

Компьютерная лингвистика и интеллектуальные технологии

По материалам ежегодной Международной
конференции «Диалог» (2015)

Выпуск 14

В двух томах

Том 1. Основная программа конференции

Computational Linguistics and Intellectual Technologies

Papers from the Annual International
Conference "Dialogue" (2015)

Issue 14

Volume 1 of 2. Main conference program

УДК 80/81; 004
ББК 81.1
К63

Программный комитет конференции выражает
искреннюю благодарность Российскому фонду
фундаментальных исследований за финансовую поддержку,
грант № 15-07-20554 Г

Редакционная
коллегия:

*В. П. Селегей (главный редактор), А. В. Байтин,
В. И. Беликов, И. М. Богуславский, Б. В. Добров,
Д. О. Добровольский, Л. М. Захаров, Л. Л. Йомдин,
И. М. Кобозева, Е. Б. Козеренко, М. А. Кронгауз,
Н. И. Лауфер, Н. В. Лукашевич, Д. Маккарти, Й. Нивре,
Г. С. Осипов, В. Раскин, Э. Хови, С. А. Шаров, Т. Е. Янко*

Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27–30 мая 2015 г.). Вып. 14 (21): В 2 т. Т. 1: Основная программа конференции. — М.: Изд-во РГГУ, 2015.

Сборник включает 69 докладов международной конференции по компьютерной лингвистике и интеллектуальным технологиям «Диалог 2015», представляющих широкий спектр теоретических и прикладных исследований в области описания естественного языка, моделирования языковых процессов, создания практически применимых компьютерных лингвистических технологий.

Для специалистов в области теоретической и прикладной лингвистики и интеллектуальных технологий.

© Редколлегия сборника «Компьютерная лингвистика и интеллектуальные технологии» (составитель), 2015

ЛИНГВИСТИЧЕСКИЙ АНАЛИЗ В СИСТЕМАХ ИДЕНТИФИКАЦИИ ДИКТОРА: ИНТЕГРАТИВНЫЙ КОМПЛЕКСНЫЙ ПОДХОД НА БАЗЕ ЭКСПЕРТОЛОГИИ

Галяшина Е. И. (galyashina@gmail.com)

Московский государственный университет
им. О. Е. Кутафина, Москва, Россия

В статье предлагается концепция экспертной технологии идентификации диктора посредством интегрирования и комплексирования перцептивных, акустико-фонетических и собственно лингвистических методов анализа звучащей речи, определяющих общую и специальную компетенцию речевода. Автор анализирует ключевые понятия и термины процедуры индивидуально-конкретного отождествления диктора в аспекте современной экспертологии. Показано, что в условиях, когда сравнению подвергаются оцифрованные речевые сигналы, полученные по неизвестному алгоритму преобразования и при невозможности исключить их фальсификацию, для достоверного и надежного решения идентификационной задачи возрастает роль профессиональных лингвистических компетенций, которые должны обладать достаточной интегративной широтой междисциплинарного знания и комплексностью познания феномена индивидуальности речевого следа.

Ключевые слова: лингвистический анализ, идентификация диктора, экспертные системы, интегрированные знания, экспертология

LINGUISTIC ANALYSIS IN THE SPEAKER IDENTIFICATION SYSTEMS: INTEGRATED COMPLEX EXAMINATION APPROACH BASED ON FORENSIC SCIENCE TECHNOLOGY

Galyashina E. I. (galyashina@gmail.com)

Kutafin Moscow State Law University, Moscow, Russia

The article proposes the concept of the integrated expert techniques for speaker identification on the domain of complex acoustic-phonetic and linguistic methods of oral speech analysis, defining general and special forensic expert competences. The result of forensic speaker identification used as evidence must exhibit a high level of reliability. The author examines the key concepts and terms of the procedure of individual-specific speaker identification in the aspect of modern expertology (forensic science). The paper states the need to take into account that the role of professional linguistic competences increases in conditions when digitized speech signals are compared, algorithm of coding is indefinite and falsification of utterances is not excluded. To solve this problem the author proposes a multistage approach consisting of a parallel application of instrument and technical methods together with aural-perceptual, waveform and sonogram investigation and sophisticated linguistic analysis. The main attention is paid to the linguistic component of the complex integrated approach based on the phonetic and semantic analyses. It is stated that individualized speech unit is formed by a system of miscellaneous formal and semantic relations of structural speech components in linguistic contents. The proposed method of integration of the multilevel speech modules was implemented in forensic linguistic methodology of speaker identification technique. This made it possible to considerably increase the reliability of the expert's decision and provided an opportunity to use it as a component of the multistage system for speech utterances authentication.

Key words: linguistic analysis, speaker identification, expert systems, integrated expert techniques, forensics

Системный анализ лингвистических методов, применяемых в современных автоматизированных системах идентификации диктора, имеющих экспертную направленность, позволил установить сложную динамическую природу индивидуальности речевого следа, выступающего в качестве объекта экспертизы. Для его полного и всестороннего исследования требуются интегрированные знания эксперта, которые должны обладать достаточной широтой и вместе с тем глубиной познания исследуемого речевого объекта. При их использовании в практической деятельности формируются дополнительные профессиональные речеведческие компетенции, представляющие собой комплекс практических навыков и умений, получаемых опытным путем, закрепление которых должно входить в систему подготовки по соответствующей экспертной специализации. Компетенция, определяемая экспертной технологией отождествления диктора, это не механическое соединение разных областей научных знаний о речи, а комплексное интегративное образование, позволяющее сформировать требуемые профессиональные компетенции на базе современной экспертологии. Применительно к лингвистическому анализу это, в первую очередь, относится к компетентной оценке индивидуально-определенной специфики фонетической и смысловой организации высказываний в речевом потоке.

