

УДК: 530.182; 57.087.1; 612.789.3

Использование метода нормированного размаха при анализе речевых патологий неврологического генеза

Н. П. Кузенков^а, В. М. Логинов

Институт математики, физики и информатики
Красноярского государственного педагогического университета им. В. П. Астафьева,
Россия, 660049, г. Красноярск, ул. Перенсона, д. 7

E-mail: ^аnk@kspu.ru

Получено 7 мая 2014 г.,
после доработки 22 октября 2014 г.

На основании модификации алгоритма вычисления нормированного размаха предлагается способ определения показателя Хёрста, а также оценки характерного времени, в течение которого показатель Хёрста остается постоянным. Разработан подход к упрощению автоматизации вычисления показателя Хёрста и увеличения его точности. Показатель Хёрста и характерное время вычислены для рядов мощностей речевых сигналов с различными моторными патологиями (афазии, дизартрии). Проведен их статистический анализ, произведена оценка корреляции между показателем Хёрста и характерным временем. Обсуждается возможность использования результатов в диагностике речевых патологий.

Ключевые слова: показатель Хёрста, нелинейная динамика, афазия, дизартрия, рандомизационный тест, диагностика

R/S method application in neurological speech disorders analyses

N. P. Kuzenkov, V. M. Loginov

*Institute of mathematics, physics and informatics of V. P. Astafiev Krasnoyarsk state pedagogical university,
7 Perensona str., Krasnoyarsk, 660049, Russia*

Abstract. — Based on modified rescaled range scale computation algorithm, the technique of Hurst exponent and its characteristic time estimation is proposed. The approach of increase the accuracy and simplification automatic Hurst exponent calculation is developed. The Hurst exponent and characteristic time is calculated for power time sets of speech signals with various motor pathologies (aphasias and dysarthrias). Results is statistically analyzed, the correlation between Hurst exponent and characteristic time is estimated.

Keywords: Hurst exponent, nonlinear dynamics, aphasia, dysarthria, permutation test, diagnostics

Citation: *Computer Research and Modeling*, 2014, vol. 6, no. 5, pp. 775–791 (Russian).

Работа выполнена при частичной поддержке гранта РГНФ 12-06-00256а.

Введение

Речь человека является главным средством коммуникации и играет ключевую роль во всех социально значимых действиях. Кроме того, именно на основе речи формируются и реализуются многие интеллектуальные процессы человека. Диапазон речевых нарушений довольно широк и простирается от едва различимых искажений звукопроизношения до полной утраты всех речевых функций, включая понимание устной и письменной речи [Морозова, 2002; Винарская, 2006; Цветкова, 2007]. При этом возникшее в результате действия каких-либо травмирующих факторов, речевое нарушение обычно не проходит самостоятельно, а требует продолжительной и кропотливой логопедической работы. В силу этих причин реабилитация больных с различными речевыми патологиями имеет большое социальное значение [Цветкова, 2007; Ковальчук и др., 2007; Ястребцева, 2009]. Одним из важных аспектов реабилитации является эффективная и объективная диагностика текущего состояния больного [Малюкова, 2012]. Диагностика должна не только фиксировать факт наличия речевых патологий, но и предоставлять шкалу оценки состояния этих патологий, основанную на результатах инструментальных измерений. Такой подход позволяет своевременно оценивать состояние речевых функций в процессе реабилитации, контролировать ее эффективность и своевременно корректировать в случае необходимости.

Использование инструментальных оценок речевых параметров в совокупности с различными математическими методами применяется при решении широкого круга задач. Помимо медицинской диагностики, также сюда можно отнести задачи идентификации (верификации) диктора и распознавания и генерации речи, имеющие большое прикладное значение [Вьюгин и др., 2012]. Подходы, используемые при решении этих задач, обычно разделяют на две группы. В первой группе строятся модели какой-либо функциональной части речевого аппарата по речевому сигналу через решение обратных задач, например определение конфигурации и размеров резонаторных полостей (носоглотка, гортань), или же через реконструкцию динамических законов для рядов длительностей речевых звуков и пауз [Вьюгин и др., 2012; Сорокин, Гераськин, 2013; Gómez et al., 2005; Herzel, Knudsen, 1995; Henríquez et al., 2009]. Во второй группе акцент делается на формальном применении каких-либо методов анализа (спектр, кепстр, вейвлет-преобразование, функция распределения вероятностей для рядов длительностей речевых звуков и пауз и т. д.) с их последующей статистической обработкой и поиском закономерностей [Сидоров, Филатова, 2012; Голубинский, 2009; Шарий, 2009; Гусев, 2013]. Очевидно, что эти два подхода не являются взаимоисключающими и могут быть объединены в разной мере.

Данная работа придерживается преимущественно второго подхода и посвящена исследованию статистических характеристик показателя Хёрста в группах с речевыми патологиями по типу моторных афазий, нейродинамических афазий и дизартрий у взрослых в сравнении с нормой (отсутствием вышеперечисленных патологий, заикания, а также выраженной дислалии).

Клинический материал

Речь человека представляет собой продукт работы речевого аппарата и специальных структур головного мозга, ответственных за речь. Условно указанные структуры подразделяют на два отдела: центральный и периферический. К центральному отделу относят головной мозг, при этом можно выделить различные участки, реализующие те или иные функции речи. Различные формы речевых патологий возникают как при поражении различных отделов головного мозга, так и при поражении различных областей коры головного мозга. За речь как высшую функцию отвечают кора премоторной, префронтальной областей лобной доли, постцентральная область теменной доли, задне-верхние отделы височной доли и примыкающее к указанным структурам подкорковое белое вещество полушарий. Непосредственно с речью обычно связывают левое полушарие у правой, при этом правое обеспечивает такие функции, как интона-

ция, мелодика речи и некоторые другие невербальные ее характеристики. Артикуляционная сторона речевой деятельности организуется при участии коры, подкорковых образований, структур ствола головного мозга [Винарская, 2006; Цветкова, 2007; Гусев и др., 2013].

К периферическому отделу речевого аппарата относят легкие с их системой подводящих дыхательных путей и дыхательными мышцами, гортань с голосовыми связками и так называемую надставную трубу, — полости носа, рта и глотки. Помимо этого, к периферическому отделу относят систему афферентных и эфферентных нервных путей, иннервирующих исполнительные органы речевого аппарата [Винарская, 2006; Беккер, 1981].

Рассмотрим речевые патологии, которые послужили предметом данного исследования. Их объединяет несколько общих свойств. Каждый из этих видов речевых патологий, как афазия, так и дизартрия, вызывается расстройствами в деятельности нервной системы и проявляется в виде нарушений плавности и темпа речи, произношения слов и фонем.

При афазии происходит полная или частичная потеря способности понимать или употреблять устную или письменную речь. Причиной возникновения афазии служат нарушения в деятельности одной или нескольких областей коры головного мозга, вызванные инсультами, опухолями, черепно-мозговыми травмами и прочими патологическими воздействиями. В зависимости от симптомов речевых нарушений и локализации поврежденных областей мозга различают несколько видов афазий, которые принято разделять на две группы: моторные и сенсорные афазии. В рамках данной работы исследовалась группа моторных афазий, связанных с нарушением временной организации речевых высказываний.

Моторные афазии существуют в двух формах: афферентной и эфферентной.

