From Production to Reception:

Advancing Subtitling for the Deaf and Hard of Hearing in China

Rongyu Wang

B.A., M.A.

Thesis submitted for the degree of Doctor of Philosophy

School of Applied Language and Intercultural Studies

Dublin City University

August 2026

Supervised by:

Dr. Joss Moorkens (Dublin City University)

Dr. Ryoko Sasamoto (Nara Women's University)

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original and have conformed to the regulations on the use and declaration of Generative AI, and does not, to the best of my knowledge, breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work. I hereby certify that no Generative Artificial Intelligence (Gen AI) tools have been used in the creation of the thesis.

Signed: Rongyu Wang

ID No.: 22267830

Date: 25-08-2026

Acknowledgements

As I reach the end of this thesis, I would like to thank the people who helped me along the way.

To Dr. Joss Moorkens, I feel very fortunate to have had you as my supervisor. Thank you for believing in me, for supporting me, and for guiding me through this journey. I have always admired your passion for research, and your ability to manage so many different things at once. Beyond your meticulous scholarship and insightful teaching, I have also learned so much from the quiet kindness you show to those around you.

To Dr. Ryoko Sasamoto, my supervisor and the person with whom my PhD journey began. Thank you for giving me the life-changing chance to pursue my dream, for guiding me across oceans, and for making time for our meetings in your busy life. But most of all, thank you for encouraging me, steadily and unfailingly. And as you know, that is something many of us East Asian kids are not always given enough of.

To Dr. Angela Leahy, my independent panel member, thank you for coming to my annual review meetings and for always listening with such care. To Dr. Áine McGillicuddy, for organising PhD student gatherings every semester and giving us space to catch up. To Dr. Patrick Cadwell, for always taking a few minutes to talk with me so kindly. To Mr. Juan Pablo Auñón and Dr. Keisuke Honda, for welcoming me whenever I dropped into the LanguaCulture Space to practise Spanish and Japanese. And to Dr. Lijie Shao, for filling the Spring Festival with the warmth of Chinese refreshments and the comfort of home.

To Dr. Lucía Pintado Gutiérrez and Prof. Anna Matamala, my thesis examiners, thank you for such an enjoyable viva discussion and for your valuable comments on my thesis.

To the subtitlers I interviewed, for trusting me with your voices and the stories behind your work. To my participants, for opening your world to me and for believing in me. The signs you taught me will stay with me, always.

To my lovely C214 people: Catalina, for making office life brighter and livelier; Shafiq, for sharing wonderful Malaysian food; Nhung, for your kind words; and Mustafa, for your generous help.

To my SALIS Chinese family: Boyi, for the Christmas meal that made Dublin feel warmer; Jing, for the eye-tracking rescue; Mengdi, for all those warm coffee moments; and Yihan, for the delicious homemade dumplings.

To Yanan, my academic twin in the Netherlands, for every phone call that carried our worries and doubts across the distance. I found my way here, so stay the course, I know you will find yours too.

To Minghao, for your generosity in sharing your experience and knowledge. Before meeting you, I had never seen someone so excited about subtitling.

To Jiaojiao, for your kindness and warmth in Bristol. I carried those days with me, when I needed them most.

To Chris, Joy and Tobi, the best roommates I could ever have asked for. Thank you for making this place feel like home.

Lastly, to my parents, thank you for giving so much of yourselves and for letting me find my own way. I would not be where I am, or who I am, without you.

Publications and Presentations

Publications

Wang, Rongyu. (2025). "Review of Szarkowska & Jankowska (2025): Introducing Audiovisual Translation". *Babel.* https://doi.org/10.1075/babel.25221.wan

Presentations

2026.07. Subtitling Sound for Chinese Deaf and Hard-of-Hearing Viewers: An Eye-tracking Reception Study. *The XVI International Symposium on Emerging Research in Translation, Interpreting, Intercultural Studies and East Asian Studies*, Universitat Autònoma de Barcelona, Spain.

2026.06. SDH Subtitlers as Mediators: Decision-making and Professional Agency Across Organisational Contexts. *The International Conference: Translation at the Crossroads of Ideologies and Cultures,* Vilnius University, Lithuania.

2026.06. Processing Subtitles for the Deaf and Hard-of-Hearing: An Eye-tracking Reception Study with Hearing-impaired Viewers in China. *The Third International Postgraduate Conference: Looking Out / Looking In: New Perspectives in Translation and Interpreting*, Queen's University Belfast, UK.

2026.05. A Qualitative Study of Professional SDH Subtitlers' Decision-Making Processes. *The 6th Translation Studies Network of Ireland (TSNI) Conference*, Trinity College Dublin, Ireland.

2026.04. Sounding Out the Silent Image: Classroom Reflections on Human and AI Description. (Joint work with Joss Moorkens and María Isabel Rivas Ginel). *The Association of Programmes in Translation and Interpreting Studies (APTIS) Conference*, Cardiff University and Swansea University, UK.

2026.02. Subtitling for the Deaf and Hard of Hearing in China: An Eye-tracking Study. *The SALIS Seminar Series*, Dublin City University, Ireland.

2025.12. Exploring the Agency Role of SDH Subtitlers in Subtitling for the Deaf and Hard of Hearing. *The 8th International Association for Translation and Intercultural Studies (IATIS) Conference*, Sultan Qaboos University, Oman.

2025.06. Behind Subtitling for the Deaf and Hard of Hearing: Exploring the Decision-Making Processes of SDH Subtitlers. *The 11th European Society of Translation Studies (EST) Congress,* University of Leeds, UK.

2025.05. Promoting Subtitling for the Deaf and Hard of Hearing in China: From Subtitlers to Users. *The 11th Media for All Conference,* University of Hong Kong, China.

2024.12. Subtitling for the Deaf and Hard of Hearing in China: A Two-Phase Study on Production and Reception. *International Postgraduate Conference in Translation and Interpreting (IPCITI),* University of Manchester, UK.

2024.10. From Subtitlers to Viewers: Subtitling for the Deaf and Hard of Hearing in China. *The 3nd Symposium for Chinese Postgraduate in Ireland*, Trinity College Dublin, Ireland.

2024.05. Promoting Subtitling for the d/Deaf and Hard of Hearing in China: Insights from Both SDH Subtitlers and SDH Users. *Open Research Network Seminar,* University of Antwerp, Belgium (online).

2024.03. Applying Relevance Theory in Multimodal Translation: Subtitling Humour for the Deaf and Hard of Hearing. *The 2nd Postgraduate Workshop in Pragmatics,* National and Kapodistrian University of Athens, Greece.

2024.02. A Relevance-Theoretic Approach in Multimodal Translation: Subtitling Humour for the Deaf and Hard of Hearing. *Relevance Researchers' Network Seminar* (online).

Table of Contents

Declaration ---- ii

Acknowledgements ---- iv

Publications and Presentations ---- vi

Table of Contents ---- viii

List of Abbreviations ---- xiii

List of Key Terms ---- xiv

List of Tables ---- xv

List of Figures ---- xvii

Abstract ---- xviii

Chapter 1 Introduction ---- 1

1.1 Motivation and background ---- 1

1.2 Research aim and questions ---- 3

1.3 Synopsis of research design ---- 4

1.4 Structure of the thesis ---- 4

Chapter 2 SDH within Audiovisual Translation and Media Accessibility ---- 6

2.1 Chapter overview ---- 6

2.2 Audiovisual translation ---- 6

2.2.1 AVT modes ---- 8

2.3 Media accessibility ---- 10

2.3.1 Audio description ---- 11

2.3.2 Audio subtitles ---- 11

2.3.3 Sign language interpreting ---- 12

2.4 Subtitling for the deaf and hard of hearing ---- 13

2.4.1 What is SDH? ---- 13

2.4.2 SDH and closed captioning ---- 16

2.4.3 Target audience ---- 17

2.4.4 The nature of SDH: making sound visible ---------- 18

2.5 Chapter summary ---------- 25

Chapter 3 The Provision of SDH Across the World ---------- 27

3.1 Chapter overview ---------- 27

3.2 Global overview of SDH development ---------- 27

3.3 Pioneering countries ---------- 28

3.3.1 The United States ---------- 28

3.3.2 The United Kingdom ---------- 30

3.3.3 Canada ---------- 31

3.3.4 Australia ---------- 33

3.4 Established countries ---------- 34

3.4.1 Spain ---------- 34

3.4.2 Italy ---------- 35

3.4.3 Poland ---------- 35

3.4.4 Portugal ---------- 36

3.4.5 Ireland ---------- 37

3.5 Early-stage countries ---------- 38

3.5.1 Japan ---------- 38

3.5.2 South Korea ---------- 39

3.6 China ---------- 40

3.6.1 Legal and policy background ---------- 40

3.6.2 Academic research and effort ---------- 41

3.6.3 SDH on Chinese television and streaming platforms ---------- 43

3.7 Chapter summary ---------- 45

Chapter 4 Analysing SDH Guidelines Across Countries ---------- 46

4.1 Chapter overview ---------- 46

4.2 SDH guidelines selected for analysis ---------- 46

4.3 Criteria in SDH guidelines ---------- 48

4.3.1 Acceptability ---------- 49

4.3.2 Legibility ---------- 57

4.3.3 Readability ---------- 65

4.3.4 Relevance ---------- 67

4.4 Implications for SDH production in China ---------- 73

4.5 Chapter summary ---------- 75

Chapter 5 Methodology ---------- 76

5.1 Chapter overview ---------- 76

5.2 Research objectives and questions ---------- 76

5.3 Mixed-methods research ---------- 78

5.3.1 Introducing mixed-methods research ---------- 78

5.3.2 Mixed-methods research design ---------- 80

5.4 Phase I: Interview study with SDH subtitlers ---------- 82

5.4.1 Participants ---------- 82

5.4.2 Tools and materials ---------- 86

5.4.3 Interview study process ---------- 89

5.4.4 Qualitative data analysis ---------- 95

5.5 Phase II: Reception Study with SDH users ---------- 106

5.5.1 Participants ---------- 106

5.5.2 Tools and materials ---------- 110

5.5.3 Reception study process ---------- 126

5.5.4 Quantitative data analysis ---------- 131

Chapter 6 Findings and Discussion of the Interview Study ---------- 137

6.1 Chapter overview ---------- 137

6.2 Emerging themes (RQ1) ---------- 137

6.2.1 Workflow and process optimisation ---------- 138

6.2.2 SDH guidelines ---- 141
6.2.3 Technical and practical constraints ---- 143
6.2.4 Audience-centred considerations ---- 145
6.2.5 Subjectivity in interpretation ---- 148
6.2.6 Differences in content and background ---- 151
6.3 Implications for the next phase (RQ2) ---- 153
Chapter 7 Results and Discussion of the Reception Study ---- 156
7.1 Chapter overview ---- 156
7.2 Pre-task questionnaire results ---- 156
7.2.1 Demographic data ---- 156
7.2.2 Viewing habits ---- 160
7.2.3 Subtitle preferences ---- 161
7.3 Eye-tracking results ---- 164
7.4 Post-task questionnaire results ---- 170
7.4.1 Comprehension test results ---- 170
7.4.2 Engagement and enjoyment results ---- 177
7.5 Correlation analyses ---- 183
7.5.1 Correlation results for the overall sample ---- 183
7.5.2 Correlation results by subtitle group ---- 185
7.5.3 Summary of correlation findings ---- 188
7.6 Discussion of results ---- 191
7.6.1 Discussion of eye-tracking results (RQ3) ---- 191
7.6.2 Discussion of comprehension results (RQ4) ---- 193
7.6.3 Discussion of engagement and enjoyment results (RQ5) ---- 195
7.6.4 Discussion of correlation analyses ---- 197
Chapter 8 Conclusions ---- 199
8.1 Chapter overview ---- 199

8.2 Addressing RQs ---------- 199

8.2.1 RQ1 ---------- 199

8.2.2 RQ2 ---------- 201

8.2.3 RQ3 ---------- 202

8.2.4 RQ4 ---------- 202

8.2.5 RQ5 ---------- 203

8.3 Contributions ---------- 204

8.4 Limitations ---------- 206

8.5 Future work ---------- 208

References ---------- 211

Appendices ---------- 243

Appendices for Phase I ---------- 243

Appendix A: Plain Language Statement ---------- 243

Appendix B: Informed Consent Form ---------- 247

Appendix C: Ethical Approval ---------- 249

Appendix D: Three Rounds of Coding ---------- 250

Appendices for Phase II ---------- 251

Appendix E: Plain Language Statement ---------- 251

Appendix F: Informed Consent Form ---------- 255

Appendix G: Ethical Approval ---------- 257

Appendix H: Pre-task Questionnaire ---------- 258

Appendix I: Post-task Questionnaire ---------- 266

Appendix J: Eye-tracking Ratios of Participants ---------- 281

List of Abbreviations

AD: Audio description

AOI: Area of interest

AST: Audio subtitles

AV: Audiovisual

AVT: Audiovisual translation

CC: Closed captioning

DHH: Deaf and hard of hearing

IQ: Interview question

M: Mean

RQ: Research question

SD: Standard deviation

SDH: Subtitling for the deaf and hard of hearing

SLI: Sign language interpreting

TOI: Time of interest

TS: Translation studies

List of Key Terms

SDH subtitles: Subtitles including dialogue and additional auditory information (e.g., music)

ST subtitles: Subtitles including dialogue only, without additional auditory information

SDH group: Participant group viewing SDH subtitles

ST group: Participant group viewing ST subtitles

SDH subtitlers: Subtitlers who specialise in producing SDH subtitles

List of Tables

Table 2-1. D/deaf and hard-of-hearing groups (based on Neves, 2008b: 172) 18

Table 2-2. Semiotic channels in AV texts (based on Chiaro, 2009: 143)..... 20

Table 3-1. Captioning requirements for new programming (Robson, 2004: 40)..... 29

Table 3-2. Captioning requirements for pre-rule programming (*ibid.*)..... 29

Table 5-1. Detailed profiles of participants (Phase I)..... 85

Table 5-2. List of ten interview questions..... 88

Table 5-3. Detailed information on the nine interviews..... 92

Table 5-4. Multiple stages of coding (based on Sigauke & Swansi, 2020: 11) 99

Table 5-5. Examples of first-cycle coding..... 101

Table 5-6. Classification of hearing impairment levels used in China..... 107

Table 5-7. AOIs defined for the study..... 114

Table 5-8. Introductory prompt in the AV stimuli 118

Table 5-9. Two subtitle versions of the main video in the AV stimuli..... 119

Table 5-10. Closing prompt in the AV stimuli 120

Table 5-11. Summary of participant groups and viewing conditions (Phase II)..... 128

Table 6-1. Key factors influencing SDH subtitlers' decision-making process..... 137

Table 6-2. Comparison of SDH production workflows across three organisations..... 139

Table 7-1. Demographic information of participants (Phase II)..... 158

Table 7-2. Viewing habits of participants (Phase II)..... 160

Table 7-3. Subtitle preferences of participants (Phase II)..... 162

Table 7-4. Descriptive statistics of eye-tracking metrics 165

Table 7-5. Shapiro-Wilk normality tests for eye-tracking metrics 167

Table 7-6. Student's *t*-tests and Mann-Whitney *U* tests for eye-tracking metrics 169

Table 7-7. Holm-Bonferroni correction for eye-tracking metrics 170

Table 7-8. KR-20 value for the comprehension test 171

Table 7-9. Descriptive statistics of comprehension test 171

Table 7-10. Shapiro-Wilk normality tests for comprehension test 173

Table 7-11. Accuracy rates and Fisher's exact tests for comprehension questions 174

Table 7-12. Questions 4 and 8 of the comprehension test 175

Table 7-13. Cronbach's α values for engagement and enjoyment 177

Table 7-14. Descriptive statistics of engagement and enjoyment 178

Table 7-15. Shapiro-Wilk normality tests for engagement and enjoyment 180

List of Figures

Figure 2-1. Classification of subtitles (based on Díaz-Cintas & Remael, 2007: 14) 15

Figure 4-1. Serif and sans-serif fonts 58

Figure 5-1. Model of the decision-making process of SDH subtitlers 106

Figure 5-2. Coding system of China's disability identification card 108

Figure 5-3. Example of fixations and saccades 115

Figure 5-4. Example of a heatmap 117

Figure 5-5. Process of the reception study 129

Figure 7-1. Boxplots of eye-tracking metrics by group 166

Figure 7-2. Boxplots of comprehension test by group 172

Figure 7-3. Boxplots of engagement and enjoyment by group 179

Figure 7-4. Distribution of engagement responses (Q11-Q18) by group 181

Figure 7-5. Distribution of enjoyment responses (Q19-Q25) by group 181

Figure 7-6. Correlation analysis (overall sample) 184

Figure 7-7. Correlation analysis (SDH group) 186

Figure 7-8. Correlation analysis (ST group) 187

Figure 7-9. Summary of correlation analysis (overall sample) 190

Figure 7-10. Summary of correlation analysis (SDH group) 190

Figure 7-11. Summary of correlation analysis (ST group) 191

Abstract

From Production to Reception:
Advancing Subtitling for the Deaf and Hard of Hearing in China

Rongyu Wang

By making sound visible on screen, subtitling for the deaf and hard of hearing (SDH) supports access to audiovisual content for people with hearing impairments. At present, SDH is already well established in many developed countries, but in China its development is still in its infancy. It was this context that motivated the present research. In the process of designing SDH subtitles for the Chinese context, it was found that little academic attention had been paid to the production side, and that little was known about how professional SDH subtitlers make decisions in practice. Yet their experience could offer useful guidance for SDH production in China.

Against this background, this thesis approaches SDH through two connected phases: production (Phase I) and reception (Phase II). In Phase I, semi-structured interviews were carried out with nine professional SDH subtitlers to explore their decision-making in practice. The qualitative analysis identified six themes: workflow and process optimisation, SDH guidelines, technical and practical constraints, audience-centred considerations, subjectivity in interpretation, and differences in content and background. These themes were then used to inform the design of the SDH subtitles tested in the next phase. In Phase II, an eye-tracking reception study was conducted with 38 Deaf and Hard of Hearing (DHH) viewers in China to compare SDH subtitles with standard subtitles in terms of cognitive load, comprehension, and viewing experience. The findings suggest that, compared with standard subtitles, SDH subtitles were associated with better comprehension, with no evidence of increased cognitive load, and produced similar levels of engagement and enjoyment.

Overall, the thesis provides empirical support for the wider use of SDH in China and, it is hoped, will encourage further research on SDH among the Chinese DHH community.

Chapter 1 Introduction

"Blindness cuts us off from things; deafness cuts us off from people."

(Helen Keller, as cited in Chartrand, 2008: 18)

1.1 Motivation and background

For people who can hear, communication is a seamless part of everyday life, something that simply happens, but for those who cannot, it can be more challenging and often relies on different ways of access. For many Deaf and Hard of Hearing (DHH) individuals, sign language is one important means of communication in everyday life, yet in audiovisual (AV) media, access often takes another form, that is, subtitling for the deaf and hard of hearing (SDH). Unlike standard subtitles (ST subtitles) that usually only contain spoken dialogue, SDH subtitles also include speaker identification, music, and descriptions of sounds (Szarkowska & Jankowska, 2025: 169).

My first encounter with SDH came during my master's studies at Universitat Pompeu Fabra, where I took a course called Audiovisual Translation. It was in that course that I first learned about SDH and created my own SDH subtitles for a short clip. Since then, I have been intrigued by how SDH makes sound visible on screen and supports DHH viewers' access to AV content. I soon realised, however, that in my home country, SDH is largely absent and invisible, despite a sizeable DHH population. This absence has also been documented in existing research (Li, 2013; Li & Looms, 2015; Yang et al., 2025). In their work, the near invisibility of SDH on Chinese television repeatedly comes to light. A more recent example comes from Wu et al. (2026). Their study of subtitle use among the Chinese deaf community shows that "while intralingual and interlingual subtitles are widely available, they are not specifically designed for deaf viewers" (Wu et al., 2026: 4). As such, it can be said that SDH, as an accessibility service, is still rarely available to the Chinese DHH community. This

situation motivated me to embark on this project, in the hope of contributing, even in a small way, to SDH development in China.

When comparing the current situation in China with that in other countries, the difference is hard to miss. In many European countries, SDH has been available for decades. In the 1980s, Germany, Italy, and the Netherlands introduced early forms of SDH, and Portugal and Spain followed with similar services in the early 1990s (Remael, 2007: 24). To standardise SDH production, many countries have developed their own standards and guidelines. Among these, the BBC Subtitle Guidelines (2026) stand out as one of the most detailed, comprehensive, and regularly updated sets of guidelines. Alongside these national efforts, it is encouraging to see that a number of European researchers have also actively advocated for the interests of DHH audiences. Relevant examples include Matamala and Orero (2010), Romero-Fresco (2015), and Zárate (2021), all of which explore various SDH parameters in depth and investigate audience reception.

Furthermore, when reviewing previous research on SDH, one can see that a considerable volume of research mainly addresses audience reception, as seen in works such as Jelinek Lewis and Jackson (2001), Szarkowska et al. (2013), Muller (2015), Aleksandrowicz (2020), Tamayo et al. (2021), Zárate (2021), Cepak and Corrius (2023), to name a few. The production side of SDH, however, has received far less attention, particularly when it comes to studies that focus on SDH subtitlers and their working practices. One such exception is Agulló and Matamala's (2019) study, where SDH users and professional subtitlers were brought together in a focus group to discuss positioning, reading considerations, speaker identification, and the use of new SDH features in immersive media. Another example is Mliczak's PhD thesis (2019), which combined subtitle file analysis with interviews of professional subtitlers to trace the evolution of SDH norms and practices in Poland over time. Apart from these few studies, little is known about how SDH subtitlers make decisions in practice. It is undeniable that they are the ones who produce SDH, and their choices directly influence not only subtitle quality but also the viewing experience of DHH audiences. Therefore, it is worth approaching this professional group and looking into their actual work.

1.2 Research aim and questions

In light of the background described above, this thesis looks at SDH from both the production and reception sides. On the production side, it intends to learn from the experiences of professional SDH subtitlers in order to better understand how decisions are made in practice and what these can tell us about producing SDH subtitles. On the reception side, it seeks to explore how Chinese DHH audiences experience SDH subtitles, especially in comparison with ST subtitles that are currently available in China, with the aim of providing empirical evidence for future SDH design and production in the Chinese context.

Accordingly, the whole research project unfolds in two phases, one on the production side and the other on the reception side. Each phase comes with its own set of research questions: RQ1-RQ2 for the first, and RQ3-RQ5 for the second, as listed below:

Phase I

RQ1: What key factors influence SDH subtitlers' decision-making processes?

RQ2: What implications can be drawn from the study with SDH subtitlers that can inform the production of SDH subtitles?

Phase II

RQ3: How do the gaze patterns of DHH participants differ when watching videos with SDH subtitles compared to those with ST subtitles?

RQ4: Do DHH participants who viewed SDH subtitles achieve better comprehension performance than those who viewed ST subtitles?

RQ5: What differences can be observed in the overall viewing experience of DHH participants, in terms of engagement and enjoyment, when watching videos with SDH subtitles versus ST subtitles?

1.3 Synopsis of research design

For this study, I will use a mixed-methods approach, organised in two phases that move from production to reception. The first phase is qualitative and looks at SDH production. Concretely, I will invite professional SDH subtitlers to take part in semi-structured interviews, so that they can share their hands-on experiences and the challenges they face. Based on the findings from these interviews, SDH subtitles will then be created for use in the second phase.

The second phase turns to reception and is mainly quantitative. Here, I will investigate how Chinese DHH audiences engage with SDH subtitles, as opposed to ST subtitles commonly used in China. By combining eye-tracking data, comprehension tests, and questionnaires, this phase will present how Chinese DHH viewers process SDH subtitles and whether they offer any advantages over ST subtitles. The findings are intended to serve as a starting point for considering the wider provision of SDH in the Chinese context.

1.4 Structure of the thesis

This thesis consists of eight chapters.

Chapter 2 reviews the literature on SDH within Audiovisual Translation and Media Accessibility. It then provides an overview of SDH, covering its definition, target audience, and its fundamental role of making sound visible. To make this clearer, the chapter goes on to explain why sound matters, what types of sounds are included, and how they are represented in SDH.

Chapter 3 outlines the provision of SDH across different countries. It begins with a global overview of SDH development, then compares countries at different stages — pioneering, established, and early-stage contexts. It then zooms in on China and outlines the current state of SDH there, including its legal and policy background, relevant academic research, and current subtitle practices on television and streaming platforms.

Chapter 4 analyses SDH guidelines across different countries. Based on existing well-developed SDH guidelines, the analysis is organised into four aspects: acceptability, legibility, readability, and relevance. Acceptability covers punctuation, symbols, and numbers. Legibility considers visual presentation such as subtitle font, colour, and speed. Readability mainly focuses on whether to use verbatim or edited subtitles. Relevance, the final aspect, concerns what information is conveyed, omitted, or clarified in SDH.

Chapter 5 details the methodology used in this thesis. In Phase I, it introduces SDH subtitlers recruitment, interview question design, and interview procedures. It then moves on to Phase II, where it describes the inclusion criteria for DHH participants recruited in China, the tools and materials used, and the experimental procedures.

Chapter 6 presents the findings of the interview study with SDH subtitlers. It begins by introducing the themes that emerged from the data and then explains each one in detail. These themes include workflow, guidelines, constraints, audience considerations, subjective interpretation, as well as content-related and background differences that may influence subtitling practices. These findings provide crucial insights for the production of SDH subtitles in the next phase of the study.

Chapter 7 reports the results of the reception study with DHH participants. It shows data from the pre-task questionnaire (on participant profile, viewing habits, and subtitle preferences), eye-tracking results, and the post-task questionnaire (on comprehension, engagement, and enjoyment). To explore the potential relationships between eye-tracking data, comprehension, and questionnaire results, correlation analyses are also included. After presenting all the results, each research question proposed for this phase is discussed.

Chapter 8 concludes the thesis by bringing together the main findings of the study. Based on results from both the production and reception phases, it revisits each research question in turn. It then highlights the main contributions of the study and acknowledges its limitations. Finally, it closes by pointing out directions for future research.

Chapter 2 SDH within Audiovisual Translation and Media Accessibility

2.1 Chapter overview

This chapter situates SDH as a mode positioned at the crossroads of Audiovisual Translation (AVT) and Media Accessibility. It begins by outlining the concept of AVT and its main modes, then turns to media accessibility and shows how the two fields intersect. This intersection gives rise to several accessibility modes: audio description (AD; Section 2.3.1), audio subtitles (2.3.2), sign language interpreting (SLI; Section 2.3.3) and SDH (Section 2.4).

As SDH is the focus of this thesis, Section 2.4 discusses it in detail. It defines the term (Section 2.4.1), distinguishes SDH from closed captioning (Section 2.4.2), introduces the target audience (Section 2.4.3), and discusses how SDH makes sound visible on screen (Section 2.4.4), including which types of sounds are featured and how they are rendered. The chapter concludes with a brief summary in Section 2.5.

2.2 Audiovisual translation

AVT began to gain recognition as a distinct research field within Translation Studies (TS) around the 1980s (Perego & Pacinotti, 2020: 33). From that point onward, the term "AVT" has undergone several terminological transformations. In the 1980s, it was referred to as "constrained translation" or "subordinate translation" (Titford, 1982). Yet, the adjectives "constrained" and "subordinate" were considered by many to carry somewhat negative connotations that may contribute to a less favourable perception in some contexts (Díaz-Cintas & Remael, 2007: 11). With the flourishing development of film industry, some scholars began to use the terms "film translation" (Delabastita, 1989) and "screen translation" (Gambier, 2003) to refer to AVT, but these appear insufficient in covering the diverse realms of AV media, including films, television, video games, opera performances, among others. Consequently, a more inclusive alternative term has emerged, namely "(multi)media translation" (Gambier & Gottlieb, 2001), which showcases its salient feature that the information is disseminated through multiple media and channels (Díaz-Cintas &

Remael, 2007: 12). Over the past decade, an even broader term has been proposed by some scholars (Chaume, 2018; Romero-Fresco & Chaume, 2022), namely "media localisation", which Chaume (2018: 93) defines as "encompassing all kinds of audiovisual translation and all kinds of media adaptation." This suggests that media localisation not only involves the translation of AV content (as typically performed in AVT) but also the adaptation of media products for global audiences catering to specific language, cultural, and market requirements. While these various terms have emerged over time, AVT has gradually become the most widely accepted and used term in the academic field. As Bolaños García-Escribano (2025: 21) argues, the expansion of research, scholarly literature, and postgraduate programmes has cemented its status as a well-rooted and stable field.

AVT, as its name implies, is a translation activity involving both sounds and images. It primarily deals with the translation of AV media, such as films, television programmes, video games, websites, and live performances. In academic discourse, Chiaro (2013: 1) defines AVT as "the term used to refer to the transfer from one language to another of the verbal components contained in audiovisual works and products." Chaume (2013: 105) further describes that "AVT is a mode of translation characterised by the transfer of AV texts either interlingually or intralingually." That is to say, AVT includes not only transfer between different languages but also transfer within the same language. However, these definitions primarily account for traditional AVT modes such as dubbing and subtitling, which mainly involve interlingual or intralingual translation of AV products. They seem insufficient to describe the expanding scope of AVT, particularly modes such as AD, SDH and SLI. For Taylor (2020), such modes position AVT within the domain of intersemiotic translation, where one system of signs is interpreted through another.

On this basis, AVT can be understood as inherently intersemiotic, given the multimodal character of AV texts. As noted by many scholars (e.g., Gottlieb, 2001; Zabalbeascoa, 2008; Gambier, 2013), AV texts consist of two types of signs, linguistic and non-linguistic, and two modes of transmission, visual and auditory. As a result, four systems of signs can be

identified in AV products: verbal auditory, non-verbal auditory, verbal visual, and non-verbal visual.

In fact, the notion of intersemiotic translation dates back to 1959, when Jakobson (1959: 233) introduced three types of translation in his article *On Linguistic Aspects of Translation*:

1) "**Intralingual translation** or *rewording* is an interpretation of verbal signs by means of other signs of the same language", which means to explain or paraphrase the meaning of a text within the same language. (e.g., from Old English to Modern English);
2) "**Interlingual translation** or *translation* proper is an interpretation of verbal signs by means of some other language", which involves a transfer of meaning from one language to another. (e.g., from English to Chinese);
3) "**Intersemiotic translation** or *transmutation* is an interpretation of verbal signs by means of signs of nonverbal sign systems", which means that a text can be translated not only through words but also through other semiotic signs, such as images, sounds or gestures (e.g., spoken language to sign language). Another common example is when a traffic light turns to be green, which can be translated as "pass" in written words.

Among these three types, intersemiotic translation differs from intralingual and interlingual translation in that it extends the scope of translation beyond linguistic boundaries, broadening the transfer of meaning across different sign systems. This feature can be clearly observed in the practices of AD, SDH and SLI. In the context of SDH, non-verbal auditory signs are transformed into verbal visual signs, thereby enabling audiences with hearing loss to access audio content that would otherwise be inaccessible to them. For this reason, SDH is widely acknowledged as a form of intersemiotic translation (Chiaro, 2013; Remael & Vercauteren, 2010; Taylor, 2020). Similarly, AD and SLI also involve shifts between different semiotic systems, albeit in different directions. These modes further reflect the intersemiotic dimension of AVT.

2.2.1 AVT modes

In the field of AVT, it is generally accepted that the various types of AVT practices can be broadly grouped into two main categories: revoicing and subtitling, the former also known

as dubbing[1] and the latter as captioning (Luyken et al., 1991; Díaz-Cintas & Orero, 2010; Chaume, 2020). Revoicing, by definition, replaces the original dialogue with a new or live recording in the target language, so that the original dialogue is no longer heard as the primary verbal track (Díaz-Cintas & Orero, 2010: 441). As for subtitling, it "consists in using written text on screen to account for the original dialogue exchanges" in AV productions (*ibid.*).

Over time, the categorisation of AVT practices within these two broad groups has been refined and expanded. In 1991, Luyken et al. (1991: 40) proposed a classification in which the category of subtitling comprised two sub-categories: (1) traditional subtitling and (2) simultaneous subtitling. The group of revoicing was further divided into three sub-categories: (1) lip-sync dubbing; (2) voiceover/narration; and (3) free commentary. More recently, Bolaños García-Escribano (2025: 23-45) introduced an updated classification. In this updated classification, revoicing includes: (1) dubbing, partial dubbing and off-screen dubbing; (2) voiceover; (3) interpreting and sign language; (4) AD; (5) free commentary; (6) remakes; and (7) amateur dubbing. Meanwhile, subtitling is also expanded to include different forms and contexts of practice: (1) interlingual subtitling; (2) intralingual subtitling; (3) surtitling; (4) live subtitling; (5) integrated subtitles; and (6) amateur subtitling.

This evolution in classification reflects the increasing diversification of AVT modes over the past few decades. To a large extent, this trend is attributed to advancements in AVT research as well as economic and technological developments, but other factors have also played a part. In particular, the rise of over-the-top platforms such as Netflix, shifts in global media consumption habits, and the growing demand for accessibility services have further

[1] It is worth noting that, although “revoicing” and “dubbing” are often considered synonymous, Chaume (2020) points out that there are still distinctions between the two terms. According to Chaume (2020: 105), revoicing is a hypernym term that includes various audio translation practices, including dubbing. Revoicing includes a broader range of practices such as voiceover, narration, fandubbing, while dubbing specifically involves replacing the original soundtrack with a new one.

contributed to the expansion and transformation of AVT practices (Bolaños García-Escribano, 2025; Szarkowska & Jankowska, 2025).

2.3 Media accessibility

According to Greco and Jankowska (2020: 62), the notion of "accessibility" began to emerge within the scholarly environment of AVT in the late 1990s and early 2000s. In 2003, Gambier noted that while many AVT products targeted the general public, minority groups such as children, older adults, people with hearing impairments, and visually impaired individuals received relatively little attention. In this context, Gambier (2003: 179) argues that "the key word in screen translation is now accessibility" and calls for better access to AV products for these target groups. Since then, media accessibility has become an important topic in AVT, receiving ongoing scholarly attention (Greco & Jankowska, 2020: 58).

As media accessibility has become an important area of research in AVT, scholars have sought to define it in different ways. Díaz-Cintas and Remael (2007: 12) describe media accessibility as "making an AV programme available to people that otherwise could not have access to it." Similarly, Chiaro (2013: 4-5) defines media accessibility as "the provision of audiovisual products such as plays, films, and opera for all members of the public, including those who are in some way sensorily challenged." Likewise, Baños (2017: 485) indicates that media accessibility seeks to "study how linguistic and sensory barriers can be overcome to make audiovisual products accessible to all." What these definitions collectively suggest is that, in essence, media accessibility aims to ensure that AV content is accessible to all individuals, regardless of language or sensory barriers.

Next, the discussion turns to selected media accessibility services that are relevant to AVT: AD for blind and partially sighted audiences, audio subtitles for viewers who cannot easily access written subtitles, and SLI for people with hearing impairments. Since SDH is the focus of this thesis, AD, audio subtitles and SLI are discussed only briefly, mainly to situate SDH within the field of media accessibility before moving on to a fuller discussion of SDH.

2.3.1 Audio description

AD, has previously been known as “video description”, "described video”, and “descriptive narration”, among other related terms (Leung, 2018: 30). In simple terms, AD translates the visual elements of a film or show into verbal narration for blind and visually impaired audiences. A more detailed definition is provided by Benecke (2004: 78), an experienced practitioner in the field, who characterises AD as:

> “a technology that enables blind and visually impaired individuals to engage with drama, films, and television programmes. Additional narration describes actions, body language, facial expressions, scenes, and costumes. The commentary is inserted between dialogues and does not interfere with crucial sound and music effects.”

AD can be traced back to as early as the 1980s, when it was already being implemented in pioneering Western countries such as the US, the UK, and Spain (Liu et al., 2022: 180). Gradually, AD became more institutionalised. In the US, efforts to regulate AD began in 2000, when the Federal Communications Commission introduced rules requiring broadcasters to provide AD. Although these regulations were soon overturned, the policy debate did not end there. A decade later, the *Twenty-First Century Communications and Video Accessibility Act* (2010) reinstated the mandate. It required major broadcasters in leading markets to provide at least four hours of description per week (Snyder, 2022: 537). In the UK, television broadcasters were required to provide AD for 4% of their programmes in 2001 (Gambier, 2003: 176). Two decades later, the four major UK broadcasters (BBC, ITV, Channel 4, and Sky) committed to increasing this provision to at least 20% of their content (Ofcom, 2023). In contrast to these more established contexts, AD in China is still at an early stage of development and largely relies on grassroots volunteers (Liu et al., 2025). In an effort to move towards professionalisation, scholars such as Leung (2018), Tor-Carroggio (2020), and Liu et al. (2023) have begun to explore AD within the Chinese context in recent years.

2.3.2 Audio subtitles

Audio subtitles (AST), also referred to as spoken subtitles, transform written subtitles into spoken output, either through human voicing or text-to-speech technology. Originally, this

service was designed primarily for individuals with visual impairments, enabling them to access subtitle information aurally, but subsequent research has broadened its potential user base. Gambier (2013: 50) points out that AST benefits not only blind people, but also dyslexic individuals, older adults, partially sighted people, and anyone unable to read subtitles at a fast speed. This inclusive perspective is reinforced by studies on ageing populations. In Poland, Jankowska (2019) explores the potential for senior citizens to serve as secondary users of AST, and Szkriba (2022) similarly reports that older adults struggled with rapid subtitles but responded positively to AST provided through a mobile application. Beyond specific user groups, Orero et al. (2020) enumerate various situational barriers to written subtitle access, including small fonts on smartphones, distant displays on public screens, excessively fast presentation, and subtitles used only for multilingual dialogue in partially subtitled films, all of which may necessitate the use of AST.

Despite the wide range of users and scenarios that AST can serve, Orero et al. (2020) caution that it should not be understood as a replacement for AD. Instead, AST complements AD by making written subtitle information available aurally.

2.3.3 Sign language interpreting

Sign language is a visual mode of communication that relies on manual signs, facial expressions, and body movements (Miller, 2007). To facilitate communication between hearing individuals and hearing-impaired individuals, SLI was introduced. Essentially, it involves “interpreting to and from a signed language from either a spoken language or another signed language” (Leeson & Vermeerbergen, 2010: 324). In the Chinese media context, SLI has also been discussed as a relatively widespread accessibility service, particularly in television news programmes (Xiao & Li, 2013).

Although SLI provides deaf individuals with barrier-free communication opportunities, it also presents certain challenges and limitations. Firstly, sign language interpreters may not always be readily available, especially in certain regions or at specific events, making it difficult for deaf individuals to access media information. This is mainly because sign

language interpreter training is still inadequate in many countries worldwide (Napier, 2009). Another limitation is relevant to AV media, where sign language interpreters are usually positioned on screen, often boxed into a small corner of the image. This requires viewers to divide their attention between the interpreter and the AV content, which may potentially result in missed information. A further limitation lies in the belief that sign language is universally used within the deaf community. However, as Neves (2007a) points out, not all pre-lingually deaf people are able to use it fluently. This means that SLI may not be suitable for all DHH viewers.

After this brief discussion of SLI, the following section now moves on to another accessibility service for the DHH community, namely SDH.

2.4 Subtitling for the deaf and hard of hearing

2.4.1 What is SDH?

As its name implies, SDH is a type of subtitling designed for viewers who are deaf or hard of hearing. Díaz-Cintas and Remael (2021: 10) define SDH as:

> "a practice that consists in presenting on screen a written text that accounts for the dialogue and its paralinguistic dimension, as well as for music, sounds and noises contained in the soundtrack, so that audiences with hearing impairments can access audiovisual material."

At first glance, SDH may appear similar to other types of subtitling (e.g., live subtitling), as they all present text on screen. However, its treatment of non-dialogue information gives SDH distinct features that set it apart from other forms of subtitling. As highlighted by Szarkowska (2020: 253), "SDH contains not only dialogue, but also other types of information such as speaker identification, descriptions of sounds, music and paralinguistic features." Regarding speaker identification, Pereira (2010: 88) concludes several techniques commonly employed to identify characters in SDH, including: 1) positioning the subtitle under the speaker; 2) using labels with the name of the speaker before the subtitle; and 3) assigning a different colour to each speaker. When it comes to descriptions of sounds,

Pereira and Lorenzo (2005: 24) distinguish between sounds made by characters on screen, such as coughing, whistling or sneezing, and sounds emitted from the external environment, such as phone ringing, doorbells or knocking at the door. Additionally, music is also conveyed visually in SDH. According to Zárate (2021: 77), once an SDH subtitler has perceived a connection between music and a specific scene or plot, it is necessary to describe this music in the subtitles, because music "describes a mood or adds to the enjoyment of the programme." Lastly, for DHH viewers, paralinguistic features, such as shouting, screaming, and whispering, are equally crucial as spoken dialogue, as they also "contribute to the development of the plot or to the creation of atmosphere" (Díaz-Cintas & Anderman, 2009: 5).

These additional layers of information also make SDH more complex in terms of subtitle content and presentation than ST subtitles. First, SDH is more commonly associated with pre-recorded content (Díaz-Cintas & Remael, 2021: 10). However, it is not limited to such content. As Bartoll's taxonomy shows, SDH can be produced not only in advance, but also simultaneously (Bartoll, 2008: 263). Second, because SDH includes additional auditory information, subtitles may extend to three or even four lines (Díaz-Cintas, 2013: 280). Third, since soundtracks are not directly accessible to viewers with hearing impairments, musical symbols are often added when music starts playing, and in some cases, emoticons or even emoji included to express the speaker's feelings (Neves, 2005). However, in that case, Zárate (2021: 99) points out that some specialised subtitling software does not fully support the use of such symbols or icons.

From a linguistic standpoint, subtitling can be intralingual or interlingual. In this sense, SDH can fall into either category, as shown in **Figure 2-1** (Díaz-Cintas & Remael, 2007: 14). In practice, Zárate (2014: 41) contends that SDH is often intralingual, reflecting its early development on television in the 1980s. Other scholars (Chaume, 2013: 113; Baños, 2017: 482; Miquel-Iriarte, 2017: 27) likewise note that SDH tends to operate within the same language. On the question of whether SDH is inherently intralingual, Szarkowska (2013: 68) explains that this perception largely stems from scholarly discussions on SDH primarily

focusing on English-speaking countries, where SDH provision has been more established. In these countries, AV productions are predominantly in English, so SDH has naturally been understood as intralingual. This Anglophone focus also reflects a broader tendency in AVT research, where English has often been treated as a major source language because of the global circulation of English-language AV content (Nikolić, 2018: 184). Nevertheless, Neves (2008a: 131), in the journal article *Ten fallacies about Subtitling for the d/Deaf and the hard of hearing*, argues that interlingual SDH is increasingly common across television, cinema, conferences and live performances. One clear example is Poland, where a significant portion of television programming is now available with interlingual SDH (Szarkowska, 2013: 69). Therefore, it can be argued that SDH can be either intralingual or interlingual, though it is more commonly associated with the intralingual form.

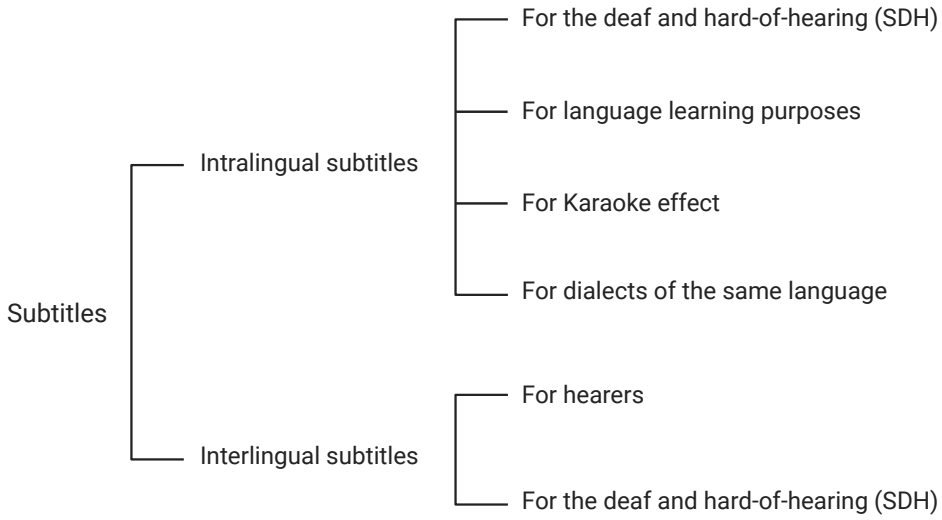


Figure 2-1. Classification of subtitles (based on Díaz-Cintas & Remael, 2007: 14)

Due to the longstanding belief that SDH is primarily intralingual and does not involve the transfer from one language to another, questions have been raised as to whether it falls within the scope of translation (Neves, 2005: 105). In response to this debate, Sala Robert (2016: 29) argues that "although SDH does not always imply the transfer from a source to a target language — which is the defining feature of translation — it always provides access." She further explains that SDH subtitlers help overcome the sensory barrier between AV

content and DHH audiences, much like translators address language barriers. In line with this view, Zárate (2021: 4) observes that scholars in TS have increasingly recognised SDH as falling within the domain of translation. This shift can be linked to Jakobson's (1959) broader conceptualisation of translation, where he proposes that intralingual, interlingual and intersemiotic translation all fall within the scope of translation. Given this, SDH can be regarded as a form of translation, more specifically, a type of intersemiotic translation. This concept has been introduced in Section 2.2.

