

**XIX МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ SPECOM'2017  
“РЕЧЬ И КОМПЬЮТЕР”**

**XIX INTERNATIONAL CONFERENCE SPECOM'2017  
“SPEECH AND COMPUTER”**

DOI: 10.31857/S241377150001114-8

С 12 по 16 сентября 2017 г. в Хатфилде (Великобритания) в университете Хертфордшир состоялась очередная XIX Международная конференция “Речь и компьютер” (SPECOM'2017), совмещенная со II-й Международной конференцией по интерактивной коллаборативной робототехнике (ICR 2017). Рабочий язык – английский. В качестве организаторов конференции традиционно выступили Московский государственный лингвистический университет (МГЛУ) и Санкт-Петербургский институт информатики и автоматки РАН (СПИИРАН). Кроме того, к этому мероприятию в 2017 году присоединился Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, а также вышеупомянутый Хертфордширский университет Великобритании, в кампусе которого проходила конференция. Следует подчеркнуть, что как МГЛУ, так и СПИИРАН являются организаторами международных конференций SPECOM с 1996 года. В соответствии с этим конференцию возглавили Родмонга Потапова (МГЛУ), Алексей Карпов (СПИИРАН) и Иосиф Борас (от университета Хертфордшира). На конференции было заслушано, а также представлено в постерном режиме 122 доклада, прошедшие рецензирование членами программного комитета, состоявшего из 120-и профессионалов по проблематике конференции из различных стран мира. Повышенные требования к качеству докладов являются условием публикации материалов в серии LNAI 10458 – Lecture Notes in Artificial Intelligence в наукометрическом формате SCOPUS, Web of Science (международное издательство Springer). Все доклады данной Международной конференции были изданы в соответствующих трудах<sup>1</sup>. В качестве докла-

дов на конференции SPECOM'2017 приняли участие представители Алжира, Аргентины, Беларуси, Бельгии, Великобритании, Венгрии, Германии, Греции, Дании, Индии, Ирана, Ирландии, Испании, Китая, Нидерландов, Палестины, России, Сербии, США, Финляндии, Франции, Чешской Республики, Швеции, Японии. Наиболее представительно выглядело участие учёных России (20 докладов), Чешской Республики (11 докладов), Германии (9 докладов), Великобритании (8 докладов). От каждой из оставшихся вышеперечисленных стран было представлено, как правило, по 2–3 доклада. Лингвистическая тематика, нашедшая отражение на конференции, включала все те области лингвистических знаний, которые связаны с коммуникацией “человек – машина” на базе информации о конкретных признаках и параметрах письменной и устной речи. Так, например, известно, что для развития современных информационных технологий в цифровой среде необходимы репрезентативные массивы лингвистических данных, что обусловило стремительное развитие корпусной лингвистики с охватом устной и письменной речи, подтверждение чему мы находим в целом ряде докладов, например, в докладе Н. Богдановой-Бегларян, Т. Шерстиновой и др., (Санкт-Петербург) “Лингвистические признаки и социолингвистическая вариативность повседневной русской речи” представлены результаты исследования, проведенного с использованием аннотированного субкорпуса повседневной русской речи, содержащего 125 000 слов, входящих в состав аннотированного корпуса повседневной русской речи, сформированного по данным 256 информантов с учетом гендера, возраста, профессии и социального статуса. В ходе исследования авторы опираются на три категории: а) диагностические признаки, которые отражают статистически надежные различия между социальными группами; б) лингвистические признаки, которые могут быть рассмотрены как общие для всех социолектов и относящиеся к постоянным универсальным проявлениям

<sup>1</sup> Proceedings of the 19th International Conference on Speech and Computer, SPECOM 2017, ser. Lecture Notes in Artificial Intelligence (including subseries Lecture Notes in Computer Science), 10458 LNAI. Karpov A., Potapova R., Mporas I. (eds). Cham; Heidelberg; New York; Dordrecht; London: Springer International Publishing, 2017. 831 p.

реализации повседневного русского языка; в) потенциально диагностические признаки, которые отражают количественные на текущий момент различия между исследуемыми группами (пока статистически не оценены). Данное исследование базируется на теоретическом фундаменте отечественных лингвистов<sup>2</sup>. Метод лонгитюдного формирования фонограмм, содержащих повседневную русскую речь, изложен в докладе Арто Мустайоки (Хельсинки), Т. Шерстиновой (Санкт-Петербург) “Ретроспективное комментирование как предмет лонгитюдного процесса записи повседневной русской речи”. В докладе рассматриваются ситуации “неудавшейся” коммуникации при наличии обоих партнеров, что связано с переспросом всех видов и эмоциональным напряжением коммуникантов.

Авторы формируют суббазу различного рода переспросов в составе основной базы данных для повседневной русской речи. В докладе рассматриваются различные виды межличностной коммуникации, реализующиеся в режиме непонимания, уточнения, эмоционального напряжения и т.д.

