [Зиновьева и др., 1984];

- особенности фонетической интерпретации «слепых» сонограмм при работе с переменным окном [Зиновьева, 1984];
- использование аппарата канонических исчислений в системе автоматического распознавания речи [Асиновский и др., 1984];
- лингвистические процедуры отождествления и интерпретации в автоматическом распознавании речи [Асиновский, Воробьева, 1984];
- алгоритм фразового ударения в слитной речи [Королёва, Могилевский, 1984];
- оптимальное параметрическое представление речевых сигналов для построения экспертных систем распознавания речи [Зиновьева и др., 1986];
- параметры автоматического выделения темпоральных характеристик эмоциональной речи [Нушикян, Трунин-Донской, 1986];
- структура экспертной системы распознавания речи [Златоустова, Зиновьева, 1986]; модель лингвистического обеспечения ИПС с устным речевым вводом [Носенко, 1986];
- модель слогового распознавания речевых сигналов [Белявский, Кузнецова, 1989];
- алгоритм распознавания характера и степени эмоциональности высказывания [Нушикян, Варбанец, 1989];
- алгоритм выделения фонетических признаков из речевого сигнала [Белявский, 1991];
- алгоритм распознавания слов по последовательности результатов текущего распознавания [Белявский, 1991];
- алгоритм сегментация слитной речи на слова, основанный на учёте ритмических характеристик слова и вероятностно-статистических характеристик слога [Галашевская, 1991];
- фреймовый подход в акустико-фонетических модулях дикторнезависимых распознающих систем [Потапова, 1991];
- методика распознавания акустического сигнала с опорой на зрительный образ [Потапова и др., 1991];
- алгоритм идентификации словесного ударения в русском языке [Виноградов и др., 1991];
- экспертная система идентификации языка речевого сообщения [Данилов, Кузнецов, 1991];
- экспертная система интерпретации вокальных участков сонограмм [Зиновьева и др., 1991];
- метод установления принадлежности речевого сообщения одному из четырёх языков (русскому, английскому, немецкому и французскому) [Кузнецов, 1991];
- алгоритмы распознавания признаков иноязычного акцента по просодическим параметрам [Михайлов, Крейчи, 1991];
- методика определения места словесного ударения на основе лексико-морфологических характеристик (для синтеза по тексту) [Чикоидзе, 1991].

Лингвисты работали также и над задачей **оценки качества и надёжности распознавания речи:**

- анализ надёжности автоматического распознавания фонетических признаков [Кузнецова, Смирнова, 1986];
- принципы оптимального тестирования системы автоматического распознавания и синтеза речи [Ковальчук и др., 1989].

IV. Синтез речи

Лингвистические модели, нацеленные на задачи синтеза речи, затрагивали **различные фонетические аспекты:**

- проблемы синтеза мелодических характеристик речи [Карневская и др., 1978];

- временная структура речи и её реализация в синтезе по правилам [Кузнецов, 1980];
- формантные характеристики гласных для многоязычного синтеза [Карневская и др., 1982];
- модели синтеза мелодического контура русских и английских фраз [Карневская, Лобанов, 1982];
- правила формантного синтеза женского голоса по данным спектрального анализа [Гусейнов, Златоустова, 1982];
- моделирование просодических характеристик связного текста [Карневская, 1984];
- гармоническая ассимиляция гласных для многоязычного синтеза речи [Стериополо, 1984];
- интонационный контур синтезированной русской речи и словарь автоматизированных информационных систем [Шпильберг, 1984];
- моделирование просодических характеристик с помощью полиномов 3-й степени [Златоустова, Краснова, Крейчи, 1984];
- моделирование характеристик эмоциональной речи на синтезаторе «Фонемофон-4» [Нушикян, Лобанов, 1986];
- анализ и синтез просодики русской выразительной речи [Бровченко, Волошин, 1986];
- фонемные границы синтезированных гласных (на материале немецкого языка) [Стериополо, 1986];
- фонетический фонд русского языка как основа для компиляционного синтеза речи [Кузнецов, Байдакова, 1991];
- проблемы синтеза мужских и женских голосов [Долотин и др., 1991];
- методика определения места словесного ударения на основе лексико-морфологических характеристик (для синтеза по тексту) [Чикоидзе, 1991].