2.4.2 SDH and closed captioning

Closed captioning (CC) is often discussed alongside SDH (Díaz-Cintas & Remael, 2007: 14; Neves, 2008a: 129; Szarkowska & Jankowska, 2025: 10). Originating in the US, CC was first demonstrated at the First National Conference on Television for the Hearing Impaired in 1971, after which the UK adopted teletext-based captioning systems. Technically, captions can be either open or closed: open captions are always visible on the screen and cannot be turned off, whereas closed captions can be turned on or off by viewers (Díaz-Cintas & Remael, 2007: 21).

In much of the literature, SDH and CC are treated as largely equivalent and differ mainly in regional usage. SDH is more often mentioned in the UK and other European countries, while CC predominates in the US, Canada and Australia (Szarkowska, 2020).

But still, strictly speaking, there are some subtle differences between SDH and CC. Firstly, according to Toledo (2018: 159), CC has been mandatory in US television programming since the *Americans with Disabilities Act* (1990), but SDH is more commonly found in other media formats, such as DVDs. Second, their encoding systems differ, which has practical implications for distribution. Unlike CC, which relies on text commands and control codes, SDH uses bitmap image encoding (Verbit, 2022). This makes SDH compatible with a broader range of media such as streaming services and Blu-ray DVDs, as HDMI (High-Definition Multimedia Interface) supports image-based subtitles but not command-based ones. Third, CC is usually presented as white text against a black background (Toledo, 2018). SDH, by

contrast, often uses colour coding to indicate different speakers (Díaz-Cintas & Anderman, 2009).

Despite these technical distinctions, Neves (2008a) cautions against overstating the divide between SDH and CC. She argues that the debate is largely terminological and regionally conditioned rather than conceptual, as both terms ultimately refer to the same accessibility service: providing DHH audiences with meaningful access to AV content. For consistency, this thesis uses the term "SDH", given its widespread use in AVT research.

2.4.3 Target audience

It is widely recognised that the target users of SDH are individuals with hearing impairments. In fact, Neves (2008b: 173) points out that three distinct groups can be identified: the Deaf, the deaf, and the hard of hearing, as summarised in **Table 2-1**. These groups differ not only in terms of hearing level, but also in relation to language background, cultural affiliation and the timing of hearing loss.

Following this distinction, it is important to recognise the diversity within SDH audiences. Previous research (e.g., Gambier, 2003: 175; Szarkowska, 2020: 251; Zárate, 2021: 21) has shown variations in the degree, type and onset of hearing loss, as well as in education, literacy and language skills within this population. Such diversity is likely to influence how DHH individuals experience SDH and what they find helpful.

D/deaf and hard of hearing people		
D/deaf people		Hard-of-hearing people
Deaf refers to people who belong to the Deaf community and who use sign language as their mother language.	deaf refers to people who have acquired deafness but use the oral language and relate to the hearing society as their own.	Hard-of-hearing people refers to people who have some degree of residual hearing and will relate to sound as hearing people do, even if they may perceive sound to a less extent.

Table 2-1. D/deaf and hard-of-hearing groups (based on Neves, 2008b: 172)

This heterogeneity is not only evident at the individual level but also at the national level. Romero-Fresco (2015: 10) states that differences in social, economic and political contexts, together with the diversity of national AV landscapes and hearing-impaired communities, make it difficult to standardise SDH practices across countries. Because of this, instead of imposing a single set of SDH criteria across Europe, Romero-Fresco (2015) worked with researchers in the UK, Spain, Italy, Poland, Denmark, France and Germany to carry out reception studies that could inform standards tailored to each national context. What this suggests is a willingness to embrace diversity rather than ignore it, and to meet different DHH audiences where they are.

2.4.4 The nature of SDH: making sound visible

2.4.4.1 Why does sound matter?

According to Kalinak (2015: 1), "sound has played an important role in the cinematic experience across the entire history of moving images." This was true even in the so-called silent era. At that time, films did not yet include recorded dialogue or soundtracks, but exhibitors still sought ways to introduce sound into screenings (Kalinak, 2015: 2).

Their efforts were reflected in several techniques. One common approach was musical accompaniment. As Cooke (2008: 1) notes, silent films were often accompanied by pianists, violinists, or other sound-producing devices. In larger cities with grand theatres, complete orchestras were sometimes engaged to create atmosphere or reflect characters' psychological states. Beyond music, exhibitors also integrated sound effects into silent films, such as doorbells, car engine sounds, and gunshots. Initially, this was done through rudimentary methods such as waving feathers to imitate birds in flight. Later, more sophisticated approaches emerged, including mechanical devices like the Kinematophone or Allefex, which were synchronised with the film to enhance the viewing experience (Cooke, 2008: 4). In some cinemas, "noisemen" (sound-effects performers) were employed to provide live effects using various instruments (Larsen, 2005: 19). Alongside these techniques, some screenings also featured film narrators or voice actors, who delivered live interpretations and dubbed dialogues during screenings. This led Larsen (2005: 19) to note that the silent film "talked". The tradition was prominent in Japan, where narrators were known as "benshi" (弁士), and also appeared in other countries under different names: "byeonsa" (변사) in Korea, "bonimenteur" in France, and "spieler" in the US (Ivarsson, 2002; Gambier, 2013: 58; Orrego-Carmona, 2013: 298).

With the advent of the talkies in the 1930s, the role of sound in cinema became more important. Synchronised dialogue, music and sound effects transformed narrative practices to some extent, allowing filmmakers to build mood and convey emotion more precisely. For instance, background music can reflect a character's emotional state (happy, angry, calm, sad) and provide information about the environment they are in (school, hospital, airport, theatre). Moreover, frequently used sounds such as gunshots, car engines and explosions have long been central to action films, as they help create a tense atmosphere that makes audiences feel apprehensive and alert.

Having reviewed historical sound practices, it is clear that sound plays an irreplaceable role in guiding how films are understood. For DHH audiences who rely on subtitles, much of this

auditory information needs to be made available in written form. The following section therefore introduces the types of sounds that are commonly included in SDH.

2.4.4.2 What types of sounds are included in SDH?

"SDH is all about sound, or to be more specific, SDH is all about making sound visible" (Neves, 2008b: 177). To understand what needs to be made visible in SDH, it is necessary to consider how sound functions within AV texts more broadly. As outlined in Section 2.2, AV texts consist of four semiotic channels: verbal audio, non-verbal audio, verbal visual and non-verbal visual. Chiaro (2009: 143) lists the elements included in each channel, as presented in the **Table 2-2**.

	VISUAL	AUDIO
NON-VERBAL	Scenery, lighting, costumes, props, etc. Gesture, facial expressions, body movement, etc.	Music, background noise, sound effects, etc. Laughter, crying, humming, body sounds (breathing, coughing, etc.)
VERBAL	Street signs, shop signs, written realia (newspapers, letters, headlines, notes, etc.)	Dialogues, song lyrics, poems, etc.

Table 2-2. Semiotic channels in AV texts (based on Chiaro, 2009: 143)

As shown in the **Table 2-2**, sound in AV texts can appear in both verbal audio and non-verbal audio channels. These two channels are directly relevant to SDH, because SDH is tasked with conveying the entirety of the audio channel to the viewers, including the words spoken but also other meaningful sounds such as paraverbal features, sound effects and music (Zabalbeascoa, 2025: 80). While spoken words are already represented through dialogue subtitles, SDH must additionally account for the information contained in the non-verbal audio channel.

Given this need, many AVT scholars (Arnáiz-Uzquiza, 2012; Szarkowska, 2020; Ogea Pozo & Talaván, 2024; Szarkowska & Jankowska, 2025) have discussed what types of non-verbal audio information should be conveyed in SDH, although it seems that no clear academic consensus has yet been reached.

In her doctoral thesis, Neves (2005) identifies **three** central elements — human voice, sound effects, and music — that are integral to SDH and should be described. Yet, later research broadens this classification and generally recognises **four** categories of non-verbal audio information that should be included in SDH: character identification, paralinguistic information, music and sound effects (Arnáiz-Uzquiza, 2012; Sala Robert, 2016; Mliczak, 2019; Szarkowska, 2020).

On the basis of this broad agreement on the four categories, Zárate (2021) suggests that the categorisation can be further expanded. She proposes including descriptions of silence, considering its role in heightening the tension between characters and emphasising the quiet ambience of crucial scenes, especially in suspenseful narratives. This view is supported by Cepak and Corrius (2023), whose reception study indicates that signalling the absence of sound can be beneficial for DHH viewers. Despite its recognised importance, silence is still frequently overlooked in practice. Khoshsaligheh et al. (2024), in their analysis of five Iranian feature films, conclude that silent scenes are often overlooked in SDH.

Based on the above discussion, non-verbal audio information in SDH can be broadly grouped into **five** categories: character identification, paralinguistic information, music, sound effects and silence. The next section considers how these sounds are conveyed visually.

2.4.4.3 How are sounds made visible in SDH?

Building on the previous section, this section presents how the literature describes the visual rendering of five types of non-verbal audio information in SDH: character identification, paralinguistic information, music, sound effects, and silence.

Character identification

Since DHH audiences, unlike hearing viewers, cannot rely on vocal qualities such as timbre or pitch to tell speakers apart, SDH needs to make it clear who is speaking whenever this is not immediately apparent on screen. In reviewing the literature, four principal strategies for identifying speakers in SDH have been proposed (Pereira, 2010; Arnáiz-Uzquiza, 2012; Szarkowska, 2020; Zárate, 2021):

- **Colours:** This method is to assign each character a specific colour that is kept consistent throughout the programme. Szarkowska and Jankowska (2025: 176) note that this method "is probably the most popular in television subtitling in Europe," and Pereira (2010: 89) mentions its wide adoption across Spanish television channels. In Spain, a national survey by Cepak and Corrius (2023) reports that both DHH and hearing viewers rated colours as the most appropriate way to identify characters on screen.
- **Labels:** With this method, speakers are identified by a name, role, or generic descriptor placed before the subtitle, usually in parentheses, square brackets, or capital letters. Depending on the context, labels may refer to a specific character by name, or indicate a role such as [Jack] or [narrator]. When the character's name or role is not known, generic labels such as [man] or [speaker 1] can be used. The main advantage of this approach is that it identifies speakers clearly and consistently, but its drawback is that longer subtitles may increase the viewer's reading load (Szarkowska & Jankowska, 2025: 176).
- **Dashes:** Some SDH formats use a dash before each line of dialogue to show speaker changes. This method makes turn-taking clearly without extra words, but when several speakers appear in the same subtitle, dashes can be less clear and labels are preferred (Zárate, 2021: 65).
- **Placement:** Another way to signal who is speaking is to place the subtitle closer to the character rather than keeping it centred at the bottom. This approach can be intuitive and is more common in the US, but it is used less frequently than other identification methods. The issue is that it can obscure parts of the image, and not all subtitling

software allow subtitles to be positioned freely on the screen (Szarkowska & Jankowska, 2025: 177).

Paralinguistic information

Paralinguistic information concerns the speaker's vocal qualities, including intonation, pitch, volume, rhythm, pauses, coughs and laughter (Sheth, 2017: 72). Unlike sound effects, which describe external auditory events, paralinguistic features are directly tied to human speech. In SDH, several common conventions for marking these features have been discussed by Szarkowska (2020) and Khoshsaligheh et al. (2024), and can be summarised as follows.

- **Uppercase in square brackets:** This method is one of the most common ways to mark paralinguistic features, including accents, emphasis, whispering and laughing, as in [LAUGHS], [SHOUTS], [WHISPERS] or [ITALIAN ACCENT].
- **Punctuation marks:** Different punctuation marks can be used to express various paralinguistic features. Zárate (2021) shows that sarcasm can be indicated through the insertion of an exclamation mark in brackets following the utterance, as in the example "That's so generous of you (!)." And also, ellipses can be used to signal an interruption or a prolonged pause.
- **Spelling:** According to Szarkowska and Jankowska (2025: 180), this method is used to represent pronunciation or speech disturbances such as stammering (e.g., P-p-pardon?).
- **Italics:** Subtitles in italics are often used to mark emphasis or shifts in speech (Khoshsaligheh et al., 2024), although Zárate (2021) recommends that they are better reserved for off-screen voices. To avoid confusion, their use should be applied consistently throughout a programme.

Music

According to Larson (2005), music in film fulfils three main functions: first, it provides formal functions that structure scenes and transitions; second, it performs narrative functions that signal time, space and character information; and third, it delivers emotional functions that

shape the moods and feelings of the story both globally and in specific moments. Given its narrative significance, music also needs to be made visible in SDH whenever necessary.

In terms of form, Aleksandrowicz (2020) shows that musical information is typically enclosed in square brackets or marked with note symbols, shown in white font, and written in either uppercase or lowercase letters, though specific stylistic choices vary across broadcasters. With regard to content, previous studies (e.g., Aleksandrowicz, 2020; Szarkowska, 2020; Khoshsaligheh et al., 2024) indicate that music can be represented in several ways, including:

- **Describing the mood, genre or instruments:** Simple labels such as [cheerful music], [jazz music] or [piano music] or a combination such as [cheerful jazz music] help DHH viewers grasp the emotional tone of a scene.
- **Naming the piece or performer:** When the music is narratively relevant, the title and composer or performer may be included, for example, [Music: Mozart 'Turkish March'].
- **Providing song lyrics:** Lyrics can be added when they contribute to plot or character development.
- **Signalling the presence of music:** Background music can be indicated simply with a musical note symbol or a brief label.
- **Omitting non-essential music:** Music information may be omitted when it is not crucial for comprehension, or when space and timing constraints limit subtitle length.

In addition to the points above, another practice, though rarely discussed in the literature, can occasionally be found in SDH practice, namely:

- **Indicating changes in the music:** Adding labels such as [music growing louder] [music approaching] or [music stops] when these details are narratively relevant.

<u>Sound effects</u>

Sound effects cover a wide range of auditory events, from environmental ambience (e.g., traffic, wind, rain) to background noises (e.g., door slams, gunshots, dog barks). Rather than

simply accompanying dialogue, these sounds can contribute to the construction of meaning in AV texts by providing contextual, narrative or emotional information.

In discussing how sound effects should be labelled in SDH, Tsaousi (2015) proposes a three-level model that considers the source of the sound, its function in the AV text, and the adequacy of its rendering for DHH audiences. In practice, both descriptive labels and onomatopoeia presented in square brackets are common strategies, such as [car horn] and [beep beep], respectively. Pereira (2010) and Gürkan (2019) show that descriptive labels tend to be more helpful than purely onomatopoeic forms for viewers born deaf, who may have limited auditory reference points. However, research with deaf children points to a different preference, as onomatopoeic forms were found to be more favourable than verbal descriptions (Lorenzo, 2010). This highlights that the choice of representation should be tailored to the intended audience. For example, in cartoons, onomatopoeia may be preferable.

Silence

As mentioned in Section 2.4.4.2, the description of silence also requires consideration in SDH. When the absence of sound carries narrative weight or precedes a sudden event, it becomes important to indicate this absence of sound for DHH audiences.

In practice, silence can be rendered in different ways depending on its narrative function. Zdenek (2015: 189) notes that the most common form is simply [silence], alongside variants such as [no audio] and [inaudible]. Zárate (2021: 82) adds further examples, including [no voice] and [no sound]. Although these distinctions may appear minor, they can be significant for DHH audiences, and appropriate labelling is essential to help them follow the scene.

2.5 Chapter summary

This chapter has centred on SDH as the main focus of this study. To set the stage for why sound needs to be visualised, it traces the role of sound in AV media from the silent film era

to the sound era. The chapter then turns to the literature to pinpoint five types of non-verbal audio information that call for representation in SDH: character identification, paralinguistic information, music, sound effects, and silence. Among these, silence deserves special mention, as it is frequently neglected in practice, although studies have acknowledged its narrative importance. Finally, it brings together and compares different strategies for rendering these elements visually. The next chapter will widen the lens to consider how SDH is practised in different parts of the world.

Chapter 3 The Provision of SDH Across the World

3.1 Chapter overview

This chapter examines SDH provision globally. Section 3.2 outlines worldwide SDH developments and introduces the three-tier grouping used for comparison: pioneering countries, established countries and early-stage countries.

Section 3.3 discusses pioneering countries (the US, the UK, Canada and Australia) where SDH was early incorporated into broadcasting policy. Section 3.4 turns to established countries (Spain, Italy, Poland, Portugal and Ireland) where SDH provision has grown through regulatory reforms. Section 3.5 considers early-stage countries in Asia (Japan and South Korea) where SDH has developed gradually over the past three decades.

Finally, Section 3.6 focuses on China, the main context of this study. It outlines the legal and policy context (Section 3.6.1), reviews current academic efforts (Section 3.6.2), and summarises SDH availability on television and streaming platforms (Section 3.6.3). A summary of the chapter is provided in Section 3.7.

3.2 Global overview of SDH development

On a global scale, the development of SDH provision has progressed at uneven rates. As sketched by Neves (2019: 84-85), "the provision of SDH on audiovisual content—and particularly on television, still seen as the preferential platform for easy open access—is progressing at three quite distinct paces."

As part of this picture, Neves (*ibid.*) identifies the US, the UK, Canada, France, Australia and Denmark as pioneering countries, where SDH has been introduced into broadcasting regulations and is now close to covering 100% of subtitled television programmes on national open-to-air channels. These countries have set specific quotas and scheduling requirements for SDH coverage, while also working to improve overall service quality. A second group comprises Italy, Spain, Portugal, Germany, Poland, Flemish-speaking Belgium,

the Netherlands and Brazil. While these countries adopted SDH later than the pioneers, they have gradually expanded their provision thanks to policy changes and rising public expectations, and for the European members, the EU's Audiovisual Media Services Directive has also contributed as a framework promoting accessibility across the continent. However, this picture is far from complete, and large parts of the world remain uncharted. In regions such as South America, Asia, Southern Africa and the Middle East, efforts to formalise SDH are still at an early stage. SDH is available only to a limited extent, and legal requirements or industry practices are either emerging or still under development.

These differences, outlined above, reflect the three-tier grouping proposed by Neves (*ibid.*). Nevertheless, this grouping should not be understood as a fixed or exhaustive classification, as SDH development does not always follow a linear path, and this global overview does not necessarily account for more recent changes in some countries. With this in mind, the following sections use this grouping only as a broad guide for examining representative examples from countries that Neves (*ibid.*) lists as pioneering, established, or early-stage contexts in SDH provision.

Before proceeding, it is worth clarifying the terminology used here. As explained in Section 2.4.2, SDH is commonly referred to as CC in the US, Canada and Australia. For the sake of consistency, this thesis uses the term SDH throughout.

3.3 Pioneering countries

3.3.1 The United States

The US has one of the longest and most influential histories of SDH development. Robson (2004: 9-16) traces its trajectory from 1940s experiments to a pivotal moment in 1980, when major broadcasters first began airing programmes with SDH on a regular basis.

Following this, progress continued with the *Television Decoder Circuitry Act of 1990*, which came into force in July 1993. This legislation mandated that all new television sets with screens 13 inches or larger, whether domestically made or imported, be equipped with

built-in decoder circuitry capable of displaying SDH (Mliczak, 2019: 113). Yet the Act's practical impact was limited, as it only required the hardware to be in place and did not regulate when SDH should be displayed or how extensively they should be provided across channels. Unsurprisingly, it did not lead to a significant rise in the availability of SDH content. A consequential step came with the *Telecommunications Act of 1996*, which introduced a concrete timetable for phasing in mandatory captioning requirements for broadcasters. The Federal Communications Commission (FCC) subsequently set out phased captioning targets that distinguished between new programming and pre-rule programming. New programming refers to programmes first broadcast after the captioning rules came into effect, while pre-rule programming covers those originally broadcast before the rules were introduced. The specific requirements for each category, as summarised by Robson (2004), are shown in **Table 3-1** and **Table 3-2**.

Hours per quarter	English-language programming	Spanish-language programming
450	1 January 2000	1 January 2001
900	1 January 2002	1 January 2004
1,350	1 January 2004	1 January 2007
100% captioning	1 January 2006	1 January 2010

Table 3-1. Captioning requirements for new programming (Robson, 2004: 40)

Captioning required	English-language programming	Spanish-language programming
30%	1 January 2003	1 January 2005
75%	1 January 2008	1 January 2012

Table 3-2. Captioning requirements for pre-rule programming (*ibid.*)

Moving beyond the realm of broadcast television, the *Twenty-First Century Communications and Video Accessibility Act* (2010) extended captioning requirements to the Internet. Specifically, it stipulated that any programme originally broadcast with SDH must retain those captions when distributed online. The Act also set new standards for hardware, mandating that all video playback devices, including set-top boxes, streaming players and mobile phones, offer user-friendly controls for displaying and customising SDH. In 2024, the FCC introduced new "readily accessible" standards to ensure that SDH display settings are easy for DHH users to locate and adjust, and compliance must be achieved no later than 17 August 2026 (FCC, 2025).

Alongside these legal requirements, the US also maintains several guideline documents intended to support SDH practice. Among these, one of the most comprehensive is the *Captioning Key*, introduced by the Described and Captioned Media Program (DCMP) in 2019. This extensive guide provides detailed recommendations on many aspects of SDH, from technical specifications (layout, reading speed, and synchronisation) to content representation (speaker identification, sound effects, music, and paralinguistic information).

3.3.2 The United Kingdom

Before introducing SDH provision in the UK, it is useful to introduce the country's shifting regulatory landscape. From 1991 to 2003, the Independent Television Commission (ITC) was responsible for licensing and regulating independent television services. In 2003, the Office of Communications (Ofcom) took over these duties and has been the UK's media regulator ever since.

In terms of SDH provision, the UK has long been ahead of other European countries, as noted by Remael (2007: 24). This early development can be partly understood in relation to the country's broader policy attention to media accessibility and the growing recognition that DHH viewers should have more equal access to television content. One important piece of legislation in this regard was the *Broadcasting Act 1990*, whose Section 35, titled

‘Subtitling for the deaf’, required public broadcasters to provide minimum amounts of SDH and meet the technical standards specified by the ITC (Zárate, 2021: 3).

These efforts were later built upon by the *Communications Act 2003*, which extended the obligation to provide SDH, in-vision signing, and AD to channels across all platforms. For digital channels, the Act also introduced a phased ten-year timeline: 60% subtitling within the first five years, rising to 80% by the end of the period (Higgs, 2006). These efforts were later built upon by the *Communications Act 2003*, which extended the obligation to provide SDH, in-vision signing, and AD to channels across all platforms. For digital channels, the Act also introduced a phased ten-year timeline: 60% subtitling within the first five years, rising to 80% by the end of the period (Higgs, 2006).

More recently, Ofcom indicates that broadcasters continue to meet these statutory obligations. In its recent access services report (Ofcom, 2025), the regulator confirms full compliance among television channels. Moreover, on-demand services, although not yet subject to mandatory quotas, also show high levels of voluntary provision, with around 95% offering SDH.

Over time, SDH guidelines in the UK have also evolved. At an early stage, the ITC issued the *ITC Guidance on Standards for Subtitling* (1999), which offered an early set of instructions for SDH production. Following this, Ofcom in 2021 updated its *Guidelines on the Provision of Television Access Services* that set out more detailed requirements for SDH and AD. Alongside these regulatory guidelines, the British Broadcasting Corporation has developed the BBC Subtitle Guidelines (2026), which are more comprehensive and detailed than those issued by Ofcom.

3.3.3 Canada

In Canada, a series of laws and regulations have been established to safeguard the needs and interests of people with sensory disabilities, the most significant of which is the *Broadcasting Act*. It has been amended several times since 1936, with the fifth revised version adopted in 1991 being the most recent amendment. In the Act, the Canadian

broadcasting system is required to ensure that accessible programming is made available to disabled audiences. The Act also grants the Canadian Radio-television and Telecommunications Commission (CRTC) the authority to regulate and supervise the country's broadcasting and telecommunications sectors.

In 2007, CRTC required English- and French-language broadcasters to provide SDH for 100% of their programming across the broadcast day, and the only exemptions were advertising and promotional material (CRTC, 2007). Several years later, this obligation was broadened to include advertising and SDH content made available online (CRTC, 2015).

With regards to SDH guidelines, the CRTC launched a public call in 2011 for ideas and comments on proposed caption quality standards for both English and French. And after several rounds of revision, the *Quality Standards for French-language Closed Captioning* and the *Quality Standards for English-language Closed Captioning* were released in 2011 and 2012, respectively. These two standards share the same fundamental requirements, although they differ slightly in the accuracy rate and the acceptable lag time between audio and captions for live programmes. Under the English standard, live captioning must achieve at least 95% accuracy with an average lag time of no more than six seconds, whereas the French standard requires a minimum accuracy of 85% and permits a lag time of up to five seconds. With respect to captioning speed, both standards state that (CRTC, 2011: 13; CRTC, 2012: 24):

> "Captioning must be verbatim representations of the audio, regardless of the age of the target audience. Speech must only be edited as a last resort, when technical limitations or time and space restrictions will not accommodate all of the spoken words at an appropriate presentation rate."

However, while the English- and French-language standards share the same general principle that captions should be verbatim representations of the audio and that editing should be used only as a last resort, only the English-language standard specifies that the "appropriate presentation rate" for children's programming is 120-130 words per minute. The French-language standard does not provide a separate numerical rate for children's

programming. Apart from this clarification, there is no significant difference between the two standards.

3.3.4 Australia

According to Carroll and Mueller (2022: 183), SDH was introduced in Australia by ABC-TV in 1983 as Teletext captions, following earlier developments in the US and UK during the 1970s. Youens (2022: 194) adds that in 1992, the Australian government passed the *Broadcasting Services Act 1992*, which requires free-to-air broadcasters to provide SDH for all programming broadcast between 6 am and midnight. However, the Act did not specify any detailed quality requirements. In response to consumer complaints and advocacy efforts, the Australian Communications and Media Authority (ACMA), Australia's media regulator, subsequently issued the *Broadcasting Services (Television Captioning) Standard 2013*. A decade later, ACMA updated the standard through the *Broadcasting Services (Television Captioning) Standard 2023*, providing a more detailed articulation of quality expectations for SDH in relation to readability, accuracy, and comprehensibility. These three aspects are defined as follows:

- **Readability** requires legible colour and font, natural line breaks, correct punctuation, and subtitles limited to a maximum of three lines.
- **Accuracy** concerns reflecting the intended meaning, considering the target audience, and accurately conveying tone, manner, sound effects and music.
- **Comprehensibility** covers a range of aspects: clear speaker identification, sufficient display time, synchronisation with speech, sound effects and music, correct spelling, explanatory captions for long pauses, and alignment with shot or scene changes.

At present, on ACMA's official captioning rules webpage (ACMA, 2025), current provisions require free-to-air commercial and national broadcasters to provide SDH for all programmes on their main channels between 6 am and midnight, as well as for all news and current affairs content at any time. Meanwhile, subscription television services are required to meet annual captioning targets depending on service category.

3.4 Established countries

3.4.1 Spain

According to Utray et al. (2009), SDH first appeared in Spain in 1990 through the Catalan public broadcaster CCRTV (Corporació Catalana de Ràdio i Televisió), the Catalan state-owned radio and television corporation. As a pioneer of SDH in Spain, CCRTV increased its subtitled output from 400 hours in 1997 to more than 5,000 hours in 2004. A few months after CCRTV introduced the service, the national broadcaster TVE (Televisión Española) also began to provide SDH. These early initiatives encouraged other Spanish television channels to adopt SDH progressively in the following years. Pereira (2005) provides an overview of how SDH provision grew across Spanish broadcasters throughout the 1990s and early 2000s. Building on this, Utray et al. (2009) bring the picture up to 2007 and show a steady increase in SDH hours on both national and regional channels.

In parallel with this progress, legislative action also began to address SDH provision. Spain's 2010 *General Law on Audiovisual Communication (Ley General de la Comunicación Audiovisual)* required that 75% of television programming be provided with SDH by the end of 2013. More recently, data from the Spanish regulatory authority, Comisión Nacional de los Mercados y la Competencia (CNMC), show that between 2021 and 2024, private broadcasters subtitled an average of 92.5% of their programming and public broadcasters 98.9%, both above this legal minimum (CNMC, 2026).

As far as national guidelines are concerned, Spain introduced its first SDH standard, *Norma UNE 153010*, in 2003 through the Spanish Association for Standardisation and Certification, AENOR (Asociación Española de Normalización y Certificación), and it was revised in 2012. In its 2012 revised version, the standard sets out criteria for timing, visual presentation, speaker identification, sound information and editorial conventions. The revised standard has been described by Arnáiz-Uzquiza (2012: 72) as the only openly accessible national stylistic standard in use for regulating SDH on Spanish television.

3.4.2 Italy

SDH in Italy was first introduced in 1986, when the feature film *Rear Window* was broadcast with teletext subtitles on the national public broadcaster, RAI (Radio Audizioni Italiane). Since then, SDH has gradually become a regular part of television programming (Morettini, 2012: 323). Throughout the 1990s, RAI played a central role in building up the service. The weekly amount of content with SDH rose from only a few hours in the early years to 43 hours by 1999, and this figure increased again by at least 10% under the next three-year (2000-2002) service agreement (De Seriis, 2006). By 2012, RAI had reached the target of 70% of its programming with SDH, and in 2020 it achieved a total of 93% (Longari, 2022: 31).

Although Neves (2019) places Italy in the established group, the actual SDH provision in the country appears to be less promising than this label suggests. As Morettini (2014) mentions, in Italy, apart from the public broadcaster RAI, there are no legal obligations for other stakeholders in the AV sector to provide accessibility services, even at a minimum level. In other words, SDH provision is still relatively scarce beyond the public service. One example given by Eugeni (2015) is the private broadcaster Mediaset, which also offers SDH, though on a much smaller scale compared with RAI. This may help explain why SDH provision in Italy is still not very developed at present. Another important reason is the country's long-standing dubbing tradition, as pointed out by Longari (2022: 21), which has made SDH less common and slowed down its use as an accessibility tool in Italy. Therefore, Italy can be seen as a case where, despite a relatively long history of SDH on public television, a fully established SDH system has yet to be developed across the broader AV sector.

3.4.3 Poland

In Poland, SDH was first introduced to the general public in 1994, when the film *Rio Grande* was broadcast on state television (Szarkowska, 2010: 144). From that point on, SDH was gradually rolled out across Poland's public television channels. Despite this, Mliczak (2015: 210) points out that the provision of SDH progressed rather slowly during the first decade

after its introduction. During this period, SDH on Polish television was scant, and the SDH provided were only available for films.

A major shift occurred in 2011, when the *National Broadcasting Council Act* was amended to introduce, for the first time, legal requirements for access services on Polish television. The revised Act, implemented in May 2011, added Article 18a, which requires broadcasters to allocate at least 10% of their annual airtime to access services, including AD, SDH and SLI (Szarkowska et al., 2015: 46; Mliczak, 2019: 122). Further regulatory measures were introduced in the years that followed. In 2018, the National Broadcasting Council (KRRiT, Krajowa Rada Radiofonii i Telewizji) raised the quota for accessible programming from 10% to 50%, with 2024 set as the final deadline. This target was to be reached through a gradual schedule: 15% in 2019, then 25% in 2020-2021, 35% in 2022-2023, and 50% in 2024 (Mliczak, 2019: 123). Recent data suggest that these targets have largely been achieved. According to the National Broadcasting Council's 2026 accessibility report (KRRiT, 2026), more than 50% of quarterly broadcasting time on Poland television now includes AD, SDH and SLI.

3.4.4 Portugal

According to Neves (2005: 275-276), the earliest provision of SDH in Portugal began in 1999, when teletext subtitles were first introduced on national television. This initiative was not driven by any legal requirement, but rather resulted from a joint effort between the Portuguese government, the public broadcaster Rádio e Televisão de Portugal (RTP), and the Portuguese Association of the Deaf (Associação Portuguesa de Surdos, APS). They together agreed on an initial target of 15 hours of SDH per week during the first year. While this early service marked an important starting point, it was still limited and fell short of meeting DHH viewers' needs. It was not until 2002 that the Portuguese Television Law formally required broadcasters to provide accessibility services, including SDH.

Following the introduction of the legal requirement, SDH in Portugal showed early promise. Even so, in the first few years, viewers began to complain about both the scheduling of subtitled programmes and the poor quality of the SDH provided (Neves, 2012: 373). In 2007,

Neves published a set of SDH guidelines (Neves, 2007b), which have since served as the official national norms in Portugal.

While Portugal has not kept pace with pioneering countries in imposing more ambitious SDH obligations, its regulatory system has developed over time. In 2025, the Portuguese Regulatory Authority for the Media, ERC (Entidade Reguladora para a Comunicação Social), approved the *Plano Plurianual de Acessibilidade das Emissões Televisivas (2026-2029)*. This plan introduces new SDH obligations in two stages. In Stage 1 (2026-2027), the public broadcaster is required to provide 30 hours of SDH per week, and in Stage 2 (2028-2029), this increases to 35 hours per week. These developments reflect that Portugal represents a case where SDH has become increasingly institutionalised, yet provision is relatively limited compared with pioneering countries.

3.4.5 Ireland

Although Ireland was not included among the established countries listed by Neves (2019), it is also worth considering here given its early development of SDH services. Much of this progress was closely tied to the efforts of its national public service broadcaster, Raidió Teilifís Éireann (RTÉ). According to RTÉ's archives (RTÉ, 1987; RTÉ, 1995; RTÉ, 2025), the broadcaster first introduced news programming specifically for deaf viewers in 1975 and then continued to expand its subtitle offerings over the following years. Teletext subtitles were introduced in the late 1980s, and SDH was subsequently rolled out from the mid-1990s.

Moving into the twenty-first century, the *Broadcasting Act 2009* set clear legal requirements for accessibility services for people with visual or hearing impairments in Irish broadcasting. To further put these requirements into practice, Ireland's Media Commission (Coimisiún na Meán) updated the *Access Rules for Television Broadcasting Services* in 2024. The rules set separate targets for each type of accessibility service (AD, SDH and SLI) over the next three years for broadcasters. For SDH, the rules require that RTÉ 1 should maintain a subtitling level of 96% between 2025 and 2027, and RTÉ 2 is expected to rise from 90% in 2025 to 92% in 2027. The final section of the rules also lays out some specific SDH

production standards, such as speaker identification, subtitling speed, line breaks, and the indication of tone of voice. A recent Irish accessibility research report by Pintado Gutiérrez (2025: 8-9) reveals that, within Ireland's accessibility services, provision for SDH and SLI has traditionally been more developed than that for AD.

3.5 Early-stage countries

3.5.1 Japan

In Japan, efforts to promote SDH began nearly three decades ago. According to Horiuchi (2022), the 1997 amendment to the Broadcasting Act already encouraged broadcasters to provide SDH and AD wherever possible. Further progress was made in 2007, when the Ministry of Internal Affairs and Communications (MIC) issued *the Guidelines for the administration of broadcasting promotions for people with visual and/or hearing disabilities* (『視聴覚障害者向け放送普及び行政の指針』). The guidelines established targets for 2008-2017 and required NHK to subtitle all programmes between 7:00 and 24:00. In 2012, the guidelines (MIC, 2012) were revised in response to the 2011 Great East Japan Earthquake. This time, the focus was on strengthening emergency broadcasting, and news programmes were expected to include captions as quickly as possible during major disasters. In 2018, the Ministry introduced a new ten-year plan (『放送分野における情報アクセシビリティに関する指針』) that required NHK and major commercial broadcasters to provide SDH for all programmes aired between 6:00 and 25:00. While the guidelines were revised again in 2023, the overall targets for SDH remained unchanged.

Beyond policy setting, it is equally important to consider how broadcasters have performed in practice. Martínez-Sirés (2016) provides an overview of the SDH rates achieved between 2008 and 2012 by Japan's major television stations, and here we present updated data published by the MIC in 2024. As shown in **Table 3-1**, NHK General and the major commercial broadcasters in Tokyo, Osaka and Nagoya provide SDH for 100% of programmes that fall within the scope of the guideline targets, and most regional stations also achieved very high levels of coverage under this metric. **Table 3-1** also presents the overall annual

SDH rate for each broadcaster, calculated across all programmes broadcast during the year, including those not legally required to carry SDH. Unsurprisingly, these annual rates are relatively lower, but they still indicate a strong commitment to SDH provision across the broadcasting sector.

Based on the above evidence, Japan shows a strong level of SDH provision and could arguably be seen as closer to an established context, even though it is not classified by Neves (2019) as a pioneering or established country. This again indicates that Neves' (ibid.) three-tier grouping is not definitive or fixed, but needs to be considered alongside specific national contexts.

【字幕放送】

	「放送分野における情報アクセシビリティに関する指針」の普及目標の対象となる放送番組における字幕番組の割合（注3）	総放送時間に占める字幕放送時間の割合
NHK（総合　放送センター）（注4）	100.0% [±0.0ポイント]	91.7% [＋1.1ポイント]
NHK（教育）	97.2% [＋1.8ポイント]	88.4% [＋2.5ポイント]
在京キー5局（注5）	100.0% [±0.0ポイント]	70.4% [＋1.1ポイント]
在阪準キー4局（注6）	100.0% [±0.0ポイント]	68.7% [＋0.2ポイント]
在名広域4局（注7）	100.0% [±0.0ポイント]	62.6% [＋0.1ポイント]
系列県域101局（注8）	89.4% [±0.0ポイント]	58.1% [＋0.5ポイント]
独立県域13局（注9）	46.5% [＋1.8ポイント]	18.2% [＋0.4ポイント]
NHK（BS/BSプレミアム）（注10）	64.6% [＋15.8ポイント] /94.8% [＋2.6ポイント]	63.6% [＋14.4ポイント] /89.2% [＋5.8ポイント]
NHK（BSプレミアム4K/BS8K）（注11）	85.0% [－6.2ポイント] /73.0% [－3.2ポイント]	88.0% [＋1.9ポイント] /82.4% [＋4.5ポイント]
民放キー局系BS事業者5社（2K放送/4K放送）（注12）	59.0% [＋1.5ポイント] /53.4% [－5.2ポイント]	30.9% [＋1.3ポイント] /32.4% [＋1.8ポイント]

Table 3-1. SDH coverage by Japanese TV broadcasters in 2023 (MIC, 2024)

3.5.2 South Korea

In South Korea, accessible services for DHH viewers began to emerge in the late 1990s. According to Song (2003: 164), MBC, one of South Korea's main broadcasters, aired its first trial SDH programme on December 1998 and launched regular SDH broadcasting on February 1999. This early move was soon followed by other major broadcasters, KBS, SBS

and EBS. Yet, at that time, to view programmes with SDH, viewers needed either a television equipped with a caption-decoding chip or an external receiver connected to a digital TV, and both options were relatively costly. In 2009, the Korea Media and Communications Commission (KMCC) introduced the *Supporting Policy for the Broadcasting Underprivileged People*, which states that "the caption broadcasting will be provided to the low-income deafs or blinds and the receiver chip for caption service is mandatory for every digital TV set by 2012" [sic] (KMCC, 2009). Several years later, the KMCC issued a notice in 2012 that set mandatory requirements for accessible broadcasting for disabled people. Under this policy, the four national broadcasters (KBS, MBC, SBS and EBS) were required to introduce SDH from 2014 (Kim & Kang, 2016: 104-105).

On the legislative front, the *Broadcasting Act* (2011), a key piece of media legislation in South Korea, obliges broadcasters to provide accessible services (AD, SDH and SLI) for viewers with disabilities. It also authorises the Korea Communications Commission, which was reorganised as the KMCC in 2025, to subsidise part or all of the associated costs through the Broadcasting Communications Development Fund. Regarding SDH standards, the KMCC released the *Guidelines for Providing Broadcasting Services for People with Disabilities* (『장애인방송 프로그램 제공 가이드라인』) in January 2026 to improve access to broadcasting content for users with hearing or visual impairments.

Overall, South Korea presents a more developed picture than the label "early-stage" might imply. Whether in the development of subsidy schemes or in the refinement of guidelines, the country shows a growing commitment to building a more accessible media environment.

3.6 China

3.6.1 Legal and policy background

At the legal level, to the best of my knowledge, there is currently no dedicated legislation governing SDH in China. Instead, accessibility for persons with hearing or visual disabilities is addressed in broader laws. For example, the *Law of the People's Republic of China on Protection of Disabled Persons* (2008) states in Article 43(3) that efforts should be made to

“offer TV programmes in sign language, set up radio programmes specially designed for persons with disabilities and add subtitles or narrations to more TV programmes and movies.” Furthermore, another relevant law, the *Barrier-free Environment Development Law of the People's Republic of China* (2023), also mandates publicly funded television stations to provide subtitles and encourages online platforms to include subtitles, SLI and AD in Article 30.

Beyond legislation, China’s media regulator, the National Radio and Television Administration (NRTA), has also promoted accessibility practices through policy documents. One example is the *Guiding Opinions on Strengthening the Construction of the Radio and Television Public Service System* (NRTA, 2020), which calls for full-time captioning on nationwide television channels and requires prime news programmes to include simultaneous SLI. Earlier, the NRTA had released the *Specification of Closed Caption System for Digital Television* (2013), but this document mainly focuses on the technical configuration of captioning systems and does not mention any broadcast quotas for SDH or requirements for SDH production.

As it stands now, China’s legal environment supports the right of DHH people to access AV content in principle, but it lacks detailed and binding requirements, not to mention clear targets for SDH provision and guidance on SDH production.

3.6.2 Academic research and effort

In China, SDH has received limited academic attention, and few scholars have dedicated themselves to the accessibility issues. One of the most representative figures in this area is Li Dongxiao from the School of Journalism and Communication at Nanjing University. In her monograph *Hear and See: Audio-Visual Media for All* (2013), Li discusses accessibility services for people with hearing and visual impairments in China. She argues that China's access service not only lags behind neighbouring countries such as Japan and South Korea, but also falls below the average level observed in Europe and the US (Li, 2013: 107). In explaining this shortfall in SDH provision, Li traces the lack of SDH provision to limitations in

technology, television production workflows and reception equipment (Li, 2013: 113-114). In addition to these technical and infrastructural barriers, Yang et al. (2022: 160) point out that the legal framework is also inadequate. Existing laws and regulations are largely generic and lack detailed implementation requirements, often using broad terms such as “strengthen” or “promote” without specifying concrete actions or timelines. In terms of SDH, current laws mention only “subtitles” in general, and subtitles specifically intended for DHH audiences have yet to enter the policy agenda.

Recognising the long-standing absence of SDH in China, Li and Looms (2015) proposed a medium- to long-term plan (2018-2025), the core of which was a regional pilot project to introduce SDH. Now that the 2018-2025 period has ended, however, the plan has seen limited progress, and the proposed pilot project has yet to materialise.

Despite the limited progress towards the proposed SDH pilot project, several research efforts have begun to appear. Yang et al. (2022) carried out a questionnaire survey and follow-up interviews with Chinese DHH participants to investigate their satisfaction with existing subtitles and their needs and expectations regarding subtitling. Building on this work, Yang et al. (2025) conducted a reception study in which DHH participants viewed videos subtitled at different presentation speeds. Their findings showed that lower presentation rates, at around 129 Chinese characters per minute, led to a more comfortable viewing experience. They concluded by calling on major media platforms to improve subtitle quality and include speaker identification and paralinguistic information. More recently, Wu et al. (2026) surveyed 228 DHH participants across China on their media use and subtitling preferences. The findings reveal a strong demand for more subtitled content, particularly verbatim subtitles, and offer practical directions for SDH production, such as speaker differentiation through subtitle placement.

In summary, a handful of researchers in China have already noted the absence of SDH and have begun to address this issue, but existing studies are few in number and mostly at an early stage.

3.6.3 SDH on Chinese television and streaming platforms

For the time being, SDH service is largely absent from both traditional television and online streaming platforms in China. Instead, what is commonly available are ST subtitles, which only include spoken dialogue and do not convey other audio information. This prevalence of ST subtitles is not a recent phenomenon. Their use can be traced back to at least 1984, when News Broadcast on China Central Television first introduced them (Yang et al., 2025). However, these ST subtitles were not originally intended to serve deaf viewers. Rather, they were designed to help hearing audiences understand programmes delivered in regional dialects (Xiao et al., 2015). For example, many mainland viewers have to rely on subtitles to follow Hong Kong films or dramas, as the Cantonese dialogue is largely unintelligible to them. This long-standing practice has made ST subtitles almost universal, and Yang et al. (2025) note that domestic television series in China now routinely carry them. More recently, ST subtitles have been provided in major live news and sports events, powered by ASR tools, but these tools are not currently used to produce SDH, and their performance sometimes can be unreliable when speakers use heavy regional dialects.

