



АНАЛИЗ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ДИКТОРА ПО ГОЛОСУ НА ОСНОВЕ ФОНЕТИЧЕСКОГО ДЕТЕКТОРА ЛЖИ

В.В. САВЧЕНКО
Р.А. ВАСИЛЬЕВ

*Нижегородский
государственный
лингвистический
университет*

*e-mail:
svv@lunn.ru
romangamma@mail.ru*

При учете проблемы многокритериальности качества устной (звучащей) речи предложен критерий минимума требуемой избыточности голосового сигнала. На его основе представлены программная реализация и результаты экспериментального исследования фонетического детектора лжи в задаче тестирования эмоционального состояния личности по голосу в комфортных и некомфортных условиях.

Ключевые слова: информатика, человеко-машинное взаимодействие, речевые технологии, информационная теория качества речи, фонетический детектор лжи, психолингвистика.

Фонетический детектор лжи (ФДЛ) – это зарегистрированная в Роспатенте программа для ЭВМ [1], предназначенная для тестирования эмоционального состояния личности по голосу. В ее названии отражены особенности принципа действия ФДЛ, а именно: анализ фонетического строя речи диктора в зависимости от текущего эмоционального состояния последнего. Общеизвестный к нему интерес со стороны не только специалистов, но и разнообразных отечественных СМИ, продиктован, прежде всего, его принципиально новыми возможностями, а именно: высокой чувствительностью к отклонениям в эмоциональном состоянии диктора при минимальных требованиях (1-2 минуты) к продолжительности анализируемого фрагмента голосового сигнала. Объяснением этого служит принципиально новый принцип действия ФДЛ, основанный на автоматической оценке качества речи диктора на базовом, фонетическом уровне по общесистемному шенноновскому критерию минимума требуемой избыточности (МТИ) речевого сигнала. По существу, это первая попытка в мире – в теории и на практике – преодолеть острейшую проблему многокритериальности устной речи с позиций строгого, теоретико-информационного подхода [2]. Тем больший интерес для специалистов широкого профиля представляют публикуемые далее результаты теоретического и экспериментального исследования ФДЛ в задаче тестирования эмоционального состояния личности по голосу.

Принцип действия большинства современных систем автоматического анализа речи на фонетическом уровне основывается [3] на последовательном членении голосового сигнала на короткие (5-10 мс) отрезки данных $x = (x_1, x_2, \dots, x_n)$ длиной в одну МРЕ с их последующим сопоставлением по тонкой, в частности, спектральной структуре [4] с соответствующим эталоном. Главной проблемой для таких систем является выбор и обоснование множества фонетических эталонов $\{x_r^*\}$.

Известно, что любой диктор в силу ряда причин, например, из-за особенностей своей речи или слуха, в принципе не в состоянии в процессе речеобразования точно воспроизвести эталон x_r^* той или иной (r -й) МРЕ. Выходом из этой ситуации может служить задание каждой МРЕ не одной, а одновременно несколькими допустимыми вариантами $x_{r,j}$, $j = \overline{1, J_r}$, где $r = \overline{1, R}$, а R – объем фонетической базы данных (ФБД). В таком случае диктору будет достаточно приблизить свое произношение к любому из них, чтобы быть правильно понятым гипотетическим наблюдателем или слушателем. Этим существенно ослабляется рассматриваемая проблема вариативности устной речи: каждый конкретный диктор в процессе своего «говорения» выбирает наиболее удобный, достижимый для себя



вариант эталонного произношения МРЕ из некоторого множества альтернатив $\{\mathbf{x}_{r,j}\}$. Одновременно становится понятным и собственно критерий качества формируемого (на выходе голосового тракта диктора) голосового сигнала к эталону: он должен войти в границы J_r -множества вариантов рассматриваемой МРЕ X_r как полноправный, $(J_r + 1)$ -й его элемент. Задача переходит, в таком случае, в сугубо предметную плоскость: сначала по каждой из R рассматриваемых МРЕ требуется сформировать множество (кластер) $X_r = \Delta\{\mathbf{x}_{r,j}\}$ ее допустимых образцов – на этапе обучения диктора. И после этого в процессе речеобразования тестировать текущий сигнал \mathbf{x} от диктора согласно правилу близости к ним

$$J_r^{-1} \sum_{j=1}^{J_r} \rho(\mathbf{x} / \mathbf{x}_{r,j}) \leq \rho_0, \tag{1}$$

