

ЭКСПЕРТНЫЙ МЕТОД ДИАГНОСТИКИ БИОЛОГИЧЕСКИХ ПАРАМЕТРОВ ДИКТОРА ПО ГОЛОСУ THE EXPERT METHOD OF THE DIAGNOSTICS OF SPEAKER'S BIOLOGICAL PARAMETERS BY SPEECH

*Коваль С.Л. (koval@speechpro.com), Панова Е.А. (panova@speechpro.com)
ООО «Центр речевых технологий», Санкт-Петербург*

В докладе представлена методика экспертной диагностики биологических параметров (размеров тела) диктора по голосу. Большое внимание в данном случае уделяется требованиям к сбору представительных речевых баз данных. В ходе диагностики эксперт использует оптимально подобранные речевые эталоны, иллюстрирующие проявление используемых аудитивных характеристик. Результаты определения экспертами биологических параметров диктора на базе данных из 289 дикторов совпадают с реальными данными измерений с точностью, приемлемой для ряда практических приложений.

Введение

Проблема опознания размеров тела человека по голосу имеет глубокую историю. Утверждения о том, что речь человека является источником суждений относительно его внешности содержатся в трудах Аристотеля, Гипократа, Дарвина, Вундта.

Современное научное изучение проблемы начинается в 60-70 годы прошлого века. Наиболее известными являются работы таких исследователей, как Г. Олпорт, Н. Ласс, Р. Ферман, К. Шерер, а также отечественных авторов: Витт Н.В. [3], Галунов В.И., Манеров В.Х. [4,5], Морозов В.П. [6,7,8], Носенко Э.Л., Рамишвили Г.С. [9].

В 70-90е годы глубокое изучение процесса идентификации и интерпретации человека по голосу и речи проводилось группой исследователей на базе НИИ «Дальняя связь». Здесь же осуществлялась координация работ по этой проблеме в исследовательских лабораториях МГУ и ЛГУ. Полученные результаты обобщены в докторской диссертации В.Х. Манерова [4,5].

Естественно-научные предпосылки работ, посвященных изучению связей «говорящий-речь», основаны на том, что речевой аппарат является и частью человеческого тела, и исполнительным органом психической деятельности.

Первая предпосылка основана на антропологическом законе позитивной связи длины тела и размеров внутренних органов. В среднем, при большем росте должны быть большие размеры артикуляторов и фонаторов, то есть для высоких дикторов следует ожидать более низкочастотных характеристик речи. Вторая предпосылка основана на данных возрастной физиологии. С возрастом в речевом аппарате происходит ряд изменений: снижается жизненная емкость легких, уменьшается эластичность мышц. В итоге старческий голос воспринимается как тихий, слабый, нечеткий.

Процесс оценивания человека по голосу и речи представляет собой сложное, многоуровневое образование, на первом этапе которого происходит идентификация опорных свойств: пол, возраст и менее адекватно, но выше случайного уровня определяются вес, рост и окружность шеи.

При изучении возможности точного оценивания антропометрических параметров говорящего было обнаружено, что средняя аудиторская ошибка для роста равна 5-7 см., для веса – 5-10 кг. При этом оценки и фактические значения параметров были слабо, но значимо коррелированы. В работе Викторова А.В. [1,2] показано, что семантика речи не влияла на точность оценок.

Метод диагностики

На основе обзора научной и методической литературы по тематике установления биологических параметров диктора по голосу была разработана следующая методология экспертной диагностики биологических свойств диктора по его речи:

1. Исследование спорной фонограммы проводится опытным экспертом, как показывает практика, более надежно – комиссией (бригадой) экспертов из 2-3 человек. Для принятия решения комиссией экспертов предла-

гается особая процедура консенсуса и компромиссов.

2. Эксперты подвергаются профессиональному отбору по психологическим и слуховым характеристикам (необходим нормативный слух, устойчивая психика, предрасположенность к абстрактному, творческому мышлению и восприятию информации в целом («правополушарность»)). Отбор оптимальных экспертов производится на основе результатов специализированных тестов.

