

XIV Международная конференция SPECOM'2011 «Speech and computer» («Речь и компьютер»)

27–30 сентября 2011 г. в Казанском (Приволжском) федеральном университете состоялась XIV Международная конференция SPECOM'2011 «Speech and computer» («Речь и компьютер»).

Конференция «Speech and computer» проводится ежегодно, начиная с 1996 г., совместно Санкт-Петербургским институтом информатики и автоматизации РАН (СПИИРАН) и Московским государственным лингвистическим университетом (МГЛУ), которые выступают в качестве ее организаторов поочередно. Дважды данное мероприятие проводилось за пределами Российской Федерации – в Румынии и Греции. В конференции принимают участие специалисты – представители различных стран, школ,

направлений, работающие в области взаимодействия человека и компьютера. Объектом исследований является как устная, так и письменная речь. В центре интереса участников конференции находятся следующие проблемы: автоматическое распознавание устной речи; кодирование и передача речевого сигнала; автоматическое понимание устной речи; обработка многоязычной речевой информации; обработка мультимодальной и мультимедийной речевой информации; идентификация и верификация говорящего по голосу и речи; идентификация эмоционального состояния говорящего по голосу и речи; идентификация языка и акцента говорящего по речи; анализ и моделирование устно-речевого дискурса; лингвокриминалистические системы; системы обработки языко-

вой и речевой информации в промышленности; речепроизводство и речевосприятие; моделирование восприятия спонтанной устной речи; системы преобразования «текст-речь»; автоматические системы устно-речевого диалога; математическое и концептуальное моделирование естественных механизмов порождения и восприятия устной речи; лингвистические и паралингвистические коммуникативные стратегии; системы автоматического перевода; новые технологии в лингводидактике; речевые и языковые базы данных; перспективы эволюции речевых технологий; автоматическая обработка текста; экспертные системы; кодирование и декодирование устной и письменной информации; акустическая и лингвистическая стеганография и криптография; тестирование систем автоматического распознавания речи; измерение разборчивости речевой информации, передаваемой по различным каналам связи; речь в шумах и помехах; использование речевой информации в робототехнике.

Ежегодно в конференции участвуют представители многих стран: России, Австрии, Белоруссии, Бельгии, Канады, Чили, Китая, Чехии, Дании, Эстонии, Финляндии, Франции, Германии, Греции, Индии, Ирана, Ирландии, Италии, Японии, Кореи, Литвы, Малайзии, Мексики, Норвегии, Польши, Португалии, Румынии, Испании, Швеции, Тайваня, Таиланда, Нидерландов, Турции, Великобритании, Украины, Вьетнама, США и др. Активными участниками являются Международная ассоциация по речевой коммуникации, Международная ассоциация по судебной фонетике и акустике, Российское акустическое общество и др. Происходит оживленный обмен мнениями, излагаются результаты исследований и принимаются решения, способствующие дальнейшему развитию отраслей фундаментальной и прикладной науки, связанных с обработкой естественной языковой информации и высокими технологиями.

Организаторами SPECOM'2011 были Московский государственный лингвистический университет и Казанский (Приволжский) федеральный университет (председатель оргкомитета – Р.К. Потапова).

В общей сложности было представлено 76 докладов, каждый из которых прорецензирован членами международного научного комитета, включающего ведущих специалистов в области человеко-машинного взаимодействия на базе естественного языка.

Большой научный интерес вызвали пленарные доклады. Четыре из них предварили открытие рабочих секций конференции в первый день ее проведения: Р.К. Потапова (Москва) «Многоязыковые устно-речевые базы данных

в России»; М.В. Хитров (Санкт-Петербург) «Синергетический эффект в речевых технологиях»; В. Делич, М. Сечуйски, М. Гнатович и др. (Сербия) «О научно-исследовательской деятельности в области речевых и языковых технологий в Сербии»; Р.С. Якушев (Казань) «Особенности компьютерной “локализации” татарского языка».