в среднем в пределах кластера в некоторой метрике $\rho(\mathbf{x} / \mathbf{x}_{r,j})$. При достаточной степени малости порогового уровня ρ_0 и выполнении требования (1) качество речи диктора можно оценить на достаточно высоком уровне. И, наоборот, при нарушении требования (1) соответствующая (текущая) МРЕ должна быть забракована наблюдателем как ошибка речеобразования.

Отметим важную отличительную особенность правила (1): в каждый момент времени решение может быть принято в пользу либо одной, либо двух, а вообще говоря, и нескольких МРЕ из используемой диктором ФБД $\{X_r\}$, либо вообще не принято – для сигналов \mathbf{x} нечеткой (маргинальной) структуры. И это точно соответствует практике речеобразования [1, 2]: в ней не исключаются сбой и, как их результат, брак.

В вычислительном отношении проще, однако, задаться аналогичным (1) условием тестирования МРЕ вида

$$\rho_r(\mathbf{x}) \leq \rho_0 \tag{2}$$

на расстояние от сигнала \mathbf{x} до «центра массы» рассматриваемого кластера

$$\mathbf{x}_r^* = \mathbf{x}_{r,v} : J_r^{-1} \sum_{j=1}^{J_r} \rho(\mathbf{x}_{r,j} / \mathbf{x}_{r,v}) = \min_{i \leq J_r} J_r^{-1} \sum_{j=1}^{J_r} \rho(\mathbf{x}_{r,j} / \mathbf{x}_{r,i}) = \rho_r^*. \tag{3}$$

В режиме реального времени (в процессе восприятия речи) вместо $J_r \gg 1$ расстояний в (1) здесь вычисляется только одно расстояние $\rho_r(\mathbf{x}) = \rho(\mathbf{x} / \mathbf{x}_r^*)$ в пределах кластера

X_r : до его центра \mathbf{x}_r^* . Здесь Δ – равенство по определению. Указанный центр – это обобщенный эталон данного кластера, или эталон соответствующей фонемы. А множество таких эталонов $\{\mathbf{x}_r^*\}$ – экономный способ задания ФБД конкретного диктора, или звукового ряда $\{X_r\}$ (строка) его разговорного языка. В информационной теории качества речи в роли расстояний между аллофонами в (3) используются величина информационного рассогласования (ВИР) по Кульбаку-Лейблеру [5].



Следуя распространенной авторегрессионной (АР) модели голосового сигнала, будем иметь следующее выражение для оптимальной решающей статистики из выражения (2) в частотной области:

$$\rho_r(\mathbf{x}) = F^{-1} \sum_{f=1}^F \frac{\left| 1 - \sum_{m=1}^P a_r(m) \exp(-j\pi m f / F) \right|^2}{\left| 1 - \sum_{m=1}^P a_x(m) \exp(-j\pi m f / F) \right|^2} - 1. \quad (4)$$

Здесь $\{a_x(m)\}$ $\{a_r(m)\}$ - векторы АР-коэффициентов тестируемого сигнала \mathbf{X} и r -го

эталона \mathbf{x}_r^* соответственно, причем оба – одного порядка $p > 1$; f – дискретная частота, F – ее верхняя граница, равная половине частоты дискретизации голосового сигнала. Это стандартная [6] формулировка метода обеляющего фильтра и критерия минимума информационного рассогласования (МИР). Его программно-аппаратная реализация основывается на применении быстрых вычислительных процедур АР-анализа, в частности алгоритма Берга-Левинсона [7].

