

## АНАЛИТИЧЕСКИЙ МЕТОД ПРИБЛИЖЕННОЙ ОЦЕНКИ РАЗБОРЧИВОСТИ ПРЕРЫВАЕМОЙ РЕЧИ

*С. Н. Терентьев, В. М. Рудой*

За последние годы в периодической печати опубликовано большое количество работ, которые свидетельствуют о значительном интересе исследователей к так называемой рваной, или прерываемой речи. Такой вид преобразования речи встречается, например, в системах с временным уплотнением каналов связи, при запираании тракта приемного устройства на время действия помех и т. п. При этом разборчивость прерываемой речи, несмотря на большую избыточность полезных сведений в естественной речи, в значительной мере становится зависимой от параметров прерывания речевого сигнала [1].

Чтобы определить разборчивость преобразованной таким образом речи, в настоящее время очень широко применяется метод артикуляционных измерений. Именно этот метод принят за основной в связи с тем, что для аналитических расчетов нет необходимых объективных характеристик речевого сигнала и соответствующих критериев, определяющих его разборчивость. Поэтому разработка методов аналитического расчета разборчивости речи при различных ее преобразованиях является весьма актуальной задачей.

Известно, что речевой сигнал с достаточной для инженерной практики точностью можно рассматривать как стационарный случайный процесс [2].

Однако при преобразованиях речи подобно тем, которые получаются при ее прерывании, преобразованный речевой сигнал представляет собой процесс вида

$$g(t) = W(t), A(t), \quad (1)$$

где

$W(t)$  — речевой сигнал;

$A(t)$  — функция прерывания.

Для процессов такого вида корреляционная функция равна произведению функций корреляции речи и программы прерывания

$$R_g(\tau) = R_W(\tau) R_A(\tau). \quad (2)$$

Корреляционную функцию речи можно найти, используя одну из возможных аппроксимаций спектра речи [2]

$$G_W(\omega) = \frac{\nu_0 \sigma^2}{\pi} \left[ \frac{1}{\nu_c^2 + (\omega_0 + \omega)^2} + \frac{1}{\nu_0^2 + (\omega_0 - \omega)^2} \right], \quad (3)$$

где  $\nu = 1,14 \cdot 10^3 \text{ сек}^{-1}$ ;  $\omega_0 = 2,98 \cdot 10^3 \text{ сек}^{-1}$ .

Подставляя (3) в формулу Хинчина—Винера, находим значение

$$R_W(\tau) = \sigma^2 e^{-\nu_0 |\tau|} \cos \omega_0 \tau. \quad (4)$$

Для простоты примем программу прерывания в виде периодической последовательности прямоугольных импульсов такую, что

$$A(t) = \begin{cases} +1 & \text{при } kT_c < t < kT_c + T_{\text{пр}} \\ 0 & \text{при } kT_c + T_{\text{пр}} < t < (k+1)T_c, \end{cases} \quad (5)$$

где

$T_c$  — время существования речевого сигнала;

$T_{\text{пр}}$  — период прерывания.

Представим  $A(t)$  в виде

$$A(t) = C_0 + \sum_{m=1}^{\infty} C_m \cos \omega_m t, \quad (6)$$

где

$$\omega_m = m\omega_{\text{пр}}; \quad m = 2l + 1; \quad l = 0, 1, 2, \dots$$

Постоянная составляющая процесса (6) равна

$$C_0 = \frac{2}{T_{\text{пр}}} \int_0^{\frac{T_c}{2}} A(t) dt = \frac{T_c}{T_{\text{пр}}} = d, \quad (7)$$

т. е. коэффициенту заполнения передачи.