В докладе М.В. Хитрова (Санкт-Петербург) «Синергетический эффект в речевых технологиях» освещалась идея анализа речи на основе синергии различных компонентов, которая достигается на двух разных уровнях: 1) среди алгоритмов обработки речи представляется возможным выделить общие компоненты; 2) информация, получаемая на выходе одного анализатора, может быть использована для уточнения результатов другого анализатора. В работе рассмотрены три основных уровня синергии применительно к извлечению самой разной информации из речи. В настоящее время они используются в мировой практике: синергия высокоуровневых компонентов обработки речи, информационная синергия и синергия общего информационного поля. Ценность доклада заключается в том, что автор указывает на важность комплексного, синергетического подхода при анализе такого мощного источника информации, как речь.

В сообщении Р.К. Потаповой (Москва) «Многоязыковые устно-речевые базы данных в России» обзор языков на территории современной мультилингвальной России с приведением полезных статистических данных позволил сделать ряд выводов: 1) для изучения вопроса о различных формах языковой миграции, а также характеристик устной речи для языков бывшего Советского Союза и на современной территории Российской Федерации представляется необходимым обращение к исследованию языков стран СНГ и целого ряда бывших советских республик; 2) особого внимания заслуживает проблема изучения процесса исчезновения языков в условиях многонациональной России; 3) языки, на которых говорят в автономных республиках и регионах Российской Федерации и странах СНГ, требуют количественных и качественных параметров описания. Отмечалась также целесообразность исследования особенностей процессов взаимовлияния между русским языком и языками бывшего Советского Союза в области судебной фонетики. Проблема мультилингвального исследования характеризуется наличием целого ряда научных направлений: анализ речи применительно к человеко-машинному взаимодействию, разработка устно-речевых диалоговых систем, анализ звучащей речи и выявление акустических характеристик речевого сигнала,

составление фонетической мультилингвальной базы данных и, в частности, фонетического корпуса вибрантов в различных языках. Устно-речевые базы данных, прежде всего, должны составить объект разработки в таких областях речеведения, как фонетика и фонология, акустика, восприятие речи и т.д.

Доклад «О научно-исследовательской деятельности в области речевых и языковых технологий в Сербии» [В. Делич, М. Сечуйски, М. Гнатович и др. (Сербия)] представлял собой обзор исследовательских программ в области речевых технологий, которые в настоящее время проводятся в Сербии. Научными проектами в рамках таких программ занимаются междисциплинарные научные организации, имеющие следующую установку: внедрить передовые междисциплинарные подходы к изучению разнообразных вопросов в области речевых технологий и разработать новейшие компьютерные продукты. В докладе освещались четыре проекта, объектами исследования в которых являются человекомашинное взаимодействие, устно-речевые диалоговые системы, автоматическое распознавание речи, синтез речи по тексту, голосовые порталы, пренатальное и доречевое развитие, восприятие речи, речеобразование, оценка вербальной коммуникации, паралингвистика, когнитивно-психосоматическая коммуникация, оценка качества речи, интернет-технологии в медицине, телемедицина, проектирование роботов, мультимодальное человекомашинное взаимодействие и т.д. Еще одной немаловажной целью таких проектов является попытка сербского научного сообщества в сфере речевых технологий заявить о себе как о надежном партнере, который может быть принят в международные научные консорциумы.

В докладе Р.С. Якушева (Казань) «Особенности компьютерной “локализации” татарского языка» отмечалось, что компьютерная «локализация» татарского языка сыграла роль катализатора в развитии этого языка как государственного во всех сферах общественной жизни и структурах управления. Помимо этого в докладе были изложены некоторые проблемы «локализации» татарского языка применительно к компьютерной технике, как то: графической реализации татарских букв для использования высококачественных шрифтов, проверке орфографии и правописания, а также единой кодировке символов татарского алфавита. Рассматривая компьютерную «локализацию» естественного языка как один из видов моделирования культурологического феномена, автор доклада поднял не менее значимый вопрос философско-методического характера – применение методов синергетики в компьютерной лингвистике.

В последний день конференции SPECOM'2011 были представлены еще два пленарных доклада: «Тенденции в автоматическом анализе текстов» (А.А. Харламов, Москва) и «Речевой портрет коммуникантов в условиях “конфликта-консенсуса”» (Р.К. Потапова, Л.Р. Комалова, Москва).