Выражение (4) совместно с (2) и (3) определяет алгоритм поэлементного (по каждой МРЕ) восприятия речи от некоторого диктора условным (идеальным) слушателем. Как всякое восприятие оно сопровождается двумя альтернативными вариантами решения: сигнал \mathbf{X} признается отвечающим требованиям к качеству r -й МРЕ, или, напротив, требования не выполнены и сигнал бракуется как ошибочный. Вероятность ошибки в общем случае [2]

$$\alpha_r = P \left\{ \rho_r(\mathbf{x}) \geq \rho_0 \mid \mathbf{x} \in X_r \right\} = P \left\{ \chi_M^2 > M(1 + \rho_0) \right\} = \alpha \quad (5)$$

определяется в терминах χ_M^2 -распределения с M степенями свободы, где $M = n - p$.

При равенстве $\rho_0 = const$ будем иметь $\alpha_r = \alpha = const$ для всех МРЕ из звукового строя $\{X_r\}$. В таком случае вероятность ошибки речеобразования (5) может служить обобщенной оценкой качества речи диктора.

Следуя общим идеям информационной теории [2], будем рассматривать каждого диктора в роли условного источника дискретных сообщений $X \subset \{X_r\}$, определенных на

R -множестве его МРЕ с ФБД $\{\mathbf{x}_r^*\}$. Исчерпывающей характеристикой коммуникативных свойств такого источника может служить по К. Шеннону его скорость создания информации, или количество информации на выходе в расчете на одну МРЕ. В предположении об идеальном голосовом механизме человека-диктора и безошибочном восприятии всего набора его МРЕ потенциальным слушателем указанная величина определяется классическим выражением для шенноновской энтропии дискретного источника сообщений вида

$$H(X) = - \sum_{r=1}^R P(X = \mathbf{x}_r^*) \log P(X = \mathbf{x}_r^*) = - \sum_{r=1}^R p_r \log p_r.$$

При этом необходимо учитывать естественное условие нормировки закона распре-

деления: $\sum_{r=1}^R p_r = 1$. В простейшем случае равновероятных МРЕ, когда $\forall r \leq R: p_r = 1/R$,

получим максимум энтропии $H(X) = \log R$, [бит]. Здесь логарифм берется по основанию 2.



Ситуация, однако, резко усложняется, если учитывать проблему вариативности устной речи [2]. В общем случае сигнал на выходе речевого тракта диктора $X' \neq X$ может существенно отличаться от его эталонного варианта \mathbf{x}_r^* в сознании диктора. Математическая модель такого источника сообщений определяется распределением вероятностей четких МРЕ

$$q_r \stackrel{\Delta}{=} P(X' = \mathbf{x}_r^*), r = \overline{1, R},$$

а также вероятностью дополнительного, (R+1)-го состояния источника

$$q_{R+1} \stackrel{\Delta}{=} P(X' \neq \mathbf{x}_r^* \forall r \leq R),$$

в которой учтены возможные ошибки речеобразования. Это типичная [8] модель дискретного канала связи с шумами (без памяти). Исчерпывающим показателем качества такого канала может служить шенноновское количество информации на его выходе в отношении источника сообщений:

$$I(X, X') \stackrel{\Delta}{=} H(X) - H(X|X').$$

Апостериорная энтропия источника сообщений $H(X|X')$ имеет в данном случае смысл величины рассеяния полезной информации в процессе речеобразования, или минимальной требуемой избыточности голосового сигнала. Чем больше рассеяние, тем выше степень искажений формируемого на выходе речевого тракта сигнала X' по сравнению с его эталоном \mathbf{x}_r^* , и тем ниже, следовательно, качество речи данного диктора. По сути, выражение (6) определяет скорость создания информации в расчете на одну МРЕ при учете действия шумов, т.е. в расчете на возможные искажения речи. Нетрудно понять, что этим строго сформулирован объективный показатель качества речи. Отталкиваясь от него, получим строго объективную оценку качества речи по конечной выборке наблюдений.

