**Notes of “Deep Learning for Siri’s Voice”**

**Type**

Article

**Topic**

Speech Synthesis

**Notes**

**Usable Quotes or Figures**

**Summary**

Apple are constantly improving their speech synthesis voice used commonly by the IOS personal assistant Siri. Deep learning algorithms are machine learning methods that use learning data sets that improve the system’s capability to perform its task. It is a new method and is becoming vastly popular, it is expected to replace Markov models. Apple have been using this new method to create multiple new TTS systems, that are capable of speech synthesis with an improved natural tone. The Article also provides a breakdown of the process where text is transformed into sound. Firstly, the text undergoes analysis when it is given a predicted acoustic value by the prosody generation model. It is here where machine learning comes into use as the values/weights that are given are training the system in the “correspondence between text and speech”. After time, it will become tuned at giving the accurate values. This creates segments of speech which is then added to by Unit Selection. A process where additional information is added such as context or language conventions. This is then translated into waveforms and omitted as a voice.