

RU

Просодические маркеры состояния «радость» (на примере языковых реализаций кабардинцев-билингвов в регионе Кавказские Минеральные Воды)

Фролова А. В.

Аннотация. Настоящее исследование посвящено изучению этноспецифических особенностей выражения эмоций в речи представителей кабардинской и русской этногрупп региона Кавказские Минеральные Воды на материале экспериментального корпуса квазиспонтанных аудиофайлов-диалогов, маркированных эмоциональным состоянием «радость», записанных в исполнении представителей русской и кабардинской этногрупп. Цель исследования – выявить национально-специфические просодические маркеры состояния «радость» на материале русской и кабардинской этногрупп региона Кавказские Минеральные Воды. Научная новизна исследования состоит в выявлении этнообусловленных признаков, манифестирующих различие эмоционально-маркированных моделей коммуникации в речи северокавказских билингвов. Акустический анализ материала показал различия в звуковых характеристиках речи между русским и интерферированным вариантами, такие как увеличение интервалов повышения и понижения тона, расширение тонального диапазона фразы и изменения в формантных значениях гласных. Полученные результаты доказывают, что как формантная картина исследуемых реплик в реализации кабардинской и русской этногрупп, так и данные уровня фразовой просодии являются этнообусловленными, и эмоционально-маркированные модели коммуникации различаются по качественным признакам.

EN

Prosodic markers of the state of "joy" (using the example of linguistic realizations of Kabardian Bilinguals in the Caucasian Mineral Waters region)

A. V. Frolova

Abstract. The present study is devoted to the study of ethnospecific features of the expression of emotions in the speech of representatives of Kabardian and Russian ethnogroups of the Caucasian Mineral Waters region based on the material of an experimental corpus of quasi-spontaneous audio files-dialogues marked with the emotional state "joy", recorded by representatives of Russian and Kabardian ethnogroups. The aim of the study is to identify nationally specific prosodic markers of the state of "joy" based on the material of the Russian and Kabardian ethnogroups of the Caucasian Mineral Waters region. The scientific novelty of the study consists in identifying ethnically conditioned signs that manifest the difference in emotionally marked communication models in the speech of North Caucasian bilinguals. Acoustic analysis of the material showed differences in the sound characteristics of speech between the Russian and the interfered variants, such as an increase in the intervals of raising and lowering the tone, an expansion of the tonal range of the phrase and changes in the formant values of vowels. The results obtained prove that both the formant picture of the studied replicas in the implementation of Kabardian and Russian ethnogroups and the data of the level of phrasal prosody are ethnically conditioned, and emotionally marked communication models differ in qualitative characteristics.

Введение

Эмоции играют важную роль в жизни человека. Межличностное общение человека включает в себя не только разговорный язык, но и невербальные сигналы, такие как жесты рук и тела, тон голоса, которые используются для выражения чувств и обратной связи. В современных исследованиях для распознавания

состояний по голосовой информации, как правило, выбираются разнообразные векторы голосовых характеристик, в которых рассчитываются различные статистические показатели, например, среднее значение, стандартное отклонение, максимум и минимум высоты тона и энергии в высказываниях являются в основном используемыми признаками в этом отношении (Lee, Narayanan, 2005; Kun, Dong, Tashev, 2014, p. 223-227).

Востребованность данного направления исследований очевидна и несомненно актуальна, например, для систем бронирования билетов, использующих автоматическое распознавание речи, способное идентифицировать раздражение или гнев пользователя, аналогичные сценарии представлены как для приложений колл-центров (Liyanage, Morrison, Wang, 2007), так и в качестве диагностического инструмента в медицине (France, Shivi, Silverman et. al., 2000, p. 829-837). Кроме того, развитие технологий AI и NLP требует учета культурно-специфических особенностей речи для создания более адаптированных и интуитивно понятных интерфейсов. Таким образом, актуальность темы данного исследования обусловлена необходимостью изучения этноспецифических особенностей выражения эмоций в речи представителей различных языковых и культурных сообществ, так как, несмотря на универсальность эмоций, их понимание и интерпретация являются культурно-обусловленными, что особенно важно для многоязычного региона Северного Кавказа (Goncharova, Zavrumov, Frolova, 2024, p. 520-535; Hévrová, 2022; Huseynli, 2022; Pan, Moisis, Styles, 2023; Васильева, 2022).

