Using eye/gaze tracking (with narrator) to improve reading ease, speed, and comprehension

Rasmus Falden, rafa@itu.dk

Supervisors:
Fabricio B. Narcizo (fabn@itu.dk), & Dan W. Hansen (witzner@itu.dk)
IT University of Copenhagen
Denmark
1 - 7 - 2022

I Abstract

This bachelors project covers eye tracking and how it can be used as a interface with a Text-to-Speech narration, the end goal being providing assistance to those who struggle with reading this could be people with reading learning disorders such as dyslexia.

Most estimates of people with reading learning disorders are between 5-20% accounting for a substantial amount of people that struggle with reading and with how integrated reading is in our societies it makes sense to try and develop assistance for those that struggle.

The project goes into detail about how the different technologies can interface to increase interactability of narration software especially during narration.

The results being that the sample size and bias within the data results in data of to little quality to prove anything regarding the hypothesis.

This exact field of research is not directly being explored, though the different aspects are. In the end this project tries to unify and explore a lot of ideas and cannot reach any resolution.

Keywords: Eye Tracking, TTS Narration, Speech Synthesis, Human Computer Interaction (HCI)

Contents

1 Introduction

The educational system of today is build on reading. Students have to read a lot of text across the years that their education spans from homework to books while reading in class. Reading is found in every facets of education and is one of the first things taught to children when they begin schooling Miller 2017. The independents of a reader is of great concern due to how much of our society structured around reading. According to a number of studies found discussed here College 2021 it is normal to read 15-30 minutes per day and that is only for books and that we encounter and estimated 105.000 words per 12 hours (the waking hours) this estimate was done in 2011 and would likely have grown larger as social media, smartphones and the internet has grown and changed over the years Andrea 2010 James E. Short 2011.

With such an emphasis on reading those who are not strong readers or even have a reading disability are severely hampered and this group is not small, while illiteracy in 1st world countries is not high Review 2022, there is still a substantial group who find reading difficult as an estimated 20% of a given population is dyslexic Dyslexia and Creativity n.d. though this 20% is not agreed on some sources estimate as little as 5% LDOnline 2004. Dyslexia is a reading learning disability and is the most common learning disability in the world, dyslexia is eloquently defined from Clinic 2017 as "Dyslexia is a learning disorder that involves difficulty reading due to problems identifying speech sounds and learning how they relate to letters and words (decoding)" .

There are a number of different tools available to people who suffer from dyslexia, though the tools changes from country to country in Denmark CD-ord is commonly provided which is a program designed for dyslexic people, and has things like text-to-speech, converting image based pdfs to a docx style documents, spelling help/predictions, and more Vitec 2022. The main take away for this project is the narration in the form of text-to-speech (TTS). There are many tools available online that can pronounce text with varying quality like NaturalReaders where you can get access to free voices or premium voices. Most operating systems (OS) have built-in voices for things like on-screen text narration to assist blind people navigate on computers these same voices can be used for TTS, though the quality changes from OS to OS and language to language meaning that languages with a smaller group of speakers might not have a TTS voice or in the case of Microsoft Helle the Danish voice for Windows being of a lesser quality (personal opinion supported by comparisons to other available voices). In addition to the quality of the pronunciation and intonation, TTS narration suffers from a host of other difficulties Canada n.d. though the industry has grown and seen more commercial success over the years especially during Covid19 Simonite 2022.

Still TTS narration is very simple in its approach: feed text, get audio. This however just means that there are a number of interesting ways to expand on the technology with say eye tracking.

Eye tracking is a comparatively speaking old technology Tobii 2022, but has seen a strong resurgence during the age of computers and now more than ever the potential of eye tracking is being explored from VR based eye tracking, screen based eye tracking, and wearable apparel eye tracking (glasses). Farnsworth 2018 These innovation on old concepts are being realised to such a degree that new possibilities are opening up as eye tracking provides exciting human computer interaction (HCI) possibilities such as combinding eye tracking and TTS narration.

1.1 The why (problem statement)

Reading is essential and reading is harder for some than others this is the problem in the shortest possible explanation. There have been other solutions to this problem like narration be it TTS, human based or things like audio books or a change in education as things like dyslexia is a learning disability not an inability to learn, the Danish site for dyslexia Børne og undervisningsministeriet n.d. and an American site for dyslexia Dyslexia and Creativity n.d. agree on this.

There exist tools that help with lessening the effects that dyslexia has the day to day life of the people who are diagnosed with it. There is design space available for improvement this technology is not finished expanding and getting improved, in fact most text-to-speech (TTS) programs works by defining a set narration speed and the user is responsible for manually highlighting or inputting the text that is then narrated for the user. These are all barriers of entry that could be eliminated or limited to some degree.

Another problem is that narration is static, once begun narration continues, in the same style as it always does, without fail, and with little to no input from the user (they can pause, start, cancel or skip to the next part) except just before, with the initial settings.

This project aims to discuss and propose solutions of how to increase interactability between the reader and narration to provide a better experience and more comfortable reading. The main component being an eye tracker as the HCI engine to facilitate the interactability during narration.

The end goal of the project is to provide assistance and to explore the eye tracking as HCI with narration and what the limitations are and how to get the most out of the interaction.

1.1.1 Hypothesis

The hypothesis is that by increasing interactability during narration the experience of using text-to-speech narration can be improved with increased comfort of reading and even text comprehension.

1.2 The how (method)

This project looks at the problem of interactability during narration by first examining, how eye tracking works in theory and then look at applying it during reading which is the end goal of the project. Before going into depth with how to leverage the system to assist with reading some knowledge of, how the eye works in in relation to, how eye tracking works, why text-to-speech narration is used and how reading works, and what can be observed during reading. The study conducted in this project revolves around a program that can adapt to the user during narration. Said program is a set of tools put together in a new form. The program is made using C# with a set of libraries that hook into different existing technologies, and combines eye tracking, file processing, UI, and TTS narration. Narration is used as the main information conveyor and is the primary technology that is iterated upon in. The evaluation is achieved by analysing data from a study conducted, where text comprehension, reading speed (average fixations), and observation during trials will form the backbone of data for the analysis.

1.3 Outline

This project is divided into 6 parts

Part 1 introduces the project. The why, and how of the project is outlined so the project can flow smoothly until the end.

Part 2 this part focuses on contemporary research and belief within eye tracking, reading and reading disabilities. The focus is on linking eye tracking with the different facets of reading and narration by way of text-to-speech. There is also a focus on discussing text-to-speech with its limitations and benefits.

Part 3 looks at the method used to examine the hypothesis and how the method changed over time. The main focus is on the study and the program designed with the hypothesis in mind. Each facet of the method is explored from theory til the practical implementations of said theory.

Part 4 consists of analysing the results generated from the study via the program. It looks at how valid the generated data is, how it can be used, and what information can be gleamed from it.

Part 5 takes a critical look at the program and study in conjunction with the generated data's limitations. The discussion is mainly about the shortcomings and how to remedy them. This part also dedicates time to defining some future works.

Part 6 is the conclusion and answers the hypothesis. The results of the analysis, as well as the discussion, is concluded here to end the project.

2 Theory

Part 2 looks at contemporary and complementary research and how it can build the foundation of understanding for the program that strives to achieve better interactability during narration.

Part 1 outlined in short the complexity of narration and why iterating on this technology is worthwhile. The different aspects of narration with text-to-speech, eye tracking, dyslexia, reading and the eye, must all be individually before being combined.