Next, we turn to China's video streaming platforms to investigate the availability of SDH on these services. The leading three platforms, namely iQIYI, Tencent Video and Youku, mainly offer professionally produced and licensed content, including dramas, films, documentaries, variety shows and sport events, supplied by established production companies. These platforms are owned by China's major internet companies, Baidu, Tencent and Alibaba, respectively. When looking at their subtitle options, we find that ST subtitles are provided for nearly all content, but SDH subtitles are still unavailable. As Li and Xiong (2017) observe, compared with the traditional media era, accessible content on online platforms in China has not progressed. Persistent issues include the lack of SDH, the limited availability of SLI and the near absence of AD on streaming sites. A positive step was taken in 2022, when Youku launched a new content section called "Barrier-Free Cinema" (Ye & Jie, 2023). Users with visual impairments can verify their accounts using their disability identification card

and gain free access to films and series with AD. However, for DHH viewers, the "Barrier-Free Cinema" section currently offers neither SDH nor SLI services.

Despite the absence of SDH on China's television and streaming platforms, some progress is being made through other forms. An accessibility company in China called Shanghai Yier Industry Co., Ltd. has been providing AD, SDH and SLI for selected films and series on its own website (https://v.yier-ea.com) in recent years. Moreover, Shanghai Television Station has produced SDH for certain films, which are occasionally screened at local special education schools. Similar small-scale initiatives can be found in other provinces, though these projects are typically sporadic and they are not part of regular broadcasting schedules.

Outside China's media environment, some international streaming platforms such as Netflix offer more mature SDH practices. Many Taiwanese films and series on Netflix include Chinese SDH in traditional Chinese characters[2]. However, since Netflix is not accessible in mainland China, most Chinese DHH viewers cannot benefit from these services. Despite this inaccessibility, Netflix has developed internal guidelines for Chinese SDH (available in both Simplified and Traditional Chinese versions): the *Chinese (Simplified) Timed Text Style Guide* and the *Chinese (Traditional) Timed Text Style Guide*, which can provide a valuable reference for the present study.

To summarise, at present, SDH is still largely absent from both traditional broadcasting and major streaming platforms in China. Although some attempts have appeared in recent years, they are limited in scope and irregular in practice. Meanwhile, overseas platforms now provide a small amount of traditional Chinese SDH, yet these services are often inaccessible to most DHH viewers in mainland China.

[2] Traditional Chinese characters are mainly used in Hong Kong, Taiwan and Macau, while Simplified Chinese characters are prevalent in Mainland China, Singapore and Malaysia. Although the two writing systems differ, the spoken language remains largely the same (Yang & Wang, 2018)

3.7 Chapter summary

After reviewing SDH provision across the countries discussed above, it becomes clear that SDH provision is indeed developing at different speeds around the world. Pioneering countries such as the UK, the US, Canada and Australia introduced SDH relatively early, largely through strong legislation and regulation. In Europe, countries including Spain, Italy, Portugal, Poland and Ireland expanded their provision at a later stage, backed by support from major broadcasters. Meanwhile, Japan and South Korea have made more recent progress within rapidly evolving media and policy environments. Compared with these countries, SDH in China is still in its infancy, and its availability is almost invisible across both broadcasting and major streaming platforms at the moment.

These differences are also reflected in how SDH guidelines are developed in different countries. By looking closely at how different contexts regulate SDH in practice, we can start to see which approaches might translate well to the Chinese context. To this end, the next chapter compares SDH guidelines from a range of countries and organisations, and the main purpose is to draw out guidance that may help inform SDH production in China.

Chapter 4 Analysing SDH Guidelines Across Countries

4.1 Chapter overview

In order to develop useful and practical suggestions that may inform future SDH production in China, this chapter reviews and compares existing SDH guidelines across different contexts.

Specifically, Section 4.2 begins by outlining the guidelines selected for analysis. Section 4.3 then lays out the four criteria used to organise the discussion, namely acceptability, legibility, readability, and relevance, and considers each in turn. Section 4.3.1 addresses acceptability, presenting conventions such as punctuation, symbols, and numbers in SDH. Section 4.3.2 turns to legibility and looks at how subtitles are visually presented, including aspects such as font, size, position, colour, number of lines, line length, and subtitle speed. Section 4.3.3 is concerned with readability, and the discussion centres on the distinction between verbatim and edited subtitles. Finally, Section 4.3.4 deals with relevance, with the aim of considering what to convey, what to omit, and what may need to be added or clarified. From there, Section 4.4 summarises the implications of these findings for SDH production in China. The chapter then closes with a summary in Section 4.5.

4.2 SDH guidelines selected for analysis

A clear set of guidelines provides a reliable baseline for SDH production. According to Gürkan (2021: 250), guidelines set standards that help ensure subtitles are readable, accurate, clear and consistent. When such guidelines are applied consistently, they can greatly improve access to AV content for viewers with hearing impairments. Without consistent guidance, SDH presentation may vary widely. Even within the same provider, if there are no uniform guidelines, different SDH subtitlers may produce SDH in their own ways, and the resulting inconsistencies could make AV materials more difficult for DHH viewers to follow and understand.

The importance of guidelines for SDH production is well recognised. At the international level, ISO/IEC 20071-23 (2018) provides detailed guidance on the visual presentation of audio information in captions and subtitles, with specific sections covering non-speech information such as sound effects, music, emotions, silence and speaker identification. However, it is also acknowledged that no single set of guidelines can "manage to apply across all contexts and scenarios" (Neves, 2019: 86). One reason, as Romero-Fresco (2015: 10) explains, is that differences in social, economic and political conditions, the AV landscape of each country, and the heterogeneity of DHH communities make it difficult to harmonise practices. Recognising this, pioneering countries in the field, such as the UK and the US, have already established standards for their own national contexts at an early stage and, over time, have been committed to refining these standards.

The situation in China is different. As mentioned in Section 3.6.1, one of the few documents related to SDH is the *Specification of Closed Caption System for Digital Television* (2013), which only covers the technical configuration of captioning systems and does not mention production requirements for SDH. The only available guidance on SDH in Chinese language comes from Netflix, where two sets of standards have been developed: *the Chinese (Simplified) Timed Text Style Guide* and *the Chinese (Traditional) Timed Text Style Guide*. But as these standards are developed for Netflix's internal use rather than for national broadcasting contexts, they cannot be regarded as official guidance. Thus, strictly speaking, it can be said that there is still no official, publicly available guideline for SDH production in China.

Given the absence of official SDH guidelines in China, it may be useful to consult existing guidelines from countries with a longer history of SDH provision and more established standards. This does not mean that these guidelines can be directly applied to the Chinese context. Rather, they are used here to identify common principles and practical considerations that may inform future SDH production in China. For this reason, this chapter analyses a number of guidelines mainly from the pioneering countries introduced in Section 3.3. The list of guidelines examined in this chapter is presented below:

- **US:** *Captioning Key* (2019) by DCMP (hereafter DCMP Guidelines)
- **UK:** *Subtitle Guidelines* (2026) by BBC (hereafter BBC Guidelines)
- **Canada:** *Closed Captioning Standards and Protocol for Canadian English Language Television Programming Services* (2012) by CAB (hereafter CAB Guidelines)
- **Australia:** *Broadcasting Services (Television Captioning) Standard 2023* (2023) by ACMA (hereafter ACMA Guidelines)
- **Netflix:** *Simplified Chinese Timed Text Style Guide* (2025) (hereafter Netflix Chinese Style Guide)

It should be noted that the final guideline in the list, the Netflix Chinese Style Guide, differs from the other guidelines selected for analysis, as it is an internal style guide developed for a specific platform rather than an official national SDH standard. That said, as a major streaming platform, Netflix has gained a certain amount of practical experience in subtitle production. More importantly, it is the only selected guideline that specifically addresses Chinese-language subtitles, making it a useful reference for considering parameters related to the Chinese language in SDH production, such as subtitle speed, font size and the number of characters per line. Given these points, it is included in the set of guidelines examined in this chapter.

4.3 Criteria in SDH guidelines

When reviewing SDH guidelines developed in different countries, it is evident that their structure, focus, features and level of detail often vary. Nevertheless, in essence, they all share a set of underlying principles or criteria that are intended to make subtitles more accessible for DHH audiences.

To facilitate the analysis, this section draws on the criteria proposed by Gambier. In his discussion of accessibility in AVT, Gambier (2003: 179) outlines several criteria that are relevant across different AVT modes, namely acceptability, legibility, readability, synchronicity, relevance and domestication strategies. Building on Gambier (2003), Neves (2005: 121) emphasises that these criteria are particularly relevant to SDH. In more recent

work, Gambier (2026: 25) revisits these criteria and screens four as relevant to subtitling. These criteria are outlined below:

- **Acceptability**, related to language norms, stylistic choice, rhetorical patterns and terminology.
- **Legibility**, defined in terms of font, position of the subtitles and subtitle rate.
- **Readability**, also defined for subtitling in terms of reading rates, reading habits, text complexity, semantic load, shot changes and speech rates.
- **Relevance**, in terms of what information is to be conveyed, deleted, added or clarified in order not to increase the cognitive effort involved in listening or reading.

Accordingly, the following sections will revolve around these four criteria to analyse the selected SDH standards. More specifically, acceptability will be used to examine how the guidelines address language norms; legibility will be related to visual presentation, including font, colour, positioning and subtitle speed; readability will be considered in relation to the use of verbatim or edited subtitles; and relevance will guide the discussion of how the guidelines determine what information should be included, reduced or clarified for DHH viewers. Through this structure, the analysis uses Gambier's criteria (2026) to examine the practical recommendations set out in existing SDH guidelines, thereby exploring their potential implications for Chinese SDH design.

4.3.1 Acceptability

Acceptability, according to Gambier (2026: 25), concerns language norms, stylistic choices, rhetorical patterns and terminology. A similar understanding appears in Pedersen's (2017) FAR model (functional equivalence, acceptability and readability), where acceptability refers to "how well the subtitles adhere to target language norms" (Pedersen, 2017: 217). In the context of SDH guidelines, acceptability can be understood as the use of accepted written conventions. Based on this understanding, this section focuses on punctuation, symbols and numbers, as these details illustrate how different guidelines regulate the

written and stylistic presentation of subtitles. The following subsections address these three aspects in turn.

4.3.1.1 Punctuation

Punctuation is an integral part of written language, and this is equally true for subtitles. A study by Datta et al. (2020) indicates that the use of punctuation in subtitles can improve the viewing experience for a wide range of television audiences, including deaf, hard of hearing, and hearing people. The findings also suggest that subtitles with punctuation are easier for viewers to follow. In addition to its basic semantic function, punctuation can also "convey some of the aspects peculiar to spoken language in the written mode" (Gürkan, 2019: 165). By an example, the use of an exclamation mark after an interjection can signal strong emotion or emphasis.

As expected, the selected SDH guidelines also provide recommendations on the use of punctuation in subtitles. These typically include brackets, ellipses, as well as dash and hyphens, which are addressed below.

<u>Round brackets, square brackets and Chinese brackets</u>

Brackets, including round and square brackets, are very common in SDH and are usually used to enclose speaker tags, sound effects and paralinguistic information (Szarkowska, 2020: 253; Talaván, 2020: 572). When looking across different guidelines, one can see that they often adopt different rules regarding the use of brackets.

Firstly, the BBC Guidelines (2026) specify the use of round brackets to deliver a sarcastic statement or question marked with an exclamation mark or a question mark, as shown in Examples 1 and 2. The guidelines also note that, when descriptive or interpretive information would take up too much time or space, round brackets may be used to indicate whispered speech, as illustrated in Example 3. Additionally, the BBC Guidelines (2026) state that brackets can also be used to indicate asides, which may or may not be whispered.

Turning to the DCMP Guidelines (2019), the document provides recommendations not only on the use of round brackets but also on square brackets. On the one hand, round brackets are used to indicate speaker identification. On the other hand, square brackets are used to describe the speaker's emotions or intonation, sound effects, background or instrumental music, lyrics, as well as the artist's name and the title, and even silence. Interestingly, DCMP also recommends explaining puns to viewers and presenting them in square brackets. Example 4 provides an illustration of this. In the other two national guidelines (CAB Guidelines, 2012; ACMA Guidelines, 2023), no detailed recommendations regarding the use of brackets were found.

The discussion above is based on national guidelines from English-speaking countries. The Netflix Chinese Style Guide (2025) introduces a particular type of Chinese bracket used to indicate albums, films, songs, TV programmes, newspapers, magazines, books, works of art and trilogies. An example of this usage is shown in Example 5.

For the use of round and square brackets, these rules can serve as a reference, as long as consistency and clarity are maintained. More importantly, Chinese brackets should not be omitted when they appear.

Example 1	It's interesting (!)
Example 2	Oh, don't you ever think about other people (?)
Example 3	(She's sleeping, let's get out of there.)
Example 4	Why do they call her "Ouisy"? ["Wheezy"]
Example 5	你看过《哈利波特》的电影吗? *English translation:* Have you watched *Harry Potter* movies?

Ellipses

As for the use of ellipses, the BBC Guidelines (2026) suggest that three dots are used to indicate pauses, hesitations, time passing and incomplete sentences, and may also represent muted or partially censored words, as shown in Example 6. Two dots may be placed at the beginning of the second line of subtitles to indicate continuation, as shown in Example 7.

In the DCMP Guidelines (2019), ellipses are mainly used in two situations: 1) to indicate a significant pause within a caption; and 2) to lead into or out of audio related to an onscreen graphic. However, in terms of continuation across captions, DCMP does not recommend using ellipses to indicate that a sentence continues into the next caption. A similar rule can be found in the Netflix Chinese Style Guide (2025), which advises against using ellipses when a sentence continues across subtitles. Instead, ellipses are only used when there is a pause (two seconds or more), an abrupt interruption occurs, or when a subtitle begins mid-sentence. In a word, regardless of the type of ellipses used, maintaining consistency in their use is important.

Example 6	Keep your f... ing nose out of it!
Example 7	I'd like... .. a cup of coffee.

Dashes and hyphens

In SDH, speaker changes are typically marked by dashes or hyphens (Zárate, 2021: 50), though the preferred symbol varies across guidelines. The BBC Guidelines (2026) use dashes, while the ACMA Guidelines (2023) opt for hyphens for this purpose. Beyond this shared function, both are also used for additional purposes. The BBC Guidelines (2026) specify the use of two dashes to mark the correction of a misspelled word during real-time subtitling, as shown in Example 8. Hyphens, in contrast, are used in a wider range of situations. Both the BBC Guidelines (2026) and the DCMP Guidelines (2019) use a single hyphen to indicate

stuttering, and the latter also uses hyphens to separate letters in spelled-out words. Although less common, three hyphens are used in specific contexts, such as the CAB Guidelines' (2012) example of using three hyphens to edit swearing (Example 9).

Concerning the use of dashes or hyphens in Chinese subtitles, the Netflix Chinese Style Guide (2025) addresses only the use of hyphens, which are used to indicate two speakers within a single subtitle. No specific guidance is provided on the use of dashes. Given that dashes and hyphens are rarely used in Chinese subtitles, it may be preferable to follow the Netflix Chinese Style Guide (2025) and use them only to indicate different speakers.

Example 8	The minster’s shrew is unchanged -- view.
Example 9	You f--- idiot!

4.3.1.2 Symbols

Compared with written words, symbols can convey certain types of information more vividly while also saving valuable subtitle space. For this reason, they are sometimes used in SDH to indicate elements such as paralanguage, sound effects, and music. The most commonly used symbols in SDH are musical notes, typically employed to signal the presence of music. In addition, some studies (Neves, 2005; Tamayo, 2015; Oomori et al., 2020) have discussed the potential use of more creative forms, such as emoticons and emojis, in SDH. Accordingly, the following discussion looks at how these two categories of symbols are represented in SDH.

Musical notation

Music in film scores can be essential in creating atmosphere, setting the mood, and conveying significance. The representation of music in SDH was introduced in Section 2.4.4.3, where the use of musical note symbols was briefly mentioned, though it was not explored in detail. Here, this issue is discussed further.

In SDH, musical notes mainly serve two functions. First, they can indicate the presence of music. Second, when the lyrics of a song are subtitled, musical notes are used to indicate that the subtitles are being sung, rather than spoken.

When looking at the practices for using musical notation, differences can be found across selected guidelines. In the US, the DCMP Guidelines (2019) recommend captioning lyrics with music icons (♪). Specifically, one quaver should be placed at the beginning and end of each caption within a song, while two quavers are used at the end of the last line of the song. An example of the latter is shown in Example 10. Similarly, in Canada, the CAB Guidelines (2012) also use quavers. A slight difference is that quavers are placed at the beginning and end of each lyrical phrase rather than each new line of text (Example 11). In the UK, a different symbol is used. The BBC Guidelines (2026) give priority to a hash mark (#) inserted at the beginning and end of the lyrics (Example 12).

Example 10	♪ The boys are back in town ♪ ♪
Example 11	♪ The rolling of the sea Is so blue ♪ ♪ I miss the sea And I miss you ♪
Example 12	# These foolish things remind me of you #

In the context of Chinese language SDH, the Netflix Chinese Style Guide (2025) also recognises the importance of using musical notes to indicate lyrics. It specifies that lyrics should be marked with a musical note (♪) at the beginning and end of each subtitle. Overall, although practices vary across guidelines, musical notes is a common and useful way to represent music in SDH.

Emoticons and emojis

Before discussing the use of emoticons and emojis in SDH, it is necessary to clarify the difference between the two. As Duraj and Szarkowska (2025) explain, emoticons are combinations of punctuation marks that visually represent facial expressions, such as :-) to indicate a smile. Emojis, on the other hand, are graphical or digital pictograms that represent not only facial expressions, but also a wide range of objects and actions.

In SDH studies, the use of emoticons and emojis has already been discussed in previous research (cf. Neves, 2005, 2009; Tamayo, 2015, 2017a; Sala Robert, 2016, 2022). One of the advantages of using these graphic forms is that they may be more efficient and effective in expressing paralinguistic and emotional information. This point is underscored by Neves (2009: 161), who explains that "sound effects (such as 'phone rings') are often straightforward to describe, but expressing what comes with the tone of voice (irony, sadness, irritation) can be difficult." In such cases, the use of graphic forms with facial expressions tends to be more effective in conveying emotions. This positive potential is further supported by Neves (2009), who shows that deaf viewers in Portugal responded favourably to the use of smileys in a prime-time soap opera. Similarly, Tamayo (2015) also supports that the presence of emoticons was welcomed by many deaf viewers in her study.

Despite their unique advantages, the use of emoticons and emojis in SDH is not without limitations. Arnáiz-Uzquiza (2015) indicates that individuals with hearing impairments may have difficulty processing emoticons, as they may take longer to identify them on screen, possibly because they appear unfamiliar or not fully integrated into the subtitles they are accustomed to. Another limitation is that emoticons or emojis may be interpreted differently across countries or communities. Neves (2005: 226-227) observes that Portuguese deaf viewers interpret the emoticon :-o as "surprise", rather than "shouting or yelling" as perceived in the Spanish context. Even within the same cultural context, the interpretation of emoticons and emojis can vary considerably. A case in point is a common smiley emoji on WeChat, a popular Chinese messaging application. It was originally

intended to convey a genuine smile, but is now often interpreted by younger users as a "mysterious smile" carrying negative connotations such as sarcasm or speechlessness. Older users, instead, tend to treat it simply as a sign of friendliness (Zhou et al., 2017: 754). Lastly, practical limitations also exist, as current subtitling software may not always support the inclusion of symbols or icons (Zárate, 2021: 99).

Perhaps due to these concerns, it is unsurprising that none of the selected guidelines address the use of emoticons or emojis. In line with this, their use in Chinese language SDH is not recommended at present.

4.3.1.3 Numbers

Almost all the English-language guidelines we looked at touch on numbers in some way. Across these documents, recommendations can be found on how numerical information should be represented, and the discussion largely revolves around when numbers should be spelled out and when numerals should be used. A shared suggestion among several guidelines (CAB Guidelines, 2012; DCMP Guidelines, 2019; BBC Guidelines, 2026) is to spell out numbers below ten and use numerals for numbers ten and above.

When it comes to Chinese subtitles, the representation of numbers also needs to be considered in order to maintain consistency. In Chinese, numbers generally appear in one of two forms: either Arabic numerals (1, 2, 3) or Chinese characters (一, 二, 三). Among the guidelines selected for analysis, only the Netflix Chinese Style Guide (2025) provides specific guidance on how numbers should be represented in Chinese subtitles. Therefore, the discussion of numerical representation in Chinese subtitles in this section draws on the recommendations provided in this guide.

According to the Netflix Chinese Style Guide (2025), these two forms are used in different contexts. On the one hand, Arabic numerals (1,2,3) are typically used in contexts involving technical or reference information. Specifically, Arabic numerals are recommended for apartment numbers, street numbers, phone numbers, measurements, chapter and page numbers, addresses, dates and flight numbers. They are also used when expressing specific

times, such as 4 点 12 分. On the other hand, Chinese characters (一, 二, 三) are generally preferred for smaller numbers and certain types of expressions that are more naturally written in Chinese. For example, numbers from one to ten should be written out in Chinese characters where space permits. Chinese characters are also preferred when expressing rounded times (e.g., 四点, *four o'clock*) and when referring to days of the week (e.g., 星期二, *Tuesday*). In addition, when representing rounded large numbers, the guide suggests using common Chinese characters rather than financial numerals. Regardless of the form used, the guide advises against mixing Arabic numerals with Chinese number characters within the same expression, as this may affect subtitle clarity.

Overall, the Netflix Chinese Style Guide (2025) provides clear and practical recommendations regarding the representation of numbers in Chinese subtitles. Given the clarity of these rules, they can be directly adopted as a useful reference for the production of Chinese SDH.

4.3.2 Legibility

When discussing legibility, it may be confused with readability, as the two terms are, more often than not, used interchangeably. Gambier (2026: 25) provides a brief definition of the two concepts, describing legibility in terms of "font, position of the subtitles and subtitle rate," whereas readability is more related to "reading rates, reading habits, text complexity, semantic load, shot changes and speech rates."

Similarly, Sanders and McCormick (1993: 102) point out that legibility relates to the visual form of characters and how easily individual characters can be distinguished, depending on factors such as font type, size, and stroke width. In comparison, readability concerns the comprehension of textual information, including writing style, sentence structure, and overall text coherence. This understanding of readability is also shared by Robert and Remael (2016: 583), who describe it as the "ease of comprehension and coherence between individual subtitles."

Based on the discussion above, legibility can be understood as relating to the visual presentation of subtitles, such as fonts, sizes, colours, subtitle position, the number and length of lines, line breaks, and subtitling speed, whereas readability relates more to the linguistic content of subtitles. The following sections discuss these aspects of legibility in turn.

4.3.2.1 Font

In subtitling norms, font selection mainly concerns the choice between serif and sans-serif typefaces. The key difference between the two lies in the presence of decorative strokes at the ends of characters: serif fonts feature small decorative lines or tapers at the ends of strokes, while sans-serif fonts do not include such flourishes (see **Figure 4-1**). As Ivarsson and Carroll (1998: 42) note, "embellishments like serifs might make the type more attractive and legible on paper, but tend to impair legibility on screen." As a result, sans-serif fonts are often preferred in subtitling. This preference is also reflected in industry practice, where television and cinema distributors routinely opt for sans-serif fonts such as Arial, Verdana or Helvetica in most Western languages to ensure legibility and visual clarity (Díaz-Cintas & Remael, 2021: 96).



Figure 4-1. Serif and sans-serif fonts

A similar preference for sans-serif fonts is evident in the SDH guidelines analysed in this study, including both English- and Chinese-language guidelines. In the English-language guidelines, both the DCMP Guidelines (2019) and the BBC Guidelines (2026) recommend

the use of sans-serif fonts. Moreover, the BBC Guidelines (2026) go a step further by specifying font styles and recommending typefaces such as Reith Sans, Verdana and Tiresias.

In Chinese-language guidelines, the Netflix Chinese Style Guide (2025) also provides specific recommendations for Chinese-language subtitling. It prescribes the use of SimHei, which is a sans-serif typeface. This choice is corroborated by previous research. Kuo (2014: 90) identifies SimSun and SimHei as among the most widely used fonts for Chinese. In line with this, font recommendations summarised by Díaz-Cintas and Remael (2021: 96), also point to SimSun and SimHei as preferred choices based on industry practices in Wincaps Q4 and OOONA. Thus, for Chinese subtitles, it is recommended to use SimSun or SimHei as the preferred fonts. But if these fonts are not available, other sans-serif typefaces should be used to support legibility.

4.3.2.2 Size

When creating or editing subtitles, in addition to selecting an appropriate font, it is also essential to choose a suitable font size. If the font size is too small, audience may struggle to read the subtitles even at a short viewing distance, but an excessively large size can obscure parts of the image and detract from the viewing experience.

On the topic of font size, Casas-Tost and Rovira-Esteva (2018: 41) propose a general range of 32 to 42 points for Chinese characters. Díaz-Cintas and Remael (2021: 96) provide a more precise recommendation, suggesting 35 points specifically for SimHei or SimSun, two commonly used Chinese fonts, when using Wincaps Q4 and OOONA.

However, in practice, the final displayed size of subtitles depends on various factors, including device type, video platform, and viewing distance. Given this variability, some guidelines avoid prescribing a fixed font size. The Netflix Chinese Style Guide (2025), for example, does not specify an exact font size, but instead requires that subtitle size be proportionate to the video resolution and allow approximately 16 Chinese characters per line. Comparable approaches are also found in the English-language guidelines examined. The BBC Guidelines (2026), for instance, focus on relative screen proportions instead,

recommending subtitle heights of around 7% - 8% of the active video height for horizontal formats, and 3.9% - 4.5% for vertical formats. Yet, other guidelines selected in this study, including the CAB Guidelines (2012), ACMA Guidelines (2023) and DCMP Guidelines (2019), do not provide specific font sizes, nor do they offer general recommendations on subtitle size.

Given the many variables involved, even though specific font sizes for Chinese have been suggested, they should not be applied rigidly. Therefore, it is suggested that subtitle size be adjusted flexibly in response to contextual factors. At the same time, maintaining legibility without obscuring the image should be the primary principle.

4.3.2.3 Position

Compared with aspects such as font type and size, subtitle positioning tends to be less controversial. In most cases, subtitles are placed at the lower centre of the screen to avoid obstructing the image, as this area is usually less relevant to the main action on the screen (Díaz-Cintas & Remael, 2021: 94). That said, there are situations where this default positioning may not be ideal. For example, in certain types of programmes, such as sports events like snooker, where most of the activity is concentrated in the lower part of the screen, subtitles are more appropriately placed at the top in such cases, as suggested in the BBC Guidelines (2026). The Netflix Chinese Style Guide (2025) also allows repositioning when bottom placement obscures important visual content.

When it comes to SDH in particular, there are also cases where subtitles are not positioned at the bottom of the screen. First, the DCMP Guidelines (2019) recommend positioning subtitles beneath the corresponding speaker on screen when multiple speakers talk simultaneously. Second, on Spanish television channels, there is a common practice of placing dialogue at the bottom of the screen, and sound descriptions appear in the top right corner. To better understand audience preferences in this regard, Bartoll and Martínez Tejerina (2010) carried out a reception study in Spain with hearing, prelingually deaf, and postlingually deaf participants, comparing three subtitle positions: bottom, mixed (dialogue

at the bottom with sound descriptions in the top-right corner), and top. The findings show that the mixed position was considered the most legible across all three groups. Interestingly, when the same three positions were tested with British viewers, the findings indicate a clear preference for the bottom position (Romero-Fresco, 2010). This difference may be related to how familiar audiences are with different subtitling positions, as well as the viewing habits they have developed over time.

While these two alternative positioning practices may work well in certain SDH contexts, they can disrupt spatial coherence, and frequent shifts in position may affect the overall viewing experience (Kuo, 2018: 423; Gürkan, 2019: 143). As Díaz-Cintas and Remael (2007: 83) point out, "viewers expect the subtitles to appear at the bottom of the screen." This expectation can also be seen among Chinese audiences, who are used to seeing subtitles at the bottom of the screen, because this has long been the dominant practice in Chinese television programmes. In this sense, it may be more appropriate to keep SDH subtitles at the bottom of the screen. For speaker identification, colour can be used or combined with name tags when needed (see Section 4.3.2.4). As for sound description, it can be included within the subtitle area alongside dialogue at the bottom, rather than being placed elsewhere on the screen.

4.3.2.4 Colour

As discussed in relation to speaker identification in Section 2.4.4.3, colour coding provides an effective means for DHH audiences to distinguish between different speakers on screen. This view is supported by existing reception studies (e.g., Arnáiz-Uzquiza, 2015; Szarkowska et al., 2015; Cepak & Corrius, 2023). The use of colour for speaker identification has also been adopted in practice, as seen in its use on television channels in Spain and Poland (Pereira, 2010; Szarkowska et al., 2015).

Among the guidelines selected for analysis, both the ACMA Guidelines (2023) and BBC Guidelines (2026) support the use of colour to distinguish different speakers. By contrast, the Netflix Chinese Style Guide (2025) recommends adding speaker labels in brackets when

the speaker cannot be visually identified. A different approach is taken by the CAB Guidelines (2012) and the DCMP Guidelines (2019), which favour white text on a black background and rely on subtitle placement to indicate speakers when necessary. It is worth noting that, as discussed in Section 2.4.4.3, both labelling and placement have inherent limitations. Labelling can increase subtitle length, and placement is constrained by software capabilities. Given that, colour coding appears to be a more effective strategy in this context.

With regard to the specific use of colour, the BBC Guidelines (2026) specify four colours for speaker identification, namely white, yellow, cyan and green, arranged in order of priority. These colours are to be used against a black background to ensure sufficient contrast. In cases where the number of characters exceeds four, the guidelines suggest reserving one colour as a "floater" for minor characters, along with other strategies such as adding labels to avoid ambiguity in speaker identification.

On the basis of the foregoing, the use of colour for speaker identification is recommended, provided that each character is consistently assigned the same colour throughout a programme. When needed, colour coding may be supplemented by additional strategies, such as name tags or subtitle placement.

4.3.2.5 Number of lines

Unlike ST subtitles, which contain only dialogue, SDH subtitles include additional information that may require more lines. As a result, Díaz-Cintas and Remael (2021: 93) note that SDH subtitles in some countries may extend to three or even four lines.

In this regard, there is no uniform limit across guidelines on the maximum number of lines in SDH. For example, the ACMA Guidelines (2023) recommend no more than three lines. The CAB Guidelines (2012) adopt a strict approach that limits subtitles to two lines as the standard but permits three lines only in exceptional cases, such as to avoid obscuring important visual content or to aid speaker identification. For similar reasons, the DCMP Guidelines (2019) acknowledge that it is acceptable for subtitles to extend to three or even four lines on occasion. The BBC Guidelines (2026) set different line limits depending on the

format: a maximum of two lines for standard horizontal formats, and up to three lines for vertical video. Nevertheless, it should be noted that all of these standards concern SDH in English. In the Chinese context, the Netflix Chinese Style Guide (2025) adopts a more flexible approach, as it limits dialogue to two lines but allows a third line for sound description. Considering that subtitles with three lines are also used in China (Casas-Tost & Rovira-Esteva, 2018: 38; Kuo, 2018: 425), they are likely to fall within an acceptable range for viewers. Therefore, three lines can be considered acceptable in Chinese SDH.

4.3.2.6 Line length

Alongside the number of lines, line length is another important factor that affects legibility. According to Díaz-Cintas and Remael (2007: 24), for Latin-based languages, line length may range from as few as 28 characters to around 37 characters per line. In their later work, they further note that this upper limit may increase slightly to 37-39 characters per line (Díaz-Cintas & Remael, 2021: 98). But these are general recommendations, and individual guidelines may adopt slightly different limits. Among the selected English language guidelines, only the CAB Guidelines (2012) and the BBC Guidelines (2026) specify exact character limits for line length. The former restricts captions to 32 characters per line, while the latter sets the maximum line length at 37 characters.

For Chinese language, Díaz-Cintas and Remael (2021: 98) state that the limit is usually set at around 16 characters per line. In the Netflix Chinese Style Guide (2025), a higher limit is proposed, with 18 characters per line. Some studies further suggest that an even higher limit may be possible in Chinese subtitles, as Kuo (2018: 425) mentions a maximum of up to 20 characters per line.

However, the number of characters that can be accommodated on a line is not fixed, as it is also influenced by font type (Casas-Tost & Rovira-Esteva, 2018). Regarding this, Kuo (2018: 426) compares monospaced and proportional fonts in Chinese and shows that, although Chinese characters themselves occupy a fixed square space, differences between font types may affect the overall display. Taking this into account, in practice, when proportional fonts

such as SimSun and SimHei are used, Netflix's recommendation of 18 characters per line can be considered appropriate, given that it is based on the use of SimHei, a proportional font.

4.3.2.7 Subtitle speed

Subtitle speed, sometimes also called "presentation rate", is usually measured in characters per second (cps) or words per minute (wpm) (Szarkowska & Gerber-Morón, 2018a). To measure this, one simply divides the number of characters by the time the subtitle stays on screen for cps or calculates the number of words per minute based on the on-screen duration (Fresno & Sepielak, 2022: 415). For example, a subtitle with 20 characters or 4 words shown for 2 seconds would result in a speed of 10 cps or 120 wpm. Since cps is less affected by differences in word length across languages, researchers prefer to use it, particularly when comparing the effects of different subtitling speeds on viewers (cf. Szarkowska & Gerber-Morón, 2018a). However, wpm is more commonly used in guidelines, as most subtitling tools calculate word rates and flag high wpm values, a point mentioned in the BBC Guidelines (2026).

In discussions of subtitle speed, the well-known "six-second rule" is often cited. This rule means that a full two-line subtitle should be displayed for six seconds, so that the average viewer has enough time to read it comfortably. Using the two measures mentioned above, "six-second rule" roughly corresponds to 12 cps or 140-150 wpm (Díaz-Cintas, 2008: 97). With this as a reference point, we now look at how subtitling speed is regulated across different SDH guidelines.

To begin with, the DCMP Guidelines (2019) outline three levels of subtitle speed: 130 wpm, 140 wpm, and a maximum of 160 wpm. The BBC Guidelines (2026) recommend a speed of 160-180 wpm. Under the CAB Guidelines (2012), subtitle speed is capped at 200 wpm, a rate considered acceptable for most adult viewers; for children's programmes, a slower pace of 120-130 wpm is advised. As for the ACMA Guidelines (2023), they do not specify a

minimum subtitle duration, but state that captions should remain on screen long enough for viewers to read them and follow the programme.

Concerning Chinese language, the Netflix Chinese Style Guide (2025) recommends a maximum subtitle speed of 11 cps for adult programmes and 9 cps for children's programmes. It is worth noting that these values are intended as upper limits for reference. As Díaz-Cintas (2020: 158) points out, subtitle speed for Chinese generally falls around 5 cps, while Kuo (2018: 419) indicates an even lower rate of about 4 cps. However, at present, there is limited empirical evidence to support these values, so it may be more appropriate to set a reasonable range by taking both existing guidelines and scholarly perspectives into consideration. By doing so, subtitle speed for Chinese can be kept around 4-5 cps, and it is better to avoid higher values approaching the upper limits of 9-11 cps.

4.3.3 Readability

In Section 4.3.2, a distinction has already been made between legibility and readability. Legibility relates to the visual form of subtitles, while readability is more concerned with their content. Among the factors associated with readability identified by Gambier (2026: 25), only text complexity and semantic load are directly relevant to subtitle content. This naturally raises the question: should subtitle content be edited and simplified to aid reading, or presented verbatim to preserve the integrity of the original speech? In the literature, this is often framed as the distinction between verbatim and edited subtitles, which is discussed further in the following section.

4.3.3.1 Verbatim or edited subtitles

Verbatim subtitles are essentially unreduced subtitles, where the dialogue is reproduced as fully and faithfully as possible. Edited subtitles, on the other hand, involve some degree of reduction, as the content is condensed and simplified to make it easier to read (Szarkowska, et al., 2011).

For DHH viewers, one might assume that edited subtitles are preferable, as edited subtitles are usually slower than verbatim subtitles and easier to follow. To explore this, Burnham et al. (2008) investigated how subtitle speed and text reduction influence comprehension among DHH viewers. The results show that comprehension was closely linked to reading ability. In other words, viewers with stronger reading skills were able to handle faster subtitles and more condensed text, and those with lower reading skills might struggle under the same conditions. In a related study, Szarkowska et al. (2011) looked at how different subtitle types are processed by tracking viewers' eye movements. They found that edited subtitles were the easiest to process, while verbatim subtitles were more demanding but led to higher comprehension. More recently, Yang et al. (2025) conducted a reception study on SDH within the Chinese DHH community. In the experimental part of the study, one of the hypotheses was that edited subtitles help comprehension more than verbatim subtitles. The results supported this assumption that, compared with verbatim subtitles, edited subtitles were found to be more helpful for DHH viewers in understanding the content, particularly for those with more limited educational backgrounds. In addition to the experimental findings, further evidence comes from the questionnaire data, which revealed a clear preference for edited subtitles, with as many as 66.13% of DHH respondents choosing them over verbatim subtitles (*ibid.*).

The findings from Yang et al. (2025) point to the advantages of edited subtitles and even suggest a preference for their use, yet this is not always mirrored in existing guidelines. Across the selected guidelines, positions can be broadly grouped into two main categories. On the one hand, many guidelines lean towards preserving fidelity to the original speech and treat verbatim subtitles as the default. This can be seen in the CAB Guidelines (2012), where verbatim subtitles are given priority and editing is introduced only as a last resort when reading becomes difficult. A similar view appears in the BBC Guidelines (2026), which encourage subtitles to remain as close to the original speech as possible and advise against unnecessary intervention. The Netflix Chinese Style Guide (2025) also advocates keeping the original dialogue and discourages any simplification or dilution of content. On the other

hand, a more flexible approach can be found in other guidelines. The DCMP Guidelines (2019) allow for edited subtitles in order to support readability and comprehension, though captions must be verbatim in specific cases such as lyrics, poems and quotations.

Regarding the debate over whether verbatim or edited subtitles are preferable, Romero-Fresco (2015) argues that the debate itself may be misplaced. Both approaches have limitations: verbatim subtitles are often too fast to ensure full access, while edited subtitles, though easier to process, may come at the expense of completeness. Rather than treating them as mutually exclusive alternatives, he proposes viewing them as complementary strategies that can be adapted to the visual context and the cognitive demands of a given scene. In this sense, he offers a potential way forward. Specifically, edited subtitles may be better suited to moments where new or visually dense information appears on screen, allowing viewers to share their attention between text and image. In other situations where the visual content is already familiar, verbatim subtitles may be more appropriate, as less attention needs to be devoted to the images (Romero-Fresco, 2015: 339).

Essentially, the question of whether subtitles should be verbatim or edited is largely a matter of speed. As discussed in Section 4.3.2.7, subtitle speed for Chinese is best kept below an upper limit of around 11 cps. Taking this upper limit into account, when the threshold is exceeded and the visual content is dense and unfamiliar to the audience, it may make more sense to introduce some degree of editing to ease processing. Conversely, when the pace is more manageable, there is less need for reduction, and the original dialogue may be more fully preserved. Nevertheless, these observations are far more suggestive than conclusive, and further empirical research on larger and more diverse samples is needed to substantiate them.

4.3.4 Relevance

Following the criteria outlined by Gambier (2026: 25), the final aspect to be considered is relevance, which concerns decisions regarding what information should be conveyed, omitted, added, or clarified so as to avoid imposing unnecessary cognitive demands on

viewers. Among the selected guidelines, the CAB Guidelines (2012), as the only set of standards to explicitly address relevance, propose a "hierarchy of relevancy" that categorises information into three levels:

- **Primary:** information that is crucial for understanding the plot, such as a ringing phone interrupting a conversation or a door slamming that signals a character's sudden departure;
- **Secondary:** information that is supportive but not essential, such as indicating whether "thanks" is expressed with affection, sarcasm, or indifference, or describing a character's tone of voice when it is not visually evident;
- **Tertiary:** optional details included where time and space permit, such as incidental sounds like distant traffic noise, birds chirping, or background chatter that does not directly contribute to the story.

Whether in Gambier (2026) or in the CAB Guidelines' three-level hierarchy (2012), relevance is conceived in operational terms from the angle of subtitlers, primarily concerning how information is selected and conveyed to viewers. Nevertheless, this is not the only way in which relevance can be understood. For scholars working with relevance theory, the concept is approached more through the question of how subtitlers guide viewers in interpreting meaning (cf. Sasamoto, 2024). This difference suggests that relevance, as used in AVT practice, does not always correspond to how relevance is understood in relevance theory. In fact, this kind of disconnect between AVT research and relevance theory in their understanding of relevance has already been noted in the literature. Some relevance theorists have described the application of relevance theory in AVT as partial, fragmented and, at times, even misunderstood (Braun, 2018; Bogucki, 2020; Sasamoto, 2024). Moreover, Bogucki (2020: 95) points out that, despite the large number of publications on relevance theory more broadly, its application within AVT is very limited, and even scarcer in the case of subtitling. Needless to say, this is even more so for SDH. Even then, this does not necessarily mean that applying relevance theory to SDH is unfeasible or unrealistic, as Bogucki (2020: 96) claims that its application to SDH "seems both possible and desirable."

Given this, it is reasonable to draw on relevance theory as a useful framework. But before applying it, a brief introduction to the theory will be provided. After that, Gambier's (2026) three main aspects will be revisited and developed further.

4.3.4.1 Relevance theory

Relevance theory was proposed by Sperber and Wilson in *Relevance: Communication and Cognition*, first published in 1986 and revised in 1995. It is a framework for understanding human communication as an ostensive-inferential process. As an early application of this framework, Gutt (1991) brought relevance theory into TS. Central to his argument is that translation is not a mere comparison of texts, but a dynamic, cognitively grounded interaction among the translator, the original author, and the audience. Soon after, relevance theory was introduced into AVT. Kovačič (1994), as one of the first scholars to apply relevance theory, used it to explain subtitling reduction as a strategy for preserving essential meaning while reducing processing effort for viewers. Beyond Kovačič (1994), relevance theory has been applied, albeit to a limited extent, in subtitling research. Bogucki (2004) addresses the constraints on subtitling, and his later work (Bogucki, 2020) explores subtitlers' decision-making processes in this domain, both grounded in relevance theory. Sasamoto (2024) explains how a relevance-theoretic approach can account for the interaction between creative subtitling and viewers' interpretation, as well as its role in engaging audiences. Along similar lines, Wang et al. (2025) use relevance theory to analyse how subtitling strategies affect the comprehension of humour and cultural references.

What makes relevance theory so appealing for AVT, and subtitling in particular, is perhaps its strong explanatory power in accounting for the communicative behaviour of both subtitlers and viewers. In relevance theory, communication is viewed as an ostensive-inferential process, in which speakers produce ostensive stimuli and hearers infer the intended meaning on the basis of contextual assumptions. This process involves two layers of intention, as noted by Wilson and Sperber (2004: 611):

1) **The informative intention:** the intention to inform an audience of something;

2) **The communicative intention:** the intention to inform the audience of one's informative intention.

If we look at the two layers of intention from the speaker's view, the two principles of relevance operate at a broader cognitive level that accounts for how utterances are interpreted by the hearer. These two principles are as follows (Sperber & Wilson, 1995: 260):

1) **The Cognitive Principle of Relevance:** human cognition tends to be geared to the maximisation of relevance;

2) **The Communicative Principle of Relevance:** every act of ostensive communication communicates a presumption of its own optimal relevance.

The "cognitive principle" reflects a general tendency of human cognition to maximise relevance. Specifically, people tend to pay attention to information that gives them sufficient cognitive effects while requiring as little processing effort as possible. These cognitive effects may include contextual implications (where new information interacts with existing assumptions), the strengthening of existing assumptions, and the contradiction and elimination of prior assumptions (Sperber & Wilson, 1995: 115).

This same tendency also applies in human communication activity, where it is reflected in the "communicative principle" through the notion of "optimal relevance". This means that an utterance is expected to provide information that is worth processing while requiring as little effort as possible from the hearer (Sperber & Wilson, 1995: 233). To put it simply, hearers are guided towards the optimally relevant interpretation, that is, the one that yields sufficient contextual effects with minimal processing effort.

In subtitling, the interaction between subtitlers and viewers can also be considered as a form of communication. By using appropriate strategies, subtitlers intend to guide viewers towards an optimally relevant interpretation with reasonable effort. This idea is further explored by Chen and Wang (2019). In their study, they examine how subtitles interact with visual information and how this influence viewers' understanding. Their findings reveal that when relevant visual messages are incorporated into subtitles, viewers can process

information more smoothly and achieve stronger contextual effects. When this is not the case, subtitles and visual images become disconnected, and viewers must move between the two to piece together the meaning. This added effort makes processing more demanding and, as a result, reduces the overall degree of relevance.