Для заданной решающей схемы (2) путем несложных вычислений будем иметь систему равенств

$$\begin{aligned} q_r &= \sum_{\nu=1}^R q_{r\nu} = \sum_{\nu=1}^R P(X' = \mathbf{x}_r^*; X = \mathbf{x}_\nu^*) = \sum_{\nu=1}^R P(X = \mathbf{x}_\nu^*) \cdot P(X' = \mathbf{x}_r^* | X = \mathbf{x}_\nu^*) = \\ &= P(X = \mathbf{x}_r^*) \cdot P(X' = \mathbf{x}_r^* | X = \mathbf{x}_r^*) = (1 - \alpha) p_r, \quad r = \overline{1, R}, \\ q_{R+1} &= \sum_{\nu=1}^R P(X' \neq \mathbf{x}_\nu^*; X = \mathbf{x}_\nu^*) = \\ &= \sum_{\nu=1}^R P(X = \mathbf{x}_\nu^*) \cdot P(X' \neq \mathbf{x}_\nu^* | X = \mathbf{x}_\nu^*) = \alpha, \end{aligned}$$

При этом выполняется, разумеется, условие нормировки распределения вида $\sum_{r=1}^{R+1} q_r = \sum_{r=1}^R p_r \equiv 1$. Отсюда вытекает окончательный результат

$$I(X, X') = (1 - \alpha)H(X), \text{ [бит]} \tag{6}$$

в отношении скорости создания информации на выходе голосового механизма человека. Отметим важную деталь: полученное выражение хорошо согласуется с известным [8] неравенством Фано вида

$$H(X|X') \leq -\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha) + \alpha \cdot \log(R - 1).$$



И этим дается еще одно подтверждение обоснованности теоретико-информационного показателя качества речи (6). При этом вероятность ошибки речеобразования α определяет относительную величину требуемой избыточности (ОВТИ) голосового сигнала:

$$\delta H(X|X') = \frac{\Delta H(X|X')}{H(X)} = \alpha. \quad (7)$$

Ее практическая реализация сводится к оценке вероятности ошибки (5) по формуле относительной частоты

$$\alpha = m_0 / N \times 100\% \quad (8)$$

случайного события (Лапласа), состоящего в том, что в серии из N последовательных наблюдений над сигналом \mathbf{x} решающее правило (2) выполнится m_r раз для каждой r -й МРЕ и, соответственно, будет не выполнено в остальных $m_0 = N - \sum_{r=1}^R m_r$ случаях.

Выражения (2)-(8) в совокупности определяют теоретико-информационную оценку качества речи диктора по конечному фрагменту его голосового сигнала. Ее аппаратно-программная реализация подробно описана в авторском патенте на полезную модель [9]. При этом точность и надежность полученной оценки нетрудно оценить с помощью классического неравенства Бернулли:

$$P\left\{m_0 / N - \alpha \leq \varepsilon\right\} > 1 - \alpha(1 - \alpha) / (\varepsilon^2 N) = \beta.$$

Задаваясь допустимой шириной доверительного интервала $\varepsilon \leq 0,05$, для случая $\alpha \leq 0,1$ будем иметь доверительную вероятность β порядка 0,99 и выше при объеме выборки N , равном 10 тысячам единиц и более. Что это значит на практике? Если учесть, что размерность вектора \mathbf{x} составляет [2] минимум 80 отсчетов голосового сигнала при частоте дискретизации 8 кГц, объем выборки N достигает требуемого значения на интервале наблюдения длиной порядка полутора-двух минут. Это говорит о беспрецедентно высоких динамических свойствах предложенной оценки. Причем, на практике именно относительная величина (7) часто представляется предпочтительной по сравнению с абсолютной величиной теоретико-информационного показателя качества речи (6). Например, это справедливо в задачах психологического тестирования личности по принципу сопоставления двух ОВТИ из выражения (8): полученных до и после нагрузки на диктора. Задача такого рода подробно рассмотрена далее – в качестве предмета экспериментальных исследований.

Для экспериментальных исследований информационной оценки качества устной речи (6), (7) был разработан экспериментальный образец информационной системы (ИС). Он реализован в виде компьютерной программы PLD (от англ. «Phonetic Lie Detector») на платформе Java SE 1.6. Программа позволяет выполнять все операции над голосовым сигналом \mathbf{x} согласно алгоритму обработки (2) – (4). Ее главное окно показано на рисунке 1.

