Научная статья
УДК 004.056: 621.391
DOI: 10.26583/bit.2025.4.02

ИССЛЕДОВАНИЕ СПЕКТРОВ ЗВУКОВ РУССКОЙ РЕЧИ

Анатолий А. Хорев[1], Илья А. Баранов[2]
Национальный исследовательский университет «МИЭТ», пл. Шокина, 1, Москва, Зеленоград, 124498, Россия
[1]horev@miee.ru, https://orcid.org/0000-0001-9074-385X
[2]sararalilya@gmail.com, https://orcid.org/0009-0006-9485-282X

Аннотация. В качестве показателя оценки эффективности защиты акустической речевой информации от ее утечки по техническим каналам используется словесная разборчивость речи W, в основу методики расчета которой положен формантный метод оценки разборчивости речи. При расчетах в качестве исходных данных по речевому сигналу используются форматный параметр речи ΔA_i и коэффициент k_i, показывающий относительную встречаемость формант в октавных полосах. Значения коэффициентов ΔA_i и k_i получены в ходе экспериментальных исследований, проведенных в середине XX века с использованием аналоговой техники. Поэтому целью данных исследований является оценка значений коэффициентов ΔA_i и k_i на основе результатов экспериментальных исследований, проводимых с использованием современных методов цифровой обработки речевых сигналов. В результате цифровой обработки аудиозаписей звуков русской речи, записанных 9 дикторами (4 мужчинами и 5 женщинами) получены: спектры звуков и спектры речи для мужских и женских голосов. На основе обработки спектров звуков установлены частотные интервалы звуков и формант мужской и женской речи. Предложен метод расчета коэффициентов встречаемости формант в октавных полосах k_i, на основе которого получены значения этих коэффициентов для мужской и женской речи. На основе обработки спектров речи определены нормированные значения уровней речевого сигнала в октавных полосах для типовых интегральных уровней громкости для мужской и женской речи. Предложен метод расчета формантных параметров речи в октавных полосах ΔA_i, на основе которого получены значения этих параметров для мужской и женской речи. На основе полученных результатов экспериментальных исследований проведен расчет разборчивости речи в зависимости от отношения сигнал/шум для мужской и женской речи. Анализ результатов расчета разборчивости речи показал, что наибольшая разборчивость наблюдается для мужской речи.

Ключевые слова: *защита речевой информации, словесная разборчивость речи, формантный метод оценки разборчивости речи.*

Для цитирования: *Хорев, Анатолий А.; Баранов, Илья А. Исследование спектров звуков русской речи. Безопасность информационных технологий, [S.l.], т. 32, № 4, с. 18–36, 2025. ISSN 2074-7136. URL: 2/3/6. DOI: http://dx.doi.org/10.26583/bit.2025.4.02.*

Scientific article

RESEARCH OF THE SPECTRA OF RUSSIAN SPEECH SOUNDS

Anatoly A. Horev[1], Ilya A. Baranov[2]
National Research University of Electronic Technology, Shokin Sq., 1, Moscow, Zelenograd, 124498, Russia
[1]horev@miee.ru, https://orcid.org/0000-0001-9074-385X
[2]sararalilya@gmail.com, https://orcid.org/0009-0006-9485-282X

Abstract. The word intelligibility of speech W is used as an indicator for assessing the effectiveness of protecting acoustic speech information from its leakage through technical channels. The calculation method is based on the formant method of speech intelligibility assessment. The calculations use the formant parameter of speech ΔA_i and the coefficient k_i, which shows the relative occurrence of formants in octave bands, as the initial data for the speech signal. The values of the coefficients ΔA_i and k_i were obtained through experimental studies conducted in the mid-20th century using analog technology. Therefore, the purpose of this study is to estimate the values of the coefficients ΔA_i and k_i based on the results of experimental studies conducted using modern methods of digital processing of speech signals. As a result of digital processing of audio recordings of Russian speech sounds recorded by 9 speakers (4 men and 5 women), the following were obtained: sound spectra and speech spectra for male and female voices. Based on the processing of sound spectra, the frequency intervals of sounds and formants of male and female speech were determined. A method for calculating the relative occurrence coefficients of formants in octave bands k_i was proposed, and the values of these coefficients were obtained for male and female speech. Based on the processing of speech spectra, normalized values of speech signal levels in octave bands are determined for typical integral loudness levels for male and female speech. A method for calculating the formant parameters of speech in octave bands ΔA_i is proposed, based on which the values of these parameters for male and female speech are obtained. Based on the results of the experimental studies, the speech intelligibility is calculated as a function of the signal-to-noise ratio for male and female speech. The analysis of the speech intelligibility calculation results showed that the highest intelligibility is observed for male speech.

Keywords: *speech information security, verbal speech intelligibility, formant speech intelligibility assessment method.*

For citation: *Horev, Anatoly A.; Baranov, Ilya A. Research of the spectra of russian speech sounds. IT Security (Russia), [S.l.], v. 32, no. 4, p. 18–36, 2025. ISSN 2074-7136. URL: https://bit.spels.ru/index.php/bit/article/view/1859. DOI: http://dx.doi.org/10.26583/bit.2025.4.02.*

Введение

В начале 2000-х годов в качестве показателя оценки эффективности защиты акустической речевой информации от ее утечки по техническим каналам вместо отношения сигнал/шум в октавных полосах стала использоваться словесная разборчивость речи.

В основу методики расчета словесной разборчивости речи был положен формантный метод оценки разборчивости речи, разработанный профессором Н.Б. Покровским в середине XX века [1].

Данный метод предполагал разбитие речевого диапазона частот на 20 равноартикуляционных полос, вносящих одинаковый вклад в разборчивость речи. Для каждой полосы определялся уровень восприятия формант в зависимости от отношения сигнал/шум. На основание этих данных сначала рассчитывалась формантная разборчивость речи, а затем – словесная разборчивость речи.

В целях сокращения объема проводимых измерений и расчетов в [2] предложено проводить измерения уровней сигналов и шумов и расчет отношений сигнал/шум не в равноартикуляционных полосах, а в октавных полосах. Аналогичный подход рассмотрен в [3].

Разработанная на основе предложенного метода методика оценки эффективности защиты акустической речевой информации от ее утечки по техническим каналам предполагала измерение отношений сигнал/шум (q_i) на входе средства разведки в октавных полосах и расчет словесной разборчивости речи (W) по эмпирическим формулам [4, 5]:

$$Q_i = q_i - \Delta A_i; \qquad (1)$$

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

$$R_i = p_i \cdot \kappa_i; \quad (3)$$

$$R = \sum_{i=1}^{7} R_i; \quad (4)$$

$$W = \begin{cases} 1,54 \cdot R^{0,25} [1 - \exp(-11 \cdot R)], \text{ если } R < 0,15; \\ 1 - \exp\left(-\dfrac{11 \cdot R}{1 + 0,7 \cdot R}\right), \text{ если } R \ge 0,15, \end{cases} \quad (5)$$

где ΔA_i – значение формантного параметра речи i-й октавной полосы, дБ;

i – номер октавной полосы;

p_i – коэффициент восприятия формант i-й октавной полосы;

R_i – спектральный индекс артикуляции речи (формантная разборчивость речи) i-й октавной полосы;

k_i – весовой коэффициент (встречаемость формант речи) i-й октавной полосы;

R – интегральный индекс артикуляции речи.

Значения коэффициентов ΔA_i и k_i для пяти октавных полос (со второй по шестую) приведены в [4], а для семи октавных полос – в [5] (см. табл. 1).

Таблица 1. Основные характеристики речи

Номер полосы	Средне-геометрическая частота полосы $f_{ср}$, Гц	Уровни речевого сигнала, дБ, в зависимости от вида речи				Весовой коэффициент полосы, κ_i	Значение формантного параметра речи, ΔA_i, дБ
		Тихая речь, 70 дБ	Речь средней громкости, 70 дБ	Громкая речь, 76 дБ	Очень громкая речь, 84 дБ		
1	125	47	53	59	67	0,01	25
2	250	60	66	72	80	0,03	18
3	500	60	66	72	80	0,12	14
4	1000	55	61	67	75	0,20	9
5	2000	50	56	62	70	0,30	6
6	4000	47	53	59	67	0,26	5
7	8000	43	49	55	63	0,07	4

В [6, 7] приведены результаты исследования влияния на разборчивость речи коэффициентов восприятия формант.

Попытки усовершенствования методики оценки эффективности защиты акустической речевой информации от ее утечки по техническим каналам, связанные с упрощением формул для расчета разборчивости речи, предприняты в [8], а связанные с изменением спектра речи при ее «форсировании» в [9–11].

Однако, во всех рассмотренных выше работах при расчете разборчивости речи использовались значения коэффициентов ΔA_i и k_i, полученные профессором Н.Б. Покровским с использованием аналоговой техники [1]. Развитие цифровых технологий и новых методов спектрального анализа речевых сигналов открывают возможности по оценке этих коэффициентов, существенно влияющих на разборчивость речи.

