

RU

Логистическое моделирование региональной вариативности
русских гласных

Гончарова О. В.

Аннотация. Статья посвящена исследованию акустических и просодико-структурных особенностей русских гласных в южнорусском и московском фонетических вариантах. Цель исследования – выявить ключевые признаки, обеспечивающие различие южнорусского и московского фонетических вариантов с применением методов машинного обучения. Исследование, основанное на материале чтения нейтральных фраз носителями из Москвы, Московской области и юга России, включало комплексный анализ речевых данных: аннотирование, акустические измерения, построение модели логистической регрессии и оценку перестановочной важности признаков. Научная новизна заключается в количественной оценке вклада различных параметров речи – темпоральных, спектральных и структурных – в задачу бинарной классификации региональных вариантов произношения. В результате установлено, что основным дифференциальным признаком выступают темпоральные характеристики гласных: абсолютная длительность значимо связана с московским вариантом, тогда как относительная длительность гласного в структуре сегмента ассоциирована с южнорусским вариантом. Спектральные параметры продемонстрировали минимальный дифференцирующий потенциал.

EN

Logistic modeling of regional variability in Russian vowels

O. V. Goncharova

Abstract. This paper investigates the acoustic and prosodic-structural features of Russian vowels in South Russian and Moscow phonetic variants. The aim of the study is to identify key features that distinguish South Russian and Moscow phonetic variants using machine learning methods. The research, based on the reading of neutral phrases by native speakers from Moscow, Moscow Oblast, and Southern Russia, involved a comprehensive analysis of speech data: annotation, acoustic measurements, construction of a logistic regression model, and assessment of feature permutation importance. Scientific novelty lies in the quantitative assessment of the contribution of various speech parameters – temporal, spectral, and structural – to the task of binary classification of regional pronunciation variants. The study revealed that the main differentiating feature is the temporal characteristics of vowels: absolute duration is significantly associated with the Moscow variant, whereas the relative duration of the vowel within the segment structure is associated with the South Russian variant. Spectral parameters demonstrated minimal differentiating potential.

Введение

Исследование региональных вариантов русского языка сохраняет свою актуальность в свете продолжающегося интереса к проблемам вариативности в фонетике и необходимости разработки объективных методов их описания. В то же время для современной теории литературного языка и орфоэпии важно опираться не только на констатацию наличия региональных вариантов, но и на количественно описанные параметры их различий, чтобы выработать объективные критерии разграничения допустимой вариативности и отклонений от кодифицированной нормы. Отсутствие таких критериев затрудняет уточнение границ литературного произношения, усложняет разработку единых орфоэпических рекомендаций и интерпретацию данных больших корпусов речи. В существующих работах по русской диалектологии и региональной фонетике, включая исследования, описывающие произносительные особенности различных территорий (Норма произношения в узусе..., 2021), основное внимание уделяется качественному описанию вариантов, тогда как систематические проверенные акустико-артикуляционные корреляты, отличающие один региональный вариант от другого, представлены фрагментарно. В условиях активного развития речевых технологий и экспертной практики это создает ситуацию, когда научному и прикладному сообществу не хватает эмпирически обоснованных моделей региональной вариативности.

Потребность в таких моделях проявляется в ряде прикладных направлений: при разработке систем автоматического распознавания и синтеза речи, в фонетической экспертизе, а также в методике преподавания русского языка как иностранного и как неродного. Во всех этих областях необходимо надежное распознавание и интерпретация речевых сигналов, производимых носителями разных регионов, что возможно только при наличии чётко сформулированных и количественно проверенных признаков, связанных с принадлежностью речи к тому или иному региональному варианту. Традиционные методы лингвистического анализа требуют дополнения инструментарием цифровой обработки сигналов и статистического моделирования, который позволяет количественно сопоставлять речевые сигналы и проверять лингвистические гипотезы. В данном контексте обращение к региональной вариативности русских гласных с использованием современных методов анализа речи отвечает потребности современного языкознания в объективных и верифицируемых описаниях произносительной нормы и её региональных модификаций.

В настоящем исследовании данный подход реализуется посредством решения следующих задач: формирование сбалансированного корпуса записей речи носителей из двух регионов; многоуровневое лингвистическое аннотирование речевого материала; акустический анализ, включающий измерение длительностей и формантных характеристик; построение и интерпретация статистической модели, классифицирующей варианты произношения на основе отобранных признаков.