В первом из них были представлены различные взгляды на автоматический анализ текстов и, в частности, на формирование смыслового «портрета» текста как одного из наиболее важных задач анализа. Данный аспект – это объединение более традиционных статистического и лингвистического подходов и когнитивного подхода, реализующихся на основе представлений об обработке информации, как речевой, так и текстовой, в головном мозге человека. Смысловой «портрет» текста при когнитивном подходе строится как ассоциативная частотная сеть, числовые характеристики которой перенормировываются в смысловые веса. В докладе была отмечена эффективность использования когнитологии в решении стандартных задач автоматического анализа текстов.

В заключительном пленарном докладе Р.К. Потаповой, Л.Р. Комаловой (Москва) «Речевой портрет коммуникантов в условиях “конфликта – консенсуса”» представлены результаты эксперимента по формированию речевого «портрета» говорящего в ситуациях конфликта и консенсуса. Исследование производительной экспликации коммуникативной модели «конфликт – консенсус» проводилось в 2009–2011 гг. в Центре фундаментального и прикладного речеведения МГЛУ. Актуальность такого исследования, прежде всего, продиктована потребностью диагностики конфликтности и выявления параметров перехода от конфликта к консенсусу и *vice versa*. Сам эксперимент включал в себя две составляющие: перцептивно-слуховой эксперимент, проводимый на материале реальных конфликтов на стадии предконфликта, и зрительный эксперимент в сочетании с перцептивно-слуховым, проводимый на материале телевизионных ток-шоу с консенсуальной коммуникацией и сценическими конфликтами. В ходе эксперимента получены данные, согласно которым определены наиболее существенные параметры для оценки социального «портрета» говорящего по голосу и речи, а также речевые маркеры перехода коммуникантов из состояния конфликта в состояние консенсуса и *vice versa*.

Пленарное заседание конференции SPECOM'2011 позволило сделать целый ряд серьезных выводов относительно настоящего положения дел в области речеведения и современных речевых технологий, а также опреде-

лить наметившиеся тенденции развития этой междисциплинарной научной области:

1. В настоящее время анализ такого многопланового феномена как звучащая речь уже не может представлять собой набор разрозненных, обособленных научных работ, результаты которых не могли бы дополнять друг друга. Автоматический анализ речи целесообразно проводить как комплексное, разностороннее исследование, где необходима синергия, совмещение самых разных научных направлений и достижений.

2. Речевые технологии имеют исключительно положительную динамику развития, которая характеризуется различными направлениями, начиная от человекомашинного взаимодействия и заканчивая технологиями других наук, как, например, медицины.

3. Сложность и неоднозначность проблемы речевых технологий в России обуславливается ее мультилингвальностью и наличием одной из насущных проблем, решение которой должно стимулировать дальнейшее развитие речевых технологий (в частности, разработку мультилингвальных устно-речевых баз данных).

Пять рабочих секций конференции были распределены по трем дням ее проведения (27–29 сентября). Первая секция **«Фундаментальные и прикладные проблемы современного речеведения»** включала ряд докладов новаторского характера. Метод стеганографии, основанный на квантовании речевого сигнала по времени с опорой на частоту основного тона и квантовании параметрических данных, скрытых в сегментах, был представлен в докладе Р.К. Потаповой, М.О. Салагай (Москва) «Разработка метода стеганографии с целью сокрытия данных на базе речевых параметров». Было отмечено, что, принимая во внимание естественную сегментацию речи и особенности восприятия речи человеком, авторам доклада удалось достичь высокого уровня защиты конфиденциальных данных от несанкционированного доступа.

Сравнительный анализ просодических элементов (тон, ритм, ударение и интонация) как средств речевого воздействия в таких языках, как английский, немецкий и нидерландский, составил содержание доклада Р.К. Потаповой, А.В. Москвиной (Москва) «Просодические признаки как средство речевого воздействия». Предпосылкой такого анализа явилась постановка следующего вопроса: схожи ли просодические единицы в упомянутых языках или же они обладают дифференциальными признаками. Результаты первого из двух частей исследования (акустический анализ ча-

стоты основного тона, уровня интенсивности и временных характеристик) показали, что, хотя английский, немецкий и нидерландский языки принадлежат к одной и той же языковой семье, в этих языках наблюдаются разные просодические особенности, выступающие как средства речевого воздействия. Результаты второй части эксперимента (во время которого испытуемые прослушивали языковой материал и описывали следующие особенности: высота голоса, темп речи, длительность пауз в речи и способы их заполнения, сила голоса, ритм и т. д.) показали, что языковая компетенция имеет отношение к эффекту речевого воздействия. Новизна проведенного анализа определяется отсутствием в настоящее время работ, освещающих сравнение сразу нескольких языков на предмет различия просодических характеристик как средств речевого воздействия.