Section 2.1 discusses the eye. Specifically, how the eyes can move since this is critical for how to understand reading. **Section 2.2** discusses reading. The focus is on how we read and the relation between the eye and reading which is important for dyslexia and eye tracking during reading **Section 2.3** discusses dyslexia. Specifically the impacts dyslexia has on reading, and how it can be detected using eye tracking. **Section 2.4** discusses eye tracking. How it can be used, its potential in HCI, and aspects to be aware of when using eye tracking for research. **Section 2.5** discusses TTS narration. How it can be used, the benefits and limitations. **Section 2.6** discusses these disparate theories and how they culminate into the methodology.

2.1 The eye

The eye is the visual sensor, most importantly for this project is the pupil and cornea. Figure 1 shows a diagram containing the pupil and cornea.

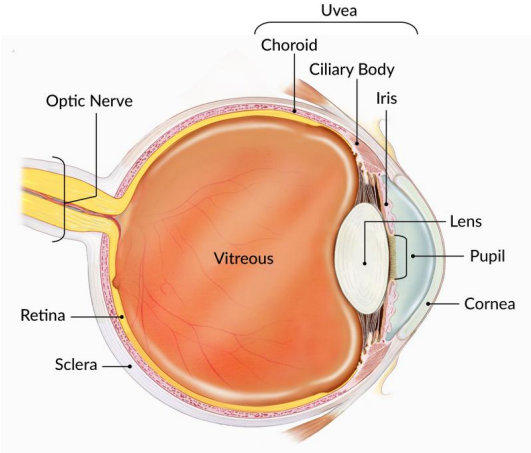


Figure 1: A side view of the anatomical depiction of the eye. Institute n.d.

The pupil is the hole to the inner eye where light gets processed into vision, the cornea is a focusing agent that also reflects some light back, Institute n.d. These parts are focused due to how it is used for eye tracking see section 2.4 for more.

Eye Movements

The eyes move a lot during any activity where visual stimulation is provided. Even if you stare straight ahead and fixate on one point, the eye still moves even if only a little, Farnsworth 2019. There are many reasons for this such as sight being based on movement and by shifting the eye itself with saccades the blind spot can be filled in by the brain.

There area a number of defined eye movements: saccades, smooth pursuit, vergence, vestibulo-ocular movements, and optokinetic response movements / postrotatory nystagmus, Farnsworth 2019 Purves D 2001.

While all of the mentioned eye movements are necessary for the eye to function for the purposes of this project only saccades, fixation and smooth pursuit will be explained in detail.

Saccades

Saccades describes the rapid movement of the eye when switching between fixations. Saccades vary in length with some smaller like during reading and then larger when switching lines or when scanning a room.

Saccades are an integral part of the eyes function, but they are not perfect they take time from the intention of a new target to the actual move. During saccades visual information is not perceived and the brain cannot send new saccade instructions and so the saccade might overshoot the target if it moved during the movement and have to refixate on the new location.Farnsworth 2019

Fixations

Fixation is not an eye movement in of itself, think of it like an absence of big eye movements. As mentioned before fixations are a series of saccades and microsaccades that shift around the fixation point to gather visual information about it. A fixation is arbitrarily defined as a series of saccades that happen close enough together in both time and space i.e. during a time period enough saccades happen within the same general area.

Smooth pursuit

Smooth pursuit is the movement that happens when you follow a point of interest as it travels horizontally or vertically across your visual field. Smooth pursuit can be combined with other types of movement such as vergence to track a point of interest in 3D space since smooth pursuit on it's own can only work in 2D up and down, left and right while vergence is in and out.

Think of smooth pursuit as X and Y and vergence as Z in a coordinate system.

Overview of the eye

A lot of information sharing is done with the use of visual perception, for most of human history oral perception was the key in information sharing. With the advent of writing visual information sharing became a powerful tool that skyrocketed what information humanity as a whole could share between ourselves, since now even deceased humans could provide knowledge indefinitely. It is only later again that we have seen an equalisation of the different methods of sharing information since we can store, visual and auditory on equal footing with electronics. Still the eyes are an incredible feat of organic engineering.

What information can we learn form the eye

There is a lot of information available in the eye, chief among them is the gaze. Other than the gaze, you can gleam information about the following: medical state, gender, identity, emotions, psychological tells, desires, age, physical state, and more, Blain 2021.

Some of these feel mystical in nature, such as being able to tell who someone is only based on their eyes, or to discern emotions like a sci-fy empath.

Since the eye hold so much information some of which we don't want others to learn like our medical information masking said data is an important part of eye tracking research this entails removing things like cornea blood veins or iris patterns which can be used for identification. Typically using the least amount of data as possible is a virtue in scientific research and eye research is no different.

For this paper only gaze points, eye to tracker distance, blinks and eye point data is gathered from the eye. Other data points are willingly shared by participants.

2.2 Reading

Reading is a learned skill, that most humans begin learning when they are kids and master throughout their life with education catered to the expected skill level for that age group. Meaning that generally you are challenged in small increments to build up an understanding of the differing structures and words used for the different levels of reading.

How do we read?

Reading is first and foremost the ability to discern meaning from writing, the writing comes in different styles mostly associated with specific letter system and languages, Essberger 2022. In the western world most writing systems are based on the Latin alphabet with variations from language to language. For this project reading is based on the English alphabet that in turn is based on the Latin one in a left to right grid with spaces between words, i.e. how this paper is written.

In this case reading is done by looking at each word from to right until we form a sentence, now most avid readers do not process each word at a time most process some difficult words faster while skipping over less important words. In conjunction with that we know that sentences come in specific structures and as such we can short cut meaning with very little context. Hanson 2022 Farnsworth 2018

It is quite common for readers to skip words below a specific length and only read the beginning and then the end of a word, since the brain is able to understand and interpolate words and structure due to our own vast experience with words and sentences.

The act of reading is a very strong ability that makes compounding and sharing knowledge possible for humans across the world.

While reading is strongly associated with a specific language the ability to read can be translated from the starting language into a foreign language's scripts with some training due to reading being the association of word to meaning. This also means that reading, understanding and language is very connected.

2.2.1 The eye during reading

During reading the eye moves around the page to different words maybe not even on the same line, this can make it harder to read if you are not able to focus completely on the task and keep superfluous eye movement to a minimum.

The eye moves in different ways during reading and in general has a lot of fixations on words, now all fixations are on one word but between words or on nothing, in general the eye follows the order of the text and sometimes we realise that something was not understood properly and we backtrack to an earlier part of the text. The eye and brain have this ability to return to the current fixture point even if the eye needs to dart around to form the actual image we are experiencing.

Why reading is harder for some

The ability to read is weird, by associating characters and words with sound we can translate something into meaning. Many young kids find reading hard at first, but quickly make associations between what they already know, speech and what they are about to learn, writing. Both produce meaning and in general can contain the same information. Reading and by extension writing is not natural to the brain it is adapted to fit within the brains existing processes such as shape recognition, that intern means that reading and writing is a social construct and not natural, Hanson 2022.

Speech and listening & writing and reading

The reason why some find reading harder is mainly due to lacking exposure to enough associations between words and sounds, the difference being is how much exposure is lacking, Rockets 2022 Hanson

2022. You can also receive conflicting exposure such as words sounding different in other languages even if they look the same and even mean the same, University 2018.

Other reasons why reading is harder for some is motivation, not all children especially find the act of reading interesting or enjoyable enough to prioritise and as with most learned skills you have to practice and maintain your reading ability, Ykema 2017.