В настоящее время существует ряд методических подходов, базирующихся на применении автоматизированных экспертных систем текстозависимой и текстонезависимой идентификации диктора. Вместе с тем их возможности ограничены при исследовании речевого сигнала, продуцированного в условиях компьютерно-опосредованной речевой коммуникации, при образовании речевого

следа путем цифровой передачи сигнала по каналам мобильной связи, а также воздействию на структуру звукового сигнала при его цифровой обработке, сжатии и иных способах модификации. Существующие автоматические системы и правила принятия идентификационного решения не позволяют достигать результатов, достоверность, надежность и воспроизводимость которых отвечает требованиям обеспечения доказательств в практике российского судопроизводства¹. Доля неавтоматизированного труда эксперта остается значительной, что снижает степень объективности вывода. Технология производства экспертизы предусматривает разрозненное использование методов акустического и лингвистического анализа, на практике реализуемого разными субъектами, обладающими различными компетенциями, независимо друг от друга. Общее решение о принадлежности речевого сигнала на проверяемой фонограмме конкретному диктору принимается двумя экспертами (инженером и лингвистом) с разным объемом специальных знаний путем суммирования результатов параметрического сравнения инженером противопоставленных сигналограмм и социолингвистического портрета, сформированного на уровне слуховой перцепции лингвиста. Лингвистические признаки структурно представлены перечнями описаний, не детерминированными по их специфичности. Не учитывается объективно-системный характер речи как бинарного процесса, опирающегося на индивидуальные психофизиологические процессы в организме человека и оформленного потоком языковой последовательности звуков с семантическим содержанием, дополняемым всевозможными интонационно-эмоциональными оттенками.

Можно констатировать, что единая общепринятая научно-обоснованная методика отождествления диктора пока отсутствует, а в экспертных организациях России изолированно применяются методы акустического и аудитивно-лингвистического анализа, механистически разрывающие неделимое знаковое единство речи как продукта языковой деятельности. Это приводит к тому, что на одних и тех же объектах сравнения одним методом могут быть получены результаты, невозпроизводимые применением иного метода, либо эксперты могут формулировать прямо противоположные выводы, устанавливая тождество одним методом и отсутствие такого — другим. В экспертных заключениях не приводятся вероятностно-статистические данные о надежности принятия идентификационного решения (вероятности совершения ошибки первого, второго рода — «пропуска цели» или «ложного захвата цели») по отдельным видам исследований. Все применяемые методические материалы носят рекомендательный характер, не содержат строгих предписаний и формальных правил, оставляя эксперту широкое поле для субъективного усмотрения как по выбору количества и комбинаторике измеряемых параметров речевого сигнала, набору вычисляемых акустических признаков, списку интерпретируемых лингвистических признаков и их оценке, и в этом смысле собственно методиками не являются.

В экспертологии методика, понимаемая как «система категорических или альтернативных научно-обоснованных предписаний по выбору и применению

¹ См., например, Обзор методов идентификации на электронном ресурсе <http://www.forenex.ru> — дата последнего посещения 08.01.2015.

в определенной последовательности и в определенных существующих или создаваемых условиях методов, приемов и средств (приспособлений, приборов и аппаратуры) для решения экспертной задачи»², занимает в настоящее время одно из центральных мест. Методика может содержать несколько вариантов решения одной экспертной задачи с учетом специфики ее условий или особенностей представленных объектов, однако, каждый этап принятия экспертом конкретного решения в случае выбора альтернативных вариантов должен быть изложен в заключении эксперта и мотивирован.

Феномен экспертной методики как научно-обоснованный алгоритм решения конкретной практически ориентированной задачи, позволяющий формировать критерии для оценки достоверности заключения эксперта как доказательства, привлекает в условиях современных реалий судопроизводства особое внимание. Это тем более важно, что получаемые по имеющимся методическим подходам идентификации диктора результаты нередко невоспроизводимы и носят неоднозначный, противоречивый характер. Видимо, поэтому в своих заключениях эксперты часто не приводят ссылок на конкретную экспертную методику либо, подменяют понятие экспертной методики понятием «методические материалы», «методические рекомендации», смешивают общенаучные и частнонаучные методы, описывают содержание и ход проведенного исследования с перечислением типовых, а не конкретно использованных технических средств, не указывают логические основания вывода и правила принятия ими того или иного решения³.

На примере, практики экспертного отождествления лица по фонограммам речи, этот тезис становится особенно актуальным. Обзор существующих методов и инструментов анализа речевого сигнала показывает, что для решения задачи идентификации диктора в отечественной экспертизе присутствуют разные подходы.

Первый реализован на основе автоматизированной системы идентификации дикторов «Диалект»⁴, изложенный в соответствующих пособиях для экспертов⁵. В структуру функциональной схемы проведения идентификационных исследований на системе «Диалект» входят измерение спектрально-временных параметров гласных и согласных звуков, вычисление акустических признаков речи неизвестного лица на спорной фонограмме и образцах речи

² Россинская Е. Р., Галяшина Е. И., Зинин А. М. Теория судебной экспертизы. — Учебник. — М.: Норма, 2013, с. 130.

³ Подробнее см. Россинская Е. Р., Галяшина Е. И. Настольная книга судьи: судебная экспертиза. — Москва: Проспект, 2014, с. 303–340.

⁴ В настоящее время данный метод реализован в программе «Диалект», ее коммерческой версии «Фонэкси», а также в системе PhonoBase. Способ идентификации личности по фонограммам произвольной устной речи в части инструментального анализа был запатентован (патент № 2107950 от 08.08.1996, G10L 5/06: Байчаров Н. В., Карлин И. П., Кураченкова Н. Б., Линьков А. Н., Попов Н. Ф., Савельев Ю. И., Тимофеев И. Н., Фесенко А. В.).

⁵ Идентификация лиц по устной речи на русском языке. Методика «Диалект». Пособие для экспертов. Издание 2-е, переработанное и дополненное / Н. Б. Кураченкова, Н. В. Байчаров, М. А. Ермакова. Под редакцией В. М. Богданова. — М., 2007.

