

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2017»

Москва, 31 мая — 3 июня 2017

СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ ПЕРЦЕПТИВНОГО ОЦЕНИВАНИЯ ГЛУБИНЫ ПРОСОДИЧЕСКИХ ШВОВ В РУССКОМ ЗВУЧАЩЕМ ТЕКСТЕ¹

Смирнова О. С. (kisaolga@mail.ru)

Московский Государственный Университет
им. М. В. Ломоносова, Москва, Россия

Ключевые слова: глубина просодического шва (PBS, сила брейка), согласованность группы экспертов (аудиторов), процедура согласования, упорядоченная классификация, медианное среднее, медиана Кемени, законы ранговых распределений

STATISTICAL ANALYSIS OF PERCEPTIVE ESTIMATION FOR DEPTH OF PROSODIC BREAKS IN RUSSIAN SPOKEN TEXT

Smirnova O. S. (kisaolga@mail.ru)

Lomonosov Moscow State University, Moscow, Russia

In our project some perception experiments were performed to establish the number of prosodic boundary strength (PBS) levels that unprofessional auditors can reliably detect in a spoken Russian text. This paper deals with the statistical analysis of these experiments' data. Nonparametric and nonnumeric statistical methods were applied. All computations were realized by means of statistical program package STATISTICA. The introduction provides a brief description of the background and discusses concept of PBS level number detection using the consistency of the auditors group as concordance (agreement) of their break indices (BI) labeling of the same text. In the next part of the paper different measures of consistency and algorithms of coordination procedures for auditors' group are discussed. An array of twenty break indices labeling of the 470-words'-spaces-long prosaic Russian text performed by 19 unprofessional auditors and one

¹ Работа выполнена при поддержке гранта РФФИ 15-06-06103 «Иерархия просодического членения звучащей речи»

highly qualified phonetician – participants in the experiment is considered as an example. By means of iterative coordination procedure of step by step removal of the “worst” auditors, i.e. those ones, whose BI labeling have the maximum deviation from current coordinate labeling, the resulting BI labeling is obtained as termwise median mean for labeled texts of most consistent resulting auditors’ group. Similar break indices labels could be obtained as a centroid of the most dense cluster whereas the formal clustering could be received by modified K-means method with cross-validation (STATISTICA, Cluster Analysis and DATA MINING sections). In the last part of the paper the dependence between the PBS levels in resulting break Indices labeling and pause localization and their timing in the same spoken text is investigated. The most probable values and confidence intervals for pause duration corresponding to various PBS levels are estimated.

Key words: prosodic boundary strength (PBS) level, consistency of the expert (auditors’) group, expertise coordination procedure, ordered classification, median mean, Kemeny median, rank distribution laws

1. Введение

В рамках исследований, поддерживаемых грантом РФФИ «Иерархия просодического членения звучащей речи», участниками проекта С. В. Князевым и Е. М. Моисеевой был поставлен перцептивный эксперимент по определению числа надежно распознаваемых уровней глубины просодических швов в звучащем тексте на базе русского языка. В данной работе описывается методы и основные результаты обработки данных, полученных в ходе этого эксперимента.

Определение числа уровней глубины ПШ, силы брейка — достаточно важная проблема, так как речь идет по существу об определении шкалы, в которой можно измерять силу брейка (глубину ПШ, PBS) в различных задачах, связанных, в том числе, с генерацией звучащей речи по тексту. Шкала должна быть достаточно подробной, чтобы обеспечить естественность звучания сгенерированной речи, но не иметь также и слишком много уровней, что может помешать построению работоспособных моделей прогноза просодического членения и определению их параметров. Существует гипотеза, что для всех языков этим требованиям лучше всего отвечает пятиуровневая шкала. Результаты наших исследований являются аргументом в пользу этой гипотезы и для русского языка.

