



ВОСПРИЯТИЕ ПРОСОДИЧЕСКОЙ ОБРАЗУЮЩЕЙ МУЛЬТИМОДАЛЬНЫХ АФФЕКТИВНЫХ СОСТОЯНИЙ

БАРАБАНЩИКОВ В.А.

Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: vladimir.barabanschikov@gmail.com

СУВОРОВА Е.В.

Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0001-8834-2037>, e-mail: esresearch@yandex.ru

МАЛИОНОК А.В.

Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0009-0001-7375-4343>, e-mail: arinamalionok333@gmail.com

Изучаются особенности выражения и восприятия вокальных экспрессий как одной из образующих мультимодальных аффективных состояний человека. В эксперименте, сконструированном на основе русскоязычной версии Женевского теста распознавания эмоций (GERT), приняли участие две группы женщин 18—45 лет по 72 человека в каждой. Одной группе демонстрировались аудиовидеоролики 14 аффективных состояний, длительностью 3—5 с, сыгранные 10 профессиональными актерами по системе К.С. Станиславского. Другой группе, через наушники, предъявлялись аудиофрагменты, извлеченные из этих же роликов. Требовалось распознать состояние натурщиков, используя интерактивное изображение Женевского колеса эмоций. Показано, что вокальные экспрессии отличаются от соответствующих мультимодальных состояний низкой средней точностью выражения и идентификации, иной конфигурацией и высоким разбросом оценок, более тесной зависимостью от степени возбуждения, индивидуальными способами выражения и нечеткой расширенной структурой категориальных полей. Обозначен круг возможностей, которым обладают вокальные экспрессии, готовые к включению в кроссmodalные интеграции.

Ключевые слова: мультимодальные динамические состояния, вокальные экспрессии, точность выражения и идентификации эмоции, структура категориальных полей, кроссmodalная интеграция.

Финансирование. Исследование выполнено при поддержке Российского научного фонда, проект № 24-18-00904 «Механизмы восприятия эмоционального состояния человека в процессах невербальной коммуникации».

Для цитаты: Барабанщиков В.А., Суворова Е.В., Малионик А.В. Восприятие просодической образующей мультимодальных аффективных состояний // Экспериментальная психология. 2024. Том 17. № 3. С. 30—51. DOI: <https://doi.org/10.17759/exppsy.2024170303>



PERCEPTION OF THE PROSODIC FORMATIVE OF MULTIMODAL AFFECTIVE STATES

VLADIMIR A. BARABANSCHIKOV

Moscow State University of Psychology & Education, Moscow, Russia

ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: vladimir.barabanschikov@gmail.com

EKATERINA V. SUVOROVA

Moscow State University of Psychology & Education, Moscow, Russia

ORCID: <https://orcid.org/0000-0001-8834-2037>, e-mail: esresearch@yandex.ru

ARINA. V. MALIONOK

Moscow State University of Psychology & Education, Moscow, Russia

ORCID: <https://orcid.org/0009-0001-7375-4343>, e-mail: arinamalionok333@gmail.com

The features of the expression and perception of vocal expressions as one of the forming multimodal affective states of a person are studied. The experiment, designed on the basis of the Russian-language version of the Geneva Emotion Recognition Test (GERT), involved two groups of women aged 18–45, 72 women each. One group was shown audio-video clips of 14 affective states, lasting 3–5 seconds, played by 10 professional actors according to the Stanislavsky system. The other group was presented with audio clips extracted from the same clips through headphones. It was demanded to recognize the state of the actors using an interactive image of the Geneva Emotion Wheel. It is shown that vocal expressions differ from the corresponding multimodal states by a low average accuracy of expression and identification, a different configuration and a high spread of estimates, a closer dependence on the degree of arousal, individual ways of expression and a fuzzy extended structure of categorical fields. The range of possibilities that vocal expressions have, ready to be included in cross-modal integrations, is indicated.

Keywords: multimodal dynamic states, vocal expression, accuracy of expression and identification of emotion, structure of categorical fields, cross-modal integration.

Funding. The work was supported by the RNF, project No. 24-18-00904 “Mechanisms of perception of the emotional state of a person in the processes of non-verbal communication”.

For citation: Barabanschikov V.A., Suvorova E.V., Malionok A.V. Perception of the Prosodic Formative of Multimodal Affective States. *Экспериментальная психология = Experimental Psychology (Russia)*, 2024. Vol. 17, no. 3, pp. 30–51. DOI: <https://doi.org/10.17759/exppsy.2024170303> (In Russ.).

Введение

Состояние человека невольно выражается в его голосе, в том, что называют просодическими, вокальными или голосовыми экспрессиями. Благодаря их восприятию осуществляется возможность более полного и глубокого понимания партнеров по общению и способность адекватно выстраивать коммуникативные действия. Эмоциональная окраска речи включена в единый поведенческий континуум и тесно переплетается с семантически однородными проявлениями мимики, жестовой активностью, движениями тела и др. Логика взаимосвязи разнотипных, унимодальных экспрессий, как и закономерности восприятия мультимодальных состояний, относятся к числу актуальных тем экспериментально-психологического исследования [1; 3; 4; 8; 11; 16; 17; 30; 34].



Наряду с собственно лингвистическим содержанием, голос передает значительные объемы информации о размерах тела говорящего, его поле, возрасте, здоровье, социальном статусе, чертах личности и, наиболее детально, об эмоциональном состоянии. Голосовые характеристики коррелируют как с мимическими выражениями базовых эмоций (радостью, печалью, удивлением, отвращением, гневом, страхом), так и с невербальными вокальными всплесками (смехом, криком, рыданиями), которые обусловлены не процессом генерации речи, а модификацией дыхания и, как правило, распознаются лучше, чем эмоционально-окрашенные высказывания, особенно в состоянии страха и отвращения [10; 12; 33; 35].

В ходе психоакустических исследований прослежена связь между физическими параметрами голоса и содержанием (категорией) переживаемой эмоции. Так, выражение радости отличается высокой средней частотой звучания, расширенным частотным спектром, большей средней интенсивностью и скоростью речи. Проявления грусти, напротив, связаны с низкой частотой и интенсивностью, нисходящим контуром высоты и замедлением речи. В число наиболее важных акустических переменных входят темп и высота звучания [29; 32; 33; 36]. Рост высоты тона в конце фразы воспринимается как неуверенность и уступчивость говорящего либо как интерес к собеседнику. Падение высоты ассоциируется с радостью и удовольствием, повышение — с удивлением или страхом. При сильном возбуждении говорящего растет сила голоса, меняются значения его высоты и тембра. Понижение высоты в просодическом контуре связывается с эмоциями положительной валентности (удовольствием, радостью), повышение — с амбивалентными или отрицательными состояниями (удивление, страх). Тембр, или частотный спектр, голоса также взаимосвязан с различным эмоциональным содержанием: для гнева характерно усиление высоких обертонов, увеличение звонкости; для страха — их падение [10; 29; 35].

Вероятность релевантного распознавания аффективных состояний по голосу широко варьирует, достигая в отдельных случаях более 80% и опускаясь ниже 20%. Многие зависит от содержания эмоции, способов ее выражения и восприятия, методов оценки, включения в текущие виды деятельности и общения и др. Например, в иноязычном общении радость, грусть и гнев распознаются легче и точнее, чем страх и спокойное состояние. При этом ответы слушателей не всегда соответствуют содержанию демонстрируемой эмоции. Например, гордость часто воспринимается как удовлетворение, а любовь — как грусть. Возникающие рассогласования интерпретируются в терминах ошибок идентификации [3; 8; 19].

В отличие от лицевых экспрессий распознавание состояний человека по его голосу выполняется значительно труднее (имеет более продолжительное время ответа) и менее точно. Вместе с тем оценки голосовых экспрессий под влиянием выражения лица могут быть изменены [15; 25].

Лучше всего на слух распознаются эмоции положительной валентности; эмоции отрицательной валентности адекватно оцениваются хуже, а оценки амбивалентных эмоций слабо согласуются между собой, характеризуя разнонаправленные тенденции [9; 12; 39].

Пол и возраст влияют на выражение и восприятие вокальных экспрессий многозначно и очень дифференцированно. Как слушатели, женщины лучше мужчин распознают эмоции, заключенные в псевдолингвистических высказываниях. Если тон голоса указывает на отвращение или гнев, точность идентификации для обоих полов нередко одинакова. Не всегда обнаруживаются преимущества женщин в распознавании радости, печали и страха. Вместе с тем точность определения состояний по голосу часто повышается, когда он принадлежит представителям слабого пола. Наиболее высокие результаты достигаются при оз-



вучивании состояния отвращения. Женщины лучше распознают проявления радости, признаваемые женщинами, при этом точнее определяют отвращение, гнев, грусть и спокойное состояние, выраженное мужчинами [9; 25; 28; 41].

Возраст молодых слушателей влияет на восприятие вокальных экспрессий в большей степени, чем зрелых и пожилых. При этом эмоции молодых дикторов распознаются лучше особенно молодыми людьми. Возраст выражающих эмоцию влияет на точность идентификации сильнее, чем возраст слушателей. Акустические параметры голоса по-разному влияют на выражение эмоций молодыми и пожилыми дикторами. На разных возрастных этапах женщины в целом лучше справляются с задачами на точность распознавания просодических экспрессий [21; 27; 41].

Вопрос адекватности выражения эмоциональных состояний по интонации голоса не получил однозначного решения. Спонтанные эмоции, возникающие в конкретных жизненных ситуациях, распознаются примерно так же, как и постановочные, сыгранные актерами. Различия проявляются в предпочтениях того или иного содержания: страх и гнев более точно выражается актерами-дикторами, грусть — в спонтанно создаваемых условиях обычными людьми. Главным требованием к адекватному выражению вокальных экспрессий является способность говорящего лично переживать демонстрируемое состояние [14; 26; 42; 43].

В большинстве экспериментальных работ обсуждаемой темы оценки эмоций по голосу изучаются на материале акустической модальности как таковой, а их соотношения с аналогичными оценками экспрессий в зрительной и других модальностях носят рядоположный внешний характер. В процессе непосредственного общения разнородные каналы сенсорной информации актуализируются синхронно, в соответствии с общей задачей текущего поведения, и взаимообуславливают друг друга. Межличностное восприятие мультимодально, а выражение и идентификация унимодальных экспрессий выступают в ином качестве — одной из образующих более сложного функционального целого. Ее свойства и закономерности на сегодняшний день изучены недостаточно.

