



# SynVis

## Darstellung von Graphem-Farb-Synästhesie mittels Augmented Reality

### MASTERARBEIT

zur Erlangung des akademischen Grades

### Master of Science

im Rahmen des Studiums

### Media and Human-Centred Computing

eingereicht von

**Christina Tüchler, BSc**

Matrikelnummer 11908107

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Mag.rer.nat. Dr.techn. Hannes Kaufmann

Mitwirkung: Projektass.in Dipl.-Ing.in Dr.in techn. Katharina Krösl, BSc

Dipl.-Ing. Dr.techn. Daniel Cornel, BSc

Wien, 24. August 2024

---

Christina Tüchler

---

Hannes Kaufmann



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# SynVis

## Digitising Grapheme-Colour Synaesthesia Through Augmented Reality

### MASTER'S THESIS

submitted in partial fulfillment of the requirements for the degree of

### Master of Science

in

### Media and Human-Centred Computing

by

**Christina Tüchler, BSc**

Registration Number 11908107

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Mag.rer.nat. Dr.techn. Hannes Kaufmann

Assistance: Projektass.in Dipl.-Ing.in Dr.in techn. Katharina Krösl, BSc

Dipl.-Ing. Dr.techn. Daniel Cornel, BSc

Vienna, 24<sup>th</sup> August, 2024

---

Christina Tüchler

---

Hannes Kaufmann



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Erklärung zur Verfassung der Arbeit

Christina Tüchler, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 24. August 2024

---

Christina Tüchler



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich während der Erstellung dieser Masterarbeit unterstützt haben.

Ein besonderer Dank gilt der **VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH**. Diese wird vom BMK, BMAW, Land Steiermark, Steirische Wirtschaftsförderung - SFG, Land Tirol und Wirtschaftsagentur Wien - Ein Fonds der Stadt Wien im Rahmen des von der FFG abgewickelten COMET - Competence Centers for Excellent Technologies (879730) gefördert. Hier konnte ich im Rahmen eines dreimonatigen Praktikums viel Wissen aus der Forschung mitnehmen und meine Masterarbeit starten.

Ich möchte meinen herzlichen Dank an **Hannes Kaufmann** aussprechen, der mich während dieser Arbeit betreut hat. Er war stets per E-Mail erreichbar und hatte immer ein offenes Ohr sowie Lösungen für meine Anliegen, sei es in Bezug auf Hard- oder Software. Ebenso bedanke ich mich bei **Katharina Krösl** und **Daniel Cornel**, die mich während meines Praktikums und darüber hinaus tatkräftig unterstützt und gefördert haben. Unsere fast zweiwöchigen Jour fixes waren äußerst hilfreich, und ich konnte mich immer auf ihr zeitnahes Feedback verlassen. Vielen Dank an euch alle, die Zusammenarbeit mit euch hat mir große Freude bereitet!

Darüber hinaus möchte ich mich bei allen **Teilnehmern** der Evaluierungsphase (Experteninterviews und Nutzerstudie) bedanken, ohne die diese Arbeit nicht möglich gewesen wäre.

Nicht zuletzt möchte ich mich bei meinem **Partner**, meiner **Familie** und meinen **Freunden** bedanken, die mich jederzeit emotional unterstützt haben. Danke euch!



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# Acknowledgements

I would like to take this opportunity to thank everyone who has supported me during writing this Master's thesis.

Special thanks go to **VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH**. This is funded by the BMK, BMAW, Province of Styria, Styrian Business Promotion Agency - SFG, Province of Tyrol and Vienna Business Agency - A Fund of the City of Vienna as part of the COMET - Competence Centres for Excellent Technologies (879730) handled by the FFG. During a three-month internship here, I was able to gain a lot of knowledge from research and start my Master's thesis.

Many thanks to **Hannes Kaufmann**, who supervised me during this work. He was always just an email away, ready to listen and provide solutions to my concerns, whether hardware or software-related. I am also grateful to **Katharina Krösl** and **Daniel Cornel**, who supported and encouraged me during my internship and beyond. Our almost bi-weekly meetings were extremely helpful, and I could always rely on their prompt feedback. Thank you all so much; working with you has been a truly enjoyable experience!

I would also like to thank all **participants** in the evaluation phase (expert interviews and user study), without whom this work would not have been possible.

Last but not least, I would like to thank my **partner**, my **family** and my **friends**, who have supported me emotionally at all times. Thank you all!



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Kurzfassung

Die häufigste Form der Synästhesie, die Graphem-Farb-Synästhesie, verursacht einzigartige Empfindungen, bei denen Buchstaben und Zahlen mit bestimmten Farben assoziiert werden. Angesichts der rasanten technologischen Entwicklung, insbesondere im Bereich der unterstützenden Technologien, untersucht diese Masterarbeit die visuelle Reproduktion der Graphem-Farb-Synästhesie mithilfe von Augmented Reality (AR).

Ziel dieser Arbeit ist es herauszufinden, ob die individuell unterschiedlichen Wahrnehmungen von Synästheten in einfache, maschinell implementierbare Regelwerke kodiert werden können, die es Synästheten ermöglichen, schwarzen Text vor dem Lesen einzufärben. Außerdem sollten die technischen Voraussetzungen für die Implementierung eines solchen Systems ermittelt werden.

Daher wird eine Literaturrecherche durchgeführt, um herauszufinden, ob es bestimmte wiederkehrende Muster in der Farbwahrnehmung von Synästheten auf Wortebene gibt. Auf der Grundlage früherer Forschungsarbeiten und Experteninterviews mit einem Synästhesieforscher und einem Synästheten wird die Identifizierung und Formalisierung von regulierenden Faktoren, die bestimmte Farben bei Synästheten hervorrufen, validiert. Dies ermöglicht die Entwicklung eines Prototyps, der mobile AR zur Darstellung von Graphem-Farb-Synästhesie-Wahrnehmungen verwendet. Die Anwendung ermöglicht die Umfärbung von Text in der realen Welt mithilfe des Kamerabildes des Geräts nach verschiedenen vordefinierten Regeln.

Zur qualitativen und quantitativen Analyse werden Experteninterviews, Benchmark-Tests zur Bewertung der Performance der Anwendung (Framezeit, Antwortzeit und Fehlerquote) sowie eine Nutzerstudie zur Ermittlung der technischen Machbarkeit durchgeführt. Die Bewertung der verschiedenen Visualisierungsalternativen ergibt eine Präferenz für die am wenigsten invasive Visualisierung. Die effektivste Methode ist die Einfärbung des Textes in den entsprechenden Farben direkt auf der Pixelbasis des Kamerabildes. Aus Gründen der Lesbarkeit bei dunklen Farben wird die Darstellung von Farbtönen als Hintergrund hinter schwarzer Schrift nicht bevorzugt.

Die Arbeit befasst sich mit den technischen Hindernissen und schlägt Optionen für zukünftige Forschung vor, die den Weg für weitere Forschung auf diesem Gebiet ebnen.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Abstract

The most common type of synaesthesia, known as grapheme-colour synaesthesia, causes unique sensations in which letters and numbers are associated with specific colours. With the rapid expansion of technology, particularly in the field of assistive technology, this Master's thesis investigates the visual replication of grapheme-colour synaesthesia using Augmented Reality (AR).

The goal of this thesis is to see if synaesthetes' different individual perceptions could be coded into simple, machine-implementable rule-sets that would allow synaesthetes to pre-colour achromatic text before reading. Furthermore, it aims to establish the technical requirements for implementing such a system.

Therefore, a literature review is conducted to find out if there are specific recurring patterns in how synaesthetes perceive colours on the word level. Based on previous research and expert interviews with a synaesthesia researcher and a synaesthete, the identification and formalisation of the regulatory factors that elicit specific colours in synaesthetes are validated. This allows for the creation of a prototype that uses mobile AR to represent grapheme-colour synaesthesia perceptions. The app enables the recolouring of real-world text using the device's camera input, based on different rule-sets provided.

To analyse this qualitatively and quantitatively, the thesis includes expert interviews, benchmark tests to assess the app performance (frame time (FT), response rate (RR), and error rate (ER)), and a user study to determine technological feasibility. Evaluating the various visualisation alternatives reveals a preference for minimum invasive visualisations. The most effective method is to outline text in the appropriate colours directly on the pixel basis of the camera picture. Visualising hues as backdrops behind black lettering, on the other hand, is disliked due to readability concerns with dark colours.

The work addresses technical hurdles and suggests options for future research, opening the door for more research in this area.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Contents

<b>Kurzfassung</b>	<b>xi</b>
<b>Abstract</b>	<b>xiii</b>
<b>Contents</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation and Problem Statement . . . . .	1
1.2 Research Questions and Approach . . . . .	2
1.3 Contribution . . . . .	3
1.4 Thesis Structure . . . . .	4
<b>2 Background and Related Work</b>	<b>5</b>
2.1 Psychological Background . . . . .	5
2.1.1 Word Origin and Definition . . . . .	5
2.1.2 Types of Synaesthesia . . . . .	6
2.1.3 Perception, Cognition and Personality . . . . .	7
2.2 Grapheme-Colour Synaesthesia . . . . .	8
2.2.1 Related Types . . . . .	9
2.2.2 Terminology . . . . .	10
Inducer and Concurrent . . . . .	10
Projective and Associative . . . . .	10
Lower and Higher Distinction . . . . .	11
Natural and Artificial . . . . .	11
2.2.3 Development . . . . .	12
Consistency . . . . .	12
Learning Synaesthesia . . . . .	12
2.2.4 Neurological Perspective . . . . .	13
2.2.5 Regulatory Factors . . . . .	14
Shared Codes . . . . .	14
Different Appearances . . . . .	15
2.3 Colouring Graphemes . . . . .	16
2.3.1 Colour Pickers . . . . .	16
	xv

2.3.2	Studies and Applications . . . . .	18
2.4	Data Augmentation and Visualisation . . . . .	20
2.4.1	Text Detection and Recognition . . . . .	20
2.4.2	Relevant Visualisation Techniques . . . . .	22
2.5	Reading Performance . . . . .	24
<b>3</b>	<b>Design</b>	<b>27</b>
3.1	Requirements . . . . .	27
3.1.1	Functional . . . . .	27
3.1.2	Non-Functional . . . . .	28
3.2	Design of the System . . . . .	28
3.3	Word Colouring Rules . . . . .	29
3.4	Visualisation and Colouring Types . . . . .	30
3.5	Application Features . . . . .	32
3.5.1	Colour Definition for Each Grapheme . . . . .	32
3.5.2	Rule-Set Definition for Word Colouring . . . . .	32
3.5.3	Visualisation Style Definition . . . . .	32
3.5.4	Text Scan Functionality . . . . .	33
<b>4</b>	<b>SynVis Implementation</b>	<b>35</b>
4.1	Tech Stack . . . . .	35
4.1.1	Unity Libraries . . . . .	35
4.1.2	Text Detection and Recognition . . . . .	36
4.1.3	Used Devices . . . . .	36
4.2	User Flow . . . . .	36
4.3	User Interface . . . . .	37
4.3.1	Intuitiveness and How-To Guidance . . . . .	37
4.3.2	Simplicity and Mode Indication . . . . .	38
4.3.3	Consistency and Colour Scheme . . . . .	38
4.3.4	User Data Persistence . . . . .	40
4.4	Algorithms and Techniques . . . . .	40
4.4.1	User Data Structure . . . . .	40
4.4.2	Detection and Colourisation Procedure . . . . .	41
4.4.3	Texture Direct Pixel Manipulation . . . . .	41
4.4.4	TMP Overlay . . . . .	43
<b>5</b>	<b>Testing and Evaluation Design</b>	<b>45</b>
5.1	Benchmark Testing . . . . .	45
5.1.1	Performance Metrics . . . . .	45
5.1.2	Implementation of Testing Environment and Tools . . . . .	46
5.1.3	Testing Procedure . . . . .	46
5.2	Pilot Study . . . . .	46
5.2.1	User Study Protocol . . . . .	47
5.3	User Study . . . . .	48



5.3.1	Participants . . . . .	49
5.3.2	Data Collection . . . . .	49
5.4	Expert Interviews . . . . .	49
5.4.1	Selection of Experts . . . . .	49
5.4.2	Interview Procedure . . . . .	50
<b>6</b>	<b>Results</b>	<b>53</b>
6.1	Quantitative Analysis . . . . .	53
6.1.1	Benchmark Tests . . . . .	53
6.1.2	Pilot Study . . . . .	56
	Input Field Adjustment . . . . .	56
	Display Adjustments . . . . .	57
	Focus Mode Activation . . . . .	57
	Questionnaire Format . . . . .	57
6.1.3	User Study . . . . .	57
	System Usability Scale . . . . .	58
	Visualisation Preferences . . . . .	58
	Readability . . . . .	59
6.2	Qualitative Analysis . . . . .	60
6.2.1	Identified Themes . . . . .	61
	Rule-Set . . . . .	62
	Visualisation . . . . .	63
	Challenges . . . . .	64
	Interaction . . . . .	65
<b>7</b>	<b>Discussion of Results</b>	<b>67</b>
7.1	Limitations . . . . .	70
7.1.1	Colouring . . . . .	70
7.1.2	Rule-Sets . . . . .	70
7.1.3	Recognition and Detection . . . . .	70
7.2	Potential Future Work . . . . .	71
7.2.1	Colouring . . . . .	71
7.2.2	Rule-Sets . . . . .	71
7.2.3	Recognition and Detection . . . . .	71
7.2.4	Performance . . . . .	72
7.2.5	Further Research and Development . . . . .	72
<b>8</b>	<b>Conclusion</b>	<b>73</b>
	<b>List of Figures</b>	<b>75</b>
	<b>List of Tables</b>	<b>79</b>
	<b>Acronyms</b>	<b>81</b>

<b>Bibliography</b>	<b>83</b>
<b>Appendix</b>	<b>95</b>
User Study Questionnaire . . . . .	95
Texts for Testing . . . . .	100

# Introduction

This thesis explores the technical feasibility of a prototype called SynVis, which enables individuals with grapheme-colour synaesthesia to modify the colour of text to align their personal perceptions with the real world text. It demonstrates the potential of achieving this in augmented reality (AR) with a System Usability Scale (SUS) score of 88.75, indicating an excellent user experience, according to the Sauro-Lewis curved grading system [LUM15].

Synaesthesia is a condition that can be experienced by humans in which the reception of regular inputs naturally results in unusual concurrent experiences. For example, one of the most prevalent and most researched forms of synaesthesia is grapheme-colour synaesthesia, which describes the phenomenon of experiencing colour when reading graphemes such as letters or digits. This happens because for people with synaesthesia, the brain's V4 (colour processing area), which is part of the visual cortex, has an abnormally high level of connection to the regions that process graphemes [BHC<sup>+</sup>10, MS22]. Around 4% of the total population experiences synaesthesia, with over 1% experiencing grapheme-colour synaesthesia, which results in more than 90 million people [SMS<sup>+</sup>06, RG19, WS20]. Although synaesthesia is linked to high levels of creativity and improved recognition recall, it also necessitates more work to resolve conflicts between sensory perceptions and contemporaneous experiences. As synaesthesia is strongly linked to learning, ignorance can lead to stigmatisation and a lack of adequate support [WAS<sup>+</sup>14].

## 1.1 Motivation and Problem Statement

While some advantages of synaesthesia, such as enhanced memory recall and increased creativity, have been identified, the benefits and downsides of grapheme-colour synaesthesia, particularly in the context of reading, remain largely unexplored [GRJ12, Sim07]. It is unknown to what extent the downsides can be alleviated with assistive technology, and if the benefits can be exploited for increased reading efficiency. Key knowledge and

technologies to conduct this research are missing. While there are already tools that allow the definition of colours for individual graphemes, like the "Synesthesia Battery" [EKT<sup>+</sup>07] and the colouring of words like the "SeeSynesthete" Chrome Extension [Chr20], these applications only function at the individual grapheme level. Given the considerable individuality of induced colour experiences, it is challenging to determine the most appropriate methodology for formalising them into a finite set of rules. It is also unknown how to use such rules to approximate colour experiences algorithmically to maximise congruence between reproduction and the synaesthetic experience it induces. It is not even studied if and how colour perceptions induced by single letters are transferred to words and full sentences. Answering these questions is hard, as synaesthesia is a complex and highly individual phenomenon. In the course of this thesis, this gap is addressed.

In this thesis, visual stimuli are aligned with synaesthetic experiences in arbitrary texts to reinforce the impression of the read text through targeted bundling of neurological signals. This is realised on a novel research platform that allows for the visualisation of individually coloured texts in AR.

### 1.2 Research Questions and Approach

This thesis aims to determine the feasibility of facilitating and reinforcing potential benefits related to synaesthesia using contemporary technology.

The research questions of this thesis can be formulated as follows:

- RQ1:** *How can we formalise and reproduce individual grapheme-colour synaesthetic experiences on a digital screen?*
- RQ2:** *What technical developments are necessary to align an AR visualisation with the experience of grapheme-colour synaesthesia?*

To answer RQ1, an extensive literature review is conducted with the objective of formalising rule-sets, which are subsequently validated through an expert interview with a psychologist specialised in synaesthesia research.

To answer RQ2, an application is created that allows users to select colours for graphemes using a colour picker widget and a variety of suggested stylised text visualisations (e.g., solid colours or outlines) to define the appearance of graphemes and reproduce their visual perception of grapheme-colour synaesthesia on a digital screen as closely as possible. These visualisation techniques are adapted from previous research [EKT<sup>+</sup>07]. The prototype is then expanded with character/text detection and AR to test its use with printed text in any environment. Finally, the recoloured and restyled text is displayed on the screen, see Figure 1.1.

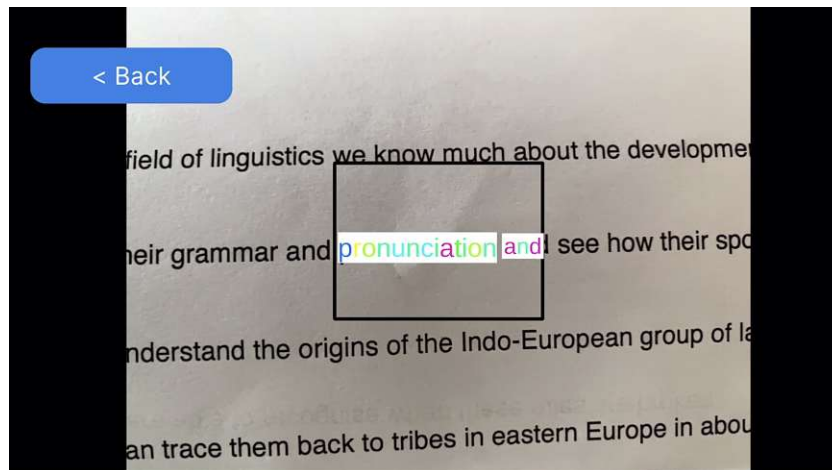


Figure 1.1: Screenshot of the prototype showing the text recoloring feature, which allows users to personalise their reading experience.

In order to evaluate this approach and ascertain its usability, user testing is conducted with twelve volunteers. Benchmark testing is carried out to test the performance and ascertain the extent of potential errors. A psychologist in synaesthesia research, as well as a synaesthete, are interviewed to validate the usefulness of the approach as a whole. This mixed methods approach is primarily designed to demonstrate technical feasibility, while leaving detailed functional assessment to psychologists.

The application of these evaluation techniques enables the determination of the technical feasibility and utility of this type of assistive technology.

### 1.3 Contribution

Since reading is an important task in everyday life, this thesis serves as a cornerstone for further research in the direction of natural synaesthetic reading. Synaesthetic experience is an essential factor in text reading that can be managed for improved reading comprehension, which can be quantified with established reading comprehension assessments at word, sentence, and text level [LFJH19a]. This thesis contributes by presenting:

- Validated formalised rule-sets that describe how synaesthetic colour perceptions can be applied to words.
- A prototype that is capable of reconstructing synaesthetic perceptions through the medium of AR.
- The results of a user study on the usability of this approach, as well as the preferred visualisation style combined with readability ratings.
- Benchmark test results and potential further development and improvement ideas.

- Insights into the general usage of assistive technology for synaesthesia.

Therefore, this project lays a solid foundation for further research and development.

### 1.4 Thesis Structure

At the beginning, Chapter 2, provides the necessary foundation knowledge on synaesthesia. Because this thesis combines psychology, computer science, user experience, and visualisation, this chapter goes over the psychological foundations of synaesthesia, with a particular emphasis on grapheme-colour synaesthesia and related types, as well as the neurological perspective. It investigates various methods of determining colours for graphemes, citing relevant research and applications. It also evaluates earlier research on text detection and recognition, as well as various possible visualisation techniques.

Chapter 3 defines the methodological approach and design decisions. It encompasses the functional and non-functional requirements, the system's design, the determination of the word colouring rules, the selection of different visualisation and colouring styles, and the application features.

The Chapter 4 focuses entirely on the technical aspects of the development phase of SynVis. It details the technology stack, the user flow, the user interface (UI), as well as specific algorithms and techniques used in the implementation.

The testing strategy to thoroughly evaluate the application is described in Chapter 5.

Chapter 6 presents the results, which are divided into two sections: quantitative results and qualitative results. The benchmark tests and the user study are evaluated statistically, whereas the expert interviews are evaluated thematically.

Chapter 7 discusses the findings, stressing the study's shortcomings and proposing possibilities for future research.

The thesis finishes with Chapter 8, which is a brief conclusion of the entire work.

# Background and Related Work

The following chapter presents the background that forms the theoretical basis for the thesis, as well as the literature already available on these topics. Additionally, the “state-of-the-art” is elaborated on in this section.

## 2.1 Psychological Background

Researchers and academics from multiple disciplines have been drawn to synaesthesia, a neurological condition that is characterised by the unintentional mixing of sensory experiences. People that possess this ability are referred to as synaesthetes, and they are able to interpret certain stimuli, such as letters, numbers, or noises, as evoking extrasensory reactions, such as colours, forms, or tastes. This section focuses on investigating the psychological aspects of synaesthesia to understand the typology, its effects on cognition, emotion, and perception.

### 2.1.1 Word Origin and Definition

The word synaesthesia comes from the Greek and is split in two parts "syn" meaning "union" or "together" and "aisthesis" meaning "sensation" or "perception", so it can be understood as joint sensation or union of the senses [Cyt89, Cyt95]. In other words, the term "synaesthesia" was created to characterise the state in which sensory experiences converge, leading to the simultaneous and uncontrollable perception of several senses in response to a single sensory stimulation. This word has been used as an umbrella term for mixing senses in various contexts [Mei22]. It was first used by the German physician and philosopher Georg Sachs in the 19th century [JDW09].

### 2.1.2 Types of Synaesthesia

Based on Sean A. Day [Day22] there are at least 75 different types of synaesthesia (see Figure 2.1) but as research goes on, more types of synaesthesia are explored, since each individual with synaesthesia may have a unique combination of sensory associations.

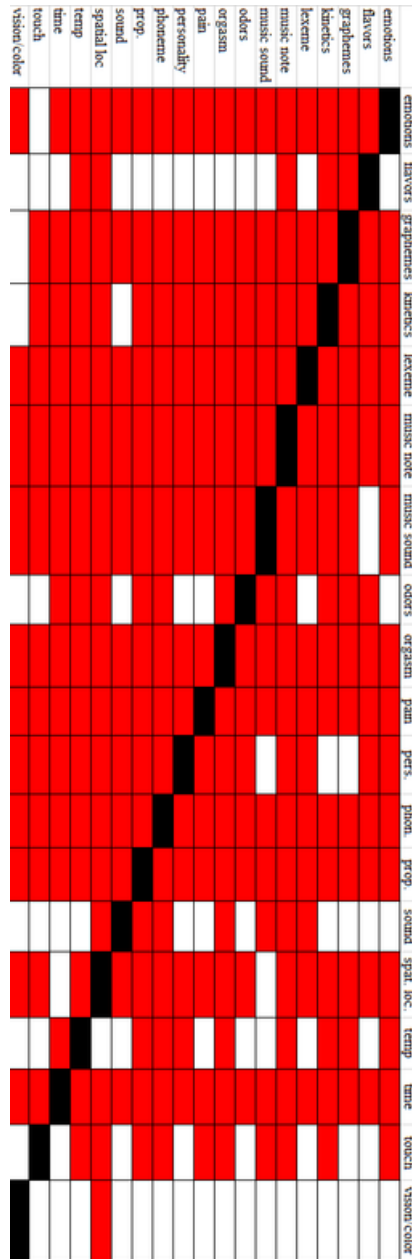


Figure 2.1: Seventy-five types of synaesthesia (Sean A. Day). Left column: inducers; top row: concurrents. White: documented; red: unrecorded; black: not a type. (Reprinted from: [Day22])



According to Sean A. Day [Day22], the most prevalent form of synaesthesia is grapheme-vision synaesthesia, in which for instance letters, numbers, or shapes evoke specific colours, which currently have 162 million people in the world. The second-most known form is time unit-vision synaesthesia, which allows people to interpret time units like days or months as visually different colours, forms, or patterns. Chromaesthesia, in which for instance sounds stimulate perceptions of colours, combining music with a spectacular visual experience, is also a common form of synaesthesia. Another common kind of synaesthesia is called spatial-sequence synaesthesia, which arranges numerals, months, and days of the week into spatial patterns. Not so common forms include for instance lexical-gustatory synaesthesia, in which words or specific sounds evoke tastes or mirror touch synaesthesia, which allows individuals to physically feel the feelings felt by others when seeing their touch or physical interactions. [SS15, MR13a]

People who have synaesthesia can, in some cases, experience numerous forms at once, a condition known as "co-occurrence" or "multiple synaesthesia" [SMS<sup>+</sup>06, NCE11, Mei22].

