

# Слоговая разборчивость в системе передачи информации речевыми сигналами в зависимости от интенсивности акустической помехи

Ю.А. Кропотов<sup>1</sup>, А.А. Белов<sup>1</sup>, А.А. Колпаков<sup>1</sup>, А.Ю. Проскуряков<sup>1</sup>

<sup>1</sup>Муромский институт (филиал) Владимирского государственного университета имени А.Г. и Н.Г. Столетовых», Орловская 23, Муром, Россия, 602264

**Аннотация.** В работе проводится исследование влияния отношения сигнал/шум на слоговую разборчивость при интенсивном воздействии внешних акустических помех при обмене речевыми сообщениями в телекоммуникационных системах громкоговорящей связи. В статье рассмотрено влияние на слоговую разборчивость отношения сигнал/внешний акустический шум, исследуется влияние интегрального индекса артикуляции, зависимость коэффициента восприятия формант от относительного уровня интенсивности формант, зависимость формантного параметра от среднегеометрической частоты  $i$ -ой полосы спектра речевого сигнала. В соответствии с результатами исследований интегрального индекса артикуляции в зависимости от отношения сигнал/шум, получена функция зависимости слоговой разборчивости от вышеуказанного отношения, по которой становится возможным определение максимального значения выходного отношения сигнал/шум в системе телекоммуникаций аудиообмена для получения заданной слоговой разборчивости. При этом экспериментальным путем определено значение отношения сигнал/шум в системе телекоммуникаций аудиообмена для получения слоговой разборчивости не менее 93%, для обеспечения полного восприятия передаваемой речевой информации.

## 1. Введение

Как известно, основным критерием эффективности системы телекоммуникационного обмена речевой информацией является слоговая разборчивость  $S\%$  или величина оценки речевого сигнала по шкале MOS (Mean Opinion Score – средняя субъективная оценка) [1].

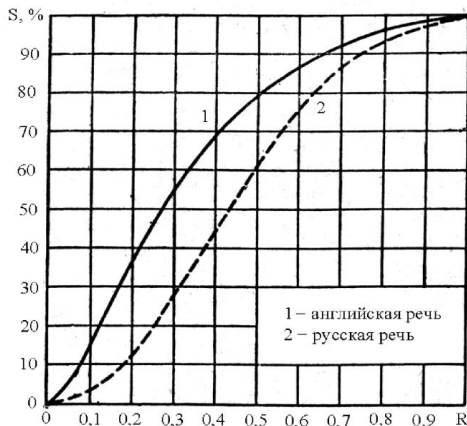
Системы телекоммуникаций аудио обмена, в частности системы громкоговорящей связи, считаются эффективными, если переданная речевая информация воспринимается объектом полностью и без затруднений, слоговая разборчивость в этом случае составляет не менее 93 % [1,2,4] или оценка MOS должна составлять не менее 3,9 баллов по пятибалльной шкале [5, 6].

## 2. Постановка задачи

Зависимость слоговой разборчивости в системе телекоммуникационного обмена речевой информацией от влияния различных факторов исследована в ряде работ [1,3]. Однако информация в известных источниках [9, 10] о влиянии отношения сигнал/шум на слоговую разборчивость на стороне приема речевых сообщений для случая оперативно-командных систем телекоммуникаций недостаточна, поэтому в данной статье рассмотрена задача

определения влияния отношения сигнал/шум на слоговую разборчивость в телекоммуникационных системах аудио обмена.

Известные результаты исследований оценивания слоговой разборчивости инструментально-расчетным методом приведены на рисунке 1 [1, 3].



**Рисунок 1.** Зависимость слоговой разборчивости S от интегрального индекса артикуляции R.

Из рисунка 1 видно, что слоговая разборчивость здесь представлена зависимостью только от значения интегрального индекса артикуляции.

### 3. Инструментально-расчетный метод оценивания интегрального индекса артикуляции и слоговой разборчивости

Значение интегрального индекса артикуляции R в зависимости от значения спектрального индекса артикуляции  $R_i$  определяется выражением

$$R = \sum_{i=1}^N R_i . \tag{1}$$

Спектральный индекс артикуляции вычисляется по выражению

$$R_i = p_i \cdot k_i , \tag{2}$$

где  $p_i$  – коэффициент восприятия формант,  $k_i$  – весовой коэффициент вероятности наличия формант речи в  $i$ -ой полосе.