Цель данной работы состоит в том, чтобы экспериментально раскрыть специфику выражения и восприятия просодической образующей мультимодальных аффективных состояний. Это предполагает:

- 1) разработку методики изучения особенностей выражения и восприятия эмоционального состояния человека по интонациям его голоса как части мультимодального целого;
- 2) анализ характеристик точности и семантической структуры выражения и идентификации вокальных экспрессий и их зависимости от условий экспозиции;
- 3) соотнесение основных метрик выражения и идентификации вокальных и мультимодальных эмоциональных состояний.

В данном контексте выражение эмоций и их идентификация являются полюсами единого акта межличностного взаимодействия, благодаря которому экспрессия натурщиков обуславливает содержание воспринимаемого, а оценка слушателей отражает состав и структуру проявления переживаний.

Методика

Основу методики составляет Женевский тест распознавания эмоций (Geneva Emotion Recognition Test) — GERT, разработанный научной группой Клауса Шерера для изучения выражения и идентификации мультимодальных динамических эмоциональных состояний и их производных [34; 38; 40]. Тест позиционируется как адекватный способ оценки переживаемых



эмоций в том виде, в котором они встречаются в реальной жизни. Для академических исследований особенно интересен расширенный набор аудиовидеоизображений ключевых эмоций. Он создавался при участии профессиональных актеров, которые проходили специальную подготовку под руководством известного режиссера, Андреа Новикова. В работе с актерами использовался метод переживания К.С. Станиславского, благодаря которому режиссер создавал настроения, инициирующие требуемое функциональное состояние. Аудиовидеоклипы включали мимику, жесты, движение глаз и головы, а также псевдолингвистические фразы, содержащие интонации выражаемой эмоции. Последние создавались специалистами по фонетике и на слух соответствовали звучанию фраз на западно-европейских языках. Использовались две формы высказываний: а) с восклицательной интонацией — «Ne kali bam sud molen!» и б) с вопросительной — «Kun se mina lod belam?». По смысловому значению: а) «Не могу в это поверить!» и б) «Ты действительно так считаешь?». Аудиозаписи актеров-дикторов различаются ударениями, паузами, тембром, темпом и тоном речи, которые в совокупности создают уникальные способы вокализации эмоциональных состояний. Подбор аудиовидеоизображений строился на следующих принципах: 1) наиболее высокий уровень правдоподобия и аутентичность проявления эмоции; 2) широкий диапазон сложности оценок, необходимый для исключения концентрации ответов на наиболее точно распознаваемых аудиовидеоизображениях; 3) сбалансированность пола актеров и двух видов псевдолингвистических предложений. Раш-тест позволил исключить сложные для распознавания видеоизображения и учесть уровень развития способности распознавать эмоцию в качестве латентного признака [6; 40].

Организация стимульного материала, средства его оценки и интерпретация данных построены на концепте «Женевского колеса эмоции» — объединении в круг 14 категорий эмоциональных состояний, проранжированных в зависимости от их валентности и степени возбуждения [17; 33].

Наше собственное исследование проводилось в рамках русскоязычной версии GERT, в которой воспроизведены схема и ключевые звенья оригинала [4–8]. Эксперимент состоял из двух серий. В одной испытуемым демонстрировались мультимодальные аффективные состояния, в другой — их просодические образующие. Аудиодорожки извлекались из аудиовидеороликов с помощью программы iMovie. Требовалось определить содержание воспринятой эмоции, указав на соответствующую категорию изображения «Женевского колеса». Техническая реализация эксперимента осуществлялась на электронных платформах Vercel, Google Sheets.

Участники исследования. В экспериментах приняли участие студентки и аспирантки российских вузов, постоянно проживающие в г. Москве. Выборка для оценки вокальных экспрессий включала 72 женщины в возрасте от 18 до 45 лет ($M=26,2$, $SD = 7,19$); выборка для оценок мультимодальных экспозиций — 72 женщины в возрасте от 18 до 45 лет ($M=22,47$, $SD = 5,7$).

Процедура. Исследование проводилось в лаборатории Института экспериментальной психологии МГППУ на компьютерах SINTO (AMD Ryzen 3 3100, Radeon RX570, 8 GB ОЗУ) с мониторами Acer XF252QXbmiiprx (23,5", 1920×1080 пх.), использовались наушники Sennheiser HD555 (15–28 кГц, 113 дб), Canyon CNR-HS01N (20–20 кГц, 106 дб), Logitech H111 (20–20 кГц, 100 дб). Дисплей располагался на расстоянии 60 см от лица испытуемого. Размер изображения — около 18×24°, лица — 8×10°. Аудиофрагменты передавались участникам исследования через наушники. Громкость и качество звука настраивались испытуемыми индивидуально.

На начальном этапе участникам исследования сообщались технические требования, инструкция, необходимая для успешного выполнения задания и определения тестируемых

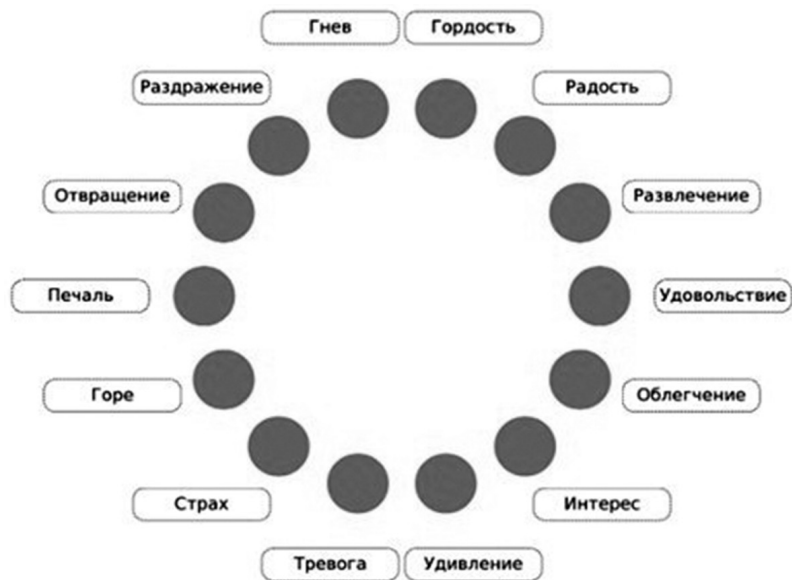


Рис. 1. Снимок экрана интерактивной версии «Женевского колеса эмоций»

эмоциональных состояний; предоставлялась возможность настроить звук. После выполнения тренировочных заданий участник имел возможность вернуться к инструкции повторно.

Во время основной сессии участникам последовательно, в зависимости от группы, в случайном порядке предъявлялись 83 аудиозаписи либо аудиовидеозаписи эмоциональных экспрессий десяти актеров (5 мужчин и 5 женщин, средний возраст — 37 лет). Распределение натурщиков по аудиовидеоклипам было неравномерным. Экспрессии отвращения, печали, горя, удивления выражены четырьмя актерами; гнев и гордость — пятью; радость, развлечение, удовольствие, облегчение, интерес, тревога, страх, раздражение — шестью. Одни и те же актеры могли участвовать в проигрывании состояний в разных объединениях — по шесть, в отдельных случаях пять эмоций [7]. Каждая экспозиция демонстрировалась один раз на 3–5 с без возможности повтора; после его демонстрации на экране дисплея предъявлялось «Колесо эмоций» (рис. 1). Переход к следующей экспозиции осуществлялся автоматически после выбора категории и касания на экране соответствующего круга. Программа исследования завершалась высвечиванием на экране усредненного результата точности распознавания, полученного испытуемым, и его отчетом об особенностях выполнения задания.

Обработка данных. Анализировались усредненные показатели распознавания предъявляемых стимулов, их отношения и зависимости. При анализе структуры категориальных полей использовались усредненные оценки каждой из категорий с порогом узнавания выше 0,05. Ответы участников перекодировались в цифровые значения категорий от 1 до 14 соответственно. Полученные данные проходили предварительную обработку в MS Office, Excel 2010. Для целей статистического анализа использовался пакет программ — SPSS 22.0. Взаимосвязи средней точности оценок демонстрируемых состояний с формами высказываний и полом натурщиков определялись с помощью коэффициентов корреляции Спирмена, $p < 0,05$. В качестве статистических критериев применялись: критерий χ^2 Фридмана, критерий Краскелла–Уолиса с уровнем значимости $p < 0,01$, критерий Вилкоксона с уровнем значимости $p < 0,05$; G-критерий Макнемара для зависимых выборок, с уровнем значимости $p < 0,05$.



Результаты и обсуждение

Обобщенные оценки (частота выбора) мульти- и унимодальных аффективных состояний натурщиков в зависимости от категории эмоций представлены в табл. 1 и 2 соответственно. По горизонтали расположены целевые состояния, по вертикали — ответы испытуемых. Вдоль основной диагонали показаны релевантные оценки, характеризующие точность идентификации. Иррелевантные оценки, расположенные вертикально, указывают на воспринимаемое сходство целевой экспрессии с нецелевыми. Совокупность искомым значений каждого из столбцов (аффективные профили) отражает категориальную структуру полей целевого аффективного состояния.

