

# АНАЛИЗ МЕТОДОВ ИССЛЕДОВАНИЯ ДЛЯ ОЦЕНКИ КАЧЕСТВА СООБЩЕНИЙ УСТНОЙ КАЗАХСКОЙ РЕЧИ

## ANALYSIS OF RESEARCH METHODS TO ASSESS THE OUTPUT QUALITY OF THE ORAL SPEECH OF THE KAZAKH

**O. Romashkova  
Y. Bakhtiyarova  
A. Kargulova  
B. Anykbaev**

*Summary.* The article deals with the analysis of research methods to assess the quality of oral Kazakh speech messages. The methods of speech quality assessment are described in detail. The algorithm with alternating priority with alternating priority through the counting of messages of oral Kazakh speech is the best of the three implementations of the stssp CONSIDERED in the work. This significant advantage can be explained by the significant impact on the SNR of losses of more than one reference in a row.

*Keywords:* voice quality, efficiency, speech information, high recognition.

**Ромашкова Оксана Николаевна**

*Д.т.н., профессор, ФГБОУ ВО «Московский государственный лингвистический университет (МГЛУ)» г. Москва, Россия  
ox-rom@yandex.ru*

**Бахтиярова Елена Ажибековна**

*К.т.н., доцент, Казахская Академия транспорта и коммуникаций им. М. Тынышпаева (КазАТК) г. Алматы, Казахстан  
baelag@mail.ru*

**Каргулова Алия Нурымовна**

*Докторант, Казахская Академия транспорта и коммуникаций им. М. Тынышпаева (КазАТК) г. Алматы, Казахстан  
kargulova84@mail.ru*

**Аныкбаев Бекзат Есжиярович**

*Докторант, Казахская Академия транспорта и коммуникаций им. М. Тынышпаева (КазАТК) г. Алматы, Казахстан  
anykbaiev@mail.ru*

*Аннотация.* В статье рассматриваются вопросы анализа методов исследования для оценки качества сообщений устной казахской речи. Подробно описаны методы оценки качества речи. Алгоритм с чередующимся приоритетом с чередованием приоритета через отсчет сообщений устной казахской речи является лучшим из трех рассматриваемых в работе реализаций СЦСП. Такое значительное преимущество можно объяснить значительным влиянием на ОСШД потерь более одного отсчета подряд.

*Ключевые слова:* речевое сообщение, методы оценки качества, отношение сигнал — шум, критерии оценки качества.

### Введение

**В** настоящее время широкое развитие получили цифровые сети, в которых все виды информации, в том числе речь, передаются в цифровом виде. При реализации цифровых преобразований речевых сигналов возникают специфические искажения, которые влияют на качество речи.

В соответствии с Рекомендацией Р.48 МККТТ эффективная полоса пропускания звукового тракта речевого (тонального) сигнала лежит в области 300–3400 Гц. Одним из критериев качества речи является разборчивость, т.е. смысловая понятность переданной информации слушателю [1,2,3].

### Постановка задачи

Разборчивость — объективная количественная величина, характеризующая способность тракта телефонной связи передавать содержащуюся в речи смысловую информацию в данных конкретных условиях акустической среды.

Под мерой разборчивости следует понимать отношение числа правильно принятых элементов речи (звуков, слогов, слов, предложений) к достаточно большому общему числу переданных, выраженное в процентах или долях единицы. В соответствии с этим различают разборчивость звуков, слогов, слов и фраз [4,5].

Таблица 1.

Класс качества (оценка разборчивости)	Характеристики класса качества	Процент совпадений с табличным материалом
Высший (отличная)	Понимание передаваемой речи без малейшего напряжения внимания	> 80
Первый (Хорошая)	Понимание передаваемой речи без затруднений	56–80
Второй (Удовлетворительная)	Понимание передаваемой речи с некоторым напряжением внимания, без переспросов и повторений	41–55
Третий (Предельно допустимая)	Понимание передаваемой речи с некоторым напряжением внимания, редкими переспросами и повторениями	25–40
Четвертый (Неудовлетворительная)	Понимание передаваемой речи с большим напряжением внимания, частыми переспросами и повторениями	< 25

Таблица 2.