В рамках настоящего исследования решались следующие задачи:

- создать корпус речевых данных;
- провести аудитивный анализ для определения естественности и однозначной интерпретируемости эмоционального состояния, передаваемого голосом, без учета текста сообщения;
- изучить относительные значения частоты основного тона (ч.о.т.) на разных этапах фразы (начало, первый ударный слог, конец фразы), а также изменения тонального диапазона и регистра;
- сравнить интервалы повышения и понижения тона между русскими и интерферированными реализациями;
- провести формантный анализ гласных, включающий измерение и анализ значений первой (F1) и второй (F2) формант гласных звуков, нормализацию значений формант для сравнительного анализа, визуализацию показателей формант для наглядного сравнения мест артикуляции гласных в обеих языковых реализациях;
- проанализировать длительности гласных: измерить длительность гласных звуков и сравнить между языковыми вариантами;
- исследовать дисперсию, стандартное отклонение и интенсивность звука (громкости) в речи, включая ядерные и первые ударные слоги, а также общую среднеслоговую интенсивность фраз.

Для реализации задач использовалось программное обеспечение Praat 5.3.32 для получения значений частоты основного тона, тонального диапазона и значений первой (F1) и второй (F2) формант гласных звуков, статистический метод для проведения количественного и корреляционного сравнения полученных показателей, в частности метод нормализации z-score, направленный на выравнивание значений показателей и устранение индивидуальных различий информантов, реализованные посредством скрипта на языке Python, разработанный сотрудниками научно-образовательного центра «Интеллектуальный анализ данных» Пятигорского государственного университета, к полученным данным применялся сравнительно-сопоставительный метод для выявления специфических этнообусловленных просодических и акустических маркеров эмоционального состояния «радость» в речи билингвов кабардинской этногруппы по сравнению с русскоязычными носителями.

Материалом настоящего исследования послужил экспериментальный корпус, сформированный сотрудниками НОЦ «Интеллектуальный анализ данных» ФГБУ ВО Пятигорский государственный университет в лабораторных условиях; полученные файлы доступны по запросу автора и хранятся в базе данных научно-образовательного центра. Корпус состоял из квазиспонтанных аудиофайлов-диалогов, маркированных эмоциональным состоянием «радость», записанных в исполнении проживающих в Кавказских Минеральных Водах представителей русской и кабардинской этногрупп, возраст которых варьировал в диапазоне от 18 до 45 лет. Дикторы участвовали в записи диалогов в ассиметричном формате: инициальная эмоционально-маркированная реплика-стимул в состоянии «радость», произнесенная женщиной / ответная реплика, произнесенная мужчиной, в реализации представителей кабардинской этногруппы; инициальная эмоционально-маркированная реплика-стимул в состоянии «радость», произнесенная женщиной / ответная реплика, произнесенная женщиной, в реализации представителей кабардинской этногруппы; инициальная эмоционально-маркированная реплика-стимул в состоянии «радость», произнесенная женщиной / ответная реплика, произнесенная женщиной, в реализации представителей русской этногруппы; инициальная эмоционально-маркированная реплика-стимул в состоянии «радость», произнесенная женщиной / ответная реплика, произнесенная женщиной, в реализации представителей русской этногруппы. Например: «Можешь меня поздравить!!! – Ура!! Вы наконец-то выиграли!!!».

Теоретической базой настоящего исследования служат работы отечественных и зарубежных фонетистов, изучавших просодические явления в области интерференции как на сегментном, так и на супraseгментном уровнях (Метлюк, 1986; Вишневецкая, 1993; Фомиченко, 2005, 2013), труды ученых пятигорской фонетической школы в области интерференции и фоновариантов языка (Дубовский, Заграевская, 2014; Гончарова О. В., 2008, 2012, 2014), распознаванию эмоций в речи (Потапова, Потапов, 2006, 2012).