2.3 Dyslexia

Dyslexia is a learning disability. Dyslexia is formally describe as an increase in difficulty of associating words/characters with sounds, Dyslexia and Creativity n.d. The Danish system informs that there are degrees of dyslexia from 'easy' to 'hard', and that dyslexia does not make it impossible for someone with the disability to become strong readers, Børne og undervisningsministeriet n.d. As a visual aid think of the degrees of dyslexia as, how much slower a person can run in a race(reading), they will finish but they get there slower than others in a comparable group; age, education and general intelligence.

All checked sources agree that dyslexia is a disability, but with proper tools and education people with the disability can become functioning students and adults. Børne og undervisningsministeriet n.d.Dyslexia and Creativity n.d.

If dyslexia is a learning disability and not something that limits the capability of understanding words, that means that by knowing how to teach people with the disability allows them to learn what is to them much harder.

The problem with dyslexia is that the people with the disability can be very intelligent in other fields, but struggle with the task of reading, and since most tests require reading, leading to many years of a bad association of bad at reading = stupid when in fact, dyslexia is not linked with intelligence, Margaret J. Snowling 2020.

2.4 Eye Tracking

Eye tracking is a powerful piece of technology with many subtle but useful uses. Chief among them is the additional data source that the eyes can provide. With this extra data much can be gained and learned.

New information can be gained and previously understood information can be recontextualised with this additional source of information. When this fact is internalised the power of eye tracking becomes apparent. Just like pictures and videos have gyroscope and gps data your device could detect your eyes focus, it could give you health advice such as "you are too close to the screen" or "you need to rest your eyes due to x and y factor that has been detected in your eyes". Ads could track whether you are looking at them or not and not play unless you have your eyes on it, or apps could optimise their layouts to suite your gaze patterns. Tobii 2022Farnsworth 2018Blain 2021

What is eye tracking

Eye tracking is the act of keeping track of the eye of someone, more specifically where that someone is looking what is commonly referred to as gaze. Most people think of gaze tracking when eye tracking is mentioned and it is the most common goal of eye tracking so the connection make sense.

The gaze indicates, but does not confirm, attention on something and having the ability to track attention is very valuable in a wide array of fields, disciplines and facets of life, "While eye tracking can tell us what people look at and what they **see**, it can't tell us what people **perceive**." Farnsworth 2018. Attention is something that is getting increasingly commodified it makes sense to have a way to track it, Mormann 2018.

A short history of eye tracking

The fascination with the eye stems far back into history many a poem tells of eyes, the tracking of the eye began in the nineteenth century with human observations and then in 1908 Edmund Huey designed a very invasive lens that participants had to use that could help with tracking their gaze. In 1937 Guy Thomas Buswell made the first light based eye tracker the idea of which is still used today.

In 1967 Alfred Lukyanovich Yarbus published a book about how interest and intention shape how the eye moves. The next big development in eye tracking comes with the computer as eye tracking can now be done in real time.

Today there are a number of different ways to track the eyes and new technology is evolving to provide better results.

What information can be gained with eye tracking? Eye trackers come in different flavours some can get most information from the eye some are designed only to track the gaze or some other attribute. There are 3 main values that can be gathered from the eye, the gaze point, distance from the eye to tracker, and eye point. A lot of information can be gathered from these 3 main values and most commercial eye trackers only provide this data.

When talking about eye information some of it is only available through looking at the area surrounding the eye, and some eye trackers are simply not able (hardware limitation) to get data from there while others are able to get information from anywhere though it might be specialised on a software level more on this here 2.4.

Different types of eye tracking

There are different types of eye trackers that vary from how intrusive they are to how they are mounted to the technology used.

There are 2 main categories the peripheral and the repurposing kind, since are designed to interface with computers you can either develop a piece of hardware or develop software that can repurpose hardware such as a webcams.

PCCR (infrared)

PCCR stands for Pupil Center Corneal Reflection and describes the technique of looking at the eye with near infrared cameras that make it easy to find the the different points of interest in and around the eye. Due to the nature of this type of eye tracker there are some inherent benefits such as not relying on external light, Farnsworth 2018. They also have very high precision due in large part to how clean the eye is visually when viewed under infrared light, this can be seen on the figure 2.

Webcam (visible light)

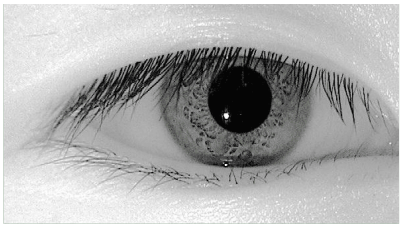
Some eye trackers are not actual hardware + software but only software meant to interface with a camera such as a webcam. This technology is still very new but carry immense potential since so many device have a webcam or front facing camera. Now there are some disadvantages from using visible light for eye data collection since the image changes depending on the environment, surrounding light, and the eyes natural reflection making this type of eye tracker environment dependent, Jensen 2019.

Screen mounted

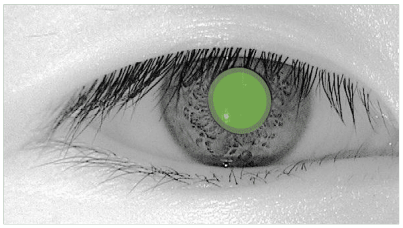
Screen mounted or screen based eye tracking is focused on tracking eyes in relation to a computer screen. These trackers have the benefit of not attaching anything to the trackee only requiring them to be within a certain 'headbox' for the tracking to work. This type of eye tracker is optimised for close screen related tasks and so are ideal for this type of task.

Head mounted

There are different types of head mounted eye trackers, such as glasses, or a camera crane. Head mounted eye trackers carry the inherent benefit of allowing free head and body movement for the trackee



Iris illuminated by infrared light.



Iris illuminated by infrared light, with iris highlighted (similar to how eye trackers follow eye movements).

Figure 2: This figure depicts an eye when photographed with an infrared camera, Jensen 2019.

this makes it possible to apply eye tracking to situations out in the world instead of confining it to one location such as with screen mounted eye tracking.

VR eye tracking

A VR headset is basically cameras and a screen that is really close to the eye making it ideal for eye tracking. With the endless scenarios that could be designed for VR this type of eye tracking could allow for a very diverse set of data with little requirement from the trackee. This also has incredible potential for HCI (human computer interaction) in VR. If done correctly you can get the benefits of both screen based and head mounted eye trackers due to the nature of a VR headset.

2.5 Narration

Narration comes in many parts, this page has a great overview of how it is done within audio descriptions, Canada n.d. In general narration as understood for this paper is the act of putting sounds to written words for the express use of those who don't want to, can't, or struggle to read.

Narration comes in two categories, by humans or by a 'voice' (synthesised). At the moment narration done by human have in general more emotion and a better understanding of context, while synthesised voices are uniform and lacks context.

The main benefits of synthesised voices is for the express purpose of text-to-speech (TTS). TTS is on demand narration of text and is most commonly used by those who suffer from disabilities, Chapuzet 2021. Originally designed for the visually impaired the technology has seen use with dyslexic people as their increased difficulty in decoding written words into sounds and meaning is lessened when you can hear the word, Todd Cunningham 2018 Sara G. Wood 2018.

TTS narration, has the following properties: consistent, on-demand, lacking context, set narration speed, and inflexible while narrating.

Some of these properties are shared with all processed audio bytes since speech does not sound natural if one word is spoken at a time, which results in a sentence being recorded or processed at a time leading to inflexibility.