Среди других значительных результатов в текстах докладов рассматриваемой рабочей секции можно отметить следующие:

– использование системы перефразирования для сокрытия информации применительно к стеганографии с учетом стилистических особенностей различных текстов может обеспечить надежную стеганографическую систему, способную предоставить обмен сообщениями со скрытой информацией, а также способствовать разработке алгоритмов для анализа текстов на наличие скрытого сообщения, в том числе и в автоматизированном режиме [А.П. Ким (Москва) «Лингвистические детерминанты средств защиты и извлечения смысловой информации»];

– большой интерес представило определение статистических данных (таких, как частота встречаемости) русских монофонов и дифонов, функционирующих в стилистически гетерогенных корпусах текстов, содержащих 460 965 словоформ, свыше 1 млн слогов и 2 500 000 фонов. Полученные результаты могут быть использованы для типологических межъязыковых сравнений и для совершенствования систем автоматического распознавания русской речи [Н.С. Смирнова, П.Г. Чистиков (Санкт-Петербург) «Статистика монофонов и дифонов в русском языке»];

– с помощью инновационного метода изучения механизмов речевой артикуляции, в частности метода на базе томографа TO MİKON S 50 (BRUKER) (MPT), были в деталях изучены артикуляционные профили русских губных согласных [б] и [м] в определенных вокалических контекстах с последующим обсуждением результатов исследования [Г.Е. Кедрова, Л.М. Захаров, Н.В. Анисимов (Москва) «О веляризованности русских губных согласных»];

– проведен перцептивно-слуховой эксперимент как часть разработанного метода анализа индивидуальных и личностных особенностей речи на основе целого набора просодических характеристик. Эксперимент включал использование специально созданного опросного листа, способствующего определению характерных особенностей говорящего, что позволило выявить специфику слухового восприятия спонтанных вербальных сообщений при отсутствии невербальных визуальных средств [Р.К. Потапова, В.В. Потапов, М.В. Хитина (Москва) «Особенности варьирования фонетических признаков спонтанного речевого материала (на основе перцептивно-слухового анализа)»].

– в докладе Р.К. Потаповой, В.В. Потапова, М.В. Хитиной (Москва) «Определение темы текста, воспринятого в затрудненных условиях (экспериментальное исследование)» были представлены результаты проведения слухового анализа речи в условиях зашумленности. Задача заключалась в определении темы устного сообщения при разных видах речевой деятельности (чтение текста, монолог, диалог и полилог без предварительной подготовки). Анализ восприятия материала в затрудненных условиях показал, что на восприятие влияет сложность в формулировании как глобальной, так и локальных тем текста; в затрудненных условиях лучше воспринимаются диалоги (полилоги); анализ глобальной и локальных тем текста свидетельствуют о том, что информанты используют для формулировок разные стратегии; восприятие текста в затрудненных условиях и определение тем текстов – задачи разного плана и т. д.;

– в результате экспериментального исследования слухового восприятия спонтанных диалогов в условиях белого шума были определены факторы, способствующие правильной формулировке глобальной и локальных тем диалогов на основе количества и качества воспринятых речевых единиц [Ю.В. Абрамов, Р.К. Потапова, М.В. Хитина (Москва) «Смысловая интерпретация спонтанного звучащего диалогического текста (в условиях зашумления речевого сигнала)»];

– в докладе С.П. Курбатова, А.А. Литвиновича, Г.С. Хахалина (Белоруссия) «Синтез выражений русского жестового языка по естественно языковому контексту» была рассмотрена проблема синтеза русского жестового языка и естественного языкового контекста. При этом приведены характеристики онтологического описания, служащего связующим звеном между лексическим анализатором, жестовым синтезатором и экспериментальной частью.