Характеристика качества речи	Баллы
Естественность звучания речи. Высокая узнаваемость. Полное отсутствие помех и искажений	4,6–5,0
Естественность звучания речи. Высокая узнаваемость. Отдельные малозаметные искажения или помехи	4,0–4,5
Естественность звучания речи. Высокая узнаваемость. Слабое постоянное присутствие отдельных видов искажений или помех	3,5–3,9
Незначительное нарушение естественности и узнаваемости. Заметное присутствие отдельных искажений или помех	3,0–3,4
Заметное нарушение естественности и ухудшение узнаваемости, присутствие нескольких видов искажений (картавость, гнусавость и др.) или помех	2,5–2,9
Существенное искажение естественности и ухудшение узнаваемости. Постоянное присутствие искажений типа картавость, гнусавость и др. или помех	1,7–2,4
Сильные искажения типа картавость, гнусавость и др. Механический голос. Наблюдается потеря естественности и узнаваемости	< 1,7

Для проведения испытаний по оценке разборчивости используются тональный и артикуляционные методы, в которых экспертные оценки выносятся экспертами. Основным является статистическая достоверность их результатов, т.е. число участвующих в испытаниях экспертов и количество прослушиваний должно быть достаточными.

Тональный метод основан на способности человеческого уха достаточно точно улавливать минимальный порог уровня громкости.

Речевой сигнал воспроизводится в виде отдельных тональных полосок. При прослушивании их уровень уменьшается до минимально воспринимаемой громкости. Полученные значения затухания с помощью справочных таблиц, используемых при расчетах разборчивости речи, пересчитываются и дают численное значение этого параметра [1,2,3].

Измерение разборчивости методом артикуляции производится бригадой операторов путем передачи и приема по телефонному тракту серии артикуляционных таблиц, составленных из элементов речи (звуков, слогов, слов, фраз). В зависимости от используемых артикуляционных таблиц (звуковые, слоговые, словарные, фразовые) измеряют следующие виды разборчивости: звуковую, слоговую, словесную и фразовую. В результате таких измерений полученные величины разборчивости являются оценкой качества телефонного тракта [4,5].

Из всех типов артикуляционных таблиц на практике нашли применение слоговые и словесные. При этом слоговые артикуляционные таблицы можно рассматривать как основные, так как на практике в основном измеряется именно слоговая разборчивость. Оценочные характеристики для слоговых артикуляционных испытаний представлены в табл. 1.

С появлением цифровых систем записи и сжатия речи выявляются специфические шумы и искажения, которые ухудшают качество речи. Для их оценки введены разные виды артикуляционных испытаний. Наиболее точными для оценки заметности искажений, вносимых кодеком, являются парные сравнения испытательных фраз. В соответствии с Рекомендацией Р.48 МККТ качество речи испытуемого тракта оценивают с эталонным трактом путем сравнения, в качестве которого используют стандартный телефонный тракт.

Качество речи оценивают по контрольным фразам, приведенным в ГОСТ Р50840–95. Каждую контрольную фразу передают один раз через оцениваемый тракт, другой — через эталонный тракт [3].

Качество речи по методу парных сравнений оценивают по 5-бальной системе с оценочным шагом в 0,1 балла. В табл. 2 приводится соответствие между качеством речи речевого тракта и оценкой в баллах для метода парных сравнений.

Классификация критериев оценки качества речевого сигнала приведена в [6]. Согласно этой классификации, выделяются показатели точности восстановления отдельной реализации и множества реализаций сообщения. Чаще используется вторая группа критериев [7,8], причем предпочтение отдается показателям среднеквадратического приближения [9].

В качестве критерия точности воспроизведения речевого сигнала используем нормированный показатель мощности — средний квадрат (усреднение по множеству реализаций) ошибки, шума воспроизведения (текущей разности между исходным и восстановленным сообщением), усредненный по времени и приведенный к дисперсии сообщения [1,10,11]

$$\gamma^{-2} = \frac{\sigma_w^2}{\sigma_c^2}, \quad (1)$$

где  $\sigma_w^2$  — мощность шума,

$\sigma_c^2$  — мощность сигнала.

Отношение мощности сигнала к мощности шумов [3,10,30,11]:

$$ОСШ = 10 \lg \left( \frac{1}{\gamma^2} \right), \text{ дБ.} \quad (2)$$

Шумы квантования по уровню (отношение мощности сигнала к мощности шумов — ОСШК) относятся к шумам, коррелированным с речевым сообщением, а шумы дискретизации и восстановления (отношение мощности сигнала к мощности шумов — ОСШД) речевого сообщением по его отсчетам — не коррелированным [12,13].