Reading with a narrator

When a narrator is reading aloud can be very hard to not follow along with the narration even if your natural reading speed is lower or higher than the narration speed. Paola Bonifacci 2021 Sara G. Wood 2018. This helps with reducing mind wandering aka getting distracted away from the current point of narration.

Reading along with a narrator is show to provide an increase in attentiveness and comprehension, Paola Bonifacci 2021 Sara G. Wood 2018. The fact that people with a learning disability also show these same increases, is why narration is so helpful and why narrators are provided to people with dyslexia.

2.6 Culmination of theory

All this theory is culminating in the method, where how reading is done, the benefits of TTS narration, the HCI component of eye tracking can be leveraged, and understanding the limitations of dyslexia can all be combined into a satisfying product and study to evaluate the product.

3 Method

Part 3 is about turning theory into practise and the steps that would be completed for the eye tracking narration interaction link. As described in the culmination section found in part 2 a lot of understanding is necessary to iterate on existing software. **Section 3.1** is about how the system is built. The systems purpose is to link narration and eye tracking, by understanding the gaze in conjunction with the current narration it can adapt narration to fit. The data extracted from the eye tracker is the limiting factor when trying to adapt. The section is split into subsections with: feature interplay, data logging, text processing, and adaptive & constant narration. **Section 3.2** is describing the study designed to test the system. The setup, procedures, and expected outcome are outlined.

3.1 The system

This section is about how the system is built. The system is built using C# 6.0 with the following frameworks: NPOI 2.5.5, System.Speech 6.0.0, Windows Forms, EyeXFramework 1.5.466.0, Tobii.EyeX.Client.Net20 1.5.466.0. The system uses windows exclusive functionality and as such cannot be used on any other operating system.

The eye tracking reading narration system is used by setting up a eye tracking profile then starting a program, that can track the user, has a file loader, text viewer, a narrator, and a data logger. The program loads an example text file and displays it, the user can then load a new file and begin narration of the new file. When a new file is chosen the logger begins logging data from the eye tracker of later evaluation.

3.1.1 Phases

There are distinct phases to using the system: **Phase 1** a eye tracking profile is setup using the EyeX's built-in calibration software. **Phase 2** the program is started and waits for a user presence. **Phase 3** a text is chosen, the text is then processed to prepare for narration and highlighting. **Phase 4** the chosen text is displayed and the user can interact with it by either reading it or getting it narrated. All gaze events are logged regardless of the chosen action.

The next phases only happen for narration: **Phase 5** the narration uses the processed text to handle visualisation and narration. **Phase 6** the gaze of the user is accumulated during each narration segment to use for narration adaption **Phase 7** the adaption and narration continues until the text file is over and the data logging finishes when the narration does. A new file can be chosen restarting the process.

3.1.2 Feature interplay

The main focus of the system is to combiework, and lastly WinForms for GUI and tying everything together.

Most calculations are done in the GUI, since they need to be handled when the user initiates the different functionality.

The flow is: User intention -> user interaction -> GUI handles the interaction -> GUI fetches necessary data -> GUI showsne various features, the TTS synthesis from System.Speech, text processing from .docx files with NPOI, eye tracking data hook with EyeXFram result of interaction.

3.1.3 Data logging

Data logging is happening from phase 4 until 7. The data logging is only for eye tracking data and consists of the system time, eye tracker timestamp, gazeX, gazeY, eyeposX, eyeposY, eyedistance, blinked (yes/no).

The data logging is setup to generate a new line each time the eye tracker sends new data in the form of an event. Eye tracking data is delivered in the form of events that trigger a set of interactions such as the gaze data culmination during phase 6 (see section 3.1.1). This approach make sure that all data points are saved no more and no less.

3.1.4 Text processing

NPOI provides tools to extract text from .docx files. During phase 4, the raw text is analysed for contextual stopping points for each narration section, the system looks for ".", line ends and file ends and chooses whichever appears next. The contextual markers can be expanded according to whatever design is needed "" like quotes or ";" semicolons.

Each section that the process finds are logged in an array which makes it flexible for both highlighting and narration. The text processing is essential for a good narration experience if the sections are too long or too short the system either can't adapt or the sentences becomes choppy.

3.1.5 Adaptive & constant narration

The eye tracking is used both to provide data for analysis but also to provide assistance with the narration. During phase 6 all gaze data points are accumulated and when the section finishes the average of he data points is calculated, that average is then computed to a character point within the text, if the location within the text is below 40% of the segment it slows down if it is above 55% it speeds up and otherwise remains the same. So the user can influence the narrator by how they are reading, this in turn means that has to break away from the narrator to enact change.

To have data points to compare to between normal narration and adaptive narration, constant narration was also implemented. It simply always has the same narration speed regardless of how the user moves their eyes during narration.

The narration is handled by the System.speech synthesizer and uses the voice "Microsoft Helle" as it is the only Danish voice available on Windows by default.

3.2 The study

A study was conducted to examine the systems ability to prove the hypothesis. The studies goal was to examine participants while reading under different conditions. To set a baseline the participants were asked to read a text by themselves, then they would be asked to read with narration sometimes constant other times adaptive.

The study was conducted in Danish meaning all the texts, information given to the participants, the pre- and post-study questionnaires, and the interviews were in Danish. All the for this project have been translated to English for the convenience of the reader.

The setup

The setup for the study is an office desk with a screen that has the eye tracker mounted below the screen. The participants can interact with the program by pressing keys on the available keyboard. The narration can be heard from a nearby speaker.

The strictly necessary elements of the setup is a desk, chair, computer, keyboard, eye tracker, and screen (with eye tracker mount). Since the area is a home office there were things that were not strictly necessary such as a desk lamp, pictures, and a secondary monitor (though it was turned off).

Procedures

This is the step by step walk through of how the study was conducted. The step 0 was ideally done before showing up while steps 1-11 was to be conducted with a study admin in the room.

The study was conducted manually meaning that choosing the next file was done by a study admin and the post text read questionnaires were asked by a study admin.

step 0: participant fills out the pre-study questionnaire

step 1: The participant is introduced to the experiment, that they are to read 5 files one withouthelp and the rest with narration. The participant is informed about the adaptive or constant narration styles only that narration would be turned on and that they should try to read at their own pace.

step 2: The participant's eye tracking profile is setup

step 3: The participant is told to read the first file and announce when they are done

step 4: the program is reset for the next part.

step 5: the participant is asked questions to make sure they have read the text, their answers are recorded

Step 6: the participant is put into one of 2 groups, group 1 starts with the adaptive speed and group 2 starts with constant speed, they then switch type between each document.

Step 7: the participant is told to read one of 4 files with narration on, they are informed to read at their own pace to the best of their ability.

Step 8: the participant is asked questions regarding the text they read.

Step 9: the narration style is changed to constant from adaptive or vice versa according to which group they are in

Step 10: repeat step 7 through 9 until all 4 text have been read by the participant

Step 11: Ask the post study questionnaire questions to the participant

The post-study questionnaire consisted of the following questions:

* Was the software intrusive?

* Side effect of the eye tracking?

* Uncomfortable using the eye tracking?

* Did the software augmentations provide any assistance?

* What did you think about the text-to-speech?

* Could you see yourself using something like this?

* Do you have any suggestions for changes/improvements?

* Was there anything else you either disliked or enjoyed about the experience?

