

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕКОТОРЫХ МЕТОДОВ ОЦЕНКИ РАЗБОРЧИВОСТИ РЕЧИ

А.В.ГАВРИЛЕНКО¹, В.С.ДИДКОВСКИЙ², А.Н.ПРОДЕУС²

¹Отдел специальных телекоммуникационных систем
и защиты информации СБ Украины, Киев

²Национальный технический университет Украины “КПИ”, Киев

A brief review of the objective methods of speech intelligibility estimation is realized. It's shown that the most of formant methods are close in the region of middle values of signal-to-noise ratio (SNR). But it's difficult to take into account the reverberation interference when formant methods are used. The concept of Modulation Transfer Function (MTF) is more powerful in this sense because it permits to take into account the reverberation interference. Empirical method %ALcons accounts the reverberation interference, but it doesn't consider noise. The most complicated task for all above mentioned methods is to determine speech privacy.

ВВЕДЕНИЕ

Первая работа Колларда, посвященная оценке разборчивости речи, датируется 1929 годом [1]. В настоящее время насчитывается несколько десятков методов расчета и измерения разборчивости речи [2-30], применяющихся для оценки качества акустики помещений, линий связи, а также защищенности речевой информации. Цель данной работы состоит в сопоставлении основных объективных (инструментальных) методов оценки разборчивости речи (рис.1).



Рис.1

ОТЕЧЕСТВЕННЫЕ ВЕРСИИ ФОРМАНТНОГО МЕТОДА

На постсоветском пространстве используются различные версии формантного метода [5-10], среди которых наибольшее распространение получила версия Н.Б.Покровского [5].

Версія Н.Б.Покровського [5]

Оценивая формантную (артикуляционную) разборчивость речи, всю анализируемую область частот разбивают на K смежных частотных полос, с центральными частотами f_{0k} и граничными частотами f_{nk} и f_{ek} , в пределах каждой из которых спектры речи и шума, а также плотность вероятностей формант, можно считать практически неизменными [5].

Формантную (артикуляционную) разборчивость A вычисляют как сумму разборчивостей A_k формант в каждой из полос:

$$A = \sum_{k=1}^K A_k = \sum_{k=1}^K p_k \cdot P(E'_k), \quad (1)$$

где p_k - вероятность пребывания формант в k -той полосе частот:

$$p_k = F_1(f_{ek}) - F_1(f_{nk}), \quad (2)$$

$F_1(f)$ - закон распределения вероятностей формант по частоте;

$P(E'_k)$ - коэффициент восприятия речи (рис.2);

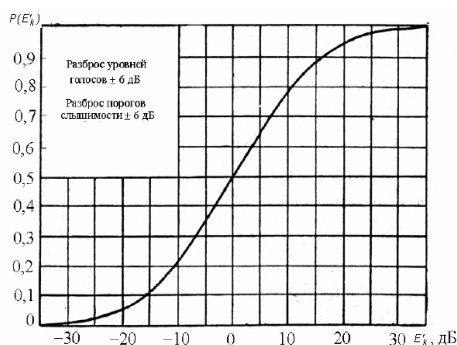


Рис.2.

E'_k - эффективный уровень ощущения формант в k -той полосе частот:

$$E'_k = B_{pk} - \Delta B_k - B_{uk}, \quad (3)$$

где B_{pk} и B_{uk} - уровни спектров мощности речевого сигнала и шума в k -той полосе частот; $\Delta B_k = B_{pk} - B'_{pk}$ - разница между спектром речи B_{pk} и спектром формант B'_{pk} .

Хотя функция $P(E'_k)$ зависит от частоты, для упрощения вычислений Покровский предложил пренебречь этой зависимостью. При этом, к сожалению, не указана возникающая погрешность вычислений.

Значительно сильнее, однако, настораживает сам вид зависимости $P(E'_k)$ (рис.2):

➤ динамический диапазон E'_k составляет 70 дБ, что значительно больше динамического диапазона уровней речевого сигнала 30-40 дБ;

➤ функция $P(E'_k)$ обладает центральной симметричностью относительно $E'_k = 0$.

В работах [17,18] показано, что наиболее вероятная причина выдвижения гипотезы о центральной симметричности функции – неверная организация Покровским экспериментальных исследований, в ходе которых использовалось линейное квантование речевого сигнала, тогда как необходимо было использовать нелинейное.