Таблица 1

Матрица оценок мультимодальных экспрессий эмоциональных состояний натурщиков

Состояния	Гордость	Радость	Развлечение	Удовольствие	Облегчение	Интерес	Удивление	Тревога	Страх	Горе	Печаль	Отвращение	Раздражение	Гнев
Гордость	0,47	0,09	0	0,04	0,02	0,01	0	0	0	0	0	0	0,01	0
Радость	0,16	0,76	0,18	0,04	0,02	0,02	0	0	0	0	0	0	0	0
Развлечение	0,06	0,06	0,68	0,01	0	0	0	0	0	0	0	0	0	0
Удовольствие	0,1	0,01	0,03	0,79	0,12	0,01	0	0	0	0	0	0	0	0
Облегчение	0,01	0,03	0,01	0,06	0,77	0,02	0,01	0,02	0	0	0,02	0	0,02	0
Интерес	0,09	0,01	0	0,05	0,02	0,57	0,11	0,06	0	0	0,01	0	0,06	0
Удивление	0,01	0,04	0,05	0	0	0,14	0,65	0,16	0,03	0	0,01	0,03	0,05	0
Тревога	0	0	0	0	0	0,02	0,1	0,58	0,12	0,11	0,02	0,03	0,04	0
Страх	0	0	0	0	0	0	0,02	0,13	0,6	0,08	0,01	0,02	0	0
Горе	0	0	0,02	0	0,01	0,02	0,02	0,02	0,13	0,61	0,22	0,08	0,02	0,01
Печаль	0	0	0	0	0	0,03	0,01	0,03	0	0,12	0,66	0,07	0,01	0
Отвращение	0,01	0	0,01	0	0,01	0,03	0,03	0	0	0,02	0,03	0,65	0,08	0
Раздражение	0,05	0	0	0	0,02	0,12	0,06	0	0,03	0,03	0,02	0,1	0,67	0,22
Гнев	0,02	0	0	0	0	0	0	0	0,08	0,03	0	0,01	0,04	0,75

Таблица 2

Матрица оценок вокальных экспрессий эмоциональных состояний натурщиков

Состояния	Гордость	Радость	Развлечение	Удовольствие	Облегчение	Интерес	Удивление	Тревога	Страх	Горе	Печаль	Отвращение	Раздражение	Гнев
Гордость	0,17	0,02	0	0,03	0,02	0,1	0,02	0,01	0	0,03	0,02	0,09	0,05	0,03
Радость	0,08	0,33	0,32	0,03	0,01	0,02	0,03	0	0,01	0,03	0	0,02	0,01	0,03
Развлечение	0,04	0,09	0,42	0,01	0	0,02	0,09	0	0	0,01	0,01	0,02	0	0,01
Удовольствие	0,02	0,02	0,01	0,71	0,18	0,15	0,06	0,02	0	0	0,04	0,1	0,01	0
Облегчение	0,01	0,01	0	0,12	0,57	0,13	0,03	0,04	0	0,02	0,08	0,05	0,07	0



Состояния	Гордость	Радость	Развлечение	Удовольствие	Облегчение	Интерес	Удивление	Тревога	Страх	Горе	Печаль	Отвращение	Раздражение	Гнев
Интерес	0,15	0	0	0,03	0,02	0,17	0,33	0,24	0	0	0,06	0,02	0,15	0
Удивление	0,17	0,04	0	0	0,01	0,03	0,33	0,17	0,01	0,01	0,02	0,05	0,1	0,01
Тревога	0,03	0,04	0	0	0,05	0,03	0,05	0,27	0,22	0,13	0,05	0,05	0,06	0,03
Страх	0,01	0,08	0	0	0,01	0	0,01	0,06	0,51	0,12	0,02	0,03	0	0,01
Горе	0,01	0,26	0,2	0,01	0,01	0,01	0	0,01	0,03	0,34	0,16	0,05	0	0
Печаль	0,02	0,05	0,03	0,03	0,06	0,17	0,03	0,15	0,01	0,2	0,52	0,06	0,04	0
Отвращение	0,07	0,01	0	0,01	0,02	0,05	0,01	0,01	0,02	0,04	0,02	0,22	0,1	0,03
Раздражение	0,14	0,01	0	0,02	0,02	0,11	0,02	0,01	0,06	0,06	0,02	0,2	0,37	0,36
Гнев	0,08	0,03	0	0	0	0	0	0	0,13	0,01	0	0,03	0,04	0,5

Процедура сепарации голосовой составляющей разнокачественных аффективных экспрессий предоставляет возможность более детально рассмотреть особенности выражения и восприятия вокальных эмоций по отношению к характеристикам мультимодальных состояний, полученных в данной и предыдущих работах [4–8].

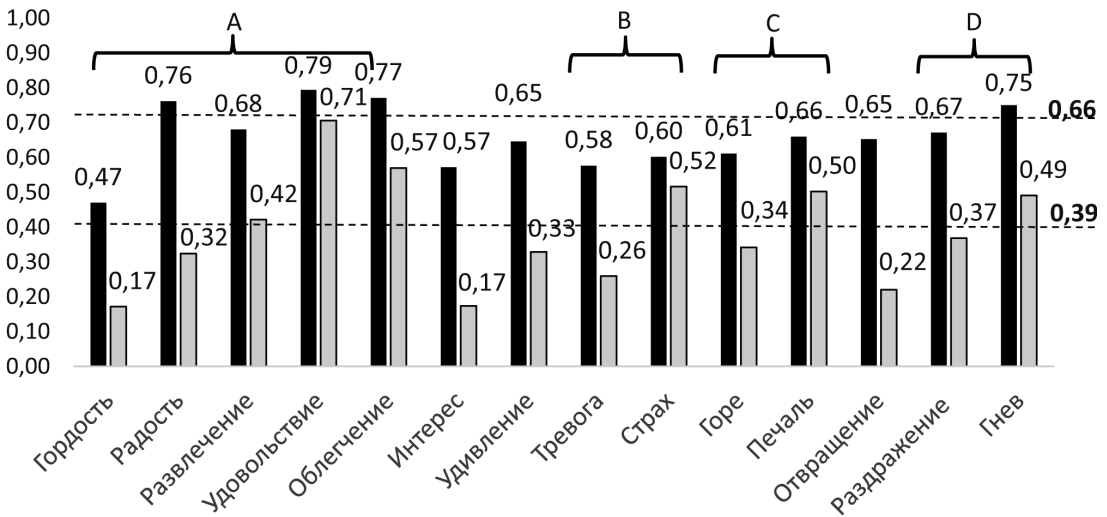


Рис. 2. Гистограммы точности распознавания эмоций по интонациям голоса и мультимодальным экспрессиям (лицо + голос). По горизонтали — экспонируемые категории эмоций, по вертикали — частота релевантных ответов испытуемых. Цифры сверху — средние значения частоты ответов для каждой категории эмоций; пунктирные линии — средние значения частоты ответов на экспозиции вокальных и мультимодальных экспрессий. А, В, С, D — аффективные группы; — вокальные экспрессии; — мультимодальные экспрессии

Точность идентификации вокальных экспрессий. Первое, что обращает на себя внимание — это низкий средний уровень точности оценок вокальных экспрессий ($M=0,39 \pm 0,16$) (рис. 2). В наших экспериментах с мультимодальным стимульным материалом, включающим точно такую же акустическую составляющую, полученные значения более чем в полтора раза



(1,6–1,8) выше. Средние значения точности идентификации ряда мультимодальных экспрессий отличаются от вокальных еще больше: выражения интереса — в 3,6 раз, отвращения — в 3 раза, гордости — в 2,8 раз. Это ожидаемо указывает на трудность выражения голосовых экспрессий по отношению к лицевым, что показывают и унимодальные сравнения, а также на их открытость кроссмодальным взаимодействиям и более полную информацию, которую несут интегрированные выражения эмоций [23; 25].

Серьезно различаются и конфигурации точности оценок вокальных и мультимодальных экспрессий. При голосовых выражениях максимально точно распознается только удовольствие (0,71), наименее точно — гордость (0,17), интерес (0,17) и отвращение (0,22). В условиях мультимодальности лучше всего, наряду с удовольствием (0,79), воспринимаются облегчение (0,77), радость (0,76) и гнев (0,75); хуже всего — гордость (0,47), интерес (0,57) и тревога (0,58). Даже радость, известная в литературе (благодаря зрительной модальности) как очень точно и устойчиво узнаваемое эмоциональное состояние, по голосу оценивается ниже средних значений. В рейтингах оценок мульти- и унимодальных (вокальных) экспрессий одни и те же эмоции могут занимать противоположные позиции, как, например, в случаях страха или печали. Дивергенция значений конфигураций точности идентификации вокальных и мультимодальных экспрессий позволяют допустить множественность сочетаний разномодальных оценок аффективных состояний людей в условиях, приближенных к экологически валидным, и соответствующие им механизмы интеграции.

Точность идентификации вокальных экспрессий соотносима с функциональными свойствами аффективных групп, что в целом также соответствует данным исследований унимодальных вокальных экспрессий. Распознавание эмоций с положительной валентностью (группа А), даже с учетом низких оценок гордости, остается выше общего среднего уровня (0,45+/-0,26), как и оценки вокальных эмоций с отрицательной валентностью — аффективных групп В, С, D, соотношения значений точности идентификации которых между собой очень похожи (0,41+/-0,18). Распознавание амбивалентных вокальных эмоций почти в два раза хуже (0,25+/-0,22); возникают вопросы о его адекватности. Складывается впечатление, что четко выражаемая валентность является фактором эффективности восприятия унимодальных эмоций данного типа. В условиях мультимодальной экспозиции резкое падение точности идентификации амбивалентных состояний не наблюдалось.

Точность идентификации и степень возбуждения диктора. В условиях мультимодальной экспозиции прогнозируемый рост точности идентификации с увеличением степени возбуждения имеет место только по отношению к состояниям группы D [5; 6]. При предъявлении унимодальных вокальных экспрессий в качестве образующих мультимодальных состояний линейная зависимость результата распознавания от возбуждения отмечается во всех тестируемых группах В, С, D. Эмоции с более высокой степенью возбуждения дикторов идентифицируются точнее: тревога (0,26) и страх (0,52), $z = -3,861$, $p < 0,01$; горе (0,34) и печаль (0,50), $z = -5,84$, $p < 0,01$; раздражение (0,37) и гнев (0,49), $z = -3,861$, $p < 0,01$. Средние значения точности распознавания вокальных эмоций для групп В, С и D статистически не различимы ($M = 0,39; 0,42; 0,43$ соответственно), а соотношения значений родственных категорий находится на одном уровне.

Точность идентификации вокальных экспрессий и акустический контур высказывания. Высказывания дикторов в каждой экспозиции содержат законченную мысль и выражают переживание соответствующих аффективных состояний. Вопросительные суждения, как правило, завершаются повышением тона голоса, восклицательные повествова-



ния — его повышением или понижением, которые определяются содержанием и контекстом эмоции. Поскольку изменение тона является одним из важных акустических признаков мелодики речи, можно ожидать, что разные акустические контуры использованных в эксперименте фраз способны повлиять на оценку аффективных состояний. Выполненные расчеты показывают, что, несмотря на отдельные эпизоды, в которых повышение тона в конце высказывания действительно увеличивает точность идентификации эмоции (об этом сообщают и участники эксперимента), значимые корреляции между типом высказывания и совокупными оценками состояния дикторов по голосу не обнаружены. Факты повышения/понижения точности восприятия эмоций в связи с изменением тона голоса требуют дополнительного исследования.

Выражение и восприятие вокальных экспрессий. Среди детерминант восприятия аффективных состояний особое место занимают способы их выражения разными натурщиками. На рис. 3 показана средняя точность вокализаций эмоций по отношению к индивидуальным способам объективации мультимодальных состояний каждого из дикторов-актеров.