### 2.1.3 Perception, Cognition and Personality

Developing synaesthesia is regarded as a typical cognitive variance in the general population [War12]. Unfortunately, synaesthetes are frequently misunderstood, which makes people avoid talking about their experiences and causes scientific research to understate the incidence of synaesthesia in the general population [Day13, SS15].

Most research on synaesthesia focuses on its causes and the neurological mechanisms involved. Studies investigating the positive and negative effects of synaesthesia on cognitive abilities mostly focus on enhanced memory abilities in grapheme-colour synaesthetes as compared to non-synaesthetes. They conclusively suggest a persistent memory advantage in memory/recall tasks [LM20, GNCHCG11, SS15]. The study by Simner and Bain [SB17] mentions benefits of synaesthesia in tasks testing processing speed and memory/recall of letters. A follow-up study by Smees et al. [SHCS19] reports significantly enhanced performance in expressive and receptive vocabulary tests compared to non-synaesthetes, but no benefits in sentence comprehension. The study by Palmeri et al. [PBM<sup>+</sup>02] also clearly shows that synaesthetes are able to recognise shapes or patterns way better because of the colouring of the graphemes than non-synaesthetes. What is found by Mannix and Sørensen [MS22] is that synaesthetes are significantly poorer in recognising people's faces. Since it has a huge impact on the cognitive function, the study by Sinke et al. [SNZ<sup>+</sup>14] demonstrates that synaesthetes are worse in speech perception, while the study by McCarthy and Caplovitz [MC14] reveals a similar finding with regard to motion perception.

Early studies indicate that synaesthesia occurs early in the processing of perception, showing real sensory connections. Based on this evidence, one study by Ramachandran and Hubbard [RH01] discovers that synaesthetes demonstrate a greater ability to recognise more geometric shapes composed of digits when compared to their non-synaesthete counterparts. Researchers found out that, synaesthetes score higher on the personality

trait of openness to experience, may score lower on the traits of agreeableness and neuroticism, and have greater levels of schizotypy and inventiveness [Mei22]. The first one mentioned suggests a stronger appreciation for novelty and creativity [LM18, WTLEK08, Mei22, SS15].

A number of studies have already been conducted with the aim of making the benefits of synaesthesia accessible to a wider audience. These studies span a variety of fields, including the work of Reif and Alhalabi [RA18] on virtual reality (VR) induced artificial synaesthesia, which seeks to guide patients' attention for medical and therapeutic purposes, such as pain relief.

### 2.2 Grapheme-Colour Synaesthesia

Grapheme-colour synaesthesia is documented as one of the most prevalent among the 75 known types [Day22]. Individuals with this form involuntarily and automatically experience non-coloured, achromatic graphemes (letters and digits/numbers) as coloured [HYS20, PVdSN11, PBM<sup>+</sup>02]. Consequently, this type involves a blending of two senses: visual perception and colour perception.

Patricia Duffy, a colour synaesthete, describes in her book "Blue Cats and Chartreuse Kittens: How Synesthetes Color Their Worlds" [Duf01] a scenario that occurred while she was learning to write the alphabet:

To make an R, all I had to do was first write a P and then draw a line down from its loop. And I was surprised that I could turn a yellow letter into an orange letter just by adding a line.

This accurately describes the experience of synaesthetes, for whom graphemes (letters and numerals) consistently elicit or trigger specific colour experiences. The "emotional meaning" of this form of synaesthesia is crucial to note, since it might make someone feel uncomfortable or disturbed to see symbols in hues that do not match their personal associations [The21]. Patrizia Puff, for instance, could feel harmonic and right when she sees the letter "P" when it is yellow, but she might feel uncomfortable and incorrect when she sees it when it is green or violet, which might slow down her processing of letters [ANS15].

What should be emphasised is that the colours of each letter/digit are highly individual perceived by each and every person, and there is no single solution that fits for everyone. Even monozygotic twins do not share the same colour associations, as found by Rich et al. [RBM05], but there are some studies on colours based on different interviews with synaesthetes that showed some similarities (see Section 2.2.5).

As mentioned in Section 2.1.3 different cognitive effects of synaesthesia, such as those on memory, creativity, and imagery, may be experienced by synaesthetes. Researchers in the field of synaesthesia often mention the "atypical cross activation" [SMS<sup>+</sup>06] or

the "hyperconnectivity" [Mei22] of the brain. This can be described by the "Semantic Representation of Synaesthesia" [Mei13] by Beat Meier (see Figure 2.2).

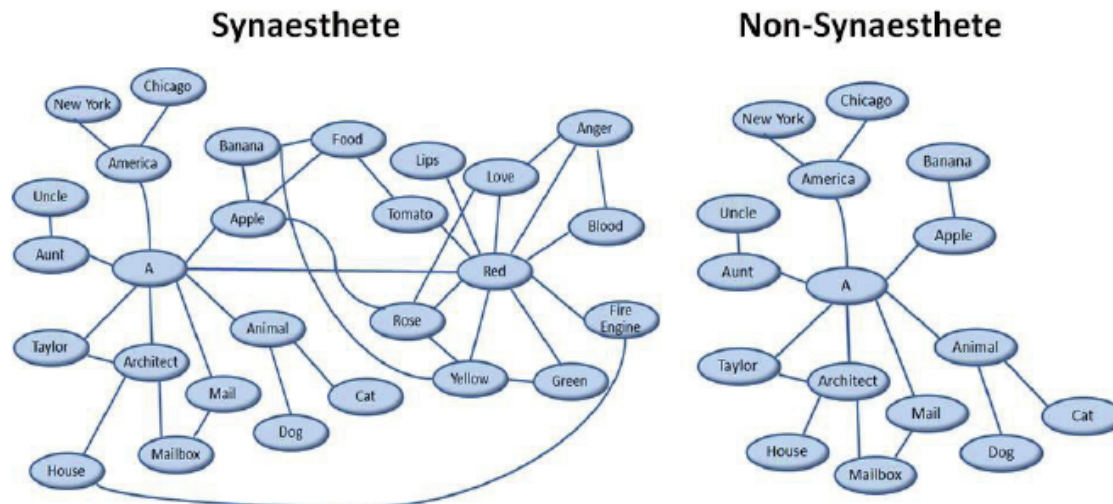


Figure 2.2: Illustration of semantic network activations in response to the letter "A": synaesthete with red colour experience (left) vs. non-synaesthetic control (right). (Reprinted from: [Mei13])

The paper demonstrates how a synaesthete who perceives colours for both letters and words could link words depending on prominent vowels or starting letters. For instance, the colour red is evoked by the letter "A", and words like "animal", or "apple" also do the same. The links between colours and other items, like a rose and a fire engine, are made more easily in synaesthetes than in non-synaesthetes thanks to these synaesthetic associations, which generate an enhanced semantic network. Compared to those without synaesthesia, the synaesthete's increased semantic network enables them to produce intriguing ideas and thoughts, enhancing the scope of experiences. [Mei13]

### 2.2.1 Related Types

There are several related types of synaesthesia which are related to each other. Beyond basic colour associations, some synaesthetes have *grapheme-shape/colour/texture/image* synaesthesia. Along to seeing colours, they might also be able to make out other sensory elements such as shapes, or textures.

One of these linked types is "phoneme-colour" synaesthesia, where colours are connected to spoken words based on how they sound rather than their written form [BMB<sup>+</sup>23, Sim07]. The difference here is, that a grapheme is the smallest unit of written language (a letter), whereas a phoneme is the smallest unit of speech distinguishing one word from another (a sound). Thus, hearing particular phonemes in a word may cause one to perceive particular colours.

In some synaesthesia forms, like in "lexeme-colour" and "morpheme-colour" synaesthesia, colours are associated with diverse portions of words in synaesthesia, not only with particular letters or phonemes [BCG16]. The "morpheme", which provide grammatical meaning, and the "lexeme", which is the word's root, both have distinctive colour connotations.

### 2.2.2 Terminology

Specific synaesthesia impressions are described in specialised vocabulary. This is explained in this section.

#### Inducer and Concurrent

There are two important terms in regard to synaesthesia, named inducer and concurrent, which describe specific aspects of perception. The inducer refers to the "triggering stimulus", meaning the information or stimulus that causes or elicits the synaesthetic perception [CR14, Mei22]. This can be for example a sound, a letter, a number, a scent, a flavour, or even concepts.

The concurrent on the other hand refers to the "resultant experience", meaning it refers to the sensation which comes from a reaction of certain exposed stimuli, so the additional sensory experience or perception [CR14, Mei22]. This can be anything related to another sense like for example a colour, a shape, a tactile sensation, a temperature sensation or for instance spatial perceptions.

Based on the type of synaesthesia, this inducer-concurrent relationship includes different sensory modalities [CR14].

#### Projective and Associative

Generally synaesthetic experiences can be present or perceived by individuals in two different ways, either projective or associative (see Figure 2.3) referring to differences in the concurrent [WLSS07]. The term "projector" is used to refer to those who perceive their synaesthetic connections as if the additional sensory experiences are projected outwardly, appearing in exterior space, or as if they were actually present in the environment, also known as "out there on the page" [WLSS07, HYS20, CWT<sup>+</sup>15]. An individual who has grapheme-colour synaesthesia, for instance, can perceive the colours connected to letters or numbers as hovering in front of or covering the actual objects. This type of experience is perceived just by the minority of synaesthetes, which is by approximately 10% [DS05].

Individuals that feel synaesthetic correlations in an internal, subjective sense, are referred to as "associators", who not actually make the associations in the real surroundings but just internalised "in their mind's eye" [WLSS07, HYS20, CWT<sup>+</sup>15]. An associator with grapheme-colour synaesthesia, for example, may mentally visualise the colours associated with letters or numbers without actually seeing them in the external world.

Individual Synaesthesia Experience Questionnaires (ISEQs) [SLM09] were developed to distinguish between projectors and associators, enabling researchers to classify people based on their subjective experiences and perception of synaesthetic relationships.



Figure 2.3: Differences in perception of letters, numbers, or words by projectors and associators. Top: two projectors; bottom: three associators. (Reprinted from: [The21] and [SLM09])

### Lower and Higher Distinction

One can distinguish between two different levels at which the synaesthetic experience is triggered, the lower and the higher distinction, referring to differences in the inducer [WLSS07]. Lower synaesthesia is triggered by sensory or perceptual elements of the stimuli that are instantaneous, whereby physical aspects of the stimuli, such as its form, colour, or texture, might influence synaesthetic perception [WLSS07].

Higher synaesthesia, on the other hand, is characterised by the involvement of abstract or conceptual variables in the initiation of synaesthetic sensations, which implies that a stimulus's symbolic or linguistic meaning affects how it is perceived by the brain, which is important for synaesthesia [WLSS07].

### Natural and Artificial

Generally, a distinction can be made between natural synaesthesia and artificial synaesthesia. Natural synaesthesia describes synaesthetic experiences that occur spontaneously in people without outside help or modification. It emerges typically in early stages of development [SHC<sup>+</sup>08]. Natural synaesthetes have constant and involuntary links between sensory inputs, allowing them to perceive extra sensory experiences that are not typically associated with the initial stimulus [RA18].

Artificial synaesthesia, on the other hand, refers to synaesthetic experiences that are generated or assisted through external methods, frequently requiring some sort of sensory stimulation or techniques, to enable anyone to experience synaesthesia. To create artificial synaesthesia, so-called sensory substitution devices (SSDs) are used, which can be for instance virtual reality or transcranial magnetic stimulation [SS15]. These seek to elicit cross-modal associations or sensations from people who do not already have synaesthesia. [RA18, WJG<sup>+</sup>21]

The study by Luke et al. [LLFT22] shows that synaesthesia can also appear in temporary altered states of consciousness in which visual and aural hallucinations co-occur, which are frequently generated by psychedelic substances [Mei22].

### 2.2.3 Development

Some researchers are investigating the development of synaesthesia in individuals. In a particular study by Witthoft et al. [WWE15], the authors explore whether there are similarities in how grapheme-colour synaesthesia develops. Some literature suggests that coloured toys, television, or even refrigerator magnets might impact the development of this type of synaesthesia [WW06, WW13, MR13b]. However, a user study by Witthoft et al. [WWE15] of 6588 synaesthetes shows that only 1 out of 6 grapheme-colour synaesthetes appear to have learned the associations through a coloured toy in a span of 10 years. Some literature indicates that the development of synaesthesia is a lengthy process, where the colours vary during the developmental phase of synaesthesia, as the longitudinal study by Simner and Bain [SB13] discovers that 34% of these letter-colour associations are fixed at the age of 6 to 7, 48% at the age of 7 to 8, and 71% at the age of 10 to 11.

### Consistency

There is also literature on its later stages, such as whether letter-colour associations remain consistent throughout a person's life or change in some way, since consistency is a fundamental characteristic of synaesthesia. For instance, a study by Meier et al. [MRW14] shows that associations with primary colours remain largely unchanged, but that as people get older, bright colours become less frequent and more subdued colours like brown and achromatic tones become more common. Uno et al. [UAY21] conduct a longitudinal study examining the letter-colour associations of alphanumeric and Japanese characters. They discover that whether the colour remains the same or not depends on the grapheme, with more frequently used grapheme associations remaining the same and less frequently used graphemes changing their colour over time.

### Learning Synaesthesia

According to studies by Colizoli et al. [CMR12, CMSR17] learning to correlate graphemes with colours is interesting when it comes to colouring text because of the potential advantages. The authors pre-colour books for the participants and run Stroop tasks



[Str35] with them before and after reading the book to compare the outcomes. It is found that this "pseudo synaesthesia" can be trained. However, because this study focuses on the outcome of the trained type of synaesthesia rather than the process and experience of reading per se, the coloured book serves just as an artefact and is not central to the study. A further study is conducted by Schwartzman et al. [SOR<sup>+</sup>23] to ascertain the extent to which the two groups, designated the induced synaesthesia-like (ISL) group and the natural occurring grapheme-colour synaesthetes (NOS), can be compared. They find that besides the fact the ISL share the similarities of NOS, but that the participants report that in the ISL the associations occur more or less "wilful" and in the NOS group automatically, that the induction or an intensive "training of letter-colour associations can alter the conscious perceptual experiences of non-synaesthetes".

### 2.2.4 Neurological Perspective

Researchers, especially neuroscientists, interested in understanding the nature of perception, cognition, and the adaptability of the human brain, identify several brain regions that are consistently active during grapheme-colour synaesthesia. Specifically, the colour area V4 and the posterior temporal grapheme areas (PTGA) are identified as regions of the brain that are consistently active during the occurrence of this condition and are researched as significant locations of interest. When viewing achromatic (black and white) graphemes, the PTGA is activated in both synaesthetes and non-synaesthetes because it is involved in processing letters and numbers, but V4 is only activated in synaesthetes because only synaesthetes perceive the sensation of colour (see Figure 2.4). [BHC<sup>+</sup>10, HARB05, SPL<sup>+</sup>06]

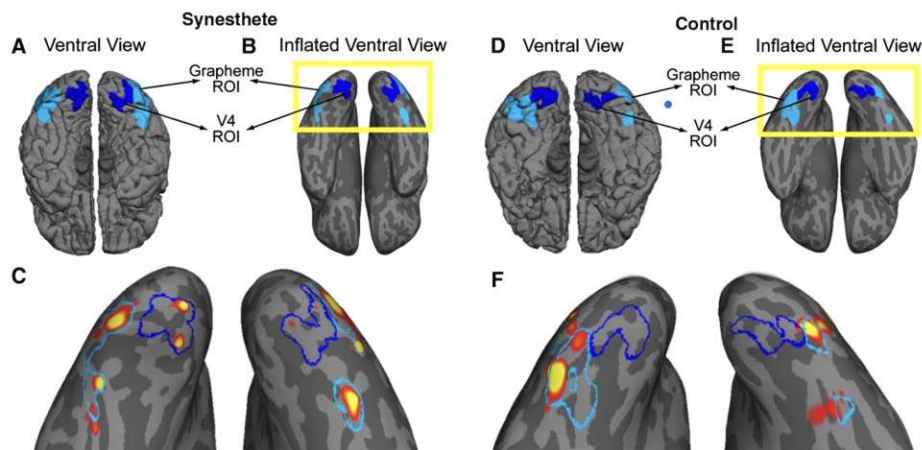


Figure 2.4: Representative synaesthete (A–C) and control (D–F) brains. Grapheme ROI (light blue) and V4 ROI (dark blue) are shown. A and D: ROIs on non-inflated cortical surfaces. B and E: ROIs on inflated brains; yellow box highlights region in C and F. Synaesthetes showed activation in both grapheme (light blue) and V4 (dark blue) ROIs when viewing achromatic letters and numbers (C). Controls showed activation only in the grapheme ROI (light blue) (F). (Reprinted from: [BHC<sup>+</sup>10])

Based on these findings, conventional studies based on the physical colours perceived through the retina cannot entirely explain the phenomenon [HYS20]. Researchers believe that this mechanism is not limited to V4, the parietal lobe, and the fusiform cortex, but may involve a more extensive network of brain regions [SNE<sup>+</sup>12].

### 2.2.5 Regulatory Factors

Regulatory Factors shape the inducer-concurrent relationship [RR21]. This section elaborates on the similarities and differences of these relationships among synaesthetes.

#### Shared Codes

It is attempted to identify potential similarities in the inducer-concurrent relationships in the perception of graphemes [Day04, WWE15]. Certain trends are observed, such as "B" often appearing blue, "C" predominantly yellow, and "I" and "O" frequently white. It is also found that colour associations are influenced by language and culture [SWL<sup>+</sup>05, BK69].

Since most research has been conducted with English-speaking synaesthetes, it is noted by Root et al. [RAM<sup>+</sup>21] that synaesthetic associations are influenced by linguistic and prelinguistic factors. This spurs interest in how the inducer-concurrent relationships of graphemes might vary across different languages. Root et al. [RRA<sup>+</sup>18] additionally focus on the association between the letter "A" and the colour red across five languages.

Given the lack of a universal approach to colouring each grapheme in a way that aligns with the experiences of all synaesthetes, the "Synesthesia Battery" [Dav24b, EKT<sup>+</sup>07] was developed. This platform serves as a tool for individuals to test for synaesthesia and as a resource for researchers, offering a comprehensive database of perceptions and their associated colours. It includes a colour picker, allowing individuals to choose colours for graphemes from a palette of 16.7 million options [BCG16] (see Figure 2.7). By collecting this data, the developers aim to identify common colour associations among synaesthetes. The aggregated data enables the authors to create graphs showing the most frequent colours associated with each grapheme across all participating synaesthetes. This platform has also been validated by other researchers [CDS<sup>+</sup>15].

Additionally, there are publicly accessible tools where synaesthetes can share their unique colour experiences, such as adding a row in a publicly accessible Google Sheets File [Goond] (see Figure 2.5).



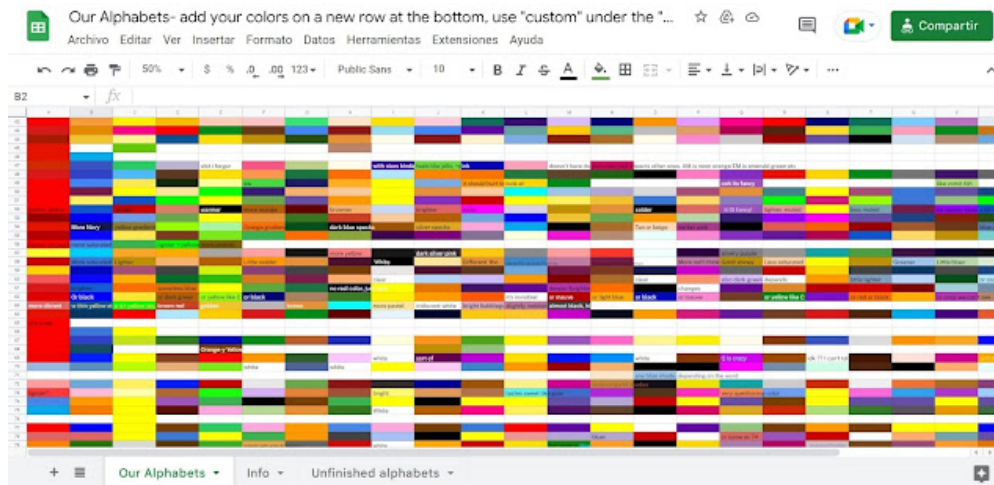


Figure 2.5: Screenshot of the publicly available shared GoogleSheets file [Goond] for colouring cells in the experienced colour. (Reprinted from: [The21])

### Different Appearances

Existing research shows the varied appearances of individual graphemes, but it has less often focused on the appearance of whole words/texts. Nevertheless, a range of manifestations exist in how words are perceived in colour. It is mentioned by Blazej et al. [BCG16] that, for entire words and texts, perceiving each letter in its individual colour is uncommon and instead, the colour of a word is often influenced by the colour of its initial letter. This fact is also stressed by Simner et al. [Sim07, SGM06] since they mention that the different colours of the letters compete and then only one colour dominates the colour of the entire word. For compound words, additional colours can be affected by the first letter of the second morpheme, indicating an interaction between linguistic structure and synaesthetic perception.

Similar findings are reported by Sidoroff-Dorso et al. [SDJDuLM20] via an interview:

[...] the meaning of a word is inseparable from its “color shell”. The first letter (I heard this is often the case with synesthetes) gives the word a dominant tone. For example, the word “trait” is vibrant, dark red, because the letter “t” is painted in this colour. The remaining letters are superimposed on the “background” tone like a mosaic, set by the first letter.

Another synaesthete, interviewed by Sidoroff-Dorso et al. [SDJDuLM20], highlights the influence of when a word is learned:

The colour of other words, typically learned later, tended to be driven by the colour of the first letter.

Additionally, an earlier study by Simner [Sim07] suggests the impact of initial vowels on word colouring. This research proposes that word colour is influenced not only by the first letter but also by the first or stressed vowel. Syllable stress is identified as a primary factor, with letter position being secondary.

Blazej et al. [BCG16] deals with appearances of whole words for people with grapheme-colour synaesthesia and finds that for their participant most of the words are coloured in the colour of the first letter, except for words starting with "I" or "O", which appear white in isolation, these words are mostly shaded in a lighter shade of another letter appearing in the word (see Figure 2.6).

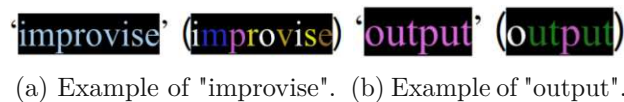


Figure 2.6: Appearance of words starting with "i" or "o". Left: experienced word; right: individual letter colours. (Source: [BCG16])

## 2.3 Colouring Graphemes

This section looks at how graphemes are coloured. Different approaches to providing colours to graphemes are investigated. In addition, light is shone on current research and applications.

### 2.3.1 Colour Pickers

The developers of the "Synesthesia Battery" [Dav24a, EKT<sup>+</sup>07] use RGB values for the colour picking task (see Figure 2.7) which is criticised by Blazej et al. [BCG16], as they mention that the sensitivity could be improved by using other measures of colour difference. Instead of using RGB values, they convert these RGB values to the CIE  $L^*a^*b^*$  colour space, which provides a more uniform colour space that better matches human perception.

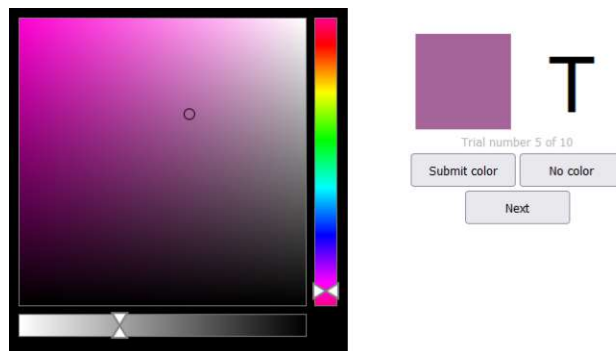


Figure 2.7: Screenshot of the synaesthesia battery's example grapheme-colour picker test. (Source: [Dav24a])

In the work by Rothen et al. [RSWW13] the authors use the presence of colour pickers to let participants choose grapheme experiences from a huge colour palette on several occasions and then measure the consistency of these selected colours. They compare RGB and HSV colour representations with CIE  $L^*a^*b^*$  and CIE  $L^*u^*v^*$  colour models and find that the latter ones can be used for maximising the sensitivity and specificity in relation to other currently used measures for assessing synaesthesia.

Hamada et al. [HYS20] conduct a user study in which synaesthetes are first asked to participate in a colour-selection task, where the researchers show them character cards and the participants have to choose the most appropriate colour according to their experience from the *Munsell Book of Color, Matte Edition* [Mun24]. This is analysed in the CIE  $L^*a^*b^*$  colour space, and then transferred to the CIE  $L^*u^*v^*$  colour space to match the Cambridge Colour Test technique.

A custom-made circular colour palette with adjustable luminance is created in the work by Ásgeirsson et al. [ANS15] which is used to track and select the colours for individual letters and digits one-by-one in the synaesthetic perception (see Figure 2.8).



Figure 2.8: Colour picker that allows the selection of a colour range and brightness adjustment. (Reprinted from: [ANS15])

Kim et al. [KHL22] mention and oppose different colour picker possibilities on different technological devices (desktop, mobile, VR) and mention the characteristics of the different ones. Additionally, they develop a three-dimensional RGB and HSV colour picker for virtual reality. They determine that, in the future, they will focus on developing a perceptually uniform colour model, such as CIE  $L^*u^*v^*$  or CIE  $L^*a^*b^*$ . This decision is based on the fact that they face challenges when selecting colours in certain regions of RGB or HSV, such as the cut-off boundaries or, for instance, in the cone of HSV, given the vast number of colours that can be represented in such a limited space.