Коэффициент восприятия формант  $p_i$  вычисляется с помощью выражения [3]

$$p_i = \begin{cases} \frac{0,78 + 5,46 \cdot \exp[-4,3 \cdot 10^{-3} \cdot (27,3 - |Q_i|)^2]}{1 + 10^{0,1 \cdot |Q_i|}} , & \text{если } Q_i \leq 0; \\ 1 - \frac{0,78 + 5,46 \cdot \exp[-4,3 \cdot 10^{-3} \cdot (27,3 - |Q_i|)^2]}{1 + 10^{0,1 \cdot |Q_i|}} , & \text{если } Q_i > 0, \end{cases} \tag{3}$$

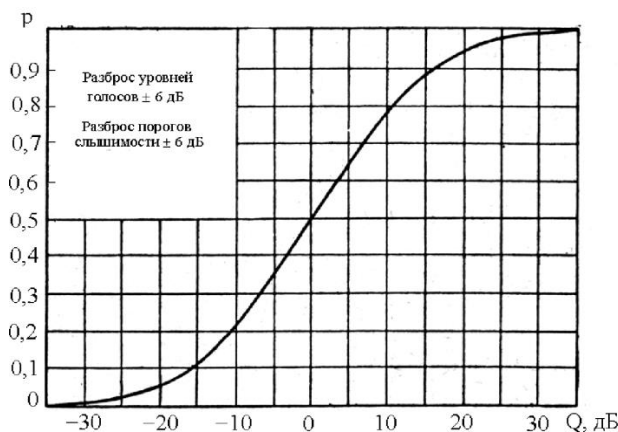
где  $Q_i = q_i - \Delta A_i$  – относительный уровень интенсивности формант.

Или значение коэффициента восприятия формант  $p_i$  можно определить по графику на рисунке 2.

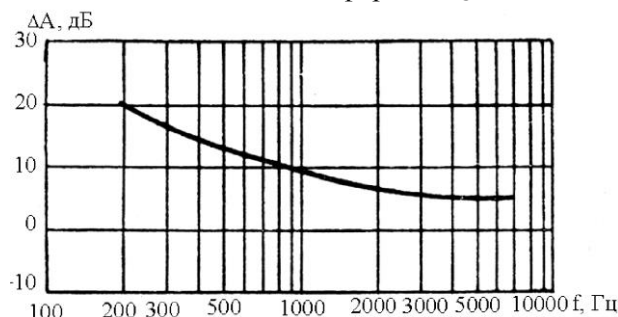
Формантный параметр  $\Delta A_i$  определяется по графику на рисунке 3 или с помощью выражения

$$\Delta A(f) = \begin{cases} 200 / f^{0,43} - 0,37, & \text{если } f \leq 1000 \text{ Гц,} \\ 1,37 + 1000 / f^{0,69}, & \text{если } f > 1000 \text{ Гц,} \end{cases} \tag{4}$$

где  $f_{cp,i} = \sqrt{f_{\sigma_i} \cdot f_{\eta_i}}$  – среднегеометрическая частота,  $f_{\eta_i}$  – нижняя частота  $i$ -ой полосы частот спектра речевого сигнала,  $f_{\sigma_i}$  – верхняя частота  $i$ -ой полосы частот спектра.



**Рисунок 2.** Зависимость коэффициента восприятия формант  $p_i$  от относительного уровня интенсивности формант  $Q_i$ .



**Рисунок 3.** Зависимость формантного параметра  $\Delta A_i$  от среднегеометрической частоты  $i$ -ой полосы спектра речевого сигнала.

Для каждой  $i$ -ой ( $i=1, 2, \dots, N$ ) частотной полосы на среднегеометрической частоте  $f_{cp,i} = \sqrt{f_{\sigma_i} \cdot f_{\eta_i}}$ , определяется формантный параметр  $\Delta A_i$ , характеризующий энергетическую избыточность дискретных составляющих речевого сигнала.

**4. Результаты экспериментов**

Примем число октавных полос  $N=5$ . Значения принятых пределов по частоте октавных полос, значения вычисленных  $f_{cp,i}$  и значений формантных параметров  $\Delta A_i$  сведены в таблицу 1.

**Таблица 1.** Значения принятых пределов по частоте октавных полос, значения вычисленных  $f_{cp,i}$  и значений формантных параметров  $\Delta A_i$ .