На рис.3 приведены корректные оценки закона распределения среднего (на интервале 0,125 с) квадрата речевого сигнала в октавных полосах с центральными частотами

500 Гц и 4000 Гц, приведенные в работе [17]. Анализ графиков Рис.3 свидетельствует, что закон распределения уровней речевого сигнала не является симметричным и заметно зависит от частоты.

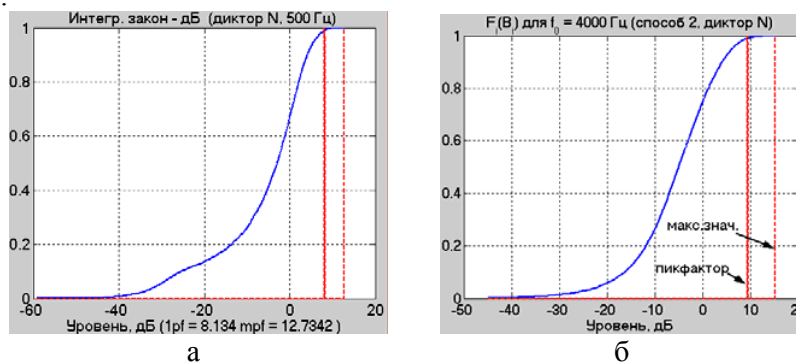


Рис.3

Распространяя на малые уровни сигналов [18] предложенную в [5] методику построения $P(E'_k)$, получаем кривые, содержащие «изломы» в точке $E'_k = 0$ (рис.5.а).

Появление таких изломов объясняется принятым в работе [5] определением спектра формант B'_p как спектра такого шума, на фоне которого происходит 50%-ная потеря артикуляционной разборчивости речи. Между тем, если бы спектр формант был определен так, чтобы точка $E'_k = 0$ находилась «на краю» графика $P(E'_k)$, тогда указанного излома удалось бы избежать.

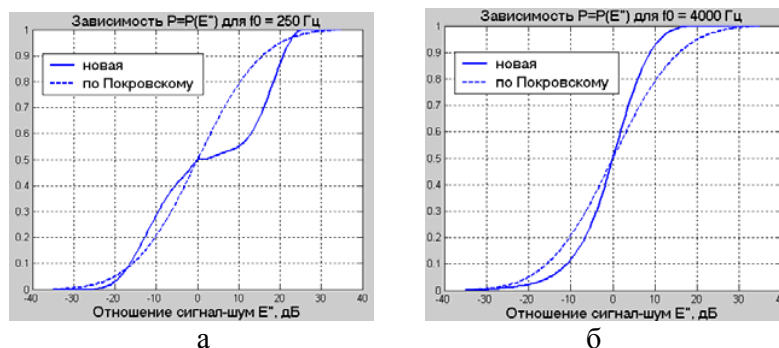


Рис.4

Версия Ю.С.Быкова [6]

Версия Быкова отличается от версии Покровского двумя моментами:

- спектр формант λ определяется как спектр такого шума, который, будучи просуммированным с речевым сигналом, приведет к полной потере разборчивости речи;
- учитывается зависимость коэффициента восприятия от частоты.

Важным следствием такого определения спектра формант является то, что нулевое значение отношения сигнал-шум попадает на левый край графика коэффициента восприятия (рис.5). Другим следствием является то, что теперь на всех частотах спектр формант λ превышает спектр речи B_p : $\Delta B_\lambda = \lambda - B_p > 0$, тогда как у Покровского спектр формант B'_p всегда меньше спектра речи: $\Delta B(f) = B'_p(f) - B_p(f) < 0$.

