

Augmented Reality app that educates the user about the lunar surface

By

Jonathan Roddy

Student ID: K00212118

Course: LC 233 Software Development

Supervisor: Carol Rainsford

Date: December 2020

**Table of Contents**

[Abstract 3](#_Toc61214610)

[1 Chapter 1: Introduction 4](#_Toc61214611)

[1.1 Background 4](#_Toc61214612)

[1.2 Aim of Project. 5](#_Toc61214613)

[1.2.1 Primary Objectives 5](#_Toc61214614)

[1.2.2 Secondary Objectives 6](#_Toc61214615)

[1.3 Research Questions 6](#_Toc61214616)

[2 Chapter 2: Augmented Reality 7](#_Toc61214617)

[2.1 Introduction 7](#_Toc61214618)

[2.2 What is Augmented Reality (AR) 7](#_Toc61214619)

[2.3 History 7](#_Toc61214620)

[2.4 Components of AR 9](#_Toc61214621)

[2.5 Architecture Diagram for AR 11](#_Toc61214622)

[2.6 Current uses 12](#_Toc61214623)

[2.7 Edutainment using AR 16](#_Toc61214624)

[3 Chapter 3: Text to Speech 19](#_Toc61214625)

[3.1 Introduction 19](#_Toc61214626)

[3.2 What is Text to Speech (TTS) 19](#_Toc61214627)

[3.3 History of TTS 19](#_Toc61214628)

[3.4 Components of TTS 20](#_Toc61214629)

[3.5 Synthesizer 21](#_Toc61214630)

[3.5.1 Concatenation 21](#_Toc61214631)

[3.5.1.1 Unit selection 21](#_Toc61214632)

[3.5.1.2 Diphone 23](#_Toc61214633)

[3.5.1.3 Domain-specific 23](#_Toc61214634)

[3.6 Formant 24](#_Toc61214635)

[4 References 26](#_Toc61214636)

**Table of Figures**

[**Figure 1:** Global Stats (GlobalStats, 2020) 5](#_Toc61212428)

[**Figure 2** History of Augmented Reality (Poetker, 2019) 8](#_Toc61212429)

[**Figure 3** Core component of Augmented Reality technology (Hui-Jeong Kim, 2020) 10](#_Toc61212430)

[**Figure 4** AR application Architecture (Herpich) 11](#_Toc61212431)

[**Figure 5** Pokemon Go(Statt, 2017) 13](#_Toc61212432)

[**Figure 6** Augmented Reality in sports (Usman, 2020) 14](#_Toc61212433)

[**Figure 7** IKEA Place (Dasey, 2017) 14](#_Toc61212434)

[**Figure 8** Face Tracking with ARKit (Apple, 2020) 15](#_Toc61212435)

[**Figure 9** Authorization with Face Recognition (BioID, 2020) 16](#_Toc61212436)

[**Figure 10** A conceptualized Framework for Edutainment Environment(Hanafizan Hussain, 2003) 18](#_Toc61212437)

[**Figure 11** Overview of a typical TTS system (Andy0101) 21](#_Toc61212438)

[**Figure 12** Example of Decision trees (Jadhav) 22](#_Toc61212439)

# Abstract

My final year project that I am proposing is an android augmented reality app that allows the user to experience what it is like to be on the lunar surface, highlighting features like Apollo landings, famous craters and drone landings. The user chooses a point on the AR moon and will be able to visually see what the surface of the Moon would be in their specific location. This project should be entertaining while being able to have an educational benefit giving the user an insight into what it would like to stand on the surface on the Moon.