$i$	Частота граничных полос, Гц	$f_{cp,i}$ Гц	Ширина полосы $\Delta f_i$	$\Delta A_i(f_{cp,i})$
1	180-355	250	175	18
2	355-710	500	355	13
3	710-1400	1000	690	9,5
4	1400-2800	2000	1400	7,5
5	2800-5600	4000	2800	5,5

С помощью выражения  $Q_i = q_i - \Delta A_i$ , определили значения уровней интенсивности формант  $Q_i$  в зависимости от отношения сигнал/шум  $q_i$ . Вычисленные значения  $Q_i$  сведены в таблицу 2. С помощью выражения (3) или по графику на рисунке 2 определяется коэффициент восприятия формант  $p_i$  в зависимости от  $Q_i$  для  $i$ -тых полос, при различных значениях отношения сигнал/шум, дБ. Вычисленные значения  $p_i$  для различных  $q_i$  сведены в таблицу 3.

**Таблица 2.** Вычисленные значения  $Q_i$ .

$Q_i = q_i - \Delta A_i$ $q_i, \text{дБ}$	$Q1$	$Q2$	$Q3$	$Q4$	$Q5$
$q_i = 0 \text{ дБ}$	-18	-13	-9.5	-7.5	-5,5
$q_i = 3 \text{ дБ}$	-15	-10	-6.5	-4,5	-2,5
$q_i = 6 \text{ дБ}$	-12	-7	-3.5	-1,5	+0,5
$q_i = 10 \text{ дБ}$	-8	-3	+0.5	+2,5	+4,5
$q_i = 20 \text{ дБ}$	+2	+7	+10,5	+12,5	+14,5
$q_i = 30 \text{ дБ}$	+12	+17	+20,5	+22,5	+24,5

**Таблица 3.** Вычисленные значения  $p_i$ .

$q_i, \text{дБ}$	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$
$q_i = 0$	0,07	0,18	2,2	0,29	0,34
$q_i = 3$	0,11	0,21	0,31	0,38	0,41
$q_i = 6$	0,2	0,3	0,40	0,48	0,51
$q_i = 10$	0,24	0,41	0,51	0,53	0,62
$q_i = 20$	0,47	0,6	0,79	0,81	0,88
$q_i = 30$	0,81	0,9	0,94	0,96	0,98

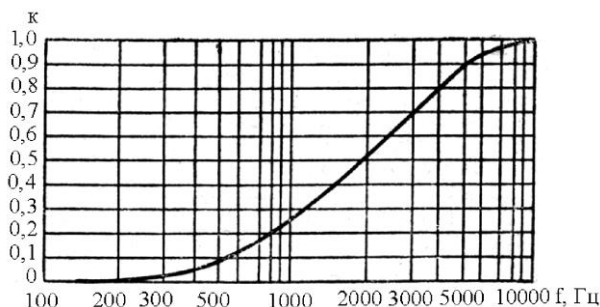
Для определения спектрального индекса артикуляции  $R_i$ , определяется весовой коэффициент вероятности наличия формант речи в  $i$ -ой полосе  $k_i$  по формуле

$$k_i = k(f_{ei}) - k(f_{ni}),$$

где  $k(f_{ei})$  и  $k(f_{ni})$  – значения весовых коэффициентов для верхней  $f_{ei}$  и нижней  $f_{ni}$  граничных частот  $i$ -ой полосы спектра речевого сигнала. Значения  $k(f_{ei})$  и  $k(f_{ni})$  определяются по формуле в виде [3]:

$$k(f) = \begin{cases} 2,57 \cdot 10^{-8} \cdot f^{2,4}, & \text{если } 100 < f \leq 400 \text{ Гц;} \\ 1 - 1,074 \cdot \exp(-10^{-4} \cdot f^{1,18}), & \text{если } 400 < f \leq 10000 \text{ Гц;} \end{cases} \quad (5)$$

или по графику на рисунке 4.



**Рисунок 4.** Значение весового коэффициента  $k(f)$  формантного распределения по частоте.

Результаты вычислений весовых коэффициентов вероятности формантной речи в  $i$ -ой полосе представлены в таблице 4.