Для анализа качества воспроизведения казахской речи при различных реализациях статистического уплотнения необходимо распределение вероятностей  $P(X)$  случайной величины  $X$  количества подряд отбракованных отсчетов в рекуррентном потоке на входе интерполирующего фильтра дополнительного комплекта [2, 3, 13].

В качестве входных данных используются два файла типа «RAW» (побайтный поток отсчетов ИКМ с параметрами, определенными рекомендацией G.711 ITU-T), полученные путем конвертирования реальных аудиозаписей телефонных разговоров двух пар абонентов. С целью определения статистики коллизий, «столкновения отсчетов» основного и дополнительного комплекта каждый отсчет основного комплекта преобразован к виду «активный/пассивный».

В работах [15, 3] определена достаточная (для имитационного моделирования) продолжительность реализации речевого сообщения ( $L$  в отсчетах и  $t$  в минутах) с использованием развития оценки вероятности  $\theta$  отбраковки отсчета дополнительного комплекта и относительной погрешности  $\delta\theta, \%$  ее определения.

Следует отметить, что значение оценки  $\theta$  практически стабилизируется (при выполнении условия  $|\delta\theta| < 2\%$ ) при  $t \geq 20$  мин. С запасом выбираем продолжительности  $L=14\ 400\ 000$  отсчетов (30 мин) для реализаций речевых сообщений основного и дополнительного комплектов.

В работах [15,16] получены результаты моделирования — распределение вероятностей случайной величины  $X$  количества подряд отбракованных отсчетов речевого сообщения основного  $p_1(X)$  и дополнительного  $p_2(X)$  комплекта (термины «основной» и «дополнительный» здесь теряют смысл, поскольку оба комплекта находятся в равноценных условиях).

Вероятность отбраковки отсчета дополнительного комплекта при чередовании приоритета через цикл передачи  $E_1$  равна 0,027, при чередовании через отсчет равна 0,127, а при фиксированном приоритете — 0,034.

Также аналогично в работах [15,16] получены распределения вероятностей  $p(X)$  для чередующегося приоритета.

На основании полученных экспериментальных распределений вероятностей  $p(X)$  случайной величины  $X$  количества подряд отбракованных отсчетов речевого сообщения можно оценить качество его восстановления, используя методику [1,5].

Отношение мощности сигнала к мощности шумов, вызванных процессами восстановления (ОСШД) [1,5] имеет вид (3)

где  $\Omega' = 2 \cdot \pi \cdot F'_{cp}$ ;  $\Omega = 2 \cdot \pi \cdot F_{cp}$ ;  $R = \alpha$ ;  $B = 2 \cdot \pi \cdot f_0$ ;  
 $F'_{cp}$  — частота среза идеального ФНЧ, используемого при восстановлении;  
 $F_{cp}$  — частота среза идеального ФНЧ, используемого при предварительном ограничении спектра;

$$C = 2 \cdot \pi \cdot f_{д} + B; D = 2 \cdot \pi \cdot f_{д} - B;$$

$$Q = \langle (x+1)(F_{cp} + F'_{cp}) / f_{д} \rangle; \quad ;$$

$\langle \rangle$  — операция округления до большего целого значения.

В формуле (2) использована аппроксимация нормированной в области положительных частот усредненной спектральной плотности мощности речи [1,5,17] (4)

с параметрами  $\alpha = 1000$  Гц и  $f_0 = 400$  Гц [18] для сообщения устной речи на русском языке (в работе [1] приведены значения этих параметров для речевых сообщений на английском, испанском и вьетнамском языках).

В работе [1] показана высокая степень соответствия этой аппроксимации спектральной плотности мощности речи, предложенной в [1,6] на основании рекомендаций P.51 ITU-T

$$S(f) = -465,75 + 465,44 \lg(f) - 157,75 \lg^2(f) + 16,71 \lg^3(f). \quad (5)$$

В работах [2,19] с использованием результатов экспериментальных исследований уточнены параметры аппроксимации (2) нормированной СПМ телефонного

сигнала на русском языке:  $f_0 = 390$  Гц,  $\alpha = 492$  Гц и  $M = 0,207$ . Для телефонного сигнала на казахском языке получаем  $f_0 = 300$  Гц;  $\alpha = 427$  Гц и  $M = 0,477$ .