Проведенное исследование имеет практическую значимость, так как информация о специфических акустических параметрах речи может быть использована для улучшения алгоритмов распознавания эмоций в речи, что актуально для разработки более чувствительных к контексту голосовых ассистентов и чат-ботов. Также

данные о том, как различные эмоции проявляются в акустических параметрах речи, могут быть использованы для создания более естественно звучащего синтезированного голоса, что улучшит пользовательский опыт при взаимодействии с AI и голосовыми интерфейсами. Кроме того, материалы исследования могут найти применение в вузах гуманитарного направления при изучении спецкурсов и спецсеминаров по фонетике, социолингвистике, межкультурной коммуникации.

Обсуждение и результаты

Вопрос влияния билингвального акцента на коммуникативный эффект высказывания уже не раз попадал в фокус внимания исследователей. Так, Г. М. Вишневская отмечает, что он крайне важен для предупреждения «коммуникативных неудач», но, к сожалению, несмотря на свою «массовость», пока недостаточно глубоко изучен (1993). Значимый вклад в описание особенностей взаимодействия просодических систем в речи билингва был сделан в работах А. А. Метлюк, заложившей теоретическую основу и предложившей методологические принципы изучения просодической интерференции (1986). С нашей точки зрения, наиболее значимых и интересных результатов в анализе феномена билингвальности и интерференции добились ученые пятигорской фонетической школы. Ю. А. Дубовский считал, что решением проблемы исследования просодической интерференции является выявление специфических просодических средств, «имеющих отклонения от эталонной речи в условиях... билингвизма» (Дубовский, Заграевская, 2014, с. 657).

Что касается фонетико-фонологических параметров речи билингвов-представителей кабардинской этногруппы, то нельзя не отметить результаты исследования О. В. Гончаровой, которая обнаружила, что для зон наиболее яркого проявления «кабардинского» акцента, а именно проминантных слогов, характерна волновая восходяще-нисходящая конфигурация тона, большая темпоральная растяженность и повышенная динамическая выделенность, показатели которых могут возрастать при усилении императивности, снижении официальности коммуникативной ситуации и с изменением социального статуса билингва (Гончарова, 2008; Гончарова, Заврумов, Левит, 2023).

Безусловно, появление новых программных инструментов анализа звучащей речи и развитие технологий AI и NLP открыли новые возможности в изучении просодической интерференции и билингвального акцента (Гончарова, Фролова, 2023, с. 3267). Это определило постановку таких задач нашего исследования, как изучение относительных значений частоты основного тона и изменений тонального диапазона и регистра, сравнение интервалов повышения и понижения тона, проведение формантного анализа гласных, нормализации и визуализации значений формант, анализ длительности гласных, исследование дисперсии, стандартного отклонения, интенсивности звука и среднеслоговой интенсивности фраз.

После создания корпуса был проведен аудитивный анализ материала, направленный на оценку естественности и однозначной интерпретируемости воспроизводимого эмоционального состояния. Важно отметить, что эмоциональную окраску предлагалось определять, опираясь исключительно на звуковую дорожку без учета произнесённого в ней текста. Итоговое количество экземпляров речи составило 500 звуковых файлов, по 250 реализаций для каждого варианта – русского и интерферированного.

Таблица 1. Относительные значения частоты основного тона в русском и интерферированном вариантах

id	per	НФ/ПУ	ПУ/КФ	НФ/КФ	rising	falling	stressed	end	start	ТД
Каб	ВШвв	Мин/0.8	Мин/1.1	Мин/0.9	Уз/26	СУ/34	СВ/64	СВ/56	СВ/52	Мал/11
Rus	ВШ	Мин/1.2	Мин/1.3	Ср/1.6	Мал/16	Уз/30	СН/50	СН/37	СВ/58	Мал/15

Условные обозначения в Таблице 1:

Id – идентификатор записи

Per – регистр

НФ – начало фразы

ПУ – первый ударный

КФ – конец фразы / соотношение по частоте основного тона

Rising – интервал понижения ч.о.т. (ядерный слог)

Falling – интервал повышения ч.о.т. (заядерный слог)

ТД – тональный диапазон

Stressed – уровень ч.о.т. первого ударного слога

End – уровень ч.о.т. конца фразы

Start – уровень ч.о.т. начала фразы

Исходя из полученных данных можно сделать вывод, что в интерферированных реализациях отмечается увеличение интервалов повышения (от малого в русской реализации к узкому в интерферированной) и понижения тона (от узкого в русской реализации к суженному в интерферированной) и незначительное расширение тонального диапазона фразы. Общий тональный регистр интерферированных реализаций выше, тональные контрасты более однородные, сопоставительно с русскими, уровни начала фразы идентичны в обеих реализациях, однако уровень завершения фразы более высокий в интерферированной, что указывает на восходящее движение тона в завершении фразы и, соответственно, большую выделенность заядерной части в реализации кабардинских билингвов (см. Таблицу 1).