It is also relevant which colour picker is used on which device since the colour selection could change and some colour representations are in some situations not so suitable and handy as others [KHL22].

In contrast to the aforementioned studies, the works by Spiller et al. [SHM<sup>+</sup>19], Simner et al. [SMS<sup>+</sup>06] and Meier et al. [MRW14] opt for an alternative approach. They pre-define

a number of colour options prior to conducting the user study, resulting in a limited range of random colours being made available for selection (see Figure 2.9).

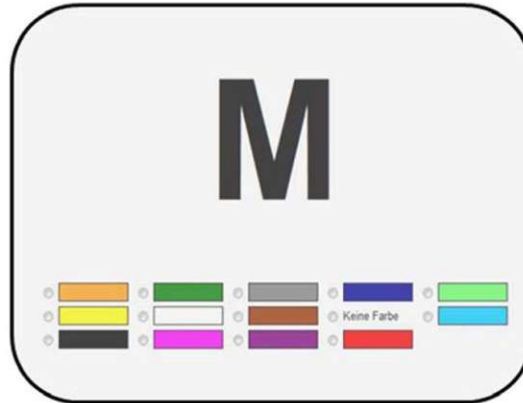


Figure 2.9: Colour picker that allows to choose from 13 colours or "no colour". (Reprinted from: [MRW14])

### 2.3.2 Studies and Applications

In two studies by Berger et al. [BHW<sup>+</sup>19, BHW<sup>+</sup>21] with grapheme-colour synaesthetes using a calculator software with personalised digit colours to perform arithmetic tasks, only marginal performance improvements are observed compared to displaying black digits. However, the feedback of participants reveals a strong preference for congruently coloured digits, with one subject considering it "life-changing". This shows that some effort has been put into the personal colouring of the digits of a calculator. This software, called SYNCalc (see Figure 2.10), is now available for download from both the App Store and Google Play [syn24].



Figure 2.10: Screenshot of the SYNCalc application by Berger and Whittingham. (Reprinted from: [The21])

In addition to this study, which focuses on natural synaesthesia, there is another similar study by Plouznikoff et al. [PPR05] that induces artificial synaesthesia, which provides synaesthetic experiences for non-synaesthetes to attempt and profit from the good features of this digit-colour synaesthesia through the usage of a wearable device.

The interviews by Sidoroff-Dorso et al. [SDJDuLM20] offer a wealth of information on synaesthetes, scientists, and artists regarding various visualisations and associations of graphemes, such as the possibility that a word might be coloured based on its first letter since this could give the word as a whole a dominant tone in the colour of the first letter, for example by visualising it as a background through the rest of the word.

Google Chrome offers a number of extensions that provide users with coloured letters while browsing the web. When one of these extensions is enabled, the letters are coloured before the text is presented to the users, so each letter is presented in its unique colour, but they partly do not support individual colouring schemes of graphemes by users (see examples in Figure 2.11 and Figure 2.12).

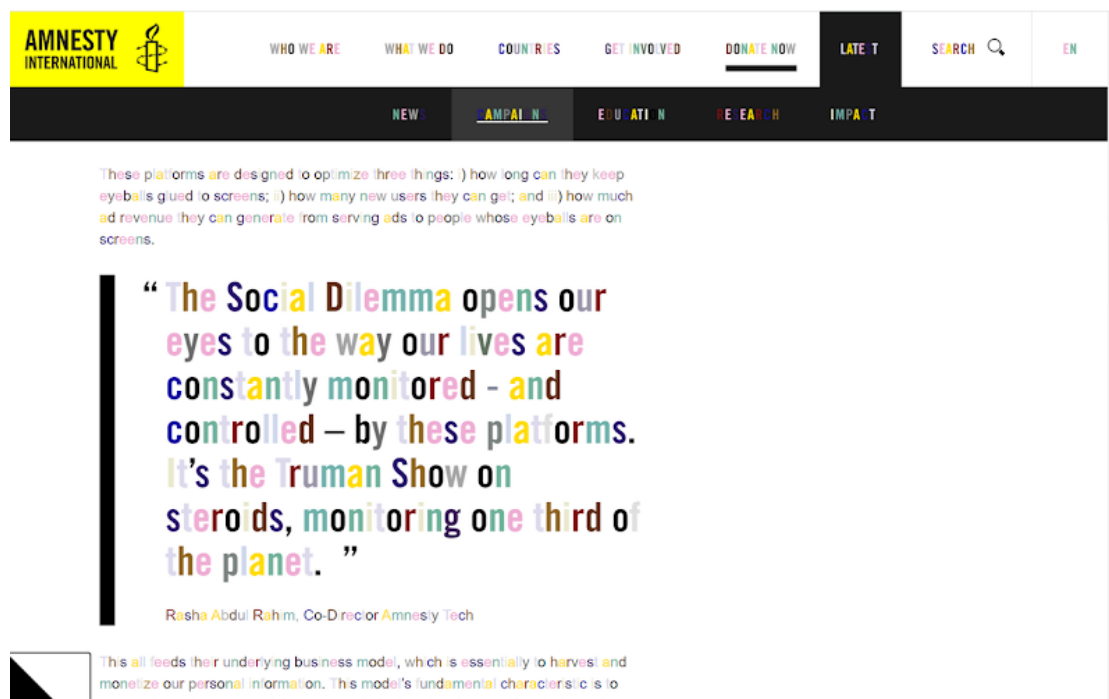


Figure 2.11: Screenshot of the "SeeSynesthete" Google Chrome extension for colouring web page fonts. (Reprinted from: [Chr20])

## 2. BACKGROUND AND RELATED WORK

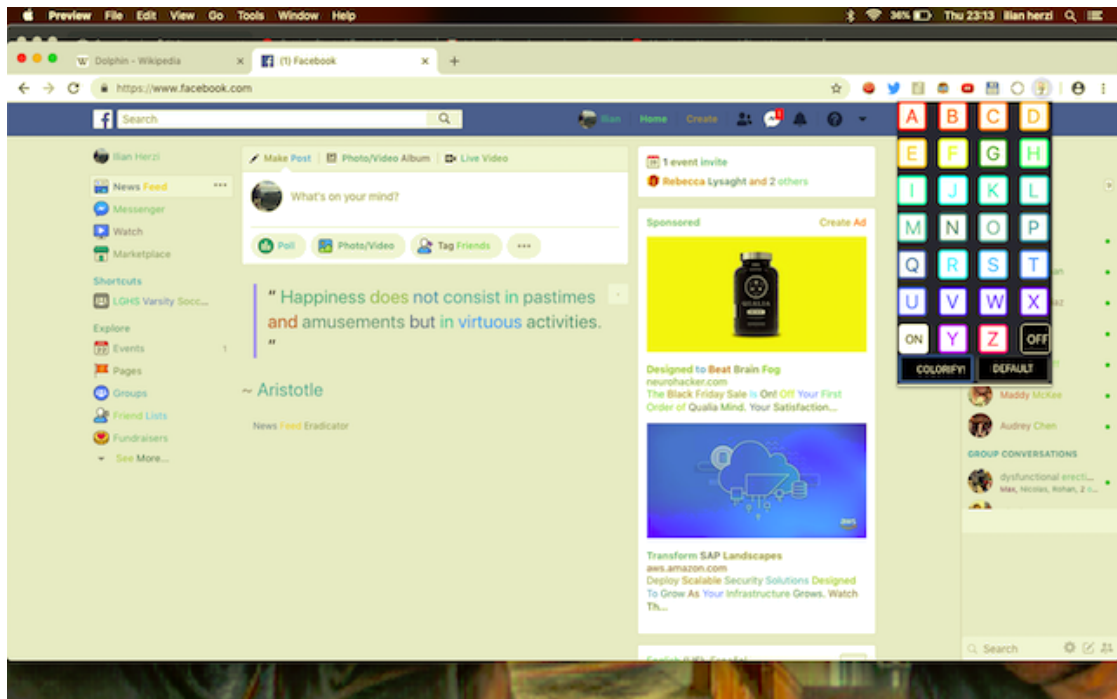


Figure 2.12: Screenshot of the "Synesthesia" Google Chrome extension for colouring web page fonts. (Reprinted from: [Chr19])

## 2.4 Data Augmentation and Visualisation

To recognise, analyse, and visualise modified text, several processes are necessary. The first step, text detection and localisation, involves identifying regions of textual content for further analysis. Following this, text extraction or recognition takes place, utilising optical character recognition (OCR) to convert visual text into machine-readable formats. Finally, the extracted and recognised text is displayed in user-friendly formats through text visualisation, such as on screens or in programs for translation and various other purposes. [TO17]

### 2.4.1 Text Detection and Recognition

Text detection and recognition can be implemented through various methods (see 2.1). Initially, it is necessary to develop different types of input sources for text detection, including those based on images or videos. Subsequently, the choice of technology for executing text detection and recognition depends on these differences; options include AR technology or deep learning (DL) techniques [OHSBHW21].

Text detection and AR/VR technologies have seen significant development in the past. For example, due to poor GPS accuracy indoors, Pivavaruk et al. [PFC22] aim to create an indoor navigation app without relying on GPS. They develop a Unity application



Visualisation device	Year	Reference	Technologies
Mobile Device	2022	[OHW22]	Vuforia ([Vuf24])
	2022	[OBHW22b]	VGG-16, Vuforia
	2022	[OBHW22a]	VGG-19, Vuforia
	2022	[PFC22]	OpenCV Efficient and Accurate Scene Text Detection (EAST), Tesseract OCR ([Git24])
	2021	[OHSBHW21]	Vuforia
	2020	[OHHW20]	Vuforia
	2018	[SH18]	Tesseract OCR
	2017	[PMI17]	Tesseract OCR
HMD	2023	[SGB <sup>+</sup> 23]	Windows Runtime OCR API ([Mic24a]), Knowledge graph
	2020	[Nik20]	OpenCV ([Ope24]), Tesseract OCR
Stationary	2023	[CVRRS23]	OpenCV, Tesseract OCR

Table 2.1: Comparison of relevant related work approaches for text detection and recognition (Table adapted from: [OBHW22b])

that utilises the AR Foundation, smartphone photography, image processing, OpenCV’s EAST DL approach, and the Tesseract OCR algorithm. This OCR algorithm runs in conjunction with Python on a separate server to determine the user’s location based on the image of a door sign. They select this OCR algorithm for its quick and accurate performance and to offload processing power to the server [PFC22, Smi07].

Another study by Strecker et al. [SGB<sup>+</sup>23] seeks to enhance the transition between non-digital and digital interactions in everyday settings by integrating AR with the Windows Runtime OCR API, enabling real-time identification of printed text characters. This project aims not to improve reading comprehension or perception, but to digitise non-digital material and visualise it in AR. The paper also discusses OCR applications in the AR domain extensively.

In order to develop a food menu translator app for translating Thai to Malay, Pu et al. [PMI17] use a mobile device for image scanning, the Tesseract OCR engine for text detection and recognition, and the Google translation service.

Meanwhile, Ouali et al. [OHW22] adopt a different approach to text detection by developing their own AR-based Arabic text detection algorithm.

In another study, Nikolov [Nik20] explores the use of OCR algorithms with AR, focusing on determining the most effective AR device based on the performance of integrated cameras for text identification.

In some literature, DL techniques like the VGG-16 [OBHW22b] or VGG-19 [OBHW22a] models are used alongside AR, the Unity 3D engine, Vuforia, and OCR Tesseract to

create a text detection and magnification application.

On the market, various tools employing OCR algorithms are available. For instance, the "Ctrl F" [Goo23b] app is an AR application readily accessible in the marketplace. It allows users to essentially use the CTRL+F function on their smartphones to search through non-digital content, such as a book page. The app's use of underlying OCR methods enables it to display the highlighted searched word on the user's screen. "SearchCam" [App21] is another application that utilises a similar approach.

### 2.4.2 Relevant Visualisation Techniques

Visualising text on a screen can be achieved in various ways. For example, some applications provide a coloured text overlay layer based on an image for purposes like proofreading (see ScribeOCR in Figure 2.13). In these cases, after the text is detected and recognised, its font and size are adjusted for readability, ensuring that the overlay aligns closely with the individual letters.

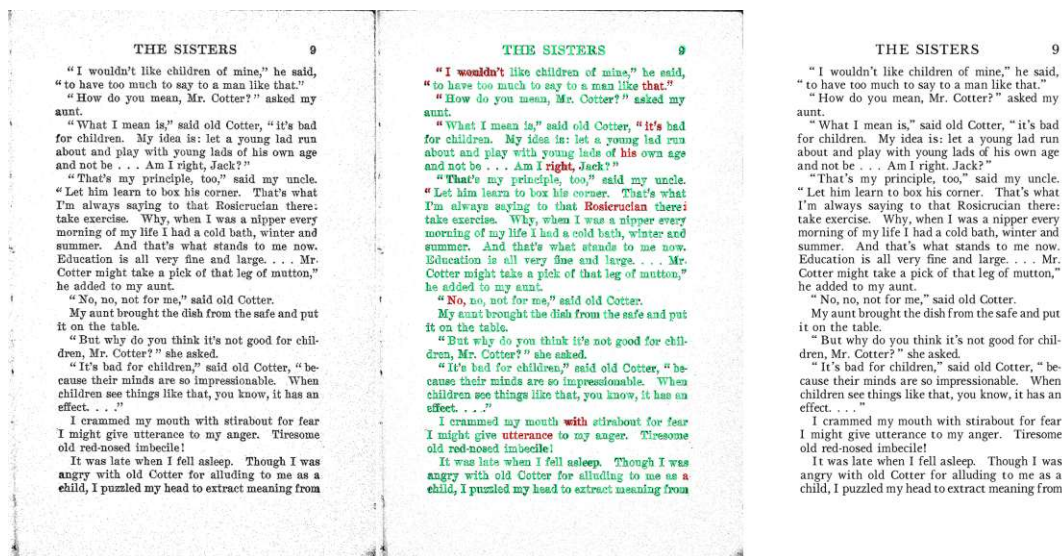


Figure 2.13: Three visualisation versions. Left: original document; center: document with overlay; right: new OCR layer. (Reprinted from: [Scrnd])

Another approach involves combining OCR algorithms with AR visualisation. This technique uses images captured by smartphone cameras as the input source, allowing for the capture of the environment through either photographing the text or recording it as a video. This method is often used in translation apps (see Table 2.1).

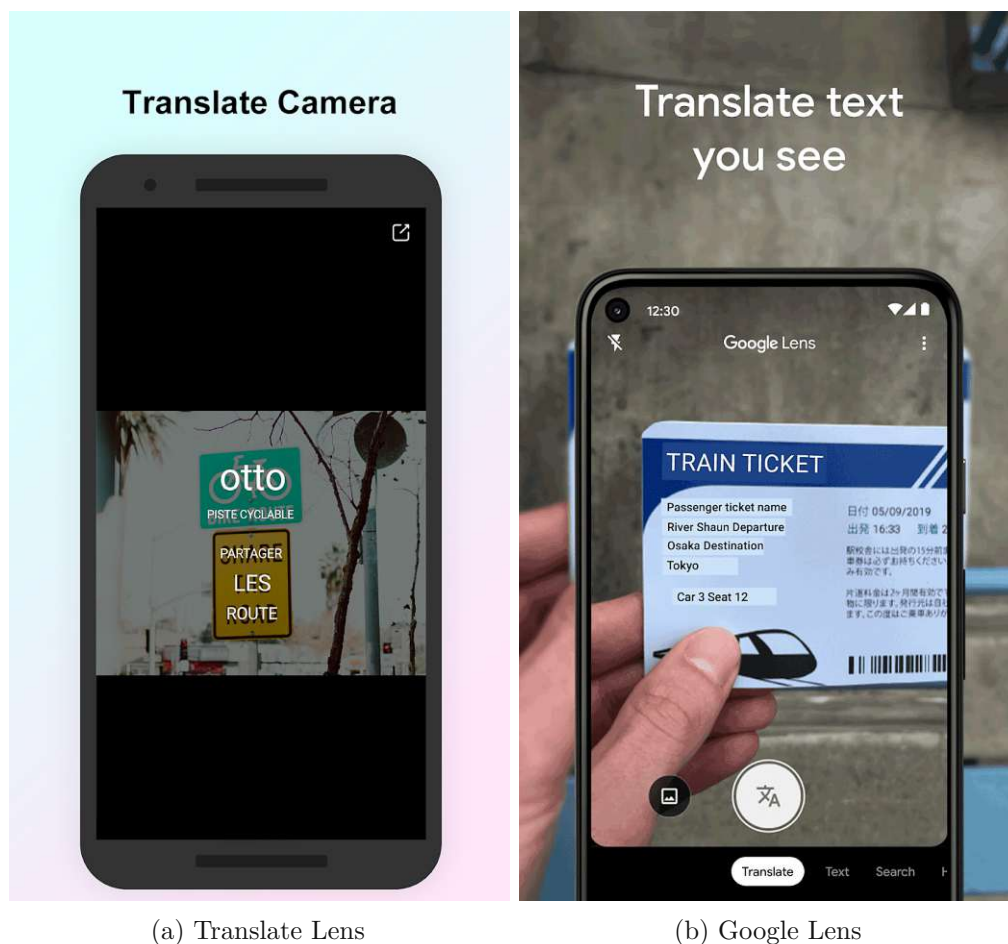
In general, there are mostly works in this direction that have anything to do with translators, as this visualisation is thus useful and relevant. For instance, the study by Tatwany and Ouertani [TO17] addresses the general application of AR in translation duties, reviews the studies and applications that are already available, and lists the



OCR technologies, tools, and programming languages that are utilised in a table. This demonstrates that the indicated studies employ Tesseract, Android, Vuforia, ABBYY, or commercial libraries for the OCR work, as well as technologies such as OpenCV, OpenGL, Matlab, or Eclipse.

Vortmann et al. [VWP22] compare several translator apps, discussing how they employ AR visuals, whether the original text is replaced, displayed as an overlay, or visualised separately. An example of visualising the translation separately beneath the original text is found in the research project by Chorfi and Tatwany [CT19], where real-time text identification and translation is provided within seconds.

Additional visualisation options include overlaying the translated text onto the image in a specific text colour or, in some cases, overlaying the text with a background to ensure sufficient contrast ratio for readability (see Figure 2.14).



(a) Translate Lens

(b) Google Lens

Figure 2.14: Two different AR translator applications using overlay and replacement as visualisation techniques. (Reprinted from: [Goo24] and [Goo23a])

### 2.5 Reading Performance

There has been extensive research on how to measure reading comprehension. In general, reading comprehension can be assessed on word, sentence, and text level [LFJH19b]. The author Fletcher [Fle06] states that there are different elements of reading comprehension based on the style of text presentation and the way the audience responds. They mention the possibility of using multiple-choice tests, fill-in-the-blank or also called cloze exams or retellings or summaries. It is found that inferences about a person's comprehension abilities can differ depending on the method of measurement, so it depends on the reader and on the text as well as the measurement technique [Fle06, CCLG20a]. In the work by Collins et al. [CCLG20b] the researchers evaluate various testing techniques and assess their different influence factors on it. During the evaluation they come to the same conclusion, so it depends on the "text, activity and reader to variance in reading comprehension test scores" and that different response formats are suggested since they are contributing to the variance in reading comprehension test scores. The differences in the various methods of measuring reading comprehension can be interpreted as varying degrees of imperfection in identifying the latent variables that comprise reading comprehension, therefore, it is essential to keep in mind that reading comprehension is a complicated construct that is impacted by a variety of elements and processes, making it challenging to fully capture with a single approach [FSA<sup>+</sup>06, Fle06]. All the above-mentioned methods can be combined with measuring the reading speed.

There are already different techniques for applying these methods, for instance developing a mobile app for assessing the reading performance using a mixture of the cloze and the multiple-choice methods [S<sup>2</sup>1].

There are nowadays also standardised measures for assessing reading comprehension like SLS-Berlin [LFJH19b] by Lüdtkke et al., Salzburger LeseScreening [MW03] by Mayringer and Wimmer, PISA [OECnd], TOEFL [TOE24], which implement a mixture out of these mentioned methods for assessing the reading performance. The "Sentence Verification Technique (SVT)" [Roy05] by Royer assesses the reading comprehension by just focusing on the sentence level.

One can then have a look on which impact coloured letters or coloured text has on the reading performance. In this sense, the studies by Uccula et al. [UEM14] and Griffiths et al. [GTHB16] have a look at the effect of coloured overlays while reading.

Smith et al. [SGMC14] as well as Gran Ekstrand and Nilsson Benfatto et al. [GENB21] use eye tracking techniques to measure the eye-movements and saccades and evaluate based on this the reading performance.

Another not so common technique for measuring reading performance is by using the mental imagery [SL22].

But what still has to be in mind is that there could be a difference in the reading performance when reading from the screen (for instance from a desktop or mobile phone)

compared to reading something from a piece of paper [SP22, KSZ18, SBRL<sup>+</sup>23, CL19, CCC<sup>+</sup>14, HE20, Par19].



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# CHAPTER 3

## Design

To answer the research questions stated in Section 1.2, an AR system is proposed with the purpose of creating a tool for recolouring real-world texts. The major goal of this system is to allow users to view texts in customised colours instead of the traditional black text on a white page.

### 3.1 Requirements

The following functional and non-functional criteria serve as a blueprint and provide a direction for the development process.

#### 3.1.1 Functional

Given the subjective nature of synaesthesia, it is essential to ensure that the approach is highly customisable in order to accurately recreate the experience. This necessitates the provision of a straightforward colour picker tool (see some potential examples in Section 2.3.1) within the system, enabling users to select colours for each grapheme with ease.

Moreover, it is essential to provide colouring rules for text on the word layer to afford users the option to personalise the appearance of colours on text. This ensures the approach is useful for a diverse range of synaesthetes, given that each person perceives this phenomenon in a unique manner. To achieve this, different rule-sets are selected and evaluated (more on that in Section 3.3).

Text scanning, detection and recognition are critical to the method's functionality. The detection and recognition algorithms have to work accurately and seamlessly to ensure that the graphemes are coloured in the appropriate colours in the colouring phase.

In order to facilitate the recolouring of real-world text in AR, it is essential to integrate text scanning, detection and recognition functionalities. This encompasses the presentation of text in AR and precise alignment and rendering in the user's actual surroundings.

To ensure that the technique can be widely deployed and used for further research, cross-platform compatibility must be a primary consideration. Consequently, SynVis is designed with this objective in mind, offering a uniform UI across multiple platforms.

#### 3.1.2 Non-Functional

Performance is a crucial non-functional criteria to use. It must operate efficiently, facilitating text detection and recognition as well as colour applying with minimal latency. This performance is vital for providing a seamless experience, thereby ensuring the uninterrupted reading of the recoloured text. Consequently, benchmark tests are conducted to evaluate the performance of SynVis.

Another crucial aspect is that of usability. The approach must possess an intuitive and straightforward UI, thus clear instructions are integrated to guide users through the process of selecting colours, establishing rules, scanning and reading recoloured text.

### 3.2 Design of the System

The system is designed and developed for usage on smartphones, taking use of its mobility and integrated camera capabilities.

Users establish colour preferences for each grapheme. This customisation is supported by a simple colour picker tool, which allows users to choose precise colours for individual letters and digits based on their needs and perceptions.

The system allows for rule-based word colouring in addition to grapheme colouring on its own. Users use predefined rules that specify how words are coloured based on the colours assigned to each grapheme (see Section 3.3).

Once all of these parameters are set up, the system is ready to recolour texts in different visualisation options, which works as shown in Figure 3.1. Users can scan a book page line by line, and the system displays the recoloured words at the moment as they are scanned. This feature enables users to read text in their preferred colours right from the app, leading to a personalised reading experience.

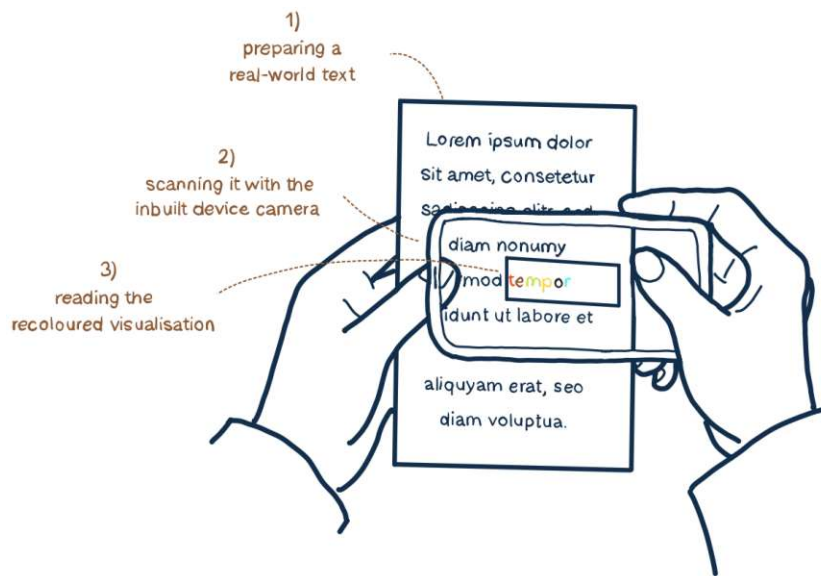


Figure 3.1: System design sketch illustrating the scanning of real-world text and recolouring via mobile AR.

### 3.3 Word Colouring Rules

The selection of word colouring rules for the application is based on an extensive literature review of current research in the field of grapheme-colour synaesthesia.

Firstly, research indicates that the majority of grapheme-colour synaesthetes perceive words in a single colour, which is primarily defined by either the initial letter, the initial vowel, or a specific letter in the stressed or dominant syllable [The21, SGM06, JWA05].

