**Rule based TTS**

1. **Text Analysis**: The system first performs text analysis by converting abbreviations, numbers, and dates into a readable format, which is typical of rule-based systems. These rules are predefined to handle specific text forms.
2. **Phonetic Analysis (Grapheme to Phoneme Conversion)**: This step uses a pronunciation lexicon and predefined rules to convert letters to sounds (graphemes to phonemes). In rule-based systems, this conversion relies heavily on linguistic rules to map text to speech sounds.
3. **Prosodic Analysis**: It assigns stress, intonation, and phrasing, based on predefined rules for how a sentence should be spoken, accounting for rhythm and melody in speech.
4. **Waveform Synthesis**: The system generates the actual speech output by concatenating pre-recorded speech units or using formant synthesis, based on the rules determined by previous steps.

**Statistical Parametric Text-to-Speech (TTS)** is another approach to speech synthesis that differs significantly from both rule-based and concatenative TTS systems. Instead of directly selecting and concatenating speech segments from a pre-recorded database, statistical parametric TTS models speech using mathematical models built from training data.

* **Deep learning models**: Utilizes **deep learning** to directly generate speech waveforms from text.
* **Improved naturalness**: Achieves significant gains in **naturalness**, **expressiveness**, and flexibility, resulting in more human-like voices.
* **Examples**: **Tacotron**, **WaveNet** are prominent examples of neural network-based TTS systems.