подозреваемого лица, добавление в имеющуюся статистическую базу акустических признаков речи разных лиц информации о векторах признаков речи данного подозреваемого лица, переобучение системы, определение «веса» (информативности) совпадающих признаков и порогов допустимой вариативности акустических признаков. Система обладает возможностями адаптации процедуры принятия решения путем произвольного исключения из сравнения совпадающих или различающихся признаков, а также изменения порога принятия идентификационного решения. Однако, по мнению некоторых специалистов: «в целом данный метод соответствует уровню развития речевых технологий конца прошлого века и не учитывает многонационального характера современной преступности, объективных изменений в русском языке, появления и развития цифровых средств речевой связи»⁶. Отметим, что лингвистические признаки представлены в комплексе «Диалект» списком воспринимаемых на слух характеристик речевого потока, слова, фразы и звука, примеры звучания которых представлены в базе эталонов в системе. При этом эксперт выбирает из предлагаемых системой списка признаков те эталоны, которые в большей степени соответствуют его слуховому восприятию.

Второй подход основан на применении в качестве инструментария для измерения и вычисления речевых параметров различных звуковых редакторов, разработанных отечественными и зарубежными специалистами⁷, которые требуют от пользователя высокой квалификации, поскольку методы обладают невысокой степенью автоматизации, в том числе блоками автоматического сравнения получаемых результатов с принятием идентификационного решения на уровне экспертной оценки. Лингвистический анализ при втором подходе сводится к слуховой перцепции и фонетическому описанию сегментных и суперсегментных особенностей произносительных навыков. Решение принимается экспертом методом субъективной оценки выявленных совпадений и различий.

В этом ряду особо следует отметить новую версию звукового редактора SIS II v2.0, в котором реализованы модули попарного и обобщенного идентификационного решения. В этом блоке пользователь имеет возможность по своему усмотрению менять границы применимости каждого из отдельных методов исследования, изменяя таким способом его весовой коэффициент в общем решении. Лингвистический анализ представлен перечислением просодических особенностей с элементами субъективного психолого-социо-речевого портрета диктора⁸.

⁶ См., например, Обзор методов идентификации [электронный ресурс]: <http://www.forenex.ru>.

⁷ Например, OTExpert (ООО «ОТ-Контакт», Москва), Justiphone (ООО «Целевые технологии», Москва), SIS II v2.0 (ООО «Центр речевых технологий», Санкт Петербург), а также Praat (Амстердамский университет), SFS (Speech Filing System) и др. Подробный список представлен в монографии Галяшина Е. И. Основы судебного речеведения. — М: Стэнси, 2003, с. 225–229.

⁸ Булгакова Е. В., Краснова Е. В. Экспертные системы и методы идентификации диктора // ИЗВ. ВУЗОВ. ПРИБОРОСТРОЕНИЕ. 2014. Т. 57, № 2, с. 58–63.

Обзор современных методических подходов, используемых в экспертной практике показал:

- отсутствие терминологического единообразия в применяемых подходах к решению задачи отождествления диктора с достаточной точностью и надежностью, приемлемой для целей доказывания в судопроизводстве;
- множественность подходов и отсутствие единой, унифицированной методики;
- изолированное использование методов акустического и слухового фонетического анализа независимо друг от друга с суммированием полученных результатов;
- применение методов акустического анализа, которые базируются на презумпции наличия у эксперта информации об источнике происхождения представленного на фонограмме следа речевого сигнала и путей (способов) его модификации от источника (диктора) до регистрации на носителе;
- лингвистический анализ представлен многомерным недетерминированным признаковым пространством;
- применение аудитивного (лингвистического) анализа базируется на презумпции, что реплицированная диалогическая коммуникация, атрибутирована по принадлежности ее участникам;
- информационная значимость алгоритмически определяется для совпадающих признаков, вес различающихся признаков устанавливается экспертом субъективно;
- наличие в автоматизированных системах адаптивных процедур, позволяющих вмешиваться в процедуру принятия решения;
- общее решение принимается эвристически, носит во многом еще субъективный характер и в значительной мере зависит от опыта и багажа знаний эксперта.

Проблема осложняется тем, что надежность идентификационного решения во многом зависит от качества речевого сигнала, записанного на материальный носитель, тогда как речь человека при прохождении через каналы цифровой мобильной связи неизбежно искажается⁹. Идентификация диктора в сигнале, подвергшемся модификации, кодированию, сжатию и иным видам цифрового (а иногда и аналогового) преобразования значительно затруднена¹⁰. При передаче речи по каналам мобильной связи экспертному исследованию подвергается

⁹ Галяшина Е. И., Галяшин В. Н. Цифровые фонограммы как судебное доказательство, Воронежские криминалистические чтения, — Воронеж: Изд-во Воронежского гос. университета, № 8, 2007, с. 71.

¹⁰ На практике нередко возникает задача установления личности анонимного свидетеля, голос которого намеренно искажается по неизвестному алгоритму с целью затруднения или исключения возможности его опознания на слух или идентификации по акустико-фонетическим параметрам. Доступным для исследования при этом остается уровень языковой организации речи, отражающий индивидуальную характеристику манеры речи говорящего, обусловленной спецификой образа мышления, понимания темы, полноты аргументации, связности, цельности, информативности речевого сообщения.

синтезированный речевой сигнал, из которого алгоритмом кодирования исключены (искажены) существенные идентифицирующие диктора признаки.

Методические требования к фонограммам, в которых сигнал подвергался дискретизации, цифровому преобразованию, сжатию по неизвестному экспертам алгоритму, ставят их в разряд объектов, в которых установить экспертными методами достоверность отображения первичной информации о голосовом источнике невозможно. Без решения вопросов о степени влияния каналов связи, алгоритмов цифровой обработки на каждую группу признаков, участвующих в процедуре идентификации диктора экспертное исследование речевого следа не может быть надежным и достоверным.