Секция «Автоматизированные системы распознавания, понимания и синтеза устной речи» включала доклады, отражающие новейшие разработки и достижения в области автоматического распознавания речи. Так, в докладе Д. Пекара, М. Янева, Н. Яковлевича и др. (Сербия) «Системы распознавания слитной речи с применением метода марковских цепей на базе дополнительного моделирования акустической вариативности» был представлен инновационный метод, улучшающий точность работы систем по распознаванию непрерывного речевого сигнала на основе скрытых марковских моделей. Метод основан на дополнительном моделировании акустической вариативности в базовой системе распознавания непрерывного речевого сигнала в форме гауссовских смешанных компонентов. Сущность метода заключается в разработке новых скрытых марковских моделей путем приписывания акустического состояния каждому гауссовскому компоненту. Экспериментальные результаты показали, что предлагаемая система имеет значительные улучшения по сравнению с базовой системой.

Доклад Ф. Стоутена, И. Иллиной, Д. Фора (Франция) «Метод кластеризации повторяющихся лексем в системе распознавания речи» освещал новейшие технические средства для обнаружения повторяющихся лексем, извлекаемых из заданного словаря в системе распознавания связной речи, основанной на словаре больших размеров и используемой для транскрибирования данных из новостных сообщений, передаваемых в устно-речевой форме. В докладе сообщалось, что предлагаемая система состоит из двух блоков: детектора слов, извлекаемых из словаря, и модуля кластеризации, который обрабатывает обнаруженные сегменты слов. Сам алгоритм кластеризации основан на оценке энтропии. В результатах было отмечено, что сочетание детектора лексем из словаря и кластеризации сегментов слов из этого словаря может предоставить большую точность в сравнении с использованием только одного детектора слов из заданного словаря.

Г. Куку, Л. Бесаки, К. Буриляну и др. (Румыния) для решения проблемы недостатка ресурсов базовых и узкоспециализированных текстов для систем распознавания румынской речи в докладе «Расширение возможностей автоматического распознавания румынской речи за счет использования корпусов текстов, полученных в результате машинного перевода и из Интернета» предложили использовать тексты, полученные на базе машинного перевода, а также тексты, доступные в интернет-среде. Авторы отметили, что благодаря данной методике в сравнительно небольшие сроки удалось

получить систему распознавания румынской речи, не зависимую от говорящего.

В докладе П.Ю. Юркова, М.Л. Корневского, К.Е. Левина (Санкт-Петербург) «Повышение помехоустойчивости автоматической системы распознавания речи с учетом варьирования уровня громкости» были приведены результаты разработки по улучшению помехоустойчивости в процессе автоматического распознавания речи при изменении уровня громкости с приведением расширенных статистических данных.

Разработка адаптивной модели языка множества предметных областей с использованием эксплицитного семантического анализа была представлена в докладе «Мультипространственная языковая адаптационная модель на базе эксплицитного семантического анализа» [К. Кильгур, Ф. Крафт, С. Штюкер, А. Вайбель (Германия)].

Создана и описана модель, прогнозирующая вероятность появления границ между предложениями исключительно на основе просодических характеристик: частоты основного тона, уровня интенсивности и длительности пауз. При этом предложенный алгоритм работает в режиме реального времени и предсказывает вероятность появления границ предложения для каждого перерыва по частоте основного тона. Полученные данные используются в качестве дополнительного источника информации при автоматическом распознавании речи (автоматическое обнаружение границ предложений в режиме реального времени применительно к системе автоматического распознавания русской речи) [П.Г. Чистиков, О. Хомицевич (Санкт-Петербург) «Обнаружение в реальном времени границ предложений на базе автоматической системы распознавания русской речи»].

В докладе Т.И. Ибрагимова, Ф.И. Салимова, М.Р. Сайхунова (Казань) «Синтезатор татарской речи по тексту» была представлена экспериментальная версия синтезатора татарской речи. При этом опыты по аудированию показали удовлетворительную разборчивость синтезированной речи, хотя по звучанию татарская речь на выходе синтезатора несколько отличается от натуральной.