Эксперименты по определению числа уровней силы брейков многократно проводились на базе различных языков, построен ряд признанных систем (шкал) измерения глубины просодических швов. Хорошо известны результаты А. Сандерман, описанные в монографии [7], по распознаванию глубины просодических швов (PBS) в звучащей речи, определению регрессионной связи между PBS (отклик) и рядом характеристик-предикторов, из которых наиболее значимым является длительность межсловной паузы. Построенная модель была использована для синтеза речи по тексту, который был предварительно размечен (положение и глубина ПШ) с помощью автоматической (3 уровня) или ручной (5 уровней) процедуры. В последнем случае группа экспертов, оценивавших

качество звучания (естественность, разборчивость) синтезированного текста, оценила его лишь немногим ниже записи, сделанной профессиональным диктором, что позволило сделать вывод о достаточности пятибалльной шкалы силы брейков. Эксперименты проводились на базе голландского языка. Пятиуровневая разметка и соответствующие оценки регрессии были получены с помощью псевдочисловой шкалы — искусственного укрупнения исходной десятиуровневой шкалы методом последовательных интервалов Эдвардса.

Наши уже проведенные и планируемые на ближайшее время перцептивные эксперименты во многом аналогичны экспериментам Сандерман, но во многом и отличаются. В первую очередь, это касается методики определения согласованности группы аудиторов, выделения из нее согласованной подгруппы, если часть результатов является «аномальной», и построения итоговой согласующей просодической разметки текста в соответствии с заданной шкалой. Мы исходим из положения, что когда число предлагаемых для распознавания уровней силы брейка близко к надежно определяемому аудиторами, уменьшается элемент угадывания, случайности оценок глубины ПШ, что влечет наибольшую согласованность результатов участников эксперимента именно на разметках по шкале с оптимальным числом уровней. Таким образом, эксперимент должен проводиться в несколько этапов с последовательным уменьшением или увеличением числа распознаваемых уровней глубины ПШ до получения наилучшей согласованности результатов.

Эксперименты, описанные в [7], проводились в два этапа. На основании 1-го делался вывод о согласованности группы аудиторов, на основании второго — строилась и изучалась модель. В первом этапе участвовало 3 диктора (профессионал и 2 непрофессионала) и 19 испытуемых, не имеющих проблем со слухом, но не имеющих также и специальной подготовки. Для разметки им был предложен короткий (175 словоразделов, 20 отдельных предложений), что позволило их тасовать во избежание эффекта привыкания) текст без знаков препинания и маркеров ударений, но с разметкой словоразделов, и его аудиозапись с возможностью повторных прослушиваний. Параллельно также производилась разметка существенно более короткого делексикализованного фрагмента того же текста. На основании высокого уровня корреляций результатов для обычного и делексикализованного текста был сделан вывод о возможности отказаться от существенно более трудоемкой работы с делексикализованным текстом. Был оставлен также только один профессиональный диктор. Длина текста для второго этапа составляла порядка 1000 словоразделов.

С учетом этих результатов в нашем первом эксперименте участвовал один диктор, 19 непрофессиональных аудиторов, к которым были предъявлены те же требования, и один профессионал — опытный фонетист. Делексикализованный текст не использовался. Предложенный аудиторами для прослушивания по аналогичной схеме материал имел такое же графическое и звуковое представление, но являлся связным текстом — отрывком из повести И. Грековой²

² И. Грекова — литературный псевдоним известного математика, специалиста в области теории вероятностей и математической статистики — Елены Сергеевны Вентцель.

«Кафедра», содержащим 470 словоразделов, на которых участникам эксперимента следовало сделать разметку для семи уровней брейка.

В экспериментах Сандерман вывод о согласованности группы аудиторов делается на основании близости к единице коэффициента конкордации (согласованности) Кендалла W , который может быть выражен через средний коэффициент парной корреляции Спирмена, \bar{R}_s , где усреднение производится по всем парам ранжировок длины n , выполненных m экспертами, $W = \frac{1+(m-1)\bar{R}_s}{m}$. Как и другие меры ранговой связи, W изначально определяется для **строгих** ранжировок, в которых число уровней ранжируемой переменной совпадает с длиной выборки. Если среди данных есть равные, так называемые связанные значения, в формулы для вычисления ранговых коэффициентов вводятся поправки, компенсирующие влияние **малого количества небольших групп** связанных рангов. В [7] вывод о хорошей согласованности группы делается на основании значений W равных всего лишь 0,71 и 0,61 при прослушивании записей, выполненных профессиональным и непрофессиональным дикторами.

