

УДК 621.391: 004.522

DOI: 10.18413/2518-1092-2020-5-2-0-2

Васильев Р.А.

**ПРИМЕНЕНИЕ МЕТОДОВ ФОНЕТИЧЕСКОГО АНАЛИЗА РЕЧИ
ДЛЯ ВЫЯВЛЕНИЯ ЭМОЦИОНАЛЬНО УСТОЙЧИВЫХ
И НЕСТАБИЛЬНЫХ СТУДЕНТОВ УНИВЕРСИТЕТА**Нижегородский государственный университет им. Н.И. Лобачевского
пр. Гагарина, д. 23, г. Нижний Новгород, 603950, Россия*e-mail: romangamma@mail.ru***Аннотация**

Данная статья посвящена разработке компьютерной системы, предназначенной для исследования эмоциональной устойчивости человека по речевому сигналу в нормальных и условиях с повышенной напряжённостью на основе методов фонетического анализа речи и критерия минимума требуемой избыточности голосового сигнала.

Основная деятельность высшего учебного заведения – образовательный процесс. Для организации учебного процесса необходимо соединить все его элементы, наладить их взаимодействие между собой, определить содержание деятельности преподавателей и студентов. Комфортная и благополучная психологическая атмосфера на занятиях в вузе, несомненно, способствует успешности обучения студентов. Одна из основных задач преподавателя современной высшей школы не только делиться научной информацией с обучающимися, но и создание психологического комфорта в процессе обучения. Исследования эмоционального состояния обучающихся на лекциях и экзаменах является актуальной задачей.

В связи с этим разработана и протестирована специальная компьютерная система «Информационная система идентификации дикторов по голосу», способная автоматизировать процесс исследования эмоционального состояния студентов по голосу в комфортных и некомфортных условиях, для выявления эмоционально устойчивых и нестабильных.

Ключевые слова: речевые технологии; информационная теория качества речи; критерий минимума требуемой избыточности; система идентификации по голосу; психолингвистика.

UDC 621.391: 004.522

Vasiliev R.A.

**APPLICATION OF METHODS OF PHONETIC ANALYSIS
OF SPEECH FOR IDENTIFICATION OF EMOTIONALLY
SUSTAINABLE AND UNSTABLE STUDENTS OF UNIVERSITY**

Nizhny Novgorod State University N.I. Lobachevsky, Gagarina Ave., 23, Nizhny Novgorod, 603950, Russia

*e-mail: romangamma@mail.ru***Abstract**

This article is devoted to the development of a computer system designed to study the emotional stability of a person according to a speech signal in normal and conditions with increased tension based on the methods of phonetic analysis of speech and the criterion for the minimum required voice signal redundancy.

The main activity of a higher educational institution is the educational process. To organize the educational process, it is necessary to combine all its elements, to establish their interaction with each other, to determine the content of the activities of teachers and students. A comfortable and prosperous psychological atmosphere in the classroom of the university, undoubtedly, contributes to the success of student learning. One of the main tasks of a teacher at a modern higher school is not only to share scientific information with students, but also to create psychological comfort in

the learning process. Studying the emotional state of students in lectures and exams is an urgent task.

In connection with this, a special computer system called “Voice Announcer Information System for Identifying Voice” was developed and tested, which is able to automate the process of studying the emotional state of students by voice in comfortable and uncomfortable conditions to identify emotionally stable and unstable ones.

Keywords: speech technologies; information theory of speech quality; minimum required redundancy criterion; voice identification system; psycholinguistics.

ВВЕДЕНИЕ

Многим известно, что в группе студентов всегда есть лидеры и аутсайдеры. Понять, какова позиция студента университета, насколько она благоприятна для его развития, можно с помощью перспективных методик [7, 15, 16]. Одной из лидирующих является методика выявления эмоционально устойчивых и нестабильных студентов на основе фонетического анализа речи и идентификации по голосу.

Информационная система идентификации дикторов по голосу («ИС ИДГ») – это программа для ЭВМ, зарегистрированная в Роспатенте [1], предназначенная для тестирования эмоционального состояния и идентификации личности по голосу. «ИС ИДГ» обладает широким спектром возможностей: высокой чувствительностью к изменениям в эмоциональном состоянии личности при минимальных требованиях к продолжительности исследуемого фрагмента голосового сигнала [3]. Данные возможности достигнуты с применением нового принципа действия «ИС ИДГ», основанного на оценке качества речи личности на фонетическом уровне по критерию минимума требуемой избыточности (МТИ) речевого сигнала [4]. По сути, решается проблема многокритериальности устной речи с позиций информационной теорией качества речи (ИТКР) [2]. Поэтому публикуемые далее результаты теоретического и экспериментального исследования «ИС ИДГ» представляют интерес для специалистов в области биометрических исследований по голосу.