При использовании цифровых каналов передачи сигнала на носителе фиксируется не сам исходный объект (например, звуковая волна) или его полное отражение (например, электрический сигнал от микрофона), а его абстрактная (математическая) модель. По каналу связи передается не нативная речь, а некий набор кодов и символов по которым на оконечном аппарате восстанавливается (синтезируется) сигнал, похожий на исходный, но который не является исходным сигналом. Причем эта абстрактная модель характеризуется двумя основными элементами — видом используемой математической модели и ее параметрами. Когда решают вопрос о выборе абстрактной модели для описания реальной акустической волны, то во главу вопроса ставят сложность устройства записи, удобство передачи полученной цифровой записи по системам связи и ее хранения на материальных носителях, но никак не требования максимально полного сохранения ее индивидуальных признаков значимых для решения задач отождествления диктора.

Речевой сигнал считается пригодным для идентификации диктора при обработке его алгоритмами сжатия с определенной нижней границей скорости цифрового потока, а именно:

- 32 кбит/с для ADPCM;
- 9,6 кбит/с для линейных предсказателей.

Речевой материал может быть признан условно пригодным для идентификации диктора, когда эти значения могут составлять:

- 16 кбит/с для ADPCM;
- 8 кбит/с для линейных предсказателей.

Во остальных случаях кодированный речевой сигнал должен быть признан, не пригодным для экспертной идентификации диктора как физическими методами, так и лингвистическими, поскольку фонетическая структура речевого отрезка, прошедшего такую обработку, существенно отличается от оригинала¹¹.

¹¹ Желудков Р. Н., Тимко Е. В., Усков К. Ю. О влиянии сжатия речи на допустимость речевой фонограммы в уголовное судопроизводство. Материалы 2-ой Всероссийской конференции «Теория и практика речевых исследований» (АРСО-2001). — М., 2001. — с. 110–116 [Электронный ресурс] http://expert.com.ua/kniise/articles/zhel1201_2.htm

Всесторонность и полнота экспертного исследования подразумевает изучение речевого следа во всех идентификационно значимых аспектах, определяющих индивидуальность и неповторимость функционально-динамических навыков звучащей речи проверяемого лица. В этой связи для того чтобы оценить, влияние искажений на пригодность речевого сигнала для идентификации, необходимо знать: тип и марку мобильного телефона, используемого абонентами при разговорах, в какой сети разговоры происходили, какой кодек и с какой скоростью передачи речевого сигнала использовался, какое устройство применялось для регистрации сигнала на материальном носителе. Без такой информации эксперт не сможет не только провести всестороннее и полное исследование речи, но даже достоверно оценить пригодность фонограмм для идентификационного исследования.

Придание большей значимости результатам акустического анализа может привести к экспертной ошибке, поскольку речь двух заведомо разных лиц, передаваемая по каналам сотовой связи, подвергшаяся разным алгоритмам цифровой обработки, по результатам акустического сравнения может случайным образом совпасть.

При отсутствии у экспертов достоверной информации о технологической цепочке записи представленных фонограмм возрастает роль собственно лингвистических признаков речевой индивидуализации.

Лингвистические идентификационные признаки, которые отражают индивидуальные речевые навыки говорящего, определяются индивидуально-стилистической манерой речи, навыками владения языком, на котором осуществляется речевое общение, а также интеллектуальными способностями и особенностями изложения мыслей. Система и классификация лингвистических признаков речевой индивидуализации для русскоязычных дикторов достаточно подробно изучена, описана, внедрена в экспертную практику¹².

Устная речевая деятельность — продукт реализации двух видов навыков, определяющих физическую и собственно языковую сторону речевого сигнала в коммуникации. Поскольку наибольшему воздействию подвержена физическая сторона сигнала, то возрастает роль устно-речевых навыков, в основе которых находятся лингво-психологические автоматизмы речевой деятельности, а также рече-двигательных навыков, базирующихся на механизме функционально-динамического стереотипа. Однако такой подход требует применения процедуры отождествления диктора, базирующейся на теоретических основаниях, отличающихся от принятых в естественнонаучных приложениях, где идентификация сводится к сравнению нескольких автоматически или полуавтоматически измеряемых параметров речевого сигнала и подразумевает вероятностное (в той или иной мере достоверное) «попадание в цель» или «пропуск цели» с учетом допускаемых пределов (ошибок первого, второго рода). Задача «верификации» (подтверждения тождества путем попарного сравнения диктора и образца его голоса) проще задачи «идентификации» (установления

¹² См., подробнее Галяшина Е. И. Судебная фоноскопическая экспертиза. — М.: Триада, 2001. с. 200–208.в

тождества многомерным сравнением неизвестного диктора с образцами речи нескольких подозреваемых дикторов»), т.к. при верификации принимается альтернативное решение.

Экспертная идентификация диктора по фонограмме речи скорее соотносится с задачей идентификации на открытом множестве дикторов, совмещающая в себе наиболее сложные случаи верификации, требующей увеличения количества образцов дикторов в эталонном множестве, объем которого соизмерим с числом населения земного шара. При том, что при бесконечно большом числе дикторов вероятность ошибки идентификации стремится к единице.

Понятие лингвистического идентификационного признака определяется как выражение индивидуализирующего свойства системы речемыслительных навыков человека. Для того чтобы лингвистический признак мог быть использован в качестве идентификационного, он должен отвечать нескольким условиям. Главным является специфичность, оригинальность признака, который наиболее точно и полно отражает свойство объекта. Под оригинальностью понимают нетипичность, отклонение признака от средних величин и норм. Такой признак имеет тем большее значение, чем реже он встречается у однородных объектов одной группы. Под специфичностью признака понимают его способность выделять объект, отграничивать его от группы схожих, похожих объектов. Второй особенностью идентификационного признака является его выраженность, способность к систематическому адекватному отображению в речевом следе, модели (когда приходится иметь дело не с самими признаками, а с их отображениями) при передаче по каналам цифровой связи. Признак должен быть воспроизводим в каждом случае образования речевого следа, и его отображение должно однозначно передавать информацию о свойствах речемыслительных навыков. Третье — признак должен быть относительно устойчивым при модификации речевого сигнала. Лингвистические признаки могут быть групповыми или частными. Групповые признаки — присущи однородным объектам, они позволяют выделить данную группу из других подобных групп, а также отнести объект по отображению групповых признаков к данной группе (например, признаки говора, диалекта, социолекта, гендера и т.п.). Частный признак позволяет выделить конкретный объект из группы однородных ему.

