

УДК 81-11

Евдокимова Вера Вячеславовна

кандидат филологических наук,
доцент
СПбГУ
г. Санкт-Петербург, Россия

Vera Evdokimova

PhD in Philology,
Associate Professor
Saint Petersburg State University
Saint-Petersburg, Russia
v.evdokimova@spbu.ru

Скрелин Павел Анатольевич

доктор филологических наук,
профессор
СПбГУ
г. Санкт-Петербург, Россия

Pavel Skrelin

Doctor of Science in Philology,
Professor
Saint Petersburg State University
Saint-Petersburg, Russia
p.skrelin@spbu.ru

Евграфова Карина Владимировна

кандидат филологических наук,
доцент
СПбГУ
г. Санкт-Петербург, Россия

Karina Evgrafova

PhD in Philology,
Associate Professor
Saint Petersburg State University
Saint-Petersburg, Russia
k.evgrafova@spbu.ru

ФОНЕТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ПСИХОФИЗИОЛОГИЧЕСКОГО СОСТОЯНИЯ ЧЕЛОВЕКА

В статье приводится краткий обзор работ, которые проводятся на кафедре фонетики и методики преподавания иностранных языков СПбГУ в рамках исследований физического, психофизиологического и эмоционального состояния человека по речевому сигналу. Описаны основные направления работ и продолжающиеся исследования. Дистанционная оценка психофизиологического и эмоционального состояния оператора, и верификация принимаемого речевого сообщения, и формализация тонких фонетических признаков диктора для качественного синтеза, распознавания его речи стала возможной в связи с развитием вычислительных возможностей и систем искусственного интеллекта. Среди направлений текущих исследований кафедры описаны работы по голосовой усталости и патологиях, когнитивной нагрузке, когнитивному, физическому и эмоциональному состояниям. В целом, перспективными являются и работы с применением систем искусственного интеллекта, и использование непосредственно фонетических признаков для получения результатов высокого уровня эффективности.

Ключевые слова: фонетика; акустический анализ речи; психофизиологическое состояние; когнитивные нарушения; речевые расстройства.

PHONETIC CHARACTERISTICS OF THE PSYCHOPHYSIOLOGICAL STATE OF A PERSON

The article provides a brief overview of the work that is being carried out at the Department of Phonetics and Methods of Teaching Foreign Languages of St. Petersburg State

University in the framework of work on determining the physical, psychophysiological and emotional state of a person by speech and voice. The main areas of work and ongoing research are described. Remote assessment of the psychophysiological and emotional state of the operator, and verification of the received speech message, and formalization of subtle phonetic features of the speaker for high-quality synthesis, recognition of his speech became possible due to the development of computing capabilities and artificial intelligence systems. Among the areas of current work at the department, studies of voice fatigue and pathologies, cognitive load, cognitive states, physical and emotional state are described. In general, both work using artificial intelligence systems and the use of direct phonetic features to obtain high-level results are promising.

Key words: phonetics; speech acoustics; psychophysiological state; cognitive impairment; speech disorders.

Человеческая речь – это сложнейшее явление. Она отражает мышление человека и прежде всего способствует информационному обмену между людьми. Каждый элемент речевого потока несет слушателю разнообразную информацию. Однако речь – не единственный информационный речевой канал. Важна не только информационная составляющая сообщения, но и как человек говорит. Это отражает его психофизиологическое состояние, его вольное или невольное стремление передать свое отношение к содержанию высказывания, к собеседнику к окружающей обстановке, к ситуации общения и т. д. Если языковые единицы передают содержание сообщения, то его фонетическая картина предельно сложна и неоднозначна [1].

Текущий уровень технического прогресса в области обработки речи делает возможным не только распознавать смысл речевого сообщения, но и анализировать фонетические признаки. Появление новых вычислительных возможностей позволило ставить новые прикладные задачи. Это и дистанционная оценка психофизиологического и эмоционального состояния оператора, и верификация принимаемого речевого сообщения, и формализация тонких фонетических признаков диктора для качественного синтеза, распознавания его речи [2].

Зафиксированная методами современной звукозаписи речь представляет собой исходный информационный массив, в котором заложена информация обо всех ее особенностях. Их выявление зависит от выбранного алгоритма обработки собранного материала. Адекватность и ценность результатов обработки определяется всем процессом работы с речевым материалом. На этапе сбора материала нужно обеспечить его достаточность и соответствие поставленной задаче, на этапе подготовки выделить и сформировать аудиофайлы, пригодные для дальнейшей работы, на этапе обработки правильно выбрать метод исследования.