3. Эксперты должны иметь базовые знания в области лингвистики, психологии, речеобразования, а также специальную подготовку (тренировку) в определении различных биологических параметров диктора по речи. Базовая тренировка определения биологических параметров диктора проводится путем прослушивания экспертами эталонных обучающих речевых примеров дикторов с известными свойствами и последующего тестового определения заданных биологических параметров неизвестных дикторов и сравнения полученных результатов тестирования с исходными данными. Тренировка продолжается вплоть до получения заданной надежности принятия решений о биологических параметрах дикторов для заданных тренировочных наборов речи разнообразных по свойствам дикторов.

4. Эксперты должны соблюдать определенный режим работы, не допуская чрезмерного утомления, не принимать решения в болезненном или не адекватном физическом или психологическом состоянии.

5. Сама процедура диагностики проходит следующим образом:

а. Эксперт многократно прослушивает речь неизвестного диктора с помощью специализированного звукового редактора

б. Эксперт описывает специфику биологических параметров диктора с помощью формального набора признаков. Основные размерности дикторской диагностики предлагаются эксперту в рамках отдельных оценочных шкал на экране специализированного АПК, ответы на которые эксперт выбирает из фиксированного набора возможных категорий оценки.

с. Эксперт сравнивает проявление тех или иных признаков облика диктора со звуковыми эталонами. Для основных вариантов проявления дикторских особенностей в АПК эксперту предоставляется возможность прослушивания набора звуковых примеров характерной речи для нескольких дикторов.

д. В случае сомнений эксперт просматривает определения и описания оцениваемых признаков облика диктора. Для всех типов категорий оценки в АПК эксперт имеет доступ к подробным текстовым описаниям специфики их проявления в речи и в текстах.

е. Эксперт составляет текст диагностического заключения для исследуемого диктора, интерпретируя в понятных не специалисту терминах результаты своего исследования.

В данной методике диагностики определяются следующие биологические параметры дикторов: пол, возраст, рост, вес и размер окружности шеи. Данные биологические параметры обнаруживают наиболее четкие корреляции с аудитивными характеристиками речи дикторов.

На основе слухового анализа образцов речи дикторов с различным набором биологических параметров, были выделены следующие аудитивные характеристики речи дикторов, значимые при диагностике биологических параметров:

- быстрый – размеренный

Аудитивное впечатление «быстрого» – «размеренного» голоса коррелирует с темпом произнесения, что характеризуется количеством произнесенных звуков за единицу времени (норма составляет 9-14 звуков за секунду), и длиной пауз.

- сильный – слабый

Аудитивное впечатление «сильного» – «слабого» голоса зависит от силы голоса, то есть типичной для диктора интенсивности выдыхаемой воздушной струи, силы смыкания голосовых складок и напряженности голосовых мышц.

- громкий – тихий

Аудитивное впечатление «громкий» – «тихий» голоса зависит от его интенсивности. Если сила голоса – величина объективная, то громкость – понятие субъективное, связанное с восприятием звука. Громкость – это управляемое качество голоса. Она изменяется в зависимости от различных обстоятельств общения. Слабый голос в определенных ситуациях может быть громким.

- глухой – звонкий

Аудитивное впечатление «глухой» – «звонкий» голос зависит от направленности голоса. «Глухим» считают голос, высокие гармоники которого быстро затухают. В противоположность к «глухому» голосу отличают «звонкий», то есть голос с большим количеством высокочастотных гармоник.

- хриплый – нехриплый

Аудитивное впечатление «хриплый» – «нехриплый» голос зависит от характера смыкания голосовых складок. Хриплый тембр голоса появляется при неполном, неплотном смыкании голосовых складок.

- шумовой – тональный

Аудитивное впечатление «шумовой» – «тональный» голос зависит от соотношения шума и тона в спектре звука голоса. При преобладании шума голос расценивается как шумовой, в обратном случае – тональный.

- монотонный – многозвучный

Аудитивное впечатление «монотонный» – «многозвучный» голос зависит от вариации ЧОТ в потоке речи. Монотонный голос характеризуется как голос с малым изменением диапазона ЧОТ. В противоположность к нему, многозвучный голос отличается вариативностью и варианностью изменения ЧОТ.