The recolouring of words into a single colour, particularly the colour of the initial letter (see at the top of Figure 3.2) or first vowel (see at the bottom of Figure 3.2), is a method that is selected due to its prevalence in grapheme-colour synaesthesia, as well as its simplicity and ease of implementation. This method does not require the use of complex natural language processing (NLP) techniques, making it a feasible approach with basic text recognition technology.



Figure 3.2: Single-colour word colouring rules selected based on literature review and relevance to this thesis.

In addition to perceiving words in a single colour, there are also individuals who perceive them in multiple colours. For instance, if the word in question is a compound word, each constituent word may be coloured according to the colour of the initial letter, as previously discussed, or other complex rules [BCG16]. Alternatively, more complex rules may be employed, whereby the word is coloured based on the colours of the vowels, with a colour gradient applied to the consonants in between them. In addition, letters that occur prior to the first vowel are represented in the colour associated with the first vowel, whereas letters following the final vowel are represented in the colour associated with the final vowel. It is less common for individuals with grapheme-colour synaesthesia to perceive words in the individual letter colours, although this does occur [The21, BCG16, SGM06].

The aforementioned technique, vowel gradient colouring (see at the bottom of Figure 3.3), is selected for its complexity and its particular relevance to this thesis. This type of synaesthetic perception is experienced by a synaesthete who is interviewed after the implementation of the prototype. They are also invited to test the prototype and provide feedback, as well as evaluate the application's effectiveness. The incorporation of this rule allows for the comprehensive testing of the application's capacity to manage sophisticated colour transitions. While infrequent, the implementation of the rule for the colouring of words in their individual letter colours (see at the top of Figure 3.3) is also included, as it serves as an additional rule for colouring words in multiple colours, thereby encompassing a broader range of individuals who perceive this form of synaesthesia.



Figure 3.3: Multi-colour word colouring rules selected based on literature review and relevance to this thesis.

### 3.4 Visualisation and Colouring Types

The AR application's visualisation approaches are selected with the objective of achieving a balance between user familiarity and control over the environmental elements.

One approach that is chosen is the direct colouring of the "paper", that is to say, the surface of the scanned real-world text (see at the top of Figure 3.4). This entails the display and manipulation of each letter in specific colours as if they are actually printed in those colours. This strategy is selected because it is consistent with the established conventions that are used in text presentation. People are familiar with this style from conventional print media, making it a convenient and straightforward alternative for users. The method's familiarity reduces the cognitive burden on the user.



An additional visualisation style is the addition of an extra layer for the presentation of the text (see at the bottom of Figure 3.4). This method involves the placement of the text on a controlled background, resulting in a uniform colour display. This style is of particular significance given that the background colour may affect colour perception [Ell15]. By isolating the text on a controlled background, the application ensures that users see the colours as intended.

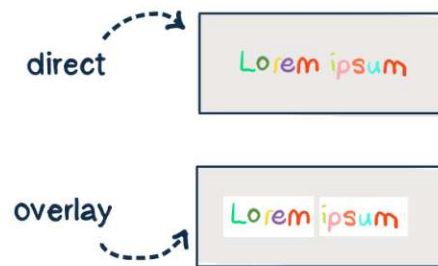


Figure 3.4: The two different visualisation types.

In order to facilitate the addressing of customers' diverse interests, a variety of colouring types is included, in recognition of the fact that associative and projective synaesthetes may have differing requirements (for instance see Figure 2.3 in Chapter 2).

A direct approach to colouring the text is to paint the letters themselves, which simulates writing the letters with a coloured pen.

Another option is colouring the background by applying colour to the area behind each letter, while the letter itself is left black, similar to how a highlighter pen is used.

The creation of a coloured outline represents a more subtle option, whereby only the letter's outline is coloured, while the interior remains black. This approach is less obtrusive, yet nevertheless serves to effectively draw attention to the text as well as to the colour.

These different colouring types (see Figure 3.5) enable users to select the style that is most aligned with their perception and preferences.

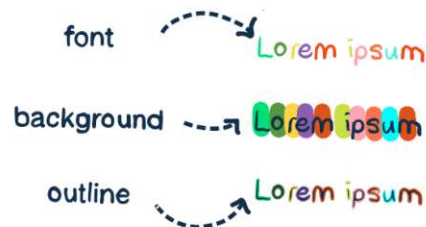


Figure 3.5: The three different colouring types.

## 3.5 Application Features

This section examines the main characteristics of the application, emphasising its features and functionality.

### 3.5.1 Colour Definition for Each Grapheme

Users first create a customised palette that affects later text rendering by using a colour picker to assign distinct hues to each grapheme (letters and digits).

### 3.5.2 Rule-Set Definition for Word Colouring

The application's key feature is its ability to apply several colouring rules to the text. The four selected rule-sets (single-colour: first letter, first vowel; multi-colour: individual, vowel gradient) are incorporated into the application.

To help with the selecting process of the above-mentioned possible rule-sets, the application shows the term "Sample Text" on both black and white backgrounds (see Figure 3.6). This enables users to quickly establish the best colour scheme for different background settings, simplifying the decision and customisation process.

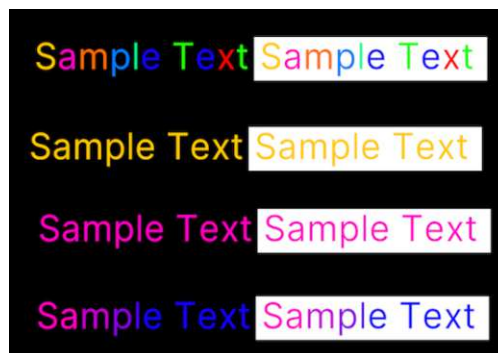


Figure 3.6: Details of a screenshot of the word-colouring screen. It displays the term "Sample Text" for each rule-set as a preview in the following order: individual, first letter, first vowel, vowel gradient.

### 3.5.3 Visualisation Style Definition

The application's visualisation and colouring capabilities cover five different types (see Figure 3.7).






	Texture Direct	Text Mesh Pro Overlay
Font		
Outline		
Background		

Figure 3.7: Various options for displaying recoloured text, organised by visualisation styles (columns) and colouring styles (rows).

The defined two visualisation styles are implemented as stated below:

- **Texture Direct:** Directly alters the scanned image's pixels, changing their colours in accordance with the rule-set specified. This solution keeps the document's original texture while applying the new colour palette.
- **Text Mesh Pro (TMP) Overlay:** Generates a new Unity game object, which serves as a container for components that define the behaviour and properties of an object, for each word. This game object is made out of a Unity image object that acts as a background with a TMP object placed on top, covering the complete bounding box for the word.

#### 3.5.4 Text Scan Functionality

The application's last major feature is its capacity to scan textual information with the device camera and display the recoloured words based on user-defined choices on the screen.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# SynVis Implementation

This chapter provides a comprehensive account of the implementation process for the application outlined in the previous chapter. It covers the technology stack, the user flow as well as the UI design, and specific algorithms and techniques.

## 4.1 Tech Stack

To implement the aforementioned system, a whole stack of technologies is used. Unity (Version 2022.3.20f1) is chosen as the major development environment, and C# as the programming language.

### 4.1.1 Unity Libraries

Several important Unity libraries are integrated into the project to provide robust performance and a state-of-the-art user experience:

- **Flexible Colour Picker:** This tool allows users to conveniently select and apply colours to each grapheme.
- **Newtonsoft JSON:** This library is used to efficiently serialise and deserialise JSON, which is essential for preserving user data such as grapheme colours and word colouring rules.
- **AR Foundation:** This package ensures compatibility between many devices and makes it easier to construct cross-platform AR experiences.
- **TextMesh Pro:** This package improves text rendering and visualisation in Unity, assuring high-quality text display. This high-quality text is critical for the in-app texts on the screen, as well as the visualisation of the recoloured text.

### 4.1.2 Text Detection and Recognition

Several options for implementing OCR and visualisation operations were investigated, with each bringing unique problems and capabilities.

Tesseract OCR was initially picked since it is open-source and has out-of-the-box capability on Windows. Tesseract was adapted for macOS and Android to provide cross-platform functionality. However, Tesseract was discovered to only support MONO scripting backend and not IL2CPP, which is required for iOS app development, resulting in its final rejection.

Another option considered was Vuforia. After discovering that the text recognition functionality was no longer supported and available, an alternative technique was tried which involved pre-scanning text and creating image targets for each word, with the aim of projecting recoloured words over the original text. However, this strategy proved unworkable because letters and words were not different enough to serve as unique image targets, resulting in projection mistakes and the decision not to proceed with Vuforia.

Finally, the "Open Source Computer Vision Library" OpenCV for Unity [Uni24] was selected due to its extensive capabilities and fit for the project's requirements. OpenCV for Unity uses the EAST method and Convolutional Recurrent Neural Network (CRNN) for text detection and recognition. Despite being a paid version, it was chosen for its dependability, and the fact that OpenCV plus Unity [Uni19] is outdated and only supports Windows and Android.

OpenCV for Unity was critical for creating numerous visualisation modalities, making the AR application as convenient and pleasant as feasible. The visualisation sought to closely mimic the original text without changing the font or background, instead recolouring exactly the words according to predetermined principles. This option provides the capability required to meet the project's objectives effectively.

### 4.1.3 Used Devices

With the determination of the requirements and capabilities for constructing the AR application, the implementation procedure is carried out on a Mac computer running macOS 14.5. The app is built with the Xcode application 15.2, which ensures compatibility and best performance on iOS devices.

The application is developed and tested on an iPhone running iOS 17.5.1. This combination of devices and software guarantees that the development environment is up-to-date with the newest tools and operating systems.

## 4.2 User Flow

The UI is designed to ensure a pleasant and straightforward user experience throughout the entire system.

Once the application is opened, the UI flow begins with an introductory screen, which prompts the user to provide their name. The user is sent to the menu screen after entering their name. In the next step, they use a colour picker tool to assign a distinct colour to each grapheme (letter and digit). After selecting the colours for the graphemes, the user is prompted to pick a rule-set that determines how the words are coloured. Following this, the user selects the visualisation mode. The final stage in the flow is to scan the text. To scan a text with the application, the user aligns the text within a specified recognition box. When the text is scanned, the application recolours it using the previously defined grapheme colours and rule-set.

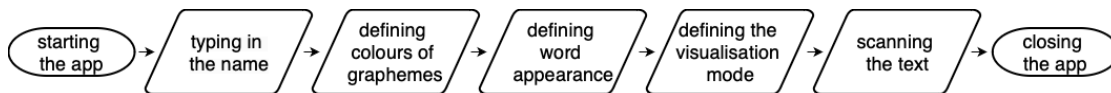


Figure 4.1: User flow diagram of the application, showing each step from starting the app to closing the app.

## 4.3 User Interface

The following points are included to offer a positive user experience and simple navigation of the application:

### 4.3.1 Intuitiveness and How-To Guidance

The AR app is designed with the goal of providing a straightforward user experience. Each touchpoint in the app is designed to be self-explanatory, allowing users to readily grasp and use the app without requiring additional instructions. In case the app's operation is unclear, a "how-to" instruction is provided (see Figure 4.2). This is especially critical for operations such as text scanning, which requires the user to rotate the phone and scan within a certain recognition area. In general, the software displays detailed information on the screen to help users through the procedure. This instructional help guarantees that users can comfortably utilise the app's functions without becoming frustrated.

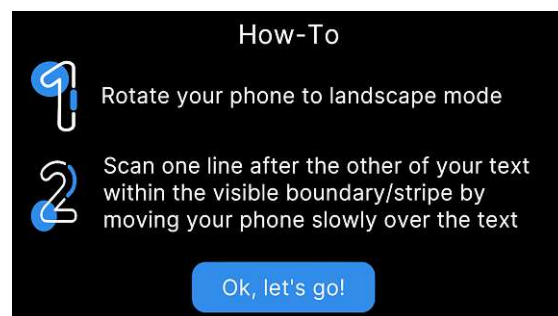


Figure 4.2: Detail of a screenshot of the how-to screen that is displayed prior to the scanning process.

### 4.3.2 Simplicity and Mode Indication

Simplicity is an important aspect in the UI design. The interface is maintained clear and uncomplicated to prevent distracting or confusing the user. This basic design (black background with white text) allows users to focus on the app's main capabilities without distractions, making interaction simple and straightforward. Additionally, to improve usability, the app displays the currently selected mode via visual hints on the buttons (see Figure 4.3). When a mode is active, the corresponding button is highlighted, ensuring that users are constantly aware of the app's current state. This feature avoids confusion and helps users understand and regulate their interactions with the app more efficiently.

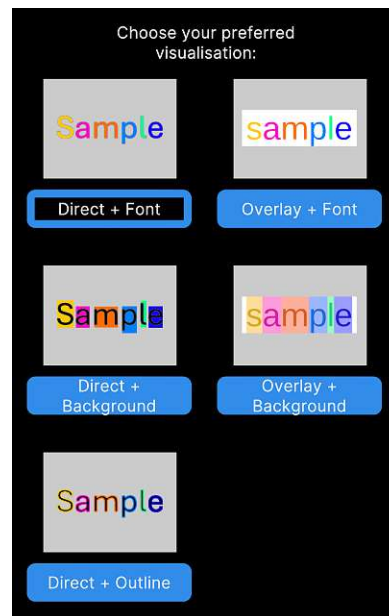


Figure 4.3: Detail of a screenshot of the visualisation mode selection screen, which shows that the *direct + font* option is currently selected.

### 4.3.3 Consistency and Colour Scheme

Consistency goes hand in hand with the aforementioned point. It is maintained across all UI components to create a cohesive design. The backdrop is always black, while the text is white, ensuring strong contrast and readability (see screens in Figure 4.4). This decision is intended to reduce discomfort caused by viewing letters in "wrong" colours while maintaining an appealing, clean appearance. The same idea is used for the app's logo, resulting in a streamlined visual identity. The text makes use of the same typeface and employs a uniform font size, based on the type of text (heading, paragraph, etc.), thus ensuring a consistent visual presentation. In addition, the colour blue (hex #308cea) is utilised frequently to highlight menu elements and buttons for the user. This consistency in button designs, font selections, and layout structures allows users to anticipate the behaviour of various elements based on past interactions.



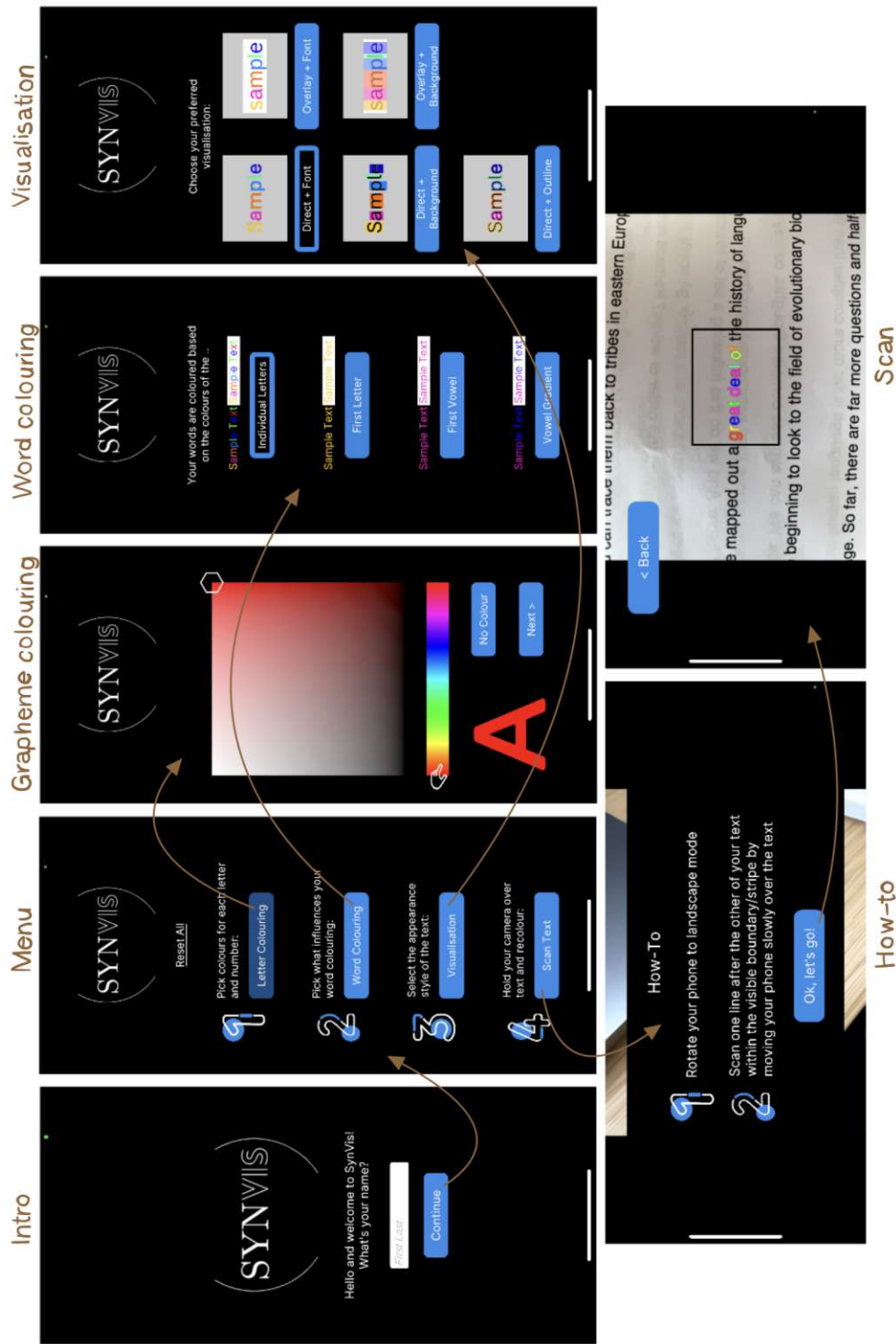


Figure 4.4: Screenshots of all application screens, labelled by name. User flow is indicated by brown arrows.

### 4.3.4 User Data Persistence

When the user closes the app, the software saves their data, using the name which is typed in when starting the app (see Figure 4.5). This feature allows users to resume their activities from where they left off without losing any progress. For example, if a user is in the middle of configuring colours, the app allows them to resume from the same position when they return. Setting graphemes to the precise observed colour is one activity where this feature is essential.

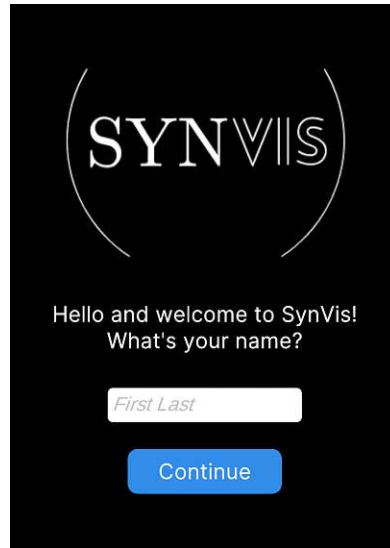


Figure 4.5: Detail of a screenshot of the introductory screen, which requires users to enter their first and last names in order to save their data.

## 4.4 Algorithms and Techniques

The code is divided into components, that include logic for colour definition, text detection, colouring, visualisation, storage, and UI testing for usability purposes. Because it is developed with design patterns in mind, such as the Factory Pattern, this structure allows for easy extension of the code, especially for the rule-sets with are of great importance.

### 4.4.1 User Data Structure

The user data file is a JSON file that contains every configuration that a user defines as shown in Listing 4.1. This consists of the colours for each grapheme, word colouring rules, and visualisation options (type and mode). The latter mentioned parameters are recorded as enumerations, guaranteeing an organised and uniform manner.

```

1 {
2   "graphemeColours": {
3     "A": {
4       "r": 1.0,
5       "g": 0.0,
6       "b": 0.7687535
7     },
8     "B": {
9       "r": 0.1432643,
10      "g": 0.0,
11      "b": 1.0
12    },
13    "C": {
14      "r": 0.0,
15      "g": 0.973161459,
16      "b": 1.0
17    },
18    .
19    .
20  },
21  "graphemeMode": 1,
22  "colouringType": 1,
23  "colouringMode": 1
24 }

```

Listing 4.1: Excerpt of a User Data JSON file, displaying colours per grapheme, saved word colouring rule (grapheme mode), visualisation style (colouring type), and colouring style (colouring mode).

#### 4.4.2 Detection and Colourisation Procedure

Text detection and recognition is carried out using the TextDetectionAndRecognition-CRNN algorithm offered by OpenCV for Unity.

Whenever a new camera frame is recognised, the text detection procedure begins. The algorithm analyses the frame for the presence of text, making use of the CRNN model's ability to reliably recognise letters and words inside the image.

After the text is identified and recognised, the OverlayManager class does word-by-word processing. This phase involves applying the desired visualisation settings, which might be *texture direct* or *TMP overlay*. The OverlayManager class renders the identified text based on these parameters.

#### 4.4.3 Texture Direct Pixel Manipulation

The *texture direct* manipulation visualisation option (see Figure 4.6) requires a thorough method to ensure that each grapheme within a word is correctly recognised and coloured.

This approach starts by inspecting the full word's bounding box and then processes each grapheme independently.

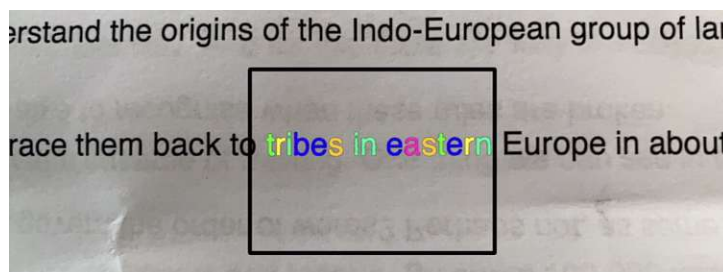


Figure 4.6: Detail of a screenshot during the text scanning process in the *texture direct* mode.

Initially, the bounding box of the word is examined to define the region of interest (ROI). A variety of image processing algorithms are used to segregate each grapheme within the bounding box.

In order to smooth out the image and minimise noise, the ROI is first processed by using a Gaussian blur, see Listing 4.2 line number 1.

Next, the colours inside the ROI are converted into greyscale to prepare for the next step, see Listing 4.2 line number 2. The black regions of the ROI are then isolated using a threshold function, which successfully isolates the text from the backdrop, see Listing 4.2 line number 3.

Then, a structural element is generated, and dilation, a morphological operation that expands the boundaries of regions in a binary or greyscale image, is used to improve the bounds of the text elements, see Listing 4.2 line numbers 4 and 5. This procedure makes items larger and fills in small holes or gaps in the image, which is necessary for combining characters, such as the body and dot of the letter "i".

Following the preparation of the text elements, each grapheme is distinguished by its contours, see Listing 4.2 line number 6. The number of contours is then compared to the number of letters in the recognised word to guarantee accuracy. If the counts are equal, the contours are arranged in an array from the leftmost contour on the  $x$ -axis to the rightmost contour.

After the contours have been successfully sorted, the user-defined word colouring rule is retrieved. According to this rule, the colours for each letter are obtained from the WordColouriser class. Each contour is then coloured using the appropriate colour from the colour array, ensuring that each grapheme is presented in the correct colour as indicated by the user's preferences.

```

1  Imgproc.GaussianBlur(roiLetterMat, roiLetterMat, new Size(5, 5), 0);
2  Imgproc.cvtColor(roiLetterMat, roiLetterMat, Imgproc.COLOR_BGR2GRAY);
3  Imgproc.threshold(roiLetterMat, roiLetterMat, 100, 255, Imgproc.
    THRESH_BINARY_INV + Imgproc.THRESH_OTSU);
4  kernel = Imgproc.getStructuringElement(Imgproc.MORPH_RECT, new Size
    (2.5, 16));
5  Imgproc.dilate(roiLetterMat, roiLetterDilateMat, kernel);
6  Imgproc.findContours(roiLetterDilateMat, contours, hierarchy, Imgproc
    .RETR_EXTERNAL, Imgproc.CHAIN_APPROX_SIMPLE);

```

Listing 4.2: Code snippet showing the OpenCV Code for extracting individual graphemes from the word's region of interest.

#### 4.4.4 TMP Overlay

The *TMP overlay* visualisation option (see Figure 4.7) starts by cloning a reusable, pre-configured game object, called a prefab which contains a TMP text object, which is an advanced text component that provides high quality text rendering and rich text formatting. This text object is displayed on a white background, which is represented as an image object in Unity. This prefab serves as the basis for presenting the identified text.

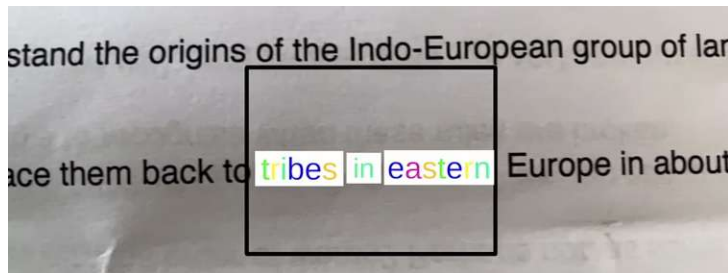


Figure 4.7: Detail of a screenshot during the text scanning process in the *TMP overlay* mode.

After cloning, the newly created game object is translated into on-screen space and positioned exactly above the bounding box of the detected word, ensuring that it is perfectly aligned with the position of the text in the camera frame.

The WordColouriser class then processes the detected word in accordance with the user's word colouring rule selections, see Listing 4.3 line number 1 and 3.

This coloured word is then set as the prefab's text using the rich text format, see Listing 4.3 line number 2 and 4. This style provides for extensive text formatting and colouring, ensuring that each letter is presented in the appropriate colour.