Поэтому целью данных исследований является оценка значений коэффициентов ΔA_i и k_i на основе результатов экспериментальных исследований, проводимых с использованием современных методов цифровой обработки речевых сигналов.

В качестве программного обеспечения для обработки речевых сигналов использовалась программа Adobe Audition CC 2017, а для обработки результатов экспериментальных исследований было разработано специальное программное обеспечение на языке программирования Python.

Основная часть исследования

Экспериментальные исследования проводились в пять этапов.

На **первом (подготовительном) этапе** осуществлялась аудиозапись 42-х звуков русской речи. В качестве дикторов использовались студенты 5-го курса университета (4 мужчины и 5 женщин), не имеющих речевых или слуховых нарушений. Запись проводилась с расстояния 1 м от диктора с использованием микрофона ZT-333 и предусилителя P410. Параметры записи: частота дискретизации 58593 Гц, разрядность 20 бит, чувствительность микрофона 48,14 мВ/Па, формат записи **.wav*. Каждый диктор произносил звуки последовательно, с короткими паузами между ними, каждый звук повторялся 10 раз.

С использованием программного обеспечения Adobe Audition для каждого диктора из аудиофайла с записями всех звуков выделялись отдельные файлы, содержащие повторения только одного звука. Каждый такой файл разбивался на сегменты фиксированной длительности, соответствующей длительности произношения звука. Длина сегментов выбиралась исходя из компромисса между временным и частотным разрешением: слишком короткие интервалы увеличивали погрешность спектральной оценки, а длинные – включали в сегмент запись неинформативного шума. Границы временных сегментов звука и шума определялись вручную.

Полученная аудиозапись каждого звука, произнесенного каждым диктором, прослушивалась аудитором. Если диктор произнёс звук разборчиво и в записи нет посторонних шумов, то запись этого звука включалась в выборку для дальнейшего анализа. Если звук был неразборчив или похож на другой звук, в программе выбирались иные временные метки, или запись данного звука исключалась из выборки.

Далее с помощью инструмента «Выделение времени» на фонограмме или спектрограмме выделялся временной интервал произношения звука, как разница между меткой конца произношения звука и меткой его начала (рис. 1).

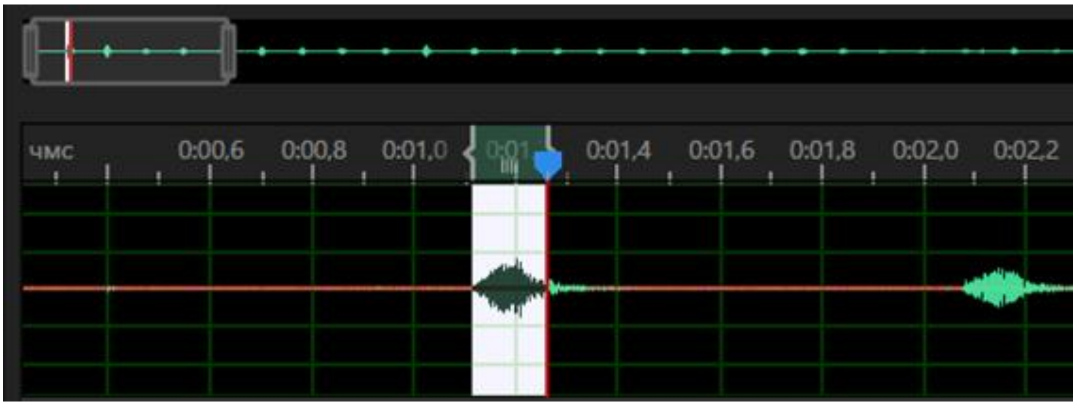


Рис. 1. Определение длительности произношения звука

Длительность произношения, метки времени звука и шума сохранялись для последующего использования.

На **втором этапе** с использованием разработанного специального программного обеспечения проводились построения спектров звуков.

Вначале проводились построения спектров звуков одного диктора. Исходные данные для построения включали временные метки сигнала и шума и длительности произношения звуков.

Для каждого сегмента, соответствующего длительности произношения звука, выполнялись операции выделения временного интервала на основе предоставленных меток и построения спектров звуков с использованием быстрого преобразования Фурье (БПФ).

Для каждой записи определялся осредненный спектр смеси «сигнал+шум» и шума, где амплитуда сигнала определялась по формуле:

$$L_i = \sqrt{\frac{\sum_{j=1}^{N} L_{i.j}^2}{N}}, \tag{6}$$

где L_i – среднеквадратическая амплитуда звука на i-й частоте;

$L_{i.j}$ – амплитуда j-го повторения звука на i-й частоте;

N – количество повторений каждого звука.

Далее строился спектр «чистого» звука, где амплитуда сигнала определялась по формуле:

$$L_{c.i} = \sqrt{L_{(c+ш).i}^2 - L_{ш.i}^2}, \tag{7}$$

где $L_{c.i}$ – амплитуда «чистого» звука на i-й частоте;

$L_{(c+ш).i}$ – амплитуда смеси «сигнал+шум» на i-й частоте;

$L_{ш.i}$ – амплитуда шума на i-й частоте.

Пример полученных спектров смеси «сигнал+шум» и шума приведен на рис. 2, а спектра «чистого» сигнала – на рис. 3.

Далее проводилась построение осредненных спектров звуков для всех дикторов. С целью нивелирования разности в громкости звуков для разных дикторов использовалась нормализация амплитуд относительно их максимального значения. Амплитуда сигнала рассчитывалась по формуле:

$$L_i = \sqrt{\frac{\sum_{n=1}^{M} L_{i.n}^2}{M}}, \tag{8}$$

где $L_{i.n}$ – амплитуда звука на i-й частоте n-го диктора; M – количество дикторов.

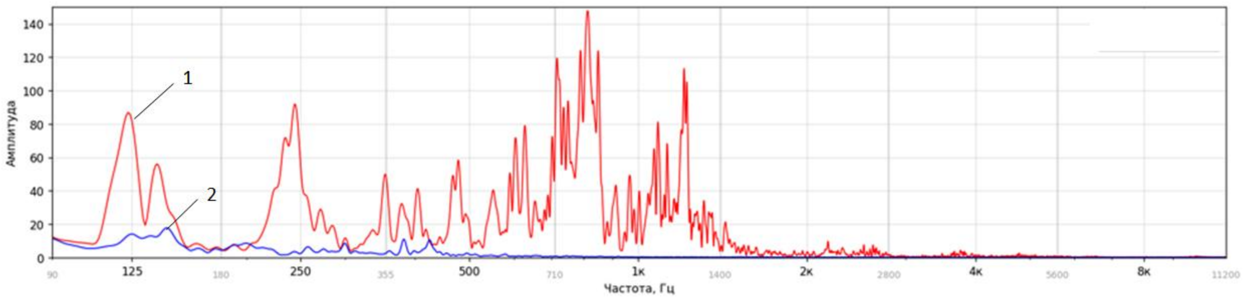


Рис. 2. Усреднённый спектр смеси «сигнал+шум» звука «А» (1) и шума (2)