Expected outcome

The expected outcome of the study is raw eye tracking data that could indicate whether adaptive is an assistance or a hindrance to the participants. Some participants have dyslexia and it is expected that they are slower to read and they may or may not have lower comprehension during constant especially if the speed of it is too fast compared to their preferred reading/narration speed.

4 Analysis

Part 4 presents the analysis and results of the study with system outline in Part 3. An analysis of the generated data from the study/system and some observation related to the study and its participants.

Section 4.1 is about who participated in the study. The two groups dyslexic and reading neurotypical.

Section 4.2 is all about the general observations made in relation to the study. Mainly focused on insights related to the process, the study, and dyslexia

Section 4.3 presents the an analysis of the data. There is a focus on average fixations and length in relation to the different markers tracked in the study such as reading disabilities, narration style, per participant, and per text.

4.1 Participants

A version of the study has been conducted with 16 people, though only 7 follow an analysable setup, read more in section 6.

The data analysis will be done with N = 7. The 7 participants contain 2 with dyslexia (participant 1 & 4) and 5 reading neurotypical.

All participants are in their early to mid 20'ties, have or is studying at a university, and study or have studied a STEM subject. In the pre-study survey no one had tried using an eye tracker before. Of the 7 only one had need of glasses. 4/7 indicated that they have read 1-3 books in their spare time and 4/7 answered that they have read 4-8 books related to their professional life.

Unfortunately there is not equalibrium between the two groups and with this sample size people with dyslexia is over represented with 2/7 being 29% which is almost 50% higher than the most generous dyslexia estimation within a population. In addition the two dyslexic participants were randomly assigned the same narration style pattern, and they are very close to time spent for each text see table 1.

4.2 Observations

In general very few realised that the narration could adapt to their reading speed, there are two conclusions from this either it is too subtle and it won't have any effect or it is having effect yet the adaption is subtle enough that the participants don't realise it.

Some participants expressed that their reading was impacted by the narration while others didn't expressly comment on it. The impact was describe as reading in defiance of the narration was hard or tedious enough that 'giving up' and reading along was easier. This mindset makes adaption during adaptive narration less likely since they are not forging ahead or lacking behind as much as they 'should'.

Both dyslexic participants spent longer on text than the other participants not by a huge margin, but enough to be consistent. The breakdown of time spent per text per participant can be seen on table 1.

Participant	Text	Time (ms)	Style	Reading Diagnosis
Participant1	Blaa	59960	Adaptive	Dyslexia
Participant2	Blaa	57773	Adaptive	Nothing
Participant3	Blaa	54117	Constant	Nothing
Participant4	Blaa	60518	Adaptive	Dyslexia
Participant5	Blaa	55577	Constant	Nothing
Participant6	Blaa	48999	Constant	Nothing
Participant7	Blaa	42403	Adaptive	Nothing
Participant1	Groenne	64843	Constant	Dyslexia
Participant2	Groenne	57720	Constant	Nothing
Participant3	Groenne	58865	Adaptive	Nothing
Participant4	Groenne	63262	Constant	Dyslexia
Participant5	Groenne	44194	Adaptive	Nothing
Participant6	Groenne	45219	Adaptive	Nothing
Participant7	Groenne	59608	Constant	Nothing
Participant1	Intro	61385	Without	Dyslexia
Participant2	Intro	31258	Without	Nothing
Participant3	Intro	33274	Without	Nothing
Participant4	Intro	64035	Without	Dyslexia
Participant5	Intro	40990	Without	Nothing
Participant6	Intro	26976	Without	Nothing
Participant7	Intro	25211	Without	Nothing
Participant1	Rosa	51572	Adaptive	Dyslexia
Participant2	Rosa	53547	Adaptive	Nothing
Participant3	Rosa	48590	Constant	Nothing
Participant4	Rosa	51181	Adaptive	Dyslexia
Participant5	Rosa	48562	Constant	Nothing
Participant6	Rosa	54413	Constant	Nothing
Participant7	Rosa	38986	Adaptive	Nothing
Participant1	Violette	88051	Constant	Dyslexia
Participant2	Violette	71704	Constant	Nothing
Participant3	Violette	61458	Adaptive	Nothing
Participant4	Violette	78115	Constant	Dyslexia
Participant5	Violette	51837	Adaptive	Nothing
Participant6	Violette	54751	Adaptive	Nothing
Participant7	Violette	74027	Constant	Nothing

Table 1: Participant ordered by text with time spent, their narration style, and reading diagnosis

4.3 Data from the study

The data from the study is been analysed with average fixation counts and lengths. For this analysis a the data was processed using Ogama with the following settings: 1920x 1080y screen, 60Hz gaze sampling rate, 20px maximum distance for a fixation with a 5 samples minimum, the fixation detection ring buffer was 31, and it should merge consecutive fixations within max distance into one fixation. From table 2 both average fixation counts and lengths are shown. The length of a fixation is in general lower than the fixation count, with the notable exception of participant 4, who expressed that their reading was very slow during the study.

What is not show on table 2 is that each participants fixation length is pretty static, the variations

between the lowest and highest fixation lengths isn't that great the same can't be said for fixation count which saw larger swings in values.

Participant	AVG fixation count	AVG fixation length
1	227.25	164.7152858
2	216	199.612939
3	191.75	154.3508247
4	210	239.1142516
5	182.25	171.1247006
6	192.75	178.4461531
7	199.5	184.948693
avg	**202.7857143**	**184.6161211**

Table 2: Participants and their average fixation counts and lengths, with narration only texts

As can be seen on table 3, there is a in general a large difference between the dyslexic and neurotypical groups, the only outlier is the Rose text which both dyslexic participants had a very low fixation count on. Though the data doesn't really support this low difference as one of the dyslexic participants (p1) had trouble remembering the Rose text as can be seen on table 4.

text	dyslexia	neurotypical	diff
B	210.5	189.8	20.7
G	224	195.8	28.2
I	203.5	124.4	79.1
R	165.5	169.4	3.9
V	274.5	230.8	43.7
avg	**215.6**	**182.04**	**35.12**

Table 3: Average fixation **counts** per text within the two groups with an absolute difference column and overall average row

Table 4 is to be taken as an indicator and not as an absolute fact, there are a couple of reasons why mainly that some of these questions have nuanced answers and multiple ones and are worth more points and the max points for each text is not the same. The questions that generated these results were used to make sure that the participant had read the text. The questions can be seen here in this section 8.1 in Danish.

participant	intro	blå	grønne	rosa	violette
1	6	5	8	1	7
2	8	7	7	11	10
3	9	6	7	7	7
4	10	6	7	9	10
5	10	7	9	13	12
6	11	7	9	8	9
7	11	8	6	7	8
max	16	10	10	13	12

Table 4: participants and their comprehension scores for each text, blue means adaptive and yellow/gold means constant.