Finally, the text is set to automatically span the whole bounding box, ensuring that the text appears as large as it can be within the available area. This assures optimum readability.

```
1 if (userData.GetColouringMode() == COLOURING_MODE.FONT)
2     colouredWord += $"<color=#{(int)(graphemeColour.r * 255):X2}{(int)
3         } (graphemeColour.g * 255):X2}{(int)(graphemeColour.b * 255):X2
4         }>{text[i]}</color>";
3 else
4     colouredWord += $"<mark=#{(int)(graphemeColour.r * 255):X2}{(int)
5         } (graphemeColour.g * 255):X2}{(int)(graphemeColour.b * 255):X2}
6         aa>{text[i]}</mark>";
```

Listing 4.3: Code snippet for applying the correct colours to letters in TMP.

# Testing and Evaluation Design

To evaluate the approach, a test strategy including both qualitative and quantitative evaluation is developed.

Benchmark tests are conducted to evaluate the technical aspects of the application, which includes for instance the time taken from scanning text to displaying the text in a recoloured visualisation on the screen. User studies are conducted to evaluate the readability and visualisation preference of users, as well as the usability of the whole application. Expert interviews are conducted as a qualitative research method to obtain in-depth knowledge and feedback about SynVis.

## 5.1 Benchmark Testing

This section explains how benchmark testing is conducted on the AR system's two visualisation modes, *texture direct* and *TMP overlay*. The tests focus on three major performance indicators.

### 5.1.1 Performance Metrics

The benchmark test metrics are (1) frames per second (FPS) / frame time (FT), (2) response rate (RR), and (3) error rate (ER).

In order to evaluate the system's rendering performance, first the FPS / FT is used, which gives a clear indication of how smoothly the visualisation is presented on screen.

Second, the RR measures the amount of time that passes between when the text is first detected and when the visualisation is entirely displayed on the screen. This statistic is essential for assessing how well the system processes and displays text in real time.

Third, the ER focuses on the accuracy of the actual visual outputs compared to expected outputs, with the assumption that the OCR is fully functioning. Because the *texture*

*direct* and *TMP overlay* approaches use different visualisation algorithms, the ER for this investigation is very relevant. The ER is calculated based on the correctness of the visual representation.

### 5.1.2 Implementation of Testing Environment and Tools

A custom logger called "StatsLogger.cs" is created particularly for Unity to record FPS (per frame) and RR (per word). Figure 5.1 shows how it is implemented to switch between these two tracking modes. Furthermore, full-screen captures using Apple's inbuilt video-screen capture feature are obtained in order to assess the ER of the visualisations and examine them later.



Figure 5.1: Checkbox added to the menu screen for selecting whether to log FPS or RR during benchmark testing.

### 5.1.3 Testing Procedure

Three sessions, each lasting two minutes, are used to assess each performance metric for each visualisation approach. To ensure that other performance logs are not weighted, only one performance statistic from each session is examined at a time. With the identical device (Apple iPhone 11 Pro) and predefined text settings (font: Helvetica Regular font size: 12, and line space: 3), a controlled test environment is set up. The text (see Appendix 8) is scanned word by word, line by line, from left to right, throughout the test.

## 5.2 Pilot Study

Before conducting the user study, a pilot study was set up to ensure all organisational aspects are in order and to see if everything runs as intended. To provide participants in the user study a clear understanding of how long it will take, the pilot study also aimed to quantify how long it takes for each participant.

Three participants (2 males, 1 female), aged between 24 and 27 years, volunteered for this study.



### 5.2.1 User Study Protocol

Before the study begins, all necessary equipment is prepared. This includes an iPhone 11 Pro with the SynVis app loaded, a text for the experiment (see Appendix 8), and a questionnaire (see Appendix 8).

Upon arriving, each participant is given a brief introduction to the app. This introduction includes subjects such as grapheme-colour synaesthesia, the app's purpose, and the study's goals. This makes sure that before starting the tasks, participants understand the context and goals clearly.

After the introduction and when the participant feels ready, they are given the smartphone with the application installed. Participants are then asked to assign colours to each grapheme (see Figure 5.2). Subsequently, they are required to select a specific word colouring rule within the app.

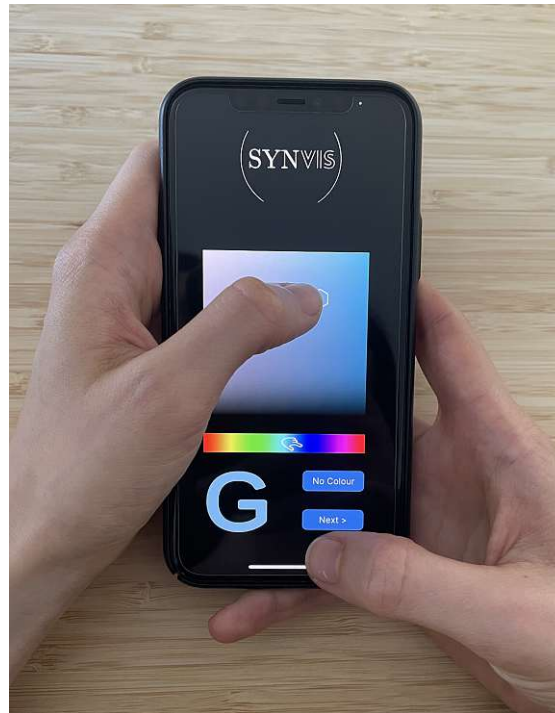


Figure 5.2: Photograph depicting the process of selecting a colour for each grapheme, with the current iteration featuring a shade of blue for the letter "g".

After that, participants are instructed to try out all the app's visualisation options independently (see Figure 5.3). This entails selecting each option and using it for scanning and reading some text from the prepared text book. Through this stage, participants are able to test and assess the various visualisation options offered by the app.

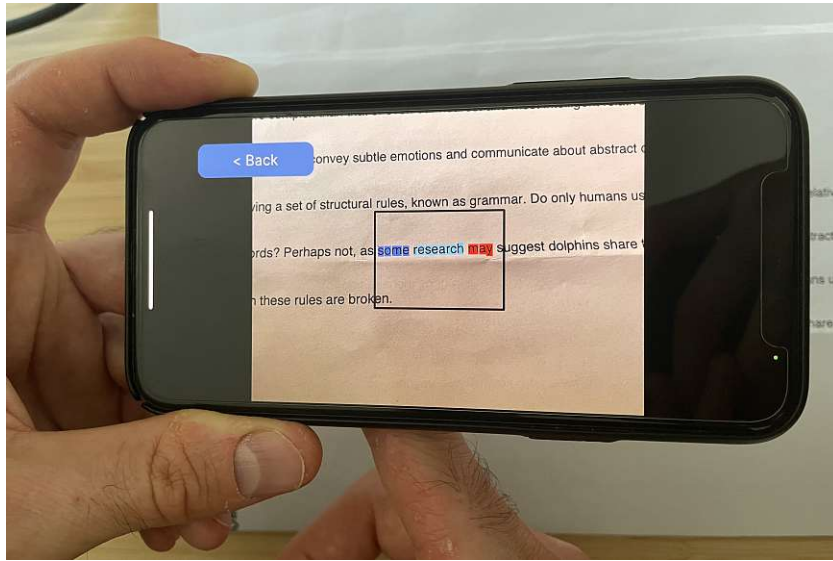


Figure 5.3: Photograph depicting a participant engaged in the text scanning process, utilising the *texture direct + background* visualisation option with colouring based on the first letter colour rule.

After trying all the visualisation modes, participants are invited to complete a short questionnaire (see Appendix 8 for the entire questionnaire) in their selected language (German or English), which contains four sections:

- Demographics
- Rating the readability of the five different visualisation modes
- SUS
- Open-ended question for individual feedback and improvement ideas

After completing the questionnaire, participants are thanked for their time and effort, which marks the end of their involvement in the study.

### 5.3 User Study

A user study was conducted to determine the app's technological viability. In addition to gathering usability input, the various visualisation styles were assessed. The methodology employed was identical to that utilised in the pilot study, see Section 5.2.1.

### 5.3.1 Participants

This study included twelve volunteers (5 females and 7 males). They were invited through word-of-mouth. The study ran from the end of June until the beginning of July in 2024.

None of the subjects experienced grapheme-colour synaesthesia. All subjects had normal or corrected-to-normal vision, with the latter group wearing glasses or lenses.

The participants ranged in age from 23 to 60 years, with a mean of 38,333 and a standard deviation of 16,289.

Every participant needed roughly 20 to 30 minutes for the session.

### 5.3.2 Data Collection

Throughout the usage of the app, participants' remarks are collected via taking notes using the think-aloud method, which involves verbalising their ideas and feelings as they interact with the app. During the try-out period, participants' behaviour and interactions with the app are also observed. This observational data gives contextual information about their experiences of and with the app. Data is also collected from completed questionnaires.

## 5.4 Expert Interviews

As an essential component of the qualitative research process, expert interviews are carried out to supplement the quantitative evaluation of the prototype and findings. The purpose of these interviews is to get contextual insights and in-depth knowledge. This method not only improves comprehension of the technical components, but it also aids in providing a full picture that is required for holistic study results.

### 5.4.1 Selection of Experts

The research questions as well as the hypotheses were carefully considered in the selection of participants for the expert interviews. This made sure that the knowledge acquired is highly useful and relevant to the study's findings.

A synaesthesia research expert was selected to mainly address the first research question, namely the formalisation of synaesthetic perceptions into rule-sets. This participant's extensive scholarly experience is critical in validating the literature findings and interpretations acquired throughout the thesis' earlier stages. The major responsibility of this expert was to conduct a critical examination of the identified concepts, assuring their validity and relevance to current scientific understanding. Furthermore, the expert was asked to evaluate the implemented prototype, providing feedback on its design and functionality. This evaluation seeks to determine the prototype's validity in reproducing synaesthetic experiences, allowing to draw findings for the research questions via a combination of theoretical insight and practical assessment.

The second participant, a person with grapheme-colour synaesthesia, was selected to offer a user-centric view on the prototype's utility and success. Engaging with a potential user who has direct experience with synaesthesia allows the research to integrate personal input into how the prototype functions in real-world scenarios. This participant assessed the prototype to see if it can accurately reproduce their synaesthetic sensations, and gave input on both its strengths and places for development. Their input is essential to understand the real-world use of the prototype and to meet the needs of the synaesthetic community. This participant's input helps to address both research questions by providing an in-depth view of the user experience.

### 5.4.2 Interview Procedure

To invite experts for interviews, a one-page overview of the project is developed and sent. The goals of the project are outlined in this document, which also emphasise the format of the interview.

The interview with the expert in synaesthesia research is done via Zoom. In order to enable a thorough discussion of the application's functioning and research implications, a demonstration of the application is given during this virtual session using video and screen-sharing capabilities. The interviewing guideline is shown in Figure 5.4.

In contrast, the interview with the person experiencing grapheme-colour synaesthesia is conducted in person. In this setting, the expert engages with the application directly and gives immediate input on its efficacy and correctness, depending upon their colour perceptions. This hands-on experience is crucial for getting real user insights and accurately analysing the application's potential for reproducing synaesthetic sensations. The interviewing guideline is shown in Figure 5.5.

The interviews are conducted in a semi-structured manner that permit both guided questions and unstructured talk. This strategy makes it possible to gather focused data while giving experts the freedom to delve deeper into subjects that come up during the discussion.

After the interviews, each participant receives a PDF file including a personalised, hand-drawn thank-you note. The purpose of the gesture is to show a heartfelt appreciation for their efforts and important contributions to the project.

Zoom is used to record the first semi-structured interview, while Photo Booth (a Mac application for taking photos and videos using the built-in camera) is used for the second.

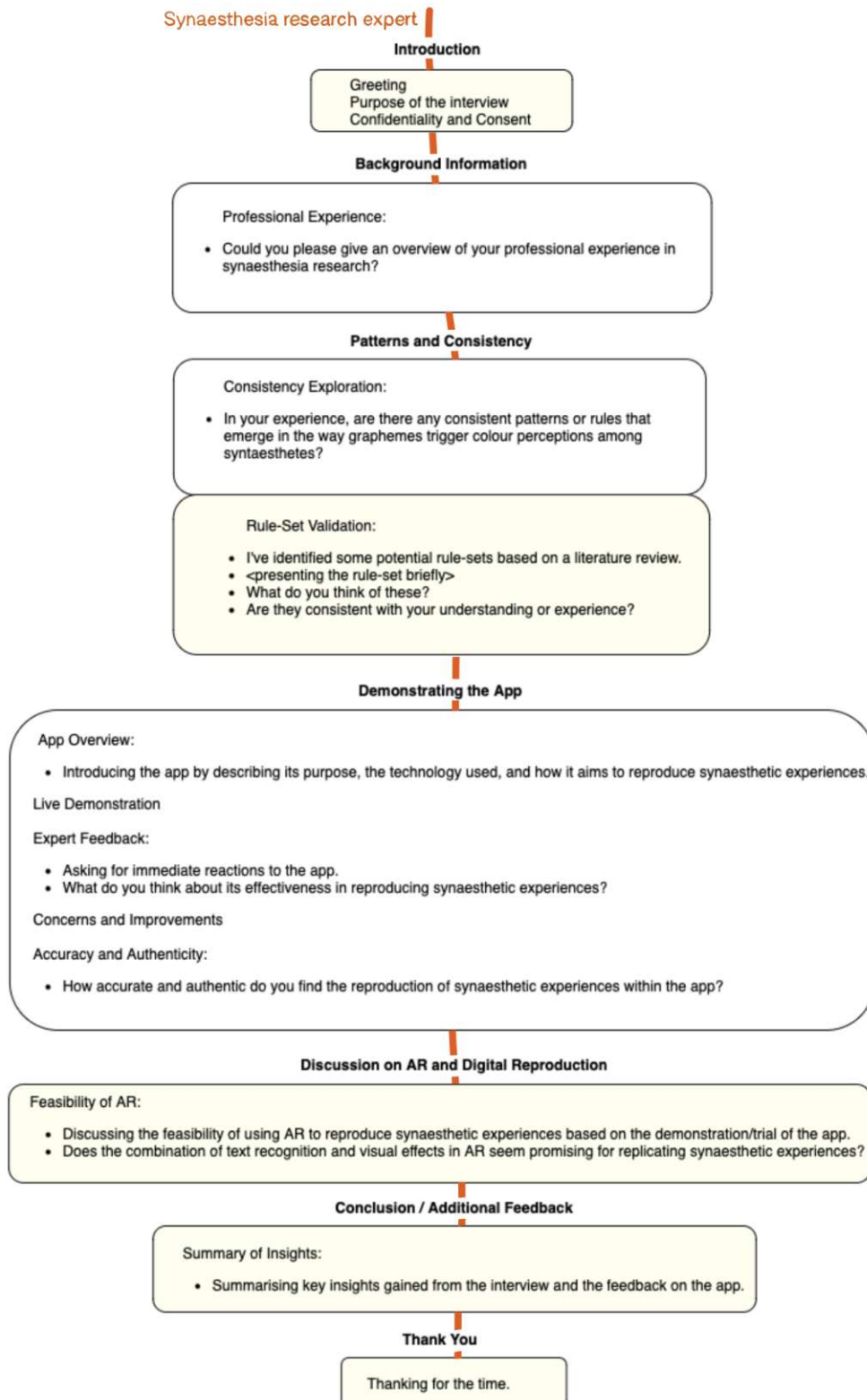


Figure 5.4: Interview guide for the semi-structured interview with the synaesthesia researcher. The yellow background indicates that the questions are identical to those posed in the second expert interview.

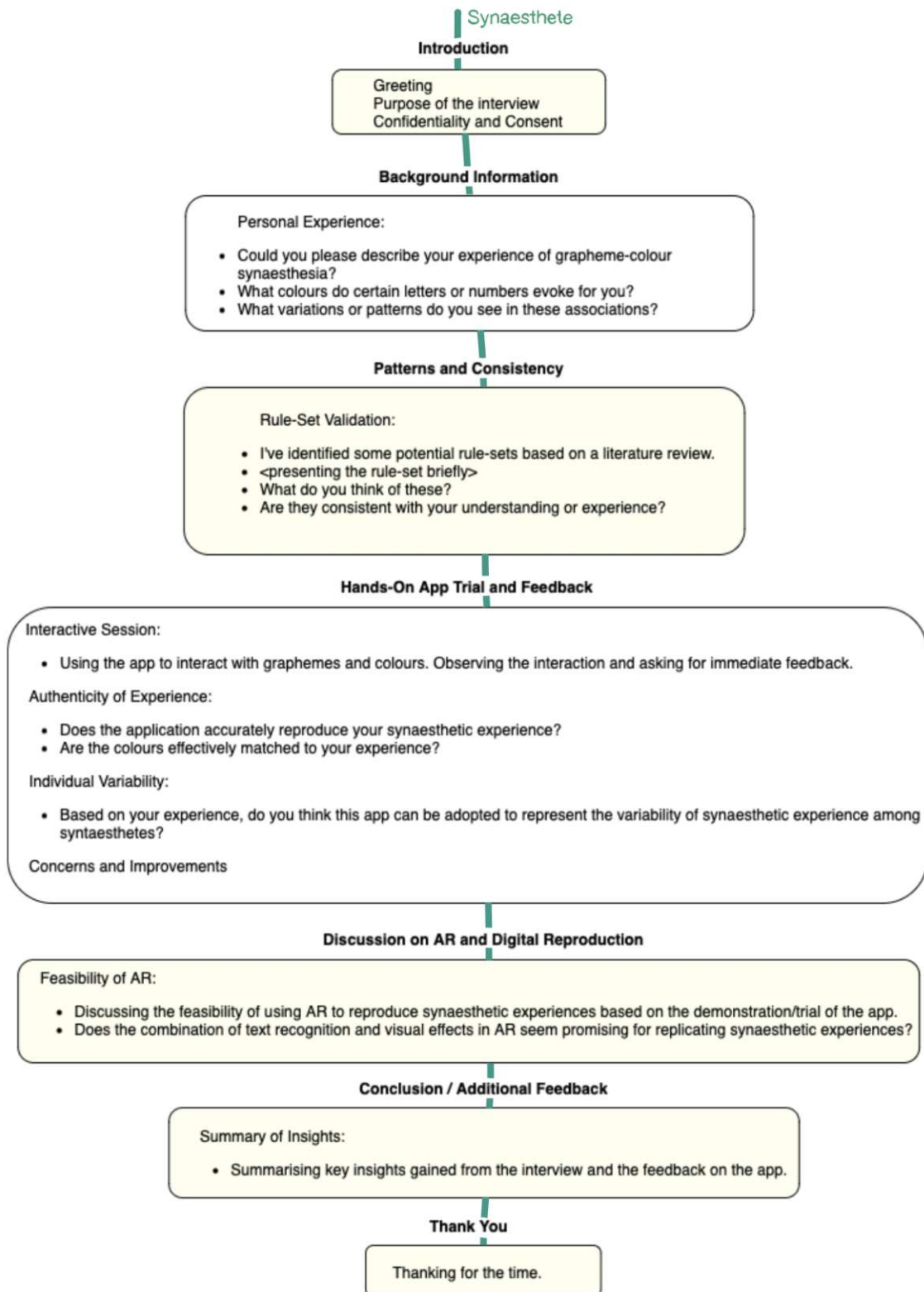


Figure 5.5: Interview guide for the semi-structured interview with the synaesthete. The yellow background indicates that the questions are identical to those posed in the second expert interview.

# CHAPTER 6

## Results

Addressing the research questions outlined in Section 1.2, this section presents the findings of the testing phase.

To address **RQ1** (*How can we formalise and reproduce individual grapheme-colour synaesthetic experiences on a digital screen?*), a comprehensive literature review is conducted, trying to find and formalise common patterns of rules, and challenging these findings through the two different expert interviews (synaesthesia researcher and synaesthete).

In order to answer **RQ2** (*What technical developments are necessary to align an AR visualisation with the experience of grapheme-colour synaesthesia?*), benchmark tests for FT, RR and ER, a user study to evaluate the experience of the application and the readability of the superimposed and recoloured text, and expert interviews, in particular the try-out session with the synaesthete, are conducted.

### 6.1 Quantitative Analysis

The quantitative data analysis is carried out with Jeffrey's Amazing Statistics Program (JASP) [JAS18].

The following results are reported as statistically significant at  $p < 0.05$ .

#### 6.1.1 Benchmark Tests

To address RQ2, it is necessary to evaluate the performance of the application. To this end, the following hypothesis is tested:



**H2-a:** “The performance of the AR application, measured in terms of frame rate and latency, will be within acceptable ranges (maintaining a frame rate above 30 FPS and a latency below 100 milliseconds (ms)) when rendering grapheme-colour synaesthetic visualisations.”

The descriptive statistics of the performance metrics defined in Section 5.1.1 are shown in Table 6.1. The metrics are divided in two sections each, naming the different visualisation modes, *texture direct* and *TMP overlay*.

	Frame Time (FT)		Response Rate (RR)		Error Rate (ER)	
	Texture	TMP	Texture	TMP	Texture	TMP
Median	258.804	288.898	133.000	131.000	30.000	7.000
Mean	255.610	278.909	137.239	132.743	30.000	7.333
Std. Deviation	44.747	47.845	54.749	55.768	0.000	0.577
Range	197.041	208.601	336.000	404.000	0.000	1.000
Minimum	136.293	124.732	8.000	8.000	30.000	7.000
Maximum	333.333	333.333	344.000	412.000	30.000	8.000

Table 6.1: Descriptive statistics of benchmark values, all of which are reported in milliseconds (ms)

FPS are converted to FT for the benchmark analysis due to the linear relationship with performance.

As FPS grows, the effect of additional frames on perceived performance diminishes. The non-linearity complicates statistical analysis and interpretation.

FPS is inversely correlated with FT, which is the amount of time (measured in milliseconds) required to render a single frame. Converting FPS to FT yields a linear measure in which each unit of time has a consistent and interpretable impact on performance. The conversion is done as followed:

$$\text{FT in (ms)} = \frac{1000}{\text{FPS}} \quad (6.1)$$

The normality tests, using the Shapiro-Wilk method, yield p-values less than 0.001 for all measurements in both options, see Table 6.2. This means that neither of the data distributions is normally distributed.

	FT		RR		ER	
	Texture	TMP	Texture	TMP	Texture	TMP
Shapiro-Wilk (W)	0.983	0.914	0.968	0.963	NaN	0.750
P-value of Shapiro-Wilk	< .001	< .001	< .001	< .001	NaN	< .001

Table 6.2: Normality tests of benchmark values (NaN means: all values are identical)



Given the divergence from normality, the metrics are evaluated using the non-parametric paired-samples statistical test called the Wilcoxon signed-rank test.

This test is used to compare the performance of *texture direct* mode with *TMP overlay* mode for the FT measure (see Table 6.3). The test shows a statistically significant difference between the two modes ( $z = -7.139$ ,  $p < 0.001$ ). What this means is that *texture direct* mode renders frames more efficiently, resulting in a smoother visual experience. The high degree of significance indicates that the differences seen in the descriptive statistics are unlikely to arise by coincidence. The rank-biserial correlation, which measures the effect size, is calculated to be  $r_{rb} = -0.452$ , with a standard error of 0.063 indicating a medium negative effect size according to Cohen's conventions [Coh88].

Measure 1	Measure 2	W	z	p	Rank-Biserial Correlation	SE Rank-Biserial Correlation
FT-Texture	FT-TMP	15144.000	-7.139	< .001	-0.452	0.063
RR-Texture	RR-TMP	555585.500	3.457	< .001	0.106	0.031

Table 6.3: Wilcoxon signed-rank test - FT and RR

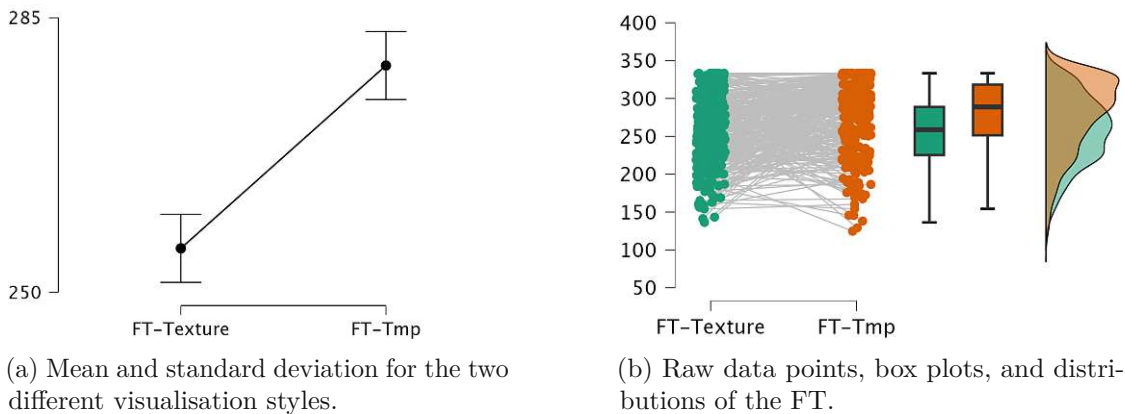
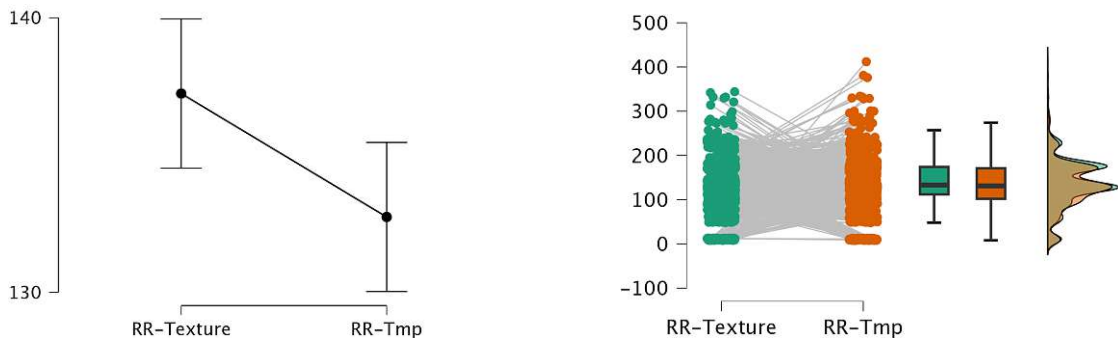


Figure 6.1: Figures illustrating details of the FT benchmark.