Много внимания уделялось также проблемам **качества синтезированной речи и методам его оценки**:

- оценка разборчивости автоматического синтеза речи [Зиновьева, Хитина, 1980];
- восприятие интонации синтезированных фраз [Урманцева, 1980];
- оценка качества синтезированной речи методом семантически противоположных пар [Кузнецов, Фролова, 1980];
- оценка разборчивости автоматического синтеза речи [Хитина, 1982];
- субъективная оценка естественности звучания речи в различных системах связи [Фролова, Кузнецов, 1984];
- особенности слитной речи для повышения качества автоматического синтеза [Игнатов, 1984];
- сопоставительный анализ интонации естественной и синтезированной речи [Бровченко и др., 1989];
- выбор мелодического контура при автоматическом синтезе речи [Карневская, 1989];
- набор тестов для обработки и проверки правил интонирования синтезированной речи [Светозарова, 1989];
- просодическое моделирование и повышение качества синтезированной речи [Карневская, 1991];
- естественность звучания вокодерной речи в зависимости от пола диктора [Кузнецов и др., 1986];

- помехоустойчивость синтезированной речи [Колесников, 1986];
- эффективность синтетической речи как канала информации в операторской деятельности [Кузнецов, Фролова, 1989];
- время реакции на речевой стимул как показатель качества синтезированной речи [Алявдина и др., 1991].

Важным результатом работы лингвистов над задачами автоматического распознавания и синтеза речи явилось развитие таких прикладных аспектов лингвистики, которые являются весьма востребованными в настоящее время.

Прежде всего, это касается **фоностилистических особенностей звучащей речи и в целом устного дискурса**. В рамках задач, связанных с распознаванием и синтезом слитной / естественной / спонтанной речи, лингвисты изучали следующие вопросы:

- проблемы распознавания просодических структур фоностилистических типов и видов устного текста [Дубовский, 1974];
- акустические корреляты фразовых ударений в высказываниях различных стилей (торжественный и непринуждённый стили в спонтанной речи) [Лавенкова, Панова, 1974];
- роль речевых пауз в фонетической дифференциации типов и стилей устной речи [Гайдучик, 1978];
- звуковой состав русской разговорной речи (сегментные характеристики) [Гейльман, 1978];
- роль высокочастотных формант в различении спонтанной и читаемой речи [Бровченко и др., 1980];
- длительность слога в спонтанной речи [Багмут, 1982];
- особенности просодики диалогической и монологической речи [Бровченко и др., 1982];
- взаимодействие акустических характеристик ядерного слога в спонтанной речи и чтении (на материале английского специального вопроса) [Еременко, 1982];
- ритмика фонетического слова в разных видах русской прозаической речи [Кривнова, 1982];
- метод последовательных процедур в распознавании спонтанной речи и чтения [Бровченко, Варбанец, 1984];
- распределение аллофонов гласных в спонтанной речи [Гейльман, Ошуйко, 1984];
- фонетические варианты согласных и их акустические корреляты (на материале речи горожан) [Сергеева, 1984];
- интонационные модели выразительной диалогической речи (анализ и синтез) [Бровченко, Волошин, 1986];
- количественные характеристики ударных гласных и звуковая стилистика (на материале английского языка) [Поплавская, Пузикова, 1986];
- форсированная речь как способ повышения разборчивости [Колесников и др., 1986];
- лингвистические корреляты социальных характеристик говорящих, уровня их речевой культуры, типа речи [Гейльман, 1991].

Фоностилистические разновидности звучащей речи неразрывно связаны с вопросом о **фонетических особенностях эмоциональной речи**, который также являлся предметом многих лингвистических исследований того времени:

- эмоциональные особенности речи как системы взаимосвязанных и взаимообусловленных элементов [Носенко и др., 1978];
- спектральные характеристики эмоционально окрашенной речи [Нушикян, 1980];
- индивидуально-варьируемые и устойчивые особенности речи в состоянии эмоциональной напряжённости [Карпов и др., 1982];

- проявление системной организации речи в условиях эмоционального стресса [Крылова, Боковиков, 1982];
- временная организация речи в зависимости от эмоционального состояния говорящего [Нушикян, 1982];
- частотные составляющие спектра ударных слогов эмоционально окрашенных фраз [Нушикян, 1984];
- частотная модуляция основного тона речи как информативный параметр эмоционального состояния человека [Дуранин и др., 1984];
- роль формантной структуры в передаче некоторых эмоциональных оттенков (на материале английского и украинского языков) [Колымба, 1986];
- эмоции и автоматическое распознавание и синтез речи: проблемы и перспективы (обзор) [Королёва и др., 1989].

В рамках изучения эмоциональной речи предлагались также **методы оценки эмоционального состояния говорящего**:

- применение теории потоков в задаче оценки эмоционального состояния человека по его речи и уровня владения иностранным языком [Карпов и др., 1980];
- исследование параметров эмоциональной речи на устройстве выделения речевых признаков [Нушикян, Трунин-Донской, 1984];
- анализ-синтез-анализ как метод выделения акустических характеристик выразительной и эмоциональной речи [Бровченко и др., 1989];
- исследование возможности тестирования эмоционального состояния говорящих по спектральным характеристикам их речевых сигналов [Ковальчук и др., 1991].