Симптомы афферентной моторной афазии возникают при поражении вторичных зон постцентральных и нижнетеменных отделов коры головного мозга, расположенных сзади от роландовой борозды. Данный вид патологии является одним из самых тяжелых речевых расстройств. Для афферентной моторной афазии наиболее характерны грубые нарушения артикуляционной стороны речи. Этот вид афазии проявляет себя в виде затруднений произношения любых речевых высказываний, от фраз или слов до отдельных звуков. При этом письменная речь может быть относительно сохранной и могут наблюдаться ситуативные высказывания в рамках контекста ситуации.

Эфферентная моторная афазия развивается на фоне поражения нижних отделов премоторной зоны левого полушария (зона Брока, поля 44, 45 коры головного мозга). Для этой формы афазии характерно чрезвычайно инертное, затрудненное переключение сценариев артикуляций. Хотя человек и способен правильно произносить отдельные звуки, слоги и слова, уже при попытке воспроизведения отдельной фразы он сталкивается с непреодолимыми трудностями, вызванными воспроизведением предыдущего слова или звука и невозможностью волевого переключения артикуляций.

Помимо этого, существуют расстройства речи, вызванные поражением глубинных, подкорковых структур, при отсутствии поражений коры головного мозга. Эти речевые расстройства способны проявлять себя симптомами различных афазий и получили в литературе названия нейродинамических речевых нарушений, подкорковых афазий, флукуативных подкорковых дисфазий, нейродинамических афазий [Прокопенко и др., 2011].

Для нейродинамических афазий, помимо стойких речевых нарушений, характерны флукуации их проявлений, качество выполнения предъявляемых больному заданий зависит как от сложности заданий, так и от его психической активности. Данный вид речевой патологии способен существовать длительное время, но при проведении коррекционных мероприятий обладает значительным потенциалом к восстановлению речи.

Дизартрия — грубое нарушение моторной динамики, расстройство артикуляции, проявляющиеся в речи в виде ее замедления, невнятности и неразборчивости. Как правило, при дизартрии грамматическая и лексическая структуры речи остаются сохранными, что позволяет больному писать, читать про себя и понимать устную речь. При данной речевой патологии часто повреждаются подкорковые структуры мозга, осуществляющие иннервацию речевого аппа-

рата либо управляющие речевыми движениями, хотя также возможно развитие дизартрии при поражении центральной нервной системы. Например, к развитию дизартрии может привести поражение следующих структур мозга: мозжечка, двигательных ядер черепных нервов бульбарной группы, подкорковых экстрапирамидных ядер и их связей, а также нижних отделов полей 1, 2, 3, 5, 6, 44 (по Бродману).

Запись экспериментальных образцов речи проводилась по уже отработанной ранее методике [Кузенков и др., 2012]. В соответствии с ней испытуемым предлагалось прочитать рассказ А. П. Чехова «Злой мальчик», а затем пересказать его в привычном для испытуемого темпе. При этом речь испытуемых фиксировалась с помощью электродинамического микрофона с кардиоидной диаграммой направленности и диапазоном частот от 50 до 15000 Гц и записывалась в файл при частоте дискретизации $\nu = 8$ кГц. Запись производилась в помещении площадью порядка 20 м² при расстоянии между микрофоном и испытуемым 10–40 см. Записанные образцы речи подвергались фильтрации с использованием фильтра верхних частот с граничной частотой 80 Гц и ослаблением сигнала 24 дБ. Также удалялась постоянная составляющая сигнала и проводилась нормировка до максимального амплитудного значения сигнала, равного 1. Из-за различия в силе голоса испытуемых и расстояния до микрофона уровень шумов после нормировки во всей группе записей находился в диапазоне от –45 до –32 дБ. Все работы по редактированию образцов речи проводились с использованием звукового редактора Audacity. В зависимости от индивидуальных свойств испытуемых и их речевых возможностей продолжительность записей варьировалась от 3 до 7 минут.

Всего было записано следующее количество образцов речи: чтение в норме (отсутствие патологии) — 69, монолог в норме — 107, чтение при афазии (моторные, без деления на эфферентную и афферентную) — 27, монолог при афазии — 54, чтение при дизартрии — 29, монолог при дизартрии — 36, чтение при нейродинамической афазии — 58, монолог при нейродинамической афазии — 69. Суммарное количество записей составило 429 образцов, суммарное время звучания 23 ч 56 мин. Несколько образцов речи с дизартрией имели сопутствующую патологию по типу нейродинамических афазий.

Модификация алгоритма вычисления показателя Хёрста

При исследовании многих природных процессов и явлений результатом наблюдений являются различные ряды зависимостей. Наиболее часто встречаются временные зависимости (зависимость какого-либо параметра от времени), затем пространственные (зависимость параметра от координаты) и в ряде случаев более абстрактные, как, например, изменение длительности земных суток (зависимость времени от времени). Подавляющее большинство таких рядов, особенно фрактального характера, целесообразно исследовать с помощью метода нормированного размаха, вычисляя показатель Хёрста [Калуш, Логинов, 2002, Апрыткина и др., 2010; Hurst, 1951; Feller, 1951]. Актуальным этот подход может оказаться и при анализе речи, поскольку речевой сигнал демонстрирует фрактальную структуру [Склярков, 2004; Henríquez et al., 2009; González, 2012; Keller, 2000; Pitsikalis, Maragos, 2009].

Классическая схема вычисления показателя выглядит следующим образом. Пусть дан временной ряд $x(t)$, определенный на множестве моментов времени $\{t_i\}$, тогда его среднее значение за период времени τ равно $\langle x(\tau) \rangle = \frac{1}{\tau} \sum_{i=1}^{\tau} x(t_i)$, а стандартное отклонение —

$S(\tau) = \sqrt{\frac{1}{\tau} \sum_{i=1}^{\tau} (x(t_i) - \langle x(\tau) \rangle)^2}$. Введем накопившееся отклонение $X(t, \tau) = \sum_{j=1}^t (x(j) - \langle x(\tau) \rangle)$ ряда $x(t)$ от его среднего значения $\langle x(\tau) \rangle$. Разница между максимальным и минимальным значениями ряда $X(t, \tau)$ называется размахом $R(\tau)$: $R(\tau) = \max_{1 \leq t \leq \tau} X(t, \tau) - \min_{1 \leq t \leq \tau} X(t, \tau)$. Размах $R(\tau)$

зависит от времени τ и имеет ярко выраженную тенденцию к росту, особенно при относительно небольших значениях τ . Эмпирическим путем было установлено, что безразмерная величина R/S подчиняется следующему закону [Hurst, 1951]:

$$\frac{R}{S} = \left(\frac{\tau}{2} \right)^H. \quad (1)$$

Величина H в литературе стала называться показателем Хёрста, а анализ рядов с помощью ее вычисления — методом нормированного размаха. Характерной особенностью показателя Хёрста является его способность различать характер анализируемых рядов. Так, например, для случайного процесса с независимыми приращениями и конечной дисперсией, или, что то же самое, для броуновского движения с отсутствием памяти, величина H строго равна 0.5 [Feller, 1951]. Отклонение показателя в ту или иную сторону от 0.5 указывает на различные особенности исследуемого ряда. В случае когда $0.5 < H < 1.0$, ряд считается персистентным и обладающим фрактальными свойствами. В таком ряду тенденция к возрастанию или убыванию ряда поддерживается порождающей ряд системой, а система, считается, обладает памятью, т. е. текущие значения ряда зависят от прошлых отсчетов за некоторый интервал времени. В случае $0 < H < 0.5$ ряд считается антиперсистентным, т. е. тенденция к возрастанию быстро изменяется на убывающую, и наоборот [Калуш, Логинов, 2002; Апрыткина и др., 2010; Бутаков, Граковский, 2005; Петерс, 2000; Hurst, 1951; Feller, 1951].