As previously conducted for the FT, the Wilcoxon signed-rank test is used to compare the two different visualisation modes for the RR (see Table 6.3). The test indicates a statistically significant better RR for the *TMP overlay* mode than for the *texture direct* mode ( $z = 3.457$ ,  $p < 0.001$ ), indicating a more responsive user experience, see Figure 6.2. Measuring the rank-biserial correlation results in an effect size of  $r_{rb} = 0.106$ ,

with a standard error of 0.031 indicating a small positive effect size according to Cohen's conventions [Coh88].



(a) Mean and standard deviation for the two different visualisation styles.

(b) Raw data points, box plots, and distributions of the RR.

Figure 6.2: Figures illustrating details of the RR benchmark.

The detailed examination of RRs reveal an interesting pattern. Specifically, there are situations where response times are nearly zero milliseconds (see Figure 6.2b). This occurs when a user moves the camera from the rightmost word on a line to the leftmost word on the next line, as no words are detected and therefore the entire detection, colouring and display algorithm does not need to be performed.

For each run, 167 words plus two digits are scanned to determine the ER. It can be seen in Table 6.1 that the *texture direct* mode has more errors or problems with an average percentage of 17.75%, while the *TMP overlay* mode has a lower ER with an average percentage of 4.34%, indicating a higher dependability.

Interestingly, across all criteria, *texture direct* mode has smaller standard deviations and narrower ranges than *TMP overlay* mode. This means that *texture direct* mode has less variability and is more consistent overall.

### 6.1.2 Pilot Study

The principal findings of the pilot study are as follows:

#### Input Field Adjustment

For ease of usage during the development process, a predetermined name was utilised in the input field on the introduction screen. However, during the study, it became clear that the pre-filled name box was inconvenient because participants had to erase the pre-filled name before typing in their own. To solve this, the name field was replaced with placeholder text, prompting users to enter their names directly, with no pre-filled information. This change shortens the procedure and minimises any early friction for participants.

## Display Adjustments

The phone was equipped with a privacy glass screen (which is a protective layer that limits the viewing angle so that only the person directly in front of you can see the display, while others see a darkened screen) from previous use, but it has been found that its darkening effect interfered with visibility and the correct colour settings in the AR app. In order to provide participants with an accurate and distortion-free view of the AR content, this privacy screen protecting glass was removed for the user study. Furthermore, to ensure optimal visibility of the AR information, the phone's brightness was adjusted to 100% throughout all experiments.

## Focus Mode Activation

In response to the observation that users were distracted by notifications from other applications, a dedicated "focus mode" was set up and enabled on the mobile device. This mode disables all alerts, and push notifications, thereby allowing users to interact with the AR software without interruption.

## Questionnaire Format

During the pilot study, it was observed that when the questionnaire was accessed in portrait mode, users were not aware that they had to swipe left in order to see the full 1-5 Likert scale, resulting in only the first three values being visible. To address this issue, participants are instructed to complete the questionnaire in landscape mode, thereby enabling them to see and select from a comprehensive range of options.

These improvements are crucial in ensuring that the user study operates as well as possible, offering a dependable and efficient experience for the participants.

### 6.1.3 User Study

In order to respond to RQ2, it is additionally necessary to evaluate two factors: firstly, the user experience of the application and secondly, the readability of the visualised, recoloured text. The following hypotheses are tested:

- H2-b:** *“Users will find the AR application intuitive and easy to use, as indicated by achieving a score of at least 80 on the SUS when visualising grapheme-colour synaesthetic experiences.”*
- H2-c:** *“Combining text detection and recognition of printed text in AR with various methods for visualising the reprinted text will successfully represent grapheme-colour synaesthetic experiences in AR.”*

Furthermore, H2-c is assessed from a psychological perspective through a trial session with a synaesthete (see 6.2).

The findings gathered from the user study, including the analysis of the SUS rating, visualisation preferences, and the readability evaluation, are presented in this section.

### System Usability Scale

Because the SUS questionnaire utilised was previously developed, refined, and verified, it is interpreted in accordance with the guidelines provided by Brooke [Bro95].

The SUS scores are determined by appropriately adding each participant's responses and multiplying them by 2.5. This gives a mean SUS score of 88.75, a standard deviation of 7.797, the range of 22.5 and the interquartile range of 10.625. These values are illustrated in Figure 6.3. According to the Sauro-Lewis curved grading system [LUM15], this score falls inside the A+ range (84.1-100), indicating excellent usability.

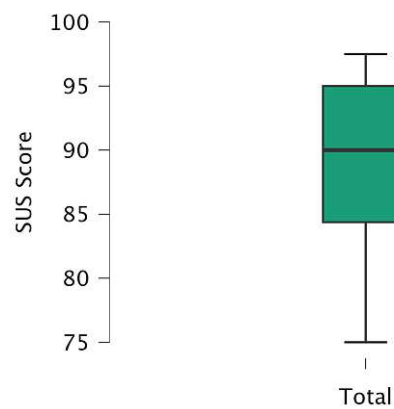


Figure 6.3: Distribution of the SUS scores of the 12 participants for SynVis.

### Visualisation Preferences

Participants were asked to choose their preferred visualisation style from two options. The majority (91.7%) picked *texture direct* visualisation. In the text field provided, they gave the following reasons for the choice of this particular type:

- *"Standard view mode"*, *"closer to the original"* and *"more appealing"*: This mode seems familiar and comfortable to participants, since the visualisation nearly matches the original text format and looks better than in the other mode.
- *"Colours show up better"*: Participants state that the colours seem more vibrant and sharp.
- *"Better readability"*: The content becomes simpler to read than with the other mode, and the letters are clearly distinguishable, according to the participants.
- *"Less cluttered background"*: Some say, that the background is not congested, which improves legibility.

In comparison, 8.3% of responders favoured the *TMP overlay* mode. The reason mentioned is:

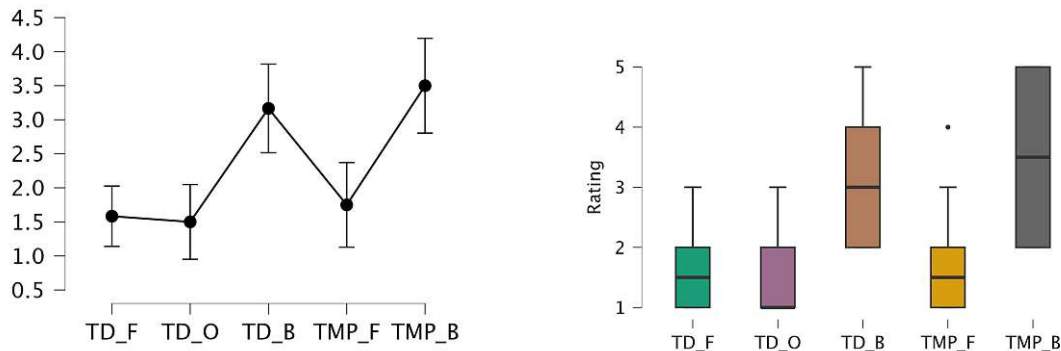
- Optimal contrast: "*The contrast is optimal due to the background field, if the background is not perfectly white.*"

## Readability

Using a school grading system, participants assess the readability of the five different visualisation options (*texture direct + font*, *texture direct + outline*, *texture direct + background*, *TMP + font*, *TMP + background*) from 1 (very good) to 5 (very poor). Having a look at the descriptive statistics in Table 6.4 and the box plots in Figure 6.4b, the *texture direct + outline* mode has the highest reading rating, with an average grade of 1.5. It is closely followed by *texture direct + font* with an average grade of 1.5833 as well as *TMP + font* with 1.75.

	Rating				
	TD_F	TD_O	TD_B	TMP_F	TMP_B
Median	1.500	1.000	3.000	1.500	3.500
Mean	1.5833	1.500	3.167	1.750	3.500
Std. Deviation	0.669	0.674	1.115	0.965	1.314

Table 6.4: Descriptive statistics of readability scores, all of which are reported in school grades ranging from 1 (very good) to 5 (very poor).



(a) Mean and confidence interval of 95% for the five different visualisation options.

(b) Box plots for the five different visualisation options.

Figure 6.4: Readability scores (on a scale from 1 to 5) where TD\_F is *texture direct + font*, TD\_O is *texture direct + outline*, TD\_B is *texture direct + background*, TMP\_F is *TMP overlay + font* and TMP\_B is *TMP overlay + background*. Lower scores indicate better readability.

Since Shapiro-Wilk tests for normality yield p-values less than 0.05 for all five visualisation options, the data deviates from normality, see Table 6.5.

	Rating				
	TD_F	TD_O	TD_B	TMP_F	TMP_B
Shapiro-Wilk (W)	0.768	0.732	0.859	0.778	0.818
P-value of Shapiro-Wilk	0.004	0.002	0.048	0.005	0.015

Table 6.5: Normality tests of readability scores

The Friedman test indicates significant differences in ratings between the five visualisation options,  $\chi^2(4) = 26.022, p = < .001$ . Post-hoc pairwise comparisons using the Conover test with Bonferroni correction reveals significant differences between *texture direct + font* and *texture direct + background* ( $p_{bonf} = 0.037$ ), *texture direct + font* and *TMP overlay + background* ( $p_{bonf} = 0.013$ ), *texture direct + outline* and *texture direct + background* ( $p_{bonf} = 0.016$ ), *texture direct + outline* and *TMP overlay + background* ( $p_{bonf} = 0.006$ ) and *TMP overlay + font* and *TMP overlay + background* ( $p_{bonf} = 0.037$ ), see Table 6.6 and Figure 6.4a.

		T-Stat	df	$W_i$	$W_j$	p	$p_{bonf}$
TD_F	TD_O	0.292	44	27.000	25.000	0.772	1.000
	TD_B	3.066	44	27.000	48.000	0.004	0.037
	TMP_F	0.365	44	27.000	29.500	0.717	1.000
	TMP_B	3.431	44	27.000	50.500	0.001	0.013
TD_O	TD_B	3.358	44	25.000	48.000	0.002	0.016
	TMP_F	0.657	44	25.000	29.500	0.515	1.000
	TMP_B	3.723	44	25.000	50.500	< .001	0.006
TD_B	TMP_F	2.701	44	48.000	29.500	0.010	0.098
	TMP_B	0.365	44	48.000	50.500	0.717	1.000
TMP_F	TMP_B	3.066	44	29.500	50.500	0.004	0.037

Table 6.6: Conover's post hoc comparisons - visualisation option

## 6.2 Qualitative Analysis

In order to address RQ1, it is necessary to obtain qualitative feedback from both an expert in synaesthesia research (R) and a synaesthete (S). During this process, the following hypothesis is tested:

**H1-a:** “It is possible to identify consistent patterns in grapheme-colour synaesthetic experiences for the majority of synaesthetes that suggest potential rules for digital reproduction.”





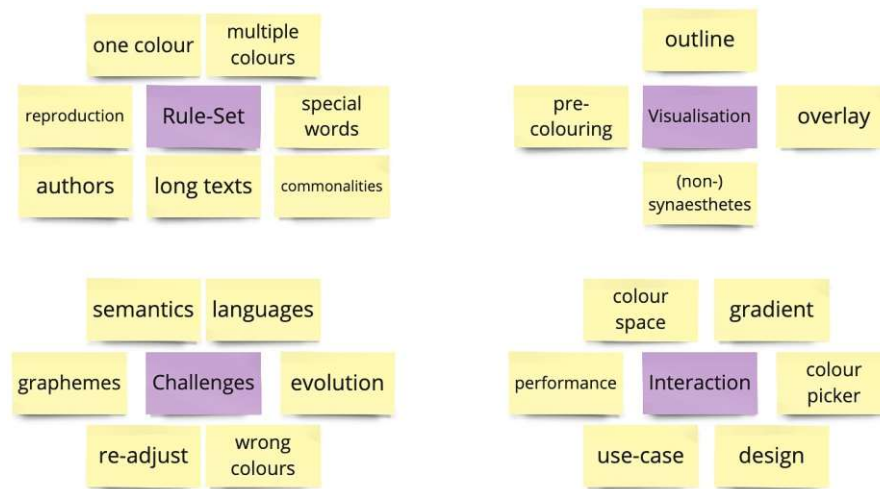


Figure 6.6: Screenshot of the Miro board displaying the identified themes (Rule-Set, Visualisation, Challenges, and Interaction) along with their grouped subtopics.

### Rule-Set

The thematic analysis of the grapheme-colour synaesthesia rule-sets reveals numerous patterns corresponding to different forms of synaesthetic experience, ranging from simple single-colour associations up to complex multi-colour mappings.

The examination reveals that the words in grapheme-colour synaesthesia *"would typically appear in one colour"* (R). In this context, the term "typical" refers to the majority of individuals who experience grapheme-colour synaesthesia. This is based on the observation that *"the colour is usually determined, either by the first letter or the prevailing vowel"* (R).

Compared to rule-sets that only include one colour, numerous colours are linked to other forms of synaesthesia, such as for instance ticker-tape synaesthesia. What needs to be noted here is that besides this, the expert who perceives grapheme-colour synaesthesia has a form with multiple colour appearance, who perceives vowels with colours and consonants without colours, so *"if you have a word with several vowels, it will be cross-faded in the consonants between them"* (S).

Special words and sequences, such as weekdays and months, typically comply with a specific colour order. Furthermore, it is mentioned as a side note, that the initial letters of male names are frequently recognised in colours of blue, whilst for female names first letters are connected with pink.

Moreover, common themes among the synaesthetic experiences of many people are discovered. For example, certain letters such as *"A is red"* (R), *"B is blue"* (R), and *"C is yellow"* (R), as well as *"0 and 1 are often perceived as black/white, since it is connected with the binary system"* (S) are reported in research and seen during the interviews. These



similarities underline the universal characteristics of synaesthetic perception (*"I myself can only confirm that these tendencies are generally the case"* (S)).

Discussing long texts reveals an interesting component of synaesthetic perception. It is found that the vividness of synaesthetic colours can fade when reading long texts, as the attention shifts to the story or content rather than individual graphemes or words - *"when they read a book, for example, it gradually fades"* (R). This observation bears similarities to the diminishing perception of unique typefaces with extended reading, indicating a potential adaptive function of perceptual attention in synaesthesia. In addition, the same phenomenon is mentioned with numbers, so if one sees a number with two dominant colours, the overall colour of the number fades somehow (an example of a violet 9 with a grass-green 6, so the number 96 is then perceived in a violet shade, even though the green 6 is perceived as more dominant, so one perceives *"some washed-out sensory impressions"* (S)).

The technological possibility of recreating synaesthetic experiences via digital interfaces using the stated rule-sets is addressed. Experts in the subject matter state that simple synaesthetic experiences (such as the appearances of the majority who experience grapheme-colour synaesthesia) might be easily recreated, *"for them, I believe the app works really well"* (S). *"You can't exactly put these thoughts into my head via the app, but when it is precisely combined, it is useful. So, when I see exactly this colour while reading the text, then it comes, I believe, relatively close"* (S).

Last but not least, some researchers (on synaesthesia and neuroscience) and their papers are mentioned in the interviews to check on their findings if they are in line with the findings discussed.

## Visualisation

This theme includes a broad variety of topics. Firstly, during the demonstration phase of the interviews, the outline method is praised for its non-intrusive nature, since *"nothing is actually changed in the text"* (S), while providing a subtle enhancement that improves legibility without overwhelming the reader. This method is described as offering a *"hint of supportive colouring"* (S), which serves to subtly highlight text without altering it, since *"I think it's important that you can still see the text somehow"* (S).

The smooth execution of *TMP overlay* visualisation is appreciated for adding to the reading experience's fluidity. Longer texts are thought to be more suited for this approach, since it improves readability without breaking the text's natural flow. Some disadvantages are noted, such the lack of punctuation and the restriction to using just lowercase characters, *"although, I don't find them so annoying now, because chat communication is now all lower case anyway"* (S). Another downside, which is also addressed, is the difficulty of reading, for example, yellow text on a white background. The suggestion is to colour the background with the maximum possible contrast of the font colour, which could on the other hand lead to a visually overwhelming display experience. Notwithstanding

these difficulties, the approach is praised for its capacity to clearly emphasise every word, which is not common, but is liked.

The idea of pre-coloured texts sparks thought-provoking conversations on how these characteristics affect perception and learning. *"That's where it gets really difficult, because emojis, for instance, themselves always bring colour with them"* (S). Research on the effects of pre-coloured components, such as logos or symbols, on cognitive processes is scarce, but early findings indicate that pre-coloured texts may help people develop colour-word connections, even those who are not synaesthetes [CMR12]. This could also lead to a possibly complementary, interplay between learnt and innate colour associations.

It is also mentioned in the interviews that non-synaesthetes can build colour-word or colour-letter connections by exposure to pre-coloured books.

### Challenges

Expert interviews show some general challenges in designing a grapheme-colour synaesthesia system. *"Significant individual variability among synaesthetes"* (R) is one of the main issues. It is critical to develop rule-sets that are wide and flexible enough to account for these variances. Furthermore, individual fine-tuning can be required to take into consideration differences in perception brought on by things like the time of day or personal circumstances. Ensuring that the system can accommodate both associator and projector synaesthetes is mentioned as a good way of providing different visualisation options.

Language variations add to the difficulty. While it is thought by the synaesthete that German and English synaesthetic sensations are comparable, there is considerable doubt concerning other languages, such as French, where *"for example, vowels are pronounced quite differently"* (S). These variations may have an influence on the synaesthetic experience.

Similar to the previous challenge, the semantic meaning of words in a text or sentence can impact the synaesthetic colour depiction, according to both the synaesthete and the synaesthesia researcher. Words having different meanings within the same context may be seen in different hues, necessitating many possibilities inside the system to accommodate this variety. For example, the synaesthete mentions that happy words may be seen differently from sad words, and words associated with the past may have different colours to those associated with the future. Another mentioned example is *"Monday, which could be the first day of the working week, might have a different colour than if you say I was born on a Monday or something, because the semantics are completely different"* (S).

Furthermore, the prototype is presently intended for the Latin alphabet, and it is uncertain how it will handle letters from other alphabets, such as those found in Asian languages.

If colour perception varies so much that users have to constantly adjust their settings, a significant difficulty arises. The synaesthete advises that in such instances, it could be

better to render the text in black on a white backdrop rather than risking inaccurate colour representation, which could lead to an even worse experience in reading.

In addition, the synaesthete emphasises the necessity of allowing users to fine-tune the colours for certain letters separately after going through the initial set-up phase. For example, if a user incorrectly sets the colour for the letter "u", they should be able to change it without resetting the entire alphabet. Screen settings and calibration can also influence colour appearance, necessitating specific modifications to assure proper representation - *"E is really a sunshine yellow, and this one is more of an orange now"* (S), *"the blue of the A, that is actually already a relatively light I would almost say sky blue, and that is now getting so dark"* (S). This is mentioned as being important not only when technical devices do not display colour correctly, but also to ensure that the system remains effective and accurate over the lifetime of the user.

### Interaction

The expert interviews reveal important details about the colour-space used. The RGB colour model is used in the prototype, but because it is non-linear, it is found to be less effective. The experts recommend using the HSV model instead, as it emphasises the hue more. The application's synaesthetic experience would probably be more accurate and relevant if the HSV model was used.

The issue mentioned above goes hand in hand with the next finding, the design of the colour picker. Experts note that brightness changes, which are frequently invisible to the human eye, take up a significant amount of the screen's space - *"the hue value, that was just this horizontal bar, I would have needed more area, more detail and less for the brightness"* (S). Furthermore, usage is made harder by the colour picker's keeping of the previous set-letter's colour, particularly when a very dark colour is selected. Experts recommend that improving visibility would include returning the picker to its basic, default, for instance, white colour after each selection.

It is observed that, rather than seeing one colour for each letter in a word, synaesthetes may see many colours for a single letter. Given this perception, it is possible that the neighbouring vowel on the left side of the grapheme exerts a greater influence on the colours of the left side than the neighbouring vowel on the right - *"left is 100% a right is 100% e"* (S). In this form of grapheme-colour synaesthesia, the colour of each pixel can vary, giving a more nuanced and complex experience than another.

Furthermore, it is discovered that no numerals are coloured in the prototype, just in the *"individual letter colouring mode"*, and that the vowel "y", which is considered a vowel in certain languages, is left out at this point.

The experts make positive comments about the prototype, such as *"very interesting how it works, how pleasant it is to read the text"* (S). They call it a prototype that is really well done, *"looks pretty cool"* (S) and *"is easy to navigate"* (S). Its usefulness as a proof-of-concept is acknowledged by professionals, who especially value *"its great potential"*

## 6. RESULTS

---

*as a research tool*" (R) to get to know more about the experiences on a personal level. If the prototype might make it possible for specialists to read research articles more rapidly in the future, then they could be more likely to utilise it on a regular basis.

Besides the positive remarks, there is some space for improvement noted. It is "*not completely responsive*" (S) on mobile devices and "*heated up a bit*" (S) when using it. Experts advise outsourcing OCR processing to a server in order to boost speed, which might increase the responsiveness and general smoothness of the application.

## Discussion of Results

This chapter contrasts and discusses the findings in relation to the hypotheses (see Chapter 6) and research questions (see Section 1.2), based on the results reported in the preceding chapter. This chapter also discusses the limitations and possible directions for further research.

The hypothesis **H1-a** proposes that the inducer-concurrent relationship of grapheme-colour synaesthesia can be formalised as rule-sets. This study's findings support this concept, which builds on earlier research and expert interviews.

There have been significant prior studies conducted to find patterns that can be used to formalise rule-sets for grapheme-colour synaesthesia (see Section 2.2.5). This has provided a framework for exploring different formalisation methods, but often encounters difficulties due to the differences in individual perceptions. It is possible to formalise distinct perspectives into rule-sets, according to the expert interviews done for this study. However, it should be noted that colour adjustments for individual graphemes may be needed in the future. As people age, their perception of colours may deteriorate, requiring occasional recalibration of the grapheme colours to maintain accuracy.

In addition, when including the semantic meaning of words in texts, a flexible approach to rule-sets is required in relation to "special words" (e.g. weekdays). The expert interviews emphasise the need of allowing users to customise rule settings for each word while reading. This adaptability is essential for dealing with special words that change colour depending on their context (see Section 6.2.1).

The synaesthesia researcher, provides extremely useful insights. They agree that inducer-concurrent relationships, particularly the "simple" ones, might be formalised into rule-sets. Simple relationships are here defined as those that do not change depending on the time of day, semantic meaning, or emotional state. This confirmation from someone with first-hand knowledge highlights the feasibility of formalising these linkages and validates the hypothesis.

## 7. DISCUSSION OF RESULTS

One noteworthy finding is that it is not always required for people to assign colours to each and every grapheme. Some synaesthetes exclusively see colours, for instance in vowels, whereas consonants do not have inherent colours but are influenced by neighbouring vowel colours. This selective perception emphasises the importance of an option for flexible colour definition that can accommodate such individual differences. Thus, the capacity to re-adjust and specify colours for certain graphemes is critical for effectively representing the synaesthetic experience.

Based on these findings, an extensive diagram is created to illustrate the various ways in which synaesthetes perceive and associate colours with graphemes (see Figure 7.1). This diagram also serves to answer **RQ1**.

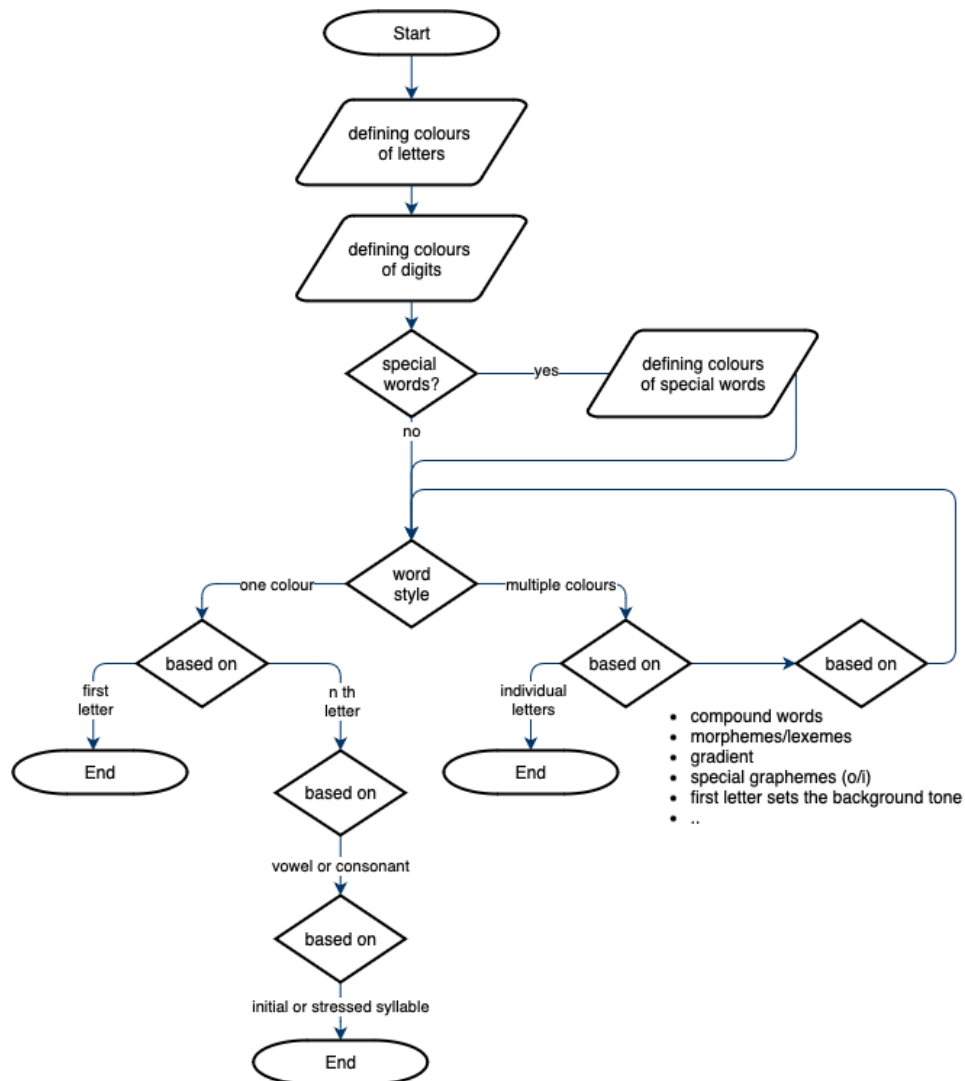


Figure 7.1: Diagram illustrating the formalisation of grapheme-colour experiences into rule-sets.