*I will update this section when I have document finished; findings etc*

# Chapter 1: Introduction

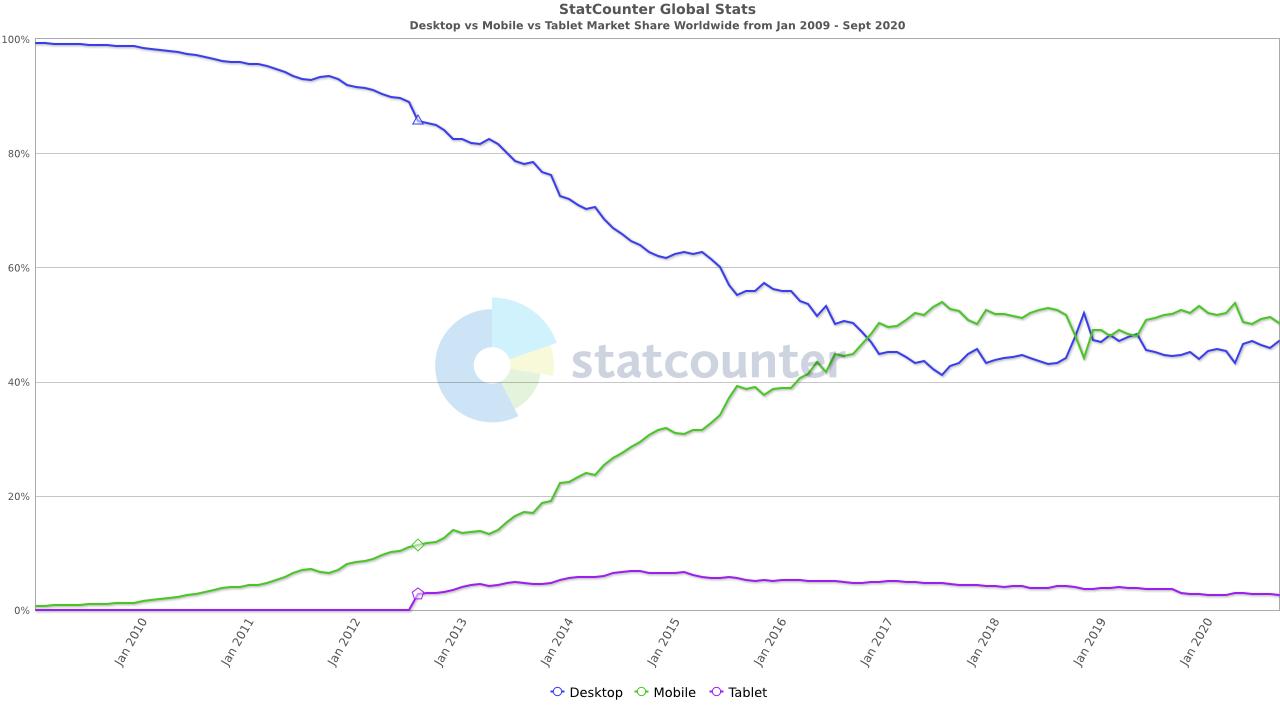
## Background

Have you have ever stood out on a clear night’s sky and watch our celestial neighbour, the Moon, as it slowly moves? I have done this several times in my life and will continue to do so at every chance that I can. Our Moon is a wonderful and scary place to find yourself alone. It is a place where no life has been found nor can a normal human being survive on the surface without external help from a spacesuit. It is 4.51 billion years old (Wall, 2017) which is a bit younger than the Earth at 4.54 billion years old. (Redd, 2019). No one knows exactly how the Earth received a moon but without it, life today would be quite different. There are a few hypotheses that scientists have agreed upon that could explain how we got out moon: Giant impact which states that a large celestial body impacted our young earth, Co-formation which states that the moon was formed around the same time as the Earth, or Capture theory which states that we inherited a floating Moon. (Reed, 2017). These types of facts are what my aim is for this project. My interest in astronomy helped me to decide on what this project will be as I would love to use an app like this.

I choose to develop a smartphone application as this allows me to have a much greater target audience. People today have access to a smartphone where they may not have access to a computer or laptop. Ever since the first smartphones were released, the mobile market scene has exploded more than anyone would imagine.

Apple released its very first iPhone back in 2007 which at the time read with them releasing the Apple App Store later in 2008. At the beginning of Apple's success, the apps that they were providing were developed by in-house developers limiting other developers to launch independent apps on IOS. Apple's first iPhone was so successful that other companies were soon to catch up and emulate their success. Google launched its equivalent, Android in 2008. Google's business model allowed all developers to publish their apps creating a burst of third-party app development. IOS and Android are the two common smartphone operating systems, however, there are many smaller options like KaiOS, Windows Phone, Symbian OS, and Blackberry.

Currently as of writing this literature review the current worldwide market share for smartphone devices as of January ‘09 – October’20 is divided as such: Mobile 50.21% Desktop 47.17% and Tablet at 2.62%. (GlobalStats, 2020) From this statistic, the mobile market share is currently in favour of mobile.

**Figure 1:** Global Stats (GlobalStats, 2020)

This project gives me the chance to be able to develop an Android application that teaches the user about the Moon and its surface. The best way for me to portray this to the user is in a way that both excites the user as well as informs them. I have chosen to use Augmented Reality (AR) in my android app as it is an immersive technology. Using AR gives a user something to look at as well as something to manipulate on their smartphones. I plan on using Text to Speech (TTS) within my application to help in narrating some text to the user. The use of both AR and TTS together helps in the immersion of the environment that I want to give to the user.

## Aim of Project.

The aim of my project is to be able to create an immersive Augmented Reality application that teaches the user about the Moon and its surface. To be able to accomplish this I need to have a further analysis of how to develop a mobile application fixating on the educational aspect.

### Primary Objectives

My primary objective for this project is to develop an Augmented Reality mobile application that educates the users about interesting features of our Moon.

1. Develop an Augmented Reality Android app that connects to a local and online instance on a database.
2. Explore the idea of simulating the lunar surface.
3. Implement and expand upon the text to speech to further aid any user.
4. Conduct an extensive literature review
5. The history of Augmented Reality and its common uses to date.
6. The development of a mobile application and its common faults
7. How can Augmented Reality help educate people on new topics
8. Utilizing immersive technologies and their benefits
9. Detailed history on the Moon, structure, and topography
10. How Text-to-Speech enables easier ways to access digital content

### Secondary Objectives

My secondary objectives for this project would focus on my personal development during this research paper.

1. I aim to meet all the deadlines on time in a professional manner.
2. Take the advice given by my supervisor and lectures and expand upon them further.
3. Learn in detail how to develop an Android app that implements augmented reality
4. To advance my knowledge and skills as a software developer.

## Research Questions

1. How can Augmented Reality be used in an effective educational environment?
2. How can Augmented Reality and Mobile applications reduce the inaccessibility of environments such as the Moon?
3. How can you explore the lunar surface utilizing Immersive technologies?