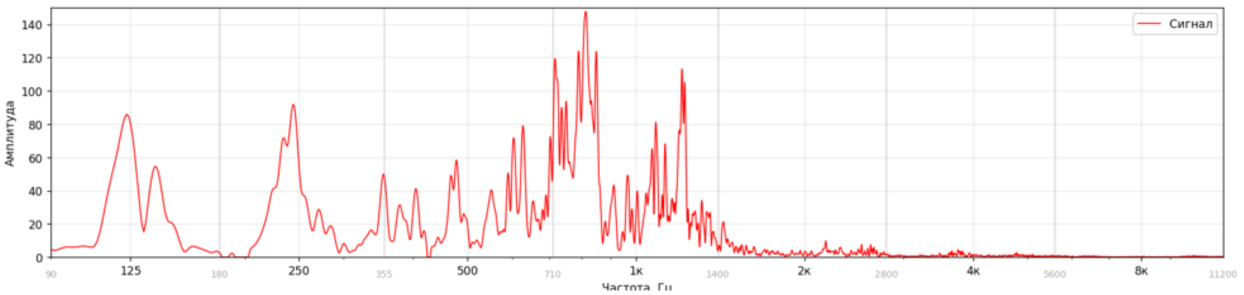
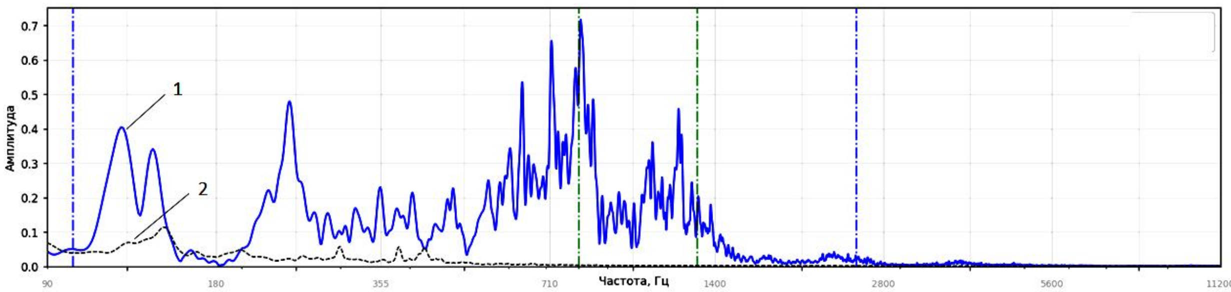
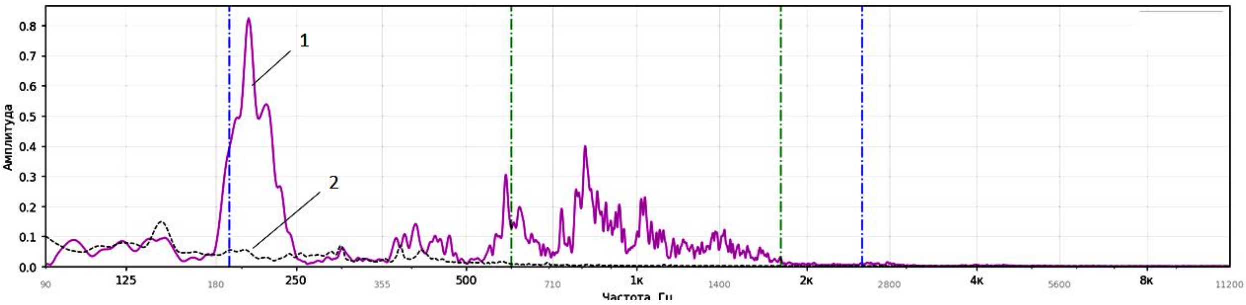


Рис. 3. Усреднённый спектр «чистого» звука «А»

Примеры полученных спектров звуков мужских и женских голосов приведены на рис. 4.



а)



б)

Рис. 4. Усреднённый спектр звука «А» (1) и шума (2) для мужских (а) и для женских (б) голосов

На **третьем этапе** проводилось определение частотных интервалов формант звуков речи.

Границы частотных областей формант определялись с использованием программного обеспечения Adobe Audition путем «обрезания» спектра звука.

Сначала выбиралась область наиболее возможного нахождения формант спектра звука. Далее полоса спектра сужалась: сначала слева, а затем справа, пока разборчивость звука не начинала снижаться.

Границы частотной области форманты фиксировались, если разборчивость звука в этой полосе сохранялась для всех дикторов рассматриваемого пола.

Примеры полученных спектров звуков и формант звуков для мужских и женских голосов приведены на рис. 5.

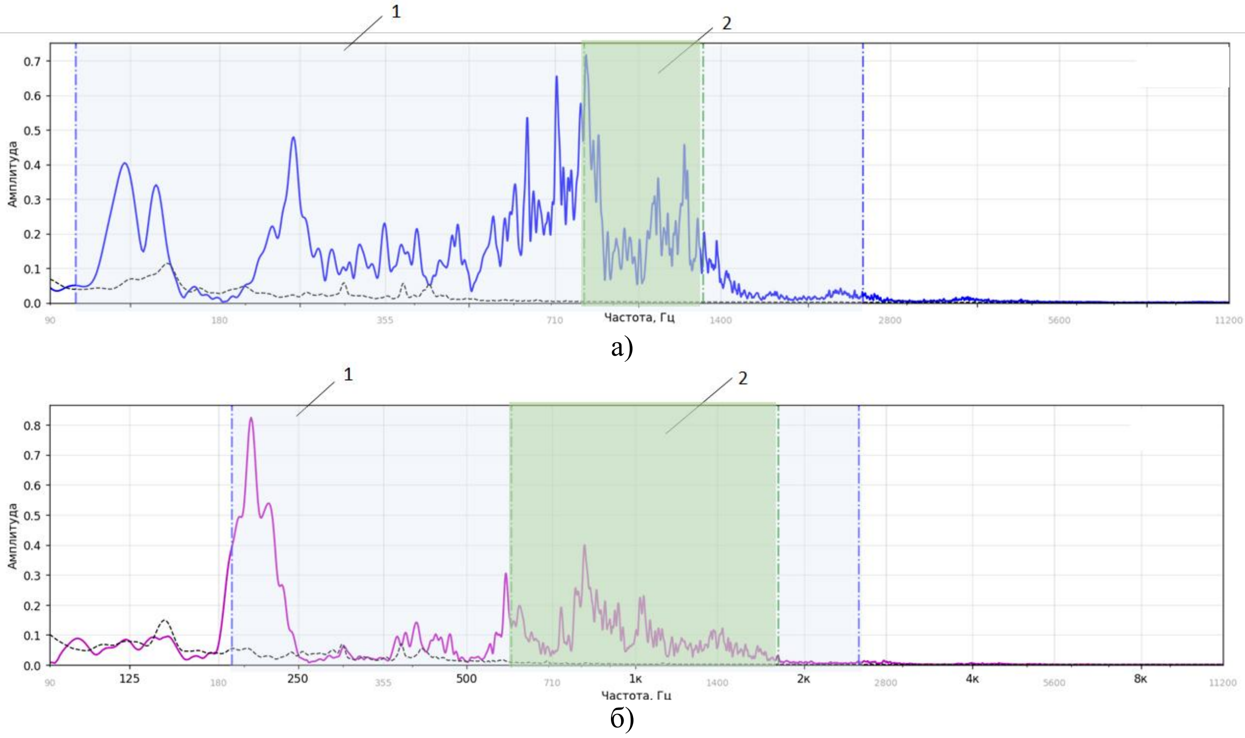


Рис. 5. Усреднённый спектр (1) звука «А» и формант (2) этого звука для мужских (а) и для женских (б) голосов

Проведенный анализ доступной литературы показал, что данные по частотным интервалам формантных областей звуков довольно противоречивы как русской речи [1, 12, 13], так и для английской [14–16].

В табл. 2 приведены границы формант звуков, полученные в результате экспериментальных исследований, в сравнении с границами формант звуков, полученными некоторыми другими авторами.

Анализ данных, представленных в табл. 2, показывает существенное отличие полученных результатов экспериментальных исследований от данных, представленных в [12], и значительное совпадение с данными, приведенными в [1].

Очевидно, эти отличия определяются прежде всего отличительными особенностями дикторов (при проведении экспериментальных исследований использовались молодые дикторы примерно одного возраста), а также разными методиками определения частотных интервалов звуков.

Таблица 2. Частотные интервалы формантных областей звуков речи

Звук	Формантные области Ф1/Ф2			
	Полученные по результатам экспериментальных исследований		Полученные по данным [1] (смешенная речь)	Полученные по данным [12] (смешенная речь)
	Мужской речи	Женской речи		
А	800-1300	600-1800	1100-1400	400-700/1100-1500
О	400-900	400-900	400-800	200-500/800-1000
У	200-740	200-700	200-600	100-200/800-1100
И	1500-4100	2500-4600	2800-4200	1800-2000/2200-2500
Ы	800-2100	800-1700	200-600/1500-2300	1800-2000
Э	250-750/1500-2300	150-700/1500-2300	600-1000/1600-2500	200-500/1900-2300
З	150-600/6600-10200	190-550/8800-10500	0-600/4200-8600	-
Ж	1600-4500	2000-4200	200-600/1350-6300	-
Л	150-1400	400-1300	200-500/700-1100	-
М	600-1200	500-1200	0-400/1600-1850	-
Н	800-1600	650-1900	0-400/1500-3400	-
Р	150-1400	500-1600	200-1500	-
С	4500-7200	5800-8000	4200-8500	-
Ф	1300-8000	1500-8000	7000-12000	-
Х	600-2600	600-2600	400-1200	-
Ш	1900-6700	1900-6700	1200-6300	-
Б	750-1300	500-1300	-	-
Бь	1600-3200	1800-3200	-	-
В	200-1300	450-1500	-	-
Вь	1800-2800	1800-3200	-	-
Г	700-1600	900-2000	-	-
Гь	1900-3200	2200-3300	-	-
Д	1200-2300	1300-2300	-	-
Дь	1800-2800	2000-3300	-	-
Зь	2100-4800	170-450/4000-8000	-	-
Й	2000-4000	2300-4100	-	-
К	800-3900	1000-4500	-	-
Кь	2300-8500	2800-7400	-	-
Ль	1000-2800	700-2400	-	-
Мь	1000-2500	1100-2800	-	-
Нь	400-2200	400-2600	-	-
П	900-4500	1000-2600	-	-
Пь	1800-5100	1000-3500	-	-
Рь	1000-2600	1700-2800	-	-
Сь	3000-7800	4000-9700	-	-
Т	1200-5000	1200-5000	-	-
Ть	2400-7700	2300-8300	-	-
Фь	2100-5600	2100-5700	-	-
Хь	2200-8200	2200-8200	-	-
Ц	2600-7300	2600-7300	-	-
Ч	2600-6800	2400-6800	-	-
Щ	2500-6800	2600-7000	-	-