---

Several grapheme-colour synaesthetes experience words in a single colour, according to the expert interview with the synaesthesia researcher. This interpretation is frequently based on the first or emphasised syllable letter/vowel, as illustrated on the left side of the picture. This is consistent with prior research findings presented in Section 2.2.5, which suggests that the first letter or stressed vowel has a substantial influence on the perceived colour of the word.

The hypothesis **H2-a** anticipates that the application will execute at a rate exceeding 30 FPS and a RR of less than 100 milliseconds (ms). The results of the benchmark analysis are 3.912 FPS and 3.585 FPS for the two visualisation options when the average FTs are converted back to FPS (using the reformulated formula in Equation (6.1)). Unfortunately, this does not reach the estimated performance requirement. Also, when having a look at the response times, both options *texture direct* and *TMP overlay*, do not meet the below 100 ms requirement with the benchmark results of 137.239 ms and 132.743 ms respectively. It can be concluded that the data does not support H2-a, therefore the null hypothesis is not rejected. User comments support this conclusion, suggesting that the app *"stalled when too much text was scanned at once"*.

Throughout the development phase, usability was a top priority. The key goal was to ensure that the app was both functional and user-friendly. The usability is of critical importance, as a confusing or difficult-to-use interface may dissuade users, thereby reducing overall effectiveness and adoption rates. Therefore, the hypothesis **H2-b** states that a SUS score greater than 80 is obtained, which is associated with an excellent usability. The results support this hypothesis, with the average SUS score of 88.75. This high score demonstrates the effectiveness of design decisions that prioritise user experience.

According to the hypothesis **H2-c**, it is possible to successfully represent grapheme-colour synaesthesia in AR via text detection and recognition of printed text. The expert interviews and try-out session support this hypothesis.

Benchmark testing demonstrates that good text detection and recognition alone are insufficient for successfully representing grapheme-colour synaesthesia in AR. The visualisation algorithm is equally important. In particular, the *texture direct* mode has a very high ER, suggesting that the encoding of grapheme-colour connections is error-prone in the absence of a robust visualisation algorithm.

One of the most significant concerns observed is the difficulty in accurately recognising contours for letters that occur close together, such as double-f or double-t. Section 4.4.3 describes the algorithm for gathering contours and matching them to the number of letters in the recognised word. The algorithm, however, finds it difficult to distinguish accurately between letters that are quite near to one another. To solve this, the algorithm may be able to recognise individual letters more correctly by increasing the distance between characters, which would lower the ER.

Punctuation marks are another major hurdle for the OCR system. Words followed by sentence marks (such as commas, periods, and semicolons) frequently produce an

additional contour, leading the OCR algorithm to misidentify the word. Furthermore, punctuation symbols like commas and hyphens are occasionally misidentified as letters.

Regarding the visualisation options, it is found that most participants do not like the background colouring options, either in *texture direct* or in the *TMP overlay* mode. Eight participants mentioned that the black graphemes of words are hardly at all readable if the grapheme colours is set to a dark colour.

Thus, based on all of the above findings, hypothesis H2-c is supported in theory, but further technological developments are needed to fully realise its potential in practice.

In regard to **RQ2** it is therefore necessary to take the findings from RQ1 and then primarily focus on the OCR and visualisation algorithms during the implementation. This is not only to ensure accurate detection, recognition and visualisation of the text, but also to guarantee optimal performance and seamless operation of the entire application.

### 7.1 Limitations

Some limitations of this work give rise to further research, which is discussed in this section.

#### 7.1.1 Colouring

The system is limited to the word colouring rule "individual" in order to realise the colouring of the digits. This means that colouring schemes, such as those based on the first or most dominating digit in a series, cannot be used. The digits in all other colouring settings are black and lack any colour.

#### 7.1.2 Rule-Sets

Furthermore, the rule-set is restricted and does not account for individual words or contextual changes. Also, the visualisation possibilities available in the system are currently restricted to basic letter-by-letter colouring of the font, outline, and background; no letter-level variants (i.e., various colours for a single letter) are offered.

#### 7.1.3 Recognition and Detection

Because of the chosen OCR algorithm, the prototype only supports texts in English at this time, which limits its applicability in multilingual settings or for users who engage with texts primarily in other languages.

As noted in Section 6.2.1, vowel recognition does not extend to the letter "y". This decision is consistent with the most frequent usage in English, but it may not fully reflect its vocalic role in many words and settings, thereby compromising the aesthetic and functional effects of the "vowel gradient colouring" mode.



Punctuation and capitalization are not supported throughout the prototype due to the chosen OCR algorithm. This affects the *TMP overlay* visualisation modes alone, where the recoloured words are superimposed over the words, rather than the *texture direct* visualisation options.

## 7.2 Potential Future Work

The system's development and assessment revealed some promise for further study and advancement. Addressing present constraints and creating more features may expand the possibilities for utilising this prototype as a research tool.

### 7.2.1 Colouring

Future developments may include redesigning the colour picker interface to use an HSV (hue, saturation, value) model, as suggested by a synaesthete, or even the CIE  $L^*a^*b^*$  or CIE  $L^*u^*v^*$  colour models (as mentioned in Section 2.3.1) to ensure maximum sensitivity to allow more accurate colour selection. This update could solve difficulties with the existing picker, which wastes a lot of screen area when making brightness adjustments, which is not as relevant as colour hues.

Furthermore, as proposed during the try-out session, after selecting colours for each grapheme, resetting the colour picker to white to ensure their visibility, would be a useful feature to add next.

Third, additional implementations might enable users to modify grapheme colours after the first configuration. The need for this capability stems from the fact that colours appear to be different on the screen, which is often not apparent to the user until they have scanned a word for the first time.

Another interesting area for improvement is dynamic colour adjustment. Modifications to the background and text colours in response to brightness and colour contrast could improve readability. This approach is exemplified by SYNCalc [syn24], which employs coloured digits on a high-contrast background. This may include automatically altering text to a brighter hue against darker backgrounds and vice versa, although thorough testing would be required to regulate the visual impact.

### 7.2.2 Rule-Sets

Using NLP techniques, more advanced rule-sets based on syllabic or morphemic structures might be implemented, allowing for more subtle colouring schemes. These systems can involve colouring words according to the first vowel of the first or stressed syllable.

### 7.2.3 Recognition and Detection

Extending the system to handle other languages would increase its worldwide applicability, making it more valuable to a wider group of users.

Due to present restrictions to the limited recognition zone, enabling text recognition across wider areas or even the full screen would be a potential future implementation. Addressing this may solve the triggering problem for some people where a letter is not fully in the recognition box and is therefore recognised as a different letter, e.g. "o" instead of "b", and appears in the wrong colour until it is fully in the box.

### 7.2.4 Performance

A detailed examination of the pipeline reveals that the FT and RR are not optimal, which can be attributed to the used OCR algorithm. Consequently, subsequent implementations may consider integrating a mobile-specific OCR algorithm, such as Google OCR, to enhance frame rate, particularly on resource-constrained devices, such as mobile devices. Moreover, the outsourcing of computationally demanding activities to a server or the integration of text tracking with the detection and recognition of text are potential avenues for consideration. This approach can also be beneficial if the system is implemented for the use of HMD's in subsequent stages.

### 7.2.5 Further Research and Development

Conducting in-depth studies, maybe in partnership with a psychological research department, could give useful insights into how colouring graphemes influences reading ability. Eye-tracking technology could offer information about reading speed, recognition rate, and understanding.

Expanding the system's functionality to incorporate HMD devices, such AR headsets, would allow for new applications in immersive learning and reading settings.

# Conclusion

This thesis expands on the idea of formalising the regulatory factors of the inducer-concurrent relationship of grapheme-colour synaesthesia, combining data from previous research with expert interviews, and assessing it directly with a grapheme-colour synaesthete.

As a proof-of-concept that these sensations may be represented using technology, notably AR, a prototype is created to obtain greater insights into individual experiences of grapheme-colour synaesthesia.

The formalisation of the various individual inducer-concurrent perceptions of grapheme-colour synaesthesia into machine-readable rule-sets is demonstrated to be a feasible undertaking. A diagram is created to illustrate the manner in which this can be achieved.

Additionally, this thesis indicates that it is possible to represent grapheme-colour synaesthesia through the use of AR. A grapheme-colour synaesthete corroborates this conclusion.

Furthermore, the findings indicate the potential for achieving this in AR with an excellent user experience (SUS of 88.75).

The technical criteria for improving the prototype are defined, ensuring that subsequent iterations are more efficient.

In terms of readability, the user study posits that it is preferable to apply colour directly to the pixels delineating the graphemes, as opposed to, for instance, colouring the background of the graphemes. This is because, in certain instances, the application of colour to the entire background of a grapheme, particularly when the colour is dark and the grapheme itself is black, can impede reading.

Considering the importance of reading in daily life, this thesis provides a solid foundation for future studies pertaining to natural synaesthetic reading.

## 8. CONCLUSION

---

On top of that, this thesis enables non-synaesthetes to benefit from the potential advantages of coloured texts while mitigating the disadvantages of achromatic text reading for synaesthetes. It enables synaesthetes to share their colour experiences and perceptions, and also provides an application that can be further improved and altered for usage with HMDs, allowing for real-time colour representation of texts in any context.

Therefore, this project lays a solid foundation for further research, serving as a "*useful research-tool*" (quote synaesthesia researcher).

# List of Figures

1.1	Screenshot of the prototype showing the text recolouring feature, which allows users to personalise their reading experience. . . . .	3
2.1	Seventy-five types of synaesthesia (Sean A. Day). Left column: inducers; top row: concurrents. White: documented; red: unrecorded; black: not a type. (Reprinted from: [Day22]) . . . . .	6
2.2	Illustration of semantic network activations in response to the letter "A": synaesthete with red colour experience (left) vs. non-synaesthetic control (right). (Reprinted from: [Mei13]) . . . . .	9
2.3	Differences in perception of letters, numbers, or words by projectors and associators. Top: two projectors; bottom: three associators. (Reprinted from: [The21] and [SLM09]) . . . . .	11
2.4	Representative synaesthete (A–C) and control (D–F) brains. Grapheme ROI (light blue) and V4 ROI (dark blue) are shown. A and D: ROIs on non-inflated cortical surfaces. B and E: ROIs on inflated brains; yellow box highlights region in C and F. Synaesthetes showed activation in both grapheme (light blue) and V4 (dark blue) ROIs when viewing achromatic letters and numbers (C). Controls showed activation only in the grapheme ROI (light blue) (F). (Reprinted from: [BHC <sup>+</sup> 10]) . . . . .	13
2.5	Screenshot of the publicly available shared GoogleSheets file [Goond] for colouring cells in the experienced colour. (Reprinted from: [The21]) . . . . .	15
2.6	Appearance of words starting with "i" or "o". Left: experienced word; right: individual letter colours. (Source: [BCG16]) . . . . .	16
2.7	Screenshot of the synaesthesia battery's example grapheme-colour picker test. (Source: [Dav24a]) . . . . .	16
2.8	Colour picker that allows the selection of a colour range and brightness adjustment. (Reprinted from: [ANS15]) . . . . .	17
2.9	Colour picker that allows to choose from 13 colours or "no colour". (Reprinted from: [MRW14]) . . . . .	18
2.10	Screenshot of the SYNCalc application by Berger and Whittingham. (Reprinted from: [The21]) . . . . .	18
2.11	Screenshot of the "SeeSynesthete" Google Chrome extension for colouring web page fonts. (Reprinted from: [Chr20]) . . . . .	19
		75

2.12	Screenshot of the "Synesthesia" Google Chrome extension for colouring web page fonts. (Reprinted from: [Chr19]) . . . . .	20
2.13	Three visualisation versions. Left: original document; center: document with overlay; right: new OCR layer. (Reprinted from: [Scrnd]) . . . . .	22
2.14	Two different AR translator applications using overlay and replacement as visualisation techniques. (Reprinted from: [Goo24] and [Goo23a]) . . . . .	23
3.1	System design sketch illustrating the scanning of real-world text and recolouring via mobile AR. . . . .	29
3.2	Single-colour word colouring rules selected based on literature review and relevance to this thesis. . . . .	29
3.3	Multi-colour word colouring rules selected based on literature review and relevance to this thesis. . . . .	30
3.4	The two different visualisation types. . . . .	31
3.5	The three different colouring types. . . . .	31
3.6	Details of a screenshot of the word-colouring screen. It displays the term "Sample Text" for each rule-set as a preview in the following order: individual, first letter, first vowel, vowel gradient. . . . .	32
3.7	Various options for displaying recoloured text, organised by visualisation styles (columns) and colouring styles (rows). . . . .	33
4.1	User flow diagram of the application, showing each step from starting the app to closing the app. . . . .	37
4.2	Detail of a screenshot of the how-to screen that is displayed prior to the scanning process. . . . .	37
4.3	Detail of a screenshot of the visualisation mode selection screen, which shows that the <i>direct + font</i> option is currently selected. . . . .	38
4.4	Screenshots of all application screens, labelled by name. User flow is indicated by brown arrows. . . . .	39
4.5	Detail of a screenshot of the introductory screen, which requires users to enter their first and last names in order to save their data. . . . .	40
4.6	Detail of a screenshot during the text scanning process in the <i>texture direct</i> mode. . . . .	42
4.7	Detail of a screenshot during the text scanning process in the <i>TMP overlay</i> mode. . . . .	43
5.1	Checkbox added to the menu screen for selecting whether to log FPS or RR during benchmark testing. . . . .	46
5.2	Photograph depicting the process of selecting a colour for each grapheme, with the current iteration featuring a shade of blue for the letter "g". . . .	47
5.3	Photograph depicting a participant engaged in the text scanning process, utilising the <i>texture direct + background</i> visualisation option with colouring based on the first letter colour rule. . . . .	48
	76	

5.4	Interview guide for the semi-structured interview with the synaesthesia researcher. The yellow background indicates that the questions are identical to those posed in the second expert interview. . . . .	51
5.5	Interview guide for the semi-structured interview with the synaesthete. The yellow background indicates that the questions are identical to those posed in the second expert interview. . . . .	52
6.1	Figures illustrating details of the FT benchmark. . . . .	55
6.2	Figures illustrating details of the RR benchmark. . . . .	56
6.3	Distribution of the SUS scores of the 12 participants for SynVis. . . . .	58
6.4	Readability scores (on a scale from 1 to 5) where TD_F is <i>texture direct + font</i> , TD_O is <i>texture direct + outline</i> , TD_B is <i>texture direct + background</i> , TMP_F is <i>TMP overlay + font</i> and TMP_B is <i>TMP overlay + background</i> . Lower scores indicate better readability. . . . .	59
6.5	Screenshot of the Miro board in progress during the iterative thematic data analysis process. . . . .	61
6.6	Screenshot of the Miro board displaying the identified themes (Rule-Set, Visualisation, Challenges, and Interaction) along with their grouped subtopics. . . . .	62
7.1	Diagram illustrating the formalisation of grapheme-colour experiences into rule-sets. . . . .	68



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# List of Tables

2.1	Comparison of relevant related work approaches for text detection and recognition (Table adapted from: [OBHW22b]) . . . . .	21
6.1	Descriptive statistics of benchmark values, all of which are reported in milliseconds (ms) . . . . .	54
6.2	Normality tests of benchmark values (NaN means: all values are identical)	54
6.3	Wilcoxon signed-rank test - FT and RR . . . . .	55
6.4	Descriptive statistics of readability scores, all of which are reported in school grades ranging from 1 (very good) to 5 (very poor). . . . .	59
6.5	Normality tests of readability scores . . . . .	60
6.6	Conover's post hoc comparisons - visualisation option . . . . .	60



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Acronyms

- AR** Augmented Reality. xi, xiii, 1–3, 20–23, 27–29, 35–37, 45, 53, 54, 57, 69, 72, 73, 76
- CRNN** Convolutional Recurrent Neural Network. 36, 41
- DL** Deep Learning. 20, 21
- EAST** Efficient and Accurate Scene Text Detection. 21, 36
- ER** Error Rate. xiii, 45, 46, 53, 54, 56, 69
- FPS** Frames Per Second. 45, 46, 54, 69, 76
- FT** Frame Time. xiii, 45, 53–55, 69, 72, 77, 79
- HMD** Head Mounted Display. 21, 72, 74
- NLP** Natural Language Processing. 29, 71
- OCR** Optical Character Recognition. 20–23, 36, 45, 66, 69–72, 76
- RR** Response Rate. xiii, 45, 46, 53–56, 69, 72, 76, 77, 79
- SUS** System Usability Scale. 1, 48, 57, 58, 69, 73, 77
- TMP** Text Mesh Pro. xvi, 33, 41, 43–46, 54–56, 59, 60, 63, 69–71, 76, 77
- UI** User Interface. 4, 28, 35–38, 40
- VR** Virtual Reality. 8, 17, 20



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Bibliography

- [ANS15] Árni Gunnar Ásgeirsson, Maria Nordfang, and Thomas Alrik S. Components of attention in grapheme-color synesthesia: A modeling approach. *PLOS ONE*, 10(8):1–19, 08 2015.
- [App21] Apple App Store - Omer Faruk Ozturk. Searchcam - ctrl-f camera app, 2021. Accessed: August 5, 2024.
- [BC06] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3:77–101, 01 2006.
- [BCG16] Laura J. Blazej and Ariel M. Cohen-Goldberg. Multicolored words: Uncovering the relationship between reading mechanisms and synesthesia. *Cortex*, 75:160–179, 2016.
- [BHC<sup>+</sup>10] David Brang, Edward Hubbard, Seana Coulson, Ming-Xiong Huang, and Vilayanur Ramachandran. Magnetoencephalography reveals early activation of v4 in grapheme-color synesthesia. *NeuroImage*, 53:268–74, 10 2010.
- [BHW<sup>+</sup>19] Joshua Berger, Irina Harris, Karen Whittingham, Zoe Terpening, and John Watson. Substantiating synesthesia: a novel aid in a case of grapheme-colour synesthesia and concomitant dyscalculia. *Neurocase*, 26:1–7, 11 2019.
- [BHW<sup>+</sup>21] Joshua Berger, Irina Harris, Karen Whittingham, Zoe Terpening, and John Watson. Sharing the load: How a personally coloured calculator for grapheme-colour synaesthetes can reduce processing costs. *PLOS ONE*, 16:e0257713, 09 2021.
- [BK69] Brent Berlin and Paul Kay. *Basic color terms: Their Universality and Evolution*. Berkeley, CA: University of California Press, 1969.
- [BMB<sup>+</sup>23] Lucie Bouvet, Cynthia Magnen, Clara Bled, Julien Tardieu, and Nathalie Ehrlé. “i have to translate the colors”: Description and implications of a genuine case of phoneme color synaesthesia. *Consciousness and Cognition*, 111:103509, 2023.

- [Brind] British Council. How humans evolved language, n.d. Accessed: August 5, 2024.
- [Bro95] John Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.
- [CCC<sup>+</sup>14] Guang Chen, Wei Cheng, Tingwen Chang, Xiaoxia Zheng, and Ronghuai Huang. A comparison of reading comprehension across paper, computer screens, and tablets: Does tablet familiarity matter? *Journal of Computers in Education*, 1:213–225, 11 2014.
- [CCLG20a] Alyson Collins, Donald Compton, Esther Lindström, and Jennifer Gilbert. Performance variations across reading comprehension assessments: Examining the unique contributions of text, activity, and reader. *Reading and Writing*, 33, 03 2020.
- [CCLG20b] Alyson Collins, Donald Compton, Esther Lindström, and Jennifer Gilbert. Performance variations across reading comprehension assessments: Examining the unique contributions of text, activity, and reader. *Reading and Writing*, 33, 03 2020.
- [CDS<sup>+</sup>15] D.A. Carmichael, M.P. Down, R.C. Shillcock, D.M. Eagleman, and J. Simner. Validating a standardised test battery for synesthesia: Does the synesthesia battery reliably detect synesthesia? *Consciousness and Cognition*, 33:375–385, 2015.
- [Chr19] Chrome Web Store - mr.bearengineer. Synesthesia chrome extension, 2019. Accessed: August 5, 2024.
- [Chr20] Chrome Web Store - emmawebdeveloper00. Seesynesthete chrome extension, 2020. Accessed: August 5, 2024.
- [CL19] Virginia Clinton-Lisell. Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, 42:288–324, 05 2019.
- [CMR12] Olympia Colizoli, Jaap M. J. Murre, and Romke Rouw. Pseudo-synesthesia through reading books with colored letters. *PLOS ONE*, 7(6):1–10, 06 2012.
- [CMSR17] Olympia Colizoli, Jaap Murre, H. Scholte, and Romke Rouw. Creating colored letters: Familial markers of grapheme–color synesthesia in parietal lobe activation and structure. *Journal of Cognitive Neuroscience*, 29:1–14, 02 2017.
- [Coh88] J. Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, 1988.

- [CR14] Rocco Chiou and Anina Rich. The role of conceptual knowledge in understanding synaesthesia: Evaluating contemporary findings from a “hub-and-spokes” perspective. *Frontiers in Psychology*, 5, 2014.
- [CT19] Henda Chorfi and Lama Tatwany. Augmented reality based mobile application for real-time arabic language translation. *Communications in Science and Technology*, 4:30–37, 07 2019.
- [CVRRS23] Abbineni Charishma, Alla Amrutha Vaishnavi, D Rajeswara Rao, and Tirumalasetti Teja Sri. Smart reader for visually impaired. In *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 1, pages 349–352, 2023.
- [CWT<sup>+</sup>15] Michael Cohen, Kathrin Weidacker, Judith Tankink, H. Scholte, and Romke Rouw. Grapheme-color synesthesia subtypes: Stable individual differences reflected in posterior alpha-band oscillations. *Cognitive neuroscience*, 6:1–12, 04 2015.
- [Cyt89] Richard Cytowic. Cytowic, r. e. synesthesia and mapping of subjective sensory dimensions. *neurology* 39, 849-850. *Neurology*, 39:849–50, 07 1989.
- [Cyt95] Richard Cytowic. Synesthesia: Phenomenology and neuropsychology a review of current knowledge. *Psyche*, 2, 01 1995.
- [Dav24a] David Eagleman. Grapheme-colour picker test, 2005-2024. Accessed: August 5, 2024.
- [Dav24b] David Eagleman. The synesthesia battery, 2005-2024. Accessed: August 5, 2024.
- [Day04] Sean A. Day. Trends in synesthetically colored graphemes and phonemes – 2004 revision. 2004.
- [Day13] Sean A. Day. 903Synesthesia: A First-Person Perspective. In *Oxford Handbook of Synesthesia*. Oxford University Press, 12 2013.
- [Day22] Sean A. Day. Types of synesthesia, 2022. Accessed: August 5, 2024.
- [DS05] Mike J. Dixon and Daniel Smilek. The importance of individual differences in grapheme-color synesthesia. *Neuron*, 45(6):821–823, 2005.
- [Duf01] Patricia Lynne Duffy. *Blue cats and chartreuse kittens*. Macmillan, 11 2001.
- [EKT<sup>+</sup>07] David Eagleman, Arielle Kagan, Steffie Tomson, Deepak Sagaram, and Anand Sarma. A standardized test battery for the study of synesthesia. *Journal of neuroscience methods*, 159:139–45, 02 2007.