Типичная схема вычислений показателя Хёрста состоит из логарифмирования выражения (1) и нахождения величины H как тангенса угла наклона графика (либо его прямолинейного участка) зависимости $\ln \frac{R(\tau)}{S(\tau)}$ от $\ln \frac{\tau}{2}$.

Однако непосредственное вычисление показателя Хёрста таким образом, когда используется фиксированное приращение τ , целесообразно лишь в редких единичных случаях для грубой оценки показателя «на глаз» в силу следующих причин. Во-первых, традиционно τ определяется на эквидистантном интервале, что приводит к экспоненциальному увеличению плотности отсчетов по мере увеличения τ при переходе к логарифмической шкале. Данный факт создает значительные неудобства при аппроксимации участка графика прямой линией, которая может отклоняться от точек графика при малых значениях τ , где плотность отсчетов минимальна. Во-вторых, при анализе достаточно длинных временных рядов в расчете показателя учитывается информация только о начале ряда, что иногда приводит к зависимости показателя от начальной фазы сигнала. В-третьих, граница линейного роста показателя определяется с большой погрешностью. Все перечисленные эффекты в совокупности значительно повышают погрешность оценки показателя Хёрста и приводят к неоправданно высоким затратам при попытке ее автоматизации.

В силу вышеизложенных причин нами был разработан оригинальный алгоритм вычисления показателя Хёрста. Поскольку наибольшие трудности создает неравномерная шкала $\ln \frac{\tau}{2}$, шкалу τ было предложено реализовать в виде геометрической прогрессии, что приводило к равномерной шкале после процедуры логарифмирования. С целью максимально эффективно использования информации о временном ряде вычислялось несколько значений показателя Хёрста на разных участках ряда с последующим усреднением.

Алгоритм вычисления теперь выглядел таким образом. Пусть N — количество отсчетов анализируемого сигнала, n — количество значений R/S , которые нужно вычислить. Тогда шкала τ имеет длину n , $\tau_1 = 3$, $\tau_{i+1} = a \cdot \tau_i$, где $a = \left(\frac{N-1}{3} \right)^{\frac{1}{n}}$. Для вычисления текущего нормированного размаха при заданном τ_i значение τ_i округляется вверх и вниз до ближайших це-

лых значений τ_i^{up} и τ_i^{down} соответственно. Затем для каждого округленного значения τ_i^* находится величина k , равная целой части от выражения $\frac{N}{\tau_i^*} + 1$, и весь исходный ряд разбивается

на k интервалов, каждый длиной τ_i^* , причем интервалы разбиения могут частично пересекаться. Расстояние от начала исследуемого ряда до середины первого интервала и от середины последнего до конца ряда примерно равны половине расстояния между центрами интервалов. Для всех интервалов, соответствующих τ_i^{up} и τ_i^{down} , находится множество значений нормированного размаха, а затем их средние значения

$\left\langle \frac{R(\tau_i^{up})}{S(\tau_i^{up})} \right\rangle$ и $\left\langle \frac{R(\tau_i^{down})}{S(\tau_i^{down})} \right\rangle$. Искомое значение $\frac{R(\tau_i)}{S(\tau_i)}$ находится путем линейной интерполяции $\left\langle \frac{R(\tau_i^{up})}{S(\tau_i^{up})} \right\rangle$ и $\left\langle \frac{R(\tau_i^{down})}{S(\tau_i^{down})} \right\rangle$. В том случае, когда $\tau_i > \frac{N}{2}$,

$k = 2$, началом первого интервала служит первый отсчет ряда, концом второго (последнего) — последний отсчет ряда, а расстояние между центрами интервалов значительно сокращается.

Для оценки свойств модифицированного алгоритма проведем сравнительную оценку его вычислительной сложности. В классическом варианте вычисляется множество значений нормированного размаха для всей последовательности значений τ от $\tau = 2$ до $\tau = N$. Для вычисления одного значения нормированного размаха требуется выполнение последовательности циклов, вычисляющих среднее, стандартное отклонение и накопившееся отклонение для текущего значения τ . Количество необходимых операций при этом пропорционально значению τ . Таким образом, вычислительная сложность операции по нахождению одного значения нормированного размаха составляет $O(n)$. Вычислительная сложность нахождения множества значений нормированного размаха для всех значений τ при этом будет пропорциональна сумме арифметической прогрессии: $\frac{\tau + N}{2}(N - \tau) = \frac{N^2 - \tau^2}{2} = \frac{N^2 - 4}{2}$, что соответствует $O(n^2)$.

В случае модифицированного алгоритма при количестве интервалов разбиения ряда, равном k , длина одного интервала составляет $\frac{N}{k}$. Количество операций с учетом удвоенного количества

вычислений для интерполяции пропорционально $2 \cdot \frac{N}{k} \cdot k = 2N$, т. е. в таком случае количество

операций для вычисления одного значения нормированного размаха пропорционально N и остается фиксированным. Количество операций для вычисления зависимости нормированного размаха пропорционально произведению nN . Следовательно, вычислительная сложность модифицированного алгоритма составляет $O(n)$ и в предельном случае $n = N$ $O(n^2)$. Таким образом, предлагаемый алгоритм является более эффективным по сравнению с традиционным при анализе достаточно длинных рядов и вычислении количества значений нормированного размаха, меньшего длины анализируемого ряда.

Кроме этого, описанный способ вычисления нормированного размаха обладает и другими преимуществами. Использование прогрессивной шкалы при вычислении значений нормированного размаха позволяет получить эквидистантный интервал после логарифмирования, что значительно упрощает последующую процедуру аппроксимации и повышает точность оценки показателя Хёрста. Использование процедуры усреднения значений нормированного размаха по нескольким интервалам ряда значительно повышает точность оценки показателя Хёрста, а также снижает его чувствительность к начальной фазе сигнала. Особенно ярко перечисленные эффекты проявляются при анализе нестационарных сигналов, например при исследовании огибающей речевого сигнала, когда имеет место чередование пауз с малоинтенсивным шумом различной длительности и импульсов звука различной формы и амплитуды.

Вычисленное множество значений $\ln \frac{R(\tau)}{S(\tau)}$ описанным способом аппроксимировалось

с помощью кусочно-линейной функции, состоящей из двух отрезков, как зависимость от $\ln \frac{\tau}{2}$.

С целью оценки работоспособности алгоритма и иллюстрации его работы данный алгоритм использовался для анализа модельных рядов. В частности, генерировался синусоидальный сигнал с периодом, равным 20 отсчетам, и продолжительностью 100000 отсчетов.

Нормированный размах находился по предложенному выше алгоритму (рис. 1).