# Chapter 2: Augmented Reality

## Introduction

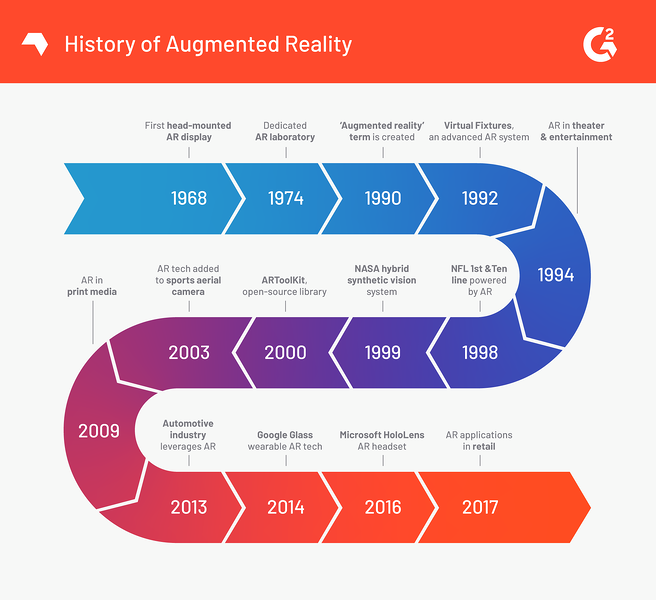
In this chapter, I will discuss what Augment Reality (AR) exactly are, the components that make up an AR application, how developers can implement AR, any limitations or drawbacks with using AR, and how to use AR in educational entertainment ‘edutainment’

## What is Augmented Reality (AR)

Augmented Reality is one of the immersive technologies that has grown in recent years with the development of stronger and faster devices. Immersive technology refers to technology that attempts to emulate a physical world through the means of a digital or simulated world. Virtual Reality (VR) is another such form of immersive technology, but it encapsulates the user with a wearable headset and controllers. AR adds digital elements to a live view often by using the camera on a smartphone (Furht, 2011). The use of AR lets the user experience the real world which has been digitally augmented or enhanced in some way. VR on the other hand, removes the user from the real-world experience, replacing it with a completely simulated one. For this reason, I will develop my project using AR as most of my users will not have the correct equipment to utiliser VR. AR can be handled and used with a common smartphone device as well as many other devices. This ability to use AR on smartphones is what is enabling me to peruse this project.

## History

AR is not a new technology by any means as the first attempt was invented by a cinematographer called Morton Heilig in 1947. Figure 1 below, shows us a detailed roadmap of where in time did a breakthrough happen (USC School of Cinematic Arts, 2020). He invented the ‘Senorama’ which delivered visuals, sounds, vibration, and smell to the viewer. It was not computer controlled but it was the first example of an attempt at adding additional data to an experience.



**Figure 2** History of Augmented Reality (Poetker, 2019)

In 1968, an American computer scientist and early internet influencer, Ivan Sutherland invented the head-mounted display as a kind of window into a virtual world. The technology used at the time made the invention impractical for commercial use.

In 1975, Myron Krueger, an American computer artist developed the first ‘virtual reality’ interface in the form of ‘Videoplace’ which allowed its users to manipulate and interact with virtual objects and to do so in real-time.

In 1980 Steve Mann, A computational photography researcher, gave the world wearable computing.

The word ‘virtual reality’ was not coined up until 1989 followed by Thomas P Caudell of Boeing coining the phrase ‘augmented reality’ in 1990

The first properly function AR system was developed at USAD Armstrong’s Research Lab by Louis Rosenberg in 1992. This was called Virtual Fixtures and was an incredibly complex robotic system that was designed to compensate for the lack of high-speed 3d graphics processing power in the early 90s. It enabled the overlay of sensory information on a workplace to improve human productivity

There are many other breakthroughs in augmented reality in recent years such as;

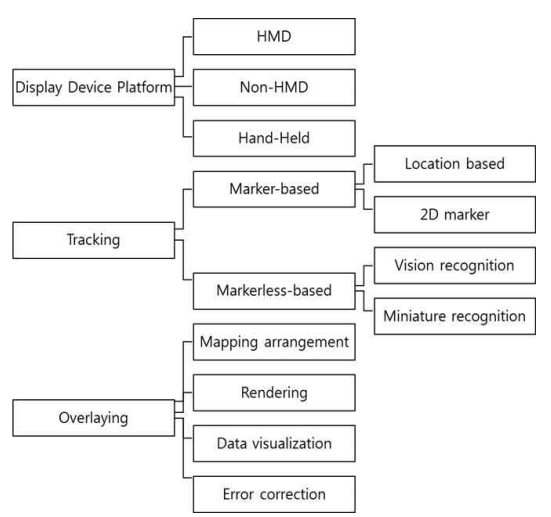
* Bruce Thomas developing an outdoor mobile AR game called ARQuake in 2000.
* ARToolKit, a design tool, being made available in Adobe Flash in 2009
* Google announced its open beta of Google Glass in 2013
* Microsoft announcing augmented reality support and their augmented reality headset HoloLens in 2015.