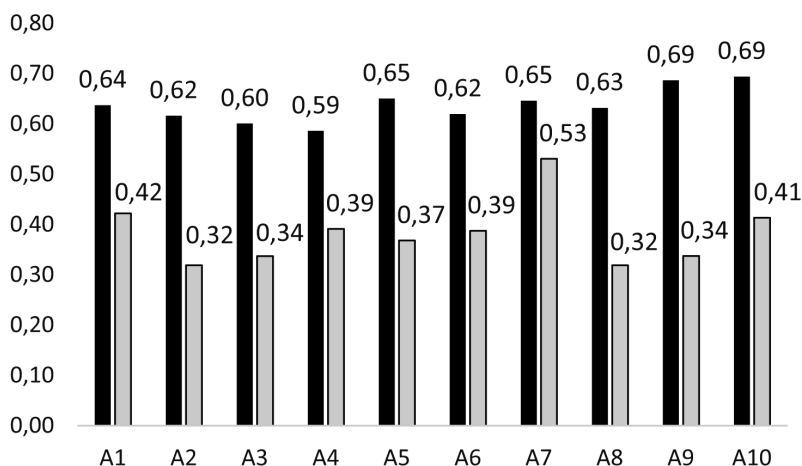


Рис. 3. Средняя точность вокализации эмоций и ее соотношение с оценками мультимодальных эмоциональных состояний, выраженных актерами (An); — вокальные экспрессии; — мультимодальные экспрессии

Максимальные значения точности вокализаций получены натурщиком A7. Это единственный диктор, достигающий сопоставимой точности со значениями мультимодальных выражений. Результаты, показанные остальными девятью дикторами, в 1,6 раза хуже. Наименее точные вокализации выполнены натурщиками A2, A3, A8, и A9. Эффект фасилитации — повышение точности идентификации эмоций при использовании дополнительных каналов однотипной информации [25; 37] продемонстрирован всеми актерами (A7 — как тенденция).

В зависимости от категории целевой эмоции рейтинги точности выражений меняются. Как правило, первые позиции занимают разные натурщики (рис. 4). С каждой новой целевой категорией они как бы вновь перемешиваются, занимая в рейтинге иные места. Также, как и выражения мультимодальных состояний, вокализации модально специфичны. В частности, при выражении голосом эмоций группы В диктор A8 может находиться в конце одного (тревога) и вначале другого (страх) списка. На уровне мультимодальных экспозиций это приводит к сочетаниям, при которых оценки одной из образующих снижа-

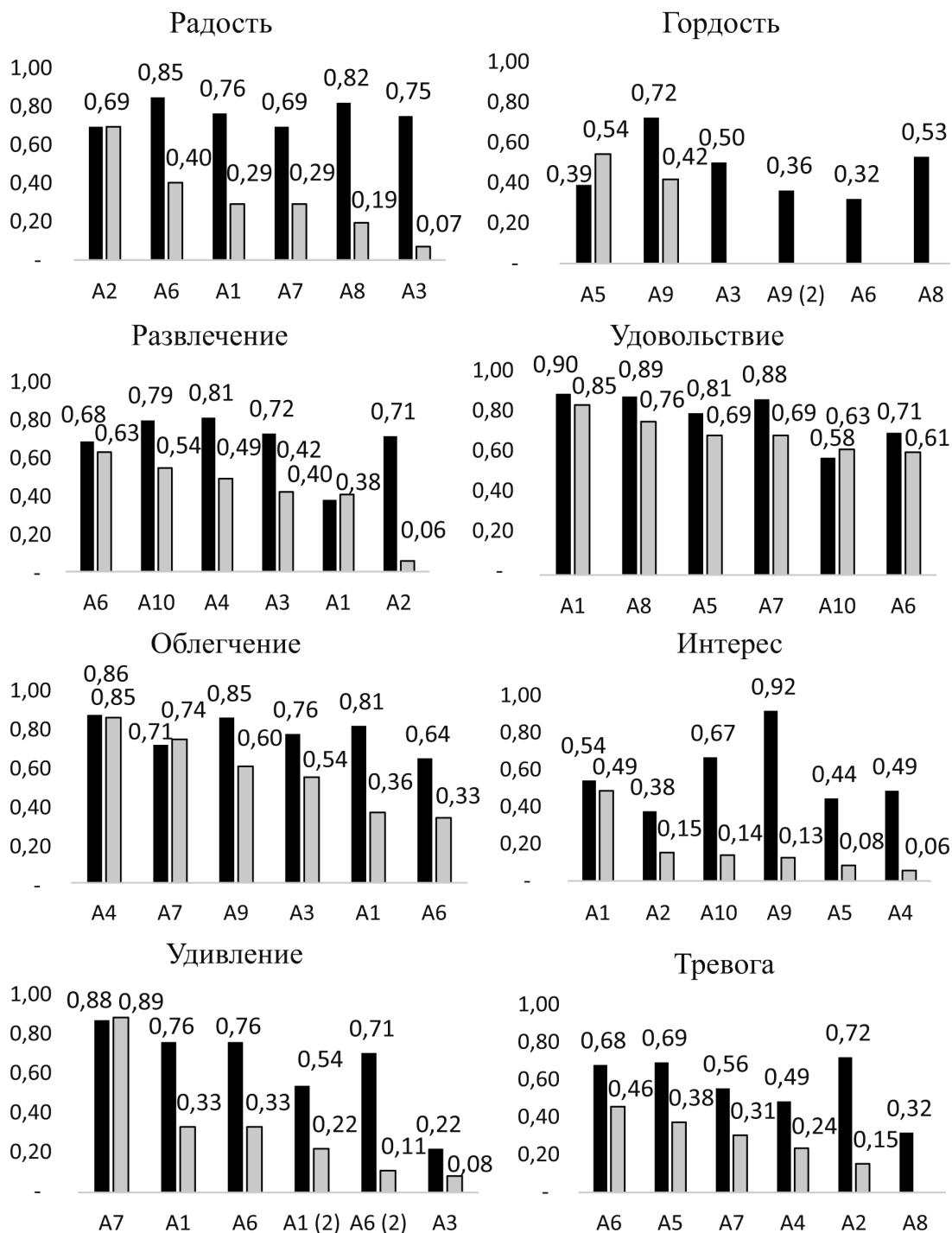


Рис. 4. Рейтинги точности выражения вокальных экспрессий разными натурщиками в зависимости от содержания эмоций; вверху — целевая категория эмоции, внизу — идентификационные номера дикторов-актеров, An; An (2) указывает на участие одного и того же актера в демонстрации разных объединений эмоций. — вокальные экспрессии; — мультимодальные экспрессии

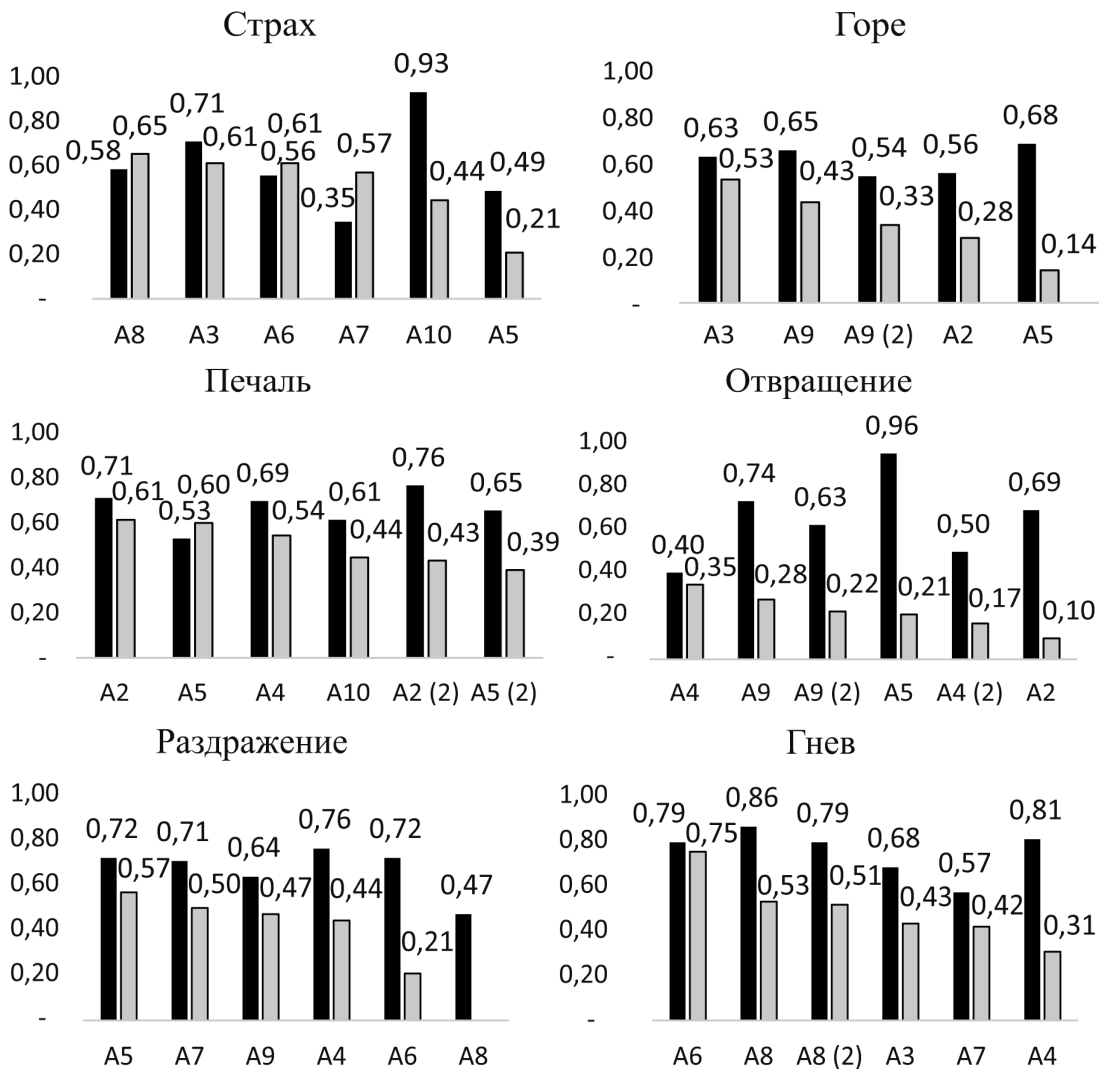


Рис. 4. Продолжение

ются до предельно низких значений или исключаются вовсе (отвращение и раздражение, выраженные A2; радость, выраженная A3).