When it comes to SDH, the interaction between subtitlers and viewers becomes even more salient, as subtitles are expected not only to represent spoken dialogue but also to render non-verbal audio information within limited time and space constraints. Under these conditions, producing SDH is less about transferring all available auditory information everything, but more about deciding what matters most for viewers and how best to convey it, so that viewers can access meaning with less processing effort. To this end, the following sections consider how this balance is handled in practice. For clarity, the discussion still follows Gambier (2026) and is organised around three aspects: what is conveyed, what is omitted, and what is added or clarified.

4.3.4.2 What information is to be conveyed

As discussed in Section 2.4.4.2, non-verbal audio information in SDH can be broadly grouped into five categories: character identification, paralinguistic information, music, sound effects, and silence. In theory, SDH subtitlers are supposed to render all of this information so as to provide access comparable to that of hearing viewers. In practice, though, time and space are limited, which means not everything can be included. This makes the task one of deciding what to prioritise.

Existing guidelines offer a helpful starting point. The DCMP Guidelines (2019) recommend including background sounds only when they are essential to the plot. The BBC Guidelines (2026) take a similar position, suggesting that music should be indicated when it is part of the action or important for the story. A more concise suggestion is offered by Szkarkowska and Jankowska (2025: 178), who propose that subtitlers should "select only the meaningful, plot-pertinent sounds that are important for following the story and which meaningfully contribute to the atmosphere but cannot be inferred from the screen or the context."

However, according to relevance theory, the issue is not simply whether something is related to the plot, but how to select information that provides sufficient contextual effects while minimising processing demands on viewers. For example, in a scene where sad background music accompanies a character's loss, the music may not directly advance the plot but still conveys emotional depth that is not accessible to DHH viewers. In this case, it may make sense to add a subtitle such as [sad music playing], so that viewers can perceive this emotional resonance with little effort. What this example illustrates is that, deciding what to convey in SDH is not simply a matter of whether information relates to the plot, but of determining what is most relevant for viewers.

4.3.4.3 What information is to be deleted

While deciding what to include is central to SDH, it is also important to know what to leave out. Bogucki (2020: 38), drawing on earlier eye-tracking studies, makes the point that nearly any subtitle added to the AV text can detract from viewers' appreciation of the polysemiotic message and increase cognitive effort. This can be seen in cases where a subtitle simply repeats what is already available, and the effort it requires becomes unnecessary. In such cases, omission is not a loss, but a way of avoiding extra processing load.

In practice, this often concerns sounds that add little new information, or can be easily inferred from what is already visible. Take a kitchen scene where a kettle is boiling in the background. The sound may matter for the story, as it implies the audience that the character is about to turn off the stove. Yet for DHH viewers, this is often already clear from the visuals, as steam is rising or the character is moving towards the stove. In this scene, if subtitlers add a description such as [kettle whistling], it may not add much, as it repeats what is already clear while still requiring processing. Instead, leaving it out would be more appropriate.

In this regard, omission in SDH is less about leaving things out and more about guiding what viewers need to pay attention to. When certain sounds are redundant or already easily

accessible, SDH subtitlers can choose to leave them out, as doing so reduces potential processing demands on viewers.

4.3.4.4 What information is to be added or clarified

Cases of addition often overlap with decisions about what to convey, since they both involve choosing which non-verbal sounds are worth bringing into the subtitles. As such, the focus here is on clarification, that is, making information more explicit for viewers. While this may seem at odds with the tight constraints of time and space, some degree of clarification is still necessary in SDH.

As a simple example, when using labels for speaker identification, it may be more helpful for viewers if SDH subtitlers refer to a character as [manager] rather than simply [man], as this provides more useful contextual information and reduces ambiguity. This is also the case in the description of music. Consider a scene where a character sits alone in a bar while a song plays softly in the background. The song's lyrics, *"I'm not ready to let you go"*, subtly reflect the character's inner state. If subtitlers choose to describe [sentimental song playing] rather than include the lyric [I'm not ready to let you go], DHH viewers may not be able to access the same emotional subtext that hearing viewers perceive from the music.

As illustrated by these examples, clarification in SDH is not intended to increase the overall amount of information, but to make contextual effects more salient. When the difference in subtitle length is minimal, it seems preferable to provide a more specific or detailed description. In short, the aim of clarification is to support interpretation and appreciation by maximising contextual effects and maintaining a manageable level of processing effort.

4.4 Implications for SDH production in China

Through the analysis of existing SDH guidelines, this chapter brings together a number of practical points that may inform the development of SDH production in China. As many of these have yet to be empirically tested, the aim here is not to propose a fixed model, but to

outline a set of workable directions grounded in both established practices and theoretical considerations.

First, in terms of acceptability, what matters here is how punctuation, symbols, and numbers are represented. While the guidelines reviewed adopt slightly different punctuation conventions, they all emphasise the need to maintain clarity and coherence. For Chinese subtitles, features specific to the language, such as Chinese brackets, should be respected. As for symbols, musical notes are recommended for indicating music, though more creative practices like emojis or emoticons may be better avoided at this stage. In relation to numbers, the Netflix Chinese Style Guide (2025) offers a useful and consistent point of reference.

Second, regarding legibility, which relates more to the visual presentation of subtitles, several aspects need to be considered. A starting point is font type and size. Sans-serif fonts are generally preferable for on-screen clarity. As for font size, given the range of influencing factors, it is difficult to prescribe a fixed value, but it should keep clear and legible without obscuring the image. In terms of positioning, subtitles are best kept at the bottom of the screen, following established viewing conventions. For speaker identification, different colours can be used, and where there are multiple speakers, labels may be introduced when necessary. With regard to subtitle layout, the number of lines is best limited to no more than three. Line length depends on font type. With proportional fonts such as SimHei, roughly 18 characters per line is a practical reference point. As for subtitle speed, it is advisable to keep it within a manageable range, ideally avoiding values that exceed around 11 cps.

Third, in relation to readability, the focus is on the use of verbatim and edited subtitles. While some guidelines favour verbatim representation, empirical evidence suggests that verbatim subtitles are not always the most effective option (cf. Szarkowska et al., 2011; Yang et al., 2025). This is closely linked to subtitle speed. Given an upper limit of around 11 cps, exceeding this threshold may require some editing to ease processing. When the pace

is manageable, there is less need for reduction, and the original dialogue can be fully retained.

Finally, the discussion of relevance draws on relevance theory, as this aspect is only briefly addressed in existing guidelines and has yet to be explored in depth in AVT research. Following relevance theory, SDH subtitlers should not only focus on conveying sounds that meaningfully contribute to the scene, but also support viewers' interpretation and appreciation, while keeping processing effort to a minimum.

4.5 Chapter summary

This chapter selected five sets of guidelines: four from pioneering countries where SDH guidelines are well developed, and one from Netflix, which specifically addresses Chinese subtitle conventions. By comparing how these guidelines deal with acceptability, legibility, readability, and relevance, the chapter summarised some practical implications for SDH production in China. Some of these can be directly applied to the SDH production for the reception study in this PhD project.

Chapter 5 Methodology

5.1 Chapter overview

This chapter outlines the research methodology and design used in the present study. To address the RQs (Section 5.2), a mixed-methods research design has been applied, specifically an exploratory sequential mixed-methods design, with a qualitative phase followed by a quantitative phase, as introduced in Section 5.3.

The qualitative phase (Phase I, Section 5.4) is an interview-based study with nine SDH subtitlers. It begins with a description of the participants (Section 5.4.1), followed by an overview of the tools and materials (Section 5.4.2). The section then details the interview process (Section 5.4.3), including preparation, data collection, and data processing, and concludes with ethical considerations for this phase. Finally, the qualitative data analysis procedures are described in Section 5.4.4.

Thereafter, the quantitative phase (Phase II, Section 5.5) is presented. This phase is a reception study conducted in China with 38 DHH individuals. It covers participant recruitment and profiles (Section 5.5.1), tools and materials (Section 5.5.2), and the reception study process (Section 5.5.3). The procedures for quantitative data analysis are presented in Section 5.5.4.

5.2 Research objectives and questions

As introduced in Section 1.2, this thesis is concerned with SDH from both the production and the reception sides. Accordingly, the study is organised into two phases and has two corresponding research objectives:

Phase I

Objective 1: To explore the decision-making processes of professional SDH subtitlers and draw implications for the creation of the SDH subtitles used in the reception study.

Phase II

Objective 2: To investigate how Chinese DHH participants engage with SDH subtitles and provide empirical evidence that may inform the future design of SDH subtitles for Chinese DHH audiences.

In line with these objectives in the two phases, the study intends to address the following RQs:

Phase I

RQ1: What key factors influence SDH subtitlers' decision-making processes?

RQ2: What implications can be drawn from the study with SDH subtitlers that can inform the production of SDH subtitles?

Phase II

RQ3: How do the gaze patterns of DHH participants differ when watching videos with SDH subtitles compared to those with ST subtitles?

RQ4: Do DHH participants who viewed SDH subtitles achieve better comprehension performance than those who viewed ST subtitles?

RQ5: What differences can be observed in the overall viewing experience of DHH participants, in terms of engagement and enjoyment, when watching videos with SDH subtitles versus ST subtitles?

To address the above RQs and achieve the stated objectives, a mixed-methods approach was adopted. The following section begins with an overview of the mixed-methods research design and then explains how this approach will be applied in the study.

5.3 Mixed-methods research

5.3.1 Introducing mixed-methods research

Simply put, mixed-methods research commonly refers to the combination of both qualitative and quantitative research methods within a particular study. Drawing from 19 definitions provided by 21 widely-cited scholars in the field, Johnson et al. (2007: 123) offer a conclusive definition of mixed-methods research:

> "Mixed-methods research is the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration."

In this context, Creswell and Plano Clark (2018: 39-40) interpret Johnson et al.'s words as framing mixed-methods research as a "methodology that spans viewpoints to inferences," rather than merely a research method. This interpretation highlights the strength of the mixed-methods approach, as it enables researchers to address research questions with sufficient depth and breadth. In detail, quantitative research facilitates the collection of data from a large number of participants, which increases the potential for generalising research findings to a broader population. Qualitative research, on the other hand, allows for a deeper exploration of the study's target issues while ensuring that participants' voices are respected (Dawadi, 2021: 27). Thus, it is fair to conclude that quantitative data provides breadth to the research, while qualitative data offers depth. Moreover, a mixed-methods approach leverages the strengths of both qualitative and quantitative methods, while minimising the weaknesses associated with relying on a single research approach (Johnson & Onwuegbuzie, 2004: 14-15).

To implement the mixed-methods approach effectively, Creswell and Plano Clark (2018: 121) identify three key mixed-methods designs: 1) the convergent design, 2) the explanatory sequential design, and 3) the exploratory sequential design. According to Creswell and Plano Clark (2018), the convergent design is a method of collecting and analysing

quantitative and qualitative data simultaneously. Its purpose is to compare and integrate the two datasets to validate or enrich the findings. The explanatory sequential design begins with the collection and analysis of quantitative data, followed by qualitative data collection to further explain or expand on the initial quantitative results. The exploratory sequential design, by contrast, starts with qualitative data collection and analysis. Building on the qualitative results, researchers may develop a feature, such as new variables, new measures, or new experimental activities, which are then used in the subsequent quantitative data collection.

It is worth noting that the difference between the explanatory sequential design and the exploratory sequential design lies not only in the order of their phases but also in whether new features are developed. The explanatory sequential design typically does not require the development of new features; instead, it interprets quantitative findings through qualitative data. By comparison, the exploratory sequential design often requires designing new features (e.g., new instruments) during the qualitative phase, which are then used in the subsequent quantitative phase. Considering this difference, Creswell and Plano Clark (2018: 149) note that the exploratory sequential design tends to take considerable time to complete. Therefore, it is recommended that researchers adopting this sequential approach "plan for extended time" and "build time into their study's plan" (*ibid.*).

Because of the distinct strengths of the mixed-methods approach and the availability of various research designs, mixed-methods research designs have become increasingly common in TS (Orero et al., 2018: 108), and AVT research is no exception. In SDH studies, examples can be found in doctoral research over the past decade, including Zárate (2014), Gürkan (2019) and Cepak (2021). These studies combine qualitative and quantitative methods to investigate various aspects of SDH, such as the reception of SDH among hearing-impaired audiences in different countries. By the same token, this study employs a mixed-methods approach to explore SDH production by SDH subtitlers and investigate SDH reception by SDH users. The following section describes how the mixed-methods research design is applied in the present study.

5.3.2 Mixed-methods research design

As introduced in the previous section, the present study adopts a mixed-methods approach, implemented in two phases, following an exploratory sequential mixed-methods design where qualitative phase is carried out prior to quantitative phase.

Phase I was a qualitative stage that collected data through semi-structured interviews conducted between May and June 2024. Participants were nine professional SDH subtitlers from a national television broadcaster, a Chinese accessibility company, and a subtitling group on a YouTube channel. The objective of these interviews was to gain insight into subtitlers' experiences and decision-making processes in the production of SDH. To explore their perspectives, ten open-ended questions were prepared to allow participants to express their views while staying focused on core topics. Findings from these interviews were applied to produce the SDH subtitles used in the reception study of Phase II.

Phase II was a reception study carried out in May 2025 with 38 Chinese DHH participants, during which quantitative data were collected. This phase began with a pre-task questionnaire designed to collect information on DHH participants' demographic background, viewing habits, and subtitle preferences. Participants were then assigned to two groups in a way that balanced their degree of hearing loss across conditions, forming a between-subjects design with subtitle condition (SDH subtitles vs. ST subtitles) as the independent variable. One group watched the video with SDH subtitles, while the other watched it with ST subtitles. During the viewing, an eye-tracking device recorded participants' gaze patterns, and each participant completed their eye-tracking session individually. After the eye-tracking sessions, participants completed a post-task questionnaire consisting of two parts: (1) a comprehension test to assess their understanding of the AV content, and (2) a Likert-scale questionnaire to evaluate their viewing experience, including levels of engagement and enjoyment. In this between-subjects design, the independent variable was subtitle condition (SDH subtitles vs. ST subtitles), and the dependent variables were: percentage of fixation duration, average

fixation duration, number of fixations, average pupil diameter, comprehension test, and self-reported questionnaire responses measuring participants' engagement and enjoyment. The results of Phase II intended to reveal how Chinese DHH participants comprehended, engaged with, and visually interacted with SDH subtitles. These findings, together with those from Phase I, could be used to formulate recommendations for improving the design of SDH subtitles in the Chinese context.

The two phases of the research, namely a qualitative phase followed by a quantitative phase, constitute an exploratory sequential mixed-methods design, as introduced in the previous section. This sequential approach, as noted by Creswell and Plano Clark (2018: 144), is suitable for studies without predefined guiding frameworks or theories, as it enables the discovery of new concepts and the development of context-specific tools. In the case of the present study, the production of SDH subtitles is regarded as a new "tool", given the absence of readily available SDH subtitles in China. This tool was developed from the first phase, drawing on the practices and experiences of professional SDH subtitlers.

In this research, complementarity is achieved by integrating qualitative and quantitative methods. The qualitative phase, involving interviews with SDH subtitlers, offers an in-depth understanding of the subtitling process and the decision-making behind it. To counterbalance the subjectivity of qualitative data, quantitative evidence was collected through eye-tracking with SDH users, providing broader and more objective perspectives. This mixed-methods design is particularly well-suited to research contexts that are complex and multifaceted (Greene, 2007: 101).

In summary, this study adopts a mixed-methods approach to investigate the production and reception of SDH. Specifically, the qualitative phase uses semi-structured interviews with SDH subtitlers to examine their decision-making processes and identify key factors influencing the creation of SDH. These findings inform SDH design in the follow-up quantitative phase, which uses eye-tracking and questionnaires to understand the

reception of SDH subtitles among hearing-impaired audiences in China. The following sections will further introduce the qualitative and quantitative phases, respectively.

5.4 Phase I: Interview study with SDH subtitlers

5.4.1 Participants

5.4.1.1 SDH subtitlers

SDH subtitlers, also known as closed captioners (Zdenek, 2015: 38), are professionals directly involved in the production of SDH subtitles. Compared to those working on ST subtitles, SDH subtitlers face a greater workload and more complex challenges. The increased workload is evident because they are required not only to transcribe dialogue but also to convey a wide range of non-verbal audio information. Even in this era of rapid advancements in artificial intelligence, where modern automation tools, such as speech recognition, have simplified certain aspects of the SDH workflow, these tools remain limited in effectively handling complex non-verbal information. Szarkowska and Jankowska (2025: 186) point out that even with future technological advancements, improvements are likely to be more evident in aspects of personalisation, such as adjustments to fonts and colours. As such, most of the time, extensive manual intervention remains essential, which also presents challenges for SDH subtitlers. These challenges, as Szarkowska and Jankowska (2025: 179) aptly note, involve not only deciding which sounds to describe but also determining how to name them. More specifically, it is not feasible to caption every sound due to limited space and time on the screen, particularly when multiple sounds overlap, making it difficult for novice SDH subtitlers to prioritise which sounds to include (Szarkowska & Jankowska, 2025: 178). Moreover, it is particularly challenging to name these sounds because subtitlers need to accurately reflect their characteristics or relevance to the scene. Meanwhile, the descriptions of sounds must be clear and understandable to the DHH audience, who rely on them for a fuller comprehension of the AV content.

5.4.1.2 Recruitment and profiles

Before recruiting participants for Phase I of the interview study, three inclusion criteria were established:

1) Participants had to be subtitlers specialising in SDH, rather than those working merely on ST subtitles;
2) Participants were required to have relevant professional experience and to have produced actual SDH content;
3) Participants had to be at least 18 years old at the time of the interview.

In order to maintain diversity among participants, I did not set specific requirements regarding the platforms where their SDH subtitles were published, the types of AV content they worked on, or whether they were involved in intralingual or interlingual SDH.

With the above criteria, I combined purposive and snowball sampling methods to recruit eligible participants in this phase of the study. Purposive sampling, in which the researcher selects participants based on specific, intentional criteria, so as to address key aspects of the RQ, rather than relying on random selection (Kemper et al., 2003: 279; Saldanha & O'Brien, 2014: 180). As described by Moorkens (2012: 89), this sampling method benefits research by obtaining detailed insights from qualified participants with in-depth knowledge relevant to the study, instead of collecting general opinions from a large sample size. On the other hand, snowball sampling is a method of participant recruitment where an initial group of participants helps the researcher to recruit additional participants through their own networks (Kemper et al., 2003: 283; Saldanha & O'Brien, 2014: 34). The advantage of this type of sampling is that it "offers real benefits for studies which seek to access difficult-to-reach or hidden populations" (Atkinson & Flint, 2001: 1). In short, these two sampling methods were chosen because they ensure that the selected participants meet the specific criteria of the study and facilitate locating hard-to-reach and hidden populations.

Using these two sampling methods, a total of nine SDH subtitlers participated in this phase of the study. The first participant was a professional SDH subtitler working at a national

television broadcaster, whom I contacted directly through my supervisor's personal networks. This participant (hereafter P1) then introduced me to two of her colleagues (P2, P3) who were also professionally engaged in SDH work. Yet, at that time, recruiting only three participants from the same organisation seemed insufficient, which prompted me to seek additional participants. I subsequently identified a Chinese accessibility company that provides AV products with accessibility services, including AD, SDH and SLI, and then reached out to its founder via the contact information provided on the website. The founder (P4), along with three of her colleagues (P5, P6, and P7), who are scriptwriters of SDH subtitles, expressed interest in participating in the interviews. Lastly, the remaining two participants were recruited from a subtitling group on YouTube, which had translated Chinese TV dramas into English and also provided SDH subtitles in English. In the same way, I contacted them using the email address listed in the credits. The leader of the subtitling group (P8) accepted my invitation for an interview and subsequently introduced me to one of his colleagues (P9). The above is the entire process of participant recruitment in this phase.

As mentioned above, nine SDH subtitlers were invited to participate in this phase of the study. The limited number of participants is due to several factors. First, the population of SDH subtitlers is relatively small and highly specialised, which limits the available pool of participants. Second, participants were required to have prior experience in SDH production, further narrowing the potential candidates. More importantly, the chief purpose of this phase is depth rather than breadth. It is designed to gain detailed insights from a smaller group rather than drawing broad conclusions from a larger sample.

Participant	Nationality	Years of SDH experience	SDH training	Employment status
Irish national broadcaster (Intralingual SDH: English to English)				
P1	Ireland	6 years	In-house training	Full-time
P2	Ireland	3 years	In-house training	Full-time
P3	Ireland	14 years	In-house training	Full-time
Chinese media accessibility company (Intralingual SDH: Chinese to Chinese)				
P4	China	5 years	In-house training	Full-time to part-time
P5	China	3 years	In-house training	Part-time
P6	China	5 years	In-house training	Full-time to part-time
P7	China	6 years	Self-taught training	Full-time
International volunteer subtitling group operating on YouTube (Interlingual SDH: Chinese to English)				
P8	China	7 years	No formal training	Volunteer
P9	United States	7 years	No formal training	Volunteer

Table 5-1. Detailed profiles of participants (Phase I)

Despite the limited sample size, the participant profiles are diverse. Nine SDH subtitlers represent various organisations, including a national broadcaster (n=3), an accessibility company (n=4), and a subtitling group operating on YouTube (n=2). Furthermore, the participants come from different countries across Asia, Europe, and North America, specifically China (n=5), Ireland (n=3), and the US (n=1). Of the nine participants, seven

specialise in intralingual subtitling (n=7), providing SDH subtitles in the same language as the original soundtrack, while two work in interlingual subtitling (n=2), translating content into another language. Additionally, these SDH subtitlers have worked on various types of AV content, such as films, teleplays, and television shows. The detailed profiles of the participants are illustrated in **Table 5-1**.

5.4.2 Tools and materials

5.4.2.1 Interview questions (IQs)

Given that the objective of this phase is to gain an in-depth understanding of the decision-making process of SDH subtitlers when creating SDH content, particularly regarding non-verbal audio information, a qualitative research approach using semi-structured interviews was adopted. In general, semi-structured interviews are characterised by open-ended questions and allow for flexibility in the order of questioning as well as the inclusion of follow-up questions (Saldanha & O'Brien, 2014: 172-173). Through semi-structured interviews, participants have the opportunity to narrate their experiences and thoughts. Moreover, since the questions are carefully tied to the research topic, there is minimal risk of deviating from the research focus (Galletta & Cross, 2013: 47). Therefore, the decision to conduct semi-structured interviews was made, and this approach also aligns with the research aims of collecting detailed qualitative data.

The semi-structured IQs were designed to address various key areas related to SDH production, including adherence to existing SDH guidelines, the selection of audio elements for subtitling, challenges encountered in various types of media, and potential innovations in SDH practices. **Table 5-2** details the ten IQs that were asked during the semi-structured interviews.

List of IQs	
IQ1	Could you walk us through the process of creating SDH for television programmes or other AV content?
IQ2	As far as I am aware, many countries have established guidelines for creating SDH. For example, in Ireland, there are the *Access Rules* and *A Guide to Commissioned Programmes*. In your experience, have you ever come across situations where you thought the current guidelines may not fully cover all the necessary aspects? If so, could you share more about which aspects you think could be improved?
IQ3	We understand that SDH is used to convey essential audio details, but in practice, not every sound gets subtitled. I've noticed that sometimes when speakers use exaggerated tones or convey irony, such information often doesn't make it into the subtitles. So, could you share some insights or thoughts on the decision-making process for selecting which sounds to subtitle?
IQ4	When describing music, do you tend to use terms like pop music, classical music, jazz music, rock music to categorise different genres? If so, to what extent do you think deaf or hard-of-hearing audiences can understand these music genres? If not, could you share with us how you or your team describe music?
IQ5	Have you considered using emojis, such as smiley faces, sad faces, or angry faces in subtitles to convey the speaker's emotions? Would this bring any technical challenges?
IQ6	What challenges do you encounter when captioning content, especially in terms of captioning some non-verbal audio information, such as music, sound effects, paralinguistic information, or even silence?
IQ7	Which type of media content do you find most challenging to create SDH, such as horror films, comedy films, or reality shows? Could you please elaborate on the specific reasons?

IQ8	I noticed that in the *Access Rules*, there are specific requirements for subtitles aimed at children, such as reducing the subtitle speed and increasing the use of three-line subtitles. Apart from these requirements, do you also adjust the language style, for example, using simpler and more easily understandable descriptions in animated films?
IQ9	Some viewers feel like SDH can be a bit subjective. In other words, the description of sounds sometimes reflects subtitler's subjective interpretation. For instance, for the same piano sound, some subtitlers may describe it as "beautiful piano sound", while others might opt for "gentle piano sound". In such cases, could you provide some recommendations on how to maintain a balance between being objective and making sure the descriptions are clear?
IQ10	Since SDH services are not yet widespread in China, and there are few subtitlers specialising in creating SDH, do you and your team have any advice for subtitlers in China who are looking to enter this field?

Table 5-2. List of ten interview questions

As can be seen, the initial questions (*IQs 1 and 2*) were designed to explain the SDH subtitler's workflow of creating SDH content. At the same time, by discussing the existing SDH guidelines, SDH subtitlers were able to share both the helpful aspects of these standards and any limitations they had encountered, which also opened the discussion for suggestions and improvements that might be considered for future SDH guidelines. As for *IQs 3, 5 and 6*, this set of questions invited participants to share their decision-making process and challenges regarding subtitling non-verbal audio information. In this regard, participants were also asked to discuss the possibility of using innovative tools such as emojis to enhance emotional representation in SDH, as well as any technical issues that they might present. *IQs 4 and 9* were intended to examine the perspectives of SDH subtitlers on describing music in SDH content and balancing objectivity with clarity in sound descriptions.

During the interview, participants were encouraged to discuss whether certain genres might resonate differently with DHH audiences and to share their views on potential subjectivity in these descriptions. *IQs 7 and 8* addressed genre-specific and age-specific issues in SDH production, respectively. For different types of media, participants were asked to share which types present the greatest challenges and to further explain the reasons involved. In the case of children's content, I sought to understand whether subtitlers adjusted language style and subtitling speed to better serve younger audiences, a practice discussed here to understand how subtitlers adapted SDH content for specific viewing needs. Lastly, *IQ 10* attempted to seek practical advice from experienced SDH subtitlers for those interested in entering the SDH field, particularly in regions like China where SDH practices are less established. This question allowed interviewees to share their experiences, technical training, and essential skills and knowledge for novice SDH subtitlers.

5.4.3 Interview study process

The discussion of the semi-structured interviews is divided into three steps:

1) **Interview preparation:** This step includes designing IQs, developing the Plain Language Statement and the Informed Consent Form, obtaining ethics approval, reaching out to suitable participants to arrange interviews, and preparing audio recording equipment.
2) **Conducting interviews:** In this step, I scheduled one-to-one semi-structured interviews with participants to gather their insights and experiences based on the prepared RQs.
3) **Data processing:** This phase involves transcribing the recorded interviews and coding the data for further analysis.

The following sections provide a detailed description of each step.

5.4.3.1 Step 1: Interview preparation

In line with the objective of this phase, which is to understand SDH subtitlers' decision-making processes when creating SDH content, ten open-ended IQs were designed (see Section 5.4.2.1). Then, I moved on to draft the Plain Language Statement and the Informed Consent Form, both essential for the ethics application. The Plain Language Statement is

intended to provide participants with clear and accessible information about the study, while the Informed Consent Form should specify critical interview details (e.g., data security, access to research findings, potential risks and benefits) and participants' rights, including their right to withdraw at any time. In the end, I submitted the ethics approval application to Dublin City University's Humanities and Social Sciences Faculty Research Ethics Committee on April 2, 2024, and received confirmation of approval on April 23, 2024 (Reference: DCU-FHSS-2024-037; see Appendix C).

After obtaining ethics approval, I began contacting participants to schedule interview times and locations. Prior to each interview, in addition to the Plain Language Statement and the Informed Consent Form, the ten IQs were sent to each participant so that they could have time to consider the topics in advance. I expressed that reviewing these questions beforehand was optional, and their initial, unprepared responses would also be valued and appreciated. With that said, most participants reviewed the questions before the interview and expressed that this approach was helpful in saving time during the interview and reducing uncertainty. Some participants even looked up information related to the IQs. For example, one participant looked into definitions of various music genres such as pop, classical, jazz, and rock to prepare for Question 4, which asked about the way of describing music in SDH.

For audio recording, I prepared a wireless lavalier microphone to ensure clear audio capture, even in environments with background noise. This microphone connects via Bluetooth and records directly to the Voice Memo application on my phone. Tests showed that it provides high sound quality and captures speakers' voices more clearly than recording without it. Prior to each interview, both the microphone and phone were fully charged and tested to ensure optimal recording performance.

5.4.3.2 Step 2: Conducting the interview

Upon meeting each participant, I began with a brief self-introduction and casual conversation to help them feel comfortable and ease into the interview. And then, a brief

explanation of key terms used in the interview was provided to each interviewee, such as clarifying the full term for SDH and noting that this would be used throughout the interview. The purpose of the interviews was also introduced, even though participants had a general idea about the study from previous communications. For participants who had not submitted a consent form beforehand, a printed copy was prepared for them to sign before the interview started. A copy of the ten IQs was also provided for reference during the interview. Next, it was explained that the interview would be recorded for transcription purposes, and consent was obtained from each participant to proceed with the recording. With their agreement, the microphone was clipped to both the participant's collar and my own, and the interview was recorded using the Voice Memo app on my phone. With all preparations complete, we moved on to the main part of the interview, where participants shared their experiences and views based on the IQs. At the end of the interview, after all questions had been asked, participants in the three in-person interviews were given a €20 One4All gift card as a token of appreciation, while those in the remaining online interviews received compensation of equivalent value. With the full cooperation of all interviewees, the nine interviews proceeded smoothly, and most expressed that they looked forward to the results of the study.

It should be noted that each interview was conducted one-on-one, either online or in person. The expected duration of each session was approximately one hour, although adjustments were made based on participants' availability. The interview process was audio-recorded for transcription purposes. The language used in the interviews depended on the participants' native languages. I now present **Table 5-3**, which contains detailed information about the nine interviews.

Participant	Date	Duration (minutes)	Form	Interview language
P1	17/05/2024	66	In person	English
P2	23/05/2024	52	In person	English
P3	29/05/2024	34	In person	English
P4	17/06/2024	65	Online	Chinese
P5	18/06/2024	55	Online	Chinese
P6	19/06/2024	43	Online	Chinese
P7	21/06/2024	78	Online	Chinese
P8	01/06/2024	55	Online	Chinese
P9	21/06/2024	53	Online	English

Table 5-3. Detailed information on the nine interviews

5.4.3.3 Step 3: Data processing

Upon completion of the interview sessions, the audio recordings of the nine interviews were uploaded in m4a format to a password-protected folder on the DCU student Google Drive. The recordings were transcribed using Happy Scribe, an audio transcription tool. Initially, the intention was to transcribe the audio recordings manually without utilising such tools, but this proved to be a very time-consuming task. As noted by Moorkens (2012: 96) in his doctoral research, “one hour of interview required eight hours of transcription” based on his experience. Thus, the decision was made to employ audio transcription tools in this study. Despite the aid of these tools, multiple listens to the recordings were necessary to verify and modify the transcriptions for accuracy. In addition, the tones and expressions of

the interviewees were recalled to improve the transcribed content. For instance, one interviewee, when discussing the challenges involved in describing non-verbal audio elements, whistled twice — once ascending and once descending. Nevertheless, the audio transcription tool could not recognise these sounds; in such cases, it is required to clarify them in the transcription as [a descending whistle sound] and [an ascending whistle sound].

Another point to clarify is that for interviews conducted in English, the audio recordings were transcribed in English. For interviews conducted in Chinese, I first transcribed them in Chinese and then translated them into English to maintain consistency with the English interviews for the sake of facilitating subsequent analysis.

Indeed, by transcribing, reviewing, and organising the collected data, certain themes relevant to the RQ emerged, while other themes also became evident (Galletta & Cross, 2013: 121). At this stage, it became necessary to employ professional tools for a systematic analysis of the data. Thus, upon the completion of audio transcription, the documents were imported in Word format into MAXQDA, a qualitative analysis software tool, in the order of P1, P2, …, P9 for further qualitative analysis. Within MAXQDA, thematic patterns can be identified and labelled, reflecting ideas evident in the data. As noted by Galletta and Cross (2013: 122), "these ideas represent a core level of meaning and are often referred to as *codes*." This term is associated with a common approach used in social science studies — *coding*. According to Kvale and Brinkmann (2015: 227), coding "involves attaching one or more keywords to a text segment in order to permit later identification of a statement." Following this, the interview data were subjected to an initial coding process.

During the actual coding process, while some codes were closely related to the research focus, other unexpected codes emerged that provide new insights into the study. To ensure that no potential information was overlooked and to uncover deeper meanings embedded in the data, this study adopted Grounded Theory (Glaser & Strauss: 1967) as the coding method for coding. A key feature of Grounded Theory is its bottom-up approach, which builds theory directly from the data through systematic comparison and analysis, rather

than relying on predefined frameworks or existing viewpoints (Charmaz, 2014: 139). Furthermore, Grounded Theory emphasises the iterative nature of data analysis (Corbin & Strauss, 2015: 55), which involves a systematic process of constant comparison to code the data in three stages: open coding, axial coding, and selective coding. During the open coding stage, the data are broken down into a number of segments, and initial codes are assigned to each meaningful segment. The axial coding stage identifies relationships among these initial codes and groups them into broader categories. Finally, the selective coding stage integrates these categories into a coherent framework that directly addresses RQ1. Throughout the process, researchers must continuously review and refine the codes. Although it is a slow and painstaking process, it ensures that the codes accurately represent the data and that no potentially significant codes are missed. Since the coding process forms the core of the data analysis, it will be discussed further in Section 5.4.4, where the results of data analysis are presented in detail.

5.4.3.4 Ethical considerations

This research project consists of two phases, and separate ethical approvals were sought for each phase. For Phase I, which worked with professional SDH subtitlers and did not include children or vulnerable adults, the ethical application was reviewed by the Faculty Research Ethics Committee rather than the University-level committee. As mentioned in Section 5.4.3.1, the application, which included the Plain Language Statement, Informed Consent Form, and IQs, received ethics approval on April 23, 2024, from Dublin City University's Humanities and Social Sciences Faculty Research Ethics Committee. The committee classified Phase I of the study as a low-risk project (Reference: DCU-FHSS-2024-037; see Appendix C).

Although Phase I was categorised as a low-risk project, several ethical considerations were necessary to address for ethical and confidential issues. To protect participant anonymity, neither the thesis nor any related publications or presentations will mention participants' full names. Interview data have been anonymised through assigned identifiers (e.g., P1,

P2, …, P9), and their real names will not be disclosed. With regard to data storage and security, all data were securely stored in a password-protected folder on my DCU student Google Drive, and only my supervisors and I had access to the data. After completing the research project, all interview data will be permanently destroyed to protect participants' privacy and confidentiality.

5.4.4 Qualitative data analysis

Phase I collected qualitative data through semi-structured interviews with nine SDH subtitlers. The analysis of the interview data was guided by the Grounded Theory approach and consisted of three main stages of coding:

1) **First-cycle coding** was conducted to break down the interview transcripts into meaningful units of data and to identify initial codes emerging from participants' accounts.
2) **Second-cycle coding** was carried out to compare, refine, and group the initial codes into more abstract categories that reflected broader concepts in the data.
3) **Third-cycle coding** was the final stage of analysis, in which the major categories were integrated into core themes explaining the decision-making processes of SDH subtitlers.

The above qualitative analyses were all conducted using MAXQDA (version 24.0.0), which helped organise, manage, and compare the codes across the interview transcripts.

5.4.4.1 Grounded Theory

<u>Rationale for using Grounded Theory</u>

When considering appropriate methods for qualitative data analysis, a review of the literature showed that there are a wide range of possible options. As listed by Silver and Lewins (2014: 23-33), five common methods of qualitative data analysis are discourse analysis, narrative analysis, framework analysis, grounded theory, and thematic analysis. These five methods are all effectively supported by qualitative data analysis software (Kuckartz & Rädiker, 2019: 7). Each method may arguably suit different types of qualitative data depending on the research aims.

To analyse the qualitative interview data collected in Phase I, Grounded Theory was adopted for two main reasons. First, it adopts a bottom-up approach that builds theory directly from the data through systematic coding and constant comparison, rather than relying on pre-existing theoretical assumptions (Charmaz, 2014: 139). Second, Grounded Theory is highly compatible with in-depth qualitative interviewing. As Charmaz (2006: 28) argues, Grounded Theory and in-depth interviewing share several key characteristics: both are "open-ended but directed, shaped yet emergent, and paced yet flexible approaches."

In practice, through in-depth interviews, researchers can collect targeted data around research topics while adapting IQs to explore new themes as they arise. In the same vein, Grounded Theory emphasises control and flexibility in data collection and analysis. Grounded theorists seek to understand what is happening from the start of their research but remain open to adjustments to avoid preconceived biases. In addition to identifying and exploring themes during interviews, grounded theorists can return to the field to generate new ideas and address conceptual gaps. This is why Charmaz (2006: 29) insists that in-depth qualitative interviewing fits the Grounded Theory method particularly well, and their combination enhances "the analytic incisiveness of the resultant analysis."

For these reasons, the present study adopted Grounded Theory to analyse the qualitative interview data obtained from SDH subtitlers.

Introduction of Grounded Theory

Grounded Theory is a method of conducting qualitative research that originated from the book *The Discovery of Grounded Theory* by Barney Glaser and Anselm Strauss in 1967. While its name includes "theory", Saldanha and O'Brien (2014: 191) explain that Grounded Theory "is actually not a theory but a research method, and more specifically a mode of data analysis." Marvasti (2004: 84) further describes it as "an inductive or ground-up approach to data analysis." The concept can be further understood through the definition provided by Bryant and Charmaz (2007: 608), who state that Grounded Theory "focuses on creating conceptual frameworks or theories through building inductive analysis from the data,"

which reflects its name of “grounded.” In practice, Grounded Theory follows two basic principles: first, data analysis must be grounded in empirical evidence and follow a systematic coding process; second, the analysis should generate a theory that goes beyond merely describing the phenomenon and offers a coherent, contextualised explanation for it (Saldanha & O’Brien, 2014: 191).

Aside from the two basic principles mentioned, Grounded Theory emphasises “the desire to step beyond the known” (Corbin & Strauss, 2015: 55); that is, researchers do not begin with preconceived notions nor draw conclusions directly from existing theoretical frameworks. As Star (2007: 90) puts it, Grounded Theory continually reminds researchers to remain “open to the data, open to being wrong, to redoing one’s own work, and actively seeking out new views and mistakes.” This approach enables researchers to uncover unexpected findings and fill conceptual gaps. Furthermore, Grounded Theory places emphasis on the interaction between researchers and participants. Researchers are encouraged to “enter into the world of participants, to see the world from their perspective, and in doing so, make discoveries that contribute to the development of empirical knowledge” (Corbin & Strauss, 2015: 55). To reflect this philosophy, the present study adopts a participatory approach in which the voices of SDH subtitlers are heard, and their thoughts regarding SDH production are prioritised for collection. Last but not least, in Grounded Theory, findings are derived from constant comparison through a multi-stage coding process (Kuckartz & Rädiker, 2019: 7). The initial data collection informs subsequent data gathering and analysis, with this iterative process continuing until a well-integrated and dense theory is constructed (Corbin & Strauss, 2015: 55).

When it comes to analysing the data, coding is considered the core process in Grounded Theory research, with the well-known three stages of coding: open coding, axial coding, and selective coding. Open coding is the initial step that involves identifying important words or groups of words within the data and assigning relevant codes to them (Boeije, 2010: 96). Axial coding, which is the second stage of data analysis, focuses on developing categories and subcategories based on the codes identified in the open coding stage and establishing

relationships among these categories (Boeije, 2010: 108). Selective coding, the final stage, aims to identify a core category or concept to describe and explain the observed phenomenon (Boeije, 2010: 115-116). Interestingly, Birks and Mills (2015: 24) observe that various terms have emerged to describe different coding phases in Grounded Theory, which can sometimes lead to confusion. To address this, Sigauke and Swansi (2020) summarise the terminology used by different Grounded Theory scholars to describe each stage of coding, as adapted in **Table 5-4**. As indicated in **Table 5-4**, Glaser and Strauss (1967) were the first to propose the three stages of Grounded Theory: coding and comparing incidents (initial coding), integrating categories and properties (intermediate coding), and delimiting the theory (advanced coding). Over time, different Grounded Theory scholars introduced various terms to name multiple stages of coding. Compared to the traditional models proposed by early grounded theorists, Saldana (2016) offers a diverse range of coding methods that may be used concurrently, particularly at the initial coding stage. At this stage, Saldana (*ibid.*) introduces several coding techniques, including "in vivo coding", which directly uses participants' own words; "process coding", which focuses on actions and sequences; and "initial or open coding" for capturing emerging patterns. Further details regarding these coding methods will be provided in the following section (Section 5.4.4.2). In this study, Saldana's data coding method was chosen to analyse the interview data because it offers various coding techniques at the initial phase, allowing me to select the most appropriate technique when dealing with extensive data during the first round of coding.

Proponent(s)	Multiple stages of coding		
	First cycle coding	Second cycle coding	Third cycle coding
Glasser & Strauss (1967)	Coding and comparing incidents	Integrating categories and properties	Delimiting the theory
Glasser (1978)	Open coding	Selective coding	Theoretical coding
Strauss & Corbin (1990)	Open coding	Axial coding	Selective coding
Böhm (2004)	Open coding	Axial coding	Selective coding
Charmaz (2006)	Initial coding	Focused coding	Theoretical coding
Holton (2010)	Substantive coding (includes both open coding and selective coding)		Theoretical coding
Saldana (2016)	In vivo coding, process coding, Initial/open coding	Axial coding	Theoretical coding

Table 5-4. Multiple stages of coding (based on Sigauke & Swansi, 2020: 11)

5.4.4.2 First cycle coding

The first cycle of coding, as described by Saldana (2016), consists of three specific coding techniques, including in vivo coding, process coding, and initial/open coding, although some scholars (Glasser: 1978; Charmaz: 2006) collectively refer to this stage as "initial coding" or "open coding". During this stage, researchers should fully immerse themselves in the data. This requires them to carefully examine the collected qualitative data, such as field notes, interview transcripts, or other documents, line by line or even word by word. The data is then broken down into different segments, such as lines, sentences, or paragraphs, to facilitate the assignment of appropriate codes. More than that, researchers should compare these segments for similarities and differences to ensure that the assigned codes accurately reflect the data (Strauss, 1987: 28; Saldana, 2016: 817). A crucial aspect of this phase is that

researchers must approach the data with an open mind, free from preconceived biases or assumptions. As Strauss (1987: 29) aptly puts it, researchers should "believe everything and believe nothing" during this coding phase. To this end, for each new piece of data, I followed Glaser's (1978: 57) guidance by repeatedly asking myself three questions to prevent preconceived notions from coding:

- What is this data a study of?
- What category does this incident indicate?
- What is actually happening in the data?

With that in mind, I started to work on the first-round coding. At the beginning of the first-round coding, the sheer volume of interview data required careful segmentation and processing to ensure effective analysis. Considering that Grounded Theory is a bottom-up approach, moving from data to theory, I decided to adopt line-by-line coding to study the data, as recommended by Charmaz (1990: 1168), to ensure that no detail was lost during this process.

With the line-by-line coding sequence, the next step is to determine how to assign appropriate codes to each segment of the data. As mentioned in the previous section, three coding techniques recommended by Saldana (2016) were adopted during this coding cycle: in vivo coding, process coding, and initial/open coding. A brief description of each technique is provided below:

- **In vivo coding:** This method uses words or phrases directly from participants' language in the data (Corbin & Strauss, 2015; Saldana, 2016). It is also referred to in some literature as "literal coding", "verbatim coding", "inductive coding", "indigenous coding", "natural coding", and "emic coding" (Saldana, 2016: 304). For consistency, the term "in vivo coding" will be used in this study.
- **Process coding:** Process coding uses gerunds ("-ing" words) to denote actions in the data. Simple observable activity (e.g., reading) and more general conceptual action (e.g., struggling) can be coded through process coding (Saldana, 2016: 319). As such, this method is sometimes labelled as "action coding" (*ibid.*).

- **Initial/open coding:** Also known as “free coding”, this method can be used alongside “in vivo coding” or “process coding” but remains flexible to incorporate other suitable methods. It often involves breaking down data into discrete parts and comparing them for similarities and differences (Saldana, 2016: 817).

Using the three coding techniques described above, I conducted line-by-line coding. **Table 5-5** presents selected excerpts from the interview transcripts and illustrates how the three coding techniques (in vivo coding, process coding, and initial or open coding) were applied during this phase.