Вопросам обработки данных в речевых технологиях была посвящена секция «Автоматическая обработка мультилингвальной, мультимодальной и мультимедийной информации». В докладе А.А. Ронжина, В.Ю. Будкова (Санкт-Петербург) «Определение и фонозапись активного говорящего в условиях помещения для совещаний» была поднята проблема автоматического подбора оптималь-

ных средств видеонаблюдения за активным выступающим в помещении для совещаний, где присутствуют более тридцати участников дискуссии. Авторы доклада провели эксперименты, в которых для записи сообщений говорящего осуществили локализацию аудиоисточника и средств видеонаблюдения. При этом для определения текущего положения говорящего использовалась локализация мультиканального аудиоисточника. Видеообработка сигналов пяти камер послужила средством регистрации участников обсуждения. При этом отслеживался главный выступающий и проходила видеозапись аудитории. Проведенные эксперименты показали, что разработанная аудиовизуальная система зафиксировала сообщения всех говорящих.

В докладе М.Л. Корневского, И.А. Пономаревой, К.Е. Левина (Санкт-Петербург) «Сегментация новостных текстов на русском языке в режиме реального времени» был предложен онлайн-метод сегментации с выделением тем в слитной речи применительно к вещанию новостей. Такой подход основан на данных о границах предложений, получаемых с помощью системы их автоматического обнаружения. Рассматриваемый в работе алгоритм сопоставляется с другим методом, при котором для деления спонтанной речи на фрагменты используется только пороговая величина количества слов. Авторы доклада отметили, что принятие в расчет данных о границах предложений для деления слитной речи на фрагменты в целях классификации выделяемых тем обеспечивает повышение их точности обнаружения и систематизации на 25–30%.

Другие доклады, представленные в данной секции, отмечены не менее значимыми и полезными научными результатами. Реализована гибридная дискриминантно-порождающая EV-HMM(GMMs) система диаризации. При этом показана эффективность итеративного взаимодействия между порождающим и дискриминантным блоками, позволяющая превзойти по качеству диаризации EV-VBA системы в случае коротких диалогов [доклад Т.С. Пеховского, А.К. Шупицы (Санкт-Петербург) «Гибрид генеративных и дискриминативных моделей для задачи диаризации в коротком телефонном диалоге»].

В докладе А.А. Жарких (Мурманск) «Канал передачи речи от одного человека к другому как нелинейный канал передачи гибридных аналого-цифровых сообщений» были представлены математические модели, которые могут быть использованы для описания некоторых механизмов речеобразования и слуха, и

изложены соответствующие выводы, среди которых: 1) канал передачи речи от одного человека к другому есть нелинейный канал передачи гибридных аналого-цифровых сообщений; 2) в речевом тракте человека входной сигнал из головного мозга преобразуется в выходной сигнал речи с расширением спектра и увеличением амплитуды; 3) слуховая система речи имеет тонкую структуру, обусловленную многоканальностью.

Доклад В.Г. Санникова, С.А. Маслова (Москва) «Оптимизация алгоритма калмановской фильтрации зашумленной речи» был посвящен представлению зашумленного речевого сигнала в виде отклика голосового тракта, задаваемого уравнениями состояния и наблюдения, синтезирован оптимальный линейный фильтр. При этом проведена оценка параметров фильтра на основе метода наименьших квадратов с конечной памятью.

Представлен эффективный вычислительный метод фонемной сегментации, не зависящей от типа текстов и дающий надежные результаты. Такой метод получен в результате эксперимента с использованием функции спектральной вариативности, а также простого непараметрического метода фонемной сегментации [доклад Дж. Вайнбергера, Х. Дауди, Х. Яхиа (Франция). «Улучшение представления речевой волны с учетом программы ДИС-ТБИК, предназначенной для фонемной сегментации»].

Рабочая секция «Автоматизированные системы идентификации/верификации, эмоционального состояния и языковой принадлежности говорящего» включала доклады, ориентированные на проблемы автоматизации определения различных характеристик говорящего. Цель, поставленная авторами сообщения «Применение упрощенного одностороннего анализа варьирования для нахождения более информативных признаков с целью определения языкового интеллекта на базе речи испытуемых» [К. Заблоская, С. Заблоский, С. Вальтер, В. Минкер (Германия)], – автоматическое определение вербального интеллекта говорящих на основе порождаемых ими высказываний. Эксперимент заключался в просьбе людей различного возраста и с различным образованием дать описание одного и того же события. На основе записанной речи уровень вербального интеллекта был измерен с использованием специального теста. Для транскрибированных монологов были определены лингвистические особенности, стиль и ключевые слова. Для обнаружения говорящих с более высоким или более низким вербальным интеллектом использованы программы опорных векторов с учетом отмеченных лингвистических

особенностей. С целью определения более информативных признаков применен простой односторонний дисперсионный анализ, давший содержательные результаты.