ТЕОРЕТИЧЕСКИЙ АНАЛИЗ

Большинство современных систем автоматического анализа речи функционируют посредством последовательного деления голосового сигнала на короткие (5-10 мс) отрезки данных $x = (x_1, x_2, \dots, x_n)$ длиной в одну минимальную речевую единицу (МРЕ) с их последующим сравнением с эталоном. Сложной проблемой для таких систем является создание и обоснование множества фонетических эталонов $\{x_r^*\}$ [2].

Проблематика выбора минимальной речевой единицы и сегодня остается актуальной. Специалисты уже давно сделали вывод о том, что элементарные звуки, из которых состоит речь, не эквивалентны буквам. В связи с чем ввели понятие фонемы для обозначения элементарных звуков речи, этой теме посвящено множество исследований [17, 18, 19, 20, 21]. Хотя и сейчас специалисты не могут решить – сколько же всего различных фонем существует. Фонема – это основная единица звукового строя языка, выделяемый линейным членением речи [8]. Она не является простейшим элементом, т. к. состоит из фреймов (реализаций), функционирующих одновременно [9]. В лингвистике фонема определяется, как минимальная речевая единица, предназначенная для различения содержания слов и определяющаяся в зависимости от местоположения – в разных своих вариантах. Речевой сигнал можно разбить на МРЕ имеющие различные реализации с помощью устойчивых параметров и объединяемые в группы одноименных речевых единиц [10]. Это показывает различия в произнесении пользователем одноименных МРЕ и особенности восприятия звуков речи. При анализе фонетического состава речи и статистических характеристик МРЕ, их суммарное число R зависит от особенностей голоса каждой конкретной личности [2].

Любой человек в силу ряда причин, например, из-за особенностей своей речи или слуха, не в состоянии в процессе произнесения звуков точно воспроизвести эталон \mathbf{x}_r^* той или иной (r -й) МРЕ. Решением данной проблемы может служить задание каждой МРЕ не одним, а одновременно несколькими допустимыми вариантами $\mathbf{x}_{r,j}$, $j = \overline{1, J_r}$, где $r = \overline{1, R}$, а R – объем фонетической базы данных (ФБД) [3]. В данном случае исследуемой личности будет достаточно приблизить свое произношение к любому из них, чтобы быть правильно понятым экспертом или слушателем. Этим решается рассматриваемая проблема вариативности устной речи: каждый конкретный человек в процессе своего «произношения звуков» выбирает наиболее удобный, достижимый для себя вариант эталонного произношения МРЕ из некоторого множества альтернатив $\left\{ \mathbf{x}_{r,j} \right\}$ [4]. Становятся понятными и параметры формируемого (на выходе голосового тракта человека) речевого сигнала к эталону: он должен войти в границы J_r -множества вариантов рассматриваемой МРЕ X_r как полноправный, $(J_r + 1)$ -й его элемент. В данном случае задача переходит в сугубо предметную плоскость: сначала по каждой из R рассматриваемых МРЕ требуется кластер $X_r = \left\{ \mathbf{x}_{r,j} \right\}$ ее допустимых образцов – на этапе обучения диктора [5].

В соответствии с ИТКР, каждый человек выступает в роли условного источника дискретных сообщений $X \subset \left\{ X_r \right\}$, определенных на R -множестве его МРЕ с ФБД $\left\{ \mathbf{x}_r^* \right\}$. Характеристикой коммуникативных свойств такого источника может служить скорость создания информации, или количество информации на выходе в расчете на одну МРЕ. В предположении об идеальном речевом механизме человека-диктора и безошибочном восприятии всего набора его МРЕ потенциальным слушателем указанная величина определяется выражением для шенноновской энтропии дискретного источника сообщений вида

$$H(X) = - \sum_{r=1}^R P(X = \mathbf{x}_r^*) \log P(X = \mathbf{x}_r^*) = - \sum_{r=1}^R p_r \log p_r \quad [3].$$

Апостериорная энтропия источника сообщений $H(X)$ имеет в данном случае смысл величины рассеяния полезной информации в процессе создания речевого сигнала, или минимальной требуемой избыточности речевого сигнала. Чем больше рассеяние, тем выше степень искажений формируемого на выходе речевого тракта сигнала X по сравнению с его эталоном \mathbf{x}_r^* , и тем ниже, следовательно, качество речи данного диктора [3].