Амплитудное значение гармоник

$$C_m = \frac{4}{T_{\text{пр}}} \int_0^{\frac{T_c}{2}} A(t) \cos \omega_m t dt = \frac{2}{\pi m} \sin(\pi m d). \quad (8)$$

Корреляционная функция программы прерывания с учетом найденных значений (7) и (8), определяется из выражения (6)

$$R_A(\tau) = d^2 + \sum_{m=1}^{\infty} \frac{2 \sin^2(\pi m d)}{\pi^2 m^2} \cos m\omega_{\text{пр}} \tau. \quad (9)$$

Подставляя (4) и (9) в (2), можно найти корреляционную функцию прерываемой речи и по формуле Хинчина—Винера вычислить ее энергетический спектр

$$\begin{aligned} G_{g_{\text{пр}}}(\omega) &= \frac{2}{\pi} \int_0^{\infty} R_g(\tau) \cos \omega \tau d\tau = \\ &= \frac{2}{\pi} \int_0^{\infty} \left[ \left( d^2 + \sum_{m=1}^{\infty} \frac{2 \sin^2(\pi m d)}{\pi^2 m^2} \cos m\omega_{\text{пр}} \tau \right) \times \right. \\ &\quad \left. \times \sigma^2 e^{-\nu_0 |\tau|} \cos \omega_0 \tau \right] \cos \omega \tau d\tau = \\ &= \frac{2\sigma^2 d^2}{\pi} \int_0^{\infty} e^{-\nu_0 |\tau|} \cos \omega_0 \tau \cos \omega \tau d\tau + \end{aligned}$$

$$\begin{aligned}
 & + \frac{4\sigma^2}{\pi} \sum_{m=1}^{\infty} \frac{\sin^2(\pi md)}{\pi^2 m^2} \int_0^{\infty} e^{-\nu_0 |\tau|} \cos \omega_0 \tau \cos m\omega_{np} \tau \cos \omega \tau d\tau = \\
 & = \frac{\nu_0 \sigma^2 d^2}{\pi} \left[ \frac{1}{\nu_0^2 + (\omega_0 + \omega)^2} + \frac{1}{\nu_0^2 + (\omega_0 - \omega)^2} \right] + \\
 & + \frac{\nu_0 \sigma^2}{\pi} \sum_{m=1}^{\infty} \frac{\sin^2(\pi md)}{\pi^2 m^2} \left[ \frac{1}{\nu_0^2 + (\omega_0 + \omega + m\omega_{np})^2} + \frac{1}{\nu_0^2 + (\omega_0 + \omega - m\omega_{np})^2} + \right. \\
 & \left. + \frac{1}{\nu_0^2 + (\omega_0 - \omega + m\omega_{np})^2} + \frac{1}{\nu_0^2 + (\omega_0 - \omega - m\omega_{np})^2} \right]. \quad (10)
 \end{aligned}$$

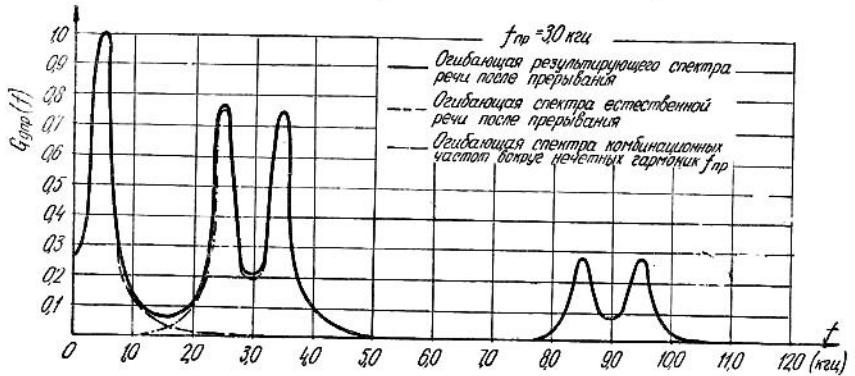


Рис. 1.

В формуле (10) первое слагаемое представляет собой исходный спектр речи до прерывания, описываемый выражением (3), умноженный на квадрат коэффициента заполнения передачи. Второе слагаемое — спектр комбинационных частот вокруг каждой нечетной гармоники частоты прерывания  $\omega_{np}$ , умноженный на свой коэффициент гармоник.