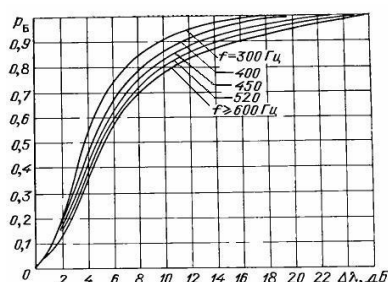


Рис. 5

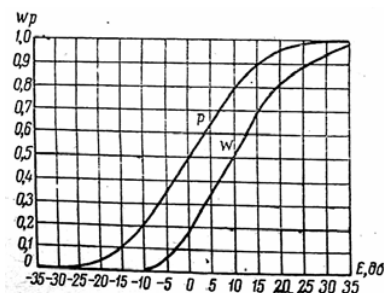


Рис. 6

Версия М.А.Сапожкова [7]

Принципиальная особенность версии Сапожкова – это фактическое отождествление спектра формант со спектром речи: «...мощность формант в диапазоне выше 300 Гц по отношению к мощности речи в этом же диапазоне составляет 98%...» [7-9]. График соответствующего коэффициента восприятия $w(E)$ приведен на рис.6. На этом же рисунке приведен коэффициент восприятия Покровского $P(E')$ (здесь сохранены авторские обозначения осей абсцисс и ординат [7]).

Сопоставление версий Покровского, Сапожкова и Быкова

Поскольку методологическая основа различных версий формантного метода одна и та же, следовало бы ожидать, что соответствующие этим версиям коэффициенты восприятия не должны существенно отличаться, если учесть:

- различие определений спектра формант;
- различие смысла, вкладываемого в понятия «уровень сигнала» и «уровень шума».

К сожалению, приходится констатировать, что даже сами, весьма авторитетные, авторы версий формантного метода совершают ошибки при сопоставлении различных коэффициентов восприятия.

Так, произведенное Сапожковым (рис.6) совмещение графиков, на наш взгляд, некорректно, поскольку функции $w(E)$ и $P(E')$ зависят от различных, хотя и тесно связанных аргументов E и E' . Действительно, Сапожков, отождествляющий спектр формант со спектром речи, оперирует *пиковыми* уровнями сигнала и шума, поэтому переменная E на графике рис.6 имеет следующий смысл:

$$E = E_{\text{пик}} = B_{p \text{ пик}} - B_{ш \text{ пик}}, \quad (5)$$

где $B_{p \text{ пик}}$ - спектр пиковых уровней речевого сигнала, $B_{ш \text{ пик}}$ - спектр пиковых уровней помехи.

Покровский не отождествляет спектр речи и спектр формант и, кроме того, оперирует с *эффективными* уровнями сигнала и шума, поэтому:

$$E' = B_p - \Delta B - B_{ш} = B_{p \text{ пик}} - \Pi_p - \Delta B - B_{ш \text{ пик}} + \Pi_{ш}, \quad (6)$$

где B_p - спектр мощности речевого сигнала; Π_p и $\Pi_{ш}$ - пикфакторы мощности речи и шума, соответственно.

Сопоставляя соотношения (5) и (6), находим связь между E и E' :

$$E' = E - \Pi_p + \Pi_{ш} - \Delta B. \quad (7)$$

Поскольку значения Π_p и Π_u принято считать равными 12 дБ и 7,5 дБ, соответственно [5,7,10], а значение ΔB при построении зависимости $P(E')$ выбрано равным 5 дБ [5], получаем:

$$E' \approx E - 10. \tag{8}$$

Следовательно, представленный на рис.6 график $P(E')$ на самом деле должен находиться правее на 10 дБ, а значит, совпадение графиков $w(E)$ и $P(E')$ оказывается весьма удовлетворительным в «средней части» отношений сигнал-шум.

На рис.7 представлены результаты еще одной попытки сопоставления различных коэффициентов восприятия, осуществленной Ю.Калинцевым [10]. Здесь кривая из «треугольников» соответствует версии Быкова, кривая из крестиков – версии Сапожкова, кривая из кружочков – версии Покровского. Заметим, что кривая версии Быкова («треугольники»), представлена не полностью – она обращается в нуль при $\Delta\lambda = 0$.

К сожалению, и здесь различие между переменными E и E' учтено неверно, составляя примерно 13 дБ (вместо 10 дБ). Неверно также отождествлены переменные E и $\Delta\lambda$. Учитывая, что Быков [6], как и Покровский [5], оперирует эффективными уровнями сигнала и шума, переменную $\Delta\lambda$ можно представить так:

$$\Delta\lambda = \lambda - B_u = B_p + \Delta B_\lambda - B_u = B_{ртик} - \Pi_p + \Delta B_\lambda - B_{штик} + \Pi_u, \tag{9}$$

где $\Delta B_\lambda = \lambda - B_p$ - разница между спектром формант λ и спектром речи B_p . Сравнивая (5) и (9), получаем:

$$\Delta\lambda = E - \Pi_p + \Delta B_\lambda + \Pi_u. \tag{10}$$

Поскольку $\Delta B_\lambda \approx 12 \div 18 \text{ dB}$, для среднего значения $\Delta B_\lambda \approx 15 \text{ dB}$ получаем:

$$\Delta\lambda \approx E + 10, \tag{11}$$

откуда следует, что представленный на рис.7 график Быкова («треугольники») нужно сместить влево на 10 дБ.