Coding methods	Selected interview transcripts	Codes
In vivo coding	*“But now, voice recognition software just in the last year, even has become very good.”*	Voice recognition software
Process coding	*“For the hearing-impaired audience, understanding whether the music is cheerful or tense is more helpful than knowing the specific genre.”*	Delivering mood or atmosphere
Initial/open coding	*“Previously, when we were working for the China Disabled Persons’ Federation, there weren’t any specific standards provided.”*	Lack of official SDH guidelines in China

Table 5-5. Examples of first-cycle coding

Although the three coding techniques were used to code the data, all codes generated are provisional and subject to continuous refinement as the analysis progresses (Saldana, 2016: 333). This point was reflected in my own process of coding. When coding data in relation to the process of SDH production, one participant stated, “So generally, what we do is we use an online software called OOONA, and you can import your video files into OOONA.” I coded this segment as “OOONA”. Later on, another participant mentioned a different software tool, saying, “For other projects, we do use Aegisub, such as with the YYeTs subtitling group.” I coded this segment as “Aegisub”. Nevertheless, as the analysis unfolded, additional subtitling software emerged in the interviews. To prevent over-fragmentation of codes, I

grouped “OOONA”, “Aegisub”, and other subtitling software into a broader code labelled “subtitling software”.

This iterative and cyclical process occurred frequently during open coding. Although coding is typically conducted in a linear manner, new and similar concepts often emerged, requiring me to revisit earlier transcripts to revise and refine the codes. Initially, I attempted to code the data by following the sequence of participants, and started with the transcription of the first participant. At that point, an unexpected challenge arose. Each participant had answered ten questions covering various topics, and it became difficult to track and recall similar codes under the same question across different interviews. To address this, I adjusted my approach by coding one question at a time. More practically, I coded the responses to a single question from all nine participants before moving on to the next question. In doing so, I found it easier to identify similar concepts within each question and to refine and organise the codes systematically.

Incorporating this adjustment to code by question, I adopted in vivo coding, process coding, and initial/open coding techniques (Saldana, 2016) to code the data, following the sequence of IQs. The whole process was conducted manually, which, although admittedly time-consuming and labour-intensive, proved to be valuable as it enabled me to gain additional insights that might have been overlooked during the initial interviews. Finally, a total of 260 codes were initially generated from the original 88 pages of interview transcripts, with the full list provided in Appendix D.

5.4.4.3 Second cycle coding

Following Saldana's (2016) data coding method (as outlined in **Table 5-4**), I embarked on the second cycle of coding, known as "axial coding.” The term takes its name from the “axis”, representing the central categories identified at this stage, around which related subcategories are organised and connected (Saldana, 2016: 691). Incidentally, while the term “axial coding” is widely used, some grounded theorists, such as Charmaz (2006; 2014),

prefer the term “focused coding” to describe a similar process of developing relationships between codes and refining categories.

Although it would have been logical to continue following Saldana's data coding method, it is important to note that Saldana (2016) primarily provides guidance for the first cycle of coding, offering limited details on specific coding techniques for the second cycle. To supplement Saldana’s method, Boeije (2010: 108-109) introduces a set of practical steps that serve as a valuable reference for axial coding. These steps include:

1) Determine whether the codes developed thus far cover the data sufficiently and create new ones when the data provide incentives to do so.
2) Check whether each fragment has been coded properly, or if it should be assigned a different code.
3) Decide which code is most suitable if synonyms have been used to create two equal categories, and merge the categories.
4) Look at the overview of fragments assigned to a certain code. Consider their similarities and differences.
5) Subdivide categories if necessary.
6) Look for evidence for distinguishing main codes and sub-codes and assign the sub-codes to the main code.

In addition to these specific steps, Boeije (2010: 109) proposes two primary goals for this stage of analysis: first, “to determine which elements in the research are dominant and which are less significant”; and second, “to reduce and reorganise the data set: synonyms are crossed out, redundant codes are removed, and the best representative codes are selected.” Thus, it is often said that, unlike the first coding stage, which mainly fragments and breaks down the data, the second stage needs to reconnect and integrate the data (Birks & Mills, 2015: 27).

Guided by the two goals and the above steps, the second cycle of coding was carried out to review and refine the codes generated from the first cycle. Low-frequency codes (only once) entirely unrelated to the research topic were excluded (e.g., code “LGBT community”).

Codes with different phrasings but identical meanings were consolidated into a single code (e.g., “no official standard in China” and “no national-level guidelines” were combined into “lack of SDH guidelines in China”). Codes with the same core meaning were abstracted into a broader code (e.g., “limitations of voice recognition software” and “constrained by subtitling software” were grouped into “software limitations”). After this review and refinement, a total of 149 valid codes were identified, and the full list can be found in Appendix D.

Thus far, the second cycle of coding remains incomplete, as it requires developing codes into categories, as explained by Boeije (2010) in their coding steps (step 3 and step 5). A *category*, defined as “a higher-level concept that represents a group of codes” (Birks & Mills, 2015: 237), is used to group similar codes together. The next step, therefore, is to sort the codes into categories by examining their similarities and connections. To achieve this, I set aside the structure of the IQs and examined all codes collectively to identify relationships among them. This process ultimately resulted in 12 categories, and the full list is provided in Appendix D.

5.4.4.4 Third cycle coding

The third cycle of coding marks the final stage in the Grounded Theory analysis process. In some Grounded Theory literature, this stage is referred to as “selective coding” (Strauss, 1987; Boeije, 2010) or “theoretical coding” (Charmaz, 2006; Saldana, 2016). Of these publications, Saldana (2016) argues that the term “theoretical coding” is more appropriate as it more accurately reflects the outcome of this analytical phase. Through theoretical coding, a theoretical explanation of the phenomenon is developed by systematically integrating categories into core themes.

Speaking of the core theme, it is defined as “a concept that encapsulates a phenomenon apparent in the categories and sub-categories constructed and the relationships between these” (Birks & Mills, 2015: 238). At this stage, Boeije (2010: 118) explains that, researchers need to identify the core theme, explore its relationships with categories through a

"storyline", interpret the findings within the context of existing literature, address the RQs, and ultimately draw meaningful conclusions.

Similar to the second stage, the third coding stage is also an abstract process and cannot rely on specific coding techniques as in the first stage. Boeije (2010: 116) suggests that, during this phase, it is helpful to address the following questions. When researchers are able to answer these questions, it indicates that they have developed an understanding of the relationships within the data:

- Which themes have turned up repeatedly in the observations?
- What is the main message that the participants have tried to bring across?
- How are the various relevant themes related?
- What is important for the description (What) and the understanding (Why) of the. participant's perspective and behaviour?

With these questions in mind, I further examined and compared 12 categories while constantly revisiting the original data. In response to RQ1, which investigates SDH subtitlers' decision-making processes, six key themes were identified as the main factors influencing their decisions, as illustrated in **Figure 5-1**. Meanwhile, the storyline gradually emerged during the analysis: in creating SDH, subtitlers begin by following SDH guidelines and standards while addressing the technical and practical constraints they face. They also take into account the diverse needs of audiences, including variations in age, cultural background, and cognitive abilities. While subtitlers strive for objectivity in their decisions, some level of subjectivity is inevitable, especially when encountering the challenges posed by video content. Moreover, their decision-making process is not static and may evolve through workflow optimisation and the integration of collaboration and feedback mechanisms.

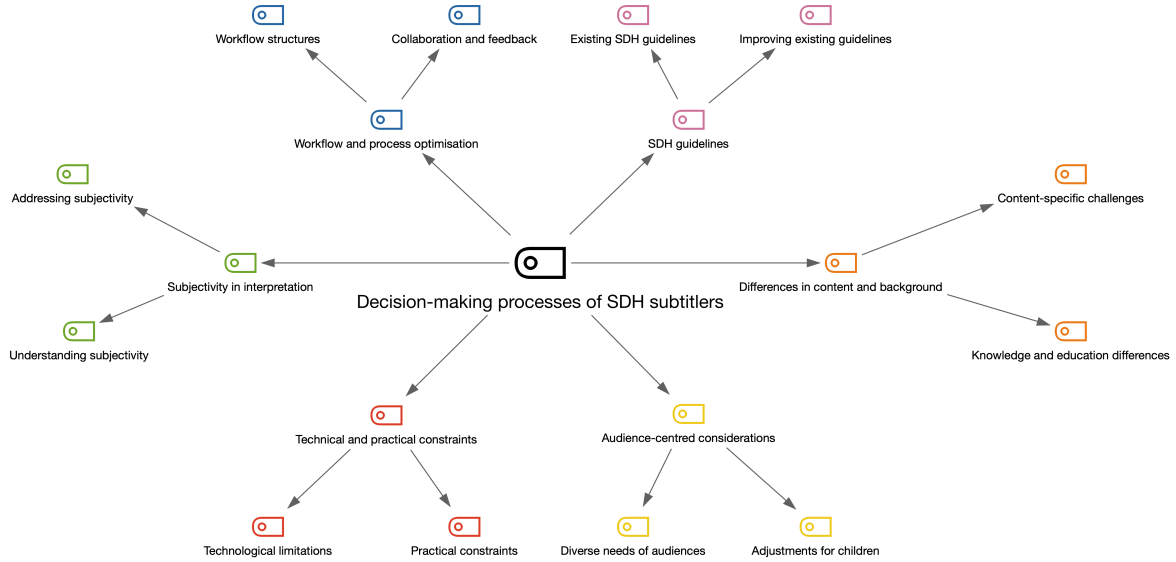


Figure 5-1. Model of the decision-making process of SDH subtitlers

Up to this point, the three cycles of coding have been completed, resulting in a theoretical model that explains the factors influencing the decision-making processes of SDH subtitlers. This marks the conclusion of the qualitative data analysis. The themes identified will be presented and discussed in detail in Chapter 6.

5.5 Phase II: Reception Study with SDH users

5.5.1 Participants

5.5.1.1 Chinese DHH individuals

According to the China Disabled Persons' Federation (CDPF), approximately 85.02 million persons with disabilities were recorded in China by the end of 2010, accounting for 6.34% of the national population. This figure is based on data from the Second National Sample Survey on Disability conducted in 2006 and the Sixth National Population Census in 2010. Although no updated official statistics have been released since then, the current number is likely to be higher, given that China's overall population has been on the rise. Among the different types of disabilities, physical disabilities were the most prevalent (24.72 million), followed by hearing (20.54 million), multiple disabilities (13.86 million), visual (12.63

million), mental (6.29 million), intellectual (5.68 million), and speech impairments (1.30 million), as reported by the CDPF (2021).

What stands out is that the population of DHH individuals represents the second-largest disability group in China, making up approximately 24.16% of the total population with disabilities. Within this large population, the degree of hearing loss varies widely, ranging from moderate to complete, all of which are recognised as hearing disabilities with a medical diagnosis. To classify the level of hearing impairments accurately, the *Practical Evaluation Criteria for Persons with Disabilities in China*, issued in 2006, classified hearing disabilities into four levels, as shown in **Table 5-6**.

Level	Hearing Loss (dB SPL[3])	Speech Recognition Rate (%)
Level 1 (Complete)	＞90	＜15
Level 2 (Profound)	71-90	15-30
Level 3 (Severe)	61-70	31-60
Level 4 (Moderate)	51-60	61-70

Table 5-6. Classification of hearing impairment levels used in China

These four levels (Levels 1 to 4) are noted in the serial number of the disability identification card used in China. First introduced in 1995, this card serves as the official registration and identification system for persons with disabilities in China (Hao & Li, 2020: 4). Each card carries a 20-digit standardised number, composed of the individual's 18-digit national ID number, followed by a single-digit code (19th) indicating the disability type and another digit (20th) representing the severity level. It is worth noting that the classification of both the disability type and severity level is diagnosed by designated medical professionals.

[3] dB SPL: Hearing loss is measured in decibels sound pressure level (dB SPL). Higher values correspond to more severe hearing loss (Dobie & Van Hemel, 2005).

Figure 5-2 illustrates the structure of this coding system. The main reason for introducing the disability identification card here is that it is directly related to the criteria for participant recruitment, which will be discussed in the next section.

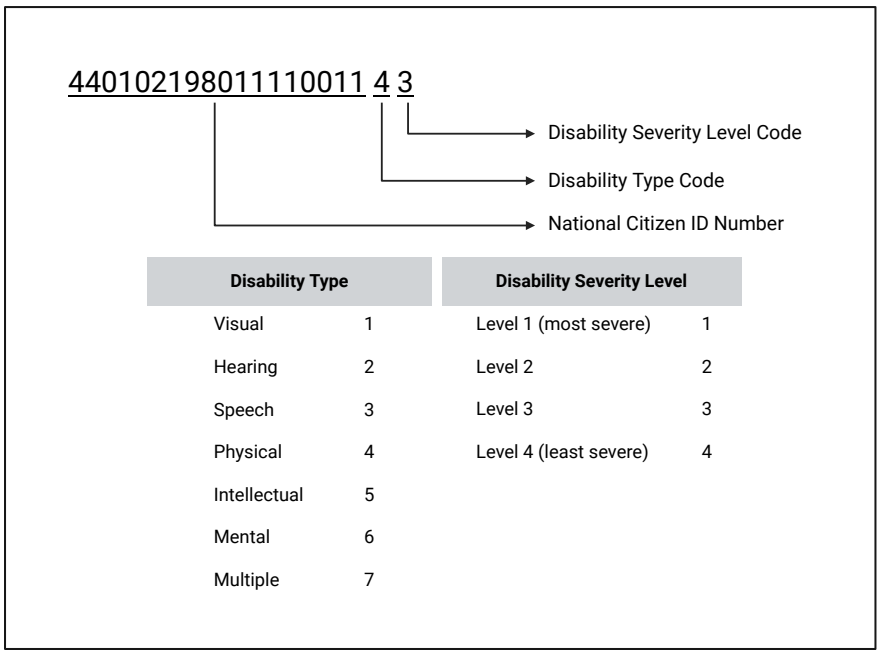


Figure 5-2. Coding system of China's disability identification card[4]

5.5.1.2 Recruitment and profiles

Before recruiting participants for Phase II of the reception study, three inclusion criteria were set out:

1) Participants had to be hearing-impaired Chinese citizens holding a valid disability identification card;
2) Participants had to be native Chinese speakers with at least nine years of compulsory education;
3) Participants had to be at least 18 years old at the time of the experiment.

[4] **Figure 5-2.** Adapted and translated from the People's Government of Chongqing (2021).

After defining the inclusion criteria, it was necessary to determine an appropriate sample size for the reception study. To this end, the study drew on recommendations from experimental research in the field of AVT. According to Orero et al. (2018: 110), "sample sizes of fewer than 25 per group are unlikely to yield statistical power." As further noted by Orero et al. (*ibid.*), data attrition is common in experimental studies (e.g., recruiting 30 participants, with the expectation that data from approximately five may be incomplete or unusable). As such, it is advisable to over-recruit participants to ensure that a sufficient number of complete datasets is obtained. In light of these two considerations, the target sample size for the present study was set at no fewer than 30 participants.

To recruit eligible participants, purposive sampling was used by directly contacting a special education school in China. The school enrols a large number of students with various types of disabilities, including those with hearing loss. According to information provided by the school, it admits students who have completed nine years of compulsory education, indicating that they generally have basic literacy skills and are able to read Chinese characters. Regarding the age criterion, the school explained that students with disabilities often begin their education later than their non-disabled peers, which means that although they may be enrolled in high school-level age ranges, in fact many were over the age of 18. To ensure that all participants met the age requirement, participants were recruited within the age range of 18 to 25.

Ultimately, a total of 38 DHH participants were recruited from the school, comprising 20 males (52.63%) and 18 females (47.37%). Their ages ranged from 18 to 25 years (M = 20.11, SD = 1.71). More detailed demographic information about the participants is presented in Section 7.2.1.

5.5.2 Tools and materials

5.5.2.1 Eye-tracking

This section is organised into four parts related to eye-tracking: previous eye-tracking studies in AVT, experimental setup, AOIs and TOIs, and the eye-tracking metrics used in this study.

Eye-tracking in AVT studies

The use of eye-tracking, as well as other experimental tools such as electroencephalography and heart-rate monitors in AVT studies, has been identified as a clear indicator that AVT is evolving into an evidence-based discipline, marking a shift from descriptivism towards experimental research (Orero et al., 2018; Díaz-Cintas & Szarkowska, 2020; Kruger & Liao, 2022). Indeed, Arnáiz-Uzquiza (2010: 171) observes that earlier reception studies in AVT primarily relied on interviews and questionnaires to gather user feedback, but the results obtained through these methods were often subjective and, at times, imprecise. Nowadays, advanced instruments such as eye-trackers, combined with sophisticated research designs (e.g., the use of linear mixed-effects models, as demonstrated in Silva et al., 2022), encourage a growing number of researchers to collect empirical evidence that yields more robust and reliable results.

Among numerous AVT studies using eye-tracking, subtitling has become a prominent and resource-intensive topic, especially when compared to dubbing. This is largely because subtitling necessarily requires viewers to divide their visual attention between the subtitles and the on-screen content to construct meaning. To capture this complex allocation of attention, eye-tracking has proved to be a well-suited instrument that helps to reveal how viewers process such multimodal texts (Sasamoto, 2024: 102-103).

Recognising the potential of eye-tracking, a growing body of studies have used this method to explore how viewers process subtitles, not limited to conventional subtitles (e.g., interlingual subtitles, intralingual subtitles), but also non-conventional subtitles (e.g.,

fansub subtitles, impact captions, SDH subtitles). In the context of conventional subtitling, Liao et al. (2022) used eye-tracking technology to examine how audio information influences viewers' reading behaviours when processing intralingual versus interlingual subtitles. The results show that viewers spent more time fixating on the subtitle area when watching interlingual subtitles compared to intralingual ones. On the side of non-conventional subtitles, Orrego-Carmona (2015), in his PhD thesis, conducted an eye-tracking study combined with questionnaires and interviews to compare professional and non-professional subtitles. He found that professional subtitles led to better comprehension, while participant satisfaction did not significantly differ between the two types. In terms of impact captions (also known as *telop*[5]), a novel form of subtitles, limited research attention has been paid to date. Nonetheless, a few empirical studies incorporating eye-tracking have been published, including those by O'Hagan and Sasamoto (2016), Sikkema (2020), and Sasamoto et al. (2021). Among these, the study by O'Hagan and Sasamoto (2016) stands out for its use of a wearable eye-tracker to investigate viewers' real-time responses to impact captions in a popular entertainment show. The findings indicate that although most participants reported reading all the captions, they tended to fixate more on facial areas than on the caption area when faces appeared on screen.

SDH subtitles, also a type of non-conventional subtitles and the focus of the present study, have likewise been investigated by AVT scholars using eye-tracking technology. One of the largest large-scale reception studies on SDH was the DTV4ALL (Digital Television for All) project conducted in Europe, coordinated by Romero-Fresco (2015). In this project, eye-tracking data were collected from 103 DHH and hearing participants across Poland, Spain, Italy, and Germany to explore how they allocated their visual attention when processing SDH subtitles. Together with data from questionnaires, the findings from this

[5] "Telop" refers to open captions commonly used in Japanese television, often featuring colourful, large-font text to highlight key dialogue or information, primarily for entertainment purposes rather than accessibility (O'Hagan & Sasamoto, 2016).

project provided empirical recommendations for improving SDH practices and standards in the participating countries.

Given the widespread application of eye-tracking in subtitling research, including studies on SDH, eye-tracking devices are also employed in this study, motivated by the following considerations: 1) all participants are DHH individuals who, unlike hearing viewers, primarily rely on visual input to access AV content; eye-tracking makes it possible to capture, in real time, how they distribute their attention and engage with what appears on screen; and 2) the experiment compares SDH subtitles with ST subtitles. Since SDH typically carries more information, it has the potential to increase cognitive effort. Eye-tracking can help determine whether this is the case and, if so, guide improvements in future SDH practice. For these reasons, eye-tracking was used in the experiment.

Eye-tracking setup

The eye-tracker used in this study was the Tobii Pro Nano, a portable, screen-based device that records gaze data at a sampling rate of 60 Hz. It was attached directly to the bottom edge of a 15.5-inch laptop screen with a resolution of 1,920 × 1,080 pixels. The laptop, a high-performance gaming model running Windows 11, had Tobii Pro Lab software pre-installed (version 1.217) and was used throughout for data collection.

All eye-tracking sessions were carried out in a quiet classroom provided by the special education school. Following Saldanha and O'Brien's (2014: 138-139) suggestion, the eye-tracker was set up in a fixed location where environmental factors such as lighting and sound could be controlled. Participants were seated individually at a fixed distance from the screen (approximately 50-60 cm), which falls within the operational range of 45-85 cm (depending on screen size), as specified in the *Tobii Pro Nano User Manual* (Tobii Technology, 2023). Each session took place with only the participant and the researcher (myself) in the room. To avoid interfering with the participant's focus, I stayed behind them and out of their direct line of sight throughout the recording.

AOIs and TOIs

In eye-tracking research, areas of interest (AOIs) are typically used to delineate specific regions of the stimulus for closer examination. More importantly, by drawing AOIs, researchers can extract eye-tracking data for different regions, making it possible to examine how viewers distribute their visual attention across the screen (Holmqvist et al., 2011: 264).

In this study, two AOIs were defined: the subtitle area (Subtitle-AOI), located at the bottom of the screen, and the full-screen area (Full screen-AOI), as shown in **Table 5-7**. The Subtitle-AOI was used to capture participants' eye movements on subtitles, and the Full screen-AOI was used to compare attention to subtitles with attention to the overall screen. For comparability, the same AOIs were applied to both video versions, one with SDH subtitles and the other with ST subtitles.

In addition to AOIs, times of interest (TOIs) were also defined in this study. TOIs are "the intervals of the recording when your stimulus or behaviour of interest are predicted to occur" (Tobii Technology, 2021: 107) and can be divided into two types: automatic TOIs, which are system-generated, and custom TOIs, which are created manually by the user (Tobii Technology, 2021: 90).

Although TOIs are rarely addressed in eye-tracking research, they are relevant in the present study because, as outlined in Section 5.5.2.2 (AV Stimuli), the experimental material included not only the main video but also an introductory prompt and a closing prompt. The function of these prompts was merely to inform viewers that the video was about to start or had just ended. As they were presented as plain text on a black background without any images or subtitles, they were excluded from the analysis. Unlike the prompts, the main video contained both visual scenes and subtitles, which were the actual focus of analysis. Therefore, the TOI for this study was manually defined to cover only the main video, spanning from 0:00:20.00 to 0:19:44.00, with a total duration of 19 minutes and 24 seconds. As with the AOIs, the same TOI was applied to both video versions.

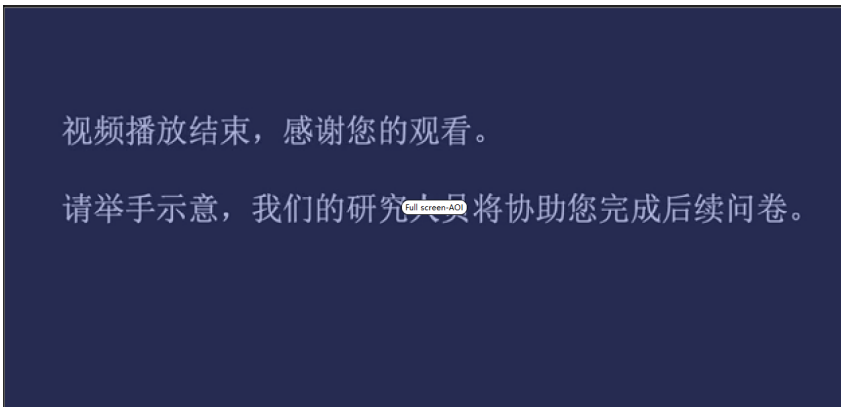
Two AOIs	
Subtitle-AOI (red shaded area)	
Full screen-AOI (blue shaded area)	

Table 5-7. AOIs defined for the study

Eye-tracking metrics

According to Rayner (1998: 373) and Holmqvist (2011: 213-214), the tradition of eye-tracking research has mainly revolved around two events: *fixations* and *saccades*. *Fixations* are relatively stable periods, typically lasting 200-300 milliseconds, during which visual information is processed. *Saccades*, in turn, are rapid shifts between fixation points, lasting about 30-40 milliseconds, which allow the eyes to move quickly from one point of interest to the next, such as from word to word in subtitle reading (Rayner, 1998: 373; Doherty & Kruger, 2018: 49). To aid understanding, **Figure 5-3** illustrates how fixations and saccades occur during subtitle viewing. As shown, the red circles denote fixation points, where the viewer's gaze remains relatively stable to extract visual information, and the lines connecting them are saccades, the rapid eye movements that link one fixation to the next.

In addition to fixation- and saccade-based measures, pupil-based metrics have also been applied in eye-tracking research. Pupil dilation, as an involuntary physiological response, reflects changes in arousal and cognitive effort, and has been accepted as a valid indicator of cognitive load (Doherty & Kruger, 2018: 49). Given that RQ3 addresses viewer's gaze patterns, three fixation-based measures (proportion of fixation duration, average fixation duration, and number of fixations) and one pupil-based measure (average pupil diameter) were selected for analysis. What follows is a detailed introduction to each of these four measures.



Figure 5-3. Example of fixations and saccades

- **Percentage of fixation duration (unit of measurement: %)**

This metric refers to the proportion of a participant's total fixation time spent on the subtitle area relative to the total fixation time on the entire screen. For example, if a viewer accumulates 50 seconds of fixation on the entire screen and 10 seconds on the subtitle area, the percentage is 20%. The calculation formula is provided as below:

$$Percentage = \frac{Total\ duration\ of\ fixations\ on\ subtitle\ area}{Total\ duration\ of\ fixations\ on\ the\ entire\ screen} \times 100\%$$

So why not simply use the total fixation duration on the subtitle area alone? The main reason is that the absolute total duration of fixations on subtitles may not account for differences in overall viewing time. Suppose two viewers both fixate on the subtitle area for

10 seconds while watching a 60-second clip: if one spends 50 seconds fixating on the screen overall and the other only 40 seconds, their percentage of fixation duration will differ, reflecting relative attention allocation more accurately. In other words, compared with total fixation duration, the percentage measure can normalise for overall viewing time, thereby reducing errors and improving comparability across participants. Moreover, the percentage measure can better represent relative attention allocation between two subtitle conditions, rather than simply recording "how long" someone looked. Finally, an additional advantage is that the percentage measure is bounded between 0 and 100%, and provides a unified scale that yields clearer, more interpretable results.

This proportional measure requires some extra calculation, which may explain why it is less widely adopted. Despite this, it has already been used in AVT eye-tracking reception studies. For instance, Black (2022) analysed percentage fixation time (and percentage fixation count) on subtitle and image regions to explore how integrated subtitles, compared to standard ones, guided viewers' attention. Another example can be found in Chen's (2024) doctoral thesis, where the "dwell time percentage of global dwell time" was computed, defined as the proportion of dwell time within a subtitle AOI or an image AOI relative to dwell time in the global screen area, to observe how attention was allocated between subtitles and visual content.

- **Average duration of fixations (unit of measurement: milliseconds)**

This is the term used in Tobii software and corresponds to what most eye-tracking studies call *mean fixation duration*. According to the *Tobii Pro Lab User Manual* (Tobii Technology, 2021), it is defined as "the average duration of the fixations inside an AOI during an interval." In the eye-tracking literature, it is widely recognised as a proxy for depth of processing and cognitive engagement with text, with higher values generally indicating more effortful processing and lower values indicating the opposite (Szarkowska et al., 2016; Fox, 2018; Orrego-Carmona et al., 2018; Liao et al., 2020; Szarkowska et al., 2024b).

- **Number of fixations (unit of measurement: count)**

Often termed *fixation counts* (Kruger et al., 2015; Doherty & Kruger, 2018), this metric refers to the total instances in which the eyes paused to process information within the AOI across a given segment. In Tobii Pro Lab, this index can be visualised through heat maps (see **Figure 5-4**), where colours indicate the number or duration of fixations within specific areas of the stimulus. Red usually represents the highest values, green the lowest, with intermediate levels shown in between (Tobii Technology, 2021: 82).

As a coarse indicator of attentional allocation, it is typically interpreted alongside duration-based indices to infer processing effort in AVT eye-tracking research. Generally speaking, more fixations point to sustained or repeated reading of subtitles and thus greater allocation of visual resources to text, which implies increased processing effort, especially when accompanied by longer average durations of fixations or greater total reading time. This interpretation has been supported in AVT eye-tracking studies (Law & Kruger, 2018; Szarkowska & Gerber-Morón, 2018a; Jiang & Doherty, 2020).

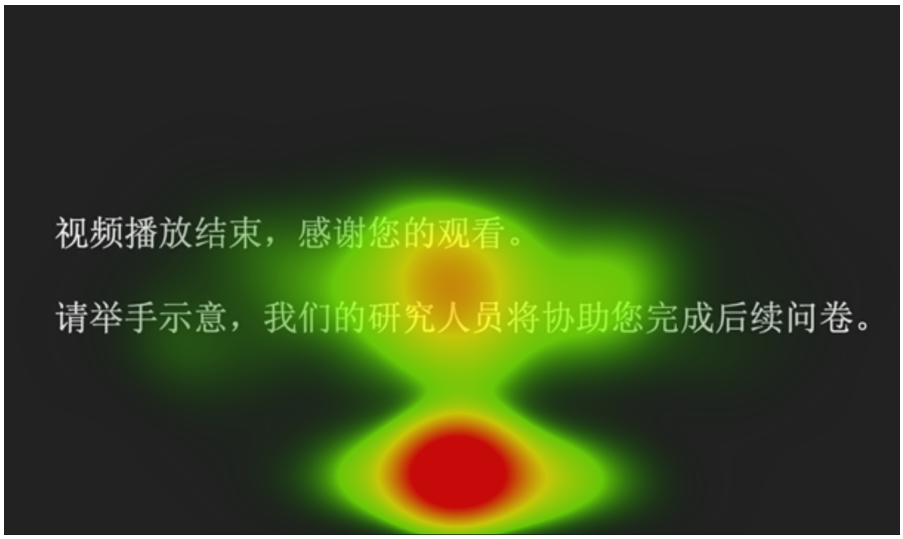


Figure 5-4. Example of a heatmap

- **Average pupil diameter (unit of measurement: millimetres)**

This measure records the mean pupil size during fixations within a defined AOI over a given interval. In the eye-tracking literature, pupil diameter has been recognised as a reliable

indicator of cognitive load (O'Brien, 2006; Kruger & Doherty, 2016; Doherty & Kruger, 2018; Teixeira & O'Brien, 2018). Empirical evidence shows that when tasks become more difficult, pupils typically dilate because greater cognitive effort is required to complete the task (Hyönä et al., 1995). It is worth noting, however, that pupil size can be influenced by various external factors, the most frequently mentioned of which is luminosity (Kruger et al., 2013: 63; Saldanha & O'Brien, 2014: 138; Kruger et al., 2018: 276). Given this limitation of the pupillometric method, the present study used average pupil diameter to complement three fixation-based indices and to make cautious inferences about viewers' cognitive responses.

5.5.2.2 AV stimuli

The AV stimuli (19 minutes and 59 seconds) used in the experiment consisted of three components, presented in the following sequence: 1) an introductory prompt; 2) the main video; and 3) a closing prompt. All three components were edited into a single continuous video, rather than separate clips.

1) The introductory prompt is a 20-second static screen with white text on a black background, as shown in **Table 5-8**. It instructs participants to keep their heads still and look straight at the screen throughout the viewing to ensure accurate eye-tracking results, and informs them that they are about to watch a video.

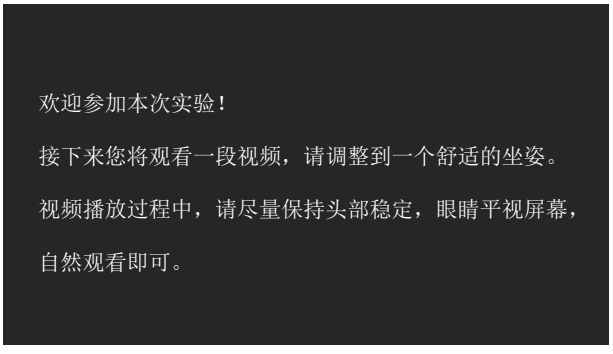
Introductory prompt (0:00:00.00"0:00:20.00, 20 seconds)	
	*(in English):* Welcome to the experiment! You are about to watch a video, please adjust yourself to a comfortable sitting position. During the video, try to keep your head still, look straight at the screen, and watch naturally.

Table 5-8. Introductory prompt in the AV stimuli

2) The main video used in this study was the first episode of a Chinese animated series titled *Yao – Chinese Folktales*, co-produced by Shanghai Animation Film Studio and Bilibili in 2023, and also available on YouTube. The series has eight stand-alone episodes, each approximately 20 minutes in length. The decision to work with an animated series was informed by findings from the Phase I interviews. SDH subtitlers pointed out that certain types of AV content, such as talk shows, reality shows, and variety shows, are more demanding for SDH production due to their fast-paced dialogue, frequent overlapping speech, and many unscripted moments. Thus, an animated series with relatively simple content was chosen as the AV material for the reception study.

For the experiment, the first episode, called *Nobody*, was selected. It tells the story of monsters inspired by the classic Chinese novel *Journey to the West.* The original runtime of this episode was 22 minutes and 32 seconds, but both the prolonged opening and ending credits were removed to minimise participant fatigue. As a result, the final version used in the experiment had a total duration of 19 minutes and 24 seconds. For the purposes of the experiment, two versions of the main video were created: one with SDH subtitles and the other with ST subtitles, as presented in **Table 5-9**.

Main video (0:00:20.00"0:19:44.00, 19 minutes and 24 seconds)	
SDH subtitles version	ST subtitles version
	
(Yelling) It's dangerous here.	It's dangerous here.

Table 5-9. Two subtitle versions of the main video in the AV stimuli

3) The closing prompt was a 15-second still screen, identical in format to the introductory prompt, displaying white text on a black background, as shown in **Table 5-10**. It asks participants to remain seated, raise their hand to notify the researcher, and refrain from moving their head to avoid affecting the tracking ratio.

Closing prompt (0:19:44.00"0:19:59.00, 15 seconds)	
视频播放结束，感谢您的观看。 请举手示意，我们的研究人员将协助您完成后续问卷。	*(in English):* The video has ended. Thank you for watching. Please raise your hand to signal the researcher, who will assist you with the post-task questionnaire.

Table 5-10. Closing prompt in the AV stimuli

5.5.2.3 Subtitles

Two subtitle versions were prepared for the main video: an ST subtitles version and an SDH subtitles version (see **Table 5-9**). Both were in Chinese language, with the original soundtrack also in Chinese. It was intended to obtain the original subtitle files directly from the production company, as suggested by implications from Phase I (see Section 6.3), which suggested that doing so could make subtitle production easier and reduce transcription time. Nevertheless, I failed in my attempt to obtain the files, and the subtitles were therefore created from scratch.

As both subtitle versions were newly prepared for the experiment, a set of technical parameters was established to ensure consistency across the two versions. With reference to the five SDH guidelines discussed in Chapter 4, the same basic parameters were adopted for both the ST and SDH versions: a sans-serif font, Arial; a font size of 50; a subtitle speed below 11 cps; a maximum of 18 characters per line; and no more than three lines per subtitle.

ST subtitles

For the ST subtitle version, the original video available on video platforms featured both Chinese and English subtitles, with the Chinese text on the first line and the English text on the second. To create a version with only Chinese subtitles, the video needed to be re-edited: the English subtitles were removed, and the Chinese subtitles were retained. But as the original subtitles were burned-in (or hard-coded) and could not be removed or switched off (Díaz-Cintas & Remael, 2021: 26), the original file could not be used directly. Because of this, the Chinese subtitles were extracted using the Subtitle Extractor tool, and a new SRT file was created. The file was then manually revised to correct typographical errors and adjust the timing using Aegisub (a subtitle editing software). Finally, the subtitles were embedded into the video using Arctime Pro (a video editing software). The video was also slightly edited, as in the original release the subtitles were displayed directly over the image. To accommodate the needs of the experiment, the subtitle and image areas were separated: a dedicated black bar was added to display the subtitles, which in turn provided a clear distinction between the subtitle area and the main image (see **Table 5-9**).

SDH subtitles

In preparing the SDH subtitles version, a series of adaptations were made. As outlined in Section 2.4.4.2, five categories of non-verbal audio information are identified in SDH: **music, sound effects, paralinguistic information, character identification, and silence**. Accordingly, the SDH subtitles in this experiment were designed with these five aspects in mind.

For **music**, some SDH guidelines, such as the CAB Guidelines (2012) and the DCMP Guidelines (2019), recommend that if a song includes lyrics, these should be written out in the subtitles, often with musical symbols. In this video, however, the songs are instrumental background music with no lyrics. Taking into account the suggestions from the SDH subtitlers interviewed in Phase I, background music in this case was described through words that convey its atmosphere and emotional tone. Based on this, when describing

background music in the clip, phrases such as [tense and uneasy music] were used, and in some scenes, this was expanded to [tense and uneasy music builds] or [tense and uneasy music continues] to reflect its narrative effect.

Regarding **sound effects**, while some guidelines, including the DCMP Guidelines (2019) and the BBC Guidelines (2026), encourage the use of onomatopoeia, sometimes in combination with descriptive labels to bring sound effects to life, this approach is not universally supported. Existing research (e.g., Pereira, 2010; Gürkan, 2019) reveals that viewers born deaf, who may lack direct auditory reference, often benefit more from descriptive expressions that explain the mood or function of the sound. This view was also shared by the SDH subtitlers interviewed in Phase I, who generally preferred descriptions over phonetic renderings. Reflecting this, the sound effects in the experiment were rendered through concise descriptive phrases such as [chopping sound], [cutting sound], [scraping sound], and [rubbing sound], all of which followed one another lasting a few seconds in the video.

With respect to **paralinguistic information** (including accent, intonation, and emotion), the subtitlers interviewed in Phase I addressed how to convey the speaker's emotions. Most of them were not in favour of using creative methods such as emojis to express emotions; instead, they preferred to describe them in words. This approach was reflected in the line "[Shouting] Ow, what are you doing?", where the tag [shouting] was used before the dialogue to indicate the speaker's emotional reaction. The scene was also included in the post-task comprehension test (Question 2) to assess whether participants noticed and understood the intended intonation. In addition, existing guidelines, particularly the BBC Guidelines (2026), offer detailed recommendations for handling paralinguistic features, including the use of punctuation to convey tone and emotion. Following that, in the subtitle "[Doubtful] Is Sun Wukong also a monster?", a question mark was added at the end to reinforce the speaker's incredulous tone. This scene was also included in Question 6 of the comprehension test to check whether participants could understand the emotion through the use of the question mark.

Character identification was not specifically addressed in the Phase I interviews. Common methods include assigning different colours to different speakers, adding a label with the speaker's name before the subtitle, and positioning the subtitle directly under the speaker. However, based on surveys of DHH viewers conducted by many scholars (Pereira, 2010; Romero-Fresco, 2015; Tamayo, 2016; Agulló & Matamala, 2019; Cepak & Corrius, 2023), the colour-coded system is generally considered the most preferable method.

In this experiment, colour-coding was also used to distinguish between speakers: yellow is used for the main character or the character with the highest density of dialogue; green for the second main character; cyan for the third; magenta for the fourth; and white for other characters. Nevertheless, since there were 11 additional speakers falling into the "other characters" category, and in some scenes the speakers marked in white were off-screen, speaker labels were added before the subtitles to identify them. To sum up, the experiment adopted a combination of colour coding (yellow, green, cyan, magenta, white, in this order) and speaker labels (white) to denote different speakers.

When it comes to subtitling **silence** in SDH, Zárate (2021: 81) states that it is generally subtitled in the most suspenseful moments of a film. In the video used for this study, however, there were no such sequences, so long absences of sound were not marked in the subtitles. Even so, subtitling filmic silence can be necessary in certain cases.

Furthermore, my approach to creating the SDH subtitles was also guided by relevance theory, as discussed in Section 4.3.4. At certain points in the subtitling process, when deciding what kinds of non-verbal audio information to include and how to present them, I tried to balance their likely cognitive value against the effort they might require from viewers. For instance, in one scene, a character speaks sarcastically, but the literal line on its own could easily be misunderstood. To make the intended meaning clearer, I added a brief description [Reply coldly] before the line. This was done on the understanding that the interpretive benefit of clarifying the tone was likely to outweigh the minimal extra reading effort.

Lastly, to minimise personal subjectivity and ensure the quality of the SDH subtitles, the initial version I created was sent to an SDH subtitler from the Chinese accessibility company who had been interviewed in Phase I. This decision was informed by implications from Phase I (see Section 6.3), where some organisations adopted peer review systems to reduce individual interpretation. The subtitler reviewed the SDH subtitles and provided suggestions for improvement. The changes mainly fell into two areas. First, some sound descriptions that were already visible on screen or had little impact on the plot were removed. This was consistent with the findings from Phase I. That said, there were a few exceptions: when there were long stretches without any subtitles, visible sounds that were less relevant to the plot were kept. Second, some of the wording was simplified. Since the material was an animated series, simpler expressions were considered a better match for the tone of the work. The final version was finalised through revision and discussion with the SDH subtitler.

5.5.2.4 Pre-task questionnaire

The pre-task questionnaire (see Appendix H) consisted of 15 questions designed to gather information on participants' basic demographic background, viewing habits, and subtitle preferences. The design of these questions was primarily informed by previous surveys conducted among DHH audiences, including Bartoll and Martínez Tejerina's research (2010) on the Spanish DHH population, Morettini's study (2012) of Italian DHH viewers, and a series of national surveys coordinated by Romero-Fresco (2015) across European countries.

More specifically, *Questions 1 to 7* aimed to collect participants' demographic characteristics, including gender, age, visual ability, degree of hearing loss, age of onset of hearing impairment, use of hearing assistive devices, and primary mode of daily communication. Notably, for the question on the degree of hearing loss, participants were instructed to refer to the final digit on their disability identification card, which indicates four levels of hearing impairment, with Level 1 representing the most severe and Level 4 the mildest. *Questions 8 and 9* examined participants' viewing habits, such as average daily television viewing time and preferred content types. *Questions 10 to 15* explored

participants' subtitle preferences, covering reading difficulty, the impact of subtitles on viewing, use of accessibility services, prior experience with SDH subtitles, helpfulness of SDH subtitles, and preferred types of sound information to be included.

With all questions prepared, the complete pre-task questionnaire was created using the Qualtrics platform. During this step of the experiment, participants completed the questionnaire on their own mobile phones by scanning a pre-prepared QR code generated by the Qualtrics system.

5.5.2.5 Post-task questionnaire

The post-task questionnaire (see Appendix I) comprised two parts: 1) a comprehension test to assess participants' understanding of the video content, and 2) a set of Likert-scale questions to evaluate their levels of engagement and enjoyment regarding the viewing experience.

In the comprehension test, there were 10 single-choice questions. The test design was informed by previous SDH reception studies, such as Szarkowska et al. (2016), which used a similar test to compare the comprehension of edited versus verbatim subtitles among Polish deaf and hearing viewers. Inspired by the work of Szarkowska et al. (2016), each question in the comprehension test was carefully designed to assess how well participants understood information that needed to be deduced from the subtitles and was barely inferable from the visual content. For example, *Question 4* asked participants to identify the speaker of a specific line and describe their tone. However, the speaker did not appear on screen, and their tone could not be inferred from the visuals alone. The correct answer could be easily found in the SDH subtitles version, where the speaker was identified using a colour-coded system and the tone was described with short textual descriptions.

In the Likert-scale questionnaire, there were 15 questions designed to assess participants' responses regarding narrative engagement (e.g., "I was emotionally moved by the ending of the story in this video") and enjoyment (e.g., "I found the video interesting, and it made me feel relaxed and happy"). These questions were developed by drawing on studies by

Guerberof-Arenas and Toral (2020), as well as Guerberof-Arenas et al. (2024). In terms of point design, according to Saldanha and O'Brien (2014: 157), a Likert scale typically offers five, seven, or nine response options to capture varying levels of agreement or disagreement. Among these, the 7-point scale was adopted in this study based on findings from Preston and Colman (2000) as well as Finstad (2010). Preston and Colman (2000), who compared a range of scale lengths, found that the 7-point scale outperforms the 9-point scale in terms of reliability, validity, and respondent preference. Finstad (2010), in a comparison between 5-point and 7-point scales, concluded that the 7-point format is more accurate, easier to use, and better reflects a respondent's true evaluation. Combining these findings, the 7-point scale was used in the Likert-scale questionnaire, and the response options were arranged from A to G as follows: strongly agree, agree, somewhat agree, neither agree nor disagree, somewhat disagree, disagree, and strongly disagree.

As with the pre-task questionnaire, the post-task questionnaire was also delivered via the Qualtrics platform, which participants completed independently on their mobile phones after the eye-tracking session.