- [Ell15] Andrew J. Elliot. Color and psychological functioning: a review of theoretical and empirical work. *Frontiers in Psychology*, 6, 2015.
- [Fle06] Jack Fletcher. Measuring reading comprehension. *Scientific Studies of Reading - SCI STUD READ*, 10, 07 2006.
- [FSA<sup>+</sup>06] David Francis, Catherine Snow, Diane August, Coleen Carlson, Jon Miller, and Aquiles Iglesias. Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *scientific studies of reading*, 10 (3), 301-322. *Scientific Studies of Reading*, 10:301–322, 01 2006.
- [GENB21] Anna Carin Gran Ekstrand, Mattias Nilsson Benfatto, and Gustaf Öqvist Seimyr. Screening for reading difficulties: Comparing eye tracking outcomes to neuropsychological assessments. *Frontiers in Education*, 6, 2021.
- [Git24] GitHub. Tesseract ocr, 2024. Accessed: August 5, 2024.
- [GNCHCG11] Veronica Gross, Sandy Nearing, Catherine Caldwell-Harris, and Alice Cronin-Golomb. Superior encoding enhances recall in color-graphemic synesthesia. *Perception*, 40:196–208, 02 2011.
- [Goo23a] Google Play Store - Google LLC. Google lens, 2023. Accessed: August 5, 2024.
- [Goo23b] Google Play Store - StuckInBasement. Ctrl-f - search text in documents, 2023. Accessed: August 5, 2024.
- [Goo24] Google Play Store - Dream Dijital. Translate lens: Photo & camera, 2024. Accessed: August 5, 2024.
- [Goond] Google Sheets. Our alphabets- add your colors on a new row at the bottom, use "custom" under the "fill color" tool to find more colors, n.d. Accessed: August 5, 2024.
- [GRJ12] Bradley Gibson, Gabriel Radvansky, and Ann Johnson. Grapheme–color synesthesia can enhance immediate memory without disrupting the encoding of relational cues. *Psychonomic bulletin review*, 19, 07 2012.
- [GTHB16] Philip Griffiths, Robert Taylor, Lisa Henderson, and Brendan Barrett. The effect of coloured overlays and lenses on reading: a systematic review of the literature. *Ophthalmic and Physiological Optics*, 36:519–544, 09 2016.
- [HARB05] Edward Hubbard, Andi Arman, Vilayanur Ramachandran, and Geoffrey Boynton. Individual differences among grapheme-color synesthetes: Brain-behavior correlations. *Neuron*, 45:975–85, 04 2005.



- [HE20] Vered Halamish and Elisya Elbaz. Children’s reading comprehension and metacomprehension on screen versus on paper. *Computers Education*, 145:103737, 2020.
- [HYS20] Daisuke Hamada, Hiroki Yamamoto, and Jun Saiki. Association between synesthetic colors and sensitivity to physical colors changed by type of synesthetic experience in grapheme-color synesthesia. *Consciousness and Cognition*, 83:102973, 2020.
- [JAS18] JASP. Jasp - a fresh way to do statistics, 2018. Accessed: August 5, 2024.
- [JDW09] Jörg Jewanski, Sean Day, and Jamie Ward. A colorful albino: The first documented case of synaesthesia, by georg tobias ludwig sachs in 1812. *Journal of the history of the neurosciences*, 18:293–303, 07 2009.
- [JWA05] Julia Simner Jamie Ward and Vivian Auyeung. A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cognitive Neuropsychology*, 22(1):28–41, 2005. PMID: 21038239.
- [KHL22] Jieun Kim, Jae-In Hwang, and Jieun Lee. Vr color picker: Three-dimensional color selection interfaces. *IEEE Access*, 10:65809–65824, 2022.
- [KSZ18] Yiren Kong, Young Sik Seo, and Ling Zhai. Comparison of reading performance on screen and on paper: A meta-analysis. *Computers Education*, 123:138–149, 2018.
- [LFJH19a] Jana Lüdtkke, Eva Froehlich, Arthur M. Jacobs, and Florian Hutzler. The sls-berlin: Validation of a german computer-based screening test to measure reading proficiency in early and late adulthood. *Frontiers in Psychology*, 10, 2019.
- [LFJH19b] Jana Lüdtkke, Eva Froehlich, Arthur M. Jacobs, and Florian Hutzler. The sls-berlin: Validation of a german computer-based screening test to measure reading proficiency in early and late adulthood. *Frontiers in Psychology*, 10, 2019.
- [LLFT22] David P. Luke, Laura Lungu, Ross Friday, and Devin B. Terhune. The chemical induction of synaesthesia. *Human Psychopharmacology: Clinical and Experimental*, 37(4):e2832, 2022.
- [LM18] Katrin Lunke and Beat Meier. Creativity and involvement in art in different types of synaesthesia. *British Journal of Psychology*, 110, 11 2018.
- [LM20] Katrin Lunke and Beat Meier. A persistent memory advantage is specific to grapheme-colour synaesthesia. *Scientific Reports*, 10, 02 2020.

- [LUM15] James Lewis, Brian Utesch, and Deborah Maher. Measuring perceived usability: The sus, umux-lite, and altusability. *International Journal of Human-Computer Interaction*, 31:150625095336004, 06 2015.
- [MC14] Dan McCarthy and Gideon Caplovitz. Color synesthesia improves color but impairs motion perception. *Trends in cognitive sciences*, 18, 02 2014.
- [Mei13] Beat Meier. Semantic representation of synaesthesia. *Theoria et Historia Scientiarum*, 10, 12 2013.
- [Mei22] Beat Meier. Synesthesia. In Sergio Della Sala, editor, *Encyclopedia of Behavioral Neuroscience, 2nd edition (Second Edition)*, pages 561–569. Elsevier, Oxford, second edition edition, 2022.
- [Mic24a] Microsoft Developer. Windows.media.ocr api, 2024. Accessed: August 5, 2024.
- [Mic24b] Microsoft Support. Transcribe your recordings, 2024. Accessed: August 5, 2024.
- [Mir24] Miro. Miro - online collaboration tool, 2024. Accessed: August 5, 2024.
- [MR13a] Beat Meier and Nicolas Rothen. Grapheme-color synaesthesia is associated with a distinct cognitive style. *Frontiers in psychology*, 4:632, 09 2013.
- [MR13b] Myrto Mylopoulos and Tony Ro. Synesthesia: a colorful word with a touching sound? *Frontiers in Psychology*, 4, 2013.
- [MRW14] Beat Meier, Nicolas Rothen, and Stefan Walter. Developmental aspects of synaesthesia across the adult lifespan. *Frontiers in human neuroscience*, 8:129, 03 2014.
- [MS22] Thea Mannix and Thomas Sørensen. Face-processing differences present in grapheme-color synesthetes. *Cognitive Science*, 46, 04 2022.
- [Mun24] Munsell Color. Munsell book of color - matte edition, 2024. Accessed: August 5, 2024.
- [MW03] H. Mayringer and Heinz Wimmer. Salzburger lese-screening (sls) für die klassenstufen. *Göttingen: Hogrefe*, pages 1–4, 01 2003.
- [NCE11] Scott Novich, Sherry Cheng, and David Eagleman. Is synaesthesia one condition or many? a large-scale analysis reveals subgroups. *Journal of neuropsychology*, 5:353–71, 09 2011.
- [Nik20] Ivan Nikolov. In *Proceedings of the 23rd International Conference on Academic Mindtrek*, AcademicMindtrek '20, page 153–156, New York, NY, USA, 2020. Association for Computing Machinery.

- [OBHW22a] Imene OUALI, Mohamed Ben Halima, and Ali WALI. Real-time application for recognition and visualization of arabic words with vowels based dl and ar. In *2022 International Wireless Communications and Mobile Computing (IWCMC)*, pages 678–683, 2022.
- [OBHW22b] Imene Ouali, Mohamed Ben Halima, and Ali Wali. *Text Detection and Recognition Using Augmented Reality and Deep Learning*, pages 13–23. 03 2022.
- [OECnd] OECD. Reading performance (pisa), n.d. Accessed: August 5, 2024.
- [OHHW20] Imene OUALI, Mohamed Saifeddine HADJ SASSI, Mohamed BEN HALIMA, and Ali WALI. A new architecture based ar for detection and recognition of objects and text to enhance navigation of visually impaired people. *Procedia Computer Science*, 176:602–611, 2020. Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 24th International Conference KES2020.
- [OHSBHW21] Imene Ouali, Mohamed Saifeddine Hadj Sassi, Mohamed Ben Halima, and Ali Wali. Architecture for real-time visualizing arabic words with diacritics using augmented reality for visually impaired people. In Leonard Barolli, Isaac Woungang, and Tomoya Enokido, editors, *Advanced Information Networking and Applications*, pages 285–296, Cham, 2021. Springer International Publishing.
- [OHW22] Imene OUALI, Mohamed BEN HALIMA, and Ali WALI. Augmented reality for scene text recognition, visualization and reading to assist visually impaired people. *Procedia Computer Science*, 207:158–167, 2022. Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 26th International Conference KES2022.
- [Ope24] OpenCV. Opencv library, 2024. Accessed: August 5, 2024.
- [Par19] Parlindungan Pardede. Print vs digital reading comprehension in efl. 5:77, 07 2019.
- [PBM<sup>+</sup>02] Thomas Palmeri, Randolph Blake, Rene Marois, Marci Flanery, and William Whetsell. The perceptual reality of synesthetic color. *Proceedings of the National Academy of Sciences of the United States of America*, 99:4127–31, 04 2002.
- [PFC22] Ilya Pivavaruk and Jorge Ramón Fonseca Cacho. Ocr enhanced augmented reality indoor navigation. In *2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pages 186–192, 2022.
- [PMI17] Muhammad Pu, Nazatul Majid, and Bahari Idrus. Framework based on mobile augmented reality for translating food menu in thai language to

malay language. *International Journal on Advanced Science, Engineering and Information Technology*, 7:153, 02 2017.

- [PPR05] N. Plouznikoff, A. Plouznikoff, and J.-M. Robert. Artificial grapheme-color synesthesia for wearable task support. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pages 108–111, 2005.
- [PVdSN11] Chris Paffen, Maarten Van der Smagt, and Tanja Nijboer. Colour–grapheme synesthesia affects binocular vision. *Frontiers in Psychology*, 2, 2011.
- [RA18] John H. Reif and Wade Alhalabi. Advancing attention control using vr-induced multimodal artificial synesthesia. *Preprints.org*, August 2018.
- [RAM<sup>+</sup>21] Nicholas Root, Michiko Asano, Helena Melero, Chai-Youn Kim, Anton V. Sidoroff-Dorso, Argiro Vatakis, Kazuhiko Yokosawa, Vilayanur Ramachandran, and Romke Rouw. Do the colors of your letters depend on your language? language-dependent and universal influences on grapheme-color synesthesia in seven languages. *Consciousness and Cognition*, 95:103192, 2021.
- [RBM05] A.N. Rich, J.L. Bradshaw, and J.B. Mattingley. A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition*, 98(1):53–84, 2005.
- [RG19] Mariagrazia Ranzini and Luisa Girelli. Colours + numbers differs from colours of numbers: cognitive and visual illusions in grapheme-colour synaesthesia. *Attention, Perception, Psychophysics*, 81, 03 2019.
- [RH01] Vilayanur Ramachandran and Edward Hubbard. Psychophysical investigation into the neural basis of synaesthesia. *Proceedings. Biological sciences / The Royal Society*, 268:979–83, 06 2001.
- [Roy05] James Royer. Uses for the sentence verification technique for measuring language comprehension. 01 2005.
- [RR21] Nicholas Root and Romke Rouw. A unifying model of grapheme-color associations in synesthetes and controls. *Annual Meeting of the Cognitive Science Society*, 43, 2021.
- [RRA<sup>+</sup>18] Nicholas B. Root, Romke Rouw, Michiko Asano, Chai-Youn Kim, Helena Melero, Kazuhiko Yokosawa, and Vilayanur S. Ramachandran. Why is the synesthete’s “a” red? using a five-language dataset to disentangle the effects of shape, sound, semantics, and ordinality on inducer–concurrent relationships in grapheme-color synesthesia. *Cortex*, 99:375–389, 2018.

- [RSWW13] Nicolas Rothen, Anil K. Seth, Christoph Witzel, and Jamie Ward. Diagnosing synaesthesia with online colour pickers: Maximising sensitivity and specificity. *Journal of Neuroscience Methods*, 215(1):156–160, 2013.
- [SĚ1] Andreas Säuberli. Measuring text comprehension for people with reading difficulties using a mobile application. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS '21*, New York, NY, USA, 2021. Association for Computing Machinery.
- [SB13] Julia Simner and Angela Bain. A longitudinal study of grapheme-color synesthesia in childhood: 6/7 years to 10/11 years. *Frontiers in Human Neuroscience*, 7, 2013.
- [SB17] Julia Simner and Angela Bain. Do children with grapheme-colour synaesthesia show cognitive benefits? *British journal of psychology (London, England : 1953)*, 109, 03 2017.
- [SBRL<sup>+</sup>23] Jennifer J. Stiegler-Balfour, Zoe S. Roberts, Abby S. LaChance, Aubrey M. Sahouria, and Emily D. Newborough. Is reading under print and digital conditions really equivalent? differences in reading and recall of expository text for higher and lower ability comprehenders. *International Journal of Human-Computer Studies*, 176:103036, 2023.
- [Scrnd] Scribe OCR. Scribe ocr documentation, n.d. Accessed: August 5, 2024.
- [SDJDuLM20] A.V. Sidoroff-Dorso, J. Jewanski, S.A. Day, and Universitäts und Landesbibliothek Münster. *Synaesthesia: Opinions and Perspectives: 30 Interviews with Leading Scientists, Artists and Synaesthetes*. Wissenschaftliche Schriften der WWU Münster / 8. 2020.
- [SGB<sup>+</sup>23] Jannis Strecker, Kimberly García, Kenan Bektaş, Simon Mayer, and Ganesh Ramanathan. Socrar: Semantic ocr through augmented reality. In *Proceedings of the 12th International Conference on the Internet of Things, IoT '22*, page 25–32, New York, NY, USA, 2023. Association for Computing Machinery.
- [SGM06] Julia Simner, Louise Glover, and Alice Mowat. Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex*, 42(2):281–289, 2006.
- [SGMC14] Nicholas Smith, Fiona Glen, Vera Mönter, and David Crabb. Using eye tracking to assess reading performance in patients with glaucoma: A within-person study. *Journal of ophthalmology*, 2014:120528, 05 2014.
- [SH18] Abdul Saudagar and Mohammed Habeevulla. Augmented reality mobile application for arabic text extraction, recognition and translation. *Journal of Statistics and Management Systems*, 21:617–629, 07 2018.

- [SHC<sup>+</sup>08] Julia Simner, Jenny Harrold, Harriet Creed, Louise Monro, and Louise Foulkes. Early detection of markers for synaesthesia in childhood populations. *Brain*, 132(1):57–64, 11 2008.
- [SHCS19] Rebecca Smees, James Hughes, Duncan Carmichael, and Julia Simner. Learning in colour: children with grapheme-colour synaesthesia show cognitive benefits in vocabulary and self-evaluated reading. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374:20180348, 10 2019.
- [SHM<sup>+</sup>19] Mary Jane Spiller, Lee Harkry, Fintan McCullagh, Volker Thoma, and Clare N. Jonas. Exploring the relationship between grapheme colour-picking consistency and mental imagery. *Philosophical Transactions of the Royal Society B*, 374, 2019.
- [Sim07] Julia Simner. Beyond perception: synaesthesia as a psycholinguistic phenomenon. *Trends in Cognitive Sciences*, 11(1):23–29, 2007.
- [SL22] Sebastian Suggate and Wolfgang Lenhard. Mental imagery skill predicts adults’ reading performance. *Learning and Instruction*, 80:101633, 2022.
- [SLM09] Richard Skelton, Casimir Ludwig, and Christine Mohr. A novel, illustrated questionnaire to distinguish projector and associator synaesthetes. *Cortex*, 45(6):721–729, 2009.
- [Smi07] R. Smith. An overview of the tesseract ocr engine. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, volume 2, pages 629–633, 2007.
- [SMS<sup>+</sup>06] Julia Simner, Catherine Mulvenna, Noam Sagiv, Elias Tsakanikos, Sarah Witherby, Christine Fraser, Kirsten Scott, and Jamie Ward. Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, 35:1024–33, 02 2006.
- [SNE<sup>+</sup>12] C. Sinke, J. Neufeld, H.M. Emrich, W. Dillo, S. Bleich, M. Zedler, and G.R. Szycik. Inside a synesthete’s head: A functional connectivity analysis with grapheme-color synesthetes. *Neuropsychologia*, 50(14):3363–3369, 2012.
- [SNZ<sup>+</sup>14] Christopher Sinke, Janina Neufeld, Markus Zedler, Hinderk Emrich, Stefan Bleich, Thomas Münte, and Gregor Szycik. Reduced audiovisual integration in synesthesia - evidence from bimodal speech perception. *Journal of neuropsychology*, 8:94–106, 03 2014.
- [SOR<sup>+</sup>23] David J Schwartzman, Ales Oblak, Nicolas Rothen, Daniel Bor, and Anil K Seth. Extensive phenomenological overlap between training-induced and naturally-occurring synaesthetic experiences. *Collabra: Psychology*, 9(1):73832, 04 2023.

- [SP22] Susanne Seifert and Lisa Paleczek. Comparing tablet and print mode of a german reading comprehension test in grade 3: Influence of test order, gender and language. *International Journal of Educational Research*, 113:101948, 2022.
- [SPL<sup>+</sup>06] Julia M. Sperling, David Prvulovic, David E.J. Linden, Wolf Singer, and Aglaja Stirn. Neuronal correlates of colour-graphemic synaesthesia: Afmri study. *Cortex*, 42(2):295–303, 2006.
- [SS15] Avinoam Safran and Nicolae Sanda. Color synesthesia. insight into perception, emotion, and consciousness. *Current opinion in neurology*, 28:36–44, 02 2015.
- [Str35] J Ridley Stroop. Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643, 1935.
- [SWL<sup>+</sup>05] Julia Simner, Jamie Ward, Monika Lanz, Ashok Jansari, Krist Noonan, Louise Glover, and David Oakley. Nonrandom associations of graphemes to colours in synaesthetic and normal populations. *Cognitive neuropsychology*, 22:1069–85, 12 2005.
- [syn24] Syncalc: Calculator for synesthetes, 2024. Accessed: August 7, 2024.
- [The21] The Synesthesia Tree. Grapheme-colour synesthesia, 2021. Accessed: August 5, 2024.
- [TO17] Lamma Tatwany and Henda Chorfi Ouertani. A review on using augmented reality in text translation. In *2017 6th International Conference on Information and Communication Technology and Accessibility (ICTA)*, pages 1–6, 2017.
- [TOE24] TOEFL Resources by Michael Goodine. Toefl reading section, 2024. Accessed: August 5, 2024.
- [UAY21] Kyuto Uno, Michiko Asano, and Kazuhiko Yokosawa. Consistency of synesthetic association varies with grapheme familiarity: A longitudinal study of grapheme-color synesthesia. *Consciousness and Cognition*, 89:103090, 2021.
- [UEM14] Arcangelo Uccula, Mauro Enna, and Claudio Mulatti. Colors, colored overlays, and reading skills. *Frontiers in Psychology*, 5, 2014.
- [Uni19] Unity Asset Store - Paper Plane Tools. Opencv plus unity, 2019. Accessed: August 5, 2024.
- [Uni24] Unity Asset Store - Enox Software. Opencv for unity, 2024. Accessed: August 5, 2024.



- [Vuf24] Vuforia Developer. Vuforia developer portal, 2024. Accessed: August 5, 2024.
- [VWP22] Lisa-Marie Vortmann, Pascal Weidenbach, and Felix Putze. Atawar translate: Attention-aware language translation application in augmented reality for mobile phones. *Sensors*, 22(16), 2022.
- [War12] Jamie Ward. Synesthesia. *Annual review of psychology*, 64, 06 2012.
- [WAS<sup>+</sup>14] Matthew R. Watson, Kathleen A. Akins, Charlotte Spiker, Lindsay Crawford, and James T. Enns. Synesthesia and learning: a critical review and novel theory. *Frontiers in Human Neuroscience*, 8:98, 2014.
- [WJG<sup>+</sup>21] Ryan Joseph Ward, Fred Paul Mark Jjunju, Elias J. Griffith, Sophie M. Wuerger, and Alan Marshall. Artificial odour-vision synesthesia via olfactory sensory argumentation. *IEEE Sensors Journal*, 21(5):6784–6792, 2021.
- [WLSS07] Jamie Ward, Ryan Li, Shireen Salih, and Noam Sagiv. Varieties of grapheme-colour synaesthesia: A new theory of phenomenological and behavioural differences. *Consciousness and Cognition*, 16(4):913–931, 2007.
- [WS20] Jamie Ward and Julia Simner. Chapter 13 - synesthesia: The current state of the field. In K. Sathian and V.S. Ramachandran, editors, *Multisensory Perception*, pages 283–300. Academic Press, 2020.
- [WTLEK08] Jamie Ward, Daisy Thompson-Lake, Roxanne Ely, and Flora Kaminski. Synaesthesia, creativity and art: What is the link? *British journal of psychology (London, England : 1953)*, 99:127–41, 03 2008.
- [WW06] Nathan Witthoft and Jonathan Winawer. Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex; a journal devoted to the study of the nervous system and behavior*, 42:175–83, 03 2006.
- [WW13] Nathan Witthoft and Jonathan Winawer. Learning, memory, and synesthesia. *Psychological science*, 24, 01 2013.
- [WWE15] Nathan Witthoft, Jonathan Winawer, and David M. Eagleman. Prevalence of learned grapheme-color pairings in a large online sample of synesthetes. *PLOS ONE*, 10(3):1–10, 03 2015.



# Appendix

## User Study Questionnaire

### SynVis - Userstudy

[In Google anmelden](#), um den Fortschritt zu speichern. [Weitere Informationen](#)

\* **Gibt eine erforderliche Frage an**

**P# \***

Meine Antwort \_\_\_\_\_

Bitte wählen Sie die bevorzugte Sprache: \*

Deutsch

Englisch

[Weiter](#) [Alle Eingaben löschen](#)

SynVis (Part 1/4)

Do you have grapheme-color synesthesia? \*

Yes

No

Please indicate any visual impairments here:

Meine Antwort \_\_\_\_\_

How old are you? \*

Meine Antwort \_\_\_\_\_

Please select your gender:

Male

Female



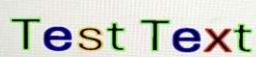


Non-binary

Sonstiges: \_\_\_\_\_

Zurück Weiter Alle Eingaben löschen

SynVis (Part 2/4)

Different modes

	Direct pixel colouring	Overlay
Font		
Outline		
Background		

Which of the following visualisation options did you prefer when carrying out the task? \*

- Direct pixel colouring
- Overlay (superimposition)

Why did you prefer this visualisation option? Please briefly explain your choice. \*

Meine Antwort \_\_\_\_\_

How do you rate the readability of the text of the visualisation options \*

	Very good - 1	2	3	4	Very poor - 5
Direct pixel colouring + font colour	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Direct pixel colouring + outline	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Direct pixel colouring + background	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
overlay + font colour	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
overlay + background	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Zurück Weiter Alle Eingaben löschen

SynVis (Part 3/4)

System Usability Scale \*

	Strongly disagree - 1	2	3	4	Strongly agree - 5
I think that I would like to use this system frequently	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the system unnecessarily complex	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought this system was easy to use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that I would need the support of a technical person to be able to use this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the various functions in this system were well integrated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought there was too much inconsistency in this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would imagine that most people would learn to use this system very quickly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found this system very cumbersome to use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt very confident using this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I needed to learn a lot of things before I could get going with this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

 Bei dieser Frage muss pro Zeile eine Antwort gegeben werden

Zurück Weiter Alle Eingaben löschen

SynVis (Part 4/4)

Do you have any comments, open questions or ideas for improvement?

Meine Antwort

Zurück Weiter Alle Eingaben löschen

Thank you for your participation!

Zurück Senden Alle Eingaben löschen

## Texts for Testing

A reading comprehension text for C1 level English reading called "How humans evolved language" [Brind] is chosen as sample text for the benchmark test as well as for the user study:

### A

Thanks to the field of linguistics we know much about the development of the 5,000 plus languages in existence today. We can describe their grammar and pronunciation and see how their spoken and written forms have changed over time. For example, we understand the origins of the Indo-European group of languages, which includes Norwegian, Hindi and English, and can trace them back to tribes in eastern Europe in about 3000 BC.

So, we have mapped out a great deal of the history of language, but there are still areas we know little about. Experts are beginning to look to the field of evolutionary biology to find out how the human species developed to be able to use language. So far, there are far more questions and half-theories than answers.

### B

We know that human language is far more complex than that of even our nearest and most intelligent relatives like chimpanzees. We can express complex thoughts, convey subtle emotions and communicate about abstract concepts such as past and future. And we do this following a set of structural rules, known as grammar. Do only humans use an innate system of rules to govern the order of words? Perhaps not, as some research may suggest dolphins share this capability because they are able to recognise when these rules are broken.

## C

If we want to know where our capability for complex language came from, we need to look at how our brains are different from other animals. This relates to more than just brain size; it is important what other things our brains can do and when and why they evolved that way. And for this there are very few physical clues; artefacts left by our ancestors don't tell us what speech they were capable of making. One thing we can see in the remains of early humans, however, is the development of the mouth, throat and tongue. By about 100,000 years ago, humans had evolved the ability to create complex sounds. Before that, evolutionary biologists can only guess whether or not early humans communicated using more basic sounds.

## D

Another question is, what is it about human brains that allowed language to evolve in a way that it did not in other primates? At some point, our brains became able to make our mouths produce vowel and consonant sounds, and we developed the capacity to invent words to name things around us. These were the basic ingredients for complex language. The next change would have been to put those words into sentences, similar to the 'protolanguage' children use when they first learn to speak. No one knows if the next step – adding grammar to signal past, present and future, for example, or plurals and relative clauses – required a further development in the human brain or was simply a response to our increasingly civilised way of living together.

Between 100,000 and 50,000 years ago, though, we start to see the evidence of early human civilisation, through cave paintings for example; no one knows the connection between this and language. Brains didn't suddenly get bigger, yet humans did become more complex and more intelligent. Was it using language that caused their brains to develop? Or did their more complex brains start producing language?

## E

More questions lie in looking at the influence of genetics on brain and language development. Are there genes that mutated and gave us language ability? Researchers have found a gene mutation that occurred between 200,000 and 100,000 years ago, which seems to have a connection with speaking and how our brains control our mouths and face. Monkeys have a similar gene, but it did not undergo this mutation. It's too early to say how much influence genes have on language, but one day the answers might be found in our DNA.