Устойчиво выделить характерные просодические и экстралингвистические особенности речи диктора можно только на основе обработки обширного речевого материала. Современные звукозаписывающие системы дают возможность сбора и стабильного хранения огромных речевых массивов. Однако в общем случае такие массивы непригодны для фонетического

анализа. Они могут представлять собой многосторонние диалоги, сопровождающиеся к тому же музыкой, производственными, транспортными и прочими шумами. Необходима процедура выделения речевых отрезков для решения поставленных фонетических задач распознавания [2].

Внедрение речевых технологий в построение новых интерактивных систем потребовало сосредоточить внимание не только на лингвистическом содержании речевой посылки, но и на его эмоциональной окраске как важном дополнительном канале передачи информации [3].

Анализ особенностей диалога в системах «человек-человек» и «человек-машина» позволяет выделить задачу учета эмоциональной и психофизиологической стороны речи и, соответственно, основные направления исследований.

Если говорить о распознавании эмоционального состояния диктора, то современная вычислительная техника открывает возможность постепенного набора информации об эмоциях и их классификации. Математический аппарат определения места эмоции в n -мерном пространстве может быть различным, но имеет общее целевое назначение [2]. Распознавание эмоций становится важной составной частью программного обеспечения центров сервисного обслуживания, систем жизнеобеспечения, систем обучения, интерактивных игр и т. д. Регистрация эмоций пользователя, таких как гнев, замешательство, страх, смущение, дает важнейшую информацию для улучшения качества обработки речевого сообщения, понимания запросов пользователя и, в конечном счете, для принятия правильного решения [3]. Часто оценка эмоций позволяет определить проблемное место в беседе, изменить стратегию диалога, заметить негативную реакцию собеседника [3].

В приложении к компьютеру изучение проявления в речи человеческих эмоций открывает возможности их моделирования и внедрения, например, при разработке эмоционально окрашенной речи для дружественных интерфейсов. Существуют программы, уже дополненные технологиями такого рода. Они обеспечивают естественность коммуникации между человеком и машиной [4, 5, 6].

Вопрос оценки физического состояния оператора по его речи лежит в области, пограничной с оценкой эмоционального состояния, и неразделим с ним. Большинство исследователей отмечают взаимопроникновение пространств признаков этих направлений, но предпочитают проводить методическую проработку на базе хорошо экспериментально классифицируемых эмоций. В то же время отмечается необходимость решения задач объективной оценки физического состояния оператора (безопасность на транспорте, авиация, медицина). Разрабатываемый при оценке эмоционального состояния математический аппарат может быть, при необходимости, переориентирован без значительных изменений на оценку физического состояния. Можно предполагать, что наибольшую трудоемкость представит сбор и классификация речевого материала.

На кафедре фонетики и методики преподавания иностранных языков СПбГУ продолжаются научные работы по анализу речи, которые могут стать основой новых способов оценки психофизиологического состояния человека [7, 8].

Продолжаются исследования голосовой усталости и патологий голоса совместно с фониатром Мариинского театра Н. В. Швалевым [9–11]. Голосовая усталость является сложным физиологическим явлением, которое проявляется в том числе и в изменении акустических параметров голоса. Его исследование представляет интерес как для уточнения моделей порождения голоса, так и для разработки методов объективной оценки качества голоса. Наиболее сильное проявление состояния голосовой усталости можно наблюдать у профессионалов голоса: преподавателей, певцов, актеров, гидов, переводчиков и т. д.

Для представителей голосо-речевых профессий особую важность представляет точная и ранняя диагностика нарушений работы голосового аппарата. Наряду с существующими инвазивными методами обследования, акустический анализ находит широкое практическое применение для безопасной и удобной оценки голосовой функции, обнаружения голосовых расстройств, а также эстетической оценки голоса. В выпускной квалификационной работе Софьи Куряковой «Проявление в голосе признаков заболевания речевого тракта» (2024 г.) особое внимание уделяется обзору существующих акустических коррелятов патологий голосового аппарата, а также анализу методов для их детектирования. Особенность данного исследования обусловлена ее направленностью на анализ вокальной и невокальной речи путем сравнения параметров джиттера, шиммера, HNR, а также спектральных характеристик при наличии или отсутствии патологии. Результаты исследования продемонстрировали существенные отличия акустических характеристик вокальной и невокальной речи в случае здорового или патологического состояния голосового аппарата. Кроме того, были отмечены значимые тенденции в связи с выбором регистра в пении (средний, высокий и низкий). Оценка результатов показала, что при фонации в высоком регистре отмечается уменьшение параметров джиттера и шиммера, увеличение HNR. Для анализа вокальной речи при мониторинге голосового здоровья был обоснован выбор в пользу спектрального анализа, а также параметров, рассчитывающих долю шумового и гармонического компонента в сигнале. Для групп пациентов были выявлены наиболее значимые параметры в зависимости от клинической картины.