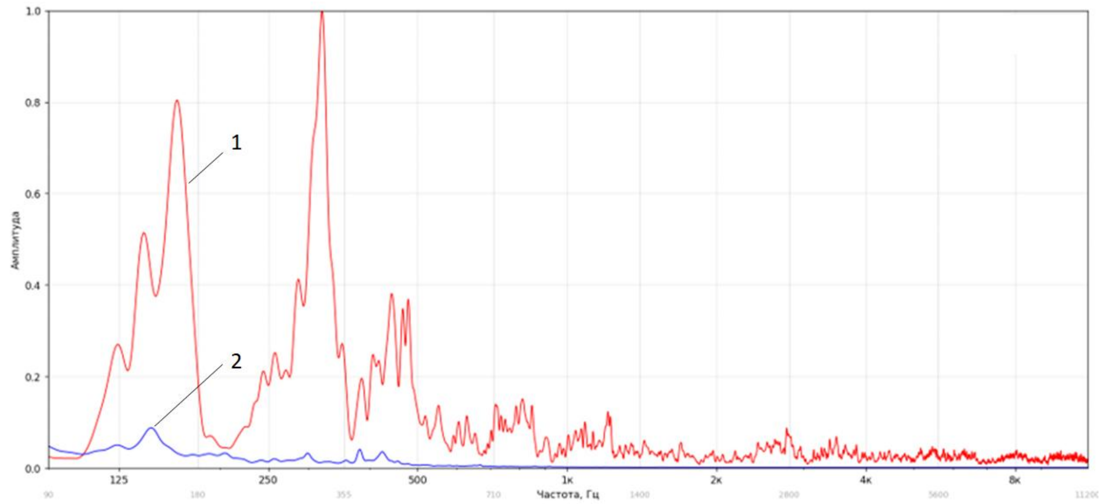
На **четвертом этапе** с использованием разработанного специального программного обеспечения проводились построения спектров речи и расчет форматного параметра в октавных полосах ΔA_i и расчет весового коэффициента k_i, показывающего относительную встречаемость формант в октавных полосах.

Вначале проводилось построение спектров речи одного диктора. Исходные данные для построения включали ненормированные спектры всех звуков речи этого диктора. При этом амплитуда сигнала рассчитывалась по формуле:

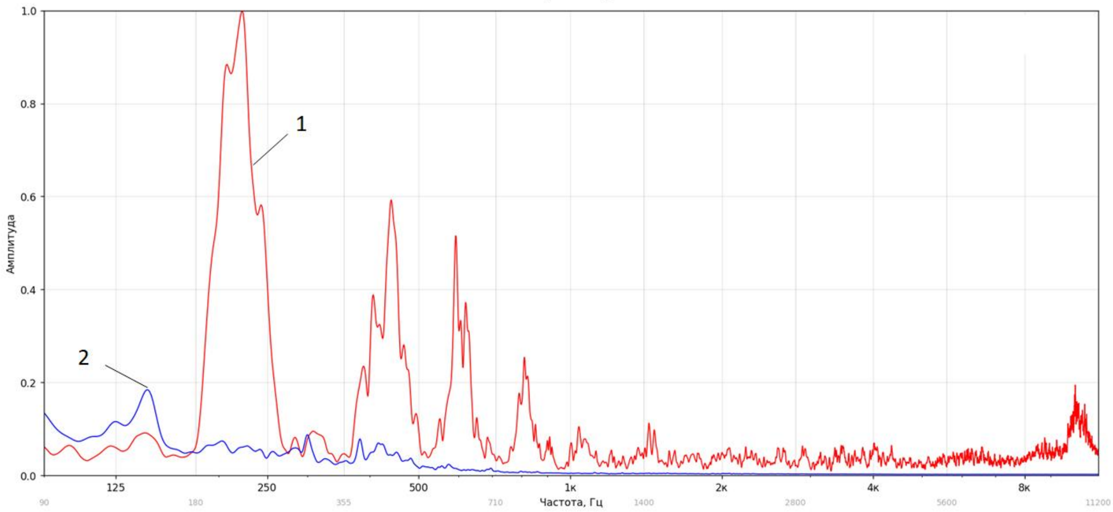
$$L_i = \sqrt{\frac{\sum_{m=1}^{K} L_{i.m}^2}{K}}, \tag{9}$$

где $L_{i.m}$ – амплитуда m-го звука на i-й частоте; K – количество звуков речи (K = 42).

Далее проводилось нормирование амплитуды сигнала относительного его максимального значения во всем частотном диапазоне. Примеры полученных нормированных спектров речи одного диктора приведены на рис. 6.



а)



б)

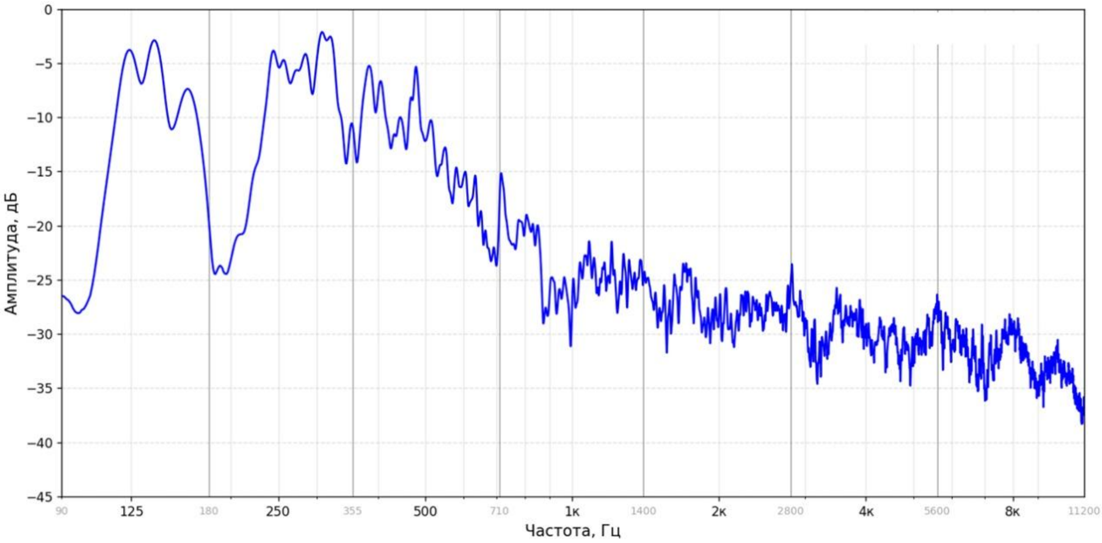
Рис. 6. Усреднённый спектр речи одного диктора для мужских (а) и для женских (б) голосов: 1 – сигнал; 2 – шум.

Далее проводилась построение осредненных спектров речи всех дикторов. Амплитуда сигнала рассчитывалась по формуле:

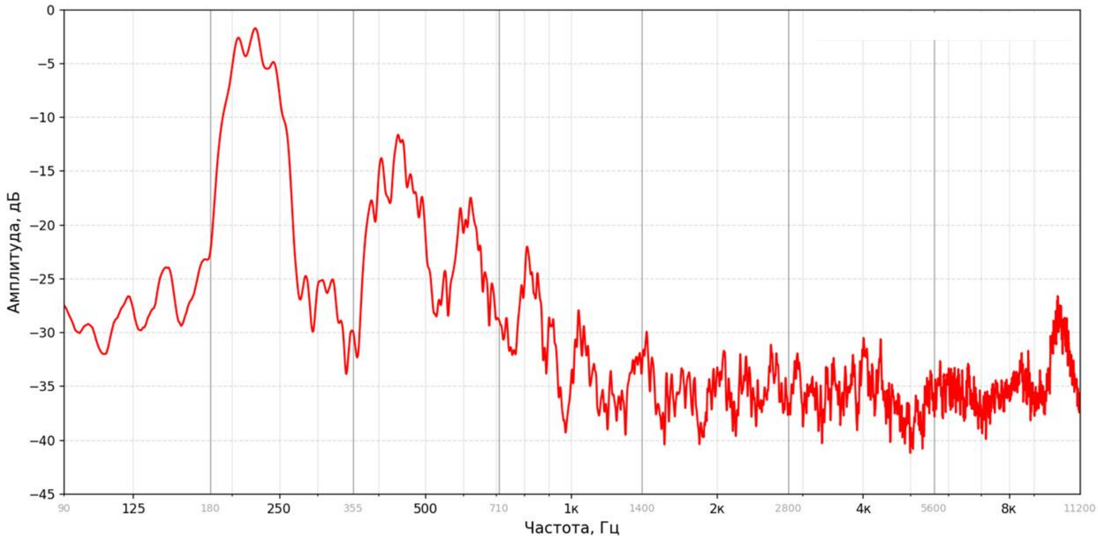
$$L_i = \sqrt{\frac{\sum_{n=1}^{M} L_{i.n}^2}{M}}, \tag{11}$$

где $L_{i.n}$ – нормированная амплитуда «чистого» сигнала (см. формулу (10)) на i-й частоте для n-го диктора; M – количество дикторов.