Индивидуализировать объект способна только индивидуальная совокупность частных лингвистических признаков, неповторимая в своей комбинаторике и локализации.

Экспертная практика показывает, что семантическая структура устного текста как отражение речемыслительной деятельности человека уникальна и неповторима. Выбор языковых средств и их упорядоченность в тексте идет от субъективного личностного смысла, отношения говорящего к действительности и его коммуникативной цели.

По мере формирования механизмов речи у человека образуется функционально-динамический стереотип устной речи, основанный на множественной системе языковых и фонационных навыков. Система языковых навыков, входящих в структуру функционально-динамического комплекса имеет сложную

структуру, но в ней можно выделить совокупность речевых средств, включающих в себя как вариации трехуровневой модели языковой личности¹³, так и ее коммуникативную способность, т. е. не только языковые компетенции, но и речевые умения.

На каждом уровне организации языковой личности (вербально-семантическом, лингвокогнитивном и мотивационно-прагматическом) можно выделить соответствующую подструктуру лингвистических признаков, характеризующих проявление языковой личности в конкретной ситуации речевого акта: индивидуальный лексикон, тезаурус концептов, признаки дискурсивного мышления и интеллектуальных навыков.

Лексикон личности понимается как словарный запас и динамическое устройство с подвижными связями и отношениями, которые обеспечивают оперирование языковыми единицами в процессе построения и понимания текста. Экспертная практика показывает, что в индивидуально-личностный лексикон входят помимо словаря лексических единиц и закрепленные в речевом опыте грамматические значения и навыки, определенные синтаксические шаблоны фраз с застывшей в них грамматикой¹⁴. Это — устойчивое употребление стереотипного набора словосочетаний и синтаксических конструкций при описании одних и тех же объектов, ситуаций, обстоятельств. Наличие речевых стереотипов указывает на ведущую роль лексико-семантической организации текста, являющейся вербализацией языковой картины мира говорящего. Особенности этой организации, выражающей вектор семантического наполнения высказывания посредством предпочитаемых личностью языковых форм и речевых оборотов, составляет тезаурус, отражающий иерархически организованные концепты. Понятийно-смысловая организация текста позволяет выявить признаки интеллектуальных и познавательных навыков человека, которые невозможно выявить только путем анализа лексико-грамматической, стилистической организации текста. В качестве интеллектуальных навыков идентификационно значимыми являются те особенности мышления личности, которые соотносимы со способом представления действительности в наглядных образах или абстрактных понятиях, т. к. по ним можно судить по соотношению конкретной и абстрактной лексики.

В качестве идентификационных признаков выделяются индивидуальные предпочтения по выбору стереотипных языковых средств (фразеологизированных и лексически ограниченных речевых оборотов), используемых для выражения эмоций, чувств, состояний человека, его отношения к действительности, поступкам или поведению иных лиц, ситуаций и т. п.

¹³ Караулов Ю. Н. Русский язык и языковая личность. — М.: УРСС, 2002, с. 60–61.

¹⁴ По Н. Ю. Шведовой — «шаблонные фразы», не требующие «комбинирования» и не подающиеся отчетливому членению — «построения в субстантивном значении этого слова». См. Н. Ю. Шведова Очерки по синтаксису русской разговорной речи. //Ран Институт русского языка. Отв. Редактор В. В. Виноградов. — М.: Азбуковник, 2003, с. 7.

Существуют два метода определения идентификационных лингвистических характеристик: 1) исследовательский анализ речевого потока по определенному набору признаков (например, по базе автоматизированной системы «Диалект» и ее производных) и принятие решения по результатам сравнения значений признаков по лингвистическим протоколам, 2) прямая аудиторская оценка экспертом текста с последующим соотнесением полученного лингвистического описания с набором лингвистических признаков, зафиксированных в базе автоматизированной системы, и оценкой полученных совпадений и различий.

Нами была поставлена задача выявить среди названных признаков наиболее специфичные, отвечающие критериям надежности, воспроизводимости и устойчивости. С этой целью были проанализированы признаки как плана выражения (сегментного и супraseгментного уровня), так и плана содержания (лексико-семантического и семантико-синтаксического уровня). Проведенный анализ позволил сделать следующие выводы.

Особенность языковой организации речевой реализации приобретает идентификационную значимость только в том случае, если, во-первых, данная реализация является нарушением языковой системы, во-вторых, отклонением от принятого в данной социальной общности употребления языковых средств в устной речи; в-третьих, нарушением узуса — несоответствием используемых языковых средств целям и условиям данной ситуации коммуникации, связанное с неумением выбрать те из них, которые наиболее адекватно отражают социальные нормы поведения в данном дискурсе, в-четвертых, речевая реализация является индивидуально предпочитаемым вариантом из ряда вариаций, допускаемых языковой системой и (или) узусом.

Значениями признака будут его конкретное проявление и частота встречаемости в речи данного диктора по сравнению с среднестатистической частотой встречаемости в данной языковой общности.

Наибольший идентификационный вес имеют те признаки, которые относятся к развернутым семантическим комплексам с присущими им фонетической организацией, синтаксическими и синтагматическими связями. К ним относятся особенности структурной организации понятийно-смысловых элементов устного текста и языковые построения, отображающие навыки дискурсивного мышления и интеллектуальных навыков.