Несомненную теоретическую и практическую значимость имеют также лингвистические исследования, направленные на выявление **индивидуальных особенностей диктора (социальных, гендерных, возрастных и пр.)**:

- способ опознания языковой принадлежности говорящего [Потапова, 1978];
- акцентно-ритмическая структура слова в речевом потоке (в зависимости от социально-возрастного статуса диктора) [Белая, 1980];
- акустические параметры, различающие женский и мужской голоса [Златоустова, 1980];
- вариативность гласных в пределах слова в зависимости от диалектных особенностей диктора [Вербицкая и др., 1982];
- сегментные характеристики, полезные для определения языковой принадлежности говорящего [Никифоров, 1982];
- вариативность гласных и согласных в пределах слова в зависимости от диалектных особенностей дикторов [Вербицкая, Купчик, 1984];
- лингвистическая база систем распознавания индивидуальности диктора [Чикоидзе, 1984];
- роль ЧОТ и F-структуры в формировании мужского / женского качества голоса [Златоустова, Гусейнов, 1986];
- экспертная и объективная оценка социолингвистической принадлежности говорящих [Андрющенко и др., 1986];
- экспертная система идентификации языка речевого сообщения [Данилов, Кузнецов, 1991];
- метод установления принадлежности речевого сообщения одному из четырёх языков (русскому, английскому, немецкому и французскому) [Кузнецов, 1991];

- акустические характеристики иноязычных акцентов с помощью метода чтения «слепых» сонограмм [Щекотова и др., 1991];
- устойчивость акустических признаков, определяющих индивидуальность диктора, при передаче речи по трактам связи [Кедрова, Ложкина, 1991];
- использование акустических параметров эмоциональной речи для верификации диктора [Нушикян, Волошин, 1991];
- исследование тенденций изменений современной произносительной нормы [Богданова и др., 1991].

В рамках данного направления велись также исследования проявлений **языковой интерференции**:

- изменение русской речи под воздействием родного языка говорящих [Вербицкая и др., 1978];
- ритмическая структура слова в русской речи украинцев [Прокопова, Тоцкая, 1980];
- вариативность сегментных характеристик русской речи при межъязыковой интерференции [Вербицкая и др., 1980];
- модели акцентно-ритмической структуры слова в говорах украинского языка [Белая, 1982];
- интерференция ритмических структур фонетического слова в русской речи украинцев [Прокопова, 1982];
- модификация русских согласных под воздействием родного языка или диалекта говорящих [Вербицкая и др., 1986].

Примечательно также, что в рамках задач АРСО едва ли не впервые в нашей стране обсуждались проблемы создания автоматических программ **обучения произношению**, в том числе и на иностранном языке:

- опыт разработки программы автоматизированного устного курса по иностранному языку [Литвинов, 1971];
- оценка уровня владения иностранным языком [Карпов и др., 1980];
- использование машинных методов обработки сигнала в исследовании произносительной нормы русского литературного языка [Вербицкая, Алексеева, 1989].

Подводя итоги, можно сказать, что лингвистические работы школы-семинара «Автоматическое распознавание слуховых образов» создали основательную научную и практическую базу для интенсивного и плодотворного развития наиболее востребованных в настоящее время направлений прикладной лингвистики и речевых технологий.

Список литературы

1. Автоматическое распознавание слуховых образов: Тезисы докладов и сообщений VI Всесоюзного семинара (АРСО-6). 7–18 сентября 1971 г. Таллинн, 1971.
2. Автоматическое распознавание слуховых образов: Тезисы докладов VIII Всесоюзного семинара (АРСО-8). 16–23 сентября 1974 г. В 2-х ч. Львов, 1974.
3. Автоматическое распознавание слуховых образов: Материалы Всесоюзной школы-семинара (АРСО-10). 19–23 октября, 1978 г. Тбилиси: Мецниереба, 1978.
4. Автоматическое распознавание слуховых образов: Тезисы XI школы-семинара АРСО-11 (Ереван, 3–9 декабря 1980 г.). Ереван, 1980.
5. Автоматическое распознавание слуховых образов: Тезисы докладов и сообщений 12-го Всесоюзного семинара (АРСО-12). Киев, 1982.

6. Автоматическое распознавание слуховых образов: Тезисы докладов и сообщений 13-го Всесоюзной школы-семинара (АРСО-13). 23–28 июля 1984 г. В 2-х ч. Новосибирск, 1984.
7. Автоматическое распознавание слуховых образов: Тезисы докладов и сообщений 14-го Всесоюзного семинара (АРСО-14). 26–28 августа 1986 г. Каунас, 1986.
8. Автоматическое распознавание слуховых образов: Тезисы докладов 15-го Всесоюзного семинара (АРСО-15). 13–17 марта 1989 г. Таллинн, 1989.
9. Автоматическое распознавание слуховых образов: Тезисы докладов 16-го Всесоюзного семинара (АРСО-16). Суздаль, 7–12 января 1991 г. М., 1991.