Поэтому формулу (10) можно представить в более наглядном виде

$$\begin{aligned}
 G_{gnp}(\omega) &= d^2 G_w(\omega) + \frac{1}{\pi^2} \sum_{m=1}^{\infty} \frac{\sin^2(\pi md)}{m^2} \times \\
 & \times [G_w(\omega_0 + \omega + m\omega_{np}) + G_w(\omega_0 + \omega - m\omega_{np}) + \\
 & + G_w(\omega_0 - \omega + m\omega_{np}) + G_w(\omega_0 - \omega - m\omega_{np})]. \quad (11)
 \end{aligned}$$

Произведя в (10) замену  $\omega$  на  $f$ , получим следующее выражение для спектра прерываемой речи:

$$\begin{aligned}
 G_{gnp}(f) &= \frac{d^2 \sigma^2 \nu_0}{\pi} \left[ \frac{1}{\nu_0^2 + 4\pi^2 (f_0 + f)^2} + \frac{1}{\nu_0^2 + 4\pi^2 (f_0 - f)^2} \right] + \\
 & + \frac{\sigma^2 \nu_0}{\pi} \sum_{m=1}^{\infty} \frac{\sin^2(\pi md)}{\pi^2 m^2} \times \\
 & \times \left[ \frac{1}{\nu_0^2 + 4\pi^2 (f_0 + f + mf_{np})^2} + \frac{1}{\nu_0^2 + 4\pi^2 (f_0 + f - mf_{np})^2} + \right. \\
 & \left. + \frac{1}{\nu_0^2 + 4\pi^2 (f_0 - f + mf_{np})^2} + \frac{1}{\nu_0^2 + 4\pi^2 (f_0 - f - mf_{np})^2} \right]. \quad (12)
 \end{aligned}$$

Для изучения вопроса о влиянии частоты прерывания на вид спектра преобразованного речевого сигнала и на разборчивость прерываемой речи, были рассчитаны по формуле (12) и построены спектры речевого сигнала для различных  $f_{пр}$ .

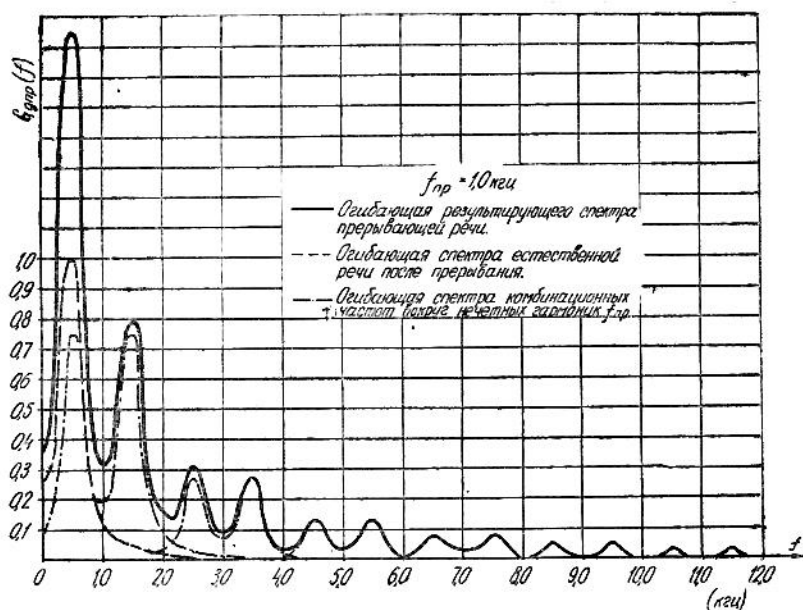


Рис. 2.