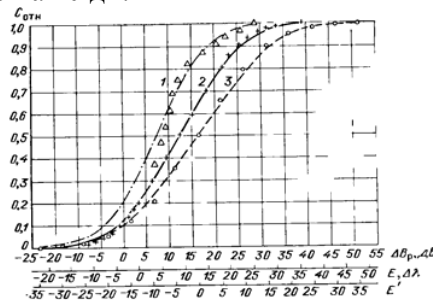


Рис.7

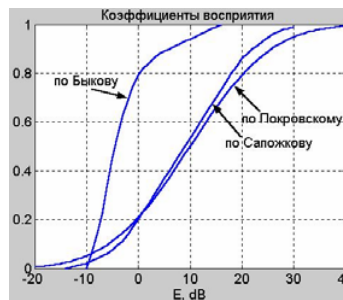


Рис.8

Правильное, по нашему мнению, совмещение коэффициентов восприятия на шкале значений переменной E должен выглядеть так, как показано на рис.8.

Версия Калинцева [10]

В работе [10] оценку относительного количества информации в речевом сообщении предложено производить с помощью соотношений, аналогичных соотношениям (1)-(2). Отличие в том, что теперь вместо закона распределения формант по частоте $F(f)$ оперируют законом $C_0(f)$ распределения информации по частоте (рис.9), а вместо $P(E')$

используют закон $C_{отн}(\Delta B_p)$ распределения относительного количества информации по уровню:

$$C_{отн}(\Delta B_p) = \lg \left[\frac{1+q}{(d+q)d} - d \right], \quad (12)$$

$$q = 10^{0,1 \cdot \Delta B_p}; \quad d = 10^{0,1 \cdot D};$$

где $\Delta B_p = B_p - B_{ш}$ - отношение сигнал-шум; $D = B_{p \max} - B_{p \min}$ - динамический диапазон речевого сигнала. Сплошная, пунктирная и штрих пунктирная линии на рис.7 – это кривые $C_{отн}(\Delta B_p)$, вычисленные для различных значений D (15, 25 и 33 дБ для линий 1, 2 и 3, соответственно). Как следует из рис.7, при $D=25$ дБ наблюдается хорошее совпадение $C_{отн}(\Delta B_p)$ с коэффициентом восприятия Сапожкова $w(E)$. К сожалению, для двух других кривых $C_{отн}(\Delta B_p)$ аналогичные совпадения с коэффициентами восприятия Покровского и Быкова не имеют места, поскольку, как уже было отмечено, графики $P(E'_k)$ и $P_B(\Delta \lambda)$ на рис.7 находятся “не на своих местах”.

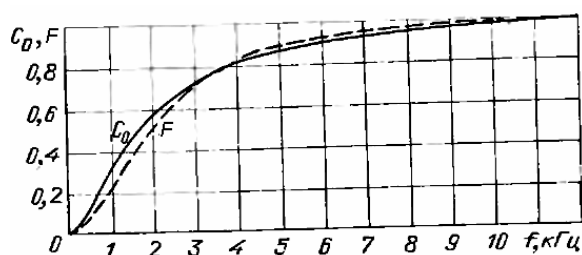


Рис.9

ЗАРУБЕЖНЫЕ ВЕРСИИ ФОРМАНТНОГО МЕТОДА

Безусловный интерес представляют зарубежные версии формантного подхода – тем более, что отечественные версии в значительной степени базируются на зарубежных работах [1-4]. Рассмотрим две такие версии: индекс артикуляции (Articulation Index - AI) и индекс разборчивости речи (Speech Intelligibility Index - SII).