Близость W к единице часто истолковываются исследователями как показатель согласованности группы экспертов. В случае сравнения строгих ранжировок это оправдано, так как $W = 1$ тогда и только тогда, когда все выборки упорядочены одинаково, и в этом случае его близость к 1 действительно говорит о том, что эксперты в большей части случаев присваивают объектам одинаковые ранги, то есть оценивают объекты согласовано. Если же речь идет о нестрогих ранжировках — упорядоченных классификациях с числом классов, много меньшим длины выборки, это не так. Близость W к парным ранговых коэффициентов корреляции к единице говорит лишь об «однаправленности» последовательностей оценок различных экспертов, но не о близости их абсолютных значений. На рисунке 1 представлены фрагменты графиков разметок глубины ПШ тремя аудиторами — наивными носителями русского языка. Разметки имеют одинаковую динамику роста убывания, но сильно отличающиеся (до четырех единиц по абсолютной шкале) значения.

Формально вычисленный коэффициент конкордации для исходной группы из 20 аудиторов, участвовавших в нашем эксперименте, которая никак не может быть признана согласованной (см. приводимую ниже таблицу частот и рисунок 1), превышает 0,8, наихудшее значение попарного коэффициента корреляции γ равно 0,85, лучшие же превышают 0,99. Это легко объясняется тем, что все участники достаточно адекватно оценивают относительную длительность пауз, которая является одним из основных факторов ([2, 3, 7]), влияющих на формирование оценки силы брейка, хотя и могут давать ей разные абсолютные значения.

При больших n величина $W' = n(m-1)W$ имеет асимптотическое распределение χ^2_{n-1} , что позволяет построить простой критерий неравенства коэффициента W нулю. Нулевая гипотеза $W = 0$ отвергается на уровне α , если $W' > \chi^2_{n-1, 1-\alpha}$. Но это значит лишь то, что неверно, что все ранжировки случайны и равновероятны, то есть вообще никак не согласованы [5, 6]. Критерия для проверки равенства коэффициента единице, к сожалению, нет.

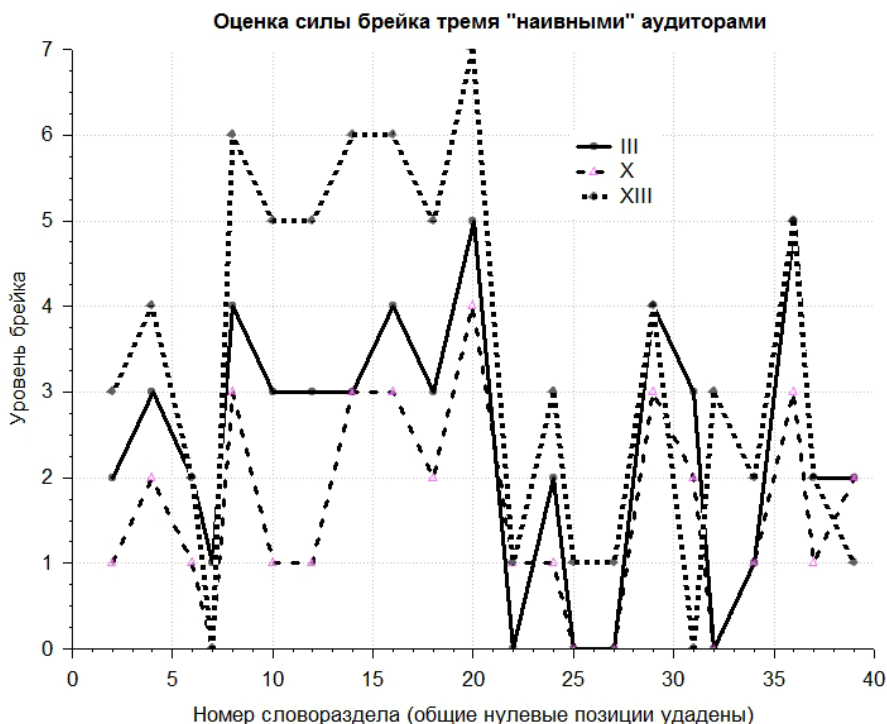


Рис. 1. Сравнительная оценка силы брейка тремя аудиторами на отрезке текста

2. Выбор наиболее согласованной подгруппы аудиторов и построение результирующей разметки

Существует много способов оценки согласованности группы экспертов, применяемых в зависимости от поставленной задачи [4, 5, 6].

