

СОЗДАНИЕ РЕЧЕПОДОБНОЙ ПОМЕХИ НА ОСНОВЕ СВЯЗНЫХ ТЕКСТОВ

В.А. Трушин, Д.Е. Попов, М.А. Кунгуров, Д.Л. Марченко
Новосибирский государственный технический университет,
г. Новосибирск

ВВЕДЕНИЕ

На сегодняшний день информационная сфера оказывает очень большое влияние на развитие нашего общества. Но, несмотря на широкое развитие средств цифровой передачи информации, речевая акустическая информация остается актуальной, потому что общение, переговоры являются основным процессом развития контактов между людьми из-за потребностей в совместной деятельности. Исследования показали, что наиболее эффективной активной мерой является речеподобная помеха, формируемая из речевых сигналов [1].

В данной работе рассматривается возможность создания речеподобной помехи на основе связанных текстов, используя программный синтезатор. Проведены эксперименты по выявлению влияния времени усреднения записей и влияния различных дикторов на форму спектра речи. Проведены артикуляционные испытания на основе полученного речевого хора и сформулированы выводы для дальнейших исследований создания речевой помехи.

ПОСТАНОВКА ЭКСПЕРИМЕНТА

Влияние времени усреднения записи речеподобной помехи на спектр речи. За основу взяты фрагменты произведения М.А. Булгакова «Мастер и Маргарита». Программа синтезатор - Vocalizer и программа для обработки аудио и получения спектров – AdobeAudition 3.0. Выбор данной программы охарактеризован возможностью поддержки таких параметров как: настройки частоты дискретизации, установка скорости речи, добавление других голосов и сохранение записи в файл - которых нет в других аналогичных программах (Acapela, RHVoice, Espeak, Festival, Говорилка и др.) Запись аудиофайлов производим с частотой дискретизации 44кГц, 16Bit Mono. Программа для обработки аудио и получения спектров - AdobeAudition 3.0

Создание речеподобной помехи будем производить на среднем уровне речи $L_s=70$ дБ. Данные интегральные уровни речи представлены в таб.1 по октавным полосам частот. Для удобства сопоставления результатов все дальнейшие данные и расчеты приведены для данного интегрального уровня речи в 70 дБ.

Таблица 1. Типовые интегральные уровни речи для уровня в 70дБ по октавам

Частота, Гц	Интегральные уровни речи $L_s=70$ дБ
125	53
250	66
500	66
1000	61
2000	56
4000	53
8000	49

Для сравнения спектров с разным временем усреднения были взяты различные отрезки по времени (1, 5, 15 и 30 минут) и произведены записи звуковых файлов одним диктором. Полученные результаты интегральных уровней спектра речи по октавным полосам представлены в таблице 2.

Таблица 2. Влияние времени усреднения на спектр речи

Октавные полосы, Гц	125	250	500	1000	2000	4000	8000
1 мин	54,21	65,76	67,23	55,40	48,74	50,38	51,76
5 мин	54,71	66,06	67,04	54,09	48,25	50,59	52,46
15 мин	55,00	66,09	66,99	54,18	47,94	50,36	52,50
30 мин	54,93	66,14	66,97	54,27	47,90	50,06	52,24

В результатах наблюдаются различия уровней спектра речи от времени усреднения не больше 1,5 дБ. Исходя из полученных результатов, можно сделать вывод, что различие по времени

усреднения записей не имеет влияния на форму спектра русской речи.

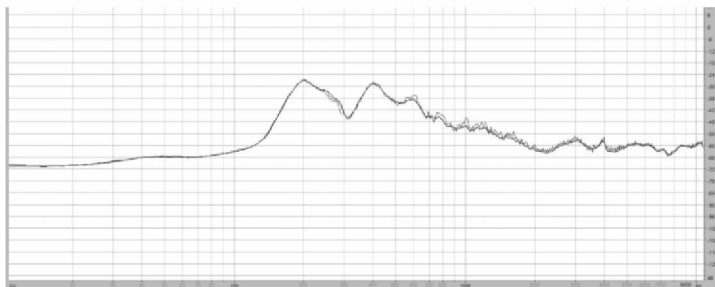


Рис. 1. График спектра речи для различного времени усреднения

Создание речевого хора.

Для создания речевого хора было проведено сравнение интегральных уровней спектров речи пяти дикторов для семи октавных полос. Для каждого диктора был выбран участок длительностью в одну минуту и записан с помощью программы-синтезатора Vocalizer с одинаковой скоростью произношения и громкостью (небольшие расхождения в спектрах были выровнены по общему интегральному уровню).

Таблица 3. Результаты спектров речи для различных дикторов в 7 октавных полосах

Октавные полосы, Гц	125	250	500	1000	2000	4000	8000
Алена	55,66	66,00	66,94	55,43	48,23	50,98	52,17
Юрий	65,70	66,00	62,23	53,10	52,64	53,09	46,24
Милена	56,59	67,12	64,79	59,58	51,05	51,13	53,28
Катя	64,75	66,66	62,57	54,79	51,61	49,89	43,84
Николай	63,37	64,62	64,43	61,54	57,92	48,68	45,96
Усредненная речь	53,00	66,00	66,00	61,00	56,00	53,00	49,00

Способ создания речевого хора. В соответствии с исследованиями И. Алдошиной [3] для создания речевого хора было решено использовать три голоса: наиболее близкий к усредненному спектру речи на низких, средних и высоких частотах. Из полученных данных видно, что на низких частотах ближе всего приближен к усредненному спектру речи спектр Алены, на средних частотах подходит спектр Николая, на высоких частотах - Юрия. Для каждого диктора был выбран участок длительностью в одну минуту и записан с помощью синтезатора Vocalizer с одинаковой скоростью произношения и громкостью (небольшие расхождения в спектрах были выровнены по общему интегральному уровню). В программе AdobeAudition 3.0 была сформирована речеподобная помеха типа речевой хор методом наложения нескольких звуковых файлов со сдвигом по времени. На выходе был записан звуковой файл речевого хора в формате wav.