В докладе И.А. Главатских, А.М. Широковой, А.П. Зыкова (Москва) «Применение N-граммной модели сглаживания и фонемной кластеризации при решении задачи автоматической идентификации языка» использован фонотактический метод, в основе которого лежит расчет вероятности принадлежности к языку по его N-граммной модели, обученной на множестве автоматически распознанных последовательностях фонем. С целью повышения точности системы идентификации языков разработан метод сглаживания, а также сокращенный инвентарь фонем. Проведенные авторами доклада эксперименты показали, что для решения задачи идентификации языка (на примере русского языка) предложенные подходы приводят к увеличению точности. Применение сокращенного инвентаря позволило увеличить точность системы на 9,8 %, а совместное использование разработанного метода сглаживания N-граммной модели привело к увеличению точности еще на 1,3 %.

Среди других полученных экспериментальных результатов следует упомянуть следующие:

- применение метода доверительного интервала для параметрических кривых оценки результатов автоматического распознавания говорящего [доклад С.Л. Коваль, А.А. Лохановой (Санкт-Петербург) «Использование метода определения величин доверительного интервала как инструмент оценки результатов автоматического распознавания говорящего»];

- представлен новый метод не зависящий от текстов автоматической идентификации говорящего. Он базируется на дифференциальном классификаторе и дает лучшие результаты, нежели другой инновационный метод, предполагающий моделирование нормальных распределений [доклад Н.А. Любимова, Д.А. Леднова, М.Ю. Андреева (Москва) «Система автоматической идентификации говорящего, разработанная на основе дискриминативного классификатора»];

- проанализировано влияние наркотического опьянения на частоту основного тона говорящего, и получены данные, согласно которым основной причиной изменений частоты основного тона является эмоциональное состояние говорящего. В работе также указано, что у людей, продолжительно принимающих наркотики героиновой группы, проявляется тенденция к снижению частоты основного тона ежегодно на 3% [доклад А.А. Раева, Ю.Н. Матвеева, Т.И. Голощаповой (Санкт-Петербург / Мо-

сква) «Эффект влияния длительного употребления наркотических средств на частоту основного тона говорящих»];

– в докладе А.А. Журавлевой (Москва) «Экспертный метод идентификации индивидуальных признаков личности по его речи» был представлен метод психологического шкалирования, позволяющий прогнозировать персональные черты говорящего по его речи. Такой метод подразумевает построение модели личности в результате процедуры обучения, отбора и поддержки экспертов;

– для построения UBM-модели GMM-систем распознавания диктора предложен алгоритм обучения, основанный на вариационном байесовском анализе вместо традиционно используемого алгоритма с учетом метода максимального правдоподобия [доклад Т.С. Пеховского, А.И. Лохановой (Санкт-Петербург) «Выбор UBM модели с помощью вариационного байесовского анализа для GMM-UBM систем распознавания диктора»];

– определены изменения, которым подвергаются различные параметры голоса и речи человека под влиянием смены языкового кода, а также установлены характеристики, которые остаются неизменными в речи носителей русского языка при переходе на немецкий язык [доклад Т.С. Платоновой, А.А. Смолиной (Москва) «Идентификация говорящего с учетом изменения языкового кода (на материале русского и немецкого языков)»].

Заключительной рабочей секцией конференции явилась секция **«Корпусная лингвистика применительно к устной и письменной речи»**, где были продемонстрированы результаты разработки корпусов текстов, на основе которых представляется возможным проводить различные исследования. В докладе А.В. Архипова, О.Ф. Кривновой, А.А. Лебедева (Москва) «Речевой корпус «Интонация русского диалога»» была представлена разработка речевого корпуса «Интонация русского диалога», который, как отмечается в работе, изначально ориентирован на звучащую речь, а не на письменный текст. Такой корпус иллюстрирует роль просодии в современном русском языке. Его основная цель – детальная просодическая транскрипция представительного массива диалогических высказываний в сочетании с их семантико-прагматическими и грамматическими аннотациями. Корпус «Интонация русского диалога» в первую очередь предназначен для исследователей русского языка, а также для обучающихся ему. От корпусов, создаваемых для целей автоматической обработки речи, он отличается специализированной классификацией и специальным под-

бором диалогических реплик, репертуаром локальных и интегральных просодических характеристик, что представляет особый интерес для лингвистов, занимающихся фонетикой и речеведением.