5.5.3 Reception study process

The reception study was conducted in three steps:

1) **Reception study preparation:** The process began by contacting the special education school and obtaining consent to conduct the study on its premises. Subsequently, the AV stimuli were prepared, and both the pre-task and post-task questionnaires were designed. Then, the Plain Language Statement and the Informed Consent Form were developed as part of the ethical approval process.
2) **Conducting the reception study:** In this step, each participant first completed a pre-task questionnaire, followed by an eye-tracking session, and finally completed a post-task questionnaire.
3) **Data processing:** Following the experiment, both questionnaire responses and eye-tracking data were exported and organised for subsequent data analysis.

The following sections provide a detailed description of each step.

5.5.3.1 Step 1: Reception study preparation

To recruit eligible participants, I contacted a special education school and provided it with all relevant information about the study. In January 2025, the school issued a formal consent letter supporting the ethics application.

Next, I selected appropriate AV stimuli and created two versions of the same video clip: one with ST subtitles and the other with SDH subtitles. At the same time, I designed the pre-task and post-task questionnaires, taking into account the characteristics of the participants and the selected AV content. The Plain Language Statement and the Informed Consent Form were also prepared at this stage for the subsequent ethics application.

Finally, the ethics application was submitted on 6 March 2025, and approval was granted by Dublin City University's Research Ethics Committee on 16 April 2025 (Reference: DCUREC/2025/043; see Appendix G).

5.5.3.2 Step 2: Conducting the reception study

The reception study was carried out in May 2025. First, the school gathered 38 DHH students who had expressed interest in participating in the study in a classroom. With the help of a sign language teacher, I explained the experimental procedures to participants using a set of prepared slides. After the introduction, the students read and signed the Informed Consent Form. Once consent was obtained, each student was given a sticker with a unique identification number (e.g., 01, 02, 03). They then scanned a pre-prepared QR code to access the pre-task questionnaire on the Qualtrics platform using their mobile phones. Students were instructed to enter their assigned numbers in the questionnaire to ensure proper matching with subsequent data.

Based on the responses provided in the pre-task questionnaire regarding their levels of hearing impairments, I temporarily divided the 38 participants into two groups: one group was assigned to watch the video with ST subtitles, while the other group watched a version with SDH subtitles. It should be noted that this grouping was provisional, as the tracking

ratio of each participant's eye-tracking data was checked immediately after they completed the task. If the ratio was below 80%, a benchmark commonly used in previous subtitling eye-tracking studies (Orrego-Carmona et al., 2018; Szarkowska & Gerber-Morón, 2018b; Szarkowska et al., 2021), the data were excluded from further analysis. In total, six participants were excluded because their eye-tracking data did not meet this minimum tracking ratio threshold (80%). The final group arrangement and corresponding viewing conditions are summarised in **Table 5-11** below.

Group	n	Degree of hearing loss		Viewing condition
SDH group	17	Complete	8	Viewed the video with SDH subtitles, including both dialogue and non-verbal audio information.
		Profound	7	
		Severe	2	
		Moderate	0	
ST group	15	Complete	6	Viewed the same video with ST subtitles, containing dialogue only.
		Profound	7	
		Severe	2	
		Moderate	0	

Table 5-11. Summary of participant groups and viewing conditions (Phase II)

To ensure the smooth running of the experiment, I worked with school staff to create a timetable that scheduled 8 to 10 DHH students to participate each day. Each participant's session was expected to last approximately 40 to 60 minutes. As scheduled in the timetable, each participant attended the eye-tracking lab individually. They were guided to their designated seat, their assigned number was confirmed, and it was used to label their eye-tracking data file in Tobii Pro Lab. Next, the eye-tracking session began with a nine-point calibration, followed by viewing the assigned version of the AV stimuli. Once the video had finished, participants raised their hand, following the on-screen instruction, to indicate they were done. I then assisted them in filling out the post-task questionnaire. The

questionnaire was also completed on the Qualtrics platform and required participants to enter their identification numbers again to ensure data consistency. Upon completion of the session, each participant received a snack pack as a token of appreciation. For clarity, the process of the reception study is summarised in the flowchart presented in **Figure 5-5**.

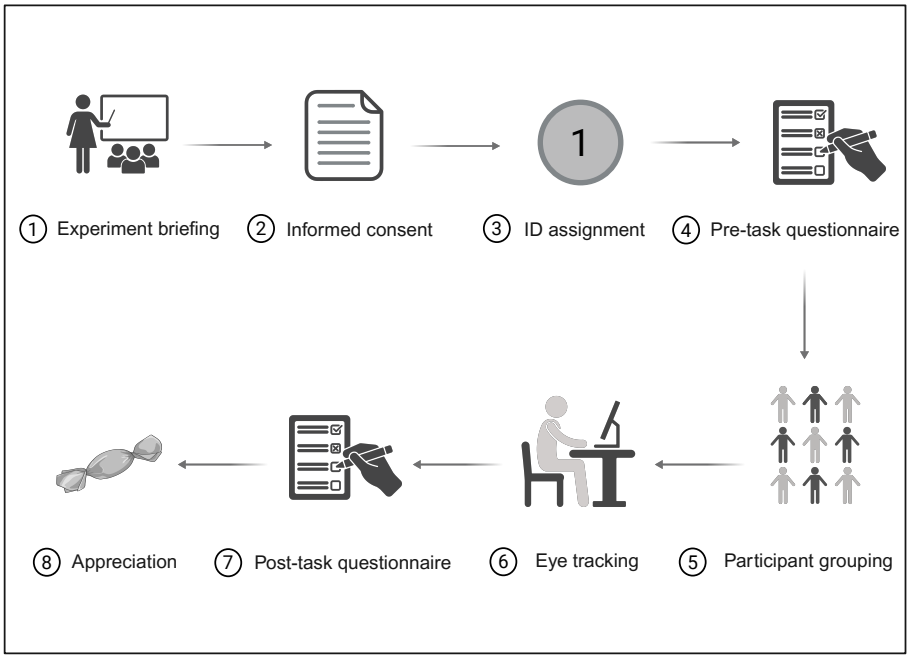


Figure 5-5. Process of the reception study

5.5.3.3 Step 3: Data processing

Once the reception study was completed, the questionnaire and eye-tracking data were exported and prepared for analysis.

Both pre-task and post-task questionnaire data were downloaded from the Qualtrics platform in XLSX format. These were then matched with the anonymous participant IDs.

The eye-tracking data were exported from Tobii Pro Lab in TSV format and paired with anonymous participant IDs from the questionnaires. Participants with a tracking ratio below 80% (six in total) were excluded (see Appendix J for detailed eye-tracking ratios of participants), leaving data from 32 participants for analysis. The pre-task and post-task questionnaire data from these six participants were also removed. As a result, the final

dataset included complete pre-task and post-task questionnaire data together with eye-tracking data from 32 Chinese DHH participants.

With the dataset organised, it was ready for analysis. The data analysis process is described in detail in Section 5.5.4.

5.5.3.4 Ethical considerations

As mentioned in Phase I, the two phases of the research project applied for ethical approval separately. Phase II was a reception study involving Chinese DHH participants who were invited to take part in an eye-tracking experiment. Since this phase involved potentially vulnerable individuals and was carried out outside Ireland, the ethical application was submitted to the university-level Research Ethics Committee. The application paperwork, included the Plain Language Statement, the Informed Consent Form, the Consent Letter from the special education school, the AV stimuli, and all materials used during the experiment. Ethical approval for this phase was obtained from Dublin City University's Research Ethics Committee on 16 April 2025, which classified Phase II of the study as a relatively low-risk project (Reference: DCUREC/2025/043; see Appendix G).

Since participants involved in Phase II were hearing-impaired individuals and the study was conducted in a school, several additional ethical considerations were put in place in advance. To obtain informed consent, all participants were first provided with a detailed explanation in the presence of a sign language teacher and school staff, and their concerns and questions were addressed at the same time. To ensure participant privacy, each participant was assigned an anonymous identification number (e.g., 01, 02, 03). These anonymous numbers were used consistently across the eye-tracking session and the online questionnaire platform (Qualtrics), and no personally identifiable information was linked to the research data. Regarding data security and confidentiality, all experimental data, including eye-tracking data and questionnaire responses, were stored in a secure, password-protected folder on the DCU student Google Drive. Access was restricted to my

supervisors and myself, and all data will be permanently deleted upon completion of the research.

5.5.4 Quantitative data analysis

The reception study (Phase II) was conducted with 38 Chinese DHH participants and collected three sets of quantitative data: a pre-task questionnaire, eye-tracking data, and a post-task questionnaire that included a comprehension test and a self-reported questionnaire on engagement and enjoyment. All analyses were conducted using IBM SPSS Statistics (version 30.0.0.0) and R (version 4.2.3) via RStudio (version 2025.05.1+513).

5.5.4.1 Eye-tracking analysis

The analysis of the eye-tracking data was carried out in the following four steps:

1) **Descriptive statistics:** Descriptive statistics were first computed to provide an overview of the eye-tracking data and to compare the general patterns between SDH and ST groups.
2) **Multivariate analysis:** A Multivariate Analysis of Variance (MANOVA) test was then performed to examine whether there were overall significant differences between two groups across the four eye-tracking metrics.
3) **Follow-up comparisons:** Once the MANOVA yielded significant differences, follow-up tests were conducted to identify which specific eye-tracking measures contributed to the multivariate effect.
4) **Multiple-comparison correction:** Finally, the Holm-Bonferroni correction was applied to control for potential Type I errors arising from multiple comparisons. The results of this step showed which eye-tracking measures remained significant and supported the conclusions related to cognitive effort.

<u>Descriptive statistics</u>

According to Saldanha and O'Brien (2014: 195), descriptive statistics are fundamental to quantitative analysis, as they summarise data and make it interpretable through measures such as minimum and maximum values, means, medians, standard deviations, and distributions.

As the first step in the data analysis, descriptive statistics were computed for the four eye-tracking metrics in both SDH and ST groups to illustrate the general characteristics of the data and to make clear comparisons between the two groups. These included the mean, standard deviation, median, minimum and maximum values, and range. Boxplots were also provided to visualise the distribution of the data and to identify potential outliers for each metric.

Multivariate analysis

Multivariate analysis of variance (MANOVA) is an extension of analysis of variance (ANOVA) that helps researchers compare two or more groups across multiple dependent variables simultaneously (Emerson, 2018: 125). Rather than carrying out multiple univariate ANOVAs on the same dataset, MANOVA reduces the risk of overestimating statistical significance (Doherty, 2012: 147).

In this study, given that four conceptually related eye-tracking dependent variables were analysed across two groups, MANOVA was used to examine whether subtitle type (SDH subtitles vs. ST subtitles) had an overall significant effect on the combined set of eye-tracking indicators. Simply put, MANOVA was conducted to test for overall group differences across all four eye-tracking measures at once. If the MANOVA result is significant ($p < 0.05$), it indicates that the two groups differed overall across the four eye-tracking measures. In that case, it is necessary to test the significance of each individual variable through follow-up univariate tests.

Follow-up comparisons

Follow-up tests, specifically Student's *t*-tests and Mann-Whitney *U* tests, were performed to determine which individual eye-tracking measures contributed to the multivariate effect. To select the appropriate test, the Shapiro-Wilk test was first applied to assess the normality of each eye-tracking metric, as it is considered reliable and powerful for small samples (Mishra et al., 2019: 71-72). According to Field (2024: 296), if the resulting *p*-value was greater than 0.05, the data were considered to be normally distributed.

Based on the Shapiro-Wilk test results, if the *p*-value was greater than 0.05, the data were considered normally distributed, and Student's *t*-test was used to compare the data between the two groups. If the *p*-value was less than 0.05, indicating that the data were not normally distributed, the Mann-Whitney *U* test was used instead as a non-parametric alternative (Mellinger & Hanson, 2017: 70-71). Additionally, for cases where *t*-tests were used, Levene's test was also carried out to check the equality of variances. A *p*-value below 0.05 was taken as evidence that this assumption had been violated (Mellinger & Hanson, 2017: 73).

Once the appropriate statistical tests were determined, the analyses included the calculation of effect sizes to provide information about the magnitude of the observed effects. For *t*-tests, Cohen's *d* was reported, with thresholds of 0.2, 0.5, and 0.8 indicating small, medium, and large effects, respectively (Cohen, 1988: 25-26). For Mann-Whitney *U* tests, the effect size *r* was computed, with benchmarks of 0.1, 0.3, and 0.5 interpreted as small, medium, and large effects (Cohen, 1988: 79-80).

Multiple-comparison correction

The Holm-Bonferroni correction is a stepwise, sequentially rejective version of the traditional Bonferroni method (Eichstaedt et al., 2013: 694). Although the Bonferroni correction is a traditional approach and has been used in previous studies (Muñoz, 2017; Liao et al., 2022; Szarkowska, 2024b), it is often regarded as overly conservative because it simply divides the desired significance level (e.g., $\alpha = 0.05$) by the number of tests, resulting in a much stricter threshold (e.g., $\alpha = 0.0125$ for four tests).

As an alternative, the Holm-Bonferroni procedure applies a step-down adjustment based on the number of remaining hypotheses. This sequential approach controls the family-wise error rate as effectively as the traditional Bonferroni method while retaining greater statistical power (Aickin & Gensler, 1996: 727). In this study, the Holm-Bonferroni correction was therefore applied to adjust the *p*-values obtained from multiple follow-up tests and to minimise the risk of Type I errors.

5.5.4.2 Comprehension test analysis

The analysis of the comprehension-test data followed a series of steps:

1) **Reliability analysis:** The internal consistency of the comprehension test was assessed using the Kuder–Richardson Formula 20 (KR-20), a variant of Cronbach's *α* designed for dichotomous items scored 0 for incorrect and 1 for correct (Kuder & Richardson, 1937). The KR-20 coefficient ranges from 0 to 1, where higher values indicate greater internal consistency. In practice, a coefficient of 0.70 is considered the minimum acceptable level of reliability (Mellinger & Hanson, 2017: 149). In SPSS, KR-20 can be obtained through Cronbach's *α*, as the calculated α value corresponds to the KR-20 coefficient.

2) **Descriptive statistics:** Descriptive statistics were produced to compare how participants in the two groups (SDH group and ST group) performed on the comprehension test.

3) **Normality test:** The Shapiro-Wilk normality test was then applied to determine whether the comprehension scores in each group were normally distributed. A significance level of $p < 0.05$ was used to identify deviations from normality (Field, 2024: 296).

4) **Between-group comparison:** Based on the normality results, a Mann-Whitney *U* test was used to compare comprehension performance between the two groups. In addition, effect-size values were calculated to indicate the magnitude of the difference. To quantify this magnitude, the effect size (*r*) was calculated following Cohen (1988: 79-80), where *r* values of 0.1, 0.3, and 0.5 indicate small, medium, and large effects, respectively.

5) **Item-level analysis:** Finally, Fisher's exact tests were conducted for each individual question to examine item-level differences and to identify which subtitle features might have contributed to variations in comprehension. Fisher's exact tests were used for each comprehension-test question. This test was chosen over the chi-square test due to the relatively small sample size ($N < 40$) and low expected frequencies (Neuhäuser & Ruxton, 2025). The purpose of this step was to identify whether certain questions showed statistically significant between-group differences. It also helped to identify which subtitle features might have influenced comprehension outcomes.

5.5.4.3 Engagement and enjoyment analysis

The engagement and enjoyment data were analysed in several steps:

1) **Reliability analysis:** The internal consistency of the questionnaire was assessed using Cronbach's α, a widely used statistical index for measuring the reliability of scales and questionnaires (Saldanha & O'Brien, 2014: 159). This coefficient ranges from 0 to 1, with higher values indicating greater internal consistency. A commonly accepted minimum threshold is 0.70, while values of 0.80 or above are considered preferable (Pallant, 2005). Separate α coefficients were calculated for the two subscales: engagement (Q11-Q18) and enjoyment (Q19-Q25).

2) **Descriptive statistics:** Descriptive statistics were then performed to provide an overview of how participants in SDH and ST groups scored on each dimension.

3) **Normality test:** The Shapiro-Wilk normality test was applied to check whether the data were normally distributed.

4) **Between-group comparison:** Depending on the normality results, either an independent-samples *t*-test or a Mann-Whitney *U* test was used to compare the two groups, and corresponding effect sizes were calculated to indicate the magnitude of any observed differences.

5) **Item-level analysis:** Finally, individual questionnaire items were examined graphically to uncover details that were not visible in the overall group comparison. This made it possible to see whether certain aspects of engagement and enjoyment (e.g., emotional involvement and willingness to rewatch) showed clearer differences between SDH and ST groups.

5.5.4.4 Correlation analyses

The correlation analysis aimed to explore the relations between eye-tracking measures and participants' comprehension, as well as their self-reported engagement and enjoyment. To achieve this, Spearman correlation tests were carried out. This non-parametric test "does not require the data to be normally distributed and better handles outliers and skewness" (Mellinger & Hanson, 2017: 142). Given that eye-tracking data often deviate from normality (Liao et al., 2020: 80), Spearman's correlation is therefore a preferable option. The Spearman correlation coefficient, written as ρ (rho) or r_s, ranges from −1 to +1, with +1 indicating a perfect positive relationship, −1 a perfect negative relationship, and 0 indicating no monotonic relationship (Howitt & Cramer, 2014: 68).

As part of this study, correlation analyses were organised in two parts:

1) **Overall correlation analysis:** correlations were calculated across all participants (N = 32) to look into the overall associations between eye-tracking measures and comprehension, engagement, and enjoyment.
2) **Separate correlation analysis:** correlations were then run separately within SDH group (n = 17) and ST group (n = 15) to see whether these associations held within each subtitle condition.

In this way, we were able to identify both the overall trends across the whole sample and the specific associations between variables under each subtitle condition.

Chapter 6 Findings and Discussion of the Interview Study

6.1 Chapter overview

This chapter presents and discusses the findings from interviews with nine professional SDH subtitlers. Section 6.2 introduces six themes identified from the interview analysis, which are further discussed in Sections 6.2.1 to 6.2.6 in relation to their impact on SDH subtitlers' decision-making. Finally, Section 6.3 summarises the implications of these findings for the next phase of the research.

6.2 Emerging themes (RQ1)

Through a qualitative analysis of the interview data from Phase I, six core themes were identified that shape how SDH subtitlers make decisions at different stages of the subtitling process. These factors included workflow and process optimisation, SDH guidelines, technical and practical constraints, audience-centred considerations, subjectivity in interpretation, and differences in content and background. Each factor was further divided into two sub-dimensions, as outlined in **Table 6-1**. The following sections discuss these factors in detail.

Factors	Sub-dimensions
Workflow and process optimisation	• Workflow structures • Collaboration and feedback
SDH guidelines	• Existing SDH guidelines • Improving existing guidelines
Technical and practical constraints	• Technological limitations • Practical constraints
Audience-centred considerations	• Diverse needs of audiences • Adjustments for children
Subjectivity in interpretation	• Understanding subjectivity • Addressing subjectivity
Differences in content and background	• Content-specific challenges • Knowledge and education differences

Table 6-1. Key factors influencing SDH subtitlers' decision-making process

6.2.1 Workflow and process optimisation

6.2.1.1 Workflow structures

Regarding IQ1 (SDH production workflow), responses from the nine interviewed SDH subtitlers indicate that the general workflow consists of transcription, editing, timing, annotation of non-verbal audio elements, and review. While these steps are broadly similar across organisations, practical differences exist in their implementation. **Table 6-2** compares the SDH workflows of the three organisations.

At the Irish television broadcaster (hereafter the national broadcaster), SDH subtitlers are responsible for the entire SDH production process independently. Each member manages multiple programmes, from transcription to editing, synchronisation, and final review, with all steps carried out individually. Due to the broadcaster's high volume of programming, one interviewee (P1) complained that they often face a heavy workload: "We had loads to do, and we had to complete everything within a couple of days."

For those working at the Chinese accessibility company (hereafter the accessibility company), a more precise designation than subtitlers might be AD and SDH scriptwriters, as they are responsible for both SDH and AD scripts. Compared to those at the national broadcaster, professionals working at the accessibility company do not need to transcribe the audio but instead work with pre-existing subtitle texts along with timecodes. These subtitle files are typically obtained from film producers or directors. Moreover, timing and synchronisation are handled by the post-production team, which reduces their overall workload and allows them to focus on crafting SDH and AD scripts. According to P4, she typically carries out preparatory work before beginning the scriptwriting process:

> "When I first get the script and the video, I watch the entire film or TV series once. I get familiar with the storylines and the relationships between characters to understand the entire plot. After that, I look for some introductions related to the film, such as the awards it has received, information about the actors, and some film reviews. If the film involves some historical aspects, I also do a bit of research on that to get a basic understanding of the relevant history. Then I start writing."

Participant	P1, P2, P3	P4, P5, P6, P7	P8, P9
SDH	English-English	Chinese-Chinese	Chinese-English
Organisation / Workflow	Irish television broadcaster	Chinese accessibility company	Subtitling group on a YouTube channel
Transcription	Combine automatic speech recognition with manual transcription.	N/A (Note: provided by film producers or directors)	Combine automatic speech recognition with manual transcription.
Translation	N/A	N/A	Translate the script into English using machine translation, followed by human post-editing.
AD creation	N/A	Write AD scripts before writing SDH scripts.	N/A
SDH creation	Create SDH subtitles while watching videos in subtitling software.	Create SDH subtitles separately in an Excel sheet as part of the scriptwriting process.	Create SDH subtitles while watching videos in subtitling software.
Timing	Adjust subtitle timing and ensure synchronisation with scenes.	N/A (Note: handled by post-production team)	Adjust subtitle timing and ensure synchronisation with scenes.
Review	Conduct an individual review.	Conduct a peer review.	Conduct a peer review.

Table 6-2. Comparison of SDH production workflows across three organisations

The workflow of the subtitling group on a YouTube channel (hereafter the YouTube subtitling group) differs from that of the other two organisations, as it focuses on interlingual SDH production. They first translate the original Chinese content into English and then create English SDH subtitles based on the translated text. According to P8, SDH is considered an additional task rather than a primary goal. Their main priority is to ensure the quality of the translated subtitles, while the refinement of SDH subtitles depends on the availability of extra time.

6.2.1.2 Collaboration and feedback

Through an analysis of workflow structures, it became evident that SDH production in the accessibility company is not an entirely isolated task but rather a process that relies on collaboration for continuous refinement. In this company, a well-structured division of labour simplifies the SDH production workflow and reduces the workload for SDH subtitlers. As previously mentioned, subtitlers in this company do not need to manually transcribe dialogue or spend a significant amount of time adjusting timing, which enables them to focus more on refining the textual content. Compared to the workflows at the other two organisations interviewed, this collaborative workflow might be considered an outlier, as neither the national broadcaster nor the YouTube subtitling group explicitly mentioned following such a division of labour in their processes.

In addition to the distinct collaborative workflow observed in the accessibility company, P7, working there, also advocated for greater involvement of SDH subtitlers in the filmmaking process when answering IQ10 (advice for novice SDH subtitlers). She argued: “We hope that SDH and AD scriptwriters can engage with the film’s making process, and ensure that they remain an integral part of the work, not just a separate technical task.” The idea behind this is to help scriptwriters better understand the film creator’s intentions. In turn, this understanding would enable them to craft SDH subtitles that are more faithful to creator’s original intent. While this concept remains largely theoretical, it has the potential to

enhance SDH production by integrating scriptwriters more closely into the filmmaking process.

Not only collaboration, but also feedback from the target audience, as indicated by the interview data, can provide valuable guidance for SDH subtitlers' decision-making. One interviewee (P4) mentioned that their team occasionally invites deaf and blind viewers to review SDH and AD content and gather their feedback. P4 recalled an experience in which she received feedback from a visually impaired audience member:

> "Once when I was writing a script for AD, I mentioned that a person was about 2 to 3 metres from the door. A visually impaired person found it hard to understand, because they don't get the concept of 2 to 3 meters well. So, we changed it to a more relatable description, like comparing the size to a familiar object such as a smartphone. This makes it easier for them to understand. Descriptions should be relatable to their experiences or familiar objects."

Although this feedback came from a visually-impaired audience rather than a deaf viewer, it shows how the feedback of the target audience helps SDH subtitlers refine word choice and influences their decision-making processes.

6.2.2 SDH guidelines

6.2.2.1 Existing SDH guidelines

To address IQ2 (limitations of existing SDH guidelines), I asked the nine SDH subtitlers about the guidelines or rules they follow in their actual work. Given that they are affiliated with three different types of organisations, their responses varied accordingly.

Firstly, three SDH subtitlers from the national broadcaster (P1, P2, P3) noted that there are national SDH guidelines in their country. However, their team primarily relied on internally developed guidelines, which not only involve elements of the national guidelines but also address additional aspects not explicitly covered. One example mentioned by P3 is that there are no specifications regarding spoken and written language in their national guidelines, such as the rule that when spelling "okay", it must be written as O-K-A-Y rather than the abbreviation "ok", which is not addressed in the national rules. This suggests that,

based on the views of these three SDH subtitlers, while the national guidelines provide a foundational reference for SDH production, their team does not necessarily follow them rigidly. Instead, they prioritise their in-house standards, which play a more practical role in guiding their actual SDH production workflow.

Similarly, four interviewees from the accessibility company (P4, P5, P6, P7) also reported following internally developed guidelines, but this is largely due to the absence of an officially published and widely adopted national SDH production standard in China. This observation aligns with existing research, as some scholars have also noted the lack of a uniform standard in China (Li, 2013: 173; Yang et al., 2022: 159). Alongside SDH requirements, P4 added that their internal rules also include other accessibility standards, such as AD and SLI.

In stark contrast to the above two organisations, the two subtitlers interviewed from the YouTube subtitling group (P8, P9) stated that they do not follow any established SDH guidelines. Instead, their subtitling practices are based on personal experience, influenced by their exposure to AV content with SDH. P8, the leader of this subtitling group, admitted that “we are not very specialised in producing SDH ourselves.” This is because their main focus is interlingual subtitling (Chinese to English), with SDH production being a secondary and exploratory endeavour. P8 further elaborated that their reliance on volunteers makes it challenging to impose additional strict requirements on members, as the group operates without financial compensation.

6.2.2.2 Improving existing guidelines

As previously mentioned, SDH subtitlers from both the national broadcaster and the accessibility company follow existing SDH guidelines in their work. In response to IQ2 (limitations of existing SDH guidelines), I further inquired about their perspectives on these guidelines and the limitations they have encountered in practice.

Since the guidelines used by the national broadcaster are developed based on national standards and supplemented with additional details, they are generally more

comprehensive and well-structured, providing stronger support for SDH subtitlers' workflows. Consequently, participants from this organisation raised fewer concerns regarding the limitations of their internal rules.

By comparison, SDH subtitlers from the accessibility company (P4, P5, P6, P7) mentioned some limitations of their internal guidelines. This is mainly because their guidelines were developed from scratch and have yet to evolve into a fully structured standard. As one interviewee (P4) explained: "Since there is no official standard in China, we are first movers in this field. It is definitely not perfect… So, for us, it is a process of trial and error." Speaking of specific drawbacks, two interviewees (P5, P6) pointed out that their current guidelines lack detailed instructions for different types of AV content. While the guidelines provide general recommendations for films and television series, they do not address how to subtitle more specialised content, such as musicals, documentaries, or reality shows. Another limitation, as highlighted by P7 in response to IQ8 (adapting SDH for children), is the absence of specific provisions for adapting SDH for children's content, such as using simpler language or reducing subtitle speed.

6.2.3 Technical and practical constraints

6.2.3.1 Technological limitations

Throughout the interviews, nearly all interviewees highlighted the role of technology in their SDH work. Undeniably, technology has facilitated their workflow. For instance, speech recognition technology reduces the need to manually generate subtitles from scratch, an otherwise time-consuming task. However, despite these advancements, technological limitations persist in practice.

Among all the IQs, the one most directly related to technology was IQ5 (emoji usage in SDH subtitles), which explores the possibility of using emojis to convey characters' emotions in SDH. Compared to text, emojis have visual intuitiveness and space-saving advantages that prompt some studies to investigate their potential integration into SDH (see Gürkan, 2019;

Oomori et al., 2020). Against this backdrop, the question was designed to gather SDH subtitlers' views on the topic.

Three SDH subtitlers from the national broadcaster (P1, P2, P3) stated that their subtitling software does not support emojis and only allows basic characters. As P1 said: "The subtitling software only takes certain characters, so you would literally just have to have like a colon and bracket to be a sad face or a smiley face."

Interestingly, other interviewees (P4, P5, P6, P7, P8, P9) noted that their subtitling software does support emoji display, yet they tend not to use them for various reasons: P4 believes that DHH audiences can infer characters' emotions from visual content; P5 doubts that emojis can accurately reflect emotions; P6 is concerned about whether DHH audiences would accept them; P7 argues that emojis could distract viewers; P8 holds the view that emojis may be inappropriate for serious AV content, such as historical dramas; and P9 points out that the interpretation of emojis varies among individuals, potentially leading to misinterpretation. Overall, these responses reflect that, despite technical feasibility, SDH subtitlers carefully consider multiple factors before deciding whether to use emojis in practice.

Aside from their responses to IQ5 (emoji usage in SDH subtitles), interviewees also frequently mentioned technological limitations when answering IQ1 (SDH production workflow). Although speech recognition has improved subtitling efficiency, as noted earlier, some interviewees (P1, P2, P3) addressed its notable downsides. P2 specifically noted that when handling dialects or strong regional accents, the accuracy of speech recognition is suboptimal. As such, manual correction is often required to ensure subtitle accuracy.

6.2.3.2 Practical constraints

One of the most significant practical constraints faced by SDH subtitlers is the spatial and temporal limitations on screen. Unlike standard subtitling, which transcribes spoken dialogue, SDH is designed to convey the full auditory experience to viewers. This includes not only speech but also other auditory elements such as paraverbal cues, sound effects,

and both diegetic and non-diegetic music (Zabalbeascoa, 2025: 80). In practice, when multiple sounds overlap with dialogue, SDH subtitlers are often forced to make difficult decisions about which elements to include and which to omit. In IQ3 (selection criteria for subtitling non-verbal audio information), the nine interviewed subtitlers shared how they prioritise non-verbal audio elements. Their responses can be summarised into two key factors:

- Whether the sound is visually represented on screen – If the audience can see the source of the sound (e.g., a door slamming, an explosion, or applause), SDH subtitlers may choose to omit the description to reduce subtitle density.
- Whether the sound is crucial to the plot – If a sound plays a critical role in the narrative (e.g., a ringing phone, a gunshot in the background, or a whispered conversation), SDH subtitlers tend to prioritise its inclusion to ensure that viewers do not miss essential information.

Yet, these two factors are only taken into consideration within the constraints of limited subtitle space and timing in most cases. P1 said that if there is a prolonged period of silence on screen with no dialogue, even if background music is not directly relevant to the plot, she would still choose to include a description of the music to avoid a complete absence of subtitles. P7 expressed a similar view, explaining that if subtitles are missing for an extended duration, DHH viewers may feel overlooked. Considering this, she would also opt to include background music descriptions where appropriate.

6.2.4 Audience-centred considerations

6.2.4.1 Diverse needs of audiences

A recurring theme in the interviews is the heterogeneity of the DHH community, a topic previously discussed in the literature review (Section 2.4.3). As outlined in that section, while DHH individuals experience hearing impairments, they vary significantly in terms of the degree of hearing loss, educational background, and cultural knowledge. Similarly, all interviewed SDH subtitlers acknowledged this diversity. As P1 noted, “There’s a big spectrum of DHH community. You could have some hearing, and you might have been born

with hearing loss." These differences result in varying needs and preferences for SDH, which, in turn, directly influence subtitlers' decision-making processes.

For instance, IQ4 (music description in SDH) asked all interviewees how they describe music to cater to diverse audience needs. More than half (P1, P3, P4, P5, P6, P9) stated that they prefer using descriptive language to convey the atmosphere or emotion of the music rather than simply labelling it by genre or instrument, such as "classical music" or "piano sound." As P1 explained: "I think it's more important to convey the mood of the music than the genre. So it's more important to say 'relaxed music,' 'jaunty music,' 'cheering music,' or 'sad music.' I think communicating these emotions is more important than specifying the genre." Other interviewees (P2, P4, P5, P6, P9) also supported this idea and believed that it makes SDH subtitles more inclusive, as not all DHH audiences are familiar with specific music genres or instruments.

That said, P7 offered a slightly different view, in which she opted for a combination of references to specific instruments and the atmosphere of the music, instead of only focusing on the latter. This preference was influenced by her interaction with a deaf viewer who expressed a desire for more musical information to enhance their knowledge while watching films. P8 also supported the idea of including instrument names or song titles and argued that certain instruments or specific songs can contribute to the audience's understanding of the narrative. For example, if a well-known song is used in a scene and carries significant emotional weight, it should be identified to improve the audience's comprehension of the storyline.

In addition to music description, the consideration of audience heterogeneity is also evident in how sound effects are conveyed in SDH. When responding to IQ6 (challenges encountered in SDH production), P2 raised concerns about the representation of sound effects. She cited "light flickering" as an example, explaining that she initially considered using the onomatopoeia "chik-chik" to mimic the flickering of the light but ultimately decided on using "light flickering." Although "flicker" itself can also be considered an

onomatopoeia, it is easier for the audience to understand compared to “chik-chik,” and it better conveys the relationship between the sound and the flickering light action. Based on this understanding, P2 made the final decision. In fact, previous SDH research, such as Gürkan’s study (2019), also discouraged the use of onomatopoeia, as deaf participants in his research found such representations difficult to comprehend and expressed a preference for textual descriptions, even though this approach results in longer subtitles.

6.2.4.2 Adjustments for children

Considering children’s comprehension levels and reading abilities, some SDH guidelines include specific adjustments for subtitles in children’s programmes. For example, the CAB Guidelines (2012) specify that subtitling speed should be kept under 200 words per minute for adult content and limited to 120 - 130 words per minute for children’s programmes. Interview data reflect similar considerations. When responding to IQ8 (adapting SDH for children), several SDH subtitlers (P1, P2, P3) indicated that they do make such adjustments. In addition to reducing subtitle speed, they also simplify the language used in SDH subtitles. As P1 cited an example: “If you’re describing music for an adult audience, you might use words like ‘poignant’ or ‘contemplative’, but if the audience is children, you would simply say ‘sad’.”

It is important to note that all the interviewees who advocated these adjustments were from the national broadcaster. When interviewing SDH subtitlers from the accessibility company, they mentioned that although they also produce SDH subtitles for animated films, these films are primarily intended for general audiences rather than specifically for children. Their company does not have specific regulations mandating adjustments for children’s content. Despite this, two interviewees (P5, P7) mentioned that when working on animated films, they consciously use simpler and more accessible language to describe non-verbal sounds to ensure that the SDH content aligns more closely with the style of animated films and audience expectations.

In response to this point, a less mainstream viewpoint was raised by P9, who works in the YouTube subtitling group and seemed to disagree with modifying language in children's programmes to match their comprehension levels. She argued that SDH can serve as a tool to improve children's literacy skills and that over-simplification may hinder language development. Her remarks indicate that SDH may also have an educational function. In fact, some AVT scholars (Díaz-Cintas & Remael, 2007; Zárate, 2014; Vanderplank, 2016) have recognised this potential. As Díaz-Cintas and Remael (2007: 15) observe: "SDH also offers great educational potential to people with limited knowledge of a country's language, e.g., immigrants and foreign students." In this context, children could also be considered among the potential audiences who benefit from SDH as a tool for language learning. Compared to the relatively short history of SDH, interlingual or intralingual subtitles have long been praised by scholars as tools for promoting language acquisition. Paivio (1979) argues that subtitles can aid language learning by engaging both visual and verbal channels, while Vanderplank (1988) demonstrates their role in enhancing language comprehension and improving literacy skills for foreign language learners.

6.2.5 Subjectivity in interpretation

6.2.5.1 Understanding subjectivity

Subjectivity in SDH production refers to the inherent personal interpretation involved in describing non-verbal audio elements. Specifically, when describing the same piece of music, one subtitler might label it as "soft piano music," while another might describe it as "gentle instrumental music." These differences in interpretation mean that the emotional impact conveyed to the target audience may vary depending on subtitlers' subjective choices. To explore how SDH subtitlers understand subjectivity, we asked participants about their views on this issue in IQ9 (balancing objectivity and subjectivity in SDH production).

SDH subtitlers from the accessibility company (P4, P5, P6, P7) and the YouTube subtitling group (P8, P9) generally recognised subjectivity as an unavoidable aspect of SDH production. P9 commented: "I don't think there is such a thing as pure objectivity. There is no way you

can avoid filtering it through your own brain if you're doing it." Similarly, P6 believed that "achieving complete objectivity is impossible, because as SDH subtitlers, we naturally make subjective judgments and choices." P6 further clarified that these subjective interpretations are not solely based on personal preferences but rather on the role of SDH subtitlers as "mediators" who strive to faithfully convey the filmmaker's intent. P7 echoed this sentiment, adding: "We aim to restore the original creator's expression rather than impose our own subjective feelings on the audience." From these responses, it seems that, for some SDH practitioners, subjectivity is not viewed as a flaw in SDH production but rather as an inherent part of the decision-making process when representing the creator's intent.

A slightly different view was held by interviewees from the national broadcaster (P1, P2, P3), who consistently emphasised the importance of maintaining objectivity in their work, even though they recognised that subjectivity is unavoidable. P2 stressed: "You have to stay away from anything that reflects a personal opinion in descriptions of sounds. Always, always remain objective." From their perspective, SDH descriptions should remain as neutral as possible, and SDH subtitlers should avoid personal bias or subjective interpretation.

It became evident that there was a minor divide in how these SDH subtitlers perceive and understand subjectivity. The reasons behind this divide likely lie in institutional culture and audience expectations. The national broadcaster operates under strict industry standards and prioritises consistency and standardisation to ensure that all viewers receive uniform information. In the other two settings, the accessibility company and the YouTube subtitling group primarily serve and cater to fan-based audiences that allows for greater flexibility in subtitling decisions.

6.2.5.2 Addressing subjectivity

In the previous section, we examined how the interviewed subtitlers perceive subjectivity in SDH. While interviewees from the national broadcaster tend to use more objective and descriptive language to minimise subjective interpretation, those from the accessibility

company and the YouTube subtitling group actively engage with subjectivity in practice and seek ways to regulate it instead of trying to eliminate it entirely.

The primary solution is to implement peer review mechanisms. As noted by P4 from the accessibility company, “Once I complete the first draft of the subtitles, another team member, who acts as an editor, reviews the draft.” In cases where discrepancies arise, discussions are held to reach a consensus before integrating the agreed-upon solutions into existing guidelines. Peer review serves not only as a quality control measure but also allows for case-by-case flexibility in decision-making. And also, the interviewee from the YouTube subtitling group (P8) described a comparable review system, in which some team members are responsible for translating subtitles, while others handle the review process. However, the feasibility of implementing this mechanism may be limited by time constraints, particularly in fast-paced environments such as the traditional television broadcasters interviewed, where there is limited flexibility to allocate more time for peer review.

In addition to the introduction of peer review mechanisms, one interviewee (P8) mentioned the potential role of AI-assisted technology in addressing subjectivity. He said: “Since I have a technical background, I often think about whether there could be a technological tool to automatically add speaker identification and intonation.” Recent research in the field of AVT has also acknowledged this trend. Szarkowska and Jankowska (2025: 186) highlight that SDH production is likely to become increasingly automated, primarily through automatic speech recognition and AI-based spotting. In practice, though, the specific technology mentioned by P8, detecting and labelling speaker intonation, does not yet appear to have been widely adopted in practice. In other words, it is still a conceptual idea, not a practical one.

Beyond that, the nine interviewees did not mention any specific training aimed at helping SDH subtitlers address subjective issues, either in response to IQ9 (balancing objectivity and subjectivity in SDH production) or IQ10 (advice for novice SDH subtitlers), both of which are highly relevant to subjectivity and training. According to the interview content, the current training system focuses on technical aspects, such as using subtitling software and adhering

to guidelines, while giving less attention to developing the ability to make appropriate subjective judgments in various contexts.

6.2.6 Differences in content and background

6.2.6.1 Content-specific challenges

Different types of AV content may influence the decision-making process of SDH subtitlers. To identify the most challenging AV genres for SDH production, I asked nine interviewees in IQ7 (the most challenging AV type for creating SDH) to specify which types of content present the greatest difficulties. According to the interview data, reality shows, variety shows, and chat shows were considered the most challenging.

One of the primary reasons reality and variety shows are particularly difficult to subtitle is their lack of structured content or, simply put, "because they're not scripted," as pointed out by P1. Other interviewees, such as P8, explained that these shows may have a general script, but actors do not adhere to it as strictly as they do in films. Most of the time, a significant portion of the dialogue and performances are spontaneous, which adds complexity to SDH production. Furthermore, P7 indicated that Chinese variety shows often use a large number of impact captions (or *telop,* see Section 5.5.2.1) to emphasise tone, enhance humour, or provide supplementary information. In such cases, SDH subtitlers must carefully consider the interaction between impact captions and SDH text to avoid redundancy or potential contradictions.

Another particularly challenging genre is chat shows, where the main difficulty lies in fast-paced conversations and excessive overlapping speech. As P3 described: "Chat shows, you know, like daytime chat shows, where you've got lots of guests on, can be challenging. If you have a presenter and guests, they often interrupt each other." Similarly, P2 noted the issue of excessive overlapping dialogue: "There's quite a lot of overlapping dialogue, and sometimes it's hard to decide what is the most important thing to actually subtitle, because obviously, you can't fit everything in the space of maybe two seconds when five people are talking." In such cases, P2 added, "We will generally pick whatever is, again, most important

to the plot, or what's most important to the conversation. So if there's like five people screaming at each other, whatever is the most relevant and important to what's happening is what we will choose to subtitle."

6.2.6.2 Knowledge and education differences

Based on the interview data analysis, the decision-making process in SDH is also influenced by differences in knowledge and education between SDH subtitlers and SDH users.

In terms of knowledge differences, particularly in specialised areas, one relevant example reflects this issue. In this Chapter, Section 6.2.4.1 (Diverse needs of audiences), I discussed how SDH subtitlers describe music in SDH subtitles. The findings indicated that most interviewed subtitlers prefer to describe the atmosphere or emotions conveyed by the music rather than specifying its genre or instrumentation. One primary reason for this approach is their concern that some deaf individuals may not have been exposed to certain music genres or instruments. Another key factor, which had not been previously mentioned, is the subtitlers' own limited knowledge of music. As P6 admitted:

> "Since we're not music experts and don't always have the time to consult professionals in every instance, we prefer to avoid specifying the genre or instrument. Instead, we tend to describe the atmosphere it conveys as a safer alternative."

Additionally, P4, who works in both SDH and AD, shared her experience of writing AD for visually impaired audiences. She stated: "In AD, we don't specify whether it's a piano, trumpet, or any specific instrument because our visually impaired audience typically have excellent hearing. Many of them are actually more knowledgeable about music than we are. An inaccurate description could mislead them, making them think our work is unprofessional and affecting their experience of the music." Although this example does not directly relate to SDH, P4 acknowledged that this experience has influenced her approach to SDH production. It has made her more cautious in providing detailed descriptions of elements she is unfamiliar with, such as musical instruments.

Beyond knowledge gaps, educational disparities between SDH subtitlers and SDH users also shape subtitling decisions. In response to IQ8 (adapting SDH for children), P6 noted that the overall educational level of the DHH population in China tends to be lower than that of hearing individuals. Considering this, P6 added that she prefers to use simpler, more accessible language in SDH subtitles to ensure that DHH audiences can easily comprehend the content.

6.3 Implications for the next phase (RQ2)

Thus far, drawing on the practical experiences of nine professionals in the field, we have explored the key factors shaping SDH subtitlers' decision-making. These findings not only contribute to our understanding of the actual SDH production process but also inform the next phase of research: the creation of SDH subtitles, which will be used in the reception study among Chinese DHH audiences.

One key finding is the notable variation in SDH production workflows across different organisations. After comparing the three interviewed organisations, it became evident that SDH production is not fully standardised but rather shaped by available resources. At the accessibility company, SDH subtitlers have access to pre-existing subtitle scripts from film producers, which greatly reduces the time required for transcription. Therefore, in the second phase of research, priority will be given to obtaining original subtitles and video materials from production teams to streamline the workflow. Furthermore, preparatory work will also include preliminary background research on the selected AV material, as proposed by interviewees.

Since the next phase of the reception study involves Chinese DHH audiences watching Chinese video clips with Chinese SDH subtitles, a well-defined SDH guideline is essential as a reference. However, as confirmed by interviewees from the accessibility company, their internal guidelines were developed from scratch and remain incomplete, with notable limitations. Given this, in the second phase of research, internationally established SDH standards will be adapted to refine the production of SDH subtitles.