За некоторым исключением первые позиции занимают дикторами, точность вокализаций целевой эмоции которых соответствует уровню выражения мультимодальных состояний, и кратно отличается от вокализаций других дикторов. В отдельных эпизодах адекватная вокализация полностью вытесняется признаками выражения других эмоций (гордость, раздражение, тревога). Наиболее полная и эффективная вокализация с участием большинства натурщиков получена на экспрессии удовольствия, результаты которой статистически неразличимы от выражения мультимодальных состояний. Очень низкая эффективность групповой вокализации связана с эмоцией отвращения. За богатой феноменологией унимодальных экспрессий стоит широкий диапазон индивидуальных средств и способов вокализаций эмоций.



Согласно полученным данным полнота и точность выражения вокальных и мультимодальных аффективных состояний в значительной степени не совпадают. При исключении лицевых экспрессий вокализации одних и тех же переживаний одними и теми же актерами воспринимаются иначе.

Точность выражения аффективных состояний и пол натуращиков. Ранее на материале аудиовидеоизображений эмоциональных состояний мы нашли, что почти в 80% случаев экспозиции актеров и свыше 90% случаев экспозиции актрис гендерная стилистика в характере оценок не проявляется. В отдельных фрагментах женщины точнее отыгрывают мультимодальные состояния и наиболее эффективно распознают эмоции, продемонстрированные другими женщинами. Восприятие мультимодальных состояний мужчин и женщин в целом совпадают [7]. В новых экспериментах сделанный вывод подтвердился. Значения коэффициента корреляции между средней точностью вокализации эмоций и полом актеров-натуращиков стремится к нулю, как и взаимодействие между точностью выражения мультимодальных экспрессий и полом натуращика. Полученный результат расходится с рядом данных унимодальных исследований голосовых эмоций.

Категориальные поля вокальных экспрессий. Категориальное поле аффективного состояния характеризует структуру семантического содержания воспринимаемой эмоции. Это объединение различных впечатлений об одной и той же целевой эмоции, полученных группой испытуемых в рамках системы оценочных категорий «Женевского колеса эмоций». Как многокомпонентное образование, категориальное поле дифференцируется на ядро — наиболее точно идентифицируемые экспрессии и периферию — ассоциированные состояния, составляющие фон.

Категориальные поля мультимодальных состояний (рис. 5) характеризуются четко структурированным контуром: имеет место выраженный эксцесс точности оценок целевых эмоций, величина которого многократно превышает значение оценок компонентов периферии. С удалением позиции компонентов от расположения целевого состояния на «Колесе эмоций» их оценки снижаются, а категории противоположной валентности отсутствуют. Периферия амбивалентных целевых эмоций включает компоненты как положительной, так и отрицательной валентности.

Анализ категориальных полей вокальных экспрессий показывает, что восприятие каждой экспонируемой эмоции имеет уникальную семантическую структуру и идентифицируется слушателями иначе. Ее объем — количество аффективных компонентов и частота их обнаружения — широко варьирует. В среднем на каждую целевую категорию приходится 3–4 дополнительных экспрессии; в некоторых случаях их может быть шесть (отвращение, раздражение) или всего одна (удовольствие, гнев); в последнем случае развитая периферия отсутствует. С увеличением числа компонентов и степени их выраженности точность идентификации целевой экспрессии снижается (облегчение, страх, гнев).

Роли компонентов в структуре категориального поля связаны с их расположением на круговой шкале «Женевского колеса» относительно целевой категории. В ранее проведенных экспериментах отмечалось, что иррелевантные ответы испытуемых, как правило, являются смежными или соседствующими с целевыми [8; 39]. Тенденция прослеживается и в условиях вокальной экспозиции, но, в отличие от мультимодальных состояний, иррелевантные оценки чаще распространяются на дальнюю периферию и включают категории противоположной валентности. Наиболее отчетливо тенденция проявляется на примере восприятия эмоций с положительной валентностью (группа А): радость (0,33) содержит



компоненты горя (0,26), страха (0,08) и печали (0,05); развлечение (0,42) — радости (0,32) и горя (0,20); гордость (0,17) наряду с амбивалентными состояниями и радостью (0,08) включает раздражение (0,14), гнев (0,08) и отвращение (0,07); облегчение (0,57) вместе с удовольствием (0,18) — печаль (0,06) и тревогу (0,05). Подобные явления наблюдались и в швейцарских экспериментах, где, например, доля эмоции горя в оценках развлечения составила 0,13 [17]. Расширение периферии категориального поля сопровождается более низкой точностью идентификации целевой эмоции. При экспозиции вокальных экспрессий гордости (0,17), интереса (0,17) и отвращения (0,22) воспринимаемое состояние диктора сливается с фоном и/или маскируется признаками других состояний. Похожий результат получен в исследованиях К. Шерера: более трети оценок гордости составили отрицательные эмоции. Вместе с тем родственные связи положительных эмоций эффективно проявляют себя во взаимовлиянии радости и развлечения или удовольствия и облегчения. В условиях акустической модальности воспринимаемые отношения экспрессий отличаются большей вариативностью и избирательностью.

Смещение оценок в сторону дальней периферии имеет место и при экспозиции отрицательных эмоций группы В (страха, раздражения), С (печали) и D (раздражения), в которых обнаруживаются положительные экспрессии. Наиболее полно родственные связи выражены в группе С.

Амбивалентные вокальные эмоции распознаются плохо и показывают низкую согласованность семантического содержания внутри группы участников. Удивление (0,33)

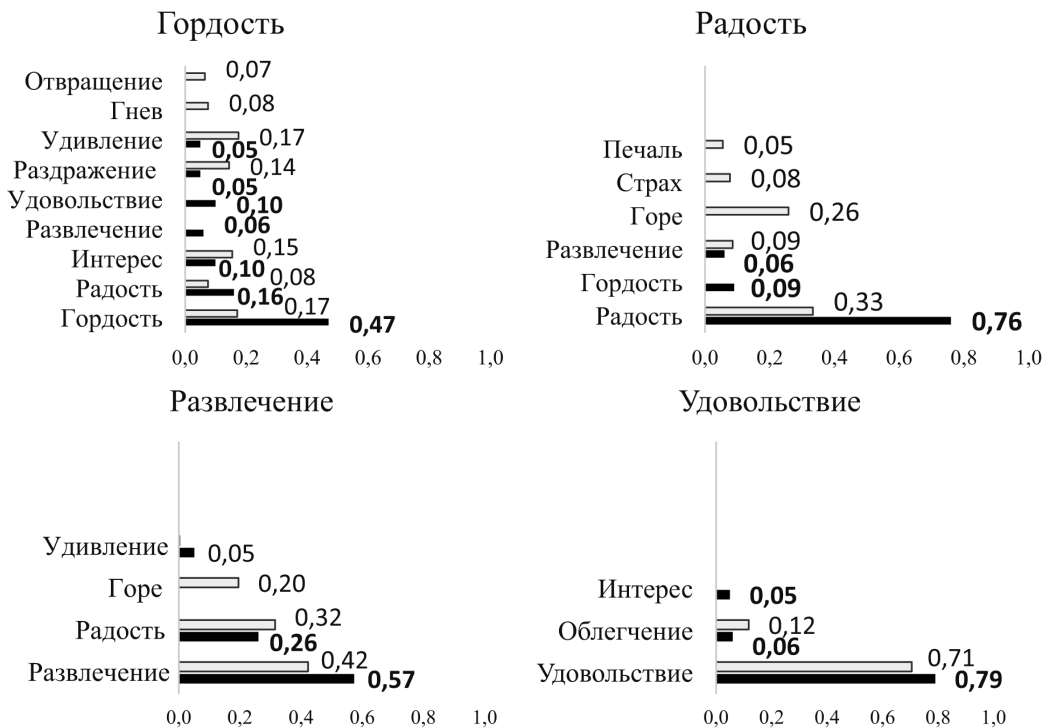


Рис. 5. Структуры категориальных полей мультимодальных и голосовых экспрессий. Вверху указаны категории экспонируемых эмоций, слева — категории ответов испытуемых, внизу — частота выбора (отмечены оценки $\geq 0,05$); — вокальные экспрессии; — мультимодальные экспрессии