Полученные спектры речи приведены на рис. 7.



а)



б)

Рис. 7. Усреднённый спектр речи для мужских (а) и для женских (б) голосов

Анализ полученных данных показал, что отличие максимальных и минимальных значений в смешенной речи может достигать 15–20 дБ (см. рис. 8).

В [1] отсутствует описание метода расчета формантных параметров речи в октавных полосах ΔA_i, поэтому была предложена следующая методика их расчета.

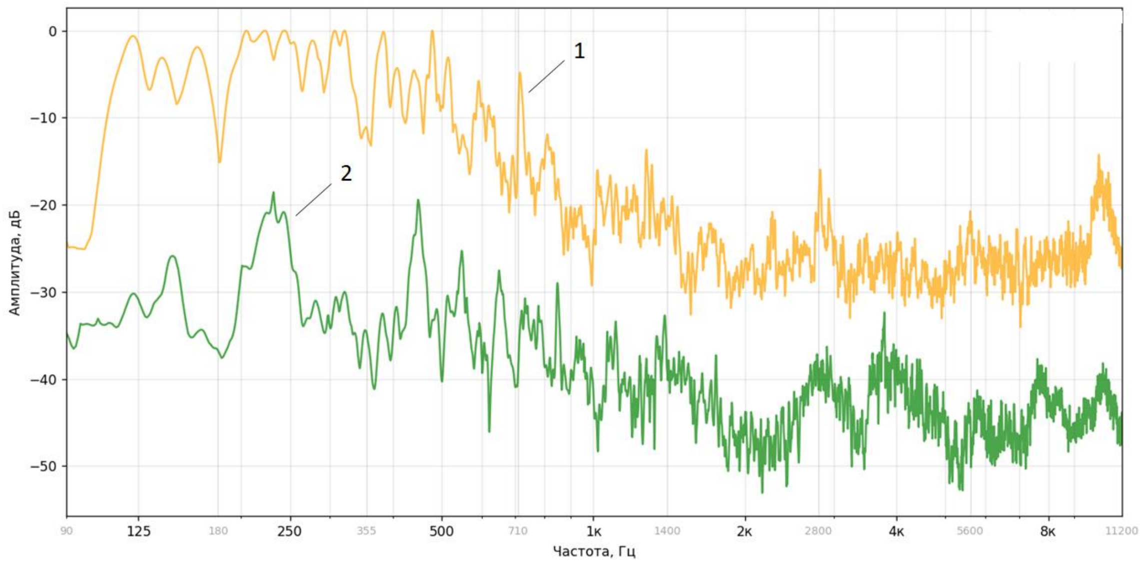
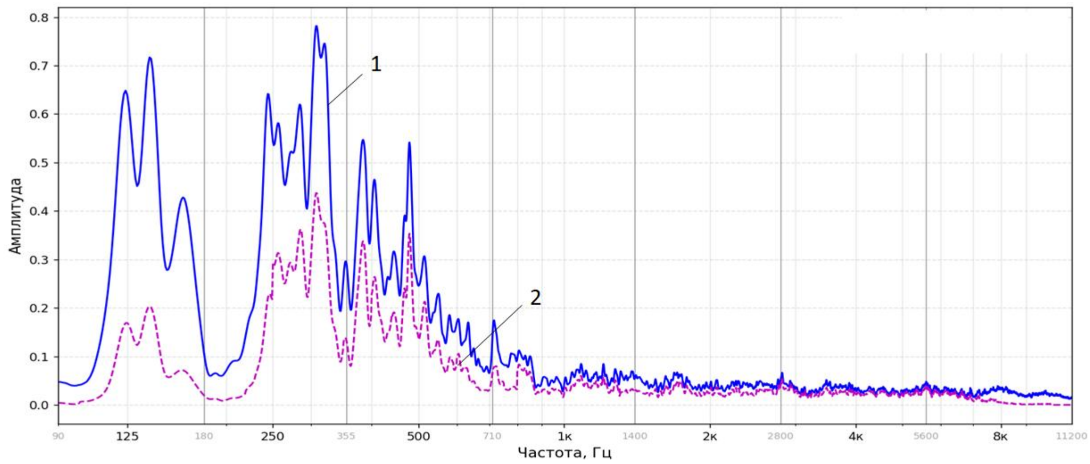
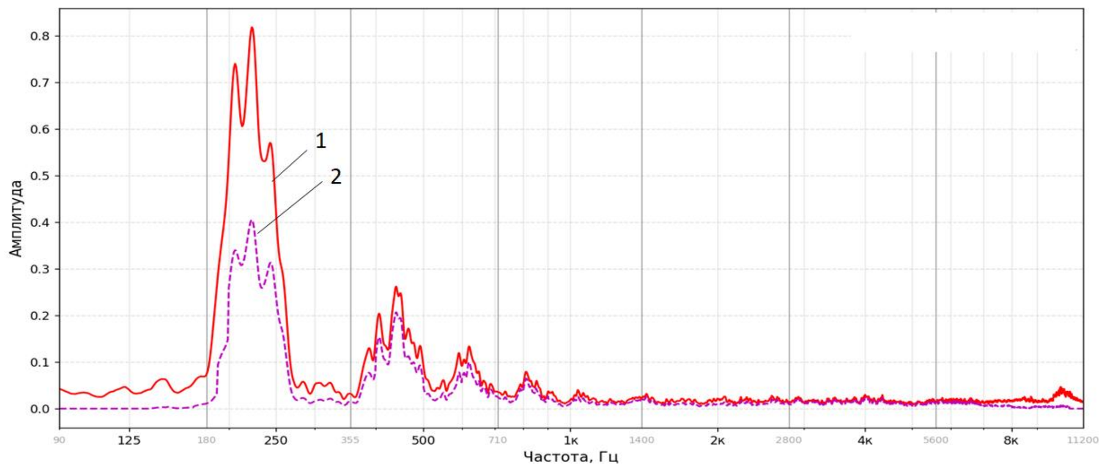


Рис. 8. Спектр максимальных (1) и минимальных (2) значений амплитуды смешенной речи

Вначале по формуле, аналогичной (11), были получены спектры формант для мужской и женской речи (см. рис. 9).



а)



б)

Рис. 9. Усреднённые спектры речи (1) и формант (2) для мужских (а) и для женских (б) голосов

Далее, на основе полученных графиков был проведен расчет формантных параметров речи в октавных полосах ΔA_i по формулам:

$$L_{c.i} = \sqrt{\sum_{f_i = f_{i.н}}^{f_i = f_{i.в}} L_c^2\left(f_i\right) \cdot \Delta F} \; ; \tag{12}$$

$$L_{ф.i} = \sqrt{\sum_{f_i = f_{i.н}}^{f_i = f_{i.в}} L_ф^2\left(f_i\right) \cdot \Delta F} \; ; \tag{13}$$

$$\Delta A_i = 20 \lg \frac{L_{c.i}}{L_{ф.i}} , \tag{14}$$

где $L_{c.i}$ – амплитуда речевого сигнала в i-й октавной полосе;

$L_{c.i}$ – амплитуда формант в i-й октавной полосе;

ΔA_i – значение формантного параметра речи в i-й октавной полосе.

Результаты расчета формантных параметров речи в октавных полосах приведены в табл. 3.

Таблица 3. Значения формантного параметра речи в октавных полосах ΔA_i

Октавная полоса частот, Гц	Значения формантного параметра речи ΔA_i, дБ		
	Мужская речь	Женская речь	Смешенная речь [4, 5]
90-180	20,98	22,36	25
180-355	12,71	16,91	18
355-710	7,48	7,13	14
710-1400	3,48	2,88	9
1400-2800	4,3	4,26	6
2800-5600	2,93	2,57	5
5600-11200	6,91	9,08	4

Как можно заметить, полученные в ходе экспериментальных исследований значения формантных параметров речи ΔA_i значительно отличаются от значений, приведенных в [4, 5]. Очевидно, данные отличия обусловлены различными методиками определения значений формантных параметров речи и отличительными особенностями дикторов.

Количество формант, попадающих в разные октавные полосы, существенно отличаются. Каждая октавная полоса вносит свой вклад в разборчивость речи. Для оценки вклада октавной полосы в разборчивость речи используется весовой коэффициент k_i, который показывает относительную встречаемость формант речи в i-й октавной полосе (см. табл. 1).

В работах [1, 3, 13, 17, 18] рассмотрено влияние коэффициентов k_i на разборчивость речи. Однако методики их расчета в работах не представлено.