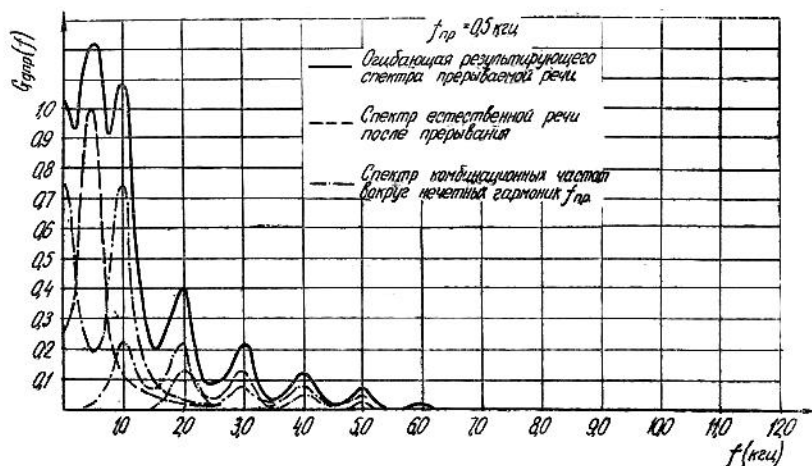


Рис. 3.

Огибающие расчетных спектров приведены на рис. 1, 2 и 3. На рисунках пунктиром изображена огибающая спектра, которая совпадает с исходным спектром речи; штрих-пунктиром — огибающая комбинационных частот вокруг нечетных гармоник частоты прерывания; сплошной линией — огибающая результирующего спектра прерываемой речи.

Рисунки дают наглядное представление о структуре спектра прерываемой речи и зависимости вида спектра от  $f_{пр}$ .

Очевидно, что спектр  $G_w(f) d^2$  несет высокую разборчивость речи, так как он полностью совпадает с исходным спектром речевого сигнала до прерывания. Однако в спектре полезного (исходного) сигнала после преобразования появляются частоты, которые до этого отсутствовали. Это значит, что спектр комбинационных частот является мешающим, так как он в зависимости от частоты прерывания в той или иной мере накладывается на полезный спектр речи, вызывая, тем самым, ухудшение ее разборчивости.

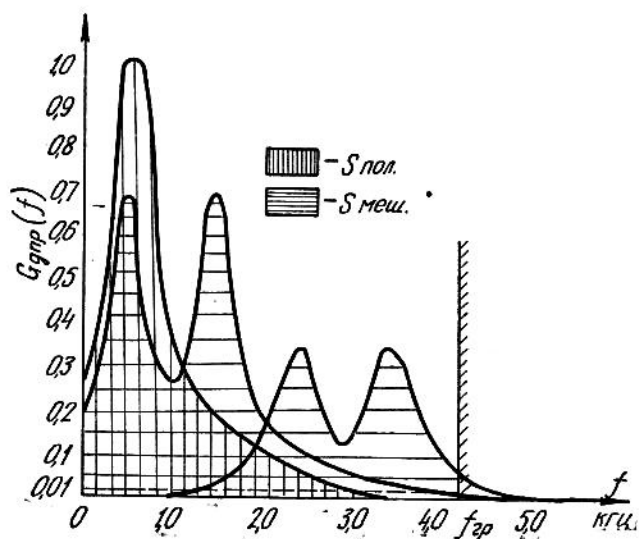


Рис. 4.

Следовательно, результирующая степень неразборчивости преобразованного сигнала сложным образом зависит от взаимодействия составляющих комбинационных частот и той его части, которая образует речевой сигнал.

Для достаточно приближенной оценки степени нарушения разборчивости речи после прерывания можно полагать, что влияние комбинационных участков спектра пропорционально отношению их энергии к энергии полезной части спектра, расположенной в полосе частот от нуля до частоты, на которой составляющие полезного спектра имеют квадрат амплитуды в 100 раз меньше максимума энергетического спектра.