Простейшей альтернативой использованию коэффициентов ранговой связи в нашем случае может быть применение критериев однородности для зависимых выборок. Действительно, если одни и те же позиции получают в основном одни и те же оценки, то и наблюдаемые частоты оценок необходимо должны быть близки. Разметки, как уже говорилось, сильно зависимы и хорошо описываются схемой последовательных наблюдений; слабой и быстро убывающей зависимостью оценок силы брейка внутри отдельных разметок мы пока пренебрегаем. Для этих критериев основной проверяемой гипотезой является именно гипотеза однородности, принадлежности выборок одному распределению. Размер критерия, то есть величина ошибки 1-го рода α , $0 \leq \alpha \leq 1$, может служить независимым от объема выборки и состава группы числовым показателем согласованности. Чем α ближе к 1, тем она выше. Наиболее известными критериями этого типа являются ANOVA Фридмана и дихотомический

критерий однородности частот Q Кохрейна, реализованные во всех статистических пакетах³. Однородность распределения частот не является достаточным условием согласованности группы, но на начальном этапе исследования ее проверка позволяет отбросить наиболее «аномальные» разметки.

Данные проведенного перцептивного эксперимента были сведены в рабочую таблицу популярного статистического пакета STATISTICA. Столбцы, обозначенные номерами участников, содержат разметки, кроме того, таблица содержит исходный текст, разметки знаков препинания, длительности пауз, итоговую разметку по согласованной подгруппе аудиторов и ряд вспомогательных данных.

Таблица 1. Частоты оценок уровней брейков

PBS	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	EM
0	88	299	321	318	308	411	311	320	341	319	321	328	317	2	16	312	318	331	300	245
1	66	10	39	48	33	35	19	23	40	28	36	35	27	202	273	34	30	30	83	52
2	182	17	49	55	46	14	52	63	26	50	63	65	53	109	56	52	46	40	34	17
3	91	35	32	25	23	5	44	49	10	47	20	34	31	67	76	27	26	23	12	13
4	32	46	18	16	15	4	16	11	10	17	18	5	16	45	14	22	6	13	4	78
5	6	13	7	3	12	1	8	2	12	7	5	2	13	20	8	16	29	30	3	19
6	1	13	3	3	18	0	16	2	10	1	3	1	8	14	23	5	9	1	0	41
7	4	37	1	2	15	0	4	0	21	1	4	0	5	11	4	2	6	2	34	5

Будем рассматривать целочисленные оценки силы брейка как категориальные числовые данные, что позволит при необходимости работать с ними и как с чисто номинальными данными, и как с ранговыми. Построим сначала таблицу частот встречаемости оценок силы бейка. Категория «0» обозначает «обычный пробел», словораздел без брейка, который далее рассматривается как брейк силы 0, то есть мы имеем дело с 8-уровневой шкалой.

Сразу можно заметить, что для большинства аудиторов частотность нуля составляет чуть больше 2/3, что хорошо согласуется с имеющимися статистическими данными и соображениями дыхательного ритма [2,3]. По результатам анализа комплексной БД (более 2000 словоразделов) текстовых признаков-предикторов, определяющих потенциальную сегментирующую силу словоразделов (по материалам коллекции прозаических текстов из книги Р. И. Аванесова «Русское литературное произношение») суммарная частота чистого пробела и брейка «-» в коллекции $\approx 72\%$. Кроме того, в распределении частот у этих аудиторов наблюдается вторая мода на брейке силы «2», которая имеется и в «аванесовской разметке». На этом основании разметки участников с номерами I,VI, IX, XIV, XV и XIX признаются «аномальными» и в дальнейшем анализе не участвуют как имеющие недопустимую частоту нуля и/или не имеющие второй моды. На первый взгляд разметки II и EM (единственного профессионала, результаты которого было бы жалко потерять) также являются аномальными: слишком мало нулей

³ Класс программного обеспечения, предназначенного для статистического анализа данных.

и мода на 4-ой позиции вместо 2-ой. Однако сравнение с другими разметками показывает, что эти аудиторы слышат очень слабые брейки, которые остальные участники практически единогласно оценивают как нулевые. После группировки и укрупнения данных эти две разметки удастся включить в дальнейший анализ. Распределение частот в укрупненной разметке ЕМ+ близко к распределению частот в «аванесовской разметке». Частоты дихотомии «пробел/брейк» признаются однородными критерием Кохрейна на высоком уровне значимости. Кроме того, весьма велика доля «чистых пробелов» (больше 80%), определяемых всеми участниками группы единогласно. Для выделяемых далее результирующих согласованных групп она превышает 95% от общего количества.