В случае использования алгоритма с фиксированным приоритетом с использованием исходных данных  $F_{cp} = F'_{cp} = 3400$  Гц,  $f_{д} = 8000$  Гц [1,2,8] при использовании распределения  $p(X)$  оценка составит ОСШД=34,79 дБ. А при использовании распределения  $p(X)$ , представленного в [17] — ОСШД=48,84 дБ соответственно.

Следовательно, при известной оценке вероятности  $\theta$  потери пакета использование геометрического распределения приводит к завышенной оценке качества восстановления речевого сообщения дополнительного комплекта.

В случае использования алгоритма с чередующимся приоритетом для сообщения устной казахской речи ОСШД=37,53 дБ при чередовании приоритета через цикл передачи Е1 и ОСШД=43,86 дБ при чередовании приоритета через отсчет.

Таким образом, также получено подтверждение преимуществу алгоритма с чередующимся приоритетом сообщения устной казахской речи (ОСШД при использовании этого алгоритма равно 37,53 дБ, а при фиксированном приоритете его значение составляет 34,79 дБ).

Однако, для чередующегося приоритета при чередовании приоритета через отсчет значение ОСШД составляет 43,86 дБ, что на 6,33 дБ больше чем при чере-

$$ОСШД = 10 \cdot \lg \left\{ \sum_x p(X) \cdot \frac{1+M}{2 \cdot \pi} \cdot \sum_{i=1}^Q \left[ \begin{aligned} & \arctg\left(\frac{\Omega'-C}{R}\right) + \arctg\left(\frac{\Omega'+D}{R}\right) - \\ & - \arctg\left(\frac{2 \cdot \pi \cdot i \cdot f_{д} / (x+1) - \Omega - C}{R}\right) - \\ & - \arctg\left(\frac{2 \cdot \pi \cdot i \cdot f_{д} / (x+1) - \Omega + D}{R}\right) \end{aligned} \right] \right\}^{-1}, \quad (3)$$

$$M = 1 - \frac{1}{\pi} \cdot \arctg\left(\frac{\Omega - B}{R}\right) - \frac{1}{\pi} \cdot \arctg\left(\frac{\Omega + B}{R}\right),$$

$$S_{\lambda}(f) = (1+M) \cdot S_{\lambda}^H(f) = 2 \cdot (1+M) \cdot \alpha \cdot \left[ \frac{1}{\alpha^2 + 4\pi^2(f-f_0)^2} + \frac{1}{\alpha^2 + 4\pi^2(f+f_0)^2} \right],$$

$$M = 1 - \int_{F_H}^{F_B} S_{\lambda}^H(f) df = 1 - \frac{1}{\pi} \left[ \arctg\left(\frac{2\pi F_B - \omega_0}{\alpha}\right) - \arctg\left(\frac{2\pi F_H - \omega_0}{\alpha}\right) + \right. \\ \left. + \arctg\left(\frac{2\pi F_B + \omega_0}{\alpha}\right) - \arctg\left(\frac{2\pi F_H + \omega_0}{\alpha}\right) \right] \quad (4)$$

довании приоритета через цикл передачи E1 и на 9,07 дБ больше чем при фиксированном приоритете.

## Вывод

В представленной работе проведен анализ методов исследования для оценки качества сообщений устной казахской речи. Подробно описаны методы оценки качества речи. Алгоритм с чередующимся приоритетом с чередованием приоритета через отсчет сообщений устной казахской речи является лучшим из трех рассматриваемых в работе реализаций СЦСП. Такое значительное преимущество можно объяснить значительным влиянием на ОСШД потерь более одного отсчета подряд.

Для стохастической цифровой системы передачи (СЦСП) со статистическим уплотнением при уточнении аппроксимации СПМ сообщения устной казахской речи [5] получены значения ОСШД, отличающиеся от значений ОСШД сообщения устной русской речи на 5,16 дБ для фиксированного приоритета, на 7,19 дБ при чередовании приоритета через цикл передачи E1 и на 1,34 дБ при чередовании приоритета через отсчет. Также получены значения ОСШД, отличающиеся от представленного в работе /20/ на 8,59 дБ для фиксированного приоритета, на 10,51 дБ при чередовании приоритета через цикл передачи E1 и на 4,47 дБ при чередовании приоритета через отсчет.