Доклад Р. Потаповой и Л. Комаловой “Лексико-семантические показатели корреляции модальности “депривация – агрессия” в социально-сетевом дискурсе” (Москва) посвящен описанию результатов дальнейшей разработки базы данных, содержащих лексику агрессивных высказываний на материале четырех языков: русского, английского, испанского, татарского – репрезентантов различных языковых семей. В докладе представлены результаты количественного анализа русскоязычной базы данных с опорой на наличие высказываний, содержащих утверждения агрессивного характера (в режиме диалога). При этом учитывалось гендерное различие коммуникантов, что позволило установить особенности ответной реакции адресата на агрессию со стороны адресанта.

Наряду с корпусной лингвистикой активно разрабатываются программные способы извлечения информации об эмоциональной насыщенности текстов в интернете в режиме дихотомии “положительная оценка события” – “отрицательная оценка события”. Речь идет о так называемой “тональности текста”. Определение “тональности текста” вызывает особый интерес, так как позволяет уста-

новить программным способом эмоциональную окраску текста в интернете, а также при передаче сообщения по различным современным каналам связи. В качестве примера можно привести доклад “Корпус высказываний на испанском языке, предназначенных для анализа тональности на материале рекламных брендов в интернете” (М. Лоро, В. Родригес-Донцель и др.) (Испания). Данный доклад по своей целевой установке объединяет два направления: корпусную лингвистику и анализ тональности на материале рекламных постов в социально-сетевом пространстве интернета. Исследование включает анализ тональности текстов с ориентацией на восемь эмоционально-модальных состояний: любовь – ненависть, счастье – печаль, доверие – боязнь (недоверие), удовлетворенность – неудовлетворенность и др. Материалом исследования послужили тексты в твиттере, проанализированные с учетом наличия вышеупомянутых состояний, определяющих положительную или отрицательную тональность текстов-реакций на функционирующие в интернете рекламные бренды в области продуктов питания, банковского дела, продажи автомобилей, спорта и т.д. Авторы подчеркивают, что исследование является первым для испанского языка относительно оценочных вербальных показателей различных рекламных брендов в твиттере.

Особое внимание уделено исследованию текстов в социально-медийных каналах с позиции лексических, синтаксических и морфологических особенностей языка. Не обошли вниманием участники конференции проблематику, связанную с исследованием лингвистических особенностей текстов в интернете. Так, например, социально-сетевой дискурс подробно анализировался в докладе “Использование нейросетевой программы Doc2vec для автоматизированного анализа тональности кратких бытовых текстов” (Н. Маслова, В. Потапов (Москва)). В целом же особенностям вербальной, паравербальной и невербальной коммуникации с позиции корпусной лингвистики на конференции 2017 года было посвящено незначительное число докладов по сравнению с прошлогодними конференциями.

Не менее активно исследуются, как и ранее, индивидуальные знаки устной речи говорящих, “пропущенной” по различным каналам связи (IP-телефонии, Skype и др.), что обусловлено поиском решения задачи автоматической идентификации говорящего по голосу и речи, атрибутики говорящего, распознавания его физического и эмоционального состояния, определения его биоме-

<sup>2</sup> Например, Современный русский язык. Активные процессы на рубеже XX–XXI веков / Отв. ред. Л.П. Крысин. М.: Языки славянской культуры, 2008. 710 с.; Потапова Р.К., Потапов В.В. Язык, речь, личность. М.: Языки славянской культуры, 2006. 496 с.

трических параметров, распознавания особенностей изменения социальных признаков личности. В непосредственной связи с исследованиями в области тональности текста находится совершенно новое направление, базирующееся на понятии “акмеология”, что отражено в докладе Р. Потаповой и В. Потапова “Личность как акмеологическая сущность в социально-сетевом дискурсе (мульти-модальный подход)” (Москва). В докладе впервые представлена концепция формирования и динамического развития индивидуального портрета личности по данным социально- сетевого дискурса в интернете, рассматриваемого во времени с учетом изменения “речевого продукта” личности с опорой на мультидимодальный подход: вербальный, паравербальный, невербальный, экстравербальный. Каждая из вышеуказанных составляющих соотносится с: а) вербальным контентом устного или письменного высказывания информанта в социально-сетевом дискурсе; б) паравербальными просодико-интонационными особенностями устной речи или эмотивной диакритикой письменных сообщений; в) сопровождающими речевое поведение информанта мимикой, жестами, проксемикой; в) экстравербальными, показателями социально-сетевого дискурса: одеждой, ситуацией и т.д.

Особое внимание уделено когнитивному аспекту перцептивно-слухового восприятия эмоционально-модальных состояний коммуникативного взаимодействия партнеров, принадлежащих к различным языкам и культурам. В определенном смежном ракурсе с вышеприведенным примером находится разработка проблемы когнитивистики применительно к восприятию эмоционально-модальных состояний, представленная в докладе “Когнитивная энтропия при перцептивно-слуховой оценке эмоционально-модального состояния иноязычного партнера по коммуникации” (Р. Потапова, В. Потапов (Москва)). В докладе представлены результаты лонгитюдного исследования, охватывающего 10 лет и проводимого в России и Германии с опорой на многоаспектную культурно-языковую идиосинкразию носителей русского и немецкого языков. В результате исследования удалось установить статистически значимую дивергенцию в перцептивно-слуховой оценке одних и тех же эмоционально-модальных речевых стимулов носителями вышеуказанных языков.