## Components of AR

AR is achieved through a variety of technological innovations; these can be implemented on their own or in conjunction with each other to create augmented reality. They include:

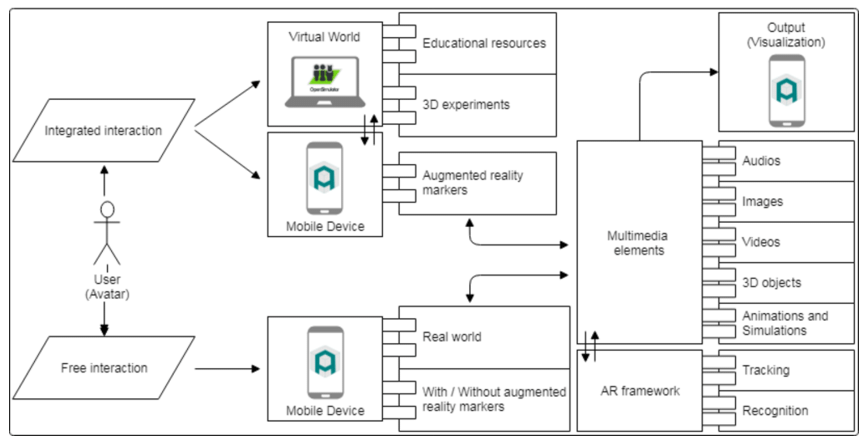
* **General hardware components –** This refers to the processor, the display, the sensors, and input devices. Typically, a smartphone contains a processor, a display, accelerometers, GPS, camera, microphone, etc., and contains all the hardware required to be an AR device.
* **Displays –** while a monitor is perfectly capable of displaying AR data there are other systems such as optical projection systems, head-mounted displays, eyeglasses, contact lenses, the heads up display (HUD), virtual retinal displays, EyeTap which is a device which changes the rays of light captured from the environment and substitutes them with computer-generated ones (EyeTap, 2020), Spatial Augmented Reality (SAR) which uses ordinary projection techniques as a substitute for a display of any kind, and handheld displays.
* **Sensors and input devices –** GPS, gyroscopes, accelerometers, compasses, Radio-frequency identification (RFID), wireless sensors, touch recognition. Speech recognition, eye tracking, and peripherals.
* **Software -** the majority of development for AR will be developing further software to take advantage of the hardware capabilities. There is already an AR Mark-up Language (ARML) which is being used to standardize XML grammar for virtual reality. There are several other software development kits (SDK) which I will further investigate in the below topic.

Below in figure 4, we can see a rough outline of how what are common variables in an AR system.

* *The display device platform* is made up of HMD (head-mounted display), non-HMD, and handheld. Knowing what device your customers will be using is key as it changes the whole development structure according to the resources in hand.
* *Tracking* can be doing using marker-based fixtures like a card or location so that the projection is fixed in place. Developing a marker-less based system allows for the application to 3d scan its surroundings for more of an open approach.
* *Overlaying* refers to the application itself and how it works to ensure it comes out error-free by dealing with map arrangement, rendering, data visualization, and error correction.

**Figure 3** Core component of Augmented Reality technology (Hui-Jeong Kim, 2020)

## Architecture Diagram for AR

Figure 3 outlines an architecture diagram for a full AR application. In the world of AR there can only be two ways in which an object can be shown to a given user, marker and marker-less interactions. Marker interactions is where the application is looking for a pre-determined picture in real-life so that it may use overlay its models on top of it. These pictures can be either QR codes, logos or an image used within a company. Marker-less however does not need an image to be able to displays its models. This form of AR is a lot more complex as the technology itself is doing a lot more then scanning the area for a picture. Marker-less uses complex algorithms to investigate the environment to assess what can be seen. This technology is called Simultaneous Localization and Mapping (SLAM) (Agrawal, 2020).

**Figure 4** AR application Architecture (Herpich, 2018)

Navigating through the above diagram the user has the option for both integrated interaction and free interaction, respectively. If they were to go down the route of integrated interaction, they would find themselves using marker-based tools to assist with the application. If they went into free interaction, they would be using the markerless based tools for the application. Multimedia elements would encapsulate all resources like audio, images, videos, 3d objects and animations, and simulations. It communicates with the AR framework which handles the tracking and recognition software. The final project is released with all data coming from the multimedia elements, to the mobile application where the application is presented to the user.

## Current uses

There are apps available for or being researched for AR in nearly every industrial sector including:

* Archaeology, Art, Architecture
* Commerce, Office
* Construction, Industrial Design
* Education, Translation
* Emergency Management, Disaster recovery, Medical and Search and Rescue
* Games, Sports, Entertainment, Tourism
* Military
* Navigation

**Interactive Gaming**

When it comes to daily AR users, the Pokémon Go app takes the gold. Based on a playing card game that became popular in the ’90s, this Nintendo gaming app skyrocketed in popularity when it was released as a free mobile app in 2016. Users simply download the app to hunt for Pokémon characters in their everyday surroundings. The app uses your GPS data to determine where you are and reveal the virtual characters within your real world. Pokémon Go has to date, 147 million active users. (Pokemon Go revenue and Usage, 2020)

**Figure 5** Pokemon Go(Statt, 2017)

**Data-Driven Sports Broadcasting**

These days there is no shortage of AR on a field, rink, court, or turf when it comes to live-action sports. These were some of the first applications of AR praised by mainstream consumers. The goal behind AR in sports broadcasting was to increase viewership by providing additional information during a game. In this case, AR applies 3D graphics and interactive visual elements on top of the live video footage of the game. It’s just another way to visualize the trajectories of a ball, puck, or a player with the help of lines and curves. This tool helps the audience see close calls, fouls, and record-breaking achievements in a much more detailed way than ever before.