Таблица 2. Формантные значения гласных в русском и интерферированном вариантах

id	Vowel	Duration	F1	F2	Norm_F1	Norm_F2
Kab	o	0.044	1001.158	1986.550	0.535	0.352
Kab	i	0.031	1523.852	2774.654	2.027	2.037
Kab	i	0.035	632.659	1853.512	-0.516	0.068
Kab	a	0.059	773.874	1541.182	-0.113	-0.599
Kab	a	0.063	582.218	1422.731	-0.660	-0.852
Kab	a	0.115	658.248	1467.931	-0.443	-0.756
Kab	i	0.049	522.498	1704.814	-0.830	-0.249
Rus	o	0.103	415.477	962.131	-0.855	-1.517
Rus	i	0.062	544.875	1772.202	-0.277	0.015
Rus	i	0.036	373.532	2150.349	-1.043	0.731
Rus	a	0.034	750.413	1642.535	0.6404	-0.229
Rus	a	0.086	861.841	1673.376	1.1382	-0.171
Rus	a	0.217	892.562	1495.301	1.275	-0.508
Rus	i	0.055	410.675	2651.541	-0.877	1.679

Условные обозначения в Таблице 2:

id – идентификатор записи

Vowel – гласный

Duration – длительность гласного

F1 – значение первой форманты гласного

F2 – значение второй форманты гласного

Norm_F1 – нормализованное значение первой форманты гласного

Norm_F2 – нормализованное значение второй форманты гласного

На основании анализа формантных значений и при визуализации показателей формант обеих языковых реализаций, прежде всего, можно сделать вывод о различном стандартном отклонении значений гласных: стандартное отклонение значений гласных в русских реализациях: F1 = 245.05, F2 = 285.12, в реализациях билингов: F1 = 157.96, F2 = 254.28.

Также существенной является разница в месте артикуляции:

– для гласной “i” мы видим, что кабардинские звуки имеют широкий диапазон значений F1 и F2, что указывает на большую изменчивость в произношении по сравнению с русскими гласными “i”, которые имеют высокие значения F2, указывающие на их переднее положение;

– гласная “a” в кабардинском языке демонстрирует более низкие значения F1 и F2 по сравнению с русским, что свидетельствует о более открытом и заднем положении артикуляции (см. Рис. 1).

Кабардинские гласные в целом имеют более длительную продолжительность по сравнению с русскими. Это может указывать на различия в ритмической структуре этих языков, где кабардинский язык может быть более склонен к использованию длительности как дистинктивного признака (см. Таблицу 2).

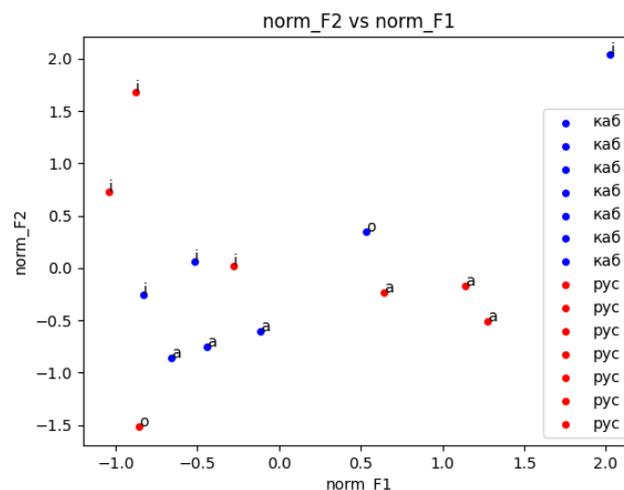


Рисунок 1. График нормализованных значений формант гласных в русском и интерферированном вариантах

Динамические показатели в речи билингов демонстрируют большую дисперсию и стандартное отклонение, что создает эффект больших перепадов громкости, значения интенсивности ядерного, первого ударного слога; общая среднеслоговая интенсивность фраз – высокая в обеих реализациях, однако числовые показатели в речи билингов немного превосходят русские реализации (см. Таблицу 3).