В работе К. В. Евграфовой, Н. В. Швалева и Н. С. Соколовой при изучении акустических характеристик голоса и речи преподавателей высшей школы в состоянии утомления до начала пандемии были получены следующие результаты: изменения значений частоты основного тона, диапазона частоты основного тона, средних длительностей гласных и согласных, изменение количества и длительности пауз, общее сокращение интенсивности артикуляции согласных и изменения в месте образования ряда согласных

[12]. В цели исследования, проведенного в течение учебных семестров 2020–2021 гг., входил сравнительный анализ голосовой нагрузки преподавателей нашего университета, вызванной необычными условиями работы в период пандемии COVID-19.

В 2023–2025 годах проходили работы по определению фонетических характеристик как маркеров когнитивной нагрузки под руководством В. В. Евдокимовой была защищена магистерская диссертация М. Р. Максимовой [13]. Цель работы заключалась в изучении фонетических особенности речи при когнитивной нагрузке. Для выполнения работы необходимо решить следующие задачи: сравнение характеристик речи в различных условиях: при когнитивной нагрузке и при ее отсутствии; при когнитивной нагрузке и при предпаузальном удлинении, изучение визуальных маркеров когнитивной нагрузки и их связи с проявлениями когнитивной нагрузки в голосе. Были сделаны выводы о том, что при когнитивной нагрузке скорость артикуляции, в целом, понижается. Человек меняет также и интонационные характеристики речи, более часто использует ровные интонационные контуры.

В большинстве случаев визуальные и фонетические признаки когнитивной нагрузки совпадали по времени начала. Сегменты с одновременным проявлением мимических и речевых показателей характеризовались большей длительностью по сравнению с участками, в которых наблюдались только визуальные признаки. Это может означать, что распознавание когнитивной нагрузки одновременно по характеристикам речи и мимики способно дать более точный результат по сравнению с распознаванием только по акустическим параметрам или только по мимическим признакам.

На кафедре фонетики и методики преподавания иностранных языков ведутся работы по работе с фонетическими признаками речевых расстройств вместе с логопедами из Ярославского государственного педагогического университета им. Ушинского – был составлен фонетически сбалансированный тестовый материал для автоматического анализа артикуляционных нарушений.

Также на кафедре ведутся работы по определению фонетических характеристик для распознавания когнитивных нарушений по типу Альцгеймера. Фонетические признаки при традиционном анализе дают более высокую эффективность, чем системы, основанные на использовании технологий искусственного интеллекта [14]. Анализ артикуляции вносит наибольший вклад и в определение характера и других нейро-дегенеративных заболеваний (Паркинсон и умственные расстройства) [15].

Таким образом, можно констатировать, что методы и инструменты анализа речевого сигнала, хорошо известные фонетистам по описаниям артикуляторных, акустических и перцептивных свойств произносительной нормы языка, его социальных и интерферированных вариантов, имеют широкие перспективы применения в исследованиях и моделировании самых разных аспектов речевого поведения человека.