Для дальнейших экспериментов подготовлены тексты большей длины — 900 и 2700 словоразделов. Грубая оценка минимальной необходимой длины текста получена из следующих соображений. Обычно при анализе таблиц частот и таблиц сопряженных признаков предполагается, что число наблюдений достаточно велико, чтобы они были «хорошо заполнены», то есть минимальные наблюдаемые частоты n_{\min} были не меньше пяти-шести⁴. Ранговое распределение PBS в «аванесовской разметке» хорошо согласуется с модификацией закона Ципфа, полученной из условия максимизации его (распределения) энтропии [1]. Вероятность p_{\min} наиболее редко встречающегося уровня признака в этой модели приблизительно равна $1/k^2$, где k — общее число уровней исследуемого показателя. Воспользуемся этой оценкой для определения необходимой длины текста. Положим $n_1 = 2n_{\min}/p_{\min} = 2n_{\min}k^2$, то есть, предположим, что все ожидаемые частоты как минимум в два раза больше n_{\min} , что позволит их надежно различать.

Положив теперь в известной формуле Лапласа для определения необходимого числа наблюдений при оценке вероятности $p = p_{\min}$, $\varepsilon = p_{\min}/2$, получим

$$n_2 \approx \frac{(u_{(1+\gamma)/2})^2 p(1-p)}{\varepsilon^2} = \frac{4(u_{(1+\gamma)/2})^2 (1-p_{\min})}{p_{\min}}$$

где γ — надежность, $u_{(1+\gamma)/2}$ — квантиль нормального распределения, ε — погрешность, p — оцениваемая вероятность. Выберем $n = \max(n_1, n_2)$ при заданном γ . При известных n_{\min} и k для текста длины n_1 мы можем оценить надежность γ , используя соотношение

$$u_{(1+\gamma)/2} \approx \left(\frac{k^2 n_{\min}}{2k^2 - 2} \right)^{1/2}$$

Для $n_{\min} = 6$, $k = 8$ и $n = 768$ она достаточно велика, не менее 0,92 и растет с длиной текста.

Выделим теперь более согласованные подгруппы из группы неаномальных разметок. Введем расстояние между разметками [5,6]. Медианой Кемени для группы, вообще говоря нечисловых, объектов называется множество объектов того же типа, сумма расстояний до которых от всех объектов группы

⁴ Это позволяет пользоваться асимптотическими распределениями (обычно χ^2) для статистик многих популяционных критериев.

минимальна. В теории экспертных оценок медиана Кемени часто выступает в роли коллективного согласованного мнения группы экспертов, а средняя величина разброса вокруг нее служит мерой согласованности группы. В качестве расстояния между строками (векторами) числовых категориальных величин часто используется линейное расстояние, так называемое расстояние городских кварталов. В этом случае наименее уклоняющимся элементом будет являться строка, составленная из покоординатных медиан элементов группы. Это так называемое согласование по медианному среднему. В отличие от арифметического среднего оно устойчиво к выбросам, в случае нечетного объема группы удобно вычисляется и измеряется в той же шкале, не имея промежуточных, дробных значений. Расстояние по числу или доле несовпадающих координат обычно применяется для двоичных строк, согласующая строка является результатом голосования по каждой позиции. Применение этого расстояния к произвольным категориальным данным может привести к неоднозначному результату голосования даже при нечетном числе экспертов. Однако, если оценка позиции принимается абсолютным большинством голосов, чего мы вправе ожидать от согласованной группы, результат голосования совпадает с медианным средним. Доля однозначных голосований является, таким образом, важной характеристикой согласованности группы, а статистика голосования за каждый уровень силы брейка, включенная в нашу рабочую статистическую таблицу, представляет отдельный интерес, поскольку характеризует надежность определения брейка каждого уровня на каждом словоразделе.