В методике оценки разборчивости речи [4, 5] значения коэффициента k_i для октавных полос взяты из графика 5.7, приведенного в [1].

Учитывая, что спектры формант могут быть распределены не в одной, а в нескольких октавных полосах, предложено коэффициенты встречаемости формант звуков в октавных полосах k_i рассчитывать по формулам:

$$k_i = \sqrt{\frac{\sum_{f_n = f_{i.н}}^{f_n = f_{i.в}} L_n^2\left(f_n\right)\cdot \Delta F}{\sum_{f_n = 90\ Гц}^{f_n = 11200\ Гц} L_n^2\left(f_n\right)\cdot \Delta F}}, \tag{15}$$

где k_i – коэффициент наличия формант звуков речи в i-й октавной полосе;

$L^2{}_n(f_n)$ – спектральная амплитуда сигнала на частоте f_n;

$f_{i.н}$, $f_{i.в}$ – нижняя и верхняя частоты i-й октавной полосы;

ΔF – шаг измерения амплитуды сигнала (ΔF=1 Гц).

Рассчитанные значения весовых коэффициентов k_i для октавных полос приведены в табл. 4.

Таблица 4. Значения весовых коэффициентов k_i для октавных полос

Октавная полоса частот, Гц	Значение коэффициента k_i для мужской речи	Значение коэффициента k_i для женской речи	Значения коэффициента k_i, приведенные в [4, 5]
90-180	0,02	0,01	0,01
180-355	0,06	0,06	0,03
355-710	0,07	0,06	0,12
710-1400	0,11	0,09	0,20
1400-2800	0,19	0,21	0,30
2800-5600	0,32	0,33	0,18
5600-11200	0,23	0,25	0,11

Как можно заметить, полученные в ходе экспериментальных исследований значения весовых коэффициентов k_i для октавных полос значительно отличаются от значений, приведенных в [4, 5]. Очевидно, данные отличия обусловлены различными методиками определения значений весовых коэффициентов k_i и отличительными особенностями дикторов.

На **пятом этапе** были определены нормированные спектры речи для различных уровней громкости.

Расчет нормированных значений уровней речевого сигнала в октавных полосах проводился по формулам:

$$L_c = \sqrt{\sum_{f_i = 90\ Гц}^{f_i = 11200\ Гц} L_c^2\left(f_i\right)\cdot \Delta F}; \tag{10}$$

$$\Delta L_i = 20\lg\frac{L_{c.i}}{L_c}, \tag{11}$$

где L_c – амплитуда речевого сигнала в семи октавных полосах;

$L_{c.i}$ – амплитуда речевого сигнала в i-й октавной полосе (рассчитывается по формуле (6));

ΔL_i – нормированное значение уровня речевого сигнала в i-й октавной полосе.

Результаты расчета нормированных значений уровней речевого сигнала в октавных полосах приведены в табл. 5.

Таблица 5. Нормированные значений уровней речевого сигнала в октавных полосах

Номер октавной полосы	Среднегеометрическая частота полосы f_{cp}, Гц	Нормированные значения уровней речевого сигнала в октавных полосах ΔL_i , дБ	
		Мужская речь	Женская речь
1	125	-7,96	-21,94
2	250	-4,15	-2,16
3	500	-5,51	-7,54
4	1000	-12,04	-13,98
5	2000	-16,48	-17,72
6	4000	-14,42	-14,89
7	8000	-12,77	-8,87

Используя нормированные значения уровней речевого сигнала в октавных полосах легко рассчитать уровни речевого сигнала в октавных полосах для его различных интегральных уровней громкости:

$$L_{c.i} = L_c + \Delta L_i \,, \quad (12)$$

где L_c – интегральный уровень речевого сигнала в семи октавных полосах, дБ.

В табл. 6 приведены результаты расчетов нормированных значений уровней речевого сигнала в октавных полосах для типовых интегральных уровней громкости.

Таблица 6. Типовой спектр речевых сигналов

Номер октавной полосы	Средне-геометрическая частота f_{cp}, Гц	Нормированные значения уровней речевого сигнала в октавных полосах, измеренные на расстоянии 1 м от источника речи, дБ							
		Тихая речь (64 дБ)		Речь средней громкости (70 дБ)		Громкая речь (76 дБ)		Очень громкая речь (84 дб)	
		М	Ж	М	Ж	М	Ж	М	Ж
1	125	56,04	42,06	62,04	48,06	68,04	54,06	76,04	62,06
2	250	59,85	61,84	65,85	67,84	71,85	73,84	79,85	81,84
3	500	58,49	56,46	64,49	62,46	70,49	68,46	78,49	76,46
4	1000	51,96	50,02	57,96	56,02	63,96	62,02	71,96	70,02
5	2000	47,52	46,28	53,52	52,28	59,52	58,28	67,52	66,28
6	4000	49,58	49,11	55,58	55,11	61,58	61,11	69,58	69,11
7	8000	51,23	55,13	57,23	61,13	63,23	67,13	71,23	75,13

Примечание: М – мужская речь, Ж – женская речь

Как можно заметить, октавные уровни громкости речи женщин и мужчин заметно отличаются в первой и седьмой октавных полосах, в остальных октавных полосах эти отличия незначительные.

На основе полученных в ходе экспериментальных исследований нормированных значений уровней речевого сигнала в октавных полосах и значений коэффициентов ΔA_i и κ_i был проведен расчет разборчивости речи в зависимости от отношения сигнал/шум для «белого шума».

Исходные данные для расчета приведены в табл. 7. Результаты расчетов приведены на рис. 10.

Анализ графиков, приведенных на рисунке 10, показывает, что при расчете разборчивости речи по формулам (1) – (5):

разборчивости мужской и женской речи несколько отличаются, что подтверждается экспериментальными исследованиями других авторов [20];

при одних и тех же отношениях сигнал/шум наибольшая разборчивость наблюдается для мужской речи, а разборчивость женской речи практически полностью совпадает с разборчивостью для смешенной речи [4, 5];

в диапазоне отношений сигнал/шум в пределах от − 15 дБ до − 5 дБ разборчивость мужской речи отличается от смешенной и женской речи на 5–7%.

Таблица 7. Основные характеристики речевого сигнала средней громкости

Номер октавной полосы	Средне-геометрическая частота $f_{ср}$, Гц	Значение формантного параметра речи в полосе ΔA_i, дБ			Весовой коэффициент полосы $к_i$			Значения уровней речевого сигнала в октавных полосах, дБ		
		М	Ж	С	М	Ж	С	М	Ж	С
1	125	20,98	22,36	25	0,00	0,00	0,01	62,04	48,06	53
2	250	12,71	16,91	18	0,03	0,03	0,03	65,85	67,84	66
3	500	7,48	7,13	14	0,09	0,10	0,12	64,49	62,46	66
4	1000	3,48	2,88	9	0,20	0,18	0,20	57,96	56,02	61
5	2000	4,3	4,26	6	0,31	0,31	0,30	53,52	52,28	56
6	4000	2,93	2,57	5	0,25	0,25	0,26	55,58	55,11	53
7	8000	6,91	9,08	4	0,11	0,13	0,07	57,23	61,13	49

Примечание: М – мужская речь, Ж – женская речь, С – смешенная речь [4, 5]

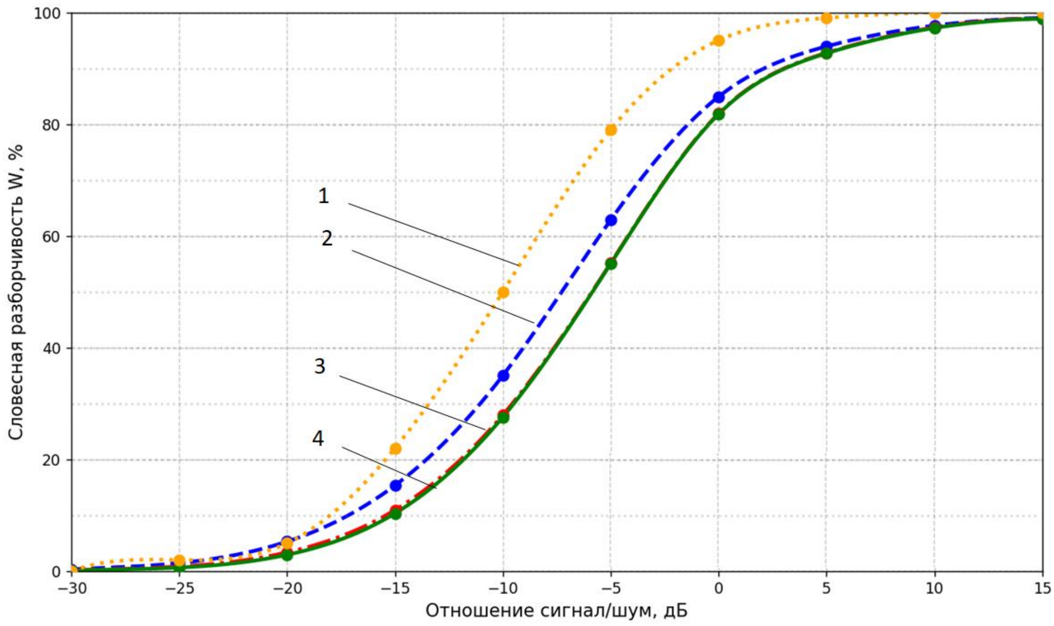


Рис. 10. Графики зависимости вероятности словесной разборчивости речи от отношения сигнал/шум («белый шум»): полученные по результатам артикуляционных испытаний (1) [19] и рассчитанные для мужской речи (2), для женской речи (3) и для смешенной речи (4) [4,5]

Отличие графиков разборчивости мужской и женской речи от графика разборчивости речи, полученного по результатам артикуляционных испытаний, возможно в первую очередь обусловлена тем фактом, что при артикуляционных испытаниях в качестве тестовых сигналов использовались не звуки речи, и артикуляционных таблицы слов и фраз.