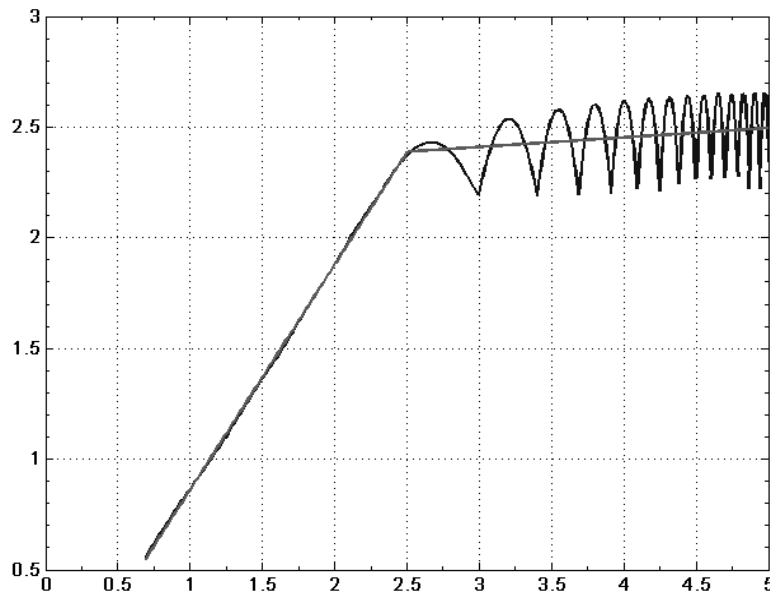


Рис. 1. Зависимость $\ln \frac{R(\tau)}{S(\tau)}$ от $\ln \frac{\tau}{2}$ для гармонических колебаний и ее аппроксимация кусочно-линейной функцией

Как видно из рисунка, существует характерное время, после которого показатель Хёрста резко меняется. При анализе такого изменения авторы иногда это объясняют временным масштабом, в течение которого система, породившая анализируемый ряд, учитывает предыдущие значения ряда. Другими словами, данное время указывает на объем памяти системы [Бутаков, Граковский, 2005; Петерс, 2000; Feller, 1951].

Вычисленное значение показателя Хёрста для синусоиды $H \approx 1.017$ близко к значениям, полученным другими авторами, $H = 1.026$ [Бутаков, Граковский, 2005], а характерное время составляет 25 отсчетов, или $5/4$ периода колебаний.

Сравнительная оценка параметров нормированного размаха в различных группах патологий

Изменение речи при речевых патологиях физически отражается в изменении тех или иных свойств речевого сигнала. Использование различных методов анализа временных рядов, в особенности методов нелинейной динамики, позволяет фиксировать такие изменения и автоматизировать процедуру диагностики речевых патологий [Кузенков и др., 2009; Скляр, 2004; Gómez et al., 2005; Henríquez, 2009; Chou, 2007; Scalassara et al., 2009; Malyska et al., 2005; Schlotthauer et al., 2010].

Поскольку речь в случае наличия патологии отличается изменением в амплитудно-временных параметрах, также можно ожидать изменения показателя Хёрста мощности речевого сигнала в группах с патологией по сравнению с нормой. В качестве исходного сигнала для вычисления показателя Хёрста использовалась огибающая речевого сигнала, вычисляемая как среднеквадратичное отклонение внутри скользящего окна длительностью 100 мс.

Характер поведения нормированного размаха мало отличается от модельных рядов. Также наблюдается линейный рост в начале графика, который резко изменяется при достижении характерного времени и часто становится нелинейным. При аппроксимации полученных зависимостей вместо линейно-кусочной функции теперь использовалась полиномиально-кусочная функция (рис. 2).

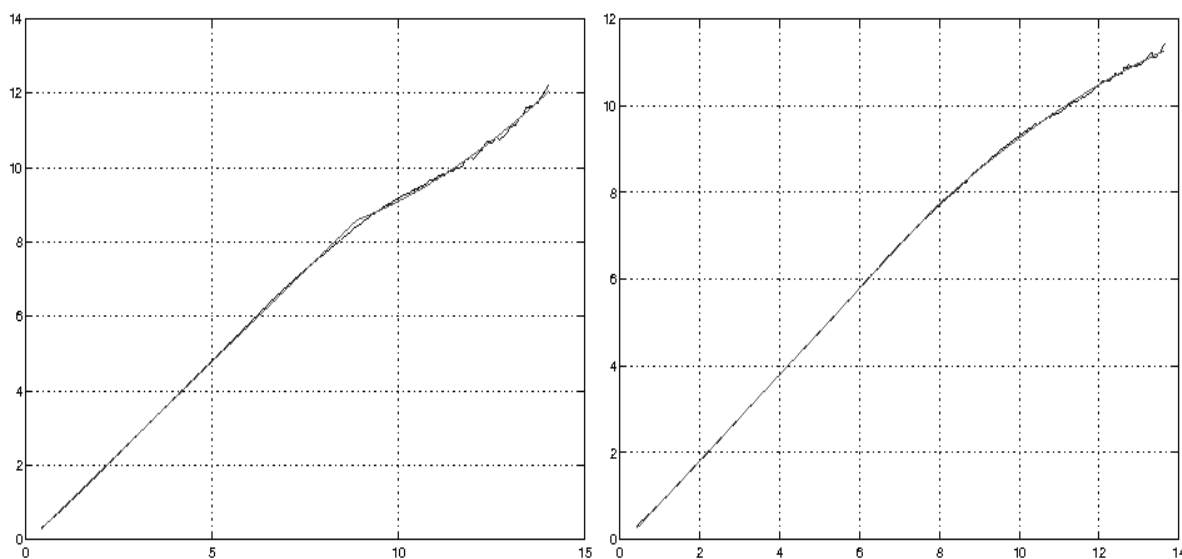


Рис. 2. Зависимость $\ln \frac{R(\tau)}{S(\tau)}$ от $\ln \frac{\tau}{2}$, τ представлена в отсчетах сигнала. Испытуемая Л., норма, слева — чтение, справа — пересказ

Такая замена аппроксимирующей функции обусловлена прежде всего требованием повышения точности вычисления характерного времени. В подавляющем большинстве случаев, характер графика нормированного размаха при временах, больших характерного времени, носит ярко выраженный нелинейный характер. В таких случаях, если его аппроксимировать линейной функцией, происходит увеличение погрешности в определении момента смены функций, а поскольку анализируемые графики строятся в логарифмических координатах и для нахождения характерного времени приходится вычислять показательную функцию, даже небольшое изменение в определении момента смены функций приводит к значительному росту погрешности в определении характерного времени.

Аппроксимирующая функция при этом представляла собой комбинацию линейной и квадратичной функций:

$$\begin{cases} f(x) = kx + d, x \leq x^* \\ f(x) = ax^2 + bx + c, x \geq x^* \end{cases}$$

Очевидно, что при оптимальных значениях параметров a, b, c, d, k должно выполняться условие минимума функции $Q(\tau) = \left(\frac{R(\tau)}{S(\tau)} - f(\tau) \right)^2 = \min$. Процедура непосредственной оптими-

зации $Q(\tau)$ по параметрам a, b, c, d, k , как выяснилось в результате экспериментов, не обладает стабильностью и часто приводит к неудовлетворительным результатам при использовании различных алгоритмов оптимизации. По этой причине для оптимизации использовался вектор параметров $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$. Начальные значения данного вектора определялись из графика зависимости $\ln \frac{R(\tau)}{S(\tau)}$ от $\ln \frac{\tau}{2}$ следующим образом: x_1, y_1, x_4, y_4 — координаты первой и последней точек графика, x_2 — такое значение абсциссы, при котором отклонение графика от отрезка, проведенного от точки (x_1, y_1) до точки (x_4, y_4) максимально, $x_3 = \frac{x_2 + x_4}{2}$, y_2, y_3 — ординаты графика, соответствующие абсциссам x_2 и x_3 . Вектор параметров a, b, c, d, k находился по значениям вектора координат $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ с учетом следующих условий: линейный участок функции проходит через точки (x_1, y_1) и (x_2, y_2) , квадратичный участок — через точки $(x_2, y_2), (x_3, y_3), (x_4, y_4)$.

После завершения процедуры оптимизации находились показатель Хёрста H как наклон линейного участка аппроксимирующей функции и характерное время $T = 2e^{x_2} / v$ (напомним, что v — частота дискретизации), а также их статистические характеристики для всех групп патологии и нормы, в том числе коэффициент корреляции между величинами H и T (таблица 1).