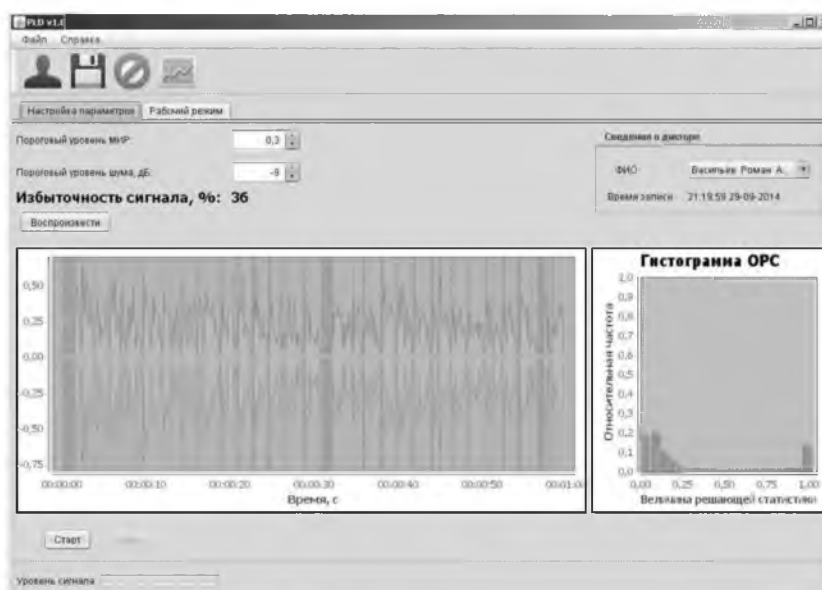


Рис. 1. Главное окно программы информационной оценки качества устной речи

Здесь под «Избыточностью сигнала» понимается выборочная оценка ОВТИ согласно выражению (8). А «Пороговый уровень МИР» $\mu_0 \leq 1$ определяет правую часть неравенства (2) согласно пропорциональной зависимости $\rho_0 = \mu_0 K - 1$, где K – коэффициент пропорциональности, равный максимальному на множестве МРЕ $\left\{ \kappa_r^* \right\}$ коэффициенту подавления голосового сигнала в r -м обесцвечивающем фильтре из выражения (4). И, наконец, «Пороговый уровень шума, дБ» – это стандартный параметр амплитудной селекции голосового сигнала из его аддитивной смеси с фоновым шумом микрофона. В процессе экспериментальных исследований он устанавливался на уровне 3-4 СКО шума в отсутствие полезного сигнала.

Для экспериментальных исследований была выбрана группа из четырех дикторов: три мужчины разного возраста и примерно одного уровня образования и одна женщина, все без явно выраженных дефектов речи. Каждым из них сначала (на этапе настройки параметров ИС) был прочитан в среднем темпе один и тот же художественный текст – из первой главы романа А.С. Пушкина "Капитанская дочка" – объемом в одну стандартную машинописную страницу. Полученные голосовые сигналы сначала были записаны в память ПК в формате wav. Его роль выполнял современный ноутбук Acer Extensa 5620, 2 Гбайт ОЗУ, Windows 7. Кроме того, использовался комплекс специальных аппаратных и программных средств, в том числе микрофон Genius и программа Audacity 1.2. Частота дискретизации встроенного АЦП была установлена равной 8 кГц – это стандартное ее значение при обработке разговорной речи.