## ЛИТЕРАТУРА

1. Горелов Г. В., Ромашкова О. Н., Чан Туан Ань. Качество управления речевым трафиком в телекоммуникационных сетях. — М.: Радио и связь, 2001.
2. Бахтиярова Е. А. Вероятностные характеристики сообщения устной казахской речи. Диссертация на соискание ученой степени канд. техн. наук. М. МИИТ. 2007. — 113 с.
3. Ефимов А. П., Рысин Ю. С., Свобода Д. Г. Акустические измерения, оценки, контроль. Учебное пособие. — М.: МТУСИ, 2005. — 113 с.
4. Быков С. Ф., Журавлев В. И., Шалимов И. А. Цифровая телефония: Учеб. пособие для вузов. — М.: Радио и связь, 2003. — 144 с.
5. Лукова О. Н. Анализ качества стохастической цифровой передачи речевой информации (Методика и ее использование при разработке информационных систем). Диссертация на соискание ученой степени канд. техн. наук. М. МИИТ. 1994. — 149 с.
6. Горелов Г. В. Нерегулярная дискретизация сигналов. М.: Радио и связь, 1982. — 256 с.
7. Новоселов О. Н., Фомин А. Ф. Основы теории и расчета информационно-измерительных систем. М.: Машиностроение, 1991. — 336 с.
8. Фомин А. Ф. Помехоустойчивость систем передачи непрерывных сообщений. — М.: Сов. радио, 1975. — 352 с.
9. Ольховский Ю. Б., Новоселов О. Н., Мановцев А. П. Сжатие данных при телеизмерениях. — М.: Сов. радио, 1971. — 304 с.
10. Горелов Г. В. Алгоритм Burst aloha в статистическом уплотнении аудиоинформационных сигналов // ЛОКСЕТЬ-90. Рига. 1990. — с. 38–41.
11. Чан Туан Ань. Разработка и исследование методов оценивания качества пакетной передачи вьетнамской речи. Диссертация на соискание ученой степени канд. техн. наук. М. МИИТ. 1995.
12. Gorelov G., Romashkova O. Influence of russian, spanish and vietnamese speech characteristics on digital information transmission quality. Proceedings of the IEEE International Symposium on Industrial Electronics. ISIE'96. Warsaw Vol 1. — p.311–313.
13. Горелов Г. В. Качество воспроизведения речи в стохастических цифровых системах передачи // Автоматика и вычислительная техника. 1993. -N5. — с. 40–43.
14. Горелов Г. В., Бахтиярова Е. А., Карпов А. В., Вероятностные характеристики речевого трафика. Мир транспорта 2007, № 1, с. 22–25.
15. Бахтиярова Е. А. Качество воспроизведения устной казахской речи в стохастических цифровых системах передачи при ее статистическом уплотнении. Вестник КазАТК, № 4, 2008. — с. 285–290.
16. Горелов Г. В., Толмачев П. Н., Бахтиярова Е. А. К оценке качества восстановления речевого сообщения при статистическом уплотнении первичной цифровой системы передачи // ВКСС Connect. — 2006. — № 2. — с. 94–97.
17. Горелов Г. В., Казанский Н. А., Лукова О. Н. Методика оценки качества пакетной передачи речи в интегральных цифровых сетях. — М.: Электросвязь, №9, 1992. — с. 31–33.
18. Горелов Г. В., Кочнов Л. Л., Пчелинцев А. В., Пчелинцева Н. М. К оценке качества полосного вокодера. Обработка сигналов в системах телефонной связи. 1998. — с. 71–74.
19. Горелов Г. В., Бахтиярова Е. А., Толмачев П. Н., Ширинский Д. А. К оценке энергетического спектра речевого сообщения // Информационно-управляющие системы на железнодорожном транспорте. Украина — Харьков.-2007, № 4. — с. 64–68.
20. Толмачев П. Н. Методы имитационного моделирования в определении качества стохастической цифровой передачи речевой информации. Диссертация на соискание ученой степени канд. техн. наук. М. МИИТ. 2006.

© Ромашкова Оксана Николаевна (ox-rom@yandex.ru), Бахтиярова Елена Ажибековна (baelag@mail.ru), Каргулова Алия Нурымовна (kargulova84@mail.ru), Аныкбаев Бекзат Есжиярович (anykbaiev@mail.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»