Индекс артикуляции (AI) [15]

Версия формантного метода, известная как индекс артикуляции (AI), сформировалась в работах [2,3,19-25]. В рамках данной версии считается, что разборчивость речи пропорциональна средней разности между пиковым уровнем речи и эффективным уровнем маскирующего шума. Формально это означает, что для оценки индекса артикуляции используется аналог соотношения (1), с тем лишь отличием, что в качестве коэффициента восприятия используется функция

$$P_{AI}(\Delta L) = \begin{cases} 0, & \Delta L \leq 0 \text{ дБ}; \\ \frac{\Delta L}{30}, & 0 < \Delta L \leq 30 \text{ дБ}; \\ 1, & \Delta L > 30 \text{ дБ}; \end{cases} \quad (12)$$

где $\Delta L = B_{p \max} - B_{ш}$ (рис.10).

Очевидно хорошее совпадение коэффициента восприятия (12) и коэффициента восприятия Сапожкова, линейаризованного соотношением (4). Действительно,

$$\Delta L = B_{p \text{ нук}} - B_{ш} = B_{p \text{ нук}} - B_{ш \text{ нук}} + \Pi_{ш} = E + \Pi_{ш}, \quad (13)$$

поэтому из (4) с учетом (13) получаем:

$$w(E) = w(\Delta L - \Pi_{ш}) \approx w(\Delta L - 7,5), \quad (14)$$

т.е., будучи представленной на рис.10, линейная часть коэффициента восприятия Сапожкова окажется правее линейной части коэффициента (12) всего на 1,5 дБ.

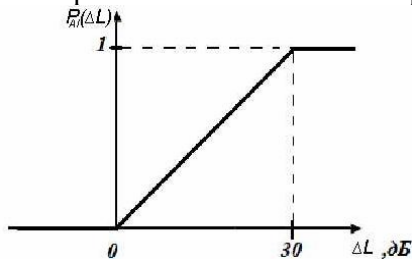


Рис. 10

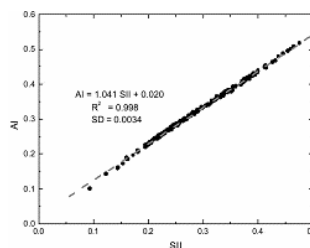


Рис. 11

Индекс разборчивости речи (SII)

Начиная с 1984 года, вычислительная схема индекса артикуляции AI неоднократно подвергалась усовершенствованиям в работах [26-29], что привело к появлению версии, стандартизованной в ANSI S3.5-1997 [30] и названной индексом разборчивости речи (SII).

Измерение разборчивости по методу SII осуществляют аналогично методу AI, поэтому неудивительно, что соответствующие значения индексов разборчивости мало отличаются численно, находясь в линейной зависимости (рис.11).

МОДУЛЯЦИОННЫЕ МЕТОДЫ

Существует несколько модуляционных методов: STIr (модифицированный STI), RASTI (быстрый STI), STITEL (STI для телекоммуникационных систем), STIPA (STI для систем звукоусиления) [3-10].

Измерение и расчет STI [4]

При использовании метода STI возможность одновременного учета шумовой и реверберационной помех обеспечивается специальным выбором тестового сигнала в виде шума со спектром, идентичным спектру долговременной речи (рис.12). Этот шум в каждой октавной полосе частот модулируется периодическим сигналом таким образом, чтобы огибающая мгновенной мощности сигнала имела форму синусоиды.

В рассмотренном на рис.12 примере (случай шумовой помехи) индекс модуляции равен:

$$m = \frac{\bar{I}_{test}}{\bar{I}_{test} + \bar{I}_{noise}} = \frac{\bar{I}_{test}/\bar{I}_{noise}}{\bar{I}_{test}/\bar{I}_{noise} + 1} = \frac{SNR}{1 + SNR}, \quad (15)$$

где $SNR = \bar{I}_{test}/\bar{I}_{noise}$ - отношение сигнал-шум «в разгах по мощности».