Такая оценка весьма удобна, так как вычисление указанного отношения может быть осуществлено довольно просто путем графического или аналитического определения отношения площадей соответствующих участков спектра.

Рассмотренный метод оценки разборчивости прерываемой речи поясняется рис. 4.

На рис. 4. площадь под вертикальной штриховкой принадлежит полезному спектру речи и обозначена  $S_{пол}$ , а площади под горизонтальной штриховкой принадлежат комбинационным частотам и обозначены  $S_{меш.}$

Результаты расчета, изложенным выше способом, представлены графиком зависимости  $\frac{S_{\text{пол}}}{S_{\text{пол}} + S_{\text{меш}}} = F(f)$ , приведенным на рис. 5 (сплошные линии).

С целью проверки изложенного метода были проведены артикуляционные измерения разборчивости слов от частоты прерывания речи. Для измерения артикуляции слов была использована методика и артикуляционные таблицы, рекомендованные в работе [3].

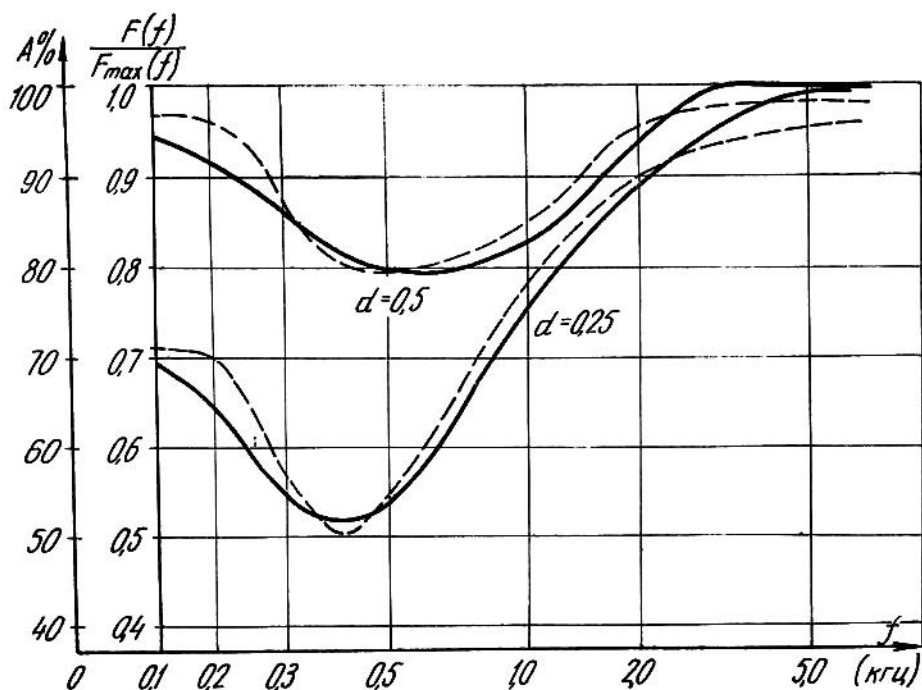


Рис. 5.

На основании полученных данных построен график зависимости разборчивости слов  $A\%$  от частоты прерывания и коэффициента заполнения передачи (рис. 5 пунктир), из которого видно, что максимальное нарушение разборчивости речи совпадает с расчетным и имеет место в районе  $f_{\text{пр}} \approx 500$  гц.

Из сравнения графиков, приведенных на рис. 5 (сплошные и пунктирные линии), можно заключить, что предлагаемый метод ориентировочной оценки нарушения разборчивости прерываемой речи обладает вполне достаточной точностью для инженерной практики.

## ЛИТЕРАТУРА

1. М. А. Сапожников. Речевой сигнал в кибернетике и связи. Связьиздат, М., 1963.
2. А. И. Величкин. Амплитудное ограничение речи. «Акустический журнал», т. VIII, вып. 2, 1962.
3. Н. Б. Покровский. Расчет и измерение разборчивости речи. Связьиздат, М., 1962.