- темный – светлый

Аудитивное впечатление «темный» – «светлый» голос зависит от общей окраски голоса – светлой или темной, что определяется наличием гармоник в общем звуковом спектре голоса выше или ниже 1500 Гц.

- тусклый – яркий

Аудитивное впечатление «тусклый» – «яркий» голос зависит от состояния голосового аппарата. С возрастом голосовые складки становятся менее эластичными, тонкими, в части случаев они смыкаются не полностью. Вследствие возникающей у некоторых людей слабости голосовых складок и малой подвижности голосовых хрящей гортани тембр голоса становится тусклым, бесцветным.

- жидкий – густой

Аудитивное впечатление «жидкий» – «густой» голос зависит от наличия дополнительных гармоник. Густой голос характеризуется большим количеством дополнительных, хорошо выделяющихся над шумовым фоном гармоник, жидкий голос, наоборот, отличается их бедностью.

- бодрый – вялый

Аудитивное впечатление «бодрый» – «вялый» голос зависит от силы и общего тонуса мышц речеобразующего аппарата диктора.

- сочный – сухой

Аудитивное впечатление «сочный» – «сухой» голос зависит от состояния голосового аппарата. С возрастом голосовые складки теряют свою эластичность, в их составе появляется много соединительной ткани, голос становится сухим, ломающимся, дребезжащим. Развивающаяся атрофия слизистых желез вестибулярного отдела приводит к сухости слизистой оболочки гортани и снижению ее защитной способности.

- гладкий – шершавый

Аудитивное впечатление «гладкий» – «шершавый» голос зависит от качества работы голосовых складок и органов произносительного аппарата. Голос воспринимается как «гладкий», если голосовые связки эластичны, плотно смыкаются, воздушная струя сфокусирована на губах. Голос воспринимается как «шершавый» при огрубении голосовых складок, их нестандартном или нестабильном смыкании, воздушная струя как бы застревает в полости рта.

- высокий – низкий

Аудитивная характеристика «высокий» – «низкий» голос взаимосвязана с высотой голоса, которая определяется частотой колебаний голосовых складок и типичным частотным положением низкочастотных резонансов вокального тракта. Частота колебания голосовых складок, в свою очередь, находится в зависимости от их длины, толщины и напряженности.

- вариативный – стабильный

Аудитивная характеристика «вариативный» – «стабильный» голос зависит от разнообразия используемых произносительных стереотипов и их повторяемости. Речь дикторов-женщин ассоциируется с впечатлением большего разнообразия, вариативности, чем у дикторов-мужчин, за счёт изрезанности мелодического контура. Женщины чаще используют интонационные средства для выражения значений, в то время как мужчины в этих же речевых ситуациях обычно прибегают к средствам лексики и грамматики.

- нежный – грубый

Голоса дикторов-женщин характеризуются более высокой ЧОТ, большей плавностью переходов мелодики и силы голоса, что составляет аудитивное впечатление «нежного голоса». Голоса дикторов-мужчин, напротив, имеют более низкую ЧОТ, более резкие перепады мелодического контура и силы голоса, что создает впечатление «грубого голоса». К данной паре признаков тесно примыкают пары признаков женственный/мужественный и мужеподобный/женоподобный. Женственный голос можно охарактеризовать как голос диктора-женщины, в котором со всей полнотой проявляются такие признаки, как нежный, мягкий, плавный голос. Если данные признаки, характерные для женских голосов, проявляются в голосе диктора-мужчины, то голос такого диктора можно назвать женоподобным (данный признак имеет отрицательную коннотацию). К мужественному голосу можно отнести как голос дикторов-мужчин, так и голоса дикторов-женщин. Такому голосу присущи сила, устойчивость мелодических контуров. Голос дикторов-женщин, который создает аудитивное впечатление «грубого голоса», можно назвать мужеподобным (данный признак имеет отрицательную коннотацию).

- глубокий – поверхностный

Аудитивное впечатление «глубокий» – «поверхностный» голоса связано с длиной голосового тракта диктора. Чем больше рост диктора, тем больше у него длина голосового тракта.