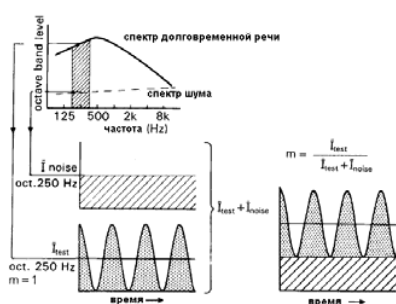


Рис.12

После несложных преобразований можно получить обратную зависимость:

$$SNR = \frac{m}{1-m} \quad (16)$$

или:

$$SNR_{dB} = 10 \log \frac{m}{1-m}. \quad (17)$$

Соотношение (17) используют для оценки разборчивости речи: вычисляют по измеренному коэффициенту m отношение сигнал-шум в точке приема, а затем, руководствуясь рис.13, пересчитывают отношение сигнал-шум в индекс передачи речи STI.

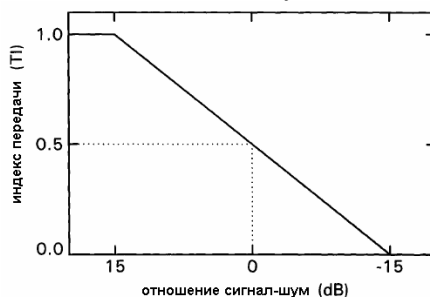


Рис.13

Таблица 1.

Значение STI, RASTI, STITEL, STIPA	0...0,3	0,3...0,4	0,45...0,6	0,6...0,75	0,75...1
Оценка разборчивости	очень плохая	плохая	посредств.	хорошая	отличная

Решение о качестве тестируемого канала передачи речи принимают, ориентируясь на шкалу табл.1.

Согласно «полной» версии метода STI, именуемой также STI_g или STI-14, определяют 98 значений коэффициента $m_{k,F}$, получаемых для 14 значений частот модуляции F (0,63; 0,8; 1; 1,25; 1,6; 2; 2,5; 3,15; 4; 5; 6,3; 8; 10; 12,5 Гц) в каждой из семи октавных полосах с центральными частотами от 125 Гц до 8 кГц (k - номер полосы частот). Вычисленные 98 значений индекса STI усредняют по специальной методике.

Теоретически влияние реверберационной и шумовой помех на коэффициент m объясняют с помощью соотношения:

$$m(F) = \frac{1}{\sqrt{1 + \left(\frac{2\pi FT}{13,8}\right)^2}} \cdot \frac{1}{1 + 10^{-0,1 \cdot (S/N)_{\text{дБ}}}} = m_R(F) \cdot m_N(F), \quad (18)$$

где первый сомножитель учитывает влияние реверберации, а второй – шума.

Методы RASTI, STIPA, STITEL [12,16]

Метод RASTI представляет собой сокращенную версию метода STI. Как и метод STI, метод RASTI позволяет учесть реверберационную помеху. Испытательный сигнал упрощен: количество 1/1 октавных полос сокращено до двух, с центральными частотами 500 Гц и 2 кГц. При таком подходе полоса пропускания ограничена и, следовательно, фоновый шум с нерегулярным спектром и нелинейные искажения не учитываются корректно. Тем не менее, метод RASTI может использоваться для прикидочной диагностики помещений.

Метод STIPA – модификация метода STI для систем звукоусиления (public address systems), позволяющая учитывать не только реверберацию, но и нелинейные искажения звука в помещениях. Испытательный сигнал упрощен в том смысле, что в каждой из семи октавных полос используют только две частоты модуляции. В остальных отношениях метод STIPA идентичен методу STIr.

В методе STITEL применяется только одна частота модуляции в каждой из семи 1/1 октавных полос. Несущий шум для каждой 1/1 октавной полосы имеет ширину спектра 1/2 октавы, во избежание влияния на смежные полосы. Метод STITEL не позволяет учитывать реверберационную помеху и нелинейные искажения.

ЭМПИРИЧЕСКИЕ МЕТОДЫ

Среди эмпирических методов наиболее популярен метод %Alcons – метод измерения величины потери артикуляции согласных выраженной в процентах [8]. Метод %Alcons широко используется, особенно в США, для приближенной оценки разборчивости речи и отображает потерю вокализованных согласных, вызванную реверберацией и поглощением звука в помещении.

Для расчёта значения %Alcons используется соотношение:

$$\%Alcons = [200 \cdot D_2^2 \cdot RT_{60}^2 \cdot (n+1)] : (V \cdot Q \cdot m)$$

где, V – объём помещения, куб. м.;

D_2 – дистанция источник речи – слушатель, м;

Q – фактор направленности;

(n+1) – общее число дикторов (громкоговорителей), создающих реверберационное поле;

$m = (1-\alpha) : (1-\alpha_c)$ – модификатор критического расстояния;

α – коэффициент поглощения;

α_c – коэффициент поглощения поверхности, перекрываемой всеми дикторами (громкоговорителями);

RT_{60} – время реверберации, сек.