Теоретической базой исследования послужили труды отечественных и зарубежных учёных, описывающих фонетическую вариативность и акустические параметры речи. В частности, положения Н. С. Трубецкого (2000) принимаются за основу при интерпретации представлений о фонологической оппозиции и системном характере корреляций в звуковом строе. Работы Л. В. Златоустовой (1962) и Л. В. Бондарко (1998), посвящённые фонетической структуре слова и артикуляционно-акустическим характеристикам русских гласных, составили основу для выбора анализируемых сегментных и просодических параметров. Подход Л. В. Златоустовой (1962) к описанию фонетической структуры слова в потоке речи лежит в основе анализа слоговой организации высказывания и учёта слоговой структуры при интерпретации длительности и просодической проминантности. При характеристике норм и вариантов современного русского произношения учитываются положения М. Л. Каленчук и Д. М. Савинова (Норма произношения в узусе..., 2021), описывающие соотношение кодифицированной орфоэпической нормы и региональных реализаций и позволяющие разграничить южнорусский и московский варианты как социально признанные разновидности литературного произношения. Акустическая интерпретация полученных данных опирается на акустическую теорию речеобразования Г. Фанта (1964), в рамках которой длительность и формантная структура гласных рассматриваются как параметры, описывающие артикуляционно-акустическую реализацию фонем. При анализе формантных характеристик используются методические принципы исследования Г. Н. Peterson и Н. L. Barney (1952), задающие представление о формантном пространстве гласных и о связи между спектральными координатами и перцептивной дифференциацией. Подходы к цифровой обработке речевого сигнала и формантному анализу, реализованные в программном комплексе Praat (Boersma, Weenink, 2025), лежат в основе используемых в исследовании процедур измерения частоты формант и длительности. Теоретическую основу статистического анализа составляют современные подходы к классификации акустических параметров речи, в число которых входят: оценка произношения на основе машинного обучения (Arkin, Abdukelim, Yilahun et al., 2025), многомерный анализ акустической вариативности гласных для дифференциации диалектов (Tanner, Sonderegger, Stuart-Smith, 2022), а также применение глубоких нейронных сетей для классификации диалектов (Themistocleous, Protosaltis, 2019).

В исследовании использовались речевые материалы, записанные носителями русского языка из Москвы, Московской области и крупных городов юга России. Все информанты были образованными говорящими в возрасте от двадцати двух до тридцати пяти лет. Стимульный материал состоял из десяти нейтральных повествовательных фраз (доступно на <https://github.com/brainteaser-ov/gramota>), которые каждый диктор зачитывал вслух.

Методами исследования выступили последовательно применяемые процедуры сбора и анализа данных. На первом этапе был проведён аудитивный анализ, в ходе которого наивные носители оценивали степень отличия записей от речи своего региона. Для дальнейшего эксперимента были отобраны реализации, воспринимаемые как совпадающие с региональной нормой или незначительно отличающиеся от неё. Затем для балансировки материала были сформированы две равные подвыборки по двести реализаций, представляющие Москву и Московскую область и юг России. Фонетисты с опытом транскрибирования провели аннотирование записей в международной фонетической записи. Аннотирование выполнялось изолированно, для контроля согласованности применялась перекрёстная верификация транскрипций. Использовалась трёхуровневая система аннотации, размечающая каждый слог по его проминантности, позиции в интонационной фразе, слоговой структуре и профилю твёрдости или мягкости согласных (Иллюстрация 1).

Акустические показатели, включая первые три форманты и длительность, извлекались из размеченных интервалов. Для сопоставления характеристик между дикторами значения формант нормализовались с использованием метода, стабилизирующего внутригрупповую вариативность. Фонемная идентификация выполнялась на основе минимального евклидова расстояния в нормализованном пространстве формант. Для анализа динамики гласных строились траектории движения формант, которые описывались через интегральную величину спектрального изменения и направленные изменения по первой и второй формантам. Все вычислительные процедуры были реализованы в рамках единого пайплайна обработки на платформе LingLab AI (доступно на <https://github.com/brainteaser-ov/PySound>).