Таблица 3. Относительные значения интенсивности в русском и интерферируемом вариантах

id	std_dev	disp	nuclear_int	stressed_int	mean_int
Rus	3.486	12.15	75/В	71/В	71/В
Kab	4.471	19.99	80/В	78/В	74/В

Условные обозначения в Таблице 3:

id – идентификатор записи

std_dev – стандартное отклонение

disp – дисперсия

nuclear_int – интенсивность ядерного слога

stressed_int – интенсивность первого ударного слога

mean_int – среднеслоговая интенсивность

Таким образом, проведенный акустический анализ показал, что эмоционально-маркированные модели коммуникации представителей кабардинской и русской этногрупп региона Кавказские Минеральные Воды различаются по ряду просодических и сегментных характеристик, что подтверждает наличие этнообусловленных параметров в выражении эмоций.

Заключение

Проведенные расчеты позволили сделать ряд выводов, а также выявить некоторые закономерности и специфические черты в акустических параметрах эмоционально-маркированной коммуникации представителей кабардинской и русской этногрупп региона КМВ. В частности, выяснено, что различия в интервалах повышения и понижения тона между русскими и интерферируемыми реализациями, а также разница в относительных значениях частоты основного тона (ч.о.т.), тонального диапазона и регистра приводят к появлению следующих просодических этнокультурных маркеров в речи билингов: 1) большая тональная вариативность; 2) большая тональная выделенность заядерной части; 3) эффект большего перепада громкости сравнительно с русскими реализациями. Анализ значений первой и второй формант гласных, и последующие статистические вычисления показали, что кабардинские гласные звуки демонстрируют различия в артикуляции по сравнению с русскими, особенно заметные в гласных “о” и “а”, где реализация гласных звуков в кабардинском языке демонстрирует более открытую и продвинутую назад артикуляцию по сравнению с русскими, а также, как правило, большую длительность. Таким образом, как формантная картина исследуемых гласных в реализации кабардинской и русской этногрупп, так и показатели тонального уровня фразовой просодии показывают, что эмоционально-маркированные модели коммуникации различаются по качественным признакам, что подтверждает наличие этнообусловленных параметров.

Результаты исследования в перспективе могут быть использованы для создания адаптивных систем распознавания и синтеза речи, способных динамически корректировать свои параметры в зависимости от культурного контекста пользователя, а также полученные данные могут быть использованы для обучения мультязычных и мультикультурных моделей, способных распознавать и правильно интерпретировать эмоциональную окраску речи на различных языках и в различных культурных контекстах.

Перспективы дальнейшего исследования связаны с расширением корпуса речевых данных, включением в анализ других эмоциональных состояний, а также изучением взаимосвязи между просодическими и сегментными характеристиками речи в условиях межэтнического взаимодействия.

Источники | References

1. Васильева А. В. Восприятие эмоциональной лексики первого и второго языков в условиях языкового контактирования (тюркско-русский херитажный билингвизм) // Русин. 2022. № 70.
2. Вишневская Г. М. Интерференция и акцент (на материале интонационных ошибок при изучении неродного языка): автореф. дисс. к. филол. н. СПб., 1993.
3. Гончарова О. В. Историко-социальные основы формирования южнорусского фоноварианта // Казанская наука. 2012. № 12.
4. Гончарова О. В. Просодия русского побуждения в условиях кабардино-черкесской интерференции (экспериментально-фонетическое исследование): дисс. ... к. филол. н. Пятигорск, 2008.
5. Гончарова О. В. Южнорусский фонорегиолект: исходные посылки // Филологические науки. Вопросы теории и практики. 2014. № 1-2 (31).
6. Гончарова О. В., Заврумов З. А., Левит А. А. Акустические маркеры состояния «гнев» в речи представителей русской и армянской этногрупп региона КМВ // Казанская наука. 2023. № 10.
7. Гончарова О. В., Фролова А. В., Просодические маркеры состояния «гнев» (на примере культурно-коммуникативных моделей общения русской и кабардинской этногрупп региона Кавказские Минеральные Воды) // Филологические науки. Вопросы теории и практики. 2023. Т. 16. № 10.