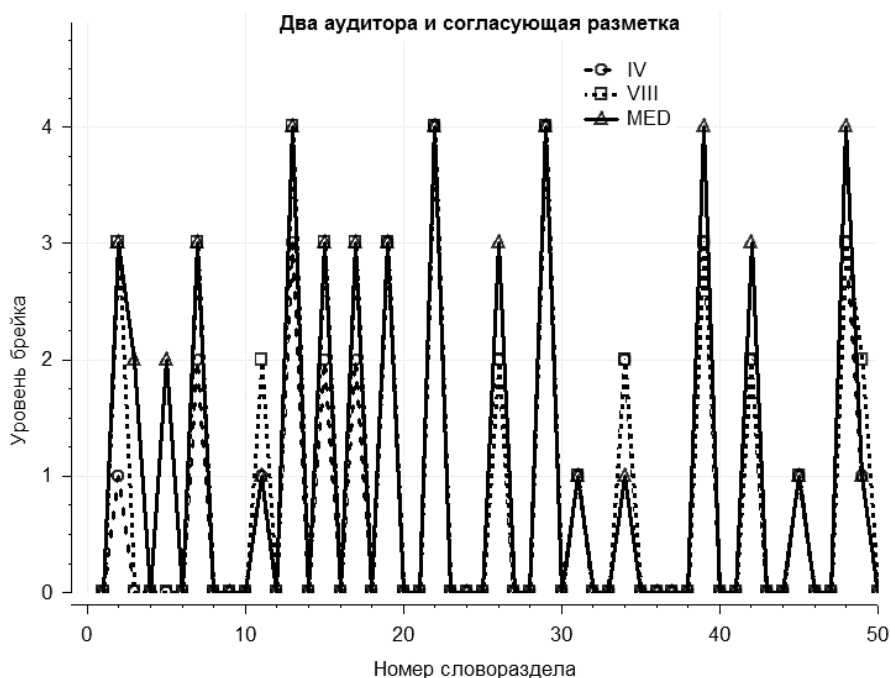


Рис. 2. Исходные разметки и результирующая (MED) последовательность

Разметки для групп из семи и пяти аудиторов практически совпадают, а неоднозначный результат голосования наблюдается лишь в 18 случаях из 470, причем голоса делятся между соседними оценками силы брейка. Группа из пяти разметок имеет неплохое согласие с гипотезой однородности по ANOVA Фридмана. На рисунке 2 показаны фрагменты итоговой разметки и разметки двух аудиторов из согласованной подгруппы.

Наконец, мы можем рассмотреть разметки как нестрогие ранжировки — классификации с упорядочиванием. Такие объекты могут быть эквивалентно представлены бинарными матрицами специального вида, расстояния между которыми определяются как число несовпадающих позиций. Для упорядоченных классификаций разработан ряд алгоритмов согласования [4,5,6], которые также могут быть применены для построения итоговой разметки, однако они существенно сложнее вычисления медианного среднего и их планируется применить при обработке результатов следующего этапа экспериментов.

Будем теперь строить подгруппу согласованных разметок с помощью следующей итеративной процедуры. На каждом шаге будем рассчитывать согласующую разметку (медианное среднее), расстояние до него для всех элементов текущей группы, и удалять из группы ту разметку, которая наиболее удалена от текущей согласующей. На первых итерациях отклонения быстро убывают, затем, начиная с группы из 9 разметок, стабилизируются. Для групп из семи и пяти наиболее согласованных разметок вычислены итоговые разметки по медианному среднему и по голосованию. Следует особо отметить, что итоговая разметка по группе из 5 аудиторов не содержит брейков уровня 7, а два вхождения уровня 6 соответствуют сверхдлинным неречевым паузам (переворачивание страницы, оговорка). То есть, уже на этом этапе эксперимента можно сделать предварительный вывод о надежном различении **пяти** уровней глубины ПШ непрофессиональными аудиторами.

Формальная кластеризация⁵ при помощи модифицированного метода К-средних с линейным расстоянием и иерархических процедур с разными метриками и мерами близости между классами выделяет один и тот же класс наиболее близких разметок из семи или восьми (с учетом укрупненной разметки II+) элементов, совпадающий с группой того же объема, полученной с помощью пошаговой процедуры. Центроид этого класса дает результирующую разметку текста близкую к полученной методом медианных средних. На дендрограмме (рис. 3) видна также подгруппа из шести наиболее близких разметок — EM+, XII, XI, VIII, IV и III, которая не входит в итоговую группу из пяти разметок при пошаговом алгоритме.

⁵ пакет STATISTICA, модули пилотного многомерного анализа (кластерный анализ) и DATA MINING, (дополнительные возможности кластерного анализа).