На подготовительном этапе была проверена работоспособность базовых функций ИС, таких как корректность процедуры загрузки и завершения работы, правильность работы с дикторами и с группами дикторов. После этого исследования выполнялись в несколько этапов:

- исследование фонетических особенностей речи контрольной группы дикторов в комфортных условиях по информационному показателю (критерию) качества речи (7);
 - исследование влияния физической нагрузки на диктора на качество его речи;
 - исследование влияния эмоционального напряжения диктора на качество его речи.
- В состав контрольной группы были включены (с их согласия) следующие физические лица:

- 1) Тузалин Юрий Алексеевич, 1953 г.р., ведущий инженер ФГУП «НПП «Гамма»,
- 2) Вахтин Дмитрий Анатольевич, 1984 г.р., нач. отдела ФГУП «НПП «Гамма»,
- 3) Васильев Роман Александрович, 1987 г.р., аспирант НГЛУ,



4) Васильева Елена Николаевна, 1987 г.р., инженер ООО «ШнЭл».

Для каждого из них в режиме настройки ИС сначала вычислялась предварительная оценка ОВТИ речи диктора α_0 . Продолжительность голосового сигнала здесь составляла примерно одну минуту. И далее, уже в рабочем режиме, с использованием предварительной оценки исследовалась динамика ОВТИ в зависимости от условий его монолога. Соответствующее окно программы показано на рисунке 2.

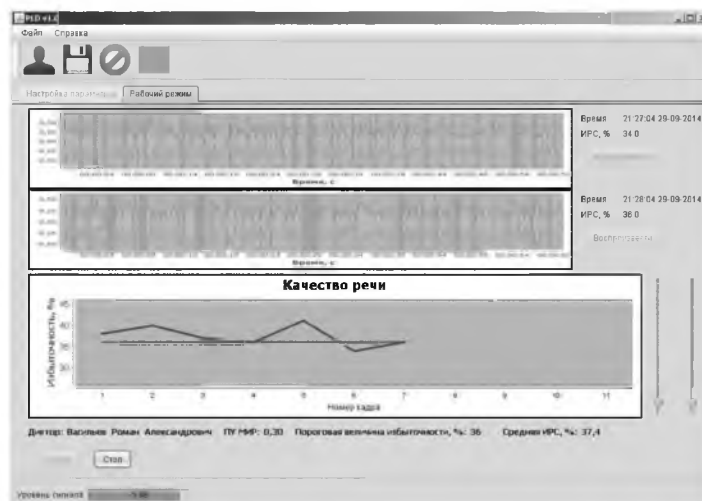


Рис. 2. Рабочий режим программы информационной оценки качества устной речи

Полученные результаты представлены ниже в виде следующих рисунков и таблиц.

В таблице 1 приведены оценки ОВТИ в зависимости от порога μ_0 по десяти реализациям голосового сигнала от первого диктора (Тузалина).

Таблица 1

Оценки относительной величины требуемой избыточности для одного диктора

Номер реализации	Порог МИР							
	0,01	0,02	0,05	0,1	0,3	0,5	0,7	1
1	80	71	49	38	40	35	31	24
2	81	69	48	37	40	33	30	25
3	83	67	45	35	42	36	31	24
4	86	70	42	36	43	35	34	26
5	83	72	43	34	41	34	33	21
6	85	63	42	34	42	37	31	21
7	84	67	44	33	42	38	35	22
8	78	66	50	40	36	32	34	26
9	81	69	43	35	43	31	34	23
10	83	65	46	37	37	38	32	21

Аналогичные результаты были получены для всех других дикторов. Хорошо видно, что предложенный показатель качества речи диктора практически инвариантен к выбору текста для чтения, времени и длительности его записи и, вместе с тем, сильно критичен по отношению к пороговому уровню МИР, а также к личности диктора. Усредненные (по множеству из десяти реализаций) оценки ОВТИ для всех четырех дикторов из нашей контрольной группы представлены в табл. 2.