Как было сказано ранее, эталоном соответствующей фонемы диктора является \mathbf{x}_r^* , а процесс сравнение множества реализаций фонемы X диктора – способ выделения эталона фонемы \mathbf{x}_r^* конкретного диктора, или звукового ряда $\left\{ X_r \right\}$ (строка) его разговорного языка [5]. В информационной теории качества речи для сравнения реализаций фонем диктора используется величина – μ_0 минимум информационного рассогласования (МИР).

Заметим, что именно относительная величина часто является более выгодной по сравнению с абсолютной величиной теоретико-информационного показателя качества речи. Это актуально в задачах речевой идентификации в системах разграничения доступа [6], в нашем случае по принципу сопоставления двух относительных величин требуемой избыточности – $\hat{\alpha}_0$ (ОВТИ), полученных в

процессе тестирования диктора. Задача такого рода подробно рассмотрена далее – в качестве предмета экспериментальных исследований.

МЕТОДЫ ИССЛЕДОВАНИЯ

Для исследований в рамках ИТКР был сконфигурирован образец «ИС ИДГ», запрограммированный для выполнения операций над речевым сигналом x_r^* по выявлению эмоционального состояния человека (студента) [11, 12, 13, 14]. Главное окно «ИС ИДГ» показано на рисунке 1.

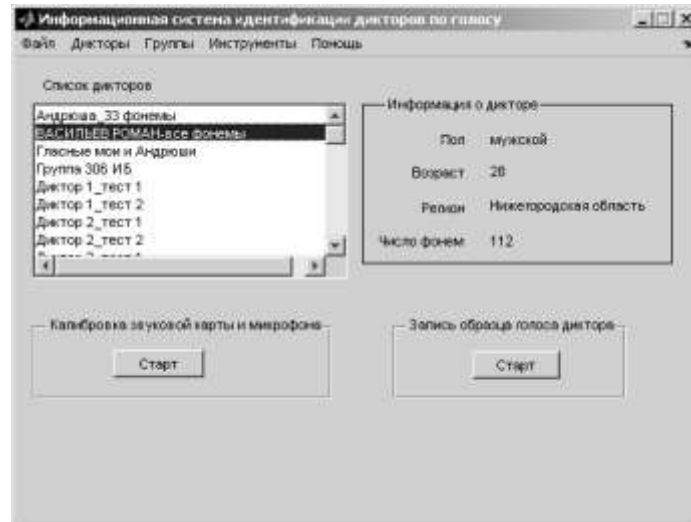


Рис. 1. Главное окно программы «ИС ИДГ»
Fig. 1. The main window of the program "IS IDG"

Испытания проводились в 2 этапа. На первом этапе были сформированы три группы обследуемых лиц в составе:

- десяти выпускников Нижегородского государственного университета им. Н.И. Лобачевского (ННГУ) кафедры «безопасность информационных систем» на заседании ГАК;

- трех студентов 2 курса радиофизического факультета ННГУ во время сдачи ими экзамена;
- трех преподавателей радиофизического факультета ННГУ – для сравнения;

Фамилии всех студентов и преподавателей в отчетных таблицах заменены.

Каждый студент (диктор) в режиме монолога отвечал последовательно на вопросы экзаменационного задания или выступил с докладом по своей дипломной работе. Им противопоставлялись устные тексты от группы преподавателей, полученные в заведомо комфортных условиях. Все выступления были записаны через систему звукозаписи в память ПК в виде соответствующих звуковых (WAV) файлов. После чего была создана база голосовых данных для проведения дальнейших исследований.

На втором этапе по каждому WAV-файлу из памяти первого ПК запускалась его система воспроизведения речи, и одновременно с ней оператор включалась в работу программа «ИС ИДГ» из памяти второго ПК. Сначала (на интервале 1-2 мин.) программа работала в режиме настройки на конкретного диктора.