**Таблица 4.** Результаты вычислений весовых коэффициентов вероятности формантной речи в  $i$ -ой полосе.

$k_i$	$k(f_e) - k(f_n)$	$k(f_{ei}) - k(f_{ni})$	$k_i$
$k_1$	$k(355) - k(180)$	0,04-0,01	0,03
$k_2$	$k(710) - k(355)$	0,18-0,04	0,14
$k_3$	$k(1400) - k(710)$	0,36-0,18	0,18
$k_4$	$k(2800) - k(1400)$	0,63-0,36	0,37
$k_5$	$k(5600) - k(2800)$	0,92-0,63	0,29

Вычисления спектрального индекса артикуляции  $R_i$  проводится по формуле (2). Вычисления  $R_i$ , при различных значения отношения сигнал/шум сведены в таблицу 5.

**Таблица 5.** Спектральный индекс артикуляции  $R_i$ , при различных значения отношения сигнал/шум.

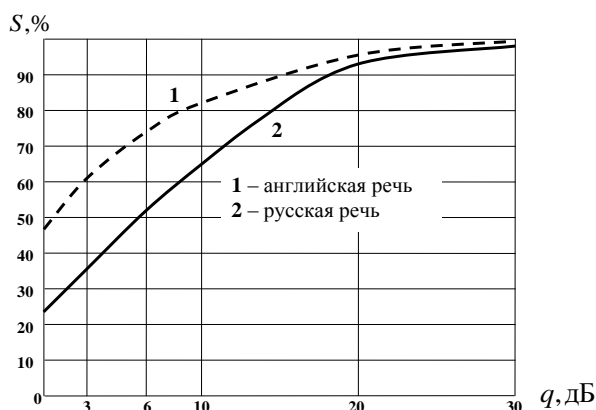
$R_i = p_i \cdot k_i$	$R_1 = p_1 \cdot k_1$	$R_2 = p_2 \cdot k_2$	$R_3 = p_3 \cdot k_3$	$R_4 = p_4 \cdot k_4$	$R_5 = p_5 \cdot k_5$
$q_i = 0$ дБ	$0,07 \cdot 0,03 = 0,0021$	$0,18 \cdot 0,014 = 0,0252$	$0,22 \cdot 0,18 = 0,0396$	$0,29 \cdot 0,37 = 0,1073$	$0,34 \cdot 0,29 = 0,0986$
$q_i = 3$ дБ	$0,11 \cdot 0,03 = 0,0033$	$0,021 \cdot 0,14 = 0,0294$	$0,31 \cdot 0,18 = 0,0558$	$0,38 \cdot 0,37 = 0,1406$	$0,41 \cdot 0,29 = 0,1189$
$q_i = 6$ дБ	$0,2 \cdot 0,03 = 0,006$	$0,3 \cdot 0,14 = 0,042$	$0,4 \cdot 0,18 = 0,072$	$0,48 \cdot 0,37 = 0,1776$	$0,51 \cdot 0,29 = 0,1479$
$q_i = 10$ дБ	$0,24 \cdot 0,03 = 0,0072$	$0,41 \cdot 0,14 = 0,0574$	$0,51 \cdot 0,18 = 0,0918$	$0,53 \cdot 0,37 = 0,1961$	$0,62 \cdot 0,29 = 0,1798$
$q_i = 20$ дБ	$0,47 \cdot 0,03 = 0,0141$	$0,6 \cdot 0,14 = 0,084$	$0,79 \cdot 0,18 = 0,0918$	$0,81 \cdot 0,37 = 0,2997$	$0,88 \cdot 0,29 = 0,2552$
$q_i = 30$ дБ	$0,81 \cdot 0,03 = 0,0243$	$0,9 \cdot 0,14 = 0,126$	$0,94 \cdot 0,18 = 0,1692$	$0,96 \cdot 0,37 = 0,3552$	$0,98 \cdot 0,29 = 0,2842$

По результатам вычислений спектрального индекса артикуляции  $R_i$ , сведенным в таблицу 5, стало возможным вычисление интегрального индекса артикуляции в зависимости от отношения сигнал/шум. Результаты вычисления интегрального индекса артикуляции позволили найти значения слоговой разборчивости в зависимости от отношения сигнал/шум, которые сведены в таблицу 6.

**Таблица 6.** Значения слоговой разборчивости в зависимости от отношения сигнал/шум.

$q_i$ сигнал/шум	0 дБ	3 дБ	6 дБ	10 дБ	20 дБ	30 дБ
$R$	0,273	0,348	0,4455	0,5323	0,7952	0,9589
$S_{англ.}$	48 %	61 %	73 %	82%	96%	99%
$S_{рус.}$	25%	35 %	53 %	65 %	93%	98,5%

График функции слоговой разборчивости  $S$  от отношения сигнал/шум представлен на рисунке 5.