Таблица 1. Статистические характеристики в различных речевых группах

	коэффициент корреляции	H			T		
		среднее	ср. кв. отклонение	медиана	среднее	ср. кв. отклонение	медиана
афазия, чтение	-0.643	0.989	0.00488	0.990	1.21	0.403	1.18
афазия, пересказ	-0.610	0.990	0.00509	0.991	1.06	0.642	0.824
дизартрия, чтение	-0.686	0.988	0.00511	0.988	1.33	0.532	1.15
дизартрия, пересказ	-0.405	0.988	0.00488	0.989	0.966	0.638	0.832
норма, чтение	-0.850	0.990	0.00289	0.990	0.902	0.259	0.851
норма, пересказ	-0.852	0.991	0.00384	0.991	0.934	0.497	0.902
нейродинамическая афазия, чтение	-0.289	0.985	0.00593	0.986	1.28	0.467	1.24
нейродинамическая афазия, пересказ	-0.391	0.984	0.00671	0.984	0.997	0.576	0.832

Как видно из таблицы 1, средние значения H слабо различаются между собой во всех группах. Типичные для речевого сигнала значения показателя Хёрста независимо от вида патологии и режима вокализации (чтение, монолог) составляют 0.98–0.99, а для характерного времени — 0.9–1.3с. Однако на плоскости параметров $T(H)$ можно увидеть, что занимаемые различными группами области совпадают лишь частично (рис. 3, 4).

Для всех групп были найдены линейные регрессии вида $T = kH + b$, где k и b вычислялись с использованием метода наименьших квадратов в соответствии с формулами

$$k = \frac{\sum_{i=1}^N H_i \times \sum_{i=1}^N T_i - N \sum_{i=1}^N H_i T_i}{\left(\sum_{i=1}^N H_i\right)^2 - N \sum_{i=1}^N H_i^2} \quad \text{и} \quad b = \frac{\sum_{i=1}^N T_i - k \sum_{i=1}^N H_i}{N},$$

где N — количество точек в анализируемой группе. Полученные уравнения регрессий для каждой группы представлены в таблице 2.

Таблица 2. Уравнения регрессий в различных речевых группах

Речевая группа	Регрессия
Норма, чтение:	$T = -76.3H + 76.5$
Норма, монолог:	$T = -110H + 110$
Афазия, чтение:	$T = -53.1H + 53.8$
Афазия, монолог	$T = -76.8H + 77.0$
Дизартрия, чтение	$T = -71.4H + 71.8$
Дизартрия, монолог	$T = -52.9H + 53.3$
Нейродинамическая афазия, чтение	$T = -22.8H + 23.7$
Нейродинамическая афазия, монолог	$T = -33.6H + 34.0$

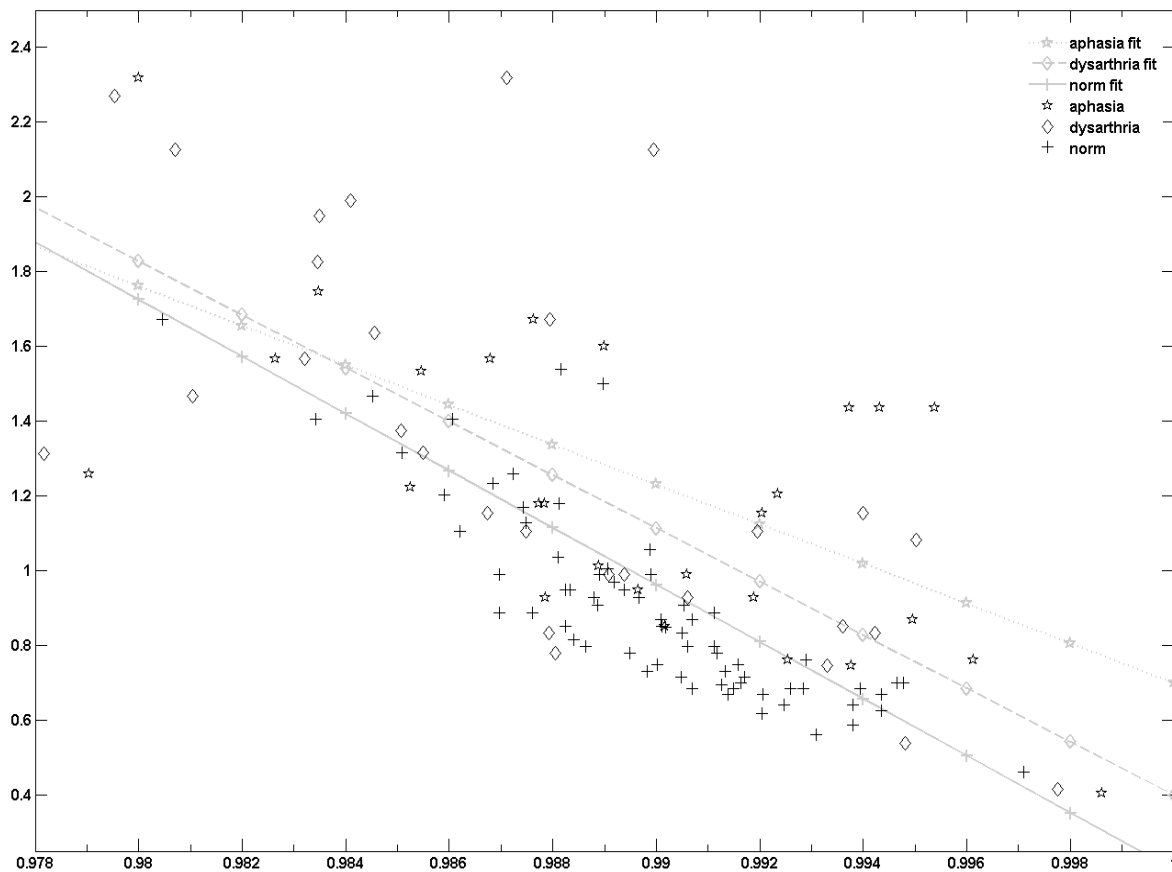


Рис. 3. Зависимость $T(H)$ в группах чтения, а также линейные регрессии для каждой группы

Полученные уравнения регрессий в силу очень близких значений k и b позволяют предположить существование закономерности вида $T = a(1 - H)$ для класса исследуемых сигналов, однако анализ этой закономерности выходит за рамки данной работы и будет проведен позже.

В силу существенных отличий распределений исследуемых характеристик от нормальных распределений во всех группах (рис. 5) применение традиционных параметрических методов

для проверки статистических гипотез при возможности проведения вычислительного эксперимента выглядит неоправданно затратным.