В результате по каждому диктору была определена его номинальная величина избыточности речевого сигнала при фиксированном значении порога МИР, равного 0,05. Затем программу переводили в рабочий режим (рис. 2), в котором оценка качества речи производилась по временной диаграмме колебаний избыточности речевого сигнала на интервале в несколько минут.

Чем больше амплитуда колебаний, тем ниже качество речи данного диктора. Окно рабочего режима программы «ИС ИДГ» показано на рисунке 2.

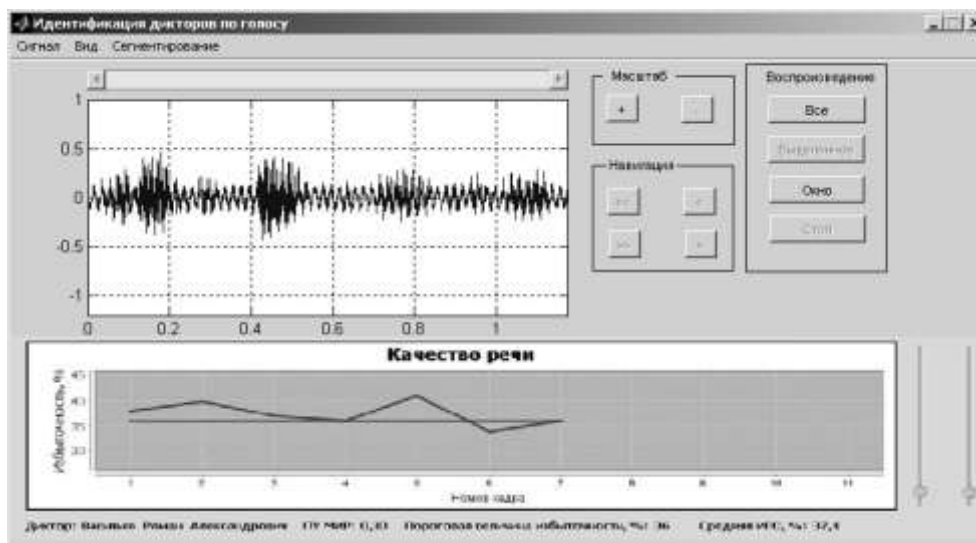


Рис. 2. Рабочий режим программы «ИС ИДГ»
Fig. 2. The operating mode of the program "IS IDH"

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Полученные показатели качества речи по каждому диктору сопоставлялись в дальнейшем с полученной им ранее экзаменационной оценкой или оценкой ГАК. Идея состояла в том, что хорошо подготовленные студенты излагают материал более качественно на фонетическом (звуковом) уровне по сравнению с плохо подготовленными студентами.

Таблица 1

Состав контрольной группы обследуемых лиц

Table 1

The composition of the control group of subjects

Дипломники	На экзамене	В комфортных условиях
1) Коваленко	2) Семенова	3) Иванов Д.Ю., проф.
4) Каштанов	5) Груничева	6) Петров И.В., аспирант.
7) Кузьмина	8) Евсеньков	9) Сидоров В.И., препод.
10) Лахтионова	11)	12)
13) Федорова	14)	15)
16) Борщовцевская	17)	18)
19) Васильев	20)	21)
22) Шумова	23)	24)
25) Николаев	26)	27)
28) Питеркин	29)	30)

Величина избыточности во всех случаях имеет характер колебаний – точно в соответствии с естественными колебаниями. Здесь красная прямая линия – это номинальная величина избыточности. А амплитуда колебаний – это главный показатель качества речи диктора.

По результатам проведенных испытаний в таблице ниже представлены сводные данные качества речи по всей контрольной группе обследованных лиц с указанием величины информационного рассогласования (ИРС). Здесь в правой колонке по каждому диктору дана оценка экзаменатора.

Таблица 2

Зависимость ИРС от экзаменационной оценки обследуемых

Table 2

The dependence of the IRS from the examination assessment of the subjects

Диктор	Номинал ИРС, %	Амплитуда колебаний ИРС, %	Экзаменационная оценка за ответ
1. Коваленко	73	16	4
2. Каштанов	85	5	3
3. Кузьмина	73	11	4
4. Лахтионова	65	6	5
5. Федорова	81	4	4
6. Борщовецкая	60	19	5
7. Васильев	50	18	5
8. Шумова	75	21	4
9. Николаев	80	13	4
10. Питеркин	74	17	5
11. Семенова	58	7	5
12. Груничева	68	17	3
13. Евсенюк	44	19	4
14. Петров И.В.	77	10	-
15. Иванов Д.Ю.	46	11	-
16. Сидоров В.И.	84	9	-

На рисунке 3 представлены оценки относительной величины информационного рассогласования для дикторов Семенова, Груничева, Евсенюк.