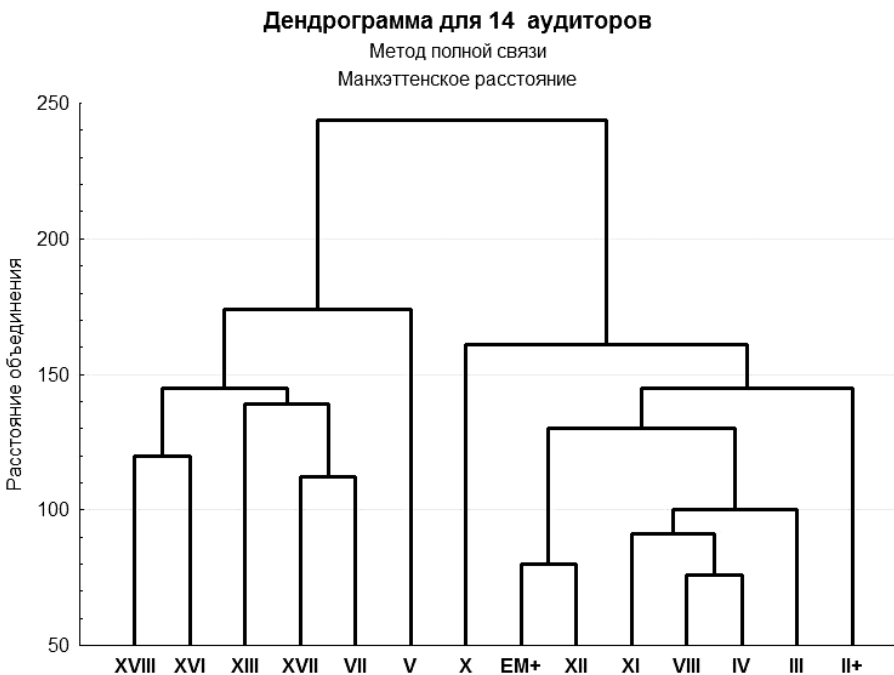


Рис. 3. Иерархическая классификация разметок

3. Связь силы брейков и длительности пауз

Полученные результирующие разметки были использованы для изучения связи между силой брейков и длительностью пауз. На рисунке 4 приведен фрагмент кумулятивной гистограммы для длительности пауз между фонетическими фрагментами с числом интервалов 300. Хорошо видны несколько участков быстрого роста, что наводит на мысль, что распределение длительностей является на самом деле смесью более простых распределений. Действительно, классификация наблюдений по силе брейка в результирующей последовательности, анализ дескриптивной статистики и гистограмм распределения пауз по классам показывает, что эти участки соответствуют наиболее вероятным значениям пауз при различных значениях силы брейка. Рисунок 4 демонстрирует наложение на общую гистограмму гистограммы плотности длительностей пауз для брейка силы 3, максимум которой попадает на участок роста в районе 800 мс. Диапазоны длительностей для разных уровней глубины ПШ пересекаются, однако не пересекаются их квартильные размахи, что позволяет построить для них диапазоны, доверительные интервалы с достаточно большими вероятностями попадания, включающие в себя наиболее вероятные значения.