Заключение

В результате цифровой обработки аудиозаписей звуков русской речи, записанных 9 дикторами (4 мужчинами и 5 женщинами) получены:

спектры звуков для мужской и женской речи;

спектры речи для мужских и женских голосов.

На основе обработки спектров звуков установлены частотные интервалы звуков и формант мужской и женской речи.

Предложен метод расчета коэффициентов встречаемости формант в октавных полосах k_i, на основе которого получены значения этих коэффициентов для мужской и женской речи.

На основе обработки спектров речи определены нормированные значения уровней речевого сигнала в октавных полосах для типовых интегральных уровней громкости для мужской и женской речи.

Предложен метод расчета формантных параметров речи в октавных полосах ΔA_i, на основе которого получены значения этих параметров для мужской и женской речи.

Некоторые отличия полученных результатов экспериментальных исследований от данных, представленных другими авторами, определяются, прежде всего, отличительными особенностями дикторов, а также разными методиками определения частотных интервалов звуков и формант, а также расчетов коэффициентов ΔA_i и κ_i.

На основе полученных в ходе экспериментальных исследований нормированных значений уровней речевого сигнала в октавных полосах и значений коэффициентов ΔA_i и κ_i был проведен расчет разборчивости речи в зависимости от отношения сигнал/шум для мужской и женской речи.

Анализ результатов расчета разборчивости речи показал, что

при одних и тех же отношениях сигнал/шум наибольшая разборчивость наблюдается для мужской речи, а разборчивость женской речи практически полностью совпадает с разборчивостью для смешенной речи;

в диапазоне отношений сигнал/шум в пределах от – 15 дБ до – 5 дБ разборчивость мужской речи отличается от смешенной и женской речи на 5–7%.

Полученные результаты могут быть использованы для определения направлений дальнейших исследований по корректировке методики расчета словесной разборчивости речи при оценке эффективности защиты акустической речевой информации от ее утечки по техническим каналам:

получение аудиозаписей звуков русской речи дикторов различного возраста и различного пола;

на основе обработки полученных аудиозаписей по методике, изложенной в статье:

определение нормированных значений уровней речевого сигнала в октавных полосах для мужской и женской речи. Использование в качестве скрываемого сигнала нормированных значений уровней речевого сигнала в октавных полосах не смешенной речи, как в действующей методике, а отдельно для мужской и женской речи;

определение частотных интервалов звуков и формант, а также расчет коэффициентов ΔA_i и κ_i для мужской и женской речи;

проведение экспериментальных исследований по распознаванию звуков речи в различных комбинациях октавных полос;

получение аналитических соотношений для расчета разборчивости речи с учетом частоты встречаемости звуков русской речи.

СПИСОК ЛИТЕРАТУРЫ

1. Покровский Н.Б. Расчет и измерение разборчивости речи. М.: Связьиздат, 1962. – 391 с.
2. Железняк В.К., Макаров Ю.К., Хорев А.А. Некоторые методические подходы к оценке эффективности защиты речевой информации. Специальная техника. М.: 2000, № 4, с. 39–45. URL: https://www.elibrary.ru/item.asp?id=29236464 (дата обращения: 07.07.2025).
3. Авдеев В.Б., Анищенко А.В., Петигин А.Ф., Трушин В.А. Оценка словесной разборчивости речи в условиях воздействия помех. Ч.1. Телекоммуникации. М.: ООО «Наука и технологии. 2024, № 11, с. 23–40. ISSN: 1684-2588. URL: https://www.elibrary.ru/item.asp?edn=eohvun (дата обращения: 00.00.2025). DOI: http://dx.doi.org/10.31044/1684-2588-2024-0-11-23-40.
4. Хорев А.А., Макаров Ю.К. К оценке эффективности защиты акустической (речевой) информации. Специальная техника. М.: 2000, № 5, с. 46–56. URL: https://aprodeus.narod.ru/Teaching/COBRAS/Art_2000_Makarov-Horev_K_ocenke_effect_zasch_acust.pdf (дата обращения: 07.07.2025).
5. Хорев А.А. Контроль защищенности речевой информации от ее утечки по техническим каналам. Специальная техника. М.: 2014, № 4, с. 43–63. URL: https://elibrary.ru/item.asp?edn=smkzpt (дата обращения: 07.07.2025).
6. Дидковский В.С., Дидковская М.В., Продеус А.Н Акустическая экспертиза каналов речевой коммуникации. Монография. Киев: Имекс-ЛТД, 2008. – 420 с.
7. Иванов А.В., Трушин В.А., Борботько Т.В. Возможность и целесообразность учета частотных зависимостей характеристик методики оценки защищенности речевой информации от утечки по техническим каналам. Научный вестник НГТУ.- Новосибирск: Изд-во НГТУ. 2017, № 4 (69), с. 85–94. ISSN 1814-1196. DOI: http://dx.doi.org/10.17212/1814-1196-2017-4-85-94.
8. Трушин В.А., Рева И.Л., Иванов А.В. Усовершенствование методики оценки разборчивости речи в задачах защиты информации. Ползуновский вестник. 2012, № 3/2, с. 238–241. ISSN: 2072-8921. URL: https://elibrary.ru/rnewbr (дата обращения: 07.07.2025).
9. Иванов А.В., Трушин В.А., Береснева А.В., Маркелова Г.В. Экспериментальные исследования защищенности речевой информации от утечки по техническим каналам с учетом эффекта форсирования речи. Актуальные проблемы электронного приборостроения (АПЭП-2014): материалы XII международной конференции, 2–4 октября 2014 г.: в 7 т. Новосибирск: Изд-во НГТУ. 2014, т. 3, с. 164–170. URL: https://elibrary.ru/tidnxx (дата обращения: 07.07.2025).
10. Иванов А.В., Трушин В.А. О модели речевого сигнала при оценке защищенности речевой информации от утечки по техническим каналам. Доклады ТУСУР. Томск. 2014, № 2(32), ч. 1, с. 87–90. ISSN: 1818-0442. URL: https://elibrary.ru/item.asp?id=21571467&ysclid=mgt0r1ujk9669688959 (дата обращения: 07.07.2025).
11. Иванов А.В., Рева И.Л., Трушин В.А., Тудэвдагва У. Корректировка методики оценки защищенности речевой информации от утечки по техническим каналам в условиях форсирования речи. Научный вестник НГТУ. Новосибирск: Изд-во НГТУ. 2014, т. 55, № 2, с. 183–189. ISSN 1814-1196. URL: https://www.elibrary.ru/item.asp?id=21978227 (дата обращения: 07.07.2025).
12. Евдокимова В.В., Кочаров Д.А., Скрелин П.А. Метод построения формантных картин для исследования фонетических характеристик гласных. Труды СПИИРАН. 2020, т. 19, № 2, с. 302–329. DOI: http://dx.doi.org/10.15622/sp.2020.19.2.3.
13. Козлачков С.Б., Дворянкин С.В., Бонч-Бруевич А.М. Ограничения формантной теории разборчивости речи в приложениях защиты речевой информации. Вопросы кибербезопасности. М.: НПО «Эшелон». 2016, № 5(18), с. 28–35. ISSN: 2311-3456 URL: https://elibrary.ru/xefzfr (дата обращения: 07.07.2025). DOI: http://dx.doi.org/10.21581/2311-3456-2016-5-28-35.
14. Yi Xu. Speech melody as articulatorily implemented communicative functions. Speech Communication. 2005, no. 46, p. 220–251. DOI: http://dx.doi.org/10.1016/j.specom.2005.02.014.
15. Santiago Barreda. Investigating the use of formant frequencies in listener judgments of speaker size. Journal of Phonetics. 2016, no. 55, p. 1–18. DOI: http://dx.doi.org/10.1016/j.wocn.2015.11.004.
16. Catford J.C. A Practical Introduction to Phonetics (2nd ed.). England, Oxford: Oxford University Press. 2001. – 229 p. ISBN 0-19-924635-1
17. Иванов А.В., Трушин В.А., Маркелова Г.В., Рева И.Л. Исследование спектра формант форсированной речи. Научный вестник Новосибирского государственного технического университета. Новосибирск, Новосибирский ГТУ. 2015, № 4(61), с. 63–73. ISSN: 1814-119. DOI: http://dx.doi.org/10.17212/1814-1196-2015-4-63-73.
18. Григорьев И.А., Казановский А.И. Методический подход к оценке эффективности защиты речевой информации. Вестник Воронежского государственного технического университета. 2010, т. 6, № 5,

с. 133–136. ISSN: 1729-6501eISSN: 2782-4624. URL: https://elibrary.ru/mjcisp (дата обращения: 07.07.2025).