В связи с развитием робототехники к исследованиям в области управления роботом техническими системами с помощью устной речи добавляются исследования, предметом которых является ис-

пользование невербалики (мимики, жестикуляции, проксемики), что объясняется расширением способов управления и контроля за поведением робототехнических систем. Особое внимание уделяется изучению общей сенсорики человека. Все эти подходы объединяются в интегративную область исследования: «Мультимодальное взаимодействие “человек–машина” с опорой на биометрические данные. В качестве примера обращения к невербалике и биометрическим данным следует упомянуть доклад М. Хлавача, И. Грубера, А. Карпова и др. (Чешская Республика; Санкт-Петербург) на тему «Формирование лицевых “ключевых” мимических точек», в котором описан полуавтоматический метод создания, считывания и формирования биометрических лицевых показателей, что необходимо для дальнейшей разработки системы управления роботами.

Из материалов докладов следует, что не менее большое внимание по сравнению с предыдущими конференциями SPECOM уделено проблеме автоматического распознавания речи. При этом наблюдается расширение круга языков, на материале которых разрабатываются новые алгоритмы.

Результаты исследований в области распознавания речи отражают два направления автоматического распознавания речевых образов на материале различных языков (например, доклады М. Мариуса, С. Штрюкера и А. Вайбеля на тему “Автоматическое распознавание мультилингвальной речи” (ФРГ); Е. Пакоци, Б. Попович и Д. Пекар “Оптимизация лингвистической модели, построенной на основе глубокой нейронной сети, предназначенной для системы распознавания сербской речи” (Сербия); Э. Эдвардса, В. Сэллоума, Г.Р. Финли и др. (США) на тему “Неавтоматическое распознавание английской речи в сфере медицины” и др.).

На конференции было представлено значительное число докладов, посвященных автоматической идентификации говорящих по голосу и речи, например, “Автоматическое обнаружение курящего говорящего по речи, передаваемой по телефонному каналу” (А. Пурьям, С. Хезараки, С. Сафави и др. (Дания; Иран; Великобритания и Бельгия)); “Различение голосов говорящих на базе фузионного подхода” (Х. Саюд, С. Квамур, З. Хамадахе (Алжир)); “Нейронно-сетевой дескриптор диарезации говорящих по телефонному каналу” (З. Заич, Я. Зелинка, Л. Мюллер (Чешская Республика)), “Эксперименты по обнаружению аномальных личностных особенностей на базе просодических при-

знаков” (Ц. Файе, А. Делэ, Д. Лолив и др. (Франция)) и др.

Одной из классических тем конференции остается тема “Преобразование текст – речь”, т.е. синтез речи с опорой на письменный текст. Были представлены результаты разработок в области синтеза речи. Например, в докладе В. Захариева, Е. Азарова и А. Петровского (Белоруссия) описана многоголосовая система синтеза речи на базе конкатенативного подхода. В докладе «Голосовой обмен в системе TTS (text-to-speech) с настройкой на “целевого говорящего”» использована, в частности, методика гауссовой, смешанной модели и метода текстонезависимого обучения, базирующаяся на модели марковских цепей и модифицированного алгоритма Витерби.

Таким образом, можно констатировать, что проведенная XIX Международная конференция SPESOM'2017 продемонстрировала широкий тематический охват и дальнейшее углубление знаний с учетом междисциплинарного подхода к решению приоритетных задач в области человеко-машинного взаимодействия на базе цифровых технологий.

Вместе с тем сугубо лингвистических тем докладов значительно меньше, чем тем, посвященных программным и цифровым концепциям с использованием информации об особенностях письменной и устной речи применительно к различным языкам мира.

*Р.К. Потапова*  
*доктор филологических наук,*  
*профессор*  
*Московского государственного*  
*лингвистического университета,*  
*Россия, 119034, г. Москва, Остоженка 38*  
*rkpotapova@yandex.ru*

*В.В. Потапов*  
*доктор филологических наук,*  
*старший научный сотрудник*  
*филологического факультета Московского*  
*государственного университета*  
*им. М.В. Ломоносова,*  
*Россия, 119991, г. Москва, Ленинские горы,*  
*д. 1, стр. 51*  
*volikpotapov@gmail.com*

*Дата поступления материала*  
*в редакцию 23 ноября 2017 г.*

*Rodmonga K. Potapova*  
*Doctor of Philological Sciences,*  
*Professor at the Moscow*  
*State Linguistic University, Ostozhenka 38*  
*Moscow, 119034, Russia,*  
*rkpotapova@yandex.ru*

*Vsevolod V. Potapov*  
*Doctor of Philological Sciences,*  
*Senior Researcher at the*  
*Lomonosov Moscow State University,*  
*1-51 Leninskie Gory,*  
*Moscow, 119991, Russia*  
*volikpotapov@gmail.com*

*Received by Editor on November 23, 2017*