**Figure 6** Augmented Reality in sports (Usman, 2020)

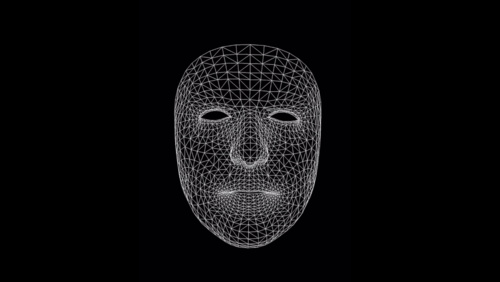
**More efficient and engaging shopping**

****Another common use of AR in our everyday lives is during the home buying process. Using a ‘virtual tours’ option, prospective homebuyers can often view a property from the comfort of their desktop computer or mobile device before making the trek to see the home in-person. These clients can also use furniture placements apps like Houss and DecorMatters to see how they would furnish one of their potential homes. IKEA is leading the charge in this respect using a digital furniture placement app called IKEA Place that also allows you to virtually ‘try before you buy’.

**Figure 7** IKEA Place (Dasey, 2017)

**Video Conferencing**

Filters on Snapchat, Instagram, and iPhone Facetime Memoji are just some of the ways consumers use AR every day. Facial recognition software within camera hardware has gotten increasingly more sophisticated. Apple even offers a solution developer kit just for your face. In this case, dedicated front camera arrays, known popularly as the ‘dreaded’ notch, which has become an industry-standard, scans your face with up to 30,000 dots near-instantly to get a very secure read. This provides a unique and in-depth analysis of your face through mathematical ratios. (Apple, 2020)



***Figure 8*** *Face Tracking with ARKit* (Apple, 2020)

Similar to gaming, family members and co-workers can use this software to build a virtual representation of themselves anyway around the world. Imagine people sitting in a conference room together even though they are in different countries. Or imagine a family gathering taking place in someone’s living room, despite being in different cities or other sides of the glove. This increased virtual presence will have a significant impact on consumer's line in the future by cutting back on travel costs.

**Authorization**

The quickest and most common way iPhoneX, or newer users, use AR today is by unlocking their phone with their faces. Long gone are the days of using a thumbprint of pushing a button to unlock your phone. These users are also able to scan their faces as a way of gaining access to password-protected sites. Instead of resetting or looking up a forgotten password, facial recognition software can be sued in place of a username and password to any website or app in just a few seconds. Facial recognition software also allows users to authorize payments through Apple Pay, Google Pay, or Samsung Pay. This applies to mobile and in-person purchases where stores have readers for mobile devices. AR payment options expedite the consumer process.

**Figure 9** Authorization with Face Recognition (BioID, 2020)

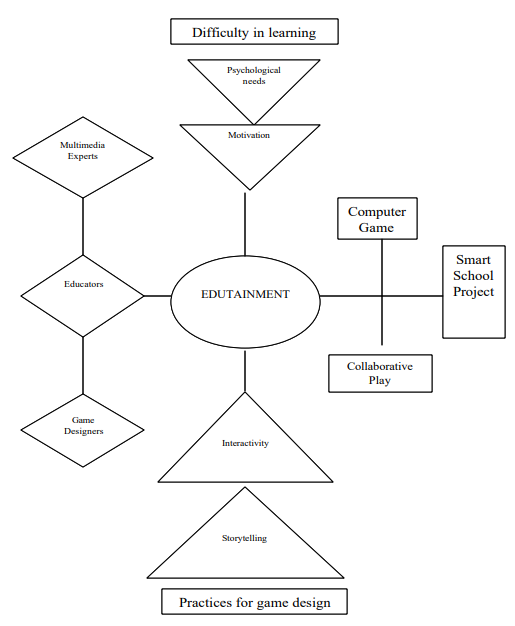
## Edutainment using AR

‘Edutainment – products such as books, television programs and especially computer software that both educate and entertain’ (Oxford Learner's Dictionaires, 2020)

Hashim and Hanafizan Hussain have created a framework that they describe fixates on the development of a computer game. (Hanafizan Hussain, 2003). They have gathered eight factors that reflect the design of an edutainment environment.

* **Meaningful learning –** computer games provide environments in which children find learning to be meaningful and useful.
* **Goal –** Goals create a sense of mission in children and often they will stay on until they could finish the game.
* **Success –** Accomplishing goals and giving achievements can provide children with a sense of success.
* **Challenge –** Challenges ensure that children would not get bored easily.
* **Cognitive artifact –** Introduction cognitive artifacts enhance our cognitive abilities.
* **Association through pleasure –** Children need to associate learning with some pleasant memory enforcing that the concepts remain with them.
* **Attraction –** Game design can create environments in which children get excited about embedded learning activities and therefore are willing to be immersed in it and spend time learning it.
* **Sensory stimuli –** Educators must have sensory stimuli so that the fun of playing the game and make the learning more enjoyable and memorable**.**

They have argued that the most common issues that children have when it comes to learning anything are motivation and psychological needs. Keeping the user’s motivation active ensures that they will be using the application or game for longer. In Figure 8, you can see that edutainment is built upon several components pulling from both the educators, game design, and multimedia. To get a complete edutainment environment it needs to allocate for each of those problems.



**Figure 10** A conceptualized Framework for Edutainment Environment(Hanafizan Hussain, 2003)

# Chapter 3: Text to Speech

## Introduction

In this chapter, we will investigate the technology of TTS, the history of TTS, different versions of TTS systems, and common problems this technology faces. I can assume that the reader has at some stage interacted with a device or application that allowed such a technology.

## What is Text to Speech (TTS)

Text-to-speech (TTS) is a speech synthesis feature that is used to create an audio version of the text in a computer document, such as a help file or a web page. Speech synthesis is the computer-generated rendering of a human voice (Allen, 1987). The use of TTS can enable a visually challenged person to read and communicate with a computer display.