**Рисунок 5.** График зависимости значения слоговой разборчивости от отношения сигнал/шум.

### 5. Выводы

Как видно из графиков на рисунке 5, слоговая разборчивость в системе телекоммуникаций обмена речевыми сообщениями обеспечивается  $S \geq 93\%$  при отношении сигнал/шум  $q \geq 20$  дБ [7, 8]. Таким образом, получена важная для практики систем телекоммуникаций оповещения и аудиообмена зависимость слоговой разборчивости от отношения сигнал/шум, которая показывает, что для эффективной передачи речевой информации оперативно-командной системой телекоммуникаций, для получения соответственно слоговой разборчивости  $S \geq 93\%$ , в системе при передаче речевых сообщений, на стороне приема сообщений необходимо обеспечить отношение сигнал/шум  $q \geq 20$  дБ.

## 6. Литература

- [1] Сапожков, М.А. Речевой сигнал в кибернетике и связи.– М.: Связьиздат, 1962. – 452 с.
- [2] ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости.
- [3] Железняк, В.К. Некоторые методические подходы к оценке эффективности защиты речевой информации / В.К. Железняк, Ю.К. Макаров, А.А. Хорев // Специальная техника. – 2000. – № 4. – С. 39-45.
- [4] Cohen, I. Speech processing in modern communication / I. Cohen, J. Benesty, S. Gannot. – Berlin Heidelberg: Springer, 2010. – 342 p.
- [5] Hansler, E. Topics in acoustic echo and noise control: Selected methods for the cancelation of acoustic echoes, the reduction of background noise, and speech processing/ E. Hansler, G. Schmidt. – Berlin, Heidelberg, Springer, 2006. – 642 p.
- [6] Kahrs, M. Applications of digital signal processing to audio and acoustics / M. Kahrs, K. Brandenburg. – New York, Kluwer Academic Publisher, 2002. – 572 p.
- [7] Кропотов, Ю.А. Алгоритм подавления акустических шумов и сосредоточенных помех с формантным распределением полос режекции / ЮА. Кропотов, А.А. Быков // Вопросы радиоэлектроники. – 2010. – Т.1, № 1. – С. 60-65.
- [8] Кропотов, Ю.А. Алгоритм вычисления сигнала управления каналом режекции многоканальной системы передачи акустических сигналов // Вопросы радиоэлектроники. – 2010. – Т.1, № 1. – С. 57-60.
- [9] Рябенький, В.С. Математическая модель устройств подавления внешнего шума в подобласти пространства // Математическое моделирование. – 2012. – Т. 24, № 8. – С. 3-31.
- [10] McAulay, R. Speech enhancement using a soft-decision noise suppression filter / R. McAulay, M. Malpass // IEEE Trans, on Acoustics, Speech, and Signal Processing. – 1980. – Vol. 28(2). – P. 137-145.

# The syllable intelligibility in the system of information transmission by speech signals depending on the intensity of acoustic noise

Yu.A. Kropotov<sup>1</sup>, A.A. Belov<sup>1</sup>, A.A. Kolpakov<sup>1</sup>, A.Yu. Proskuryakov<sup>1</sup>

<sup>1</sup>Murom Institute (branch) Vladimir State University named after Alexander and Nicholay Stoletovs, Orlovskaya street 23, Murom, Russia, 602264

**Abstract.** The paper investigates the effect of the signal-to-noise ratio on syllable intelligibility under the intense influence of external acoustic interference when exchanging voice messages in telecommunication systems of public address systems. The article discusses the effect on the syllable intelligibility of the signal / external acoustic noise ratio, examines the effect of the integral articulation index, the dependence of the perception coefficient of formants on the relative level of formant intensity, the dependence of the formant parameter on the geometric mean frequency of the  $i$ -th spectrum of the speech signal. In accordance with the results of studies of the integral articulation index depending on the signal-to-noise ratio, a function of syllable intelligibility depending on the signal-to-noise ratio was obtained, using which it is possible to determine the maximum value of the output signal-to-noise ratio in the audio exchange telecommunications system to obtain a given syllable intelligibility. In this case, experimentally determined the value of the signal-to-noise ratio in the telecommunications system of audio exchange to obtain a syllable intelligibility of at least 93%, to ensure full perception of the transmitted speech information.