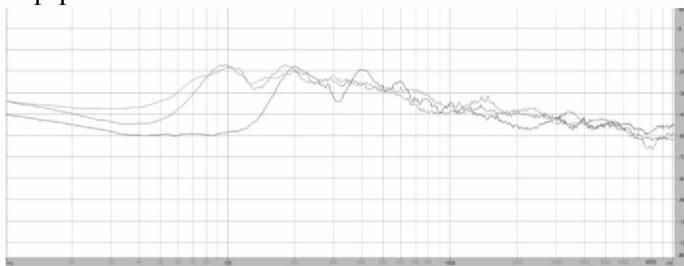


Рис. 2. График спектров трех дикторов, использованных для создания речевого хора

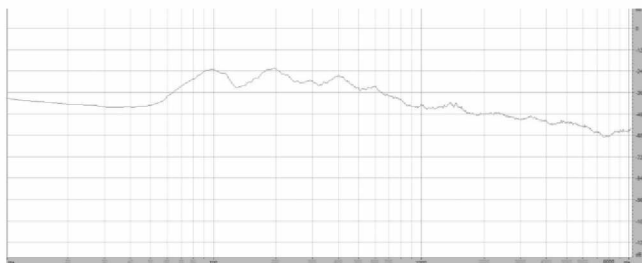


Рис. 3. График спектра полученного речевого хора

Сравнение таблиц 3 и 4 показывает, что речевой хор слабо приближен к значениям усредненного спектра речи на низких

частотах, чем женский голос. На отдельных частотах определенные голоса дают лучшее приближение, однако на всей полосе частот речевой хор дает максимально близкие значение, чем отдельные голоса. Таким образом, спектр речевого хора является самым приближенным к усредненному спектру русской речи.

Таблица 4. Уровни спектров речи речевого хора по октавным полосам

Октавные полосы, Гц	125	250	500	1000	2000	4000	8000
Интегральные уровни речевого хора, дБ	63,76	65,58	64,61	57,22	54,69	52,74	50,51
Интегральные уровни усредненного спектра русской речи, дБ	53,00	66,00	66,00	61,00	56,00	53,00	49,00

Артикуляционные испытания. Артикуляционные испытания – это процесс субъективной оценки разборчивости речи, заключающийся в наложении на исследуемую запись речи шумов с определенными отношениями сигнал/шум, которые в дальнейшем прослушиваются группой auditors. Разборчивость речи - относительное количество (в процентах) правильно принятых элементов (слов, слов, фраз) артикуляционных таблиц. [4]

Для эксперимента были взяты несколько голосов, включающих в себя мужской, женский а также три голоса, содержащихся в помехе (два мужских и один женский). В аудиоредакторе AdobeAudition было произведено наложение аудиозаписей, после чего группа auditors в составе трех человек прослушала речь пяти дикторов с определенными отношениями сигнал/шум (-15дБ, -12дБ, -10дБ, -8дБ, -5дБ, -2дБ, 0дБ, 5дБ) с наложенной на них помехой. Усредненные результаты, полученные от группы auditors, представлены на рисунке 4.

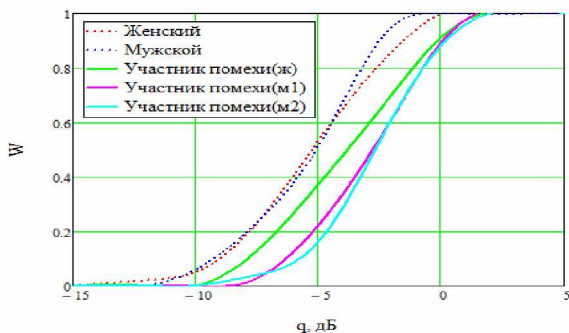


Рис. 4. График зависимости разборчивости W от интегрального отношения сигнал/шум q для помехи из словесных таблиц.

Лучшие результаты, а именно наименьшую разборчивость, имеют аудиозаписи дикторов, чьи голоса участвуют в помехе. Также заметно, что мужские голоса показывают меньшую разборчивость, чем женские.

ЗАКЛЮЧЕНИЕ

- Созданный речевой хор на всей полосе частот больше приближен к усредненным уровням русской речи, чем отдельные записи дикторов.
- Время усреднения записей не имеет влияния на форму спектра речи.
- Результаты артикуляционных испытаний показали, что записи, содержащие голоса участников, включенных в помеху, имеют наименьшую разборчивость.

Библиографический список

1. Хорев А.А., Макаров Ю.К., К оценке эффективности защиты акустической (речевой) информации. М.: Специальная техника №5, 2000. 14 с.
2. ГОСТ 16600-72. Передача речи по трактам радиотелефонной связи. Требования к разборчивости речи и методы артикуляционных измерений. М: ИПК Издательство стандартов, 1974. – 76 с.
3. Алдошина И., Основы психоакустики. М.: Оборонгиз, 2000. 154 с.
4. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. М: ИПК Издательство стандартов, 1996. 198 с.