The scale in which a good speech synthesizer has been developed when it can mimic the similarities of a human’s speech and on how well it can be understood.

A smart TTS program that can help people who have visual impairments or reading disabilities interact and read along with the computer with little to no resistance. Fortunately, a lot of time and research has been developed to enable us with a smarter and more diverse TTS system. Most of our devices we use daily have some form of TTS enable API’s built-in.

## History of TTS

Humans have tried to mimic human speech for a couple of hundred years, long before the invention of electricity. The most popular origin inventors would be ‘Brazen Heads’ who involved Pope Silvester II (d. 1003AD), Albertus Magus(1198-1280), and Roger Bacon(1214-1294).

A German-Danish scientist in 1779, Christian Gottlieb Katzenstein won the first prize in a competition for the models he built of the human vocal tract that could produce the five long vowel sounds (a, e, i, o, u) (Helsinki University of Technology, 2006) . From this invention, the bellowed-operated speech machine came from Wolfgang von Kempelen of Pressburg, Hungary. (J. B. Degen). This machine was different from the first iteration as this model was fitted a tongue with lips so that they could sound out constants and vowels.

The next iteration of this design came in 1837, where Charles Wheatstone produced a design based on the last design, and in 1846, Joseph Faber exhibited the ‘Euphonia’. (G.Mattingly, 1974)

In 1930 Bell Labs developed a new way forward for TTS as with his invention, the vocoder could change sounds into its fundamental tones and resonances. The inventor Homer Dudley then proceeded to continue his work by creating a keyboard-operated voice-synthesizer called The Voder(Voice Demonstrator), which he exhibited in 1939 at the New York World’s Fair.

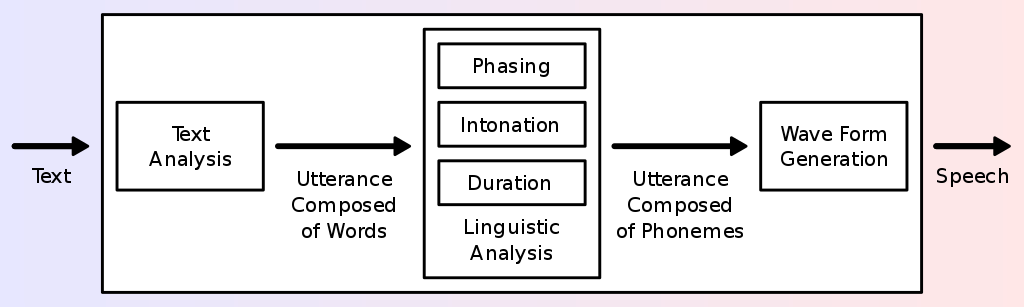
At Haskins Laboratories, Dr. Franklin S.Cooper and his colleagues developed the Pattern playback in the early 1940s but didn’t complete it in 1950. This machine was different from all other forms as it approached the design differently. They investigated the method of using the acoustic wavelength of the word to be recorded and played back. This way of design allowed scientists to discover acoustic cues for the perception of different important phonetic sentences.

## Components of TTS

A complete TTS system is made up of two parts, a front, and a back-end. (Jan P.H. van Santen, 1997) In figure 11, we can see what is described as an overview of a TTS system.

The front end has two major operations that it needs to handle, it must first translate the raw text passed into it into each symbol like its corresponding numbers and symbols. The second job it must do is to assign each letter, symbol, or letter into its respective phonetic transcriptions and divides and marks the text into each of its prosodic units, like phrases, clauses, and sentences. Phonetic transcriptions and prosody information together make up the symphonic linguistic representation that is output by the front-end.

The back end is called the synthesizer as this part is the output from the front end. The back end converts the symbols it has been received into what it believes that symbol sounds like. (Santen, 1994)



**Figure 11** Overview of a typical TTS system (Andy0101)

## Synthesizer

Naturalness and intelligibility are the most significant properties of a speech synthesis device (Taylor, 2009). Naturalness defines how closely the output sounds like a human voice, while the ease with which the output is understood is intelligibility. The perfect synthesizer for speech is normal as well as intelligible. Systems of speech synthesis typically aim to optimize all characteristics.

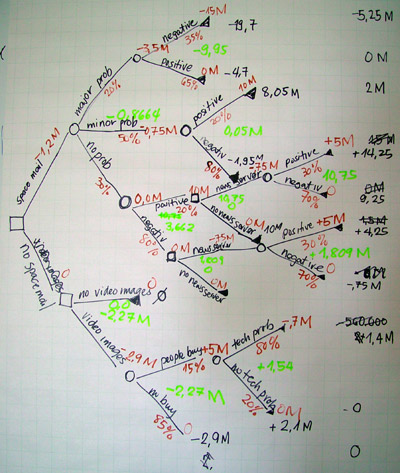
Concatenative synthesis and formant synthesis are the two primary technologies that generate synthetic speech waveforms. There are strengths and disadvantages in each technology, and the expected uses of a synthesis method will usually dictate which approach is used.

### Concatenation

Concatenation synthesis is a process for synthesizing sounds by concatenating short samples (called units) of captured sound. The unit length is not precisely specified and can vary depending on the implementation, varying from approximately 10 milliseconds to 1 second. Concatenative synthesis has three major sub-types.

#### Unit selection

Unit selection synthesis uses broad recorded speech databases. Every reported utterance is segmented into any or all of the following during the creation of the database: individual phones, phones, half-phones, syllables, morphemes, sentences, phrases, and phrases. "Usually, with some manual correction afterward, the division into segments is carried out using a specially adapted speech recognizer set to a "forced alignment" mode, using visual representations such as the waveform and spectrogram (Black, 2002).