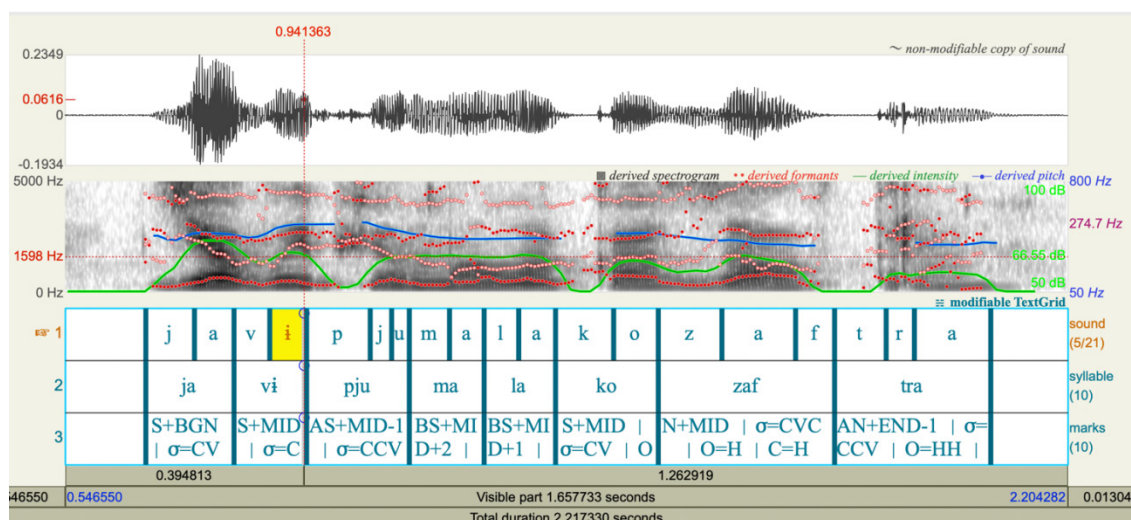


Иллюстрация 1. Пример файла аннотации TextGrid

Теоретическая значимость исследования состоит в разработке и апробации комплексной методологии анализа региональной фонетической вариативности, которая интегрирует методы теоретической фонологии, инструментальной фонетики и Data Science. Полученные результаты вносят вклад в теорию фонетической вариативности, конкретизируя представления о том, какие именно параметры речевого сигнала – темпоральные, спектральные или структурные – несут основную диагностическую нагрузку при различении региональных вариантов.

Практическая ценность работы связана с возможностью применения разработанного пайплайна обработки речевых данных для решения прикладных задач. Методика может быть использована при создании систем автоматической идентификации и верификации диктора, а также в системах оценки и коррекции произношения при изучении русского языка как иностранного или как неродного. Алгоритмы аннотирования и анализа представляют интерес для разработчиков лингвистического программного обеспечения.

Полученные результаты могут быть использованы в процессе преподавания таких дисциплин, как современный русский язык, фонетика и фонология, диалектология, лингвистическая экспертиза речи, а также в курсах, посвящённых применению математических методов и информационных технологий в лингвистике. Материалы и выводы исследования могут служить иллюстративной базой при изучении темы фонетической вариативности и методов акустического анализа речи.

Обсуждение и результаты

Для того чтобы установить, какие именно акустические и структурные признаки позволяют наиболее надёжно различать два региональных варианта произношения – южнорусский (ЮФВ) и московский (МФВ), была применена модель логистической регрессии с L2-регуляризацией (код выполнения доступен на <https://github.com/brainteaser-ov/gramota>), которая оценивала вклад различных предикторов. Постановка задачи классификации сводилась к бинарному различению двух вариантов (ЮФВ/МФВ) по акустическим и слоговым признакам гласных. В качестве числовых предикторов использовались длительность в секундах (duration), длительность сегмента в процентном отношении ко всей длине фразы (percent_duration), нормализованные показатели первой и второй формант (F1z, F2z), дистанция до ближайшего эталона (distance to nearest/margin) и целочисленная дистанция слога до словесного ударения (distStress); числовые переменные предварительно стандартизовались. Категориальные факторы соответствовали принятой слоговой аннотации и включали тип слога по проминантности (role), позиции в пределах интонационной фразы (pos), тип слога (syllable type), профили согласных по твёрдости/мягкости (onset, coda). Для сопоставимости вкладов применялась перестановочная важность: для каждого признака оценивалось снижение качества предсказания при случайной перестановке его значений на фиксированной обученной модели, после чего значения усреднялись по повторам и сопровождалась стандартным отклонением. Направление эффекта для числовых предикторов задавалось знаком соответствующего регрессионного коэффициента, интерпретируемого в шкале логитов после стандартизации. Для факторных групп вычислялись суммарные значения перестановочной важности путём аддитивного объединения важностей отдельных индикаторов внутри каждой группы, что позволяло сопоставлять информативность просодико-структурных параметров с акустическими длительностно-спектральными показателями.