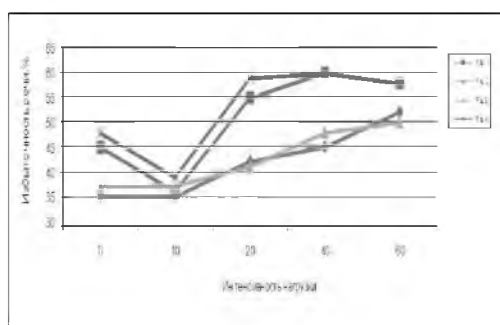


Таблица 2

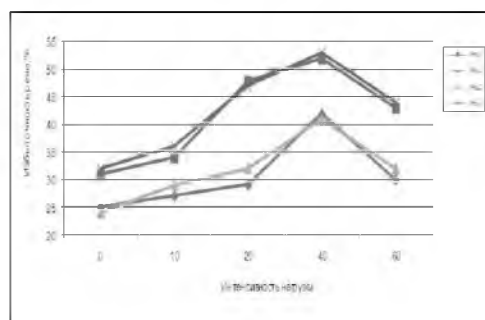
Оценки относительной величины требуемой избыточности для группы дикторов

Диктор	Порог МИР							
	0.01	0.02	0.05	0.1	0.3	0.5	0.7	1
1. Тузалин	80	71	49	38	40	35	31	24
2. Вахтин	79	68	59	43	36	31	22	30
3. Васильев	89	64	58	39	36	37	33	23
4. Васильева	86	78	60	53	28	39	29	28

Влияние физической нагрузки на качество речи дикторов – цель второго этапа экспериментальных исследований. Здесь каждый диктор сначала приседал несколько раз и только после этого читал свой текст в течение одной-двух минут. Оценки показателя качества речи всех четырех дикторов в зависимости от интенсивности нагрузки для двух значений порогового уровня МИР 0,1 и 0,7 отражены семействами кривых на рисунках 3, а) и 3, б) соответственно.



а)



б)

Рис. 3. Зависимость относительной величины требуемой избыточности от физической нагрузки

Здесь номера кривых отвечают номерам дикторов в контрольной группе.

Видно, что при увеличении нагрузки избыточность в общем случае возрастает, причем, сначала резко, а потом, по мере вовлечения организма диктора в режим интенсивной физической нагрузки, медленнее. У спортсменов (диктор Васильев и Васильева) процесс адаптации протекает быстрее.

На заключительном этапе каждый диктор читал в течение одного часа роман А.С. Пушкина «Евгений Онегин». Во второй половине часа каждые 5 минут дикторы проводили измерения ОВТИ при фиксированном пороге МИР 0,1. Усредненные (на множестве реализаций) результаты по всей группе дикторов отражены семейством кривых на рис. 4.

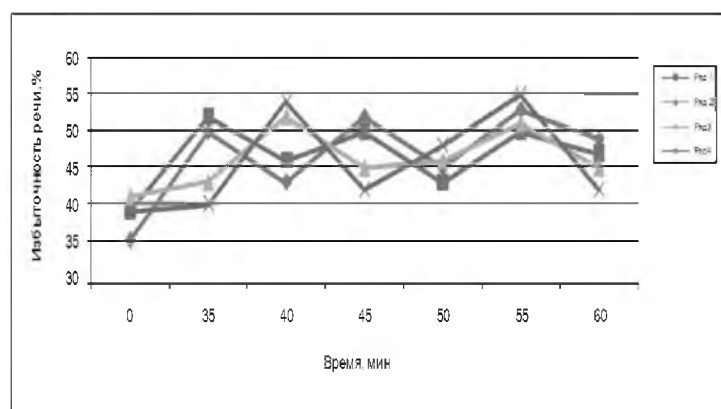


Рис. 4. Оценки относительной величины требуемой избыточности в зависимости от эмоциональной нагрузки на дикторов



Здесь хорошо видна тенденция увеличения избыточности речи при длительном эмоциональном напряжении диктора. При этом динамика избыточности имеет характер колебаний – синхронно с колебаниями степени сосредоточенности диктора на конкретном тексте. Причем, у молодого диктора Васильева (кривая 3) колебания имеют большую амплитуду: до $(52-41)/41 \times 100 = 26,8\%$ и длятся дольше, чем у диктора Тузалина (кривая 1), в силу его (Васильева) естественной (для этого возраста) недостаточной сосредоточенности.