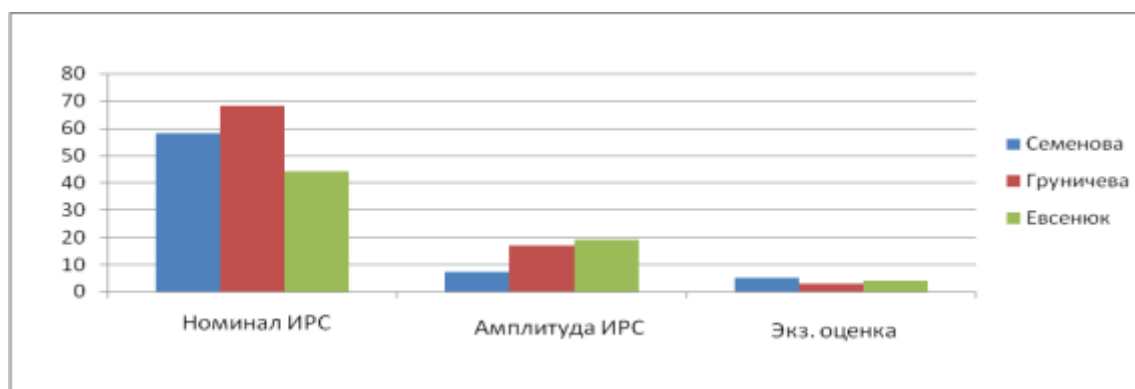


Рис. 3. Оценки относительной величины информационного рассогласования для дикторов Семенова, Груничева, Евсенюк

Fig. 3. Estimates of the relative magnitude of the information mismatch for the announcers Semenov, Grunichev, Yevsenyuk

Здесь хорошо видна тенденция к увеличению амплитуды колебаний избыточности речи относительно ее номинального значения при уменьшении качества предметной подготовки диктора (чем больше номинал и амплитуда колебаний величины информационного рассогласования, тем более нестабильна речь диктора).

ЗАКЛЮЧЕНИЕ

По результатам проведенных исследований сделаны следующие выводы:

- Подтверждена устойчивость информационного показателя качества речи: скорости создания информации на фонетическом уровне – относительно личности диктора на различных вариациях текстовой информации и вариативном временном диапазоне;

- Исследована зависимость величины ИРС от интенсивности эмоциональной нагрузки личности;
- Дано экспериментальное обоснование принципа стабильной индивидуальной избыточности речи диктора и скорости создания информации на выходе голосового тракта в комфортных условиях говорения или в отсутствие внешних раздражителей;
- Для дикторов в спокойном состоянии амплитуда колебаний ИРС составляет величину до 10% и практически не изменяется для разных значений порога;
- Показано, что изменение величины избыточности речи для диктора в состоянии стресса стабильно выше, чем для него же в спокойном состоянии; поэтому «ИС ИДГ» может быть использован как эффективное средство для оценки эмоционального состояния человека;
- У менее подготовленных дикторов разброс величины избыточности речи намного сильнее, чем для дикторов с хорошей подготовкой. Это доказывает, что менее подготовленные дикторы испытывали больший стресс при выступлении.

В общем случае даже относительно небольшое по накалу стрессовое состояние диктора оказывает весьма заметное влияние на его речь на фонетическом уровне в теоретико-информационном смысле, что позволяет с высокой точностью определять эмоционально неустойчивых и нестабильных студентов университета.