После обучения модели были получены результаты, которые демонстрируют, что темпоральные характеристики гласного оказываются главным носителем дифференциальной информации между вариантами. Признак duration демонстрирует наибольшую перестановочную важность (0.1374; SD=0.0180) и положительное направление эффекта, что указывает на более длительную реализацию гласных в записях МФВ. Показатель

percent_duration занимает второе место по вкладу (0.0847; SD=0.0180) и имеет отрицательное направление, то есть относительно больший вклад гласного в суммарную длительность сегмента ассоциирован с классификацией в сторону ЮФВ. Спектральные координаты в z-нормировке дают минимальные различия: для F1z перестановочная важность невелика (0.0021) при отрицательном направлении, что интерпретируется как слабая тенденция к большей открытости (повышенному F1) в наблюдениях, относимых к ЮФВ; для F2z вклад практически нулевой (0.0001), что говорит об отсутствии существенных межвариантных сдвигов по горизонтальному измерению. Метрика distance to nearest (0.0003) также почти не влияет на классификацию, следовательно, различие вариантов не сводится к систематическому отдалению реализаций от ближайших эталонных облаков в пространстве формант. Параметр distStress имеет по величине близкую к нулю важность (–0.0002), при этом знак коэффициента положительный, что указывает на отсутствие самостоятельного позиционного эффекта дистанции до удара после учёта остальных предикторов. В совокупности эти результаты описывают межвариантную дифференциацию как преимущественно темпоральную: увеличение абсолютной длительности повышает вероятность отнесения к МФВ, тогда как увеличение относительной доли гласного в структуре длительности – индикатор ЮФВ, при слабых и нестабильных спектральных различиях по F1 и отсутствии систематического вклада по F2 (Таблица 1).

Таблица 1. Важность числовых предикторов (mean ± SD) и направление эффекта в задаче классификации ЮФВ/МФВ

Признак	Вклад	SD
duration	0.1374	0.0180
percent_duration	0.0847	0.0180
F1z	0.0021	0.0019
distance to nearest	0.0003	0.0009
F2z	0.0000	0.0020
distStress	–0.0002	0.0023

По факторным группам были рассчитаны суммы перестановочной важности, то есть агрегированные оценки вклада всех индикаторов внутри каждой группы. Среди категориальных параметров наибольшую суммарную информативность демонстрирует группа role (0.0357); далее по убыванию следуют coda (0.0138), pos (0.0123), onset (0.0118) и syllable type (0.0108), что интерпретируется как относительный вклад просодико-позиционных и слоговых характеристик в задачу различения вариантов. Существенно, что доминирующие числовые предикторы – duration (0.1374) и percent_duration (0.0847) – превосходят по абсолютной величине любую из перечисленных категориальных групп, что указывает на первичный статус темпоральных показателей в рассматриваемой классификации и на вспомогательную роль структурно-просодических факторов.

Внутри группы role наибольший вклад дают уровни заударный (AS) и ударный (S), что означает, что распределение токенов по сильным и слабым позициям относительно удара связано с вероятностью отнесения к одному из вариантов. В группе pos фиксируется заметный, но меньший по величине вклад позиционных отметок начала/середины/конца фразы. Группа onset демонстрирует сопоставимый с pos вклад: профиль по твёрдости/мягкости даёт дополнительную информацию, вероятно через коартикуляционные ограничения на траекторию гласной. Группа syllable type вносит умеренный вклад, что указывает на ограниченное, но стабильное влияние слоговой схемы (например, CVC по сравнению с открытыми типами) на распределение длительности и спектральных параметров (Таблица 2).