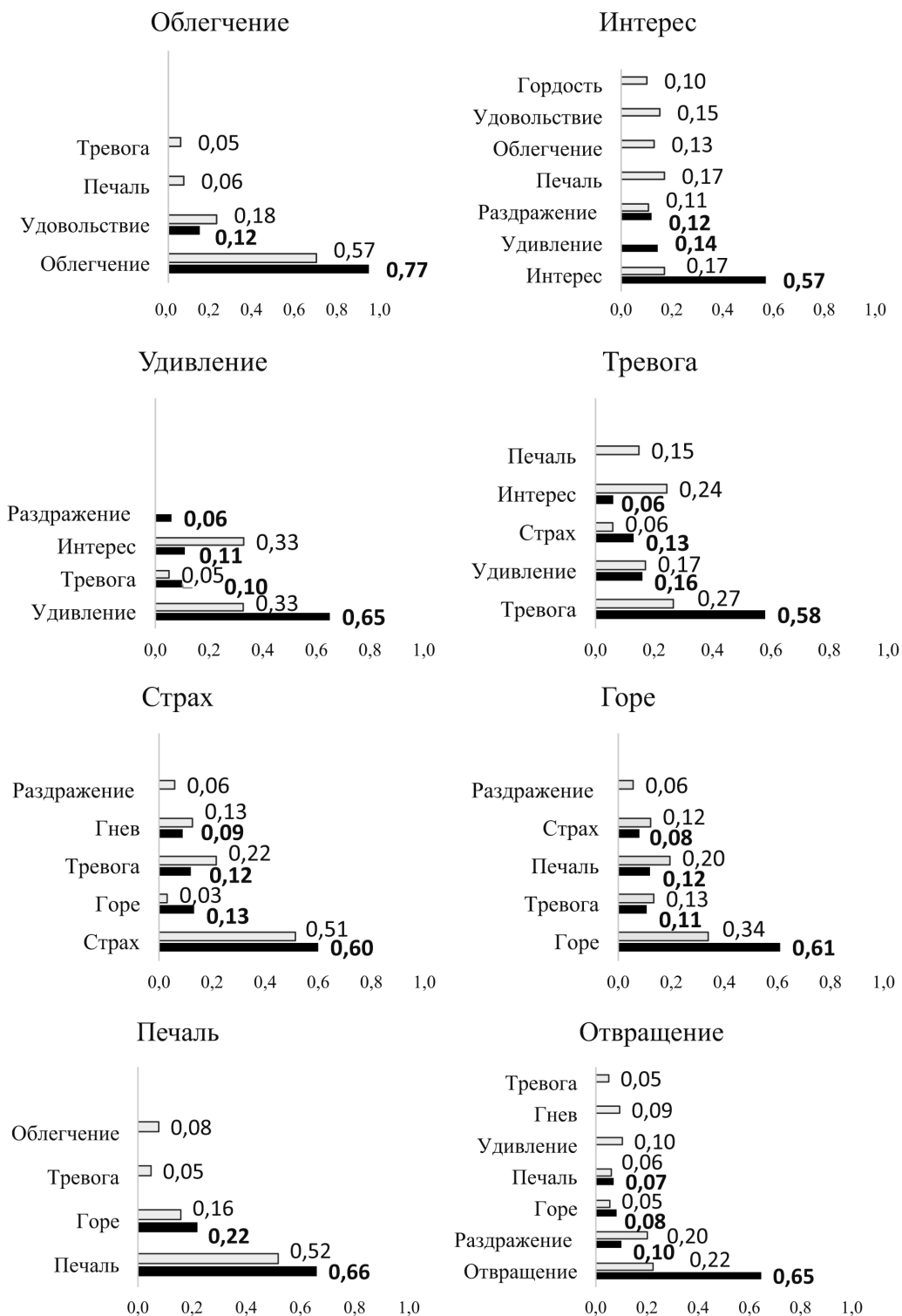


Рис. 5. Продолжение

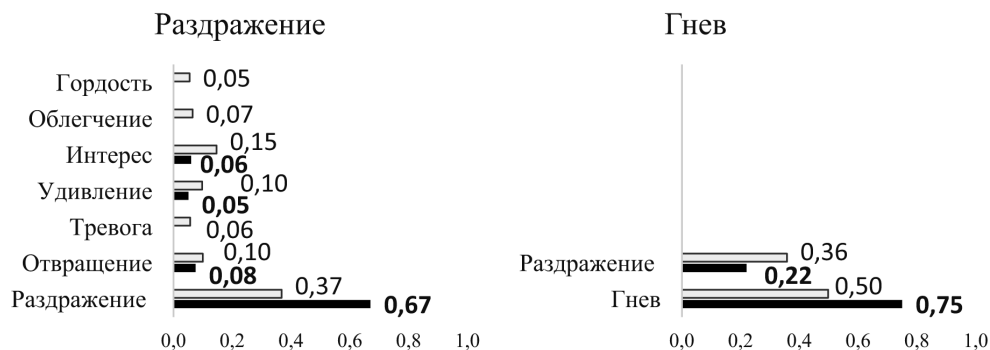


Рис. 5. Продолжение

практически неотличимо от интереса (0,32); интерес (0,17) — от печали (0,17), удовольствия (0,15) и облегчения (0,13); отвращение (0,22) — от раздражения (0,20). Значения точности идентификаций смежных экспрессий чаще носят ассиметричный характер.

В целом, по сравнению с восприятием мультимодальных состояний структура категориальных полей вокальных экспрессий выглядит менее определенной, более размытой и смещенной в сторону дальней периферии, а сами эмоции — менее точными и более микшированными.

Резюмируя сказанное, отметим, что вокальные экспрессии, включаясь в более объемное функциональное целое, сохраняют специфику и отличаются от соответствующих мультимодальных состояний рядом основных метрик: низкой средней точностью выражения и идентификации, иной конфигурацией и высоким разбросом оценок, более тесной зависимостью от возбуждения, индивидуальными способами выражения эмоций, семантическим содержанием и нечеткой структурой категориальных полей. Это характеристика собственных возможностей просодической образующей, ее *экспрессивного потенциала*. Когда, как и насколько полно он сможет реализоваться, во многом зависит от природы других образующих, входящих в систему мультимодальных отношений. Они проистекают из одного и того же переживания натурщика и благодаря этому обстоятельству сливаются в единый поток, подчиняющийся закономерностям кроссmodalного целого.

Предложенный метод сепарации образующих мультимодальных экспрессий соответствует задачам выполненного исследования и имеет неплохую перспективу. Очевидно, что в рамках мультимодальных отношений вокализация эмоций занимает подчиненное положение, а доминирует динамика выражений лица. Но всегда ли? Существуют ли условия, при которых происходит обратное. Если да, то какие? Встают вопросы избирательности восприятия и динамической концентрации внимания испытуемых на лице или голосе натурщиков. Их решение ведет к пониманию конкретных механизмов интеграции унимодальных экспрессий в качественно новое целое.

Выводы

Выполненное исследование просодической образующей мультимодальных аффективных состояний человека позволяет охарактеризовать особенности ее выражения и восприятия.

- Средний уровень точности идентификации вокальных экспрессий более чем в 1,6 раза ниже точности мультимодальных состояний. Конфигурации оценок существенно различны. Их дивергенция указывает на множественность типов кроссmodalных взаимодействий и соответствующих им механизмов интеграции.



• В оценках вокальных экспрессий отражаются характеристики функциональных свойств аффективных состояний. Эмоции с положительной (группа А) либо отрицательной (группы В, С, D) валентностью воспринимаются почти в два раза точнее, чем амбивалентные. Подобное снижение в условиях мультимодальной экспозиции не наблюдается. Эффект имеет методическое значение в исследованиях и/или диагностике мультимодальных состояний.

• Идентификации вокальных экспрессий В, С, D зависят от степени возбуждения диктора: с ее ростом точность идентификации значимо повышается. В условиях мультимодальности линейные отношения между точностью оценок и степенью возбуждения воспроизводятся только в группе D.

• Изменение тона в вопросительных и отрицательных фразах, произносимых дикторами-актерами, эпизодически влияют на оценку слушателей, но с общей точностью идентификации не коррелируют.

• Как и проявления мультимодальных состояний, выражения эмоций голосом для каждого натурщика модально-специфичны, обусловлены ее категорией. С изменением содержания целевой эмоции рейтинги точности вокализации натурщиков меняются. В отдельных эпизодах адекватная вокализация полностью вытесняется признаками других экспрессий.

• Значения коэффициентов корреляции между средней точностью вокализаций эмоций и полом натурщиков, а также между средней точностью выражения мультимодальных экспрессий и полом актеров стремятся к нулю.

• Выражение каждой эмоции имеет уникальный экспрессивный состав, который по-разному идентифицируется слушателями. Количество экспрессивных компонентов и частота их обнаружения широко варьируют. В среднем на каждую целевую категорию приходится 2–4 дополнительные экспрессии; с увеличением их числа и степени выраженности точность идентификации эмоции снижается. В условиях мультимодальности тенденция сохраняется.

• Категориальные поля мультимодальных состояний отличаются четко выраженным ядром и анизотропной периферией, ограниченной валентностью целевой эмоции. Структура категориальных полей вокальных экспрессий менее определена, более размыта и смещена в сторону дальней периферии, охватывающей категории противоположной валентности.

• Обозначен экспрессивный потенциал просодической образующей в системе мультимодальных отношений интонаций голоса и динамических экспрессий лица.

Литература

1. Андреева Е.В. Феномен кроссмодальных взаимодействий: современное состояние проблемы в зарубежной психологии // Российский научный журнал. 2013. Том 1. № 32. С. 219–225. DOI:10.17759/rsj.20130132219
2. Барабанщиков В.А., Жегалло А.В., Королькова О.А. Перцептивная категоризация выражений лица. М.: Когито-центр, 2016. 376 с. DOI:10.17759/9785990501513
3. Барабанщиков В.А., Королькова О.А. Восприятие экспрессий «живого» лица // Экспериментальная психология. 2020. Том 13. № 3. С. 55–73. DOI:10.17759/exppsy.2020130305
4. Барабанщиков В.А., Суворова Е.В. Оценка эмоционального состояния человека по его видеоизображениям // Экспериментальная психология. 2020. Том 13. № 4. С. 4–24. DOI:10.17759/exprsy.2020130401