- мягкий – жесткий

Аудитивное впечатление «мягкого» голоса характеризуется общей плавностью артикуляции, плавностью мелодических переходов речи, относительно небольшим диапазоном варьирования ЧОТ.

- тонкий – толстый

«Толстый» голос характеризуется большим количеством дополнительных гармоник и большими размерами произносительного аппарата, тонкий голос характеризуется меньшим количеством дополнительных гармоник и меньшими размерами произносительного аппарата.

- четкий – расплывчатый

Аудитивное впечатление «четкий» – «расплывчатый» голос зависит от характера артикуляции. Голос расценивается как «четкий» при четком характере артикуляции, у «расплывчатого» голоса отмечается некая смазанность артикуляции.

- шумное – неслышное речевое дыхание

По данным аудитивного исследования у дикторов весом до 80 кг дыхание бесшумно, у дикторов весом от 80 до 100 кг дыхание расценивается как среднее, либо средне-шумное. В голосе дикторов-мужчин от 100 кг, либо дикторов-женщин от 90 присутствует шумное речевое дыхание, у некоторых дикторов наблюдается одышка.

Для каждой пары признаков была введена пятибалльная шкала оценки. То есть, например, для признаков «высокий» – «низкий» голос введена следующая шкала: «высокий», «средне-высокий», «средний», «средне-низкий» и «низкий» голос. Для каждого элемента шкалы были подобраны речевые эталоны, на основе которых проводится дальнейшая диагностика дикторов.

Используемый речевой материал

Создание автоматизированной системы оценки биологических параметров неизвестного диктора по речи требует разработки специализированных РБД. При создании речевых баз данных учитываются следующие цели:

- сбор представительной информационной БД, включающей максимальное количество дикторов с, по возможности, максимально разнообразными биологическими параметрами, потенциально определяемыми по голосу и речи;

- создание на основе созданных БД эталонов (образцов) речи дикторов, которые отражают различные типичные комбинации биологических параметров дикторов. Данные эталоны доступны пользователям разрабатываемого АПК и используются как для тренировки и тестирования экспертов, так и для прямого аудитивного и/или инструментального сравнения с эталонами спорного речевого материала при диагностике биологических параметров дикторов по фонограммам их речи.

Разработка представительной базы данных включает следующие аспекты:

1. выделение основных биологических параметров диктора, которые предположительно определяются по фонограммам речи диктора
2. отбор дикторов для записи представительной РБД
3. выработка требований и обеспечение технических условий для записи представительной РБД, отвечающим всем необходимым требованиям
4. составление речевого материала для записи представительной РБД
5. составление системы анкет, тестов и процедур антропометрических измерений для получения объективной информации об обликовых дикторских свойствах
6. составление инструкций оператору сбора баз данных и диктору
7. запись речевого материала, анкетирование, антропометрическое тестирование дикторов
8. обработка записанных фонограмм для РБД, создание файловой структуры РБД, которая будет отвечать требованиям автоматизированной системы оценки биологических параметров неизвестного диктора по речи
9. обработка результатов тестирования, анкетирования и измерения дикторов, перевод их в формат удобный для последующего использования

На основании анализа современных научных исследований по данной тематике можно сделать вывод, что для сбора минимально репрезентативного для данной задачи речевого корпуса необходимо записать не менее 40-50 дикторов. Речевой корпус средних размеров включает в себя около 150-200 дикторов, более расширенная база данных должна включать 300-400 дикторов и выше. В РБД должны быть, по возможности, максимально представлены разные возрастные, половые, весовые и ростовые категории, возможно включение дополнительных замеров окружности шеи.

При сборе представительной РБД для определения биологических свойств диктора по речи достаточно сложно найти дикторов, имеющих нестандартное сочетание антропометрических параметров (например, большой вес тела при очень маленьком росте, окружность головы меньше нормы при высоком росте). В таком случае применяется следующая стратегия: на этапе сбора основной РБД производится запись дикторов со стандартными параметрами (которые представлены большинством людей), а на последующих этапах, после анализа собранной информации и реальной оценки ее представительности и достаточности для разработки методов обличевой диагностики проводится целенаправленный поиск и запись информантов с нестандартным сочетанием антропометрических параметров.