Таблица 2. Суммарная перестановочная важность по факторным группам и числовым предикторам в модели логистической регрессии

Группа факторов	Суммарный вклад
duration	0.1374
percent_duration	0.0847
role	0.0357
coda	0.0138
pos	0.0123
onset	0.0118
syllable type	0.0108
F1z	0.0021
distance to nearest	0.0003
F2z	0.0001

Сопоставление агрегированных оценок между группами и ведущими числовыми предикторами показывает, что длительность гласного и его относительная доля в структуре сегмента/слова остаются ключевыми индикаторами варианта, а категориальные признаки уточняют предсказание через просодическую роль слога, консонантное окружение и позицию в фразе. Спектральные координаты и метрика ближайшего эталона в суммарной картине играют второстепенную роль.

Заключение

Проведенное исследование позволило сформулировать следующие выводы: 1) различие южнорусского и московского фонетических вариантов возможно на основе ограниченного набора акустических и просодико-структурных параметров, причем ключевую роль играют темпоральные характеристики гласных; 2) абсолютная длительность гласного является наиболее информативным признаком, увеличение которого статистически значимо связано с московским вариантом, тогда как относительная длительность гласного в структуре сегмента ассоциирована с южнорусским вариантом; 3) спектральные характеристики (нормализованные значения F1/F2) и метрика близости к эталонным точкам показали минимальный дифференцирующий потенциал.

Таким образом, разработанный методический подход, сочетающий лингвистическую аннотацию, акустический анализ и методы машинного обучения, подтвердил свою эффективность для решения задач фонетической таксономии и может быть использован при изучении других типов произносительной вариативности. Перспективным представляется включение в анализ большего количества региональных вариантов русского языка, а также исследование консонантизма и просодических характеристик речи. Дальнейшее развитие метода предполагает применение более сложных моделей машинного обучения для выявления нелинейных зависимостей между параметрами речи.

Материалы исследования | Research materials

1. Boersma P., Weenink D. Praat: doing phonetics by computer (Version 6.1.03) (Computer software). 2025. <http://www.praat.org/>

Источники | References

1. Бондарко Л. В. Фонетика современного русского языка: учеб. пособие для студентов вузов. СПб.: Изд-во Санкт-Петербургского университета, 1998.
2. Златоустова Л. В. Фонетическая структура слова в потоке речи. Казань: Изд-во Казанского университета, 1962.
3. Норма произношения в узусе и кодификации: коллективная монография / отв. ред. М. Л. Каленчук, Д. М. Савинов. М., 2021.
4. Трубецкой Н. С. Основы фонологии. М.: Аспект пресс, 2000.
5. Фант Г. Акустическая теория речеобразования / пер. с англ. Л. А. Варшавского и В. И. Медведева; под ред. В. С. Григорьева. М.: Наука, 1964.
6. Arkin G., Abdukelim T., Yilahun H., Hamdulla A. Machine learning-driven acoustic feature classification and pronunciation assessment for Mandarin learners // Applied Sciences. 2025. Vol. 15 (11). <https://doi.org/10.3390/app15116335>
7. Peterson G. H., Barney H. L. Control methods used in a study of the vowels // Journal of the Acoustical Society of America. 1952. Vol. 24 (2). <https://doi.org/10.1121/1.1906875>
8. Tanner J., Sonderegger M., Stuart-Smith J. Multidimensional acoustic variation in vowels across English dialects // Proceedings of the 19th SIGMORPHON Workshop. 2022. <https://doi.org/10.18653/v1/2022.sigmorphon-1.8>
9. Themistocleous C., Protopsaltis P. Dialect classification from a single sonorant sound using deep neural networks // Frontiers in Communication. 2019. Vol. 4. <https://doi.org/10.3389/fcomm.2019.00064>

Информация об авторах | Author information



Гончарова Оксана Владимировна¹, к. филол. н., доц.

¹ Российский университет дружбы народов им. П. Лумумбы, г. Москва



Oxana Vladimirovna Goncharova¹, PhD

¹ Peoples' Friendship University of Russia named after Patrice Lumumba, Moscow

¹ goncharova_oxvl@pfur.ru

Информация о статье | About this article

Дата поступления рукописи (received): 09.11.2025; опубликовано online (published online): 28.11.2025.

Ключевые слова (keywords): региональная вариативность; русские гласные; акустическая фонетика; логистическая регрессия; темпоральные характеристики гласных; regional variability; Russian vowels; acoustic phonetics; logistic regression; temporal characteristics of vowels.