 Based on the segmentation and acoustic parameters such as the fundamental frequency (pitch), length, location in the syllable, and neighbouring phones, an index of the units in the speech database is then established. The desired target utterance is generated at run time by deciding from the database the best chain of candidate units (unit selection). As shown in Figure 10, this approach is usually performed using a specially weighted decision tree.

**Figure 12** Example of Decision trees (Jadhav)

The unit selection offers the greatest naturalness since the recorded speech only applies a limited amount of digital signal processing (DSP). DSP also makes recorded speech sound less natural, although some devices, at the point of concatenation, use a small amount of signal processing to smooth the waveform. Sometimes the output from the best unit-selection systems is indistinguishable from real human voices, especially in contexts where the TTS system has been tuned for. (John Kominek, 2004).

Unit selection algorithms have also been known to choose segments from a position that results in less than optimal synthesis (e.g. minor terms are unclear) even though the database has a better choice. Researchers have recently proposed numerous automated methods for detecting abnormal segments in speech synthesis systems for unit selection. (Zhang, 2004).

Researchers have recently proposed numerous automated methods for detecting abnormal segments in speech synthesis systems for unit selection. (William Yang Wang).

#### Diphone

Diphone synthesis uses a minimal speech database containing all the diphones (sound-to-sound transitions) occurring in a language is used in diphone synthesis. The number of diphones depends on the language's phonotactic: for example, there are about 800 diphones in Spanish and about 2500 in German. In diphone synthesis, the speech database only contains one instance of each diphone. The goal prosody of a sentence is superimposed on these minimal units. At runtime, these minimal units are superimposed on the target prosody of a sentence using automated signal processing techniques such as linear predictive coding. PSOLA (ppgb, 2001) or MBROLA.

More recent techniques, such as pitch modulation using discrete cosine transformation in the source domain (R.Muralishankar, 2004). Diphone synthesis suffers from concatenative synthesis acoustic glitches and formant synthesis's robotic-sounding nature and has few of the advantages of either method other than a small scale. As such, its use in commercial applications is decreasing, although it continues to be used in research since many software implementations are readily accessible.

A teaching robot, Leachim, invented by Michael J. Freeman, is an early example of diphone synthesis (cyberneticzoo.com, n.d.). Leachim had information about the curricular class and some biographical information about the 40 students it was expected to teach. It was tested in the Bronx, New York, in a fourth-grade classroom (World Future Society, 1978).

#### Domain-specific

To construct full utterances, domain-specific synthesis concatenates pre-recorded terms and phrases. It is used in systems where, including transit service announcements or weather forecasts, the variety of texts the device can generate is restricted to a certain domain (L.F. Lamel, 1993). The system is very easy to incorporate and has long been in commercial use in devices such as talking clocks and calculators, for a long time. The level of naturalness of these systems can be very high because the variety of sentence types is limited, and they closely match the prosody and intonation of the original recordings. Since the number of sentence forms is small, and they closely follow the prosody and intonation of the original recordings, the degree of naturalness of these structures can be very high. They are not general-purpose and can only synthesize the variations of terms and phrases for which they have been preprogrammed because these systems are constrained by the words and phrases in their databases.

However, the blending of words within the language spoken spontaneously will also create difficulties until the differences are taken into account. "In non-rhotic English dialects, for example, the "r" is usually pronounced in words such as "clear" /klə/ only when the following word has a vowel as the first letter (e.g. "clear out" is realized as /ˌklɪəɹˈʌʊt/). In non-rhotic English dialects, for instance, the "r”. Likewise, in French, if preceded by a word that begins with a vowel, an effect called liaison, several final consonants become no longer silent. A basic word-concatenation scheme, which would require extra sophistication to be context-sensitive, will not replicate this alternation.

## Formant

Formant synthesis, at runtime, does not use individual speech recordings. Instead, using additive synthesis and an acoustic model (physical modelling synthesis) (Darmouth College, 2011) the synthesized speech output is generated. To construct a waveform of artificial speech, parameters such as fundamental pitch, voice, and noise levels differ over time. This approach is often referred to as rules-based synthesis, but there are also rules-based components in many concatenative schemes. Artificial, robotic-sounding speech is created by multiple systems based on formant synthesis technology that can never be mistaken for human speech. Maximum naturalness, however, is not always the goal of a method of speech synthesis, and formant synthesis systems have benefits over concatenative systems. Also at extremely high levels, formant-synthesized speech may be accurately intelligible, eliminating the auditory glitches that usually affect concatenative systems. The visually impaired use high-speed synthesized speech to easily access machines using a screen reader. Formant synthesizers are normally smaller programs than concatenative systems since they do not have a speech sample index.

In embedded systems, where memory and microprocessor capacity are particularly small, they can also be used.

Because formant-based systems provide full control of all facets of the speech production, it is possible to generate a wide range of prosodies and intonations, conveying not only questions and declarations, but a spectrum of emotions and tones of voice.

# References

Agrawal, S. (2020, 09 01). *What is the Difference Between Marker-based AR and Marker-less AR?* Retrieved from brsoftech.com: https://www.brsoftech.com/blog/marker-based-ar-and-marker-less-ar/

Allen, J. (1987). *From text to speech.* Cambridge : Cambridge University Press.