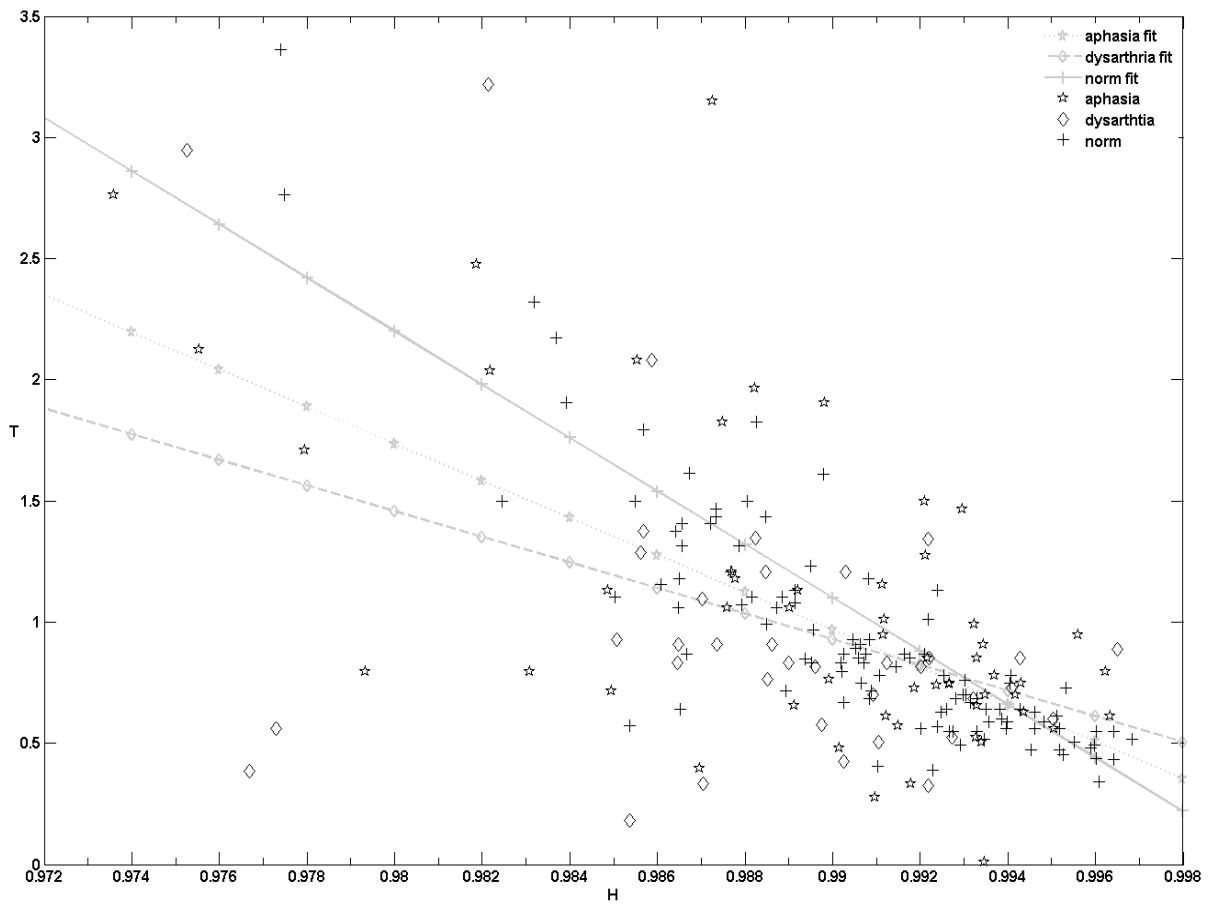


Рис. 4. Зависимость $T(H)$ в группах монолога, а также линейные регрессии для каждой группы

Вследствие этого для анализа результатов использовались непараметрические статистические методы, в частности рандомизационный тест [Шитиков, Розенберг, 2013; Good, 2004].

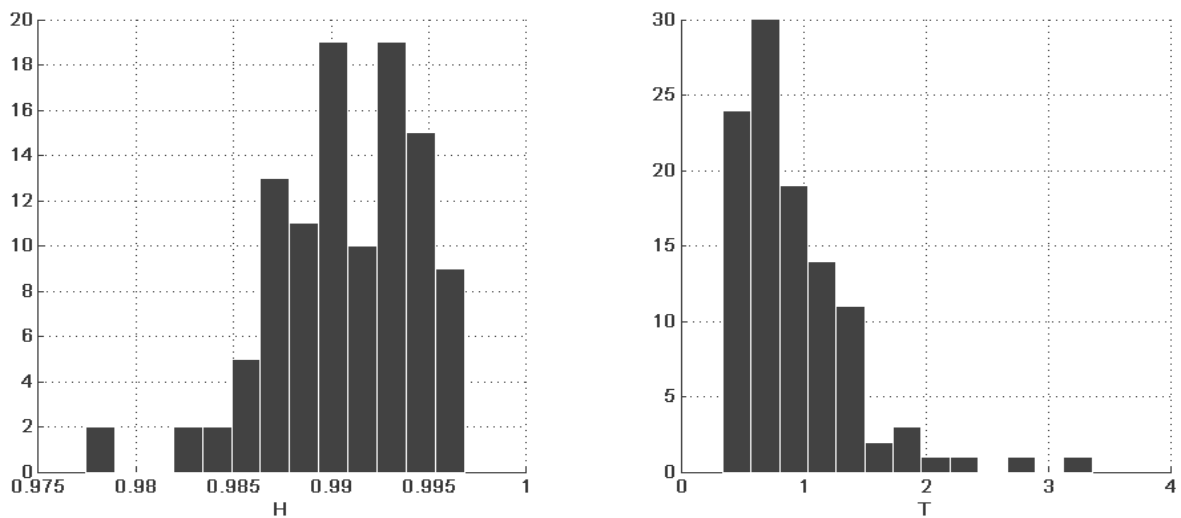


Рис. 5. Распределения испытуемых в группе, норма, монолог в зависимости от показателя Хёрста (слева) и характерного времени (справа)

Проверка статистической гипотезы с помощью рандомизационного теста заключается в генерации большого количества синтетических выборок на основании исходных выборок. С этой целью данные двух выборок объединяются, а затем случайным образом разделяются на две группы в таком же количестве, как исходные. Затем для этих групп находится разность каких-либо параметров δ (математическое ожидание, дисперсия, наклон линии регрессии и т. п.) и процедура многократно повторяется. Далее по результатам такого вычислительного эксперимента строится экспериментальная функция распределения δ и находится количество случаев m , когда значение δ больше (меньше, в зависимости от вида гипотезы) δ_0 , вычисленной для пары исходных выборок. Отношение найденного значения m к общему количеству проведенных экспериментов по перемешиванию выборок M является численной оценкой уровня значимости проверяемой гипотезы: $p = \frac{m}{M}$.

Очевидно, что при таком подходе генерируемые выборки не будут являться полностью независимыми, однако при достаточно большом количестве экспериментов этот факт не имеет большого значения. Главные преимущества рандомизационного теста перед классическими параметрическими тестами заключаются в отсутствии требований к случайным величинам (вид распределения, равенство дисперсий и т. д.), а также универсальности процедуры проверки различных статистических гипотез. Всего для каждой пары групп выполнялась генерация выборок в количестве 5000 раз, что достаточно для проверки статистических гипотез с уровнем значимости 0.01 [Шитиков, Розенберг, 2013; Good, 2004].

В данной работе проверялись статистически значимые различия между угловыми коэффициентами наклона линий регрессий k и средними значениями показателя Хёрста H и характерного времени T в группах патологии, чтения с группой нормы, чтения, а также в группах патологии, монолога с группой нормы, монолога. При этом были получены следующие значения уровня значимости:

сравнение групп чтения с нормой

афазия:	$p_k = 0.153,$	$p_H = 0.509,$	$p_T = 0;$
дизартрия:	$p_k = 0.388,$	$p_H = 0.0086,$	$p_T = 0;$
нейродинамическая афазия:	$p_k = 0.0168,$	$p_H = 0,$	$p_T = 0;$

сравнение групп монолога с нормой

афазия:	$p_k = 0.0538,$	$p_H = 0.092,$	$p_T = 0.089;$
дизартрия:	$p_k = 0.0762,$	$p_H = 0.0022,$	$p_T = 0.622;$
нейродинамическая афазия:	$p_k = 0,$	$p_H = 0,$	$p_T = 0.778.$

Кроме того, при сравнении чтения и монолога в норме был получен уровень значимости $p_k = 0.0412, p_H = 0.081, p_T = 0.311$.