Technical and practical constraints are also key considerations in SDH production. Interviewed SDH subtitlers suggested that while automatic speech recognition technology improves transcription efficiency, it remains prone to errors and requires manual review and correction. In the second phase of research, I plan to prioritise obtaining original subtitle scripts from production teams, but if these scripts are unavailable, speech recognition will be combined with human post-editing to ensure the accuracy of subtitles. Another technical consideration is whether to use emojis in SDH to convey character emotions. Since most interviewees expressed reservations about using emojis in SDH, they will not be included in the next phase of SDH production. On the practical side, spatial and temporal limitations are identified as major challenges in SDH production. Based on interview data, two key factors for determining the inclusion of non-verbal sounds in SDH subtitles have been established: (1) the visibility of the sound and (2) its relevance to the plot. Nevertheless, exceptions apply. According to interviewed subtitlers, when there is a prolonged absence of dialogue, descriptions of visible but less plot-relevant sounds, such as background music, may be included to enhance audience immersion. In the next phase, I will create the SDH subtitles following these two key factors while allowing for exceptions where necessary.

Another insight is that all interviewees emphasised audience-centred considerations, particularly in their approach to representing music and sound effects. Regarding music, most interviewees preferred to describe its atmosphere or emotional impact rather than specifying genres or instruments. For sound effects, they recommended avoiding onomatopoeic representations to prevent confusion among deaf audiences. Both suggestions will be incorporated into the second phase of SDH production. As for adjustments for child audiences, most interviewees indicated they would use simpler language and reduce subtitle speed. While the next phase does not specifically target children, the selected AV material is an animated film. To align with its style and audience expectations, I will apply simplified language and adjust subtitle speed accordingly.

For SDH subtitlers, subjectivity remains a major concern in SDH production. As all interviewees acknowledged, some degree of subjectivity is inevitable, but efforts can be directed towards minimising individual interpretation biases. To this end, some organisations have adopted peer review systems to reduce bias through collective discussion. Accordingly, in the second phase of research, I plan to invite professional SDH subtitlers to review SDH subtitles that I created.

Lastly, certain types of AV content, particularly reality shows, variety programmes, and chat shows, prove especially challenging for SDH subtitlers due to their fast-paced dialogue, overlapping speech, and unscripted nature. Given these complexities, the second phase of research will prioritise more manageable AV materials and exclude these genres. Additionally, considering differences in knowledge and background between SDH subtitlers and their audience, interviewees opt for simpler, more accessible language in SDH subtitles, a practice that will inform the next phase of SDH production.

Chapter 7 Results and Discussion of the Reception Study

7.1 Chapter overview

This chapter presents the results of the reception study, which draws on three quantitative datasets: responses to the pre-task questionnaire, eye-tracking measures, and responses to the post-task questionnaire. Section 7.2 provides the pre-task questionnaire results through a series of tables that summarise the DHH participants' demographic profiles, viewing habits, and subtitle preferences. In Section 7.3, the results from the analysis of the four eye-tracking metrics are presented. Section 7.4 then outlines the post-task questionnaire results, including participants' comprehension performance as well as their reported levels of narrative engagement and enjoyment. Section 7.5 follows with the correlation analyses that include both overall and group-based results and conclude with a summary of the key findings. Finally, Section 7.6 discusses these results in relation to the three RQs.

7.2 Pre-task questionnaire results

In the pre-task questionnaire (see Appendix H), Questions 1-7 gathered information on participants' demographic characteristics, Questions 8-9 examined their viewing habits, and Questions 10-15 investigated their subtitle preferences. The following sections present participants' responses for these three parts.

7.2.1 Demographic data

Demographic information was obtained from Questions 1-7 of the pre-task questionnaire, which asked about participants' gender, age, eyesight, degree of hearing loss, and related background information. The results are represented in **Table 7-1**.

	SDH group (n=17)		ST group (n=15)		Total
	n	%	n	%	n
Gender					
Male	8	47.1	7	46.7	15
Female	9	52.9	8	53.3	17
Age					
Mean	20	/	20.13	/	20.06
SD	1.73	/	1.94	/	1.82
Eyesight[6]					
Normal vision	10	58.8	15	100.0	25
Myopia (below 5.0 dioptres)	4	23.5	0	0.0	4
Myopia (above 5.0 dioptres)	3	17.6	0	0.0	3
Degree of hearing loss[7]					
Complete (Level 1)	8	47.1	6	40.0	14
Profound (Level 2)	7	41.2	7	46.7	14
Severe (Level 3)	2	11.8	2	13.3	4
Moderate (Level 4)	0	0.0	0	0.0	0

[6] The term “myopia” refers to near-sightedness, with higher values indicating more severe levels. Eyesight information was collected to help ensure reliable eye-tracking results, as visual conditions can affect data quality during recording.

[7] A detailed description of the degree of hearing loss is provided in Section 5.5.1.1.

Onset age of hearing loss					
At birth	6	35.3	7	46.7	13
Before 2	5	29.4	2	13.3	7
2-6	5	29.4	4	26.7	9
7-12	0	0.0	1	6.7	1
13-18	1	5.9	1	6.7	2
Use of hearing assistive devices					
No	1	5.9	4	26.7	5
Yes (cochlear implants)	8	47.1	5	33.3	13
Yes (hearing aids)	8	47.1	6	40.0	14
Primary communication mode					
Spoken language	8	47.1	4	26.7	12
Sign language	4	23.5	5	33.3	9
Sign language & spoken language	4	23.5	4	26.7	8
Text-based communication	1	5.9	1	6.7	2
Other	0	0.0	1	6.7	1

Table 7-1. Demographic information of participants (Phase II)

In terms of gender distribution, the SDH group consisted of eight males (47.1%) and nine females (52.9%), while the ST group included seven males (46.7%) and eight females (53.3%). In total, the sample comprised 15 males (46.9%) and 17 females (53.1%).

Regarding age, the mean age of participants in the SDH group was 20 years (SD = 1.73), while the ST group had a slightly higher mean age of 20.13 years (SD = 1.94). The overall mean age across both groups was 20.06 years (SD = 1.82).

With regard to eyesight, all participants in the ST group reported normal vision, whereas the SDH group included participants across all three eyesight categories.

With respect to the degree of hearing loss, both groups showed similar distributions, with the majority of participants classified under complete or profound hearing loss. Across all participants, 43.8% had complete hearing loss, 43.8% had profound hearing loss, and 12.5% had severe hearing loss.

Concerning the onset of hearing loss, 13 participants (40.6%) described being deaf or hard of hearing since birth, including six from the SDH group and seven from the ST group. Seven participants (21.9%) indicated onset before the age of 2, with a higher proportion in the SDH group (n = 5, 29.4%) compared to the ST group (n = 2, 13.3%). Another nine participants (28.1%) reported onset between ages 2 and 6 (five in the SDH group and four in the ST group). Only one participant (3.1%) showed onset between ages 7 and 12, who belonged to the ST group. Finally, two participants (6.3%) noted onset between ages 13 and 18, one in each group.

When it comes to the use of hearing assistive devices, only five participants (15.6%) reported not using any device, with the majority in the ST group (n = 4, 26.7%) and only one in the SDH group (5.9%). Thirteen participants (40.6%) had cochlear implants, distributed across both groups (SDH: n = 8, 47.1%; ST: n = 5, 33.3%). A further 14 participants (43.8%) used hearing aids, with similar proportions in the SDH group (n = 8, 47.1%) and the ST group (n = 6, 40.0%).

Lastly, on communication modes, spoken language was the most common mode, used by 12 participants (37.5%). This was more frequent in the SDH group (n = 8, 47.1%) than in the ST group (n = 4, 26.7%). Nine participants (28.1%) preferred to use sign language exclusively (SDH: n = 4, 23.5%; ST: n = 5, 33.3%). Eight participants (25.0%) used a combination of sign language and spoken language, with equal distribution across both groups (n = 4 in each). Text-based communication was the primary mode for two participants (6.3%), one from each group. Only one participant (3.1%) in the ST group selected “other” as their primary

communication mode. After further checking the responses, it was found that this participant used a combination of text-based communication and lip reading.

7.2.2 Viewing habits

Questions 8 and 9 were designed to understand participants' viewing habits, involving their average daily video-viewing time and preferred types of video content. A summary of the responses is provided in **Table 7-2**.

	SDH group (n=17)		ST group (n=15)		Total
	n	%	n	%	n
Video viewing time per day					
Less than 1 hour	2	11.8	2	13.3	4
1-2 hours	12	70.6	6	40.0	18
2-4 hours	2	11.8	4	26.7	6
More than 4 hours	1	5.9	3	20.0	4
Preferred video type					
Movies	4	23.5	7	46.7	11
TV series	6	35.3	2	13.3	8
Animation	3	17.6	3	20.0	6
Sports	2	11.8	0	0.0	2
Other	2	11.8	3	20.0	5

Table 7-2. Viewing habits of participants (Phase II)

According to the questionnaire, participants showed different amounts of daily video-viewing time. In the SDH group, the majority (n = 12, 70.6%) watched videos for 1-2 hours per day, followed by two participants (11.8%) who watched less than 1 hour, two

participants (11.8%) who watched 2-4 hours, and one participant (5.9%) who watched more than 4 hours per day. In the ST group, six participants (40.0%) watched videos for 1-2 hours per day, four participants (26.7%) for 2-4 hours, three participants (20.0%) for more than 4 hours, and two participants (13.3%) for less than 1 hour per day.

Participants' preferences for video content also varied across groups. In the SDH group, the most frequently selected type was TV series (n = 6, 35.3%), followed by movies (n = 4, 23.5%), animation (n = 3, 17.6%), sports (n = 2, 11.8%), and other content (n = 2, 11.8%). In the ST group, movies were the most common preference (n = 7, 46.7%), followed by TV series (n = 2, 13.3%), animation (n = 3, 20.0%), and other content (n = 3, 20.0%). Notably, participants in both groups who selected "other" specified that this included livestreaming, game commentary, and reaction videos.

7.2.3 Subtitle preferences

Participants' subtitle preferences were collected through Questions 10-15 of the pre-task questionnaire, which addressed their general experiences with subtitles, attitudes towards SDH, and preferred subtitle features. **Table 7-3** presents the results.

	SDH group (n=17)		ST group (n=15)		Total
	n	%	n	%	n
Subtitle reading difficulty					
Yes	3	17.6%	3	20.0%	6
No	6	35.3%	5	33.3%	11
Sometimes	8	47.1%	7	46.7%	15

Choosing videos by subtitles					
Only with subtitles	8	47.1%	12	80.0%	20
Prefer subtitles when available	3	17.6%	2	13.3%	5
No preference	6	35.3%	1	6.7%	7
Subtitles/sign language preference					
Subtitles only	8	47.1%	7	46.7%	15
Sign language only	3	17.6%	2	13.3%	5
Both	5	29.4%	5	33.3%	10
Neither	1	5.9%	1	6.7%	2
Used SDH before					
Yes	11	64.7%	12	80.0%	23
No	6	35.3%	3	20.0%	9
Helpfulness of SDH					
Yes	11	64.7%	9	60.0%	20
No	2	11.8%	3	20.0%	5
Depends	4	23.5%	3	20.0%	7
Preferred SDH content					
Speaker identification	3	17.6%	2	13.3%	5
Paralinguistic information	1	5.9%	4	26.7%	5
Background music	0	0.0%	2	13.3%	2
Sound effects	0	0.0%	0	0.0%	0
Silence or quiet moments	0	0.0%	0	0.0%	0
All of the above	13	76.5%	7	46.7%	20

Table 7-3. Subtitle preferences of participants (Phase II)

When asked whether they found it difficult to read subtitles when watching videos, similar proportions were observed in the SDH group (n = 3, 17.6%) and the ST group (n = 3, 20.0%). Eleven participants (34.4%) noted no difficulty, comprising six participants (35.3%) in the SDH group and five (33.3%) in the ST group. The majority of participants (n = 15, 46.9%) admitted that they sometimes found it difficult to read subtitles, although not consistently. This included eight participants (47.1%) from the SDH group and seven (46.7%) from the ST group.

Regarding the influence of subtitles on video selection, this preference was stronger in the ST group (n = 12, 80.0%) than in the SDH group (n = 8, 47.1%). Five participants (15.6%) indicated that they tended to choose videos with subtitles when available (SDH: n = 3, 17.6%; ST: n = 2, 13.3%). A smaller proportion (n = 7, 21.9%) expressed no preference for subtitles, including six participants from the SDH group (35.3%) and one from the ST group (6.7%).

As for accessibility preferences, subtitles were the most frequently preferred method, chosen by eight participants from the SDH group (47.1%) and seven from the ST group (46.7%). Five participants (15.6%) preferred sign language (SDH: n = 3, 17.6%; ST: n = 2, 13.3%). Ten participants (31.3%) reported a combination of subtitles and sign language, evenly distributed across both groups (SDH: n = 5, 29.4%; ST: n = 5, 33.3%). Three participants (6.3%) indicated no preference for either option (SDH: n = 1, 5.9%; ST: n = 1, 6.7%).

Interestingly, a majority of participants (n = 23, 71.9%) reported having previously used SDH services, despite the inadequate provision of SDH in China. This included 11 participants (64.7%) from the SDH group and 12 (80.0%) from the ST group. Nine participants (28.1%) indicated that they had never used SDH before (SDH: n = 6, 35.3%; ST: n = 3, 20.0%).

The next question asked participants whether they considered SDH subtitles helpful. Twenty participants (62.5%) acknowledged their helpfulness (SDH: n = 11, 64.7%; ST: n = 9, 60.0%), while five participants (15.6%) felt that SDH was not helpful (SDH: n = 2, 11.8%; ST: n = 3, 20.0%). Neutral responses were also present in both groups, with seven participants

(21.9%) indicating that the helpfulness of SDH depended on the situation (SDH: n = 4, 23.5%; ST: n = 3, 20.0%).

The final question addressed participants' preference for the types of non-verbal information to be included in SDH. The majority (n = 20, 62.5%) selected "all of the above," which covered speaker identification, paralinguistic information, background music, sound effects, and silence (SDH: n = 13, 76.5%; ST: n = 7, 46.7%). Five participants (15.6%) selected speaker identification alone (SDH: n = 3, 17.6%; ST: n = 2, 13.3%), and another five (15.6%) selected only paralinguistic information (SDH: n = 1, 5.9%; ST: n = 4, 26.7%). Two participants, both from the ST group, chose background music as their sole preference.

7.3 Eye-tracking results

Descriptive statistics

Following the procedure outlined in Section 5.5.4.1, the results are first presented with descriptive statistics for each eye-tracking metric across the two groups in **Table 7-4**.

Overall, the SDH group showed higher mean values than the ST group in two indicators: the percentage of fixation duration (M = 36.30, SD = 16.14 vs. M = 26.07, SD = 7.50) and the number of fixations (M = 1217.41, SD = 311.62 vs. M= 835.07, SD = 167.22). To the contrary, the ST group recorded higher mean values than the SDH group for the average duration of fixations (M = 296.47, SD = 111.34 vs. M = 243.06, SD = 69.84) and the average pupil diameter (M = 4.11, SD = 0.72 vs. M = 3.89, SD = 0.76).

	Mean	SD	Median	Min-Max	Range
Percentage of fixation duration					
SDH group	36.30	16.14	34.02	6.92-80.68	73.76
ST group	26.07	7.50	26.30	15.29-42.52	27.23
Average duration of fixations					
SDH group	243.06	69.84	229	131-441	310
ST group	296.47	111.34	269	129-542	413
Number of fixations					
SDH group	1217.41	311.62	1266	251-1620	1369
ST group	835.07	167.22	812	485-1083	598
Average pupil diameter					
SDH group	3.89	0.76	3.90	2.67-5.33	2.66
ST group	4.11	0.72	3.94	3.04-5.83	2.79

Table 7-4. Descriptive statistics of eye-tracking metrics

As a complement, a set of boxplots (**Figure 7-1**) is provided to illustrate details that were not evident from the descriptive statistics. First, the SDH group outperformed the ST group in two measures: the percentage of fixation duration and the number of fixations. In both metrics, the SDH group displayed wider interquartile ranges, suggesting greater dispersion and within-group variability. Second, the SDH group recorded lower values than the ST group in two measures: the average duration of fixations and the average pupil diameter. For these two measures, the ST group showed higher means and medians. Lastly, outliers appeared in all measures, most notably in the percentage of fixation duration and the

average duration of fixations. This indicates that some participants fixated on subtitles for unusually long or short periods.

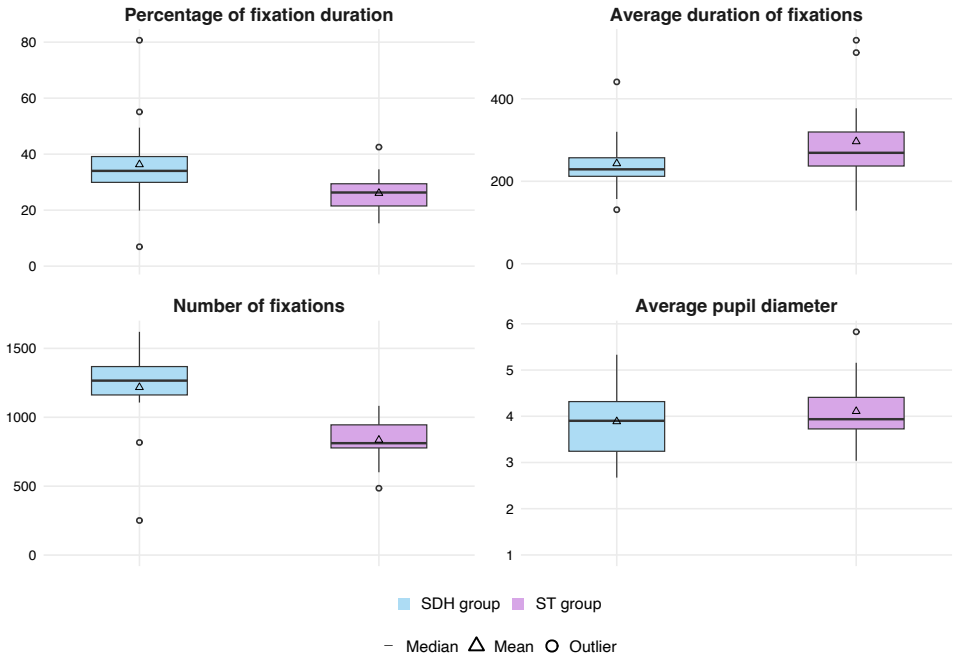


Figure 7-1. Boxplots of eye-tracking metrics by group

Multivariate analysis

A MANOVA test was carried out to see whether the four eye-tracking measures jointly differed between SDH and ST groups. Accordingly, the MANOVA yielded a significant difference between the two groups across the four combined eye-tracking measures, *Pillai's Trace* = 0.412, $F(4, 27) = 4.73$, $p = 0.005$, $\eta^2 = 0.41$. This result ($p = 0.005 < 0.05$) indicates that, overall, the four eye-tracking measures jointly distinguished between the two groups.

Since the MANOVA yielded a significant effect, follow-up tests were subsequently conducted to determine which individual measures contributed to the group differences.

Follow-up comparisons

The follow-up analyses began with the Shapiro-Wilk tests to evaluate the normality of each eye-tracking variable. Accordingly, Shapiro-Wilk tests were performed for each metric, and the results are summarised in **Table 7-5**.

Shapiro-Wilk Normality Test ($p > 0.05 \rightarrow$ Normally distributed)			
	Group	*p*	Normally distributed
Percentage of fixation duration	SDH group	0.103	No
	ST group	0.701	No
Average duration of fixations	SDH group	0.070	No
	ST group	0.100	No
Number of fixations	SDH group	**0.003**	Yes
	ST group	0.507	No
Average pupil diameter	SDH group	0.710	No
	ST group	0.396	No

Table 7-5. Shapiro-Wilk normality tests for eye-tracking metrics

Based on the Shapiro-Wilk results, variables meeting the assumption of normality were analysed using Student's *t*-tests, whereas those violating this assumption were examined using the Mann-Whitney *U* test instead. For variables analysed with *t*-tests, the assumption of homogeneity of variances was also checked using Levene's test. The results of the Student's *t*-tests and Mann-Whitney *U* tests are listed in **Table 7-6**, and the corresponding effect sizes are presented below.

- **Percentage of fixation duration:** Levene's test indicated that the assumption of homogeneity of variances was met ($p = 0.119 > 0.05$). A Student's t-test showed that the SDH group devoted a significantly higher proportion of fixation time to the subtitle area compared to the ST group, $t(30) = 2.25$, $p = 0.032$, mean difference = 10.24, 95% CI [0.94, 19.54]. The effect size was large (Cohen's $d = 0.80$), which indicated that the magnitude of this difference was substantial in practice. This implies that SDH viewers devoted more visual attention to subtitles and showed higher engagement with subtitle information.

- **Average duration of fixations:** Levene's test confirmed homogeneity of variances ($p = 0.165 > 0.05$). A Student's *t*-test suggested no significant difference in the average duration of fixations between SDH and ST groups, $t(30) = -1.65$, $p = 0.110$, mean difference = − 53.41, 95% CI [−119.66, 12.85]. The calculated effect size was medium (Cohen's $d = -0.58$). Although the effect size was moderate, this difference did not reach statistical significance. Therefore, it cannot be inferred that the two groups showed a substantial difference in the average duration of fixations.

- **Number of fixations:** A Mann-Whitney *U* test indicated that the SDH group (mean rank = 22.71) made significantly more fixations than the ST group (mean rank = 9.47), $U = 22.00$, $Z = -3.98$, $p < 0.001$, $r = 0.70$, indicating a large effect size. This demonstrates that the SDH group made significantly more fixations on the subtitle area than the ST group, which indicates greater visual attention to subtitle information in the SDH condition.

- **Average pupil diameter:** Levene's test showed that the assumption of homogeneity of variances was met ($p = 0.724 > 0.05$). A Student's *t*-test found no significant difference in average pupil diameter between the SDH group (M = 3.89, SD = 0.76) and the ST group (M = 4.11, SD = 0.72), $t(30) = -0.84$, $p = 0.406$, mean difference = –0.22, 95% CI [–0.76, 0.32]. The effect size was small (Cohen's $d = -0.30$), which indicates that the difference in pupil size between the two groups was negligible. From this non-significant difference, it can be inferred that average pupil diameters were similar across the two subtitle conditions, which in turn suggests comparable levels of processing effort, as pupil diameter is an established indicator of cognitive load.

Eye-tracking metric	Student's *t*-test/ Mann-Whitney U test	*p*-value	Significance
Percentage of fixation duration	Student's *t*-test	**0.032**	**Significant**
Average duration of fixations	Student's *t*-test	0.110	Not significant
Number of fixations	Mann-Whitney *U* test	**0.001**	**Significant**
Average pupil diameter	Student's *t*-test	0.406	Not significant

Table 7-6. Student's *t*-tests and Mann-Whitney *U* tests for eye-tracking metrics

Multiple-comparison correction

Lastly, it was necessary to apply the Holm-Bonferroni correction to adjust the *p*-values obtained from multiple follow-up tests. The adjusted *p*-values were then calculated and are presented in **Table 7-7**.

It can be seen from **Table 7-7** that only the effect for the number of fixations remained significant (adjusted $p = 0.004$). This finding further confirms that the SDH group devoted more visual attention to subtitle information than the ST group. However, the effects for the percentage of fixation duration (adjusted $p = 0.096$), average fixation duration (adjusted $p = 0.220$), and average pupil diameter (adjusted $p = 0.406$) were not significant. These three metrics, particularly average fixation duration and average pupil diameter, which are direct indicators of cognitive load, showed no statistically significant group differences.

Based on these results, it can be argued that participants in the SDH group paid greater visual attention than those in the ST group, but there is insufficient evidence to conclude that they invested higher cognitive effort.

Eye-tracking metrics	Raw *p*-value	Adjusted *p*-value (after Holm-Bonferroni correction)	Significance
Percentage of fixation duration	0.032	0.096	Not significant
Average duration of fixations	0.110	0.220	Not significant
Number of fixations	0.001	**0.004**	**Significant**
Average pupil diameter	0.406	0.406	Not significant

Table 7-7. Holm-Bonferroni correction for eye-tracking metrics

7.4 Post-task questionnaire results

After completing the eye-tracking experiment, each participant filled in a post-task questionnaire (see Appendix I) including two parts: a comprehension test (Q1-Q10) and a self-reported 7-point Likert-scale questionnaire (Q11-Q25) on their viewing experience. The latter addressed two main dimensions: narrative engagement (Q11-Q18) and enjoyment (Q19-Q25).

7.4.1 Comprehension test results

The post-task comprehension test comprised ten questions (Q1-Q10), each with four options and only one correct answer. Each correct answer was scored as one, and incorrect answers as zero, giving a maximum possible score of ten. During the experiment, both SDH and ST groups completed the same test after viewing the video with their respective subtitle types. The purpose of the test was to address RQ4:

> Do DHH participants who viewed SDH subtitles achieve better comprehension performance than those who viewed ST subtitles?

Before moving on to the main analyses, the internal consistency of the comprehension test was assessed using the KR-20 to confirm its reliability. The obtained coefficient was 0.700 (see **Table 7-8**), which falls within the acceptable range for internal consistency. Thus, the comprehension scores were considered reliable for subsequent analysis.

	Cronbach's α	Cronbach's α based on standardised items	Number of items
Comprehension test (N=32)	**0.700**	0.709	10

Table 7-8. KR-20 value for the comprehension test

With the reliability of the test established, **Table 7-9** presents the descriptive statistics for the comprehension test by group. It can be seen that the SDH group achieved a higher mean score (M = 8.06, SD = 2.436) than the ST group (M = 6.53, SD = 1.685). The median score was also higher in the SDH group (Median = 9) than in the ST group (Median = 7). The SDH group further exhibited a broader score range (1-10) compared with the ST group (4-10). Finally, it is also notable that eight participants in the SDH group obtained full scores, whereas only one participant in the ST group achieved a full score.

	Mean	SD	Median	Min-Max	Range	Number of participants achieving full scores
SDH group	8.06	2.436	9	1-10	9	8
ST group	6.53	1.685	7	4-10	6	1

Table 7-9. Descriptive statistics of comprehension test

In addition, **Figure 7-2** visualises the distribution of comprehension test scores by group. It is evident that the SDH group showed a broader interquartile range, reflecting greater

variability among participants, while the ST group displayed a more compact distribution, with scores clustering more closely but at a generally lower level. Outliers were observed in both groups: an unusually low score in the SDH group and an unusually high score in the ST group.

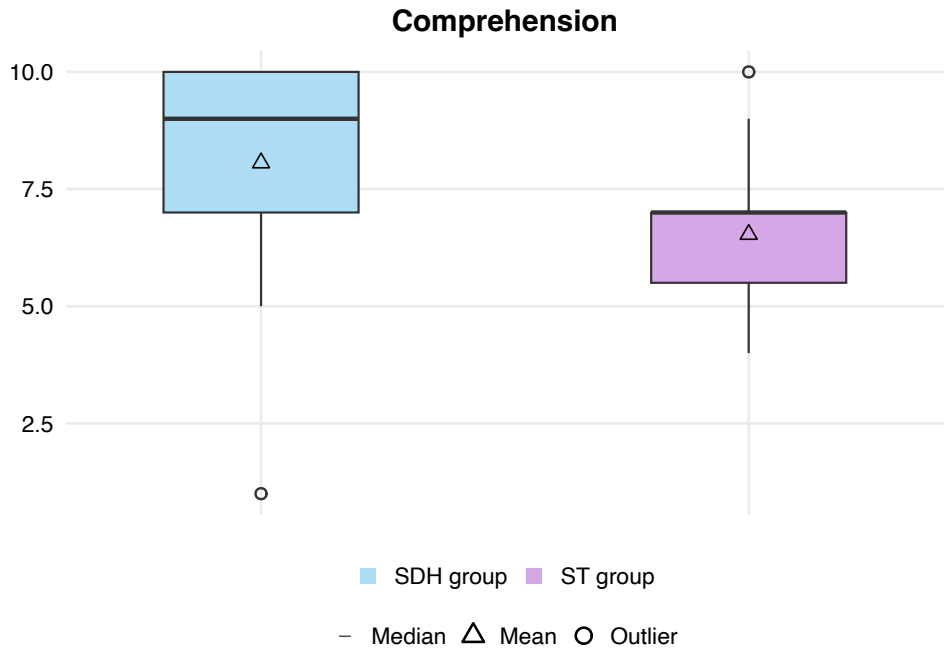


Figure 7-2. Boxplots of comprehension test by group

Following the descriptive summaries of comprehension scores, the next step was to determine the appropriate statistical test for between-group comparison. To this end, the normality of comprehension scores was first assessed. The results (see **Table 7-10**) revealed the SDH group significantly deviated from normality (p = 0.002), while the ST group did not (p = 0.479).

Shapiro-Wilk Normality Test ($p > 0.05$ → Normally distributed)			
	Group	*p*	Normally distributed
Comprehension scores	SDH group	**0.002**	Yes
	ST group	0.479	No

Table 7-10. Shapiro-Wilk normality tests for comprehension test

Given this violation in the SDH group, the non-parametric Mann-Whitney *U* test was applied to examine whether the two groups differed significantly in comprehension performance. The test revealed a statistically significant difference between the groups ($U = 64.50$, $Z = -2.42$, $p = 0.015$), with the SDH group (mean rank = 20.21) outperforming the ST group (mean rank = 12.30). It confirms that the advantage of SDH subtitles reached statistical significance.

To assess the magnitude of this difference, the effect size (*r)* for the Mann-Whitney *U* test was computed. The obtained value ($r = 0.43$) corresponds to a medium effect (Cohen, 1988: 79-80). With such an effect size, it can be said that, in real-world viewing contexts, Chinese DHH viewers with hearing loss may experience a notable improvement in comprehension when SDH subtitles are provided.

Building on the overall score comparison, I looked more closely at each individual question to see how the two groups performed. This analysis added detail to the overall findings by revealing whether the SDH group's advantage was consistent across the test items or specific to certain ones. Moreover, for questions where the between-group difference was statistically significant, it was worth revisiting the question content to identify what types of non-verbal audio information in the SDH subtitles may have supported DHH participants in making sense of the scene.

Towards this end, the accuracy rates for each item were calculated for both groups (see **Table 7-11**). The results showed that, out of the ten questions, the SDH group achieved higher accuracy rates on eight. That said, the ST group outperformed the SDH group on two questions (Q5 and Q8).

Test question number	SDH group (n=17)		SDH group (n=15)		Fisher's exact test
	Correct responses	Accuracy rate	Correct responses	Accuracy rate	*p*
Q1	13/17	76.5 %	8/15	53.3 %	0.266
Q2	14/17	82.4 %	11/15	73.3 %	0.678
Q3	11/17	64.7 %	9/15	60 %	0.536
Q4	16/17	94.1 %	7/15	46.7 %	**0.005**
Q5	16/17	**94.1 %**	15/15	**100 %**	1.000
Q6	13/17	76.5 %	9/15	60.0 %	0.450
Q7	13/17	76.5 %	10/15	67.7 %	0.699
Q8	11/17	64.7 %	4/15	26.7 %	**0.042**
Q9	14/17	**82.4 %**	13/15	**86.7 %**	1.000
Q10	16/17	94.1 %	12/15	80.0 %	0.319

Table 7-11. Accuracy rates and Fisher's exact tests for comprehension questions

Subsequently, Fisher's exact tests were performed for each item to determine whether the differences in accuracy between the two groups were statistically significant and to identify the items meeting this criterion. The results, as displayed in **Table 7-11**, showed that only Q4 ($p = 0.005 < 0.05$) and Q8 ($p = 0.042 < 0.05$) reached statistical significance. Accordingly, these two items were revisited in the comprehension test for closer examination, as

significant differences at the item level may point to specific subtitle features that influence comprehension. To illustrate this, the corresponding questions are presented in **Table 7-12.**

Question 4 in the comprehension test	
In this scene, who said the line “Bro! My bristles are going to fall off” and what was their tone? A.The Crow, feeling shocked B.The Boar, feeling pain C.The Bear, feeling sad D. The Bull, feeling happy	
SDH subtitles	ST subtitles
(惨叫)大哥 毛都没了 *In English:* [Screaming] Bro, my bristles are going to fall off.	大哥 毛都没了 *In English:* Bro, my bristles are going to fall off.
Question 8 in the comprehension test	
When the Boar asked the Bear if the Crow could come back, how did Bear respond? A. Crow will definitely come back B. Crow cannot come back C. Crow might not come back D. Crow has already come back	
SDH subtitles	ST subtitles
(冷漠地反问）回来？被狼大人抓走还能回来？ *In English:* [Reply coldly] Back? Can you come back after being captured by Lord Wolf?	回来 被狼大人抓走还能回来 *In English:* *Back, can you come back after being captured by Lord Wolf.*

Table 7-12. Questions 4 and 8 of the comprehension test

Q4 was designed to test viewers' ability to identify the speaker and to recognise how the line was delivered. In the ST version, all dialogue appeared in the same white font, with no indication of tone. By contrast, in the SDH version, the speaker was colour-coded in green and their tone was signposted in brackets. These additional cues may have helped participants in the SDH group to locate the speaker more quickly and to interpret the intended delivery more accurately, which could in turn explain their stronger performance on this question.

Q8 required viewers to infer the correct answer from the character's tone of voice, which in this case was a cold, rhetorical question. If viewers failed to recognise the tone, the line could easily have been misinterpreted as a neutral statement rather than one carrying sarcasm. In the SDH version, however, the subtitle was prefaced with [Reply coldly], which gave viewers advance access to the speaker's emotion and tone. The use of a question mark further reinforced that the line was rhetorical rather than straightforward. In the ST subtitles, instead, no tone indication was provided, leaving viewers to rely solely on the line itself, which increased the risk of misinterpretation.

Looking at Questions 4 and 8 together, it is clear that the way tone is marked can make a decisive difference to how DHH viewers understand a scene. In both cases, the SDH subtitles gave viewers clear and explicit access to paralinguistic information, described in brackets, so that they could locate the speaker more easily, grasp how the line was delivered, and fit that delivery into the unfolding story. Furthermore, colour-coding and punctuation further supported comprehension. Without these hints, as in the ST subtitles, DHH viewers had to rely on the words themselves to infer tone, an approach that left greater room for misinterpretation.

To sum up, the comprehension test results demonstrate that the SDH group achieved superior performance compared with the ST group, thus addressing RQ4. Yet, a closer look at the individual items found that this advantage did not extend to all comprehension questions.

7.4.2 Engagement and enjoyment results

The post-task questionnaire (see Appendix I) consisted of 15 Likert-scale items (Q11-Q25), with the first eight measuring engagement (e.g., "I was emotionally moved by the ending of the story in this video") and the remaining seven measuring enjoyment (e.g., "I found the video interesting, and it made me feel relaxed and happy"). Each item was rated on a 7-point scale: strongly agree, agree, somewhat agree, neither agree nor disagree, somewhat disagree, disagree, and strongly disagree. For analysis, responses were coded numerically from 7 to 1, with strongly agree assigned a score of 7 and strongly disagree a score of 1. The results of this questionnaire were used to address RQ5:

> What differences can be observed in the overall viewing experience of Chinese DHH participants, in terms of engagement and enjoyment, when watching videos with SDH subtitles versus ST subtitles?

To assess the reliability of the Likert-scale questionnaire, Cronbach's α coefficients were calculated for the two subscales (engagement and enjoyment). As shown in **Table 7-13**, the engagement subscale achieved a Cronbach's α of 0.835, indicating good internal consistency, while the enjoyment subscale achieved a Cronbach's α of 0.840, also reflecting good reliability. With such reliability levels, the data can confidently be used for the subsequent analyses.

	Cronbach's α	Cronbach's α based on standardised items	Number of items
Engagement (N=32)	**0.835**	0.840	8
Enjoyment (N=32)	**0.853**	0.852	7

Table 7-13. Cronbach's α values for engagement and enjoyment

Then, descriptive statistics were used to summarise the engagement and enjoyment scores of both groups. As presented in **Table 7-14**, for engagement, the SDH group (M = 5.81, SD = 0.94) had a slightly higher mean score than the ST group (M = 5.35, SD = 0.69). The median engagement score was 5.75 for the SDH group and 5.63 for the ST group, with scores ranging from 4 to 7 and from 4.13 to 6.25, respectively. For enjoyment, the SDH group (M = 6.12, SD = 0.90) also showed a slightly higher mean than the ST group (M = 6.00, SD = 0.74). The median enjoyment score was 6.43 for the SDH group and 6.00 for the ST group, with ranges of 3.86-7 and 4.86-7, respectively. These descriptive results suggest slightly higher engagement and enjoyment in the SDH group compared to the ST group.

	Mean	SD	Median	Min-Max	Range
Engagement					
SDH group	5.81	0.94	5.75	4-7	3
ST group	5.35	0.69	5.63	4.13-6.25	2.12
Enjoyment					
SDH group	6.12	0.90	6.43	3.86-7	3.14
ST group	6.00	0.74	6.00	4.86-7	2.14

Table 7-14. Descriptive statistics of engagement and enjoyment

As a visual complement to these descriptive results, **Figure7-3** displays the boxplots of engagement and enjoyment scores for both groups. As shown, for engagement, the SDH group showed a broader interquartile range, suggesting greater variability among participants. In contrast to the SDH group, the ST group presented a more compact distribution, indicating more consistent responses. For enjoyment, the two groups performed very similarly. Both groups had comparable interquartile ranges, though the SDH group displayed a wider overall spread.

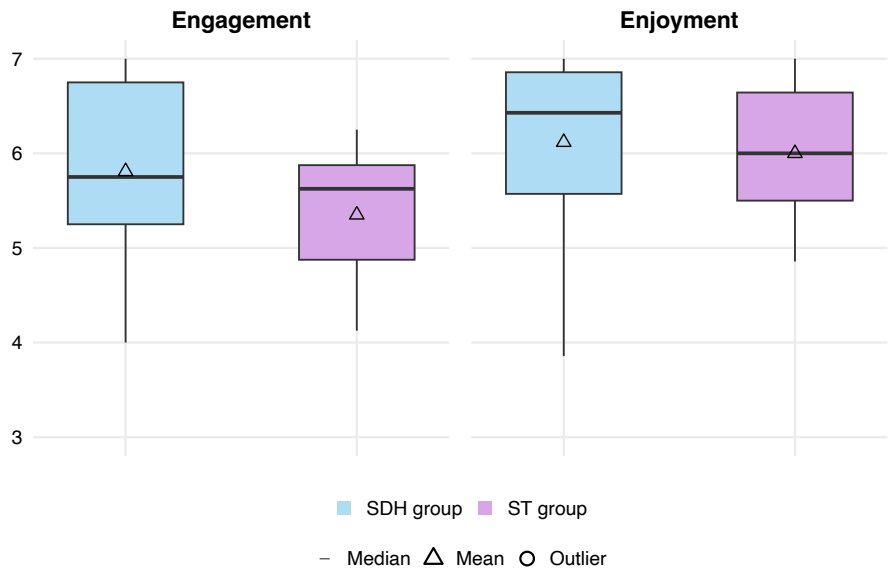


Figure 7-3. Boxplots of engagement and enjoyment by group

Next, the analysis aimed to test whether these differences between the two groups were statistically significant. As a preliminary step, the normality of the engagement and enjoyment scores was assessed using the Shapiro-Wilk test, based on which either Student's *t*-test or the Mann-Whitney *U* test were applied.

- **Engagement:** As presented in **Table 7-15**, the engagement scores for both groups met the normality assumption (SDH group: $p = 0.205$; ST group: $p = 0.200$; both > 0.05). Consequently, the Student's t-test was conducted to assess whether mean engagement scores differed between SDH and ST groups. Prior to the *t*-test, Levene's test confirmed the assumption of homogeneity of variances ($p = 0.396 > 0.05$). The *t*-test yielded $t(30) = 1.55$, $p = 0.131$, which is greater than the conventional significance threshold of $\alpha = 0.05$, indicating no statistically significant difference between the groups. The corresponding effect size, calculated as Cohen's *d*, was 0.55, representing a medium effect (Cohen, 1988: 25-26). This suggests a medium practical difference, although the result did not reach statistical significance.
- **Enjoyment:** As shown in **Table 7-15**, the Shapiro-Wilk test for enjoyment scores presented a mixed outcome: the SDH group violated the normality assumption ($p = 0.020 < 0.05$), whereas the ST group satisfied it ($p = 0.278 > 0.05$). Because one group showed a violation of the normality assumption, a non-parametric Mann-Whitney *U* test was selected to compare the two groups. This analysis produced $U = 112.00$, $Z =$

−0.59, and $p = 0.557$ ($p > 0.05$), indicating no statistically significant difference. Again, the effect size was calculated, yielding $r = 0.10$, which corresponds to a small effect (Cohen, 1988: 79-80). This means that, in practice, the difference between the two groups was minimal and did not reach statistical significance.

Shapiro-Wilk Normality Test ($p > 0.05$ → Normally distributed)			
	Group	*p*	Normally distributed
Engagement scores	SDH group	0.205	No
	ST group	0.200	No
Enjoyment scores	SDH group	**0.020**	Yes
	ST group	0.278	No

Table 7-15. Shapiro-Wilk normality tests for engagement and enjoyment

According to the above statistical analysis, there was no significant difference in narrative engagement and enjoyment between the two groups. In view of this, it was considered relevant to look beyond the overall analysis. Thus, responses to each individual questionnaire question were graphically presented in order to discover some details that were not apparent in the overall results. Two figures were then produced (**Figure 7-4** and **Figure 7-5**) to present the distribution of responses for narrative engagement (Q11-Q18) and enjoyment (Q19-Q25) across the two groups, respectively.

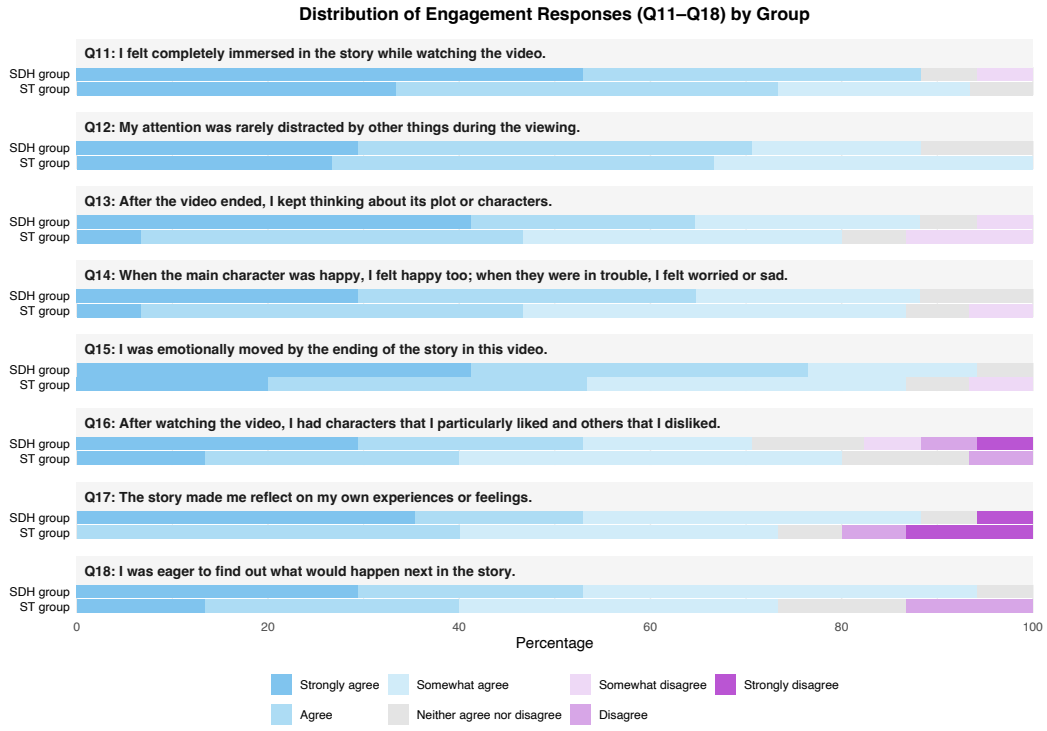


Figure 7-4. Distribution of engagement responses (Q11-Q18) by group

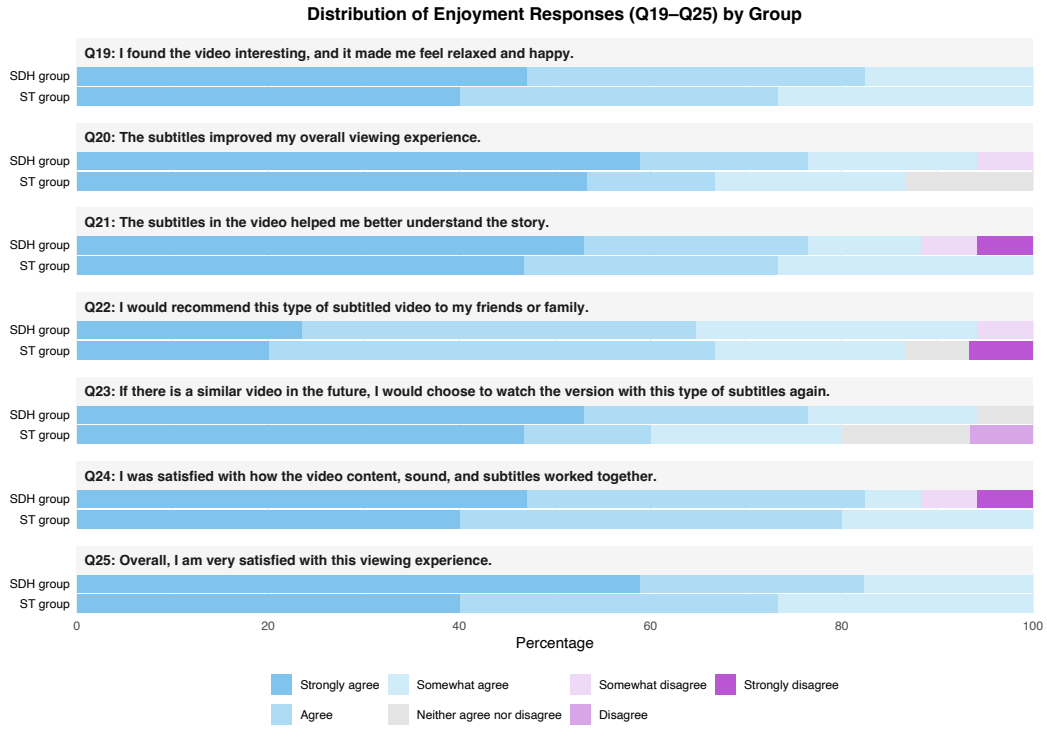


Figure 7-5. Distribution of enjoyment responses (Q19-Q25) by group

In terms of narrative engagement, the SDH group generally provided stronger positive responses across a greater number of questions compared with the ST group. For immersion (Q11, Q13, Q18), a higher proportion of SDH participants selected “agree” or “strongly agree,” and the ST group’s responses were more evenly distributed across the moderate positive categories. On attention (Q12), however, the ST group showed a higher proportion of agreement, while the SDH group presented lower levels. This difference may indicate that the presence of SDH subtitles introduced some distraction. With regard to emotional engagement (Q14, Q15), both groups tended towards agreement, although the SDH group contained a higher proportion of “strongly agree” responses. Regarding character preference (Q16), the SDH group included some “strongly disagree” responses, which did not appear in the ST group. Finally, for self-reflection (Q17), a considerable proportion of participants in the SDH group selected “strongly agree,” but none of the ST participants did so. It may imply that viewers in the SDH group were better able to establish a stronger connection with the characters and relate the story to their own experiences. Altogether, the SDH group tended to agree more strongly than the ST group across most questions. However, two exceptions stood out: participants in the SDH group found it harder to maintain attention (Q12) and were more critical in their evaluation of character preference (Q16).