К дискурсивным навыкам мышления традиционно относят: использование того или иного вида определения, деления понятия; того или иного вида силлогизма (категорического, условно-категорического, условного и т. д.), энтитемы (сокращенного силлогизма), например предпочтительное использование энтитемы с пропущенной большей или меньшей посылкой или заключением; того или иного вида доказательства и т. д.; особенности общей логической структуры (формы аргументации); наличие акцентуации — той или иной формы выделения важных для говорящего понятий, суждений; характер исправлений смыслового характера. К интеллектуальным навыкам относят: признаки интеллектуальных навыков восприятия действительности (характер восприятия), характеризующий аналитический, синтетический или смешанный

тип восприятия; признаки навыков переработки информации, характеризующие описывающий, объясняющий или смешанный тип, признаки интеллектуальных навыков акцентуации и аргументации, на какие аспекты (авторитеты, факты, обстоятельства) говорящий акцентирует внимание для обоснования своего мнения, признаки навыков оценки описываемых предметов, ситуаций.¹⁵

К этим признакам относятся: характер выражения субъективной модальности (вводные слова и конструкции, междометия, частицы, используемые в значении акцентирования или, оценки, отражающие чувства и переживания, отношение говорящего к предмету речи или действительности); характер и способ использования когнитивных элементов, структур и форм связи между ними (характер и способ аргументации, наличие или отсутствие логики изложения, характер акцентуации), а также способ отображения характера оценки (этической, эстетической, прагматической, эмоционально-экспрессивной).

Проведенное исследование позволило сделать следующие выводы. Индивидуальные особенности лексико-семантического и семантико-синтаксического уровней речи дают возможности судить об индивидуальных навыках речемыслительной деятельности, стилистике речи, языковой и коммуникативной компетенции говорящего. Устный текст как системное целое с иерархической организацией составляющих его структур (языковых, логических) может рассматриваться как проекция языковой личности, которая включает наряду с лексиконом, тезаурус личности, где отображен образ мира с прагматикой, системой целей и мотивов, установок и ценностных ориентаций.

Языковой компонент речевого сигнала, интерпретируемый как устный текст, а, следовательно, и содержащаяся в нем информация о дикторской индивидуальности независимы от акустической составляющей. Лингвистическая информация о языковой личности диктора, заключенная в тексте, является чисто аддитивной величиной относительно соответствующей голосовой информации. Лингвистический анализ при идентификации говорящего по фонограммам речи имеет свой объект и при исследовании этого объекта может опираться на комплекс вербально-семантических, лингво-когнитивных и мотивационно-прагматических признаков, независимых от параметрических характеристик акустического сигнала. Иными словами, информация о дикторе, содержащаяся в лингвистическом компоненте речевого сигнала, дополняет информацию, содержащуюся в параметрическом представлении сигнала, независима от нее, может использоваться самостоятельно. Для установления тождества языковой личности диктора необходимо выявление индивидуального комплекса индивидуализирующих лингвистических признаков, определенных на всех элементах, составляющих структуру речи на трех уровнях гносеологической модели языковой личности.

¹⁵ См. Роль человеческого фактора в языке. Язык и картина мира. — М., Наука, 1988; Леонтьев А. А. Лингвистическое моделирование речевой деятельности // Основы теории речевой деятельности. — М.: Наука, 1974; Вул С. И. Судебно-автороведческая идентификационная экспертиза. Методические основы. — Х.: ХНИИСЭ, 2007, с. 14–16.

Литература

1. Булгакова Е. В., Краснова Е. В. (2014), Экспертные системы и методы идентификации диктора // Известия вузов. Приборостроение. Т. 57, № 2, с. 58–63.
2. Галяшина Е. И. (2001), Судебная фоноскопическая экспертиза, Триада, Москва.
3. Галяшина Е. И. (2003), Основы судебного речеведения, Стэнси, Москва.
4. Караулов Ю. Н. Русский язык и языковая личность. — М.:УРСС, 2002.
5. Кураченкова Н. Б., Байчаров Н. В., Ермакова М. А. (2007), Идентификация лиц по устной речи на русском языке. Методика «Диалект», Институт Криминалистики ЦСТ, Москва.
6. Россинская Е. Р., Галяшина Е. И., Зинин А. М. (2013), Теория судебной экспертизы, Норма, Москва.
7. Россинская Е. Р., Галяшина Е. И. (2014), Настольная книга судьи: судебная экспертиза, Проспект, Москва.

References

1. Bulgakova E. V., Krasnova E. V. (2014) Expert Systems and Methods of Speaker Identification [èkspertnyie sistemy i metody identifikatsii diktora].// Izvestia vuzov. Priborostroenie. Vol. 57, № 2, с. 58–63.
2. Galyashina E. I. (2001), Forensic Phonogram Investigation. [Sudebnaja Ffonoskopicheskaja èkspertiza], Triada, Moscow.
3. Galyashina E. I. (2003), The Fundamentals of Forensic Speech Science [Osnovy sudebnogo rechevedenija], Stènsi, Moscow.
4. Karaulov Ju. N. Russkij jazik I jazykovaja lichnost. URSS, Moscow, 2002.
5. Kurachenkova N. B., Bajcharov N. V., Jermakova M. A. (2007), Oral Russian Language Speaker Identification. Method “Dialect” [Identifikatsija lits po ustnoj rechi na russkom jazyke. Metodika “Dialekt”, Institut Kriminalistiki TsST], Moscow.
6. Rossinskaja E. R., Galyashina E. I., Zinin A. M. (2013), The theory of forensic science. [Teorija sudebnoj èkspertizy, Norma, Moscow.
7. Rossinskaja E. R., Galyashina E. I., Zinin A. M. (2014), Forensic expertise: a judge concise book [Nastol'naja kniga sudji: sudebnaja èkspertiza, Prospekt, Moscow.