Каждый диктор перед записью в РБД заполняет анкету, в которой указывает свои антропометрические данные по возрасту, росту, весу, размеру окружности шеи. Если диктор не может точно указать какие-либо антропометрические данные, оператор по сбору РБД производит антропометрическое тестирование (замеры) информантов.

Условия записи РБД для определения основных биологических параметров должны удовлетворять следующим требованиям: запись осуществляется в тихом помещении, на высококачественный микрофон и соответствующее оборудование для записи. Материал для записи РБД для определения основных анатомо-физиологических параметров должен содержать различные типы произнесения: спонтанную речь, чтение, монолог и диалог с оператором по сбору РБД. Однако на основе исследований литературы по данной тематике, можно сделать вывод, что на диагностику биологических параметров диктора не оказывает влияние как тип речи диктора (спонтанная речь, чтение), так и лингвистический уровень материала (изолированные слова, словосочетания, текст). Составленная РБД на данный момент включает в себя следующие диапазоны значений биологических параметров:

- В РБД представлено 168 дикторов-мужчин и 121 диктор-женщина.
- Возраст дикторов на момент записи в РБД составил от 18 до 84 лет.
- Рост дикторов колебался в пределах от 135 до 208 см.
- Вес дикторов колебался в диапазоне от 40 до 163 кг.
- Размер окружности шеи дикторов составил от 30 до 48 см.

Результаты

После составления эталонов на основе РБД и тренировки группы четырех экспертов в течение 2 месяцев были проведены тестовые испытания методики. В целях тестирования предложенной методики диагностики биологических параметров дикторов было выбрано 20 дикторов: 10 дикторов-мужчин и 10 дикторов-женщин по 2 подхода записей, биологические параметры которых были продиагностированы комиссией четырех экспертов. В ходе данной диагностики были получены следующие результаты: возраст дикторов был определен в среднем с точностью до 5 лет, рост дикторов был определен в среднем с точностью до 6 см, вес дикторов был определен с точностью до 6 кг, окружность шеи дикторов была определена с точностью до 2 см.

Заключение

Субъективно и объективно биологические параметры диктора определяются по его речи с точностью, намного превышающей случайное угадывание. Однако экспертная диагностика биологических параметров дикторов, имеющих нестандартные параметры конституции, затруднена и требует специальной подготовки экспертов.

Список литературы

1. Викторов А.Б., Остроухов А.В., Лобанова М.А. О возможности создания автоматизированного комплекса диагностирования обличевых признаков. // Информатизация и информационная безопасность правоохранительных органов. Сб. трудов ХУ международной научной конференции. 23-24 мая 2006г. М.: Академия управления МВД России. 2006. Стр. 346-351.
2. Викторов А.Б., Остроухов А.В., Лобанова М.А. О возможности создания автоматизированного комплекса диагностирования обличевых признаков// Информатизация и информационная безопасность правоохранительных органов: Тез. докл. Межд. Науч. Конф. М.: Акад. Упр. МВД. России, 2006
3. Витт Н.В. Эмоциональная регуляция речевого поведения при общении. М.: Наука. 1983.
4. Манеров В.Х. Проблематика и перспективы решения задачи диагностики свойств говорящего по речи. – В сб. Возможности судебной видеофоноскопической экспертизы. ВНИИ судебных экспертиз. М. 1989. С 100-108

5. Манеров В.Х. Экспериментально-теоретические основы социальной идентификации и интерпретации говорящего. Автореф. док. дисс.СПб., 1993
6. Морозов В.П. Биофизические основы вокальной речи Л.: Наука. 1977
7. Морозов В.П. Занимательная биоакустика. М.: Знание. 1987
8. Морозов В.П. Невербальная коммуникация: экспериментально-теоретический и прикладной аспекты. Психологический журнал. Т.14, №1,1993, С. 18-31
9. Рамишвили Г.С. Автоматическое опознавание говорящего по голосу. М.: Радио и связь. 1981