Список литературы

1. Свид. о гос. регистрации программы для ЭВМ №2015663306 Программа идентификации дикторов по голосу / Васильев Р.А. Зарег. 15.12.2015г. – М.: Роспатент, 2015.
2. Савченко В.В. Информационная теория качества речи // Изв. вузов России. Радиоэлектроника. 2011. Вып. 1. С. 17-27.
3. Савченко В.В., Васильев Р.А. Анализ эмоционального состояния дикторов по голосу на основе фонетического детектора лжи // Научные ведомости Белгородского государственного университета. 2014. Вып. № 21(192) 32/1. С. 186-195.
4. Савченко. В. В., Васильев Р. А. Автоматическая оценка качества речи по критерию минимума требуемой избыточности речевого сигнала // Материалы XI Международной научно-технической конференции посвященной памяти Б.И. Рамеева. Новые информационные технологии и системы. Пензенский государственный университет. 2014. С. 15-19.
5. Васильев Р.А. Биометрическая идентификация пользователей информационных систем на основе кластерной модели элементарных речевых единиц: дис. ... к-та тех. наук. М., 2017. 153 с.
6. Николаев Д.Б., Васильев Р.А. Анализ возможности применения голосовой идентификации в системах разграничения доступа к информации // Научный результат // Серия: Информационные технологии. Белгородский государственный университет. 2016. Вып. 1. С 48-57.
7. Алимуратов А. К., Тычков А. Ю., Чураков П. П. // Оценка психоэмоционального состояния человека на основе декомпозиции на эмпирические моды и кепстрального анализа речевых сигналов // Вестник Пензенского государственного университета. – 2018. – № 2 (22). – С. 89–95.
8. Герасимов А.В., Фидельман В.Р. Применение методов классического и модифицированного линейного предсказания для определения порядка линейной модели в задаче акустического кодирования речи // XXIV научные чтения имени академика Н.В. Белова. Тезисы докладов. Нижний Новгород. 2005. С. 142-144.
9. Кульбак С. Теория информации и статистика. М.: Наука, 1967. 408 с.
10. Савченко В.В., Акатьев Д.Ю. Автотестирование качества произношения речи по принципу минимального информационного рассогласования // Сборник научных трудов «Современные тенденции компьютеризации процесса изучения иностранных языков». Луганск: Восточно-украинский национальный университет. 2005. Вып.3. С.205-206.
11. Васильев Р.А. Исследование особенностей фонетического строя речи и текстонезависимая идентификация дикторов по непрерывной речи // Информационная безопасность регионов. 2012. № 2 (11). С. 57-63.
12. Васильев Р.А. Исследование фонетического строя речи и идентификация дикторов по голосу // Вопросы защиты информации. 2013. № 1 (100). С. 43-51.
13. Васильев Р.А. Исследование особенностей идентификации дикторов по голосу при различиях в произношении дикторов// Безопасность информационных технологий. 2013. № 1. С. 85-86.

14. Васильев Р.А. Исследование особенностей фонетического строя речи и определение национальности диктора при проведении процедуры идентификации по голосу // Информация и безопасность. 2012. Т. 15. № 4. С. 487-494.
15. Райгородский, Д.Я. Практическая психодиагностика / Д.Я. Райгородский. – Самара: Бахрах-М, 2011. – 672 с.
16. Сонин. В.А. Психодиагностическое познание профессиональной деятельности / В.А. Сонин. – СПб.: Речь, 2004. – 408 с.
17. Gray A., Markel J. Distance measures for speech processing. // IEEE Trans. On Acoust., Speech and Lang. processing. Vol. 24 (5), oct. 1976. p. 380 – 291.
18. Garofolo J., Auzanne G., and Voorhees E. The trec spoken document retrieval track: A success story. // In proceedings of the Recherche d'Informations Assiste par Ordinateur: Content Based Multimedia Information Access Conference, 2000. pp. 1-20.
19. Huijbregts M., Ordelman R., Jong F. Annotation of heterogeneous multimedia content using automatic speech recognition // In Proceedings of the second international conference on Semantics And digital Media Technologies (SAMT). Lecture Notes in Computer Science. Berlin. Springer Verlag. December 2007. pp. 78-90.
20. Mamou J., Mass Y., Ramabhadran B., Sznajder B. Combination of multiple speech transcription methods for vocabulary independent search // In proceedings of the ACM SIGIR Workshop `Searching Spontaneous Conversational Speech. Singapore. 2008. pp. 20-27.
21. Rabiner L. A. Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition // Proceedings of the IEEE. 1989. Vol. 77 № 2. pp. 257-285.