Авторский указатель

Аванесов В.	т. 2: 34	Кашкин Е. В.	т. 1: 426
Адаскина Ю. В.	т. 2: 1	Кибрик А. А.	т. 1: 231, 487
Акинина Ю. С.	т. 1: 41	Кипяткова И. С.	т. 1: 240
Алимова И. С.	т. 2: 22	Киселева К. Л.	т. 1: 272
Андреев И.	т. 2: 133	Киселёв Ю. А.	т. 1: 254
Андрианов И.	т. 2: 34	Клячко Е.	т. 2: 119
Антонова А.	т. 1: 548	Князев С. В.	т. 1: 284
Апресян В. Ю.	т. 1: 2	Козлов И.	т. 2: 34
Арефьев Н. В.	т. 2: 105	Козлова Е. А.	т. 1: 169
Астраханцев Н.	т. 2: 34	Константинова Н.	т. 2: 88
Баранов А. Н.	т. 1: 19	Коротаев Н. А.	т. 1: 294
Бергельсон М. Б.	т. 1: 41	Котельников Е. В.	т. 2: 2
Бердичевский А.	т. 1: 30	Котов А. А.	т. 1: 308
Блинов П. Д.	т. 2: 2	Крейдлин Г. Е.	т. 1: 321
Богуславский И. М.	т. 1: 61	Кривнова О. Ф.	т. 1: 338
Бонч-Осмоловская А. А.	т. 1: 80	Крижановская Н. Б.	т. 1: 254
Браславский П. И.	т. 1: 254	Крижановский А. А.	т. 1: 254
Вепрева И. Т.	т. 1: 453	Крылова Т. В.	т. 1: 352
Вилинбахова Е. Л.	т. 1: 626	Кудинов М.	т. 1: 369
Вишнёва Е. А.	т. 1: 560	Кузьменко Е. А.	т. 1: 388
Галицкий Б. А.	т. 1: 141	Кустова Г. И.	т. 1: 376
Галяшина Е. И.	т. 1: 156	Кутузов А.	т. 2: 133
Гаращук Р. В.	т. 1: 169	Лазурский А. В.	т. 1: 61
Гончарова М. Б.	т. 1: 169	Левонтина И. Б.	т. 1: 104
Гришина Е. А.	т. 1: 183	Лесота О. О.	т. 2: 105
Грозин В. А.	т. 1: 202	Лобанов Б. М.	т. 1: 414
Гусарова Н. Ф.	т. 1: 202	Лопухина А. А.	т. 2: 145
Даниэль М. А.	т. 1: 95	Лопухин К. А.	т. 2: 145
Диконов В. Г.	т. 1: 61	Луканин А. В.	т. 2: 105
Добренко Н. В.	т. 1: 202	Лукашевич Н. В.	т. 2: 2, 88
Добровольский Д. О.	т. 1: 104	Лютикова Е. А.	т. 1: 696
Добрушина Н. Р.	т. 1: 118	Ляшевская О. Н.	т. 1: 426
Драгой О. В.	т. 1: 41	Майоров В.	т. 2: 34
Загулова М. А.	т. 2: 65	Малафеев А. Ю.	т. 1: 441
Зализняк Анна А.	т. 1: 683	Малых В. А.	т. 2: 65
Захаров В. П.	т. 1: 667	Мейер К. М.	т. 2: 88
Зинина А. А.	т. 1: 308	Меньшиков И. Л.	т. 1: 254
Иванов В. В.	т. 2: 2, 22, 65	Мингазов Н. Р.	т. 2: 22
Иомдин Б. Л.	т. 1: 214	Мисюрев А.	т. 1: 548
Иомдин Л. Л.	т. 1: 61	Музычка С.	т. 1: 468
Искра Е. В.	т. 1: 41	Мустайоки А.	т. 1: 453
Карпов А. А.	т. 1: 240	Мустакимова Э. Г.	т. 1: 388
Калинина М. В.	т. 2: 44	Мухин М. Ю.	т. 1: 254

Николаева Ю. В.	т. 1: 487	Сизов В. Г.	т. 1: 61
Нин Тао	т. 1: 202	Смирнов И. В.	т. 1: 560
Носырев Г. В.	т. 2: 145	Тарасов Д. С.	т. 1: 595, т. 2: 53
Падучева Е. В.	т. 1: 500	Татевосов С. Г.	т. 1: 272
Паничева П. В.	т. 2: 1	Тимошенко С. П.	т. 1: 61
Панченко А. И.	т. 2: 88, 105	Турдаков Д.	т. 2: 34
Паперно Д.	т. 2: 88	Тутубалина Е. В.	т. 2: 2, 22, 65
Пасюков А. В.	т. 1: 169	Урысон Е. В.	т. 1: 603
Плешко В. В.	т. 2: 44	Усталов Д. А.	т. 1: 616, т. 2: 88
Пионтковская И.	т. 1: 369, 468	Федорова О. В.	т. 1: 131, 487
Пиперски А. Ч.	т. 1: 515	Хесед Л. А.	т. 1: 321
Подлеская В. И.	т. 1: 523	Худякова М. В.	т. 1: 41
Поляков П. Ю.	т. 2: 44	Циммерлинг А. В.	т. 1: 696
Попов А. М.	т. 2: 1	Шелманов А. О.	т. 1: 560
Протопопова Е.	т. 1: 548	Шмелев А. Д.	т. 1: 584
Романов П. В.	т. 2: 105	Экхофф Х.	т. 1: 30
Рубцова Ю. В.	т. 2: 2	Яковлева И. В.	т. 1: 638
Селегей В. П.	т. 1: 169	Янко Т. Е.	т. 1: 650

