**Вопросы:**

1. Основная задача систем преобразования текста в речь (TTS – Text-ToSpeech) – создание голосового сообщения без участия человека напрямую.

2. Основные критерии для оценки качества синтеза речи: разборчивость, т.е. возможность человеком декодировать речевой сигнал, и естественность – автоматически созданное высказывание не должно значительно отличаться от речи человека.

Под естественностью речи подразумевается наличие следующих свойств: соответствие синтезированных звуков речи входному тексту; определенная длительность и скорость произнесения каждого звука; наличие интонации; наличие тона соответствующей высоты; наличие акцента, что особенно важно для языков с несколькими вариантами (например, английский или итальянский).

3. В основе компилятивного синтеза речи заложены заранее записанные сегментные единицы (от звуков речи до целых высказываний), т.е. на вход системы подаются подготовленные человеком образцы речевых сигналов, которые система объединяет и озвучивает в требуемой последовательности

4. Формантный синтез осуществляется посредством искусственного создания акустических свойств звуков. Для создания качественного речевого сигнала необходимо около 60 акустических параметров. Данные параметры подразумевают всю информацию о частоте, амплитуде и ширине формант отдельных звуков. Процесс создания такого синтезатора требует много времени и подготовки.

Результатом работы подобного синтезатора является имитация речевых колебаний воздуха. Получаемый сигнал весьма сложно перепутать с человеческой речью, он имеет характерный «электронный» голос. Несмотря на то, что генерируемая формантными синтезаторами речь не естественна, она весьма удачно распознается человеком.

Статистический синтез речи подразумевает использование скрытых Марковских моделей, т.е. систем нейронных сетей. Их работу можно описать двумя этапами: во-первых, на вход системе подается достаточно большой объем фрагментов человеческой речи (минимум 5-10 часов), что позволяет нейронным сетям «обучиться» на речи диктора; во-вторых, система пытается самостоятельно воспроизвести речь по акустическим параметрам.

5.Артикуляционный синтез основывается на создании модели человеческого речевого аппарата и ее реализации с помощью некоторого механизма – «говорящей» машины, которая порождает речевой сигнал путем артикуляции. Акустические свойства речи на выходе производятся с учетом особенностей речевого аппарата человека, которые мы подробно рассмотрели в параграфе “Артикуляционная фонетика”. Создание аналога системы человеческого речевого аппарата все еще сложная задача и требует множества исследований в сферах нейрофизиологии, механики, аэродинамики и биоакустики. Имеющиеся на данный момент разработки способны порождать разборчивую речь, но назвать ее естественной пока нельзя.