Andy0101. (n.d.). *Overview of a typical TTS system.*

Apple. (2020, 11 2). *Face tracking with ARKit*. Retrieved 11 2, 2020, from developer.apple.com: https://developer.apple.com/videos/play/tech-talks/601/

BioID. (2020, 11 2). *BioID Facial Recognition App with face login*. Retrieved 11 2, 2020, from BioID.com: https://www.bioid.com/facial-recognition-app/

Black, A. W. (2002). Perfect Synthesis for all of the people all of the time. *Unit Selection Synthesis*.

cyberneticzoo.com. (n.d.). *1960 – Rudy the Robot – Michael Freeman (American)*. Retrieved from cyberneticzoo.com: http://cyberneticzoo.com/robots/1960-rudy-the-robot-michael-freeman-american/

Darmouth College. (2011, 08 06). *Music and Computers*. Retrieved from https://web.archive.org/web/20110608035309/http://digitalmusics.dartmouth.edu/~book/MATCpages/chap.4/4.4.formant\_synth.html

Dasey, D. (2017, 09 12). *IKEA Place App*. Retrieved 11 2, 2020, from IKEA.com: https://www.ikea.com/ms/en\_JP/this-is-ikea/ikea-highlights/2017/ikea-place-app/index.html

EyeTap. (2020, 11 1). *EyeTap*. Retrieved 11 1, 2020, from eyetap.org: http://www.eyetap.org/video/

Furht, B. (2011). *Handbook of Augmented Reality.* Springer Science & Business Media.

G.Mattingly, I. (1974). Speech Synthesis for phonetic and phonological models. *Current trends in linguistics*.

GlobalStats. (2020, October 19). *Desktop vs Mobile vs Tablet Market Share Worldwide*. Retrieved from gs.statcounter.com: https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet/worldwide

Hanafizan Hussain, E. (2003). A Conceptualized Framework for Edutainment. *Where Parelles Intersect*.

Helsinki University of Technology. (2006). History and Development of Speech Synthesis. *History and Development of Speech Synthesis*.

Herpich, F. (n.d.). *AR Application Architecture .*

Hui-Jeong Kim, J. C. (2020). Augmented Reality Based Edutainment Content Design about. 46.

J. B. Degen, W. (n.d.). *Mechanism of the human speech with description of its speaking machine.*

Jadhav, G. (n.d.). *A manually drawn decision tree diagram drawn on a flip chart.*

Jan P.H. van Santen, R. W. (1997). *texts.* New York : Springer.

John Kominek, A. W. (2004). the CMU Arctic speech databases.

L.F. Lamel, J. G. (1993). Generation and Synthesis of Broadcast Messages.

Oxford Learner's Dictionaires. (2020, 11 2). Retrieved 11 2, 2020, from Oxford Learner's Dictionaires: https://www.oxfordlearnersdictionaries.com/definition/english/edutainment?q=edutainment

Pantazis, Y. (n.d.). *Modules of a Unit Selection Synthesis System. .*

Poetker, B. (2019, 08 22). *A brief history of Augmented REality*. Retrieved 11 1, 2020, from Learning Hub: https://learn.g2.com/history-of-augmented-reality

*Pokemon Go revenue and Usage*. (2020, 10 30). Retrieved 11 2, 2020, from BusinessofApps: https://www.businessofapps.com/data/pokemon-go-statistics/#:~:text=Indeed%2C%20Pok%C3%A9mon%20GO%20stats%20from,starting%20from%20this%20May%20bump.

ppgb. (2001, 1 1). *PSOLA synthesis*. Retrieved 11 6, 2020, from fon.hum.uva.nl: http://www.fon.hum.uva.nl/praat/manual/PSOLA.html

R.Muralishankar. (2004). Modification of pitch using DCT in the source domain. *Speech Communication*, 143-154.

Redd, N. T. (2019, February 7). *How Old Is Earth?* Retrieved October 19, 2020, from space.com: https://www.space.com/24854-how-old-is-earth.html

Reed, N. T. (2017, November 16). *How Was the Moon Formed?* Retrieved October 19, 2020, from space.com: https://www.space.com/19275-moon-formation.html

Santen, J. P. (1994). Assignment of segmental duration in text-to-speech synthesis. *Computer Speech & Language*, 95-128.

Statt, N. (2017, 12 20). *Pokemon Go IOS ipdate will deliver more advanced augmented reality thanks to ARKit*. Retrieved 11 2, 2020, from The Verge: https://www.theverge.com/2017/12/20/16798388/pokemon-go-niantic-apple-iphone-ios-arkit-augmented-reality-ar-plus

Taylor, P. (2009). *Text-to-speech synthesis.* Cambridge: Cambridge University Press.

USC School of Cinematic Arts. (2020, 11 1). *USC School of Cinematic Arts*. Retrieved 11 1, 2020, from uschefnerachive.com: http://uschefnerarchive.com/mortonheilig/

Usman, F. (2020, 1 22). *Augmented Reality in sports*. Retrieved 11 2, 2020, from glidarone.com: https://www.glidarone.com/augmented-reality-in-sports/

Wall, M. (2017, January 11). *How Old Is the Moon? Scientists Say They Finally Know*. Retrieved October 19, 2020, from Space.com: https://www.space.com/35291-moon-age-pinned-down.html

William Yang Wang, K. G. (n.d.). Automatic Detection of Unnatural Word-Level. *2011*, 54.

World Future Society. (1978). *The Futurist.*

Zhang, J. (2004). Language Generation and Speech Synthesis in.