Table 5 shows that again the dyslexic group spent longer on average for their fixations, this could indicate that they need longer to decipher the words for each fixation which is very likely given the nature of their diagnosis see section 2.3 for more info.

text	dyslexia	neurotypical	diff
b	202.9386	178.0533	24.88531
g	201.5722	176.7333	24.83898
i	226.0616	163.679	62.38255
r	207.556	175.2819	32.27412
v	195.5923	180.7183	14.87401
avg	**206.7441**	**174.8931**	**31.851**

Table 5: Average fixation **length** per text within the two groups with a absolute difference column and overall average row

Table 6 has some very interesting data points, mainly that the Blå and Rosa texts did not see a very large difference from constant to adaptive while Grønne and Violette saw a large difference. This could mean that the Grønne and Violette texts adaptive leads to less needed fixations, what is most likely is that the two dyslexic participants were reading with constant narration style for Grønne and Violette and adaptive for Blå and Rosa.

text	adaptive	constant	difference
b	195.25	196.3333	1.083333
g	184.6667	218.25	33.58333
r	166.75	170.3333	3.583333
v	204.3333	272.5	68.16667

Table 6: The average fixation **count** for each text in the different narration styles and the absolute difference between them

Table 7, the avg. fixation length for each text with comparing adaptive and constant goes in a checker pattern of which is higher. The pattern matches that of group 1 which has both 4/7 participants and both dyslexic participants.

text	adaptive	constant	diff
b	202.2897	162.3283	39.96138
g	164.2624	198.5059	34.24341
r	193.4271	172.6042	20.82288
v	172.7006	194.1685	21.46793

Table 7: The average fixation **length** for each text in the different narration styles and the absolute difference between them

5 Results

It is hard to reach any real conclusions with such a biased and low sample size data set. There are inklings that adaptive narration helps with reducing average number of fixations. The average fixation length is very undecided as the difference comes in pairs Blå and Grønne, and Rosa and Violette and swings from lower fixation lengths to higher. If the factor of just as good comprehension with less time spent fixating was positive in comparison to dyslexia some conclusions could be made, but the bias is felt here.

6 Discussion

Part 5, discusses the limitations of the system and study, with a focus on what could be done to remedy the bias and non-conclusive results.

Section 6.1 discusses the limitations within the system and what could have been done differently as well as some future improvements.

Section 6.3 discusses the study and the bias within the data set and what could have been done differently to make conducting the study easier.

Section 6.4 discusses where this avenue of research could be taken in the future. Some interesting facets of this research.

This project tried to develop an iteration upon an already existing technology that being the TTS narration by increasing the interactability during narration the thought was that this could create a better experience.

The problem is there were a number of different approaches that could have been taken and the narration speed adaption is the simplest and unless it is iterated upon can only serve as a calibration tool for setting the default speed of conventional TTS narrators.

Now the above statement might not be true but there was way of knowing if the data that indicated that adaptive helped was not biased due to the low sampling size and over representation of dyslexic in the participation group.

6.1 System limitations and improvements

This section is dedicated to how the program could evolve to provide more assistance to it's users. There are different factors that the original implementation did not account for and features that could not be ready in time for the study.

In general there are a number of oddities with the program, such as the text taking up the full width of the screen, making for an atypical reading situation combined with making it more likely that participants would move their head resulting in less accurate data due to the nature of screen mounted eye trackors see section 2.4 for more.

The system and the study was only matter of executing a set of tasks in succession, meaning that a program was not designed to make the study as easy as possible to conduct.

The study was setup in such a way that the narration dictated when the trial for that text ended, this could have been remedied by allow the participant to end a trial early when they had finished. It is a matter of what you want to test for, the study tested for single read, speed with a flavouring of comprehension assertion. If you allowed the participant to re-read section with the narrator or gave them a time limit and said comprehend as much as possible you could run different types of studies that might or might not be more suited to proving the hypothesis.

The programs logging was focused on eye tracking data only but there was a number of different aspects that could be interesting to know such as the narration speed, when a section was finished being narrated the attention placed on a narrated segment for the purposes of area of interest calculations.

One interesting change would be to implement rereading by utilising the gaze to indicate what word they are looking at while the users is holding down a key and then when they release it the narration begins from the point/word on wards. This would increase interactability for the purposes of comprehension and to test for this the study would have to change.

There are many interactions that eye tracking can make easier that could help facility interactability during narration.

Here are some design ideas for alternate functionality and how you would go about implementing them.

6.2 Alternate program features

Panning Read A panning read is the idea that the user could pan their eyes over some text and then get that text read for them. There are a couple of different ways to handle a panning read, there is the trailing read that queues up words to be narrated as the eye passes, the real time read where each word that the gaze passes over is read, the predictive read where after the initial eye movement it predicts the next word to be passed over and can queue it.

There are some problems that need to be addressed for panning read one of which is the text-to-speech engine, which functions best when it has a set of words to read instead of single words. Another is eye movement, a pan is something the eye does but the typical pan, slowly moving your eyes from one side to another which is hard, is done in rough movements, the eye can also chase something like a moving vehicle and in general a smooth pursuit 2.1 as it is called provides a more **smooth** set of data points provided the eye trackers frequency is high enough. Regardless of how to trigger the pan some interpolation is needed since some small words are bound to be skipped.

With all these things in order the different types of panning read become clearer, trailing read solves the single word problem by caching words to then narrate. Real time read is the weakest implementation since it solves non of the issues since it is always focused on narrating the current word, this could use interpolation to a very minor degree. Predictive read would have to guess which sections the user would want to read, the predictive read might start as trailing but then slowly move towards matching the user to become as close to a real time read as possible while still making the narration sound natural.

Implementing a simple panning read 'algorithm': snake read

The simplest form of panning read is a starting point and then like the snake from the snake game you add on by eating the next fruit/word. You start somewhere we can call that the seed which also starts as the 'head' and then you can only get the next word either backwards or in front and you wait to 'consume' the next word until the current head word has been pronounced. The next word to consume is decided by where the user is looking this has some considerations ideally the lines of text would separated by a fair amount of line spacing since the accuracy of the eye tracker is not perfect, so by prioritising the current line unless the difference between current word location and gaze point is vast enough, or the current head is close to the end of the line to ensure that switching lines is still possible.

The problem is that single word pronunciation does not carry the same flow and makes the voice sound even more synthetic than usual. To combat this issue you could make x amount of characters until a word boundary each 'fruit' in the snake algorithm which would make flow better and probably carry less strain on the user but this approach makes it so the user might read too much.

A benefit from this is that a new seed can be picked from anywhere and so a user can start a new 'snake' or chain anywhere and get a word they missed repeated to them. The snake read can even over time increase the length of subsequent sections until it switches over to the section based version found in the adaptive/constant narration calculations.

Repeat section

It is very common when reading to lose our focus on the text and having to read it again or if a section is particularly hard rereading the section would prove beneficial for understanding its meaning. This could have been implemented in different ways one simple one is to cue a reread if the user pressed a button, this reread could be handled differently depending on the adaptive narration, it could stop the

adaptation temporarily or it could automatically slowdown the reread so that the user could get all the information that they wanted.

Optimised section calculations

Currently the sections are calculated after first found of the following ".?point, end of line or end of file. It would however be beneficial if particularly long sections would be split up into parts. This would serve two functions it is easier for the user to focus on smaller sections than if the section spans multiple lines, the other being that the adaptive narration would have more chance to adapt.

In addition many eye tracking studies define areas of interest and each section could be calculated as such potentially identifying problematic sentences or sections that need changing.

Word difficulty mapping

Some words are longer or more difficult to understand making it harder for the reader to comprehend them in time. By knowing which words are harder the program can adapt narration speed to give the user the time they need to understand such words before continuing on with the narration.

6.3 Study bias and sampling size

There were 7 viable data sets from participants due to how the studies structure changed overtime due to getting a better understanding of what the study needed to contain to generate useful date. The study did not have a large sample size and as mentioned in section 4 this creates bias especially when both dyslexic participants were assigned the same and larger narration style pattern group. This bias is not something that can't be overcome or accounted for if the sampling size was larger.

The sampling size could have been larger if the program was ready earlier and more time was spent designing the study.