References

1. Testimonial. about state registration of a computer program No. 2015663306 Voice recognition speaker identification program / Vasiliev R.A. Zareg. 12/15/2015 – М.: Rospatent, 2015.
2. Savchenko V.V. Information theory of speech quality // Izv. Russian universities. Radio Electronics 2011. Issue. 1, pp. 17-27.
3. Savchenko V.V., Vasiliev R.A. Analysis of the emotional state of voice announcers based on a phonetic lie detector // Scientific Bulletin of Belgorod State University. 2014. No. 21 (192) 32/1. P. 186-195.
4. Savchenko. V.V., Vasiliev R. A. Automatic assessment of speech quality by the criterion of the minimum required redundancy of a speech signal // Materials of the XI International scientific and technical conference dedicated to the memory of B.I. Rameeva. New information technologies and systems. Penza State University. 2014. P. 15-19.
5. Vasiliev R.A. Biometric identification of users of information systems based on a cluster model of elementary speech units: Dis. ... to those. sciences. Moscow, 2017. 153 p.
6. Nikolaev D.B., Vasiliev R.A. Analysis of the possibility of using voice recognition in systems for differentiating access to information // Scientific Result // Series: Information Technologies. Belgorod State University. 2016. Issue. 1. P. 48-57.
7. Alimuradov A.K., Tychkov A.Yu., Churakov P.P. // Assessment of the psycho-emotional state of a person based on decomposition into empirical modes and cepstral analysis of speech signals // Bulletin of Penza State University. – 2018. – No. 2 (22). – P. 89–95.
8. Gerasimov A.V., Fidelman V.R. The use of classical and modified linear prediction methods to determine the linear model order in the problem of acoustic speech coding // XXIV scientific readings named after academician N.V. Belov. Abstracts of reports. Nizhny Novgorod. 2005. P. 142-144.
9. Kullback S. Information Theory and Statistics. M.: Nauka, 1967. 408 p.
10. Savchenko V.V., Akatiev D.Yu. Autotesting the quality of speech pronunciation on the principle of minimal information mismatch // Collection of scientific papers "Modern Trends in Computerization of the Process of Studying Foreign Languages". Lugansk: East-Ukrainian National University. 2005. Issue 3. P.205-206.
11. Vasiliev R.A. Study of the features of the phonetic system of speech and text-independent identification of speakers by continuous speech // Information Security of the Regions. 2012. No. 2 (11). P. 57-63.
12. Vasiliev R.A. The study of the phonetic structure of speech and the identification of voice announcers // Issues of information security. 2013. No. 1 (100). P. 43-51.
13. Vasiliev R.A. The study of the identification of voice announcers with differences in pronunciation of speakers // Security of information technology. 2013. No. 1. P. 85-86.
14. Vasiliev R.A. Study of the phonetic system of speech and determining the nationality of the speaker during the voice identification procedure // Information and Security. 2012.Vol. 15. No. 4. P. 487-494.

15. Raigorodsky, D.Ya. Practical psychodiagnostics / D.Ya. Raigo Rodsky. – Samara: Bahrakh-M, 2011. – 672 p.
16. Sonin. V.A. Psychodiagnostic knowledge of professional activity / V.A. Sonin. – SPb.: Speech, 2004. – 408 p.
17. Gray A., Markel J. Distance measures for speech processing // IEEE Trans. On Acoust., Speech and Lang. processing. Vol. 24 (5), oct. 1976. p. 380-291.
18. Garofolo J., Auzanne G., and Voorhees E. The trec spoken document retrieval track: A success story // In proceedings of the Recherche d'Informations Assiste par Ordinateur: Content Based Multimedia Information Access Conference, 2000. pp. 1-20.
19. Huijbregts M., Ordelman R., Jong F. Annotation of heterogeneous multimedia content using automatic speech recognition // In Proceedings of the second international conference on Semantics And digital Media Technologies (SAMT). Lecture Notes in Computer Science. Berlin Springer Verlag. December 2007. pp. 78-90.
20. Mamou J., Mass Y., Ramabhadran B., Sznajder B. Combination of multiple speech transcription methods for vocabulary independent search // In proceedings of the ACM SIGIR Workshop `Searching Spontaneous Conversational Speech. Singapore. 2008. pp. 20-27.
21. Rabiner L. A. Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition // Proceedings of the IEEE. 1989. Vol. 77 No. 2. pp. 257-285.

Васильев Роман Александрович, кандидат технических наук, доцент кафедры «Безопасность информационных систем» Нижегородского государственного университета им. Н.И. Лобачевского
Vasiliev Roman Aleksandrovich, Candidate of Technical Sciences, Associate Professor of the Department “Security of Information Systems” of Nizhny Novgorod State University named after N.I. Lobachevsky