Author's Index

- Adaskina Yu. V. v. 2: 1
Akinina Yu. S. v. 1: 41
Alimova I. S. v. 2: 22
Andreev I. v. 2: 133
Andrianov I. v. 2: 34
Antonova A. v. 1: 548
Apresjan V. Ju. v. 1: 2
Arefyev N. V. v. 2: 106
Astrakhantsev N. v. 2: 34
Avanesov V. v. 2: 34
Baranov A. N. v. 1: 19
Berdičevskis A. v. 1: 30
Bergelson M. B. v. 1: 41
Blinov P. D. v. 2: 3, 12
Bogdanov A. V. v. 1: 52
Boguslavsky I. M. v. 1: 62
Bonch-Osmolovskaya A. A. v. 1: 80
Braslavski P. I. v. 1: 254
Daniel M. A. v. 1: 95
Denisenko A. A. v. 2: 76
Dikonov V. G. v. 1: 62
Dobrenko N. V. v. 1: 202
Dobvol'enskij D. O. v. 1: 104
Dobrushina N. R. v. 1: 118
Dragoy O. V. v. 1: 41
Eckhoff H. v. 1: 30
Fedorova O. V. v. 1: 131, 488
Galitsky B. A. v. 1: 141
Galyashina E. I. v. 1: 156
Garashchuk R. V. v. 1: 169
Goncharova M. B. v. 1: 169
Gorbunova I. M. v. 1: 52
Grishina E. A. v. 1: 183
Grozin V. A. v. 1: 202
Gusarova N. F. v. 1: 202
Iomdin B. L. v. 1: 214
Iomdin L. L. v. 1: 62
Iskra E. V. v. 1: 41
Ivanov V. V. v. 2: 3, 22, 65
Kalinina M. V. v. 2: 44
Karpov A. A. v. 1: 241
Kashkin E. V. v. 1: 427
Katinskaya A. Y. v. 1: 398
Khesed L. A. v. 1: 321
Khudyakova M. V. v. 1: 41
Kibrik A. A. v. 1: 231, 488
Kipyatkova I. S. v. 1: 241
Kiselev Y. A. v. 1: 254
Kisseleva X. L. v. 1: 272
Klyachko E. v. 2: 119, 159
Knyazev S. V. v. 1: 284
Konstantinova N. v. 2: 89
Korotaev N. A. v. 1: 294
Kotelnikov E. V. v. 2: 3, 12
Kotov A. A. v. 1: 308
Kozlov I. v. 2: 34
Kozlova E. A. v. 1: 169
Kreydlin G. E. v. 1: 321
Krivnova O. F. v. 1: 338
Krizhanovskaya N. B. v. 1: 254
Krizhanovsky A. A. v. 1: 254
Krylova T. V. v. 1: 352
Kudinov M. v. 1: 369
Kustova G. I. v. 1: 376
Kutuzov A. v. 2: 133
Kuzmenko E. A. v. 1: 388
Lagutin M. B. v. 1: 398
Lazursky A. V. v. 1: 62
Levontina I. B. v. 1: 104
Lesota O. O. v. 2: 106
Lobanov B. M. v. 1: 414
Lopukhina A. A. v. 2: 145
Lopukhin K. A. v. 2: 145
Loukachevitch N. V. v. 2: 3
Loukachevitch N. V. v. 2: 89
Lukanin A. V. v. 2: 106
Lyashevskaya O. N. v. 1: 427
Lyutikova E. A. v. 1: 696
Malafeev A. Yu. v. 1: 441
Malykh V. A. v. 2: 65
Mayorov V. v. 2: 34
Menshikov I. L. v. 1: 254
Meyer C. M. v. 2: 89
Mingazov N. R. v. 2: 22
Misyurev A. v. 1: 548
Mukhin M. Yu. v. 1: 254

Mustajoki A.	v. 1: 453	Sheremetyeva S. O.	v. 1: 573
Mustakimova E. G.	v. 1: 388	Shmelev A. D.	v. 1: 584
Muzychka S.	v. 1: 468	Sizov V. G.	v. 1: 62
Nedoluzhko A.	v. 1: 474	Smirnov I. V.	v. 1: 560
Nikolaeva Y. V.	v. 1: 488	Solovyev D. A.	v. 2: 76
Ning Tao	v. 1: 202	Sorokin A. A.	v. 1: 398
Novák M.	v. 1: 474	Tarasov D. S.	v. 1: 595, v. 2: 53
Nosyrev G. V.	v. 2: 145	Tatevosov S. G.	v. 1: 272
Paducheva E. V.	v. 1: 500	Timoshenko S. P.	v. 1: 62
Panchenko A. I.	v. 2: 89, 106	Toldova S.	v. 1: 474
Panicheva P. V.	v. 2: 1	Turdakov D.	v. 2: 34
Paperno D.	v. 2: 89	Tutubalina E. V.	v. 2: 3, 22, 65
Pasyukov A. V.	v. 1: 169	Uryson E. V.	v. 1: 603
Piontkovskaya I.	v. 1: 369, 468	Ustalov D. A.	v. 1: 616, v. 2: 89
Piperski A. Ch.	v. 1: 515	Vasilyev V. G.	v. 2: 76
Pleshko V. V.	v. 2: 44	Veprva I. T.	v. 1: 453
Podlesskaya V. I.	v. 1: 523	Vilinbakhova E. L.	v. 1: 626
Polyakov P. Yu.	v. 2: 44	Vishneva E. A.	v. 1: 560
Ponomarev S. V.	v. 1: 536	Yakovleva I. V.	v. 1: 639
Popov A. M.	v. 2: 1	Yanko T. E.	v. 1: 651
Protopopova E.	v. 1: 548	Zagulova M. A.	v. 2: 65
Romanov P. V.	v. 2: 106	Zakharov V. P.	v. 1: 667
Rubtsova Y. V.	v. 2: 3	Zalizniak Anna A.	v. 1: 683
Selegey V. P.	v. 1: 169, 398	Zimmerling A. V.	v. 1: 696
Sharoff S.	v. 1: 398	Zinina A. A.	v. 1: 308
Shelmanov A. O.	v. 1: 560		

Научное издание

Компьютерная лингвистика и интеллектуальные технологии

По материалам ежегодной
Международной конференции «Диалог»

Выпуск 14 (21). 2015

Том 1. Основная программа конференции

Ответственный за выпуск **А. А. Белкина**
Вёрстка **К. А. Климентовский**

Подписано в печать 08.05.2015
Формат 152 × 235
Бумага офсетная
Тираж 250 экз. Заказ № 52

Издательский центр «Российский
государственный гуманитарный университет»
125993, Москва, Миусская пл., д. 6
Тел.: +7 499 973 42 06

Отпечатано с готового оригинал-макета в типографии
ООО «Издательско-полиграфический центр Маска»
117246, Москва, Научный пр-д, д. 20, стр. 9