The study was also conducted with no mention of how the adaptive narration worked meaning that the participants were in the dark when it came to utilising it. It is not certain if this blind style study is a benefit or a hindrance there are multiple sides of this, you get participants who don't realise that they are reading faster and you get others who just follow along with the narration even if they could read faster or should read slower due to them not realising what is available to them. There is an argument to be made that informing the participant of the systems capabilities would help increase interactability since the participant could actively engage with the narration instead of accidentally/passively as was the case in the study. The counter argument is that this increases the amount of variables that the study has to account for namely learning or relearning how a tools works meaning that even more data bias has to be accounted for.

6.4 Future work

There are many avenues that this research could go down, one path is to try another style of interactability during narration like the reread outlined in section 6.1 by changing the studies goal to being increasing the ease of which the participants achieve comprehension in comparison to another TTS narrator. This change would shift the focus away from dyslexia and would require people who use TTS and introducing them to the new tools functionality and getting their input on the tool and then having a group of novices that could try both tools to see what gives the best results.

Another approach is to implement free form narration highlight like discussed in the panning read section 6.2, again this would change the nature of the study to be about the ease of using narration tools in day to day scenarios and for the best result would have to be a long form study where the participant has access to the tool on a daily basis.

The last future work is to implement narration with eye tracking glasses to allow for on the fly narration, this would increase the flexibility narration and would the study would most likely be in comparison to pen readers. This is form of AR (augmented reality) which is not very common and as such the idea might not be feasible with current technology.

7 Conclusion

This projects goal was to increase intractability during TTS narration which would lead to increase in comfort and ease of use. Comfort is hard to define and when the data had as much bias and low sampling size (N=7) any conclusions could not compensate for the datas low quality.

The discusses then focused on what type of work could be done in the future to better focus the study towards the hypothesis and which improvements and changes the system could be implemented for the future.

The conclusion for this study is that more and varied data is required to eliminate the bias and that while primitive this style of narration assistance could still have merit even if this project could not prove it.

The possibilities of an eye tracking in HCI (human computer interaction) is a worthwhile pursuit.

References

[And10] Harris Andrea. *The Human Brain is Loaded Daily with 34 GB of Information*. 2010. URL: https://www.tech21century.com/the-human-brain-is-loaded-daily-with-34-gb-of-information/.

[Bla21] Loz Blain. *Eye tracking can reveal an unbelievable amount of information about you*. 2021. URL: https://newatlas.com/science/science/eye-tracking-privacy/.

[Bør] Socialstyrelsen Børne og undervisningsministeriet Nota. *Ordblindhed*. URL: https://www.ordblindhed.dk/om-ordblindhed/ordblindhed.

[Can] Described Video Canada. *Can Synthetic Voices be Used for Audio Description?* URL: https://describedvideocanada.com/can-synthetic-voices-be-used-for-audio-description/.

[Cha21] Aurélien Chapuzet. *Speech synthesis (TTS), how to use it and why is it so important ?* 2021. URL: https://vivoka.com/how-to-speech-synthesis-tts/.

[Cli17] Mayo Clinic. *Dyslexia*. 2017. URL: https://www.mayoclinic.org/diseases-conditions/dyslexia/symptoms-causes/syc-20353552.

[Col21] Mississippi College. *Importance of Reading for Elementary Students*. 2021. URL: https://online.mc.edu/degrees/education/med/elementary/reading-for-elementary-students/.

[DC] The Yale Center for Dyslexia and Creativity. *What Is Dyslexia?* URL: https://dyslexia.yale.edu/dyslexia/what-is-dyslexia/.

[Ess22] Josef Essberger. *What is Reading?* 2022. URL: https://www.englishclub.com/reading/what.htm.

[Far18] Bryn Farnsworth. *Eye Tracking: The Complete Pocket Guide*. 2018. URL: https://imotions.com/blog/eye-tracking/.

[Far19] Bryn Farnsworth. *Types of eye Movements [Saccades and Beyond]*. 2019. URL: https://imotions.com/blog/types-of-eye-movements/.

[Han22] Joe Hanson. *The Unbelievable Science of How We Read*. 2022. URL: https://www.youtube.com/watch?v=Wt7rR0MCYsg&list=WL&index=22&ab_channel=BeSmart.

[Ins] National Eye Institute. *How eyes work*. URL: https://www.nei.nih.gov/learn-about-eye-health/healthy-vision/how-eyes-work.

[Jam11] Chaitanya Baru James E. Short Roger E. Bohn. *How Much Information?* 2011. URL: https://web.archive.org/web/20110515161545/http://hmi.ucsd.edu/pdf/HMI_2010_EnterpriseReport_Jan_2011.pdf.

[Jen19] Ole Baunbæk Jensen. *Webcam-Based Eye Tracking vs. an Eye Tracker [Pros & Cons]*. 2019. URL: https://imotions.com/blog/webcam-eye-tracking-vs-an-eye-tracker/.

[LDO04] LDOnline. *Dyslexia: What Brain Research Reveals About Reading*. 2004. URL: https://www.ldonline.org/ld-topics/reading-dyslexia/dyslexia-what-brain-research-reveals-about-reading.

[Mar20] Kate Nation Margaret J. Snowling Charles Hulme. "Defining and understanding dyslexia: past, present and future". In: *Oxford Review of Education* 46.4 (2020), pp. 501–513. DOI: https://doi.org/10.1080/03054985.2020.1765756.

[Mil17] E. Ce Miller. *How Much Time You Should Actually Spend Reading Each Day, According To Science*. 2017. URL: https://www.bustle.com/p/how-much-should-i-read-a-day-these-studies-have-some-answers-on-the-ideal-amount-of-time-70040.

[Mor18] Milica Mormann. *The Economics Of Attention: A Valuable Commodity That's Hard To Measure*. 2018. URL: https://www.smu.edu/cox/Learning-Culture/Research-Papers/20180501_Mormann.

[Pao21] et al. Paola Bonifacci Elisa Colombini. "Text-to-speech applications to reduce mind wandering in students with dyslexia". In: *Journal of Computer Assisted Learning* 38.2 (2021), pp. 440–454. DOI: https://doi.org/10.1111/jcal.12624.

[Pur01] et al. Purves D Augustine GJ. "Neuroscience 2nd edition". In: Sunderland (MA): Sinauer Associates, 2001. Chap. 20.3. URL: https://www.ncbi.nlm.nih.gov/books/NBK10991/.

[Rev22] World Population Review. *Literacy rate by country*. 2022. URL: https://worldpopulationreview.com/country-rankings/literacy-rate-by-country.

[Roc22] Reading Rockets. *Why Some Kids Struggle*. 2022. URL: https://www.readingrockets.org/helping/struggle.

[Sar18] et al. Sara G. Wood Jerad H. Moxley. "Does Use of Text–to–Speech and Related Read–Aloud Tools Improve Reading Comprehension for Students With Reading Disabilities? A Meta–Analysis". In: *Journal of learning disabilities* 51.1 (2018), pp. 73–84. DOI: https://doi.org/10.1177/0022219416688170.

[Sim22] Tom Simonite. *Synthetic Voices Want to Take Over Audiobooks*. 2022. URL: https://www.wired.com/story/audiobooks-synthetic-voices/.

[Tob22] Tobii. *Eye Tracking*. 2022. URL: https://www.tobiipro.com/blog/what-is-eye-tracking/.

[Tod18] Harrison McNaughtan Todd Cunningham. *Text-to-Speech (TTS)*. 2018. URL: hhttps://www.readingrockets.org/article/text-speech-tts.