19. Хорев А.А., Порсев И.С. Вероятностный метод обоснования показателей и критериев эффективности защиты речевой информации от ее утечки по техническим каналам. Вестник УрФО «Безопасность в информационной сфере». Челябинск, УрФО. 2024, № 2(52), с. 27–36. ISSN: 2225-5435. URL: https://elibrary.ru/nuswcv (дата обращения: 07.07.2025).
20. Иванов А.В., Рева, И.Л., Шемшетдинова Э.Э. Исследование влияния различий в спектрах речи на результат оценки разборчивости. Динамика систем, механизмов и машин. Омск, Омский ГТУ, государственный технический университет. 2017, т. 5, № 4, с. 65–70. ISSN: 2310-9793. URL https://elibrary.ru/zueiar (дата обращения: 07.07.2025). DOI: http://dx.doi.org/10.25206/2310-9793-2017-5-4-65-70.

REFERENCES

[1] Pokrovsky N.B. Calculation and measurement of speech intelligibility. M.: Svyazizdat, 1962. – 391 p.
[2] Zheleznyak V.K., Makarov Yu.K., Horev A.A. Some methodological approaches to evaluating the effectiveness of speech information protection. M.: 2000, no. 4, p. 39–45. URL: https://www.elibrary.ru/item.asp?id=29236464 (accessed: 07.07.2025) (in Russian).
[3] Avdeev V.B., Anishchenko A.V., Petigin A.F., Trushin V.A. Evaluation of verbal intelligibility of speech under the influence of interference. Part 2. Telecommunications. M.: OOO "Science and Technology. 2025, no. 1, p. 29–35 URL: https://www.elibrary.ru/item.asp?edn=eohvun (accessed: 07.07.2025). DOI: http://dx.doi.org/10.31044/1684-2588-2024-0-11-23-40 (in Russian).
[4] Horev A.A., Makarov Yu.K. To evaluate the effectiveness of acoustic (speech) information protection. Special technique. M.: 2000, no. 5, p. 46–56. URL: https://aprodeus.narod.ru/Teaching/COBRAS/Art_2000_Makarov-Horev_K_ocenke_effect_zasch_acust.pdf (accessed: 07.07.2025) (in Russian).
[5] Horev A.A. Control of voice information leakage prevention via technical channels. Special equipment. M.: 2014, № 4, p. 43–63. URL: https://elibrary.ru/item.asp?edn=smkzpt (accessed: 07.07.2025) (in Russian).
[6] Didkovsky V.S., Didkovskaya M.V., Prodeus A.N. Acoustic examination of speech communication channels. Monograph. Kiev: Imex-LTD., 2008. – 420 p. (in Russian).
[7] Ivanov A.V., Trushin V.A., Borbotko T.V. Possibility and expediency of accounting frequency dependences of characteristics of the valuation method of voice information protection from leakage through technical information leakage channels. Scientific Bulletin of NSTU. Novosibirsk: NSTU Publishing House. 2017. № 4(69), p. 85–94. ISSN 1814-1196. DOI: http://dx.doi.org/10.17212/1814-1196-2017-4-85-94 (in Russian).
[8] Trushin V.A., Reva I.L., Ivanov A.V. Improving the methodology for assessing speech intelligibility in information security tasks. Polzunovsky Vestnik. 2012, no. 3/2, p. 238–241. ISSN: 2072-8921. URL: https://elibrary.ru/rnewbr (accessed: 07.07.2025) (in Russian).
[9] Ivanov A.V., Trushin V.A., Beresneva A.V., Markelova G.V. Experimental studies of the security of speech information from leakage through technical channels, taking into account the effect of speech forcing. Actual problems of electronic instrumentation (APEP-2014): proceedings of the XII International Conference, October 2-4, 2014: in 7 v. Novosibirsk: NSTU Publishing House.2014, v. 3, p. 164-170. URL: https://elibrary.ru/tidnxx (accessed: 07.07.2025) (in Russian).
[10] Ivanov A.V., Trushin V. About model of a speech signal at an assessment of security of speech information by leaking from technical channels. TUSUR reports. Tomsk. 2014, № 2(32), Part 1, p. 87–90. ISSN: 1818-0442. URL: https://elibrary.ru/item.asp?id=21571467&ysclid=mgt0r1ujk9669688959 (accessed: 07.07.2025) (in Russian).
[11] Ivanov A.V., Reva I.L., Trushin V.A., Tudevdagva U. Corrected methods for the assessment of audio information security against leakage through engineering channels for forced speech. Scientific bulletin of NSTU. Novosibirsk: NSTU Publishing House. 2014, v. 55, no. 2, p. 183–189. ISSN 1814-1196. URL: https://www.elibrary.ru/item.asp?id=21978227 (accessed: 07.07.2025) (in Russian).
[12] Evdokimova V.V.,. Kocharov D.A, Skrelin P.A. Method of constructing formant patterns for studying phonetic characteristics of vowels. Proceedings of SPIIRAN, 2020, v. 19, no. 2, p. 302-329. DOI: http://dx.doi.org/10.15622/sp.2020.19.2.3.
[13] Kozlachkov S.B., Dvoryankin S.V., Bonch-Bruevich A.M. Restrictions formant theory of speech intelligibility in speech applications, information security. Cybersecurity issues. Moscow: NGO "Echelon". 2016, no. 5(18), p. 28–35. ISSN: 2311-3456 URL: https://elibrary.ru/xefzfr (accessed: 07.07.2025). DOI: http://dx.doi.org/10.21581/2311-3456-2016-5-28-35 (in Russian).
[14] Yi Xu. Speech melody as articulatorily implemented communicative functions. Speech Communication. 2005, no. 46, p. 220–251. DOI: http://dx.doi.org/10.1016/j.specom.2005.02.014.

[15] Santiago Barreda. Investigating the use of formant frequencies in listener judgments of speaker size. Journal of Phonetics. 2016, no. 55, p. 1–18. DOI: http://dx.doi.org/10.1016/j.wocn.2015.11.004.
[16] Catford J.C. A Practical Introduction to Phonetics (2nd ed.). England, Oxford: Oxford University Press. 2001. – 229 p. ISBN 0-19-924635-1.
[17] Ivanov A.V., Trushin V.A., Markelova G.V., Reva I.L. Research on the spectrum of forced speech formants. Scientific Bulletin of Novosibirsk State Technical University. Novosibirsk, Novosibirsk State Technical University. 2015, no. 4(61), p. 63–73. ISSN: 1814-119. DOI: http://dx.doi.org/10.17212/1814-1196-2015-4-63-73 (in Russian).
[18] Grigoriev I.A., Kazanovsky A.I. Methodical approach to efficiency evaluation of voice data protection. Bulletin of the Voronezh State Technical University. 2010, v. 6, no. 5, p. 133–136. ISSN: 1729-6501eISSN: 2782-4624. URL: https://elibrary.ru/mjcisp (accessed: 07.07.2025) (in Russian).
[19] Horev A.A., Porsev I.S. A probabilistic method for substantiating indicatorsand criteria for the effectiveness of protecting speechinformation from its leakage. Bulletin of the Ural Federal District "Security in the information sphere". Chelyabinsk, Ural Federal District. 2024, no. 2(52), p. 27–36. ISSN: 2225-5435. URL: https://elibrary.ru/nuswcv (accessed: 07.07.2025) (in Russian).
[20] Ivanov A.V., Reva, I.L., Shemshetdinova E.E. Investigation of the effect of differences in speech spectra on the result of the assessment of intelligibility. Dynamics of systems, mechanisms and machines. Omsk, Omsk State Technical University, Omsk State Technical University. 2017, v. 5, no. 4, p. 65–70. ISSN: 2310-9793. URL https://elibrary.ru/zueiar (accessed: 07.07.2025). DOI: http://dx.doi.org/10.25206/2310-9793-2017-5-4-65-70 (in Russian).

Статья поступила в редакцию 10.07.2025; одобрена после рецензирования 10.08.2025; принята к публикации 01.09.2025
The article was submitted 10.07.2025; approved after reviewing 10.08.2025; accepted for publication 01.09.2025