Величина p здесь имеет смысл вероятности получить наблюдаемое значение статистики (разности k , а также средних значений H и T) или еще менее вероятных значений при условии, что верна гипотеза о равенстве этих значений [Шитиков, Розенберг, 2013].

Значения $p = 0$ означают, что уровень значимости не превышает 0.01.

Если в качестве признака достоверного различия групп взять превышение общепринятым значением уровня значимости $p = 0.05$ вычисленного уровня, тогда можно утверждать, что гипотеза о равенстве статистик в группах не подтвердилась в следующих случаях:

афазия, чтение — норма, чтение различаются по среднему значению характерного времени;
дизартрия, чтение — норма, чтение различаются по средним значениям показателя Хёрста и характерного времени;
дизартрия, монолог — норма, монолог различаются по среднему значению показателя Хёрста;

нейродинамическая афазия, чтение — норма, чтение различаются по средним значениям показателя Хёрста и характерного времени;

нейродинамическая афазия, монолог — норма, монолог различаются по среднему значению показателя Хёрста.

Помимо различий в средних значениях показателей Хёрста и характерного времени, также в трех случаях наблюдаются различия в наклоне линий регрессии. В первых двух случаях различия наблюдаются в группах нейродинамических афазий по сравнению с нормой. В третьем случае различия наблюдаются между монологом и чтением в норме.

Единственной группой патологий, которая с высокой вероятностью совпадает с нормой, оказалась афазия, монолог.

Заключение

На основании статистической обработки речевых сигналов можно сделать вывод о возможности применения показателя Хёрста при решении задач диагностики речевых патологий. В большинстве случаев наличие речевой патологии приводит к достоверному изменению значений показателя Хёрста и (или) характерного времени на графиках нормированного размаха. Примечательной особенностью метода явилось уверенное отличие как монолога, так и чтения при нейродинамической афазии от нормы. Данный факт является достаточно важным результатом, поскольку симптомы речевых нарушений при нейродинамической афазии часто походят на симптомы моторных афазий, а предпринятая нами ранее попытка диагностики речевых патологий по речевому сигналу не позволила достоверно различить нейродинамические афазии и дизартрии [Кузенков и др., 2009]. Поскольку часть речевых патологий на плоскости параметров $T(H)$ совпадает с областями, занятыми речевыми сигналами в норме, для повышения эффективности диагностики данный метод может применяться в комбинации с другими методами анализа речевых сигналов.

В более общем случае предложенный способ анализа может быть использован в исследовании временных рядов произвольного происхождения. Наиболее перспективным использование данного метода по сравнению с традиционным представляется в следующих случаях: при анализе достаточно длинных временных рядов с целью снижения времени вычислений, при анализе сигналов с нестационарными характеристиками, результат обработки которых может зависеть от начальной фазы сигнала, а также в других случаях, когда требуется увеличение точности оценки показателя Хёрста, в особенности характерного времени.

Список литературы

- Апраткина М. Л., Бородин А. С., Васильченко М. Н. Анализ показателя Херста для вариаций периода сокращений сердца человека // Известия высших учебных заведений. Физика. — 2010. — Т. 53. — № 9–3. — С. 233–234.
- Беккер К.-П., Совак. М. Логопедия: Пер. с нем. — М.: Медицина, 1981. — 288 с.
- Бутаков В., Граковский А. Оценка уровня стохастичности временных рядов произвольного происхождения при помощи показателя Хёрста // Computer Modelling and New Technologies. — 2005. — Vol. 9, № 2 — P. 27–32.
- Винарская Е. Н. Дизартрия. — М.: АСТ, 2006. — 141 с.
- Вьюгин В. В., Сорокин В. Н., Тананыкин А. А. Распознавание личности по голосу: аналитический обзор // Информационные процессы. — 2012. — Т. 12, №1. — С. 1–30. URL: <http://www.jip.ru/2012/1-30-2012.pdf> (дата обращения 06.10.2013).
- Голубинский А. Н. Выявление эмоционального состояния человека по речевому сигналу на основе вейвлет-анализа // Вестник Воронежского института МВД России. — 2011. — № 3. — С. 144–153.

- Голубинский А. Н. Математическая модель речевого сигнала, основанная на аппроксимации спектра набором постоянных составляющих в соответствующих полосах частот // Безопасность информационных технологий. — 2009. — № 2. — С. 12–17.
- Гусев Е. И., Коновалов А. Н., Никифоров А. С. Клиническая неврология. Т. 1. — М.: Медицина, 2002. — 704 с.
- Гусев М. Н. Расширенная модель длительности звуков для системы распознавания речи. Современное машиностроение // Наука и образование. — 2013. — № 3. — С. 1112–1121.
- Калуш Ю. А., Логинов В. М. Показатель Хёрста и его скрытые свойства // Сибирский журнал индустриальной математики. — 2002. — Т. V, № 4. — С. 29–37.
- Ковальчук В. В., Скоромец А. А., Высоцкая М. Л. Пациент после инсульта. психоэмоциональное состояние и физическое восстановление // Вестник восстановительной медицины. — 2007. — № 4. — С. 85–88.
- Кузенков Н. П., Логинов В. М., Никольская О. Н., Прокопенко С. В. Алгоритм классификации речевых патологий при органических поражениях головного мозга // Информационные процессы. — 2009. — Т. 9, № 3. — С. 121–137. URL: <http://www.jip.ru/2009/121-137-2009.pdf> (дата обращения 06.10.2013).
- Кузенков Н. П., Логинов В. М., Прокопенко С. В., Можейко Е. Ю. Статистическое описание речи человека при органических нарушениях центрального отдела речевого аппарата // Нелинейный мир. — 2012. — № 4. — С. 247–255.
- Малюкова Н. Г. Значение методов нейропсихологической диагностики нарушений когнитивных функций в системе современной нейрореабилитации // Психологическая наука и образование. — 2012. — № 4. — С. 20–30.
- Морозова Н. Ю. Как преодолеть заикание. — М.: ЭКСМО-Пресс, 2002. — 192 с.
- Петерс Э. Хаос и порядок на рынках капитала. — М.: Мир, 2000. — 333 с.
- Прокопенко С. В., Можейко Е. Ю., Визель Т. Г., Никольская О. Н. Нейродинамические нарушения речи в постинсультном периоде: патогенез, клиника, диагностика // Бюллетень сибирской медицины. — 2011. — Т. 10, № 2. — С. 154–160.
- Сидоров К. В., Филатова Н. Н. Автоматическое распознавание эмоций человека на основе реконструкций аттракторов образцов речи // Программные системы и вычислительные методы. — 2012. — № 1. — С. 67–79.
- Скляров О. П. V/U — ритм речи при чтении как индикатор состояния функции речевого дыхания у заикающихся // Техническая акустика. — 2004. — Т. 4. — С. 163–173.
- Скляров О. П. Фракталы и крупномасштабная временная структура акустического речевого сигнала и музыки // Техническая акустика. — 2004. — Т. 4. — С. 224–233.
- Сорокин В. Н., Гераськин И. В. Оценка длины речевого тракта // Информационные процессы. — 2013. — Т. 13, № 2. — С. 35–47. URL: <http://www.jip.ru/2013/35-47-2013.pdf> (дата обращения 06.10.2013).
- Цветкова Л. С. Введение в нейропсихологию и восстановительное обучение. — М.: Московский психолого-социальный университет, 2007. — 184 с.
- Шарий Т. В. Об одном методе автоматической сегментации речевых сигналов // Бионика интеллекта. — 2009. — № 2(71). — С. 61–65.
- Шитиков В. К., Розенберг Г. С. Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R. — Тольятти: «Кассандра», 2013. — С. 305.
- Ястребцева И. П. Терапия аффективных расстройств после инсульта головного мозга // Вестник новых медицинских технологий. — 2009. — Т. XVI, № 3. — С. 113–114.
- Chou M., Jiang J. J., Rahn III D. A., and Zhang Y. Phonatory Impairment in Parkinson's Disease: Evidence from Nonlinear Dynamic Analysis and Perturbation Analysis. J. Voice. 2007. — Vol. 21. P. 64–71.
- Diana Cristina González, Lee Luan Ling, Fábio Violaro. Analysis of the multifractal nature of speech signals. Progress in pattern recognition, image analysis, computer vision, and applications lecture notes in computer science. — 2012. — Vol. 7441. — P. 740–748.