For enjoyment, both groups gave generally positive responses, although the SDH group’s answers were more clearly concentrated at the higher end of the scale (“agree” and “strongly agree”). In terms of overall satisfaction (Q19, Q25), more participants in the SDH group selected “strongly agree” than in the ST group, indicating higher levels of satisfaction with the viewing experience. Regarding the role of subtitles (Q20, Q21, Q24), the SDH group also presented a greater proportion of positive responses, yet this trend was not consistent, as a noticeable share of participants selected “disagree” or even “strongly disagree,” a tendency less frequent in the ST group. This pattern suggests that this type of subtitles (i.e., SDH subtitles) was not fully accepted by all viewers in the SDH group. On willingness to recommend or rewatch (Q22, Q23), the SDH group showed a greater concentration of

"agree" and "strongly agree," while negative ratings appeared more often in the ST group. Collectively, the SDH group reported higher levels of enjoyment than the ST group, particularly in terms of satisfaction and willingness to recommend or rewatch, yet some negative feedback from the SDH group suggests that SDH subtitles were not embraced by all viewers.

To conclude this section, the comparison between SDH and ST groups showed no significant differences in overall engagement and enjoyment. Nonetheless, beyond the statistical results, a closer look at the questionnaire data provided additional observations related to RQ5, specifically that SDH subtitles appeared to offer a slight advantage in both dimensions. The SDH group tended to give stronger agreement across most survey questions, and this was evident in their responses related to immersion, emotional involvement, overall satisfaction, and a stronger willingness to recommend or rewatch the video. At the same time, their answers also pointed to certain reservations: some SDH participants indicated greater difficulty in maintaining attention during viewing, and their views on character preference and on the role of SDH subtitles were more divided, with a few even expressing negative feedback. In comparison, the ST group gave more moderate ratings, with responses clustering around the middle-to-positive categories.

7.5 Correlation analyses

7.5.1 Correlation results for the overall sample

Across the whole sample (N = 32), a Spearman correlation test was run for seven independent variables: percentage of fixation duration, average duration of fixations, number of fixations, average pupil diameter, comprehension, engagement, and enjoyment. The correlation heat map of these variables is presented in **Figure 7-6**. To be clear, values in the upper triangle are Spearman correlation coefficients (deeper blue = stronger positive; deeper purple = stronger negative), while values in the lower triangle (grey-shaded cells) are two-tailed *p*-values; and cells marked with an asterisk are considered statistically

significant at the 0.05 level. To put it simply, deeper blue or purple colours indicate stronger relationships between two measures, and an asterisk marks cases where the relationship is statistically reliable. The interpretation of these correlations is organised in three aspects:

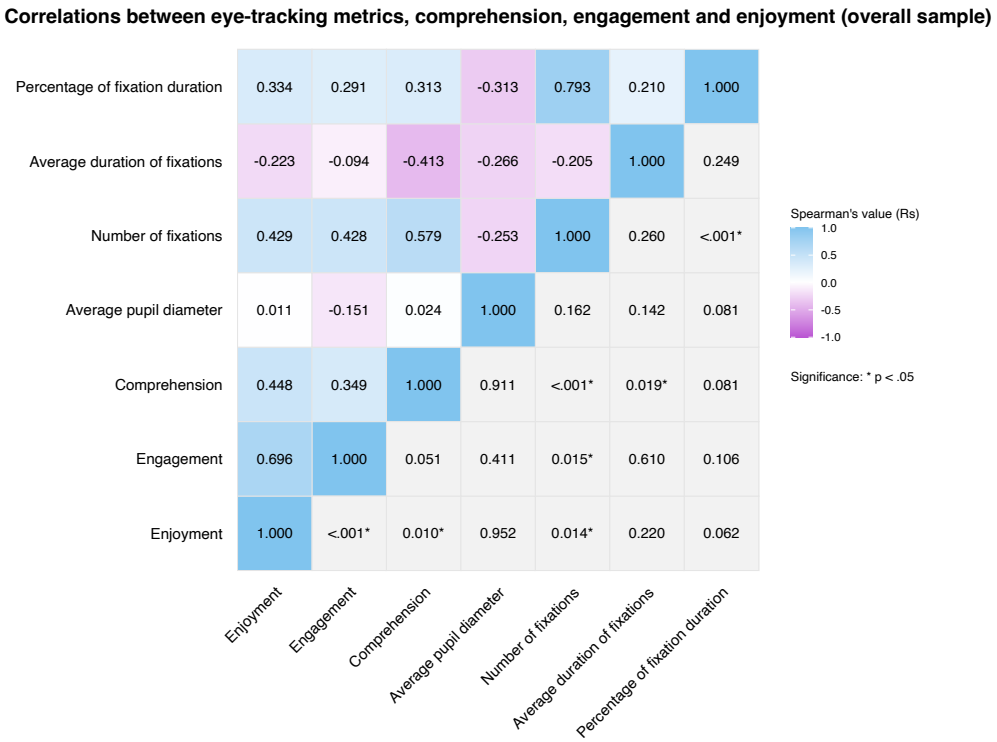


Figure 7-6. Correlation analysis (overall sample)

Correlation between eye-tracking and comprehension: Overall, participants with higher comprehension scores tended to make a greater number of fixations, and those with lower scores usually showed longer average fixation durations. More specifically, comprehension was positively correlated with the number of fixations (r_s = 0.579, p = 0.001). Participants who looked at the subtitles more often were generally better at understanding the AV content. In contrast, comprehension was negatively correlated with average fixation duration (r_s = –0.413, p = 0.019). Participants who spent more time on each fixation often showed poorer comprehension, possibly due to the extra processing effort required. Finally, neither the percentage of fixation duration (r_s= 0.313, p = 0.081) nor the average pupil diameter (r_s = 0.024, p = 0.911) showed any significant association with comprehension.

Correlation between eye-tracking and engagement: In general, a higher level of engagement was associated with a greater number of fixations (r_s = 0.428, p = 0.015). This result suggests that participants who produced more fixation counts on the subtitles tended to report higher engagement with the AV content. Regarding other eye-tracking metrics, the percentage of fixation duration (r_s = 0.291, p = 0.106), the average duration of fixations (r_s = –0.094, p = 0.610), and the average pupil diameter (r_s = –0.151, p = 0.411) showed no significant associations with engagement.

Correlation between eye-tracking and enjoyment: Similarly, a higher level of enjoyment was associated with a greater number of fixations (r_s = 0.429, p = 0.014). This result indicates that participants who made more fixations on the subtitles tended to experience greater enjoyment. Again, the percentage of fixation duration (r_s = 0.334, p = 0.062), the average duration of fixations (r_s = –0.223, p = 0.220), and the average pupil diameter (r_s = 0.011, p = 0.952) showed no significant correlations with enjoyment.

7.5.2 Correlation results by subtitle group

As noted in Section 5.5.4.4, it was necessary to conduct separate Spearman correlations by group to verify whether these associations persisted under different subtitle conditions. The Spearman correlation results for the SDH group are presented first (see **Figure 7-7**). Likewise, the discussion of the correlations is organised into three aspects, as follows:



Figure 7-7. Correlation analysis (SDH group)

Correlation between eye-tracking and comprehension: Within the SDH group, a higher level of comprehension was associated with a larger average pupil diameter (r_s = 0.486, p= 0.048). In other words, participants who exhibited greater pupil dilation performed better on the comprehension test. The remaining eye-tracking metrics did not show significant correlations with comprehension: average fixation duration (r_s = –0.421, p = 0.093), fixation counts (r_s = 0.266, p = 0.301), and percentage of fixation duration (r_s = –0.158, p = 0.545). Hence, these three fixation-based measures cannot be considered reliably associated with comprehension.

Correlation between eye-tracking and engagement: In the SDH group, a higher level of engagement was correlated with a greater number of fixations (r_s = 0.516, p = 0.034). This suggests that participants who produced more fixation counts tended to show higher engagement while watching the video. In contrast, no significant correlations were found between engagement and the other eye-tracking metrics, namely the percentage of fixation duration (r_s = 0.104, p = 0.692), the average duration of fixations (r_s = –0.094, p = 0.720), and

the average pupil diameter (r_s = 0.117, p = 0.654). Given these non-significant results, no reliable links can be established between these three measures and engagement.

Correlation between eye-tracking and enjoyment: In the SDH group, no significant correlations were found between enjoyment and any of the four eye-tracking measures. The number of fixations (r_s = 0.446, p = 0.073), the average duration of fixations (r_s = –0.159, p = 0.541), the percentage of fixation duration (r_s = 0.167, p = 0.522), and the average pupil diameter (r_s = 0.269, p = 0.296) all failed to show reliable associations with enjoyment.

With the SDH group completed, attention now turns to the ST group. The Spearman correlation results for the ST group are shown in **Figure 7-8**, and the analysis again follows the same three aspects:



Figure 7-8. Correlation analysis (ST group)

Correlation between eye-tracking and comprehension: In the ST group, better comprehension was associated with a higher number of fixations and smaller pupil sizes. This was demonstrated by the results showing that comprehension correlated positively

and significantly with the number of fixations (r_s = 0.547, p = 0.035), and negatively and significantly with the average pupil diameter (r_s = –0.596, p = 0.019). The correlations between comprehension and the percentage of fixation duration (r_s = 0.477, p = 0.072), as well as the average duration of fixations (r_s = –0.064, p = 0.821), were positive but did not reach statistical significance.

Correlation between eye-tracking and engagement: Within the ST group, engagement was significantly associated with smaller pupil sizes, as indicated by a negative and significant correlation with average pupil diameter (r_s = –0.558, p = 0.031). However, no significant relationships were found between engagement and the other eye-tracking metrics, including the number of fixations (r_s = 0.206, p = 0.461), the percentage of fixation duration (r_s = 0.244, p = 0.381), and the average duration of fixations (r_s = 0.063, p = 0.823).

Correlation between eye-tracking and enjoyment: In the ST group, higher levels of enjoyment were associated with a greater number of fixations (r_s = 0.587, p = 0.021). That is, participants who made more fixation counts on the subtitles showed higher levels of enjoyment while watching the video. Nevertheless, this relationship did not extend to the other eye-tracking measures: no significant associations were found between enjoyment and the average duration of fixations (r_s = –0.267, p = 0.336), the percentage of fixation duration (r_s = 0.501, p = 0.057), or the average pupil diameter (r_s = –0.367, p = 0.178).

7.5.3 Summary of correlation findings

This section summarises how eye-tracking measures were related to comprehension, engagement, and enjoyment, both in the overall sample and within each group.

As shown in **Figure 7-9**, across the overall sample, comprehension was associated with two eye-tracking metrics: negatively with average fixation duration and positively with the number of fixations. It is reasonable to assume that viewers who produced more fixation counts but for shorter periods tended to achieve higher comprehension. In addition, both engagement and enjoyment showed significant positive correlations with the number of

fixations. This suggests that the more fixation counts viewers made on the subtitles, the more engaged and enjoyable their viewing experience tended to be.

In the SDH group **(see Figure 7-10)**, the number of significant correlations was notably lower than in the overall sample. Comprehension was significantly and positively correlated with average pupil diameter, which shows that better comprehension performance was associated with larger pupil sizes. Engagement showed a significant positive correlation with fixation counts, which indicates that more frequent attention to subtitles was linked to a stronger sense of involvement. Enjoyment, however, did not show any meaningful relationships with the eye-tracking measures.

In the ST group (see **Figure 7-11**), the pattern of correlations differed from that observed in the SDH group. Comprehension was significantly and positively correlated with the number of fixations but negatively with average pupil diameter. This finding indicates that viewers who achieved higher comprehension tended to look at the subtitles more frequently while simultaneously exerting less cognitive effort, as reflected by smaller pupil sizes. Similarly, engagement showed a significant negative correlation with average pupil diameter, meaning that higher levels of engagement were associated with smaller pupil sizes. Furthermore, viewers who made more fixation counts on the subtitles reported higher levels of enjoyment, as evidenced by the positive correlation between enjoyment and fixation counts.

In summary, the correlation analyses identified the number of fixations as the common factor linking eye movements with comprehension and viewing experience, although its role varied across subtitle types. In the SDH group, engagement increased with more fixation counts, and in the ST group, both comprehension and enjoyment were enhanced by more fixation counts on subtitles.

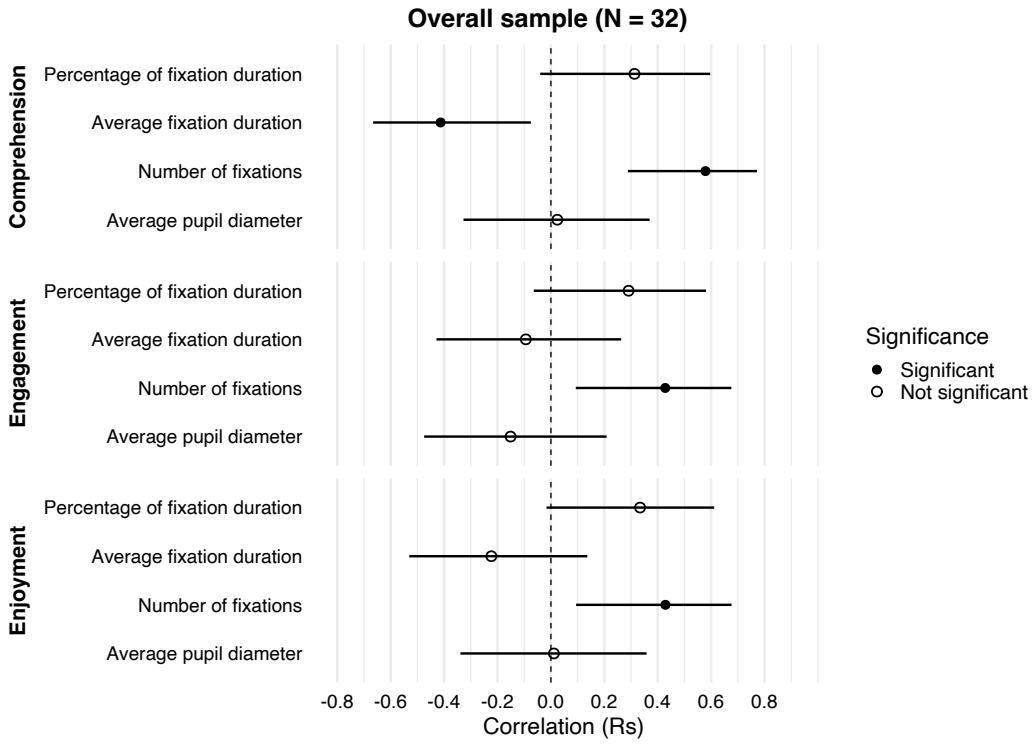


Figure 7-9. Summary of correlation analysis (overall sample)

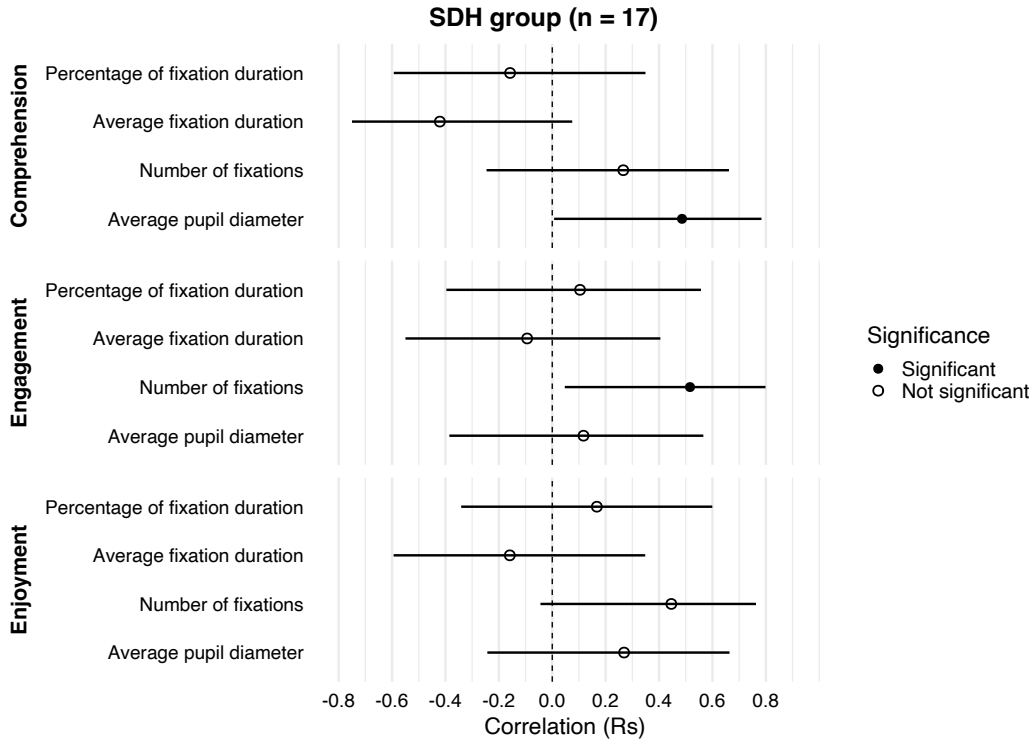


Figure 7-10. Summary of correlation analysis (SDH group)

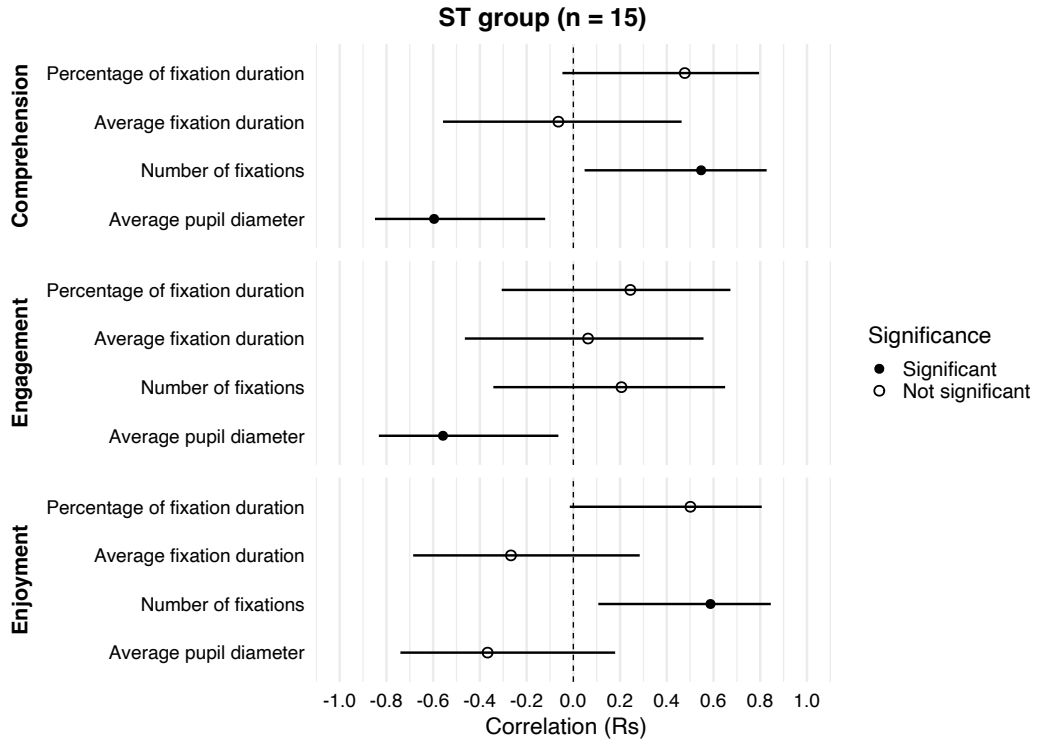


Figure 7-11. Summary of correlation analysis (ST group)

7.6 Discussion of results

7.6.1 Discussion of eye-tracking results (RQ3)

The eye-tracking results were presented in Section 7.3. We found that, of the four eye-tracking measures analysed, only the number of fixations yielded a significant difference between the two groups, where viewers in the SDH group produced higher fixation counts than those in the ST group. As for the other three eye-tracking measures (the percentage of fixation duration, average fixation duration, and average pupil diameter), no significant differences were observed between the two groups.

The higher fixation counts observed in the SDH group can be explained in that SDH subtitles attracted greater visual attention or, alternatively, required viewers to invest more processing effort. However, the number of fixations should be interpreted with caution, as this measure reflects multiple aspects of viewing behaviour and is often considered together with other eye-tracking indices. In studies of cognitive load, for example, fixation

counts are frequently discussed together with fixation duration, since subtitles that require more time or elicit a higher number of fixations are generally regarded as more cognitively demanding (Orrego-Carmona, 2016: 173). In the present data, however, no group differences were found in average fixation duration or pupil diameter, both of which are well-established indicators of cognitive load (see Section 5.5.2.1 for the introduction of these eye-tracking indicators). Therefore, the possibility that SDH subtitles increased viewers' cognitive load can be ruled out. Rather, a more plausible explanation is that SDH subtitles primarily promoted greater visual attention.

This finding is somewhat different from the common assumption that subtitles containing additional visual information may lead to heavier processing demands, as viewers need to read and process extra information, which may increase cognitive effort (Kruger et al., 2016: 172). Despite this assumption, the question of whether the addition of subtitles induces cognitive overload has also been empirically investigated. Perego et al. (2016) compared subtitled and dubbed versions of videos and found that the presence of subtitles did not result in cognitive overload. Another reception study, conducted by Yoon and Kim (2011), compared sign language and subtitled versions among Korean deaf students and likewise found no statistically significant difference in cognitive load. These studies provide further evidence that the addition of a visual source of information does not appear to impose cognitive overload on viewers, thus pointing to an important benefit of SDH subtitles.

Such an outcome invites further consideration of why the additional auditory information in SDH subtitles did not lead to increased cognitive load. One possible explanation is that viewers are capable of processing subtitles in parallel with other visual inputs on screen. As observed by Perego et al. (2018), viewers' subtitle processing is semi-automatic and cognitively efficient. This ability to coordinate information from different channels has been repeatedly confirmed in empirical research (D'Ydewalle et al., 1991; Perego et al., 2010; Perego et al., 2015). Yet, unlike the empirical studies mentioned above, which were limited to hearing participants and ST subtitles, the present study extended this line of research to

DHH participants and SDH subtitles and provided evidence that DHH viewers were likewise able to process SDH subtitles efficiently.

Based on the above discussion, the higher number of fixations observed in the SDH condition should not be interpreted as evidence of increased cognitive load, but rather as an indication of greater visual attention elicited by this type of subtitles. This interpretation is consistent with previous studies suggesting that subtitles do not necessarily cause cognitive overload (Kruger et al., 2014), as viewers are generally able to process subtitle information alongside other visual inputs on screen (D'Ydewalle et al., 1991; Perego et al., 2010). In this sense, SDH subtitles seem to encourage more active engagement with the visual channel instead of imposing additional cognitive effort.

7.6.2 Discussion of comprehension results (RQ4)

The results of the comprehension test were presented in Section 7.4.1, where participants who viewed the SDH subtitles obtained higher comprehension scores than those who watched the same AV material with ST subtitles, and this difference was statistically significant. This finding answers RQ4 by showing that SDH subtitles can indeed facilitate the understanding of AV content among DHH participants.

This finding also corroborates the view that SDH subtitles offer greater advantages than ST subtitles in enhancing comprehension among DHH viewers. It lends empirical support to Tamayo's (2017b: 79) argument, drawn from her research on SDH for deaf children, that making the auditory visual can help DHH viewers achieve better comprehension. While no empirical research has yet directly contrasted SDH and ST subtitles within DHH populations, previous studies (Yoon & Kim, 2011; Gernsbacher, 2015) have shown that the presence of subtitles improves comprehension of AV material among DHH audiences. Since DHH individuals tend to rely more heavily on visual perception to compensate for hearing limitations (Agrawal & Peiris, 2021), SDH subtitles, which are richer in informational content than ST subtitles, better help viewers to interpret non-dialogue sounds and other auditory

information through visual means, consequently improving their overall understanding of AV material to a greater extent.

Although SDH subtitles have been shown to be beneficial for the DHH community, their advantages were not shared equally among viewers. Within the SDH group, comprehension scores varied widely, ranging from as low as 1 to as high as 10, compared with a narrower range of 4 to 10 in the ST group (see **Table 7-9**). This implies that not all DHH participants in the SDH group were able to take advantage of the additional visual information to the same extent. Some relied on it to aid comprehension, whereas others struggled to do so. Such variation may be attributed to individual differences among DHH viewers, including reading speed (Romero-Fresco, 2009; Perego et al., 2016), proficiency in written or different caption styles (Szarkowska et al., 2011; Zárate & Eliahoo, 2014), and literacy skills (Tamayo, 2017b; Aleksandrowicz, 2019), as well as other unexplored factors.

Meanwhile, the advantage of SDH subtitles did not extend to all questions in the comprehension test. This is because the item-level analysis indicated significant group differences in only two out of ten questions (Q4 and Q8), both of which involved paralinguistic information, one related to speaker identification and the other to tone of delivery. In Q4, colour-coded speaker names and bracketed tone cues in the SDH version guided viewers in recognising who was speaking and how each line was meant to sound. In Q8, the tag [Reply coldly] explicitly conveyed sarcasm and prevented misinterpretation. Without such cues, the ST version left tone to be inferred from words alone, often at the cost of meaning. These examples illustrate how making paralinguistic information visible through SDH can make comprehension more attainable for DHH viewers.

To conclude, the comprehension findings demonstrate that SDH subtitles have a clear advantage over ST subtitles in improving comprehension among DHH viewers. However, not all DHH individuals were able to benefit equally from SDH, as the DHH community is diverse, and it is therefore difficult to ensure that every viewer can make full use of such subtitles. Moreover, the effectiveness of SDH may depend on the specific type of

information conveyed. The present results highlight the need for future SDH design to prioritise the inclusion of paralinguistic information related to on-screen characters, as these elements have been shown to be particularly conducive to enhancing comprehension.

7.6.3 Discussion of engagement and enjoyment results (RQ5)

Based on the analysis presented in Section 7.4.2, it was found that there were no significant differences between SDH and ST groups in either engagement or enjoyment.

In fact, such non-significant results concerning the viewing experience are not uncommon in subtitling reception research. Previous studies, such as that by Szarkowska and Gerber-Morón (2018a), also found no significant effects of subtitle speed or language on viewers' enjoyment. In a later study, the same authors (Szarkowska & Gerber-Morón, 2019) concluded that the number of subtitle lines (two or three) had little impact on how much viewers enjoyed the content. That said, this does not imply that measures of viewing experience have never yielded significant results. Some studies have shown statistically significant differences, particularly when the viewing conditions differed more noticeably. For instance, when comparing videos with and without subtitles (Kruger et al., 2016) or with and without sound (Szarkowska et al., 2024b), the effects on viewers' experience were found to be statistically significant. Returning to the present study, the differences between SDH and ST subtitles were likewise relatively subtle, which may explain why no statistically significant effects were found.

Although no significant differences were observed in the overall viewing experience, a more detailed analysis of the questionnaire data on engagement and enjoyment (see **Figures 7-4** and **Figure 7-5**) brought to light some aspects that merit further discussion.

<u>Engagement:</u>

Having analysed the distribution of engagement-related questionnaire items, it became evident that participants in the SDH group expressed stronger agreement than those in the ST group. These positive responses were most apparent in questions related to immersion

and emotional connection with the story. From this, it can be inferred that including sound descriptions may help DHH viewers follow the unfolding of the narrative more closely and, in doing so, foster a stronger sense of engagement with the plot.

At the same time, the analysis also uncovered what might be considered a potential drawback of SDH subtitles. Some participants showed lower levels of attention when watching the SDH version. A possible reason is that the additional on-screen information may have diverted their attention and made it harder to stay focused on the visual content itself. However, this finding can be read together with the eye-tracking results discussed in Section 7.6.1. The SDH group produced more fixations than the ST group, suggesting greater rather than reduced visual attention. Thus, the questionnaire responses may reflect participants' subjective feeling that their attention was divided between the subtitles and the image, instead of indicating an actual decrease in attention. As there was no evidence of increased processing effort, the higher fixation count may instead imply more active visual engagement with SDH subtitles.

Enjoyment:

When it comes to enjoyment, item-level responses showed that both groups gave generally positive ratings, although the SDH group scored slightly higher overall. Participants in this group expressed greater satisfaction with the viewing experience and a stronger willingness to rewatch or recommend the video. These findings are consistent with those of Zárate (2021: 109), whose questionnaire results indicated that nearly 80% of DHH participants agreed that SDH subtitles improve access to enjoyment.

Even so, a certain degree of ambivalence was evident in the present data. Most viewers in the SDH group rated their experience positively, but a small number expressed reservations about the role and necessity of SDH subtitles. This means that not all DHH viewers perceive them as equally beneficial, possibly due to individual differences in reading habits, linguistic competence, or a preference for less textual information. Such mixed responses reflect the

diversity within the DHH community and point to the need for more adaptive subtitling approaches that can accommodate different viewing preferences.

7.6.4 Discussion of correlation analyses

A series of correlation analyses were conducted to investigate the relationships between the four eye-tracking metrics and the three outcome measures (comprehension, engagement, and enjoyment). The results of these analyses are presented in Section 7.5.

In general, the number of fixations stood out as the most consistent indicator. It was found to have significant positive correlations with comprehension, engagement, and enjoyment. That is to say, DHH participants who made more fixation counts on the subtitles tended to process the AV information more attentively and efficiently. This may explain why viewers in the SDH group achieved a better understanding and experienced a more engaging viewing process. This interpretation is in line with the eye-tracking findings discussed earlier (Section 7.6.1), which concluded that a greater number of fixations does not necessarily indicate higher cognitive effort. Rather, it can be interpreted as reflecting stronger visual attention or a heightened interest in the subtitles themselves.

With regard to comprehension, it was positively related to the number of fixations but negatively related to average fixation duration. More specifically, better comprehension performance was associated with more fixation counts but shorter average fixation durations, rather than deeper processing through fewer and longer fixations. This reading pattern implies that viewers may have adopted an efficient and flexible strategy when processing subtitles. Similar findings were documented by Perego et al. (2010), whose eye-tracking evidence also points to cognitively efficient subtitle reading.

As for engagement and enjoyment, both variables showed similar patterns, with higher self-reported scores corresponding to a greater number of fixations. This reinforces the idea that immersion or emotional engagement during viewing is reflected in increased visual attention to the subtitles. Turning to pupil-based measures, no significant associations were found between average pupil diameter and either engagement or enjoyment. In this case,

it seems that physiological arousal may not have varied greatly across participants, or that pupil-based measures were influenced by external factors such as screen brightness (Kruger et al., 2013; Doherty & Kruger, 2018).

When the two groups were analysed separately to verify whether the above overall correlations held across the two subtitle conditions, the results showed distinct differences between them. The most striking difference concerned average pupil diameter, which showed opposing trends and may reflect variation in how processing effort translated into comprehension under the two subtitle conditions. In the SDH group, larger pupil diameters were associated with higher comprehension scores. This means that although viewers devoted more effort while watching, this effort was efficiently channelled into understanding the AV material. In the ST group, however, the processing mechanism was opposite to that observed in the SDH group. Greater pupil dilation was associated with lower levels of both comprehension and engagement. This suggests that the effort invested by participants in the ST group did not translate into better understanding or enhance the viewing experience. It may instead indicate unproductive or misdirected processing, as viewers attempted to infer contextual meaning but ultimately struggled to construct a coherent understanding in the absence of auditory information.

In closing, it is worth summarising some practical implications drawn from the correlation analyses. First, the number of fixations can act as a useful metric for exploring the relationship between eye-movement behaviour, comprehension, and the subjective viewing experience, since consistent correlations were found between this measure and several indicators in this study. Second, group-based correlation analyses also proved valuable, as they helped identify effects that were not evident in the overall sample.

Chapter 8 Conclusions

8.1 Chapter overview

This chapter brings together the main findings of the thesis and reflects on their broader implications. Section 8.2 revisits the five research questions in light of the findings from both phases of the study. Section 8.3 outlines the main contributions of the thesis, and Section 8.4 acknowledges the main limitations of the research. Finally, Section 8.5 points to several directions for future work.

8.2 Addressing RQs

8.2.1 RQ1

RQ1: What key factors influence SDH subtitlers' decision-making processes?

To answer this question, I spoke with nine professional SDH subtitlers through one-to-one semi-structured interviews. Their responses were then analysed using a Grounded Theory approach. After three rounds of coding, six key themes gradually emerged, and each one represents a factor that influences how decisions are made in practice. These themes are: workflow and process optimisation, existing SDH guidelines, technical and practical constraints, audience-centred considerations, the role of subjectivity in interpretation, and differences in content and background.

The first factor identified is workflow and process optimisation. Based on our interviews, we found that subtitlers from different organisations work within different production structures, and that collaboration, time pressure, and feedback cycles all influence how decisions are made. Thus, it can be said that, rather than following a fixed process, their choices are influenced by the realities of the workflow they are part of.

The second factor is related to existing SDH guidelines. It was initially assumed that SDH subtitlers would follow their guidelines strictly, but according to the interviews it turned out

that they do not always follow them rigidly. Instead, they interpret and adapt them depending on the context and, in some cases, identify areas where current guidelines could be improved. In some organisations where SDH guidelines are absent, subtitlers either have to develop their own standards or rely on their own judgement in less strict environments.

Third, technical and practical constraints also have a strong impact. More specifically, limitations related to subtitling software, time, and space on the screen often restrict what can be included in SDH. Under such constraints, subtitlers tend to decide what to include based on whether a sound is visible on screen and whether it is relevant to the plot. The use of emojis in subtitles may also be avoided due to limitations of the subtitling software.

Fourth, audience-centred considerations relate to the needs of DHH viewers. The interviewed SDH subtitlers were clearly aware of the diversity within this audience, and their decisions often reflected an effort to describe music in terms of emotion and atmosphere, rather than naming instruments or musical genres. Additionally, for young viewers, some interviewees mentioned that they would intentionally use simpler language or slower subtitle speeds in children's programmes such as cartoons.

The next factor is one that seems almost unavoidable in SDH production, that is, subtitlers may bring a certain level of subjective interpretation into their descriptions. To deal with this, some interviewees mentioned that they tend to use more objective and descriptive language. It was also found that peer review can be a useful way to minimise personal bias, while also serving as a form of quality control.

The final factor relates to the type of video content, as well as the subtitlers' own background knowledge. On the one hand, programmes such as chat shows, reality shows, and variety shows often involve fast paced and overlapping speech, which makes them more challenging to subtitle. On the other hand, decisions may also depend on the subtitlers' own knowledge. For example, when it comes to describing music, some interviewees mentioned that limited familiarity with music often made them more cautious about naming musical genres or instruments.

8.2.2 RQ2

> RQ2: What implications can be drawn from the study with SDH subtitlers that can inform the production of SDH subtitles?

Looking back at the interviews with nine subtitlers, several practical implications can be summarised, and these helped guide me in creating the SDH subtitles used in the reception study.

At the preparation stage, priority can be given to obtaining original subtitles and video materials from production teams in order to reduce transcription work. Where this is not possible, speech recognition combined with human post-editing can be used to save time while still ensuring accuracy. It is also helpful to do some background research on the video content beforehand, so as to gain a basic understanding of the plot and characters. This may also be useful when subtitlers are dealing with content areas they are less familiar with. Given that there are no national SDH guidelines in China, it is helpful to draw on internationally established standards for SDH subtitle production, as discussed in Chapter 4 and summarised in Section 4.4.

During the production stage, time and space limitations often affect decisions about what should be included. In cases where it is unclear whether a non-verbal sound should be included, two principles can be considered: whether the sound is visible on screen and whether it is relevant to the plot. That said, some room for flexibility may still be needed in exceptional cases. When describing music, it is often more effective to convey its atmosphere or emotional effect rather than specifying genres or instruments. For sound effects, onomatopoeic descriptions are best avoided where clarity and readability are the main priorities. For content aimed at children, simpler language and slower subtitle speeds would be more appropriate. Even beyond children's content, it is important to keep the language clear and accessible, as DHH viewers come from different educational backgrounds.

After subtitle production, peer review can be a useful way to reduce personal interpretation and bias, while also providing an additional layer of quality assurance. More broadly, incorporating review and discussion into the workflow can help reduce overly individualised decisions and improve consistency in subtitle production.

8.2.3 RQ3

> RQ3: How do the gaze patterns of DHH participants differ when watching videos with SDH subtitles compared to those with ST subtitles?

When we examined the eye-tracking data in Section 7.3, we found that the only significant difference between the two groups was in the number of fixations. Participants in the SDH group made more fixations than those in the ST group.

This raises two possible interpretations: SDH subtitles either attracted greater visual attention or required DHH viewers to invest more processing effort. However, as discussed in Section 7.6.1, the latter interpretation is not supported by the other eye-tracking measures, as no significant differences were found in average fixation duration or pupil diameter, two indicators often associated with cognitive load. From this, it can be inferred that the higher number of fixations in the SDH group is more likely to reflect increased visual attention to the subtitles rather than greater processing difficulty. Thus, in effect, SDH subtitles seem to encourage viewers to engage more actively with the visual content without necessarily increasing cognitive demands.

8.2.4 RQ4

> RQ4: Do DHH participants who viewed SDH subtitles achieve better comprehension performance than those who viewed ST subtitles?

The results of the comprehension test, as shown in Section 7.4.1, provide a clear answer to this question: DHH participants who viewed SDH subtitles did perform better than those who viewed ST subtitles in terms of comprehension. This demonstrates that SDH subtitles can help DHH viewers better understand AV content.

Further analysis of the comprehension test showed that SDH subtitles were particularly beneficial for questions related to speaker identification and tone of delivery. This points to the value of including speaker identification and paralinguistic information in supporting viewers' understanding of AV material. In addition, it was found that not all DHH viewers in SDH group benefited from SDH subtitles to the same extent. As discussed in Section 7.6.2, this variation may be due to individual differences among DHH viewers, including reading speed, familiarity with subtitle conventions, literacy skills, or other relevant factors.

8.2.5 RQ5

> RQ5: What differences can be observed in the overall viewing experience of DHH participants, in terms of engagement and enjoyment, when watching videos with SDH subtitles versus ST subtitles?

Based on the results presented in Section 7.4.2, no statistically significant differences were found between SDH and ST groups in terms of overall engagement and enjoyment. That said, when we looked at engagement and enjoyment separately, we did find some differences between the two groups.

In terms of engagement, participants in the SDH group showed higher levels of engagement, especially in relation to immersion in and emotional connection with the story in the video. But at the same time, some participants in the same group mentioned that they found it harder to stay focused compared to the ST group. As discussed in Section 7.6.3, this is likely due to the fact that SDH subtitles included additional information that drew attention away

from the visual content and made it harder for viewers to stay focused on the visuals themselves.

As for enjoyment, again, participants in the SDH group gave more positive feedback than those in the ST group. This could be seen in their greater satisfaction with the viewing experience and their stronger willingness to rewatch or recommend the video. Despite these positive responses, a few participants had reservations about the role of SDH subtitles, which seems to reflect individual differences in subtitle preferences. As already indicated in Section 7.6.3, these mixed responses from participants suggest that standardised SDH design is not sufficient to meet the diverse needs of the DHH community. This reinforces the need for more flexible and tailored ways of designing and presenting SDH. It also resonates with Neves' recent call for greater personalisation in media accessibility, as accessibility services and products should adapt to "a diversity of users while respecting individual traits, social standings and cultural perceptions" (Neves, 2025: 12).

8.3 Contributions

This thesis takes a step towards exploring SDH from both the production and reception sides. First, it looks at SDH subtitlers through interviews to understand their real-world decision making and offer some practical guidance for SDH production. It then turns to SDH users in China, where SDH is relatively new, in the hope that the findings can feed back into and help guide future SDH design in China. Therefore, the contributions of this thesis are considered in two parts: the first relates to SDH production, and the second to SDH reception.

SDH production

In the field of SDH research, most attention so far has been given to how SDH is received by DHH audiences, while much less is known about how it is created or produced. Even where SDH production is considered, previous research is often concerned with guidelines, standards, or broader policy dimensions (e.g., O'Sullivan, 2016; Toledo, 2018; Gürkan, 2021). This study instead focuses on practice by directly speaking with professional SDH subtitlers to understand how decisions are made in real working environments.

Through interviews with these professional practitioners, this study reveals that SDH production is affected by multiple factors, including workflows, the use of SDH guidelines, technical and practical constraints, audience considerations, subjective interpretation, and limitations in subtitlers' own knowledge. Some of these factors are worth highlighting here.

First, the four SDH subtitlers from China (P4-P7) confirmed that there are still no national SDH guidelines in China. This is consistent with what we saw in the literature review (Section 3.6.2), and, given this, there is a clear need to develop such guidelines so that SDH production in China can be better regulated. Second, some interviewees mentioned that current technologies can help in the initial stage through speech recognition to generate transcripts, but they still need humans to check for accuracy, and decisions about what to include and how to describe sounds still rely on human judgement. Choices about which sounds to include are mainly based on two considerations: whether the sound is visible on screen and whether it matters to the plot. When it comes to music description, it was found that conveying its atmosphere or mood is more preferable, not only to improve audience understanding but also because subtitlers themselves may not have much knowledge in specific domains such as music. Finally, in the final stage, peer review turned out to be a useful way to support consistency and quality, and it can also help minimise subjective interpretation to some extent.

In addition to these empirical contributions, this thesis also makes a theoretical contribution by introducing relevance theory into SDH research. As mentioned in Section 4.3.4, its application in AVT has been very limited, and even more so in SDH, but this study demonstrates its value in understanding SDH subtitlers' decision making. By drawing on relevance theory, the study argues that decisions about what information should be conveyed, omitted, or clarified can be better understood in terms of relevance. In practice, this means that SDH subtitlers need to weigh the likely cognitive effort required of viewers against the potential value of the information being presented.

In sum, by speaking with SDH subtitlers and giving space to their voices, this study opens up dialogue between research and professional practice and contributes to closer connections between the two in future work.

SDH reception

To date, there have been quite a few SDH reception studies, but not many have been concerned with the Chinese DHH community, and this thesis manages to do so by focusing on this under-researched group. By combining objective eye-tracking data with subjective measures, the reception study provides empirical evidence on how Chinese DHH viewers experience SDH subtitles compared to ST subtitles. In terms of research methods, it extends existing SDH research in China by moving beyond earlier studies (e.g., Yang et al., 2022; Yang et al., 2025; Wu et al., 2026) that mainly relied on self-reported data collected through surveys, questionnaires, or