[Uni18] Florida Atlantic University. *Children take longer to learn two languages at once compared to just one – don't fret*. 2018. URL: https://www.sciencedaily.com/releases/2018/08/180823092018.htm.

[Vit22] Vitec. *CD-ord*. 2022. URL: https://www.vitec-mv.com/dk/produkter/cd-ord/.

[Yke17] Sophia Ykema. *Increasing Reading Motivation in Young Struggling Readers: Tips for Teachers*. 2017. URL: https://www.uwo.ca/fhs/lwm/teaching/dld2_2017_18/Ykema_ReadingMotivation.pdf.

8 Appendix A

8.1 Questions for each text

PDF on next page.

Sektioner

Template:

[Points worth] question? – answer

[! x points per correct answer] question? – answer list[]

Intro

[1] Hvor mange segmenter? – 5

[!1] Hvilke farver? – [blå, grønne, rosa, violette, og grå]

[!1] Hvilke procenter blev nævnt? – [25%, 20% og 10%]

[!3] Hvilke farver repræsenterede dem som var idealister? – [grønne og rosa]

(1) Hvad var det grå segments profil? – De har ikke nogen klar profil

Blå

[!1] Hvad er der antaget om det blå segment? – at de er [moderne og materialistiske]

[!1] Hvad ligger det blå segment stor fokus på? – [karriere og tjene penge]

[2] Hvordan skal folk belønnes ifølge det blå segment? – hvis man gør en indsats

[1] Flest mænd eller kvinder? – mænd

[3] Forklar hvad en højreorienteret person er? – konservativ - social og økonomisk ulighed er uundgåeligt, noget med politik evt. mere.

Grønne

[!1] Hvad regner man det grønne segment for at være? – [moderne og idealistiske]

[1] Hvilket andet segment bliver det grønne sammenlignet med? – det blå

[!1] Hvilke interesser har man i det grønne segment? – [kultur, skønlitteratur og naturoplevelser]

[1] Flest mænd eller kvinder? – kvinder

[3] Spiller dette segment godt med andre? – ja og nej, godt med dem der ligner dem, men skepsis mod folk uden for deres gruppe.

Rosa

[!1] Hvad regner man det rosa segment for at være? – [traditionelle og idealistiske]

[!1] Hvem er i det rosa segment? – [ældre, lavere indkomst, lavere uddannelse/ufaglærte]

[!1] Hvad går de op i? – [familien, haven, indkøbsture og ugentlige lodtrækningen]

[!1] Hvad vægter dette segment højt? – [traditioner og tryghed]

[!1] Hvilke partier stemmer dette segment ofte på? – [Dansk Folkeparti DF, Socialdemokratiet A]

Violette

[!1] Hvad regner man det violette segment for at være [hvilke overordnet kategorier falder de ind under]? – [traditionelle og materialister]

[!1] Hvilke eksempler bliver der givet om de traditionelle træk? – [havearbejde, gør-det-selv-byggerier og fiskeri]

[2] Hvad har dette segment stadig til trods for deres store vægt på traditioner? – [forbrugstrang]

[!1] Hvilken køn og aldersgruppe er mest repræsenteret i dette segment? – [unge, mænd]

[!1] Hvilke 3 partier svinger dette segment i mellem? – [Dansk Folkeparti DF, Venstre V, og Konservative C]

8.2 The pre-study questionnaire

The list is as follows:

- Consent to gather data from them with eye trackers
- Name (will be anonymized for the report)
- Age
- Sex
- Disabilities (with focus on mental, and eye related disabilities)
- Self reported reading level
- Familiarity with eye trackers
- Familiarity with text-to-speech/narrators
- Consent to being recorded
- Eye sight issues (glasses / contact lenses)
- books read in the last year

The questions from the pre-study survey (in Danish):

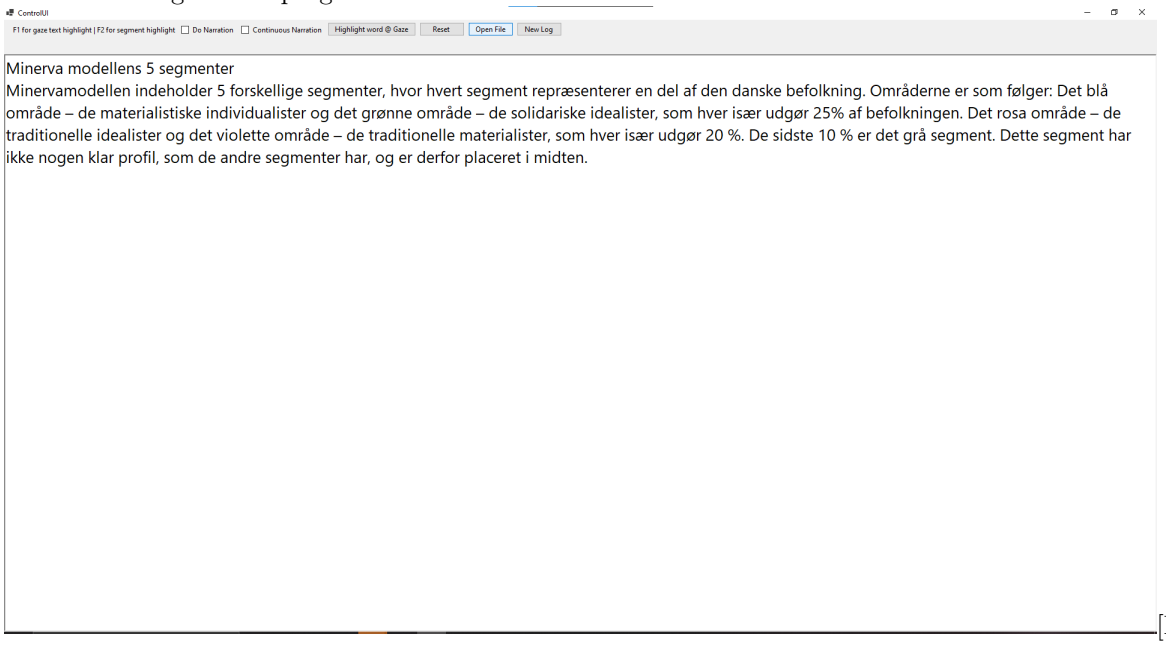
Tilladelse til at gemme og bruge information som er anskaffet ved hjælp af eye tracker. Dette data er: Dine øjnes distance fra eye trackeren, x og y position af dine øjne og der hvor du kigger henne på skærmen.

Tilladelse til at optage interview/samtalen mellem dig og forsøgsansvarlig? Hvad er dit navn? Hvornår blev du født? Hvad er dit køn? Har du synsbesvær? (briller, kontaktlenser, etc.) Har du nogen mentale handikap? (Autisme, ADHD, ordblindhed etc.) Hvad er dit dansk læse niveau? (vælg det lavest passende niveau) Har du brugt / kender til eye trackers? Har du brugt text-to-speech (tts) og eller oplæsningssoftware? (fritid) Hvor mange bøger har du læst i din fritid (ikke relateret til dit professionelle liv) det seneste år? (professionelt) Hvor mange bøger har du læst pga. dit professionelle liv det seneste år?

9 Appendix P

9.1 Image of program

This is an image of the program loaded with the file called 'Minervamodellen intro.docx'.



[H]

9.2 Github link

https://github.itu.dk/rafa/EyeNarrator.git