Аналогичные выводы можно сделать по дикторам Вахтину и Васильевой: у молодого диктора Васильевой (кривая 4) колебания имеют большую амплитуду и длятся дольше, чем у диктора Вахтина (кривая 2). Отметим, что в общем случае указанные колебания затухают во времени, причем, на определенном, повышенном уровне избыточности речи диктора.

По результатам проведенных исследований можно сделать следующие выводы:

- подтверждена устойчивость предложенного информационного показателя качества речи диктора (7), (8) на разных текстах и в разное время ее записи;
- экспериментально подтверждена прямо пропорциональная зависимость ОВТИ от интенсивности физической нагрузки на диктора;
- установлена высокая чувствительность ОВТИ по отношению к эмоциональным нагрузкам на диктора в процессе его монолога.

Таким образом, в результате проведенного исследования дано экспериментальное обоснование принципа минимума требуемой избыточности в роли информационного показателя качества речи диктора, который нацелен не на сравнение речи разных дикторов между собой, а на исследование влияния разного рода факторов на качество речи конкретного диктора. По-видимому, именно по направлению психолингвистики предложенная информационная оценка качества речи может найти наиболее широкое практическое применение. Анализируя колебания ОВТИ в процессе речеобразования относительно ее значения $\bar{\alpha}_0$ в заведомо комфортных условиях, мы можем установить как факт отклонения психологического состояния диктора от нормы, так и степень указанного отклонения, причем, с одновременной оценкой допустимых пиковых значений нагрузки на данного диктора.

Список литературы

1. Информационная система тестирования эмоционального состояния личности по голосу: Программа для ЭВМ /А.В. Савченко, В.В. Савченко, Д.Ю. Акатьев, И.В. Губочкин / Роспатент: рег. № 2013611003 от 09.01.2013.
2. Савченко В.В. Информационная теория качества речи // Изв. вузов России. Радиоэлектроника. 2011. Вып. 1. С. 17-27.
3. Белов А.С., Курлов А.В., Фирсова А.А. О различии концентрации энергии по частотным диапазонам на отрезках сигналов, соответствующих шипящим звукам русской речи и шумам // журнал «Научные ведомости БелГУ»: Серия «История. Политология. Экономика. Информатика». 2011. №13(108). Вып. 19/1. С. 186-190.
4. Белов С.П., Белов А.С. О различиях частотных свойств информационных и неинформационных звуковых сигналов речевого диапазона // журнал «Научные ведомости БелГУ»: Серия «История. Политология. Экономика. Информатика». 2008. №10(50). Вып. 8/1. С. 86-93.
5. Кульбак С. Теория информации и статистика. М.: Наука, 1967. 408с.
6. Савченко В.В. Автоматическая обработка речи по критерию минимума информационного рассогласования на основе метода обесцвечивающего фильтра // Радиотехника и электроника. 2005. Т50. №3. С.309-314.
7. Марпл С.Л.-мл. Цифровой спектральный анализ и его приложения. М.: Мир, 1990. 584с.
8. Файнштейн. А. Основы теории информации: Пер. с англ. М.: Изд-во иностр. лит-ры, 1960. 136 с.
9. Патент на полезную модель № 102138. Устройство для фонетического анализа и тестирования речи. / В.В. Савченко, Д.Ю. Акатьев / Роспатент: по заявке № 2010135569 от 25.08.2010. Оп. в БИ, 2011, №3.



THE ANALYSIS OF THE EMOTIONAL CONDITION OF THE ANNOUNCER ON THE VOICE ON THE BASIS OF THE PHONETIC LIE DETECTOR

V. V. SAVCHENKO
R. A. VASILYEV

*Nizhny Novgorod State
Linguistic University*

e-mail:
svv@lunn.ru
romangamma@mail.ru

When taking into account the quality of oral problems multicriteriality (sounding) speech proposed criterion required minimum redundancy of the voice signal. On the basis of its software implementation and presents the results of experimental study of the phonetic lie detector test in the problem of the emotional state of the person 's voice in a comfortable and uncomfortable conditions .

Keywords: computer science, human-computer interaction, speech technology, information theory of speech quality phonetic lie detector, psycholinguistics .