ЛИТЕРАТУРА

1. Евдокимова В. В. Акустический анализ речевого сигнала., Санкт-Петербург: Скифия-принт, 2023. 128 с.
2. Евдокимова, В. В. Фонетические методы распознавания психофизиологического состояния человека. Санкт-Петербург: Скифия-принт, 2023. 108 с.
3. А. А. Поволоцкая, В. В. Евдокимова, П. А. Скрелин. Запись и апробация набора речевых данных для распознавания негативных эмоций в речи // Terra Linguistica. – 2023. – Т. 14, № 2. – С. 59–76.
4. Emotional Speech Recognition of Holocaust Survivors with Deep Neural Network Models for Russian Language. Bukreeva, L., Guseva, D., Dolgushin, M., Evdokimova, V., Obotnina, V. (2023). In: Karpov, A., Samudravijaya, K., Deepak, K.T., Hegde, R.M., Agrawal, S.S., Prasanna, S.R.M. (eds) Speech and Computer. SPECOM 2023. Lecture Notes in Computer Science(), vol 14338. Springer, Cham.
5. Developing a Question Answering System on the Material of Holocaust Survivors' Testimonies in Russian. Bukreeva, L., Guseva, D., Dolgushin, M., Evdokimova, V., Obotnina, V. (2023). In: Karpov, A., Samudravijaya, K., Deepak, K. T., Hegde, R. M., Agrawal, S. S., Prasanna, S. R. M. (eds) Speech and Computer. SPECOM 2023. Lecture Notes in Computer Science(), vol 14339. Springer, Cham.
6. Автоматический анализ эмоциональной мимики и интонации в русской речи. Щербаков, П. П. Кочеткова, У. Е. Скрелин, П. А. Борисов, Н. В. Федькин, П. С. Евдокимова, В. В. & Долгушин, М. Д. 2024, Язык-Музыка-Жест: информационные перекрестки (LMGIC-2022): МАТЕРИАЛЫ МЕЖДУНАРОДНОЙ КОНФЕРЕНЦИИ Санкт-Петербург, 18–20 апреля 2024 г. Эйсмонт, П. М. & Алексеева-Нилова, Т. Е. (eds.). Издательство Санкт-Петербургского университета, с. 102–104.
7. THE EFFECT OF PHYSICAL ACTIVITY ON THE PHONETIC CHARACTERISTICS OF SPEECH. Evdokimova, V. V. & Zakharchenko, E. A., 2019, Models and Analysis of Vocal Emissions for Biomedical Applications: 11th International Workshop, December, 17–19, 2019. Manfredi, C. (ed.). Firenze: Firenze University Press, p. 237–240 4 p. (Proceedings e report; vol. 122).
8. Phonetic changes in speech in various physical states / A. D. Ananeva, D. S. Gineva, V. V. Evdokimova [et al.] // Proceedings of the 6th Saint Petersburg Winter Workshop on Experimental Studies of Speech and Language (Night Whites 2022) : Materials of the symposium, Saint Petersburg, 15–16 декабря 2022 года. – Saint Petersburg: Skifia-print, 2023. – P. 14–15.
9. Acoustic analysis of vocal fatigue in professional voice users. Evgrafova, K. V. & Evdokimova, V. V., 2011, Models and Analysis of Vocal Emissions for Biomedical Applications – 7th International Workshop, MAVeBA 2011. Manfredi, C. (ed.). Firenze University Press, p. 153–156.
10. Detection of the Frequency Characteristics of the Articulation System with the Use of Voice Source Signal Recording Method. Evdokimova, V., Evgrafova, K.,

Skrelin, P., Chukaeva, T., Shvaley, N. (2013). In: Železný, M., Habernal, I., Ronzhin, A. (eds) *Speech and Computer. SPECOM 2013. Lecture Notes in Computer Science()*, vol 8113. Springer, Cham.

11. Речевой корпус для исследования голосовой усталости, связанной с профессиональной нагрузкой. Евграфова, К. В., Евдокимова, В. В., Скрелин, П. А. & Чукаева, Т. В., 2017, Анализ разговорной русской речи (АРЗ – 2017): Труды седьмого междисциплинарного семинара. Политехника-принт, с. 26–31.

12. К. В. Евграфова, Н. С. Соколова, Н. В. Швалев. Особенности артикуляции и голоса преподавателей высшей школы в процессе онлайн-обучения // *Современные векторы развития специального и инклюзивного образования: Сборник научных статей II международной научно-практической конференции, Ярославль, 13–14 февраля 2024 года.* – Ярославль: Ярославский государственный педагогический университет им. К. Д. Ушинского, 2024. – С. 199–205.

13. Максимова, М. Р. Фонетические характеристики речи при когнитивной нагрузке / М. Р. Максимова // *Фонетический лицей: Сборник статей.* – Санкт-Петербург : ООО "Скифия-принт", 2024. – С. 44–48.

14. Costantini, G., Cesarini, V., Di Leo, P., Amato, F., Suppa, A., Asci, F., Pisani, A., Calulli, A., Saggio, G.: Artificial intelligence-based voice assessment of patients with parkinson's disease off and on treatment: machine vs. deep-learning comparison. *Sensors* 23(4), 2293 (2023).

15. Garcia, A.M., Arias-Vergara, T., C Vasquez-Correa, J., Noth, E., Schuster, M., Welch, A. E., Bocanegra, Y., Baena, A., Orozco-Arroyave, J. R.: Cognitive determinants of dysarthria in parkinson's disease: an automated machine learning approach. *Movement Disorders* 36(12), 2862-2873 (2021).