Еще одним эмпирическим методом является применение коэффициента C50, позволяющего оценить четкость (ясность) звучания музыки. Этот коэффициент вычисляется как отношение ранней энергии (50 мс) отраженных звуков к поздней. При этом, как и в случае с коэффициентом %Alcons, не учитывается влияние шумовой помехи.

СЕКРЕТНОСТЬ И РАЗБОРЧИВОСТЬ РЕЧИ

Используемый в зарубежной литературе «индекс секретности» (privacy index - PI) связан с индексом артикуляции AI простым соотношением [31]:

$$PI = (1 - AI) * 100\%.$$

Очевидно, аналогичную связь можно установить и для других показателей разборчивости речи, используемых при формантном подходе.

В отличие от метода AI, учитывающего только шумовую помеху, метод RASTI позволяет учесть реверберационную помеху и эхо. Однако при малых значениях индекса RASTI результирующая модуляция будет очень малой, соизмеримой с неравномерностью уровня фонового шума. Для разрешения этой проблемы процедуру оценки $m(F)$ «расщепляют» на две ветки:

- 1) задавая большое отношение сигнал-шум, экспериментально определяют первый сомножитель соотношения (18), учитывая таким образом влияние реверберации;
- 2) задавая «нормальное» отношение сигнал-шум (соответствующее нормальному уровню речи) аналитически определяют второй сомножитель соотношения (18).

ЗАКЛЮЧЕНИЕ

Если ограничиться рассмотрением объективных (инструментальных) методов, можно выделить три подхода к оценке разборчивости речи: формантный, модуляционный и эмпирический.

Различие рассмотренных версий формантного метода главным образом сводится к различию определений понятия «формантный спектр», а также различию смысла понятий «уровень сигнала» и «уровень шума». Если корректно учесть эти различия, оказывается, что почти все версии формантного метода (за исключением версии Быкова) обеспечивают получение близких результатов для области «средних» и «больших» отношений сигнал-шум. В этой связи понятие формантного спектра представляется малопродуктивным.

При малых отношениях сигнал-шум результаты оценивания разборчивости могут существенно отличаться для различных версий формантного метода. Объяснить это можно принципиальной трудностью организации исследований в условиях малых отношений сигнал-шум.

Главным достоинством модуляционного подхода к оценке разборчивости речи является сравнительная легкость учета влияния таких помех как реверберация и эхо, нелинейные искажения. Однако надежность оценки разборчивости речи при малых отношениях сигнал-шум и в этом случае остается сомнительной. А такой эмпирический метод, как оценка потери артикуляции согласных, обладает невысокой точностью даже при больших отношениях сигнал-шум.

ЛИТЕРАТУРА

1. Collard J.A. Theoretical Study of the Articulation and Intelligibility of a Telephone Circuit // Electrical Communication. – 1929. - Vol.7. – P. 168.
2. French N., Steinberg J. Factors Governing the Intelligibility of Speech Sounds // J. Acoust. Soc. Am. – 1947. - Vol.19, #1.