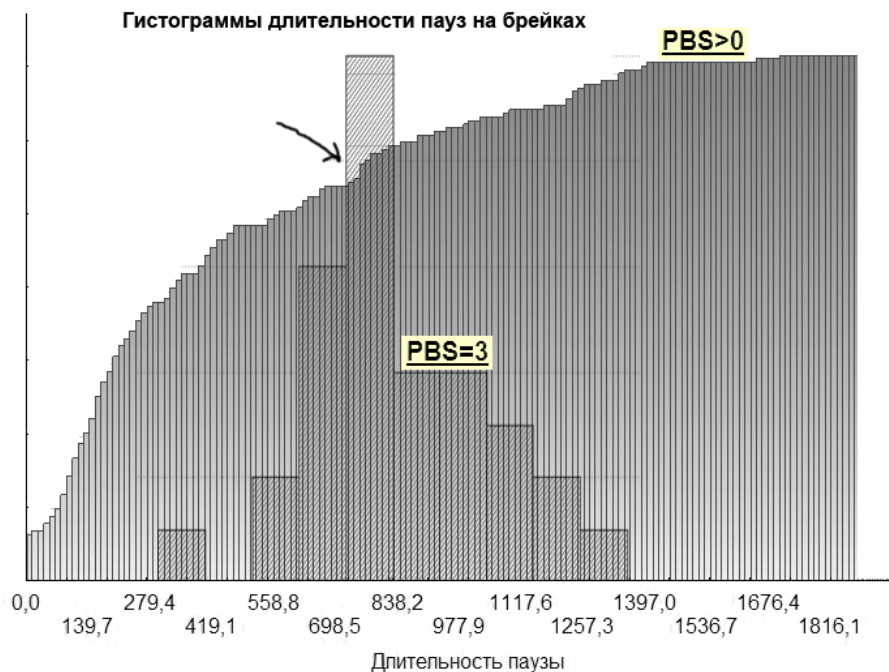


Рис 4. Гистограммы длительности пауз в миллисекундах

Это 104–450 мс с вероятностью попадания 0,93 и наиболее вероятным значением порядка 400 мс для силы брейка меньше или равной 2, 450–1150 (0,82, ≈800 мс) для силы брейка 3, больше 1150 (0,83, ≈1200 мс) для силы брейка больше или равной 4. Эти оценки хорошо согласуются с результатами О. Ф. Кривновой. Разделить диапазоны для слабых брейков таким методом не удастся. По-видимому, это объясняется тем, что для их формирования более значимы факторы, отличные от длительности пауз. В частности, с большой вероятностью наблюдаются непаузальные брейки силы 1 и даже силы 2.

Кластеризация длительности пауз методом К-средних с растущим числом классов позволяет выделить устойчивые классы со следующими центроидами: 0, 800, 1200, 2000 миллисекунд, которые почти не меняются при увеличении числа классов вручную или при работе алгоритма в режиме cross validation. Эти классы хорошо ассоциируются с отсутствием брейка, брейками силы 3, 4, >4 (5), а их центроиды близки к соответствующим наиболее вероятным значениям. С ростом числа классов происходит более подробная классификация только коротких пауз. Выделяются классы с центроидами порядка 400, 200, 100 мс (все значения округлены), соответствующие брейкам малой силы. При дальнейшем увеличении числа классов результаты становятся трудноинтерпретируемыми.

4. Заключение

В результате проведенной работы опробована методика проверки согласованности группы аудиторов, выделения согласованной подгруппы для участия в дальнейших экспериментах и получения итоговой разметки. В результате сокращения исходной 7-бальной шкалы до 5-бальной при построении согласующей разметки уже на первом этапе эксперимента можно сделать предварительный вывод, что гипотеза об оптимальности 5-бальной шкалы верна и для русского языка. Оценены наиболее вероятные значения и получена система доверительных интервалов для длительности пауз при разных уровнях брейка.

В настоящее время готовится серия новых перцептивных экспериментов с использованием существенно более длинного текста, что позволит сделать выводы статистического анализа более достоверными. Для согласования разметок предполагается использовать алгоритмы для упорядоченных классификаций.

Литература

1. *Gusejn-Zade S. M.* (1987), On occurrence of key words and another ranked series. [O vstrechaemosti klyuchevykh slov i o drugih ranzhirovannykh ryadah], Scientific and technological information, ser. 2 [NTI Ser. 2], №1, pp. 28–32.
2. *Krivnova O. F.* (2015), The Depth of Prosodic Breaks in Spoken Text (Experimental Data) [Glubina prosodicheskikh shvov v zvuchashchem tekste (ehksp'imental'nye dannye)], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2015” [Komp'yuternaya Lingvistika i Intel'ektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2015”], Moscow, v. 1, pp. 338–351.
3. *Krivnova O. F.* (2016), Prosodic phrasing in spoken text: localization of breathing pauses [Prosodicheskoe chlenenie zvuchashchego teksta: tekstovaya lokalizatsiya dyhatel'nykh pauz]], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2015” [Komp'yuternaya Lingvistika i Intel'ektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2015”], Moscow, pp. 340–354.
4. *Litvak B. G.* (1982), Expert information. Methods for the preparation and analysis [Ekspertnaya informatsiya. Metody polucheniya i analiza] Radio and Communications, Moscow [Radio I Svyaz', Moskva].
5. *Orlov A. I.* (2004), Nonnumeric statistics [Nechislovaya statistika], MZ-Press, Moscow [MZ-Press, Moskva].
6. *Orlov A. I.* (2004), Applied statistics [Prikladnaya statistika], Exam, Moscow [EHkzamen, Moskva].
7. *Sanderman A.* (1996), Prosodic phrasing (production, perception, acceptability and comprehension), Eindhoven.