5. Барабанщиков В.А., Суворова Е.В. Оценка мультимодальных экспрессий лица в лаборатории и онлайн // Лицо человека в контекстах природы, технологий и культуры / Отв. ред. К.И. Ананьева, В.А. Барабанщиков, А.А. Демидов. М.: Когито-центр, 2021. С. 310–322. DOI:10.17759/kogito.202103220
6. Барабанщиков В.А., Суворова Е.В. Гендерный фактор в распознавании эмоционального состояния человека по его аудиовидеоизображениям // Российский психологический журнал. 2022. Том 19. № 2. С. 6–20. DOI:10.17759/rpj.2022190201
7. Барабанщиков В.А., Суворова Е.В. Индивидуальные формы выражения и идентификация мультимодальных динамических состояний человека // Познание и переживание. 2022. Том 3. № 2. С. 6–35. DOI:10.17759/kogito.202203220
8. Барабанщиков В.А., Суворова Е.В. Выражение и восприятие мультимодальных эмоциональных состояний // Национальный психологический журнал. 2023. Том 18. № 3(51). С. 106–127. DOI:10.11621/npj.2023.0311
9. Борисова А.А. Восприятие эмоционального состояния человека по интонационному рисунку речи // Вопросы психологии. 1989. № 1. С. 117–124. DOI:10.17759/voppsy.1989010117
10. Burt H. Личностно-ситуационная опосредованность выражения и распознавания эмоций в речи // Вопросы психологии. 1991. № 2. С. 95–107. DOI:10.17759/vp.1991020095
11. Лупенко Е.А., Королькова О.А. Семантическая эквивалентность как основа интермодальной интеграции // Экспериментальная психология. 2022. Том 15. № 3. С. 159–177. DOI:10.17759/exppsy.20221503159
12. Морозов В.П. Невербальная коммуникация: экспериментально-психологические исследования: монография. М: Институт психологии РАН, 2011. 528 с. DOI:10.17759/rusbook.20110528
13. Потапова Р.К., Потапов В.В., Комалова Л.П. Восприятие мультимодальной моно- и полиэтнической коммуникации: монография / Отв. ред. Л.П. Комалова. М: ИНИОН РАН, 2020. 210 с. DOI:10.17759/iniom.2020210
14. Юрьева М.В., Хозе Е.Г. Создание классификатора вокальных эмоциональных экспрессий: персонализированные базы данных // Лицо человека в контекстах природы, технологий и культуры / Отв. ред. К.И. Ананьева, В.А. Барабанщиков. М: Когито-центр; Московский институт психоанализа, 2020. С. 364–372. DOI:10.17759/kogito.2020364372
15. Baart M., Vroomen J. Recalibration of Vocal Affect by a Dynamic Face // Experimental Brain Research. 2018. Vol. 236. № 1. P. 1911–1918. DOI:10.1007/s00221-018-5270-5
16. Barkhuysen P., Kraemer E., Swerts M. Crossmodal and Incremental Perception of Audiovisual Cues to Emotional Speech // Language and Speech. 2010. Vol. 53. № 1. P. 3–30. DOI:10.1177/0023830910362212
17. Bänziger T., Mortillaro M., Scherer K.R. Introducing the Geneva Multimodal Expression Corpus for Experimental Research on Emotion Perception // Emotion. 2012. Vol. 12. № 5. P. 1161–1179. DOI:10.1037/a0025827
18. Bänziger T., Patel S., Scherer K.R. The Role of Perceived Voice and Speech Characteristics in Vocal Emotion Communication // Journal of Nonverbal Behavior. 2014. Vol. 38. № 1. P. 31–52. DOI:10.1007/s10919-013-0165-x
19. Bryant G.A. The Evolution of Human Vocal Emotion // Emotion Review. 2021. Vol. 13. № 1. P. 25–33. DOI:10.1177/1754073919870985
20. Chronaki G., Hadwin J.A., Garner M., Maurice P., Sonuga-Barke E.J.S. The Development of Emotion Recognition from Facial Expressions and Non-Linguistic Vocalizations During Childhood // The British Journal of Developmental Psychology. 2015. Vol. 33. № 1. P. 218–236. DOI:10.1111/bjdp.12075
21. Cortes D.S., Tornberg C., Bänziger T., Elflein H.A., Fischer H., Laukka P. Effects of Aging on Emotion Recognition from Dynamic Multimodal Expressions and Vocalizations // Scientific Reports. 2021. Vol. 11. № 1. P. 1–12. DOI:10.1038/s41598-021-94993-4
22. Cowen A.S., Elflein H.A., Laukka P., Keltner D. Mapping 24 Emotions Conveyed by Brief Human Vocalization // The American Psychologist. 2019. Vol. 74. № 1. P. 698–712. DOI:10.1037/amp0000399
23. Dolan R.J., Morris J.S., de Gelder B. Crossmodal binding of fear in voice and face // Proceedings of the National Academy of Sciences of the United States of America. 2001. Vol. 98. № 17. P. 9465–9470. DOI:10.1073/pnas.171288598
24. ElAyadi M., Kamel M.S., Karray F. Survey on speech emotion recognition: Features, classification schemes and databases // Pattern Recognition. 2011. Vol. 44. № 3. P. 572–587. DOI:10.1016/j.patcog.2010.09.014



25. Gelder B. De, Vroomen J. The Perception of Emotions by Ear and by Eye // *Cognition and Emotion*. 2000. Vol. 14. № 3. P. 289–311. DOI:10.1080/026999300378824
26. Jürgens R., Grass A., Drolet M., Fischer J. Effect of Acting Experience on Emotion Expression and Recognition in Voice: Non-Actors Provide Better Stimuli Than Expected // *Journal of Nonverbal Behavior*. 2015. Vol. 39. № 1. P. 195–214. DOI:10.1007/s10919-014-0201-6
27. Lambrecht L., Kreifelts B., Wildgruber D. Age-Related Decrease in Recognition of Emotional Facial and Prosodic Expressions // *Emotion*. 2012. Vol. 12. № 1. P. 529–539. DOI:10.1037/a0026827
28. Lausen A., Schacht A. Gender Differences in the Recognition of Vocal Emotions // *Frontiers in Psychology*. 2018. Vol. 9. № 1. Article 882. DOI:10.3389/fpsyg.2018.00882
29. Patel S., Scherer K.R., Björkner E., Sundberg J. Mapping Emotions into Acoustic Space: The Role of Voice Production // *Biological Psychology*. 2011. Vol. 87. № 1. P. 93–98. DOI:10.1016/j.biopsycho.2011.02.010
30. Paulmann S., Pell M.D. Is There an Advantage for Recognizing Multi-Modal Emotional Stimuli? // *Motivation and Emotion*. 2011. Vol. 35. № 1. P. 192–201. DOI:10.1007/s11031-011-9211-y
31. Pittam J., Scherer K.R. Vocal Expression and Communication of Emotion // *Handbook of Emotions* / M. Lewis, J.M. Haviland (Eds.). The Guilford Press, 1993. P. 185–197.
32. Sauter D.A., Eisner F., Calder A., Scott S. Perceptual Cues in Nonverbal Vocal Expressions of Emotion // *Quarterly Journal of Experimental Psychology*. 2010. Vol. 63. № 1. P. 2251–2272. DOI:10.1080/17470211003721642
33. Scherer K.R. Vocal Affect Expression: A Review and a Model for Future Research // *Psychological Bulletin*. 1986. Vol. 99. № 2. P. 143–165. DOI:10.1037/0033-2909.99.2.143
34. Scherer K.R. What Are Emotions? And How Can They Be Measured // *Social Science Information*. 2005. Vol. 44. № 1. P. 695–729. DOI:10.1177/0539018405058216
35. Scherer K., Banse R. Acoustic Profiles in Vocal Emotion Expression // *Journal of Personality and Social Psychology*. 1996. Vol. 70. № 1. P. 614–636. DOI:10.1037/0022-3514.70.3.614
36. Scherer K.R., Banse R., Wallbott H.G., Goldbeck T. Vocal Cues in Emotion Encoding and Decoding // *Motivation and Emotion*. 1991. Vol. 15. № 2. P. 123–148. DOI:10.1007/BF00995674
37. Schirmer A., Adolphs R. Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence // *Trends in Cognitive Sciences*. 2017. Vol. 21. № 1. P. 216–228. DOI:10.1016/j.tics.2017.01.001
38. Schlegel K., Scherer K.R. The nomological network of emotion knowledge and emotion understanding in adults: evidence from two new performance-based tests // *Cognition and Emotion*. 2017. Vol. 32. P. 1514–1530. DOI:10.1080/02699931.2017.1414687
39. Schlegel K., Grandjean D., Scherer K.R. Emotion Recognition: Unidimensional Ability or a Set of Modality- and Emotion-Specific Skills? // *Personality and Individual Differences*. 2012. Vol. 53. № 1. P. 16–21. DOI:10.1016/j.paid.2012.01.026
40. Schlegel K., Grandjean D., Scherer K.R. Introducing the Geneva Emotion Recognition Test: An Example of Rasch-Based Test Development // *Psychological Assessment*. 2014. Vol. 26. № 2. P. 666–672. DOI:10.1037/a0025827
41. Sen A., Isaacowitz D., Schirmer A. Age Differences in Vocal Emotion Perception: On the Role of Speaker Age and Listener Sex // *Cognition and Emotion*. 2018. Vol. 32. № 1. P. 1189–1204. DOI:10.1080/02699931.2017.1390725
42. Williams C., Stevens K. Emotions and speech: some acoustical correlates // *The Journal of the Acoustical Society of America*. 1972. Vol. 52. № 4(2). P. 1238–1250. DOI:10.1121/1.1913238
43. Wilting J., Kraemer E., Swerts M. Real vs acted emotional speech // *Proceedings of Interspeech 2006*. Pittsburgh, 2006. DOI:10.21437/Interspeech.2006-250

References

1. Andreeva E.V. Fenomen krossmodal'nykh vzaimodeistvii: sovremennoe sostoyanie problemy v zarubezhnoi psikhologii. *Rossiiskii nauchnyi zhurnal = Russian Scientific Journal*, 2013. Vol. 1, no. 32, pp. 219–225. DOI:10.17759/rsj.20130132219 (In Russ.).
2. Barabanshikov V.A., Zhegallo A.V., Korolkova O.A. *Pertseptivnaya kategorizatsiya vyrazhenii litsa [Perceptual categorization of facial expressions]*. Moscow: Publ. Kogito-Tsent, 2016. 376 p. DOI:10.17759/9785990501513 (In Russ.).