3. *Fletcher H., Galt F.* Perception of Speech and its Relation to Telephony // J. Acoust. Soc. Am. – 1950. - Vol.22, #2.
4. *Беранек Л.* Расчет речевых систем связи // Proceedings of the IRE. – 1947. – September. – P. 880-890.
5. *Покровский Н.Б.* Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. – 390 с.
6. *Быков Ю.С.* Теория разборчивости речи и повышение эффективности радиотелефонной связи. – М.-Л.: Госэнергоиздат, 1959. – 350 с.
7. *Сапожков М.А.* Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963. – 472 с.
8. *Сапожков М.А.* Звукофикация помещений. Проектирование и расчет. – М.: Связь, 1979. – 144 с.
9. *Сапожков М.А., Михайлов В.Г.* Вокодерная связь. – М.: Радио и связь, 1983. – 246с.
10. *Калинцев Ю.К.* Разборчивость речи в цифровых вокодерах. – М.: Радио и связь, 1991. с. 81 – 111.
11. *Алдошина И., Приттс Р.* Музыкальная акустика. – С-Пб.: Композитор • Санкт-Петербург, 2006. - 717 с.
12. *Steeneken H.J.M., Houtgast T.* RASTI: A Tool for Evaluating Auditoria // Bruel & Kjaer Technical Review No.3 – 1985. – P.13-39.
13. *Железняк В.К., Макаров Ю.К., Хорев А.А.* Некоторые методические подходы к оценке эффективности защиты речевой информации // Специальная техника. – М.: 2000, № 4.
14. *Макаров Ю.К., Хорев А.А.* К оценке эффективности защиты акустической (речевой) информации (<http://st.ess.ru/publications/articles/tspi/tspi.htm>)
15. *Рашевский Я.И., Каргашин В.Л.* Обзор зарубежных методов определения разборчивости речи. - Специальная техника, №№ 3-6 за 2002 год, № 1 за 2003 год (http://www.detektor.ru/publication/d2003_1.htm)
16. *Steeneken H.J.M., Houtgast T.* RASTI: The Modulation Transfer Function in Room Acoustics // Bruel & Kjaer Technical Review No.3 – 1985. – P.1-12.
17. *Гавриленко О.В., Дидковский В.С., Продеус А.Н.* Расчет и измерение разборчивости речи при малых отношениях сигнал-шум. Часть 1. Корректное измерение функции распределения речевого сигнала // Электроника и связь, Тематический выпуск «Проблемы электроники», ч.1. - 2007. – С. 137-141.
18. *Гавриленко О.В., Дидковский В.С., Продеус А.Н.* Расчет и измерение разборчивости речи при малых отношениях сигнал-шум. Часть 2. Коррекция коэффициентов восприятия // Электроника и связь, Тематический выпуск «Проблемы электроники», ч.1. - 2007. – С. 142-147.
19. *Kryter K.D.* Methods for the calculation and use of the articulation index // J. Acoust. Soc. Am. - 1962. – Vol. 34. – P. 1689–1697.

20. *Kryter K.D.* Validation of the articulation index // *J. Acoust. Soc. Am.* - 1962. – Vol. 34. – P. 1698–1702.
21. *Rankovic C.M.* Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural hearing loss // *J. Acoust. Soc. Am.* – 1998. - Vol. 103. P. 1043–1057.
22. *Turner C.W., Henry B.A.* Benefits of amplification for speech recognition in background noise // *J. Acoust. Soc. Am.* - 2002. – Vol. 112. – P. 1675–1680.
23. *Musch H., Buus S.* Using statistical decision theory to predict speech intelligibility. II. Measurement and prediction of consonantdiscrimination performance // *J. Acoust. Soc. Am.* - 2001. – Vol. 109. – P. 2910–2920.
24. *Brungart D.S.* Informational and energetic masking effects in the perception of two simultaneous talkers // *J. Acoust. Soc. Am.* - 2001. – Vol.109. – P. 1101–1109.
25. *Dubno J.R., Horwitz A.R., Ahlstrom J.B.* Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing // *J. Acoust. Soc. Am.* - 2003. – Vol. 113. – P. 2084–2094.
26. *Pavlovic C.V.* Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment // *J. Acoust. Soc. Am.* - 1984. – Vol.75. – P. 1253–1258.
27. *Studebaker G.A., Taylor R., Sherbecoe R.L.* The effect of noise spectrum on speech recognition performance-intensity functions // *J. Speech Hear. Res.* - 1994. – Vol.37. P. 439–448.
28. *Dirks D.D., Bell T.S., Rossman R.N., Kincaid G.E.* Articulation index predictions of contextually dependent words // *J. Acoust. Soc. Am.* - 1986. – Vol. 80. – P. 82–92.
29. *Kamm C.A., Dirks D.D., Bell T.S.* Speech recognition and the Articulation Index for normal and hearing-impaired listeners // *J. Acoust. Soc. Am.* - 1985. – Vol. 77. – P. 281–288.
30. *ANSI S3.5-1997*, American National Standard Methods for Calculation of the Speech Intelligibility Index - American National Standards Institute, New York. - 1997.
31. *Privacy Index* (<http://www.diracdelta.co.uk/science/source/p/r/privacyindex/source.html>)