В докладе В. Пилипенко, В. Робейко, М. Сашок и др. (Украина) «Формирование речевого корпуса на базе украинских новостных сообщений» было представлено описание этапов разработки корпуса устно-речевых новостных сообщений на украинском языке. При этом дается обзор настоящего положения дел в создании украинских речевых корпусов. Украинский корпус устно-речевых новостных сообщений содержит спонтанную украинскую речь, морфологически аннотированную и транскрибированную. Помимо этого, в докладе освещаются основные идеи и технические решения, использованные в записи и компьютерном аннотировании для уже существующих речевых корпусов.

В других докладах, представленных на секции, затрагивались различные научные проблемы и методика их решения с применением данных корпусов. Создан алгоритм, позволяющий реализовать процесс автоматической предварительной обработки текстовых данных (взятых из русских текстовых корпусов) для языкового моделирования [доклад С. Заблоцкого, К. Заблоцкой, В. Минкера (Германия) «Автоматическая предварительная обработка корпусов русскоязычных текстов с целью последующего моделирования языка»].

Разработана программа, преобразующая разметку звукового сигнала из текстового формата в формат данных xml и позволяющая представить отраженные в разметке сигнала уровни в виде иерархии, каждый уровень которой соответствует определенному уровню языка; такой подход делает возможным поиск средствами языка запросов LINQ тех сегментных и супraseгментных единиц и их статическое описание, которые необходимы для дальнейшего исследования [доклад С.О. Тананайко, Д.А. Кочарова, К.Р. Садуртиновой (Санкт-Петербург). «Программа статистической обработки корпуса речевых данных»].

Организаторы конференции SPECOM'2011 констатировали, что уровень представленных вниманию участников конференции докладов по-прежнему высокий, тематика исследований разнообразна и достаточно полно разработана. Несомненно, SPECOM'2011 позволит продолжить традицию регулярного обмена мнениями и результатами исследований, что положительно повлияет на дальнейшую эволюцию компьютеризации естественных языков в России и

на развитие продуктивных научных контактов ученых разных стран мира.

Обоснованность организации конференции в Казани на базе Казанского университета не вызывает сомнения. Лингвистическая школа Казанского университета, восходящая к Бодуэну де Куртенэ, всемирно признана. С конца 70-х гг. XX в. произошло объединение усилий казанских лингвистов и специалистов в области компьютерных технологий в разработке проблем компьютерной лингвистики и организации естественного языкового интерфейса в диалоге «человек – ЭВМ». После принятия Закона о языках народов Республики Татарстан работа в данном направлении интенсифицировалась. В условиях многонациональной России интерес к опыту компьютеризации исследований татарского языка вполне обоснован.

В частности, в Академии наук Республики Татарстан существует целостная концепция компьютерной поддержки основных языков населения республики. Такой подход обеспечивает повышение активности функциональной роли языка и его значения как хранителя этнических и культурных традиций. Создание машинного фонда того или иного языка открывает широкую перспективу для разнообразных исследований, имеющих практическую направленность, что актуально для решения

определенных проблем, обусловленных миграционными процессами, взаимопроникновением и интерференцией языков.

Для изучения языка и речи Международная конференция «Речь и компьютер» (SPECOM) является прекрасной возможностью для координации и уточнения направлений исследований в международном масштабе.

Анализ обсужденных докладов показал, что наиболее активная работа ведется в направлении корпусной лингвистики, автоматизированных систем анализа – синтеза речи, распознавания и понимания речи, эмоционального состояния говорящего, математической обработки данных. Заслушанные доклады в большинстве случаев носили практический характер. Представлены были также и глубокие концептуальные теоретические обобщения.

Р.К. Потапова, В.В. Потапов

Сведения об авторах:

Родмонга Кондратьевна Потапова
Московский государственный лингвистический университет

Всеволод Викторович Потапов
МГУ им. М.В. Ломоносова
rkpotapova@yandex.ru