3. Barabanshchikov V.A., Korolkova O.A. Vospriyatie ekspressii “zhivogo” litsa. *Ekspperimental'naya psikhologiya = Experimental Psychology*, 2020. Vol. 13, no. 3, pp. 55–73. DOI:10.17759/exppsy.2020130305 (In Russ.).
4. Barabanshchikov V.A., Suvorova E.V. Ocenka emocional'nogo sostoyaniya cheloveka po ego videoizobrazheniyam. *Ekspperimental'naya psikhologiya = Experimental Psychology*, 2020. Vol. 13, no. 4, pp. 4–24. DOI:10.17759/exppsy.2020130401 (In Russ.).
5. Barabanshchikov V.A., Suvorova E.V. Ocenka multimodal'nykh ekspressii litsa v laboratorii i onlain. *Lico cheloveka v kontekstakh prirody, tekhnologii i kul'tury = The Human Face in Contexts of Nature, Technology, and Culture* / Eds. K.I. Ananyeva, V.A. Barabanshchikov, A.A. Demidov. Moscow: Kogito-Tsentr, 2021. Pp. 310–322. DOI:10.17759/kogito.202103220 (In Russ.).
6. Barabanshchikov V.A., Suvorova E.V. Gendernyi faktor v raspoznavanii emocional'nogo sostoyaniya cheloveka po ego audiovideoizobrazheniyam. *Rossiiskii psikhologicheskii zhurnal = Russian Psychological Journal*, 2022. Vol. 19, no. 2, pp. 6–20. DOI:10.17759/rpj.2022190201 (In Russ.).
7. Barabanshchikov V.A., Suvorova E.V. Individual'nye formy vyrazheniya i identifikatsiya multimodal'nykh dinamicheskikh sostoyanii cheloveka. *Poznanie i perezhivanie = Cognition and Experience*, 2022. Vol. 3, no. 2, pp. 6–35. DOI:10.17759/kogito.202203220 (In Russ.).
8. Barabanshchikov V.A., Suvorova E.V. Vyrazhenie i vospriyatie multimodal'nykh emocional'nykh sostoyanii. *Natsional'nyi psikhologicheskii zhurnal = National Psychological Journal*, 2023. Vol. 18, no. 3(51), pp. 106–127. DOI:10.11621/npj.2023.0311 (In Russ.).
9. Borisova A.A. Vospriyatie emocional'nogo sostoyaniya cheloveka po intonatsionnomu risunku rechi. *Voprosy psikhologii = Questions of Psychology*, 1989. No. 1, pp. 117–124. DOI:10.17759/voppsy.1989010117 (In Russ.).
10. Vitt N. Lichnostno-situatsionnaya oposredovannost' vyrazheniya i raspoznavaniya emocii v rechi. *Voprosy psikhologii = Questions of Psychology*, 1991. No. 2, pp. 95–107. DOI:10.17759/vp.1991020095 (In Russ.).
11. Lupenko E.A., Korolkova O.A. Semanticheskaya kvivalentnost' kak osnova intermodal'noi integratsii. *Ekspperimental'naya psikhologiya = Experimental Psychology*, 2022. Vol. 15, no. 3, pp. 159–177. DOI:10.17759/exppsy.20221503159 (In Russ.).
12. Morozov V.P. Neverbal'naya kommunikatsiya: kspperimental'no-psikhologicheskije issledovaniya: monografiya. Moscow: Institut psikhologii RAN, 2011. 528 p. DOI:10.17759/rusbook.20110528 (In Russ.).
13. Potapova R.K., Potapov V.V., Komalova L.R. Vospriyatie multimodal'noi mono- i polietnicheskoi kommunikatsii: monografiya / Ed. L.R. Komalova. Moscow: INION RAN, 2020. 210 p. DOI:10.17759/inion.2020210 (In Russ.).
14. Yureva M.V., Khoze E.G. Sozdanie klassifikatora vokal'nykh emocional'nykh ekspressii: personalizirovannye bazy dannykh. *Lico cheloveka v kontekstakh prirody, tekhnologii i kul'tury = The Human Face in Contexts of Nature, Technology, and Culture* / Eds. K.I. Ananyeva, V.A. Barabanshchikov. Moscow: Kogito-Tsentr; Moskovskii institute psikhoanaliza, 2020. Pp. 364–372. DOI:10.17759/kogito.2020364372 (In Russ.).
15. Baart M., Vroomen J. Recalibration of Vocal Affect by a Dynamic Face. *Experimental Brain Research*, 2018. Vol. 236, no. 1, pp. 1911–1918. DOI:10.1007/s00221-018-5270-5
16. Barkhuysen P., Kraemer E., Swerts M. Crossmodal and Incremental Perception of Audiovisual Cues to Emotional Speech. *Language and Speech*, 2010. Vol. 53, no. 1, pp. 3–30. DOI:10.1177/0023830910362212
17. Bänziger T., Mortillaro M., Scherer K.R. Introducing the Geneva Multimodal Expression Corpus for Experimental Research on Emotion Perception. *Emotion*, 2012. Vol. 12, no. 5, pp. 1161–1179. DOI:10.1037/a0025827
18. Bänziger T., Patel S., Scherer K.R. The Role of Perceived Voice and Speech Characteristics in Vocal Emotion Communication. *Journal of Nonverbal Behavior*, 2014. Vol. 38, no. 1, pp. 31–52. DOI:10.1007/s10919-013-0165-x
19. Bryant G.A. The Evolution of Human Vocal Emotion. *Emotion Review*, 2021. Vol. 13, no. 1, pp. 25–33. DOI:10.1177/1754073919870985
20. Chronaki G., Hadwin J.A., Garner M., Maurage P., Sonuga-Barke E.J.S. The Development of Emotion Recognition from Facial Expressions and Non-Linguistic Vocalizations During Childhood. *The British Journal of Developmental Psychology*, 2015. Vol. 33, no. 1, pp. 218–236. DOI:10.1111/bjdp.12075



21. Cortes D.S., Tornberg C., B nziger T., Elfenbein H.A., Fischer H., Laukka P. Effects of Aging on Emotion Recognition from Dynamic Multimodal Expressions and Vocalizations. *Scientific Reports*, 2021. Vol. 11, no. 1, pp. 1–12. DOI:10.1038/s41598-021-94993-4
22. Cowen A.S., Elfenbein H.A., Laukka P., Keltner D. Mapping 24 Emotions Conveyed by Brief Human Vocalization. *The American Psychologist*, 2019. Vol. 74, no. 1, pp. 698–712. DOI:10.1037/amp0000399
23. Dolan R.J., Morris J.S., de Gelder B. Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. Vol. 98, no. 17, pp. 9465–9470. DOI:10.1073/pnas.171288598
24. El Ayadi M., Kamel M.S., Karray F. Survey on speech emotion recognition: Features, classification schemes and databases. *Pattern Recognition*, 2011. Vol. 44, No. 3, pp. 572–587. DOI:10.1016/j.patcog.2010.09.014
25. Gelder B. De, Vroomen J. The Perception of Emotions by Ear and by Eye. *Cognition & Emotion*, 2000. Vol. 14, no. 3, pp. 289–311. DOI:10.1080/026999300378824
26. Jürgens R., Grass A., Drolet M., Fischer J. Effect of Acting Experience on Emotion Expression and Recognition in Voice: Non-Actors Provide Better Stimuli Than Expected. *Journal of Nonverbal Behavior*, 2015. Vol. 39, no. 1, pp. 195–214. DOI:10.1007/s10919-014-0201-6
27. Lambrecht L., Kreifelts B., Wildgruber D. Age-Related Decrease in Recognition of Emotional Facial and Prosodic Expressions. *Emotion*, 2012. Vol. 12, no. 1, pp. 529–539. DOI:10.1037/a0026827
28. Lausen A., Schacht A. Gender Differences in the Recognition of Vocal Emotions. *Frontiers in Psychology*, 2018. Vol. 9, no. 1, Article 882. DOI:10.3389/fpsyg.2018.00882
29. Patel S., Scherer K.R., Björkner E., Sundberg J. Mapping Emotions into Acoustic Space: The Role of Voice Production. *Biological Psychology*, 2011. Vol. 87, no. 1, pp. 93–98. DOI:10.1016/j.biopsycho.2011.02.010
30. Paulmann S., Pell M.D. Is There an Advantage for Recognizing Multi-Modal Emotional Stimuli? *Motivation and Emotion*, 2011. Vol. 35, no. 1, pp. 192–201. DOI:10.1007/s11031-011-9211-y
31. Pittam J., Scherer K.R. Vocal Expression and Communication of Emotion / In Lewis M., Haviland J.M. (Eds.). *Handbook of Emotions*. The Guilford Press. 1993. Pp. 185–197.
32. Sauter D.A., Eisner F., Calder A., Scott S. Perceptual Cues in Nonverbal Vocal Expressions of Emotion. *Quarterly Journal of Experimental Psychology*, 2010. Vol. 63, no. 1, pp. 2251–2272. DOI:10.1080/17470211003721642
33. Scherer K.R. Vocal Affect Expression: A Review and a Model for Future Research. *Psychological Bulletin*, 1986. Vol. 99, no. 2, pp. 143–165. DOI:10.1037/0033-2909.99.2.143
34. Scherer K.R. What Are Emotions? And How Can They Be Measured. *Social Science Information*, 2005. Vol. 44, no. 1, pp. 695–729. DOI:10.1177/0539018405058216
35. Scherer K., Banse R. Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, 1996. Vol. 70, no. 1, pp. 614–636. DOI:10.1037/0022-3514.70.3.614
36. Scherer K.R., Banse R., Wallbott H.G., Goldbeck T. Vocal Cues in Emotion Encoding and Decoding. *Motivation and Emotion*, 1991. Vol. 15, no. 2, pp. 123–148. DOI:10.1007/BF00995674
37. Schirmer A., Adolphs R. Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence. *Trends in Cognitive Sciences*, 2017. Vol. 21, no. 1, pp. 216–228. DOI:10.1016/j.tics.2017.01.001
38. Schlegel K., Scherer K.R. The nomological network of emotion knowledge and emotion understanding in adults: evidence from two new performance-based tests. *Cognition and Emotion*, 2017. Vol. 32, pp. 1514–1530. DOI:10.1080/02699931.2017.1414687
39. Schlegel K., Grandjean D., Scherer K.R. Emotion Recognition: Unidimensional Ability or a Set of Modality- and Emotion-Specific Skills? *Personality and Individual Differences*, 2012. Vol. 53, no. 1, pp. 16–21. DOI:10.1016/j.paid.2012.01.026
40. Schlegel K., Grandjean D., Scherer K. R. Introducing the Geneva Emotion Recognition Test: An Example of Rasch-Based Test Development. *Psychological Assessment*, 2014. Vol. 26, no. 2, pp. 666–672. DOI:10.1037/a0025827
41. Sen A., Isaacowitz D., Schirmer A. Age Differences in Vocal Emotion Perception: On the Role of Speaker Age and Listener Sex. *Cognition & Emotion*, 2018. Vol. 32, no. 1, pp. 1189–1204. DOI:10.1080/02699931.2017.1390725
42. Williams C., Stevens K. Emotions and speech: some acoustical correlates. *The Journal of the Acoustical Society of America*, 1972. Vol. 52, no. 4(2), pp. 1238–1250. DOI:10.1121/1.1913238



43. Wilting J., Krahmer E., Swerts M. Real vs acted emotional speech. *Proceedings of Interspeech 2006*. Pittsburgh, 2006. DOI:10.21437/Interspeech.2006-250

Информация об авторах

Барабанчиков Владимир Александрович, доктор психологических наук, профессор, член-корреспондент РАО, директор Института экспериментальной психологии, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: vladimir.barabanschikov@gmail.com

Суворова Екатерина Владимировна, научный сотрудник Института экспериментальной психологии, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0001-8834-2037>, e-mail: esresearch@yandex.ru

Малионok Арина Витальевна, магистр психологии, Институт экспериментальной психологии, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0009-0001-7375-4343>, e-mail: arinamalionok333@gmail.com

Information about the authors

Vladimir A. Barabanschikov, Dr.Sci. in Psychology, Professor, Corresponding Member of Russian Academy of Education, Director, Institute of Experimental Psychology, Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: vladimir.barabanschikov@gmail.com

Ekaterina V. Suvorova, Research Scientist, Institute of Experimental Psychology, Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0001-8834-2037>, e-mail: esresearch@yandex.ru

Arina V. Malionok, Master in Psychology, Institute of Experimental Psychology, State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0009-0001-7375-4343>, e-mail: arinamalionok333@gmail.com

Получена 29.07.2024

Received 29.07.2024

Принята в печать 01.09.2024

Accepted 01.09.2024