- Feller W.* The asymptotic distribution of the range of sums of independent variables // *Ann. Math. Statist.* — 1951. — Vol. 22. — P. 427–432.
- Gastón Schlotthauer, María E. Torres, Hugo L. Rufiner.* Pathological Voice Analysis and Classification Based on Empirical Mode Decomposition. *Development of Multimodal Interfaces: Active Listening and Synchrony Lecture Notes in Computer Science.* — 2010. — Vol. 5967. — P. 364–381.
- Gómez P., Martínez R., Díaz F., Lázaro C., Álvarez A., Rodellar V., and Nietro V.* Voice Pathology Detection by Vocal Cord Biomechanical Parameter Estimation. In: *Nonlinear Analyses and Algorithms for Speech Processing*, Berlin: Springer, 2005. — P. 242–256.
- Good P. I.* *Permutation, Parametric and Bootstrap Tests of Hypotheses.* NY: Springer Science+Business Media, 2004. — P. 315.
- Henriquez P., Alonso J. B., Ferrer M. A., Travieso C. M., Godino-Llorente J. I., Diaz-de-Maria F.* Characterization of healthy and pathological voice through measures based on nonlinear dynamics // *IEEE Trans Audio Speech Lang Processing.* — 2009. — No. 17(6). — P. 1186–1195.
- Herzel H., Knudsen C.* Bifurcations in a vocal fold model // *Nonlinear Dynamics.* — 1995. — Vol. 7, No. 13. — P. 53–64.
- Hurst H. E.* Long-term storage capacity of reservoirs // *Trans. Amer. Soc. Civ. Engrs.* — 1951. — Vol. 116. — P. 770–808.
- Malyska N., Quatieri T. F., Sturim D.* Automatic dysphonia recognition using biologically-inspired amplitude-modulation features // *Acoustics, Speech, and Signal Processing.* — 2005. — Proceedings. (ICASSP '05). IEEE International Conference on (Vol. 1). — P. 873–876.
- Henriquez P., Alonso J. B., Ferrer M. A., Travieso C. M., Godino-Llorente J. I., and Diaz-de-Maria F.* Characterization of healthy and pathological voice through measures based on nonlinear dynamics // *Ieee transactions on audio, speech, and language processing.* 2009. — Vol. 17, No. 6. — P. 1186–1195.
- Scalassara P. R., Dajer M. E., Maciel C. D., Guido R. C., Pereira J. C.* Relative entropy measures applied to healthy and pathological voice characterization // *Applied Mathematics and Computation.* — 2009. — Vol. 207. — Issue 1. — P. 95–108.
- Pitsikalis V., Maragos P.* Analysis and classification of speech signals by generalized fractal dimension features // *Speech Communication* — 2009. — No. 51. — P. 1206–1223.
- Zellner Keller B. and Keller E.* The chaotic nature of speech rhythm: hints for fluency in the language acquisition process. In *Delcloque, Ph., Holland, V. M. (eds.) Integrating Speech Technology in Language Learning.* — Swets & Zeitlinger, In Press (2000). — P. 30.

ПРИЛОЖЕНИЕ

Скрипт для вычисления нормированного размаха был разработан в среде Scilab 5.5.0 и представлен ниже. Обращение к нему происходит в виде вызова функции $y = \text{avrRSsetRange}(s, n1, n2, L)$, где s — анализируемый временной ряд (одномерный массив), $n1$ и $n2$ соответственно нижняя и верхняя границы (выраженные в количестве отсчетов) анализируемого интервала времени, L — количество вычисляемых значений нормированного размаха внутри интервала $[n1, n2]$. Функция возвращает двумерный массив размером $L \times 2$, в котором первый столбец содержит значения времен (выраженные в количестве отсчетов), а второй — вычисленные для этих времен значения нормированного размаха. Отсутствие целых значений для количества отсчетов не должно вводить в заблуждение, поскольку они получены с использованием интерполяции и посредством частоты дискретизации сигнала связаны со шкалой времени.

В случае необходимости данный скрипт может быть легко адаптирован для сред Octave, Freemath, Matlab, а также с их помощью транслирован на другие языки программирования.

```

function y=avrRSsetRange(s, n1, n2, L)
    // s - data set; n1,n2 - lower and upper bounds;
    // L - length of RS data set
    winH=waitbar(0);
    t=zeros(L,1);
    x=zeros(L,1);
    N=length(s);
    a=(n2/n1)^(1/(L-1));
    t(1)=n1;
    for i=2:L
        t(i)=a*t(i-1);
    end;
    for i=1:L
        t1=floor(t(i));
        t2=ceil(t(i));
        k=fix(N/t1)+1;
        delta=N/(2*k);
        pos=zeros(k,1);
        pos(1)=delta;
        for j=2:k
            pos(j)=pos(j-1)+2*delta;
        end;
        boroders=zeros(k,2);
        for j=1:k
            boroders(j,1)=floor(pos(j)-delta)+1;
            boroders(j,2)=boroders(j,1)+t1-1;
        end;
        if boroders(k,2)>N
            tt=boroders(k,2)-N;
            boroders(k,1)=boroders(k,1)-tt;
            boroders(k,2)=boroders(k,2)-tt;
        end;
        x1=0;
        for j=1:k
            x1=x1+RS(s(boroders(j,1):boroders(j,2)));
        end;
        x1=x1/k;
        k=fix(N/t2)+1;
        delta=N/(2*k);
        pos=zeros(k,1);
        pos(1)=delta;
        for j=2:k
            pos(j)=pos(j-1)+2*delta;
        end;
        boroders=zeros(k,2);
        for j=1:k
            boroders(j,1)=floor(pos(j)-delta)+1;
            boroders(j,2)=boroders(j,1)+t2-1;
        end;
        if boroders(k,2)>N
            tt=boroders(k,2)-N;
            boroders(k,1)=boroders(k,1)-tt;
            boroders(k,2)=boroders(k,2)-tt;
        end;
        x2=0;

```



```
for j=1:k
    x2=x2+RS(s(boroders(j,1):boroders(j,2)));
end;
x2=x2/k;
if t2==t1
    x(i)=(x1+x2)/2;
else
    x(i)=x1+(x2-x1)*(t(i)-t1)/(t2-t1);
end;
waitbar(i/L,'Computing, please wait...',winH);
end;
y(:,1)=t; y(:,2)=x;
close(winH);
endfunction

function y=RS(x)
    lx=length(x);
    avrx=sum(x)/lx;
    devx=x-avrx;
    xcum=cumsum(devx);
    stdx=(sum(devx.^2)/lx)^0.5;
    if stdx~=0
        y=(max(xcum)-min(xcum))/stdx;
    else
        y=0;
    end;
endfunction
```