8. Дубовский Ю. А., Заграевская Т. Б. Просодическая интерференция в условиях искусственного билингвизма (накопление экспериментальных данных для целей лингводидактики конца XX в.) // Известия Самарского научного центра Российской академии наук. 2014. № 2-3.
9. Метлюк А. А. Взаимодействие просодических систем в речи билингва Минск: Вышэйша школа, 1986.
10. Потапова Р. К., Потапов В. В. Речевая коммуникация: от звука к высказыванию. М.: Яз. славян. культур, 2012.
11. Потапова Р. К., Потапов В. В. Язык, речь, личность. М.: Яз. славян. культуры, 2006.
12. Фомиченко Л. Г. Когнитивные основы просодической интерференции. Волгоград: Изд-во ВолГУ, 2005.
13. Фомиченко Л. Г. Психолингвистические основы исследования речевой экспрессии при билингвизме // Вестник Волгоградского государственного университета. Серия 2: Языкознание. 2013. № 3.
14. France D., Shiavi R., Silverman S., Silverman M., Wilkes D. Acoustical properties of speech as indicators of depression and suicidal risk // IEEE Trans. Biomed. Eng. 2000. № 47.
15. Goncharova O. V., Zavrurnov Z. A., Frolova A. V. Emotive Expression In Political Discourse: Analyzing Gender And Social Influences On Speech In The Russian And Kabardian Ethnic Groups Of The Caucasian Mineralnye Vody Region // Journal of Law and Political Sciences. 2024. Vol. 40.
16. Hévrová M. Phonetic attrition and cross-linguistic influence in L1 speech of late Czech-French bilinguals. Linguistics. Université Toulouse le Mirail – Toulouse II; Univerzita Karlova (Prague). 2022. URL: <https://theses.hal.science/tel-03603032/document>
17. Huseynli S. The impact of the Azerbaijani language vocalism on Baku residents Russian pronunciation: the bilingual aspect // Journal of Philology. 2022. Vol. 2.
18. Kun H., Dong Yu., Tashev I. Speech emotion recognition using deep neural network and extreme learning machine // INTERSPEECH. 2014. September.
19. Lee Ch., Narayanan S. Toward Detecting Emotions in Spoken Dialogs // Transactions on speech and audio processing. 2005. Vol. 13. № 2.
20. Liyanage C., Morrison D., Wang R. Ensemble methods for spoken emotion recognition in call-centres // Speech Communication. 2007. Vol. 49. Iss. 2.
21. Pan L., Moiskis, S., Styles S. Being a round /y/: An acoustic description of high front vowels in Singapore Mandarin elicited by speakers with different bilingual balance in Mandarin and English. 2023. URL: <https://osf.io/preprints/psyarxiv/rdc6f>

Финансирование | Funding

RU Публикация подготовлена в рамках поддержанного РФФИ и Министерством образования Ставропольского края научного проекта № 23-28-10124 «Квантитативно-статистическая модель анализа эмоционально-маркированной коммуникации в условиях межэтнических взаимодействий в регионе Кавказские Минеральные Воды».

EN The publication was prepared within the framework of the scientific project No. 23-28-10124 supported by the Russian Academy of Sciences and the Ministry of Education of the Stavropol Territory "Quantitative Statistical Model for the Analysis of Emotionally Marked Communication in the Context of Interethnic Interactions in the Caucasian Mineral Waters region".

Информация об авторах | Author information

RU **Фролова Анастасия Вадимовна**¹
¹ Пятигорский государственный университет

EN **Anastasia Vadimovna Frolova**¹
¹ Pyatigorsk State University

¹ frolova-nastya19@yandex.ru

Информация о статье | About this article

Дата поступления рукописи (received): 18.04.2024; опубликовано online (published online): 04.06.2024.

Ключевые слова (keywords): фразовая просодия; формантные значения гласных; акустический анализ материала; этнообусловленные параметры; эмоционально маркированные модели коммуникации; phrasal prosody; formant values of vowels; acoustic analysis of the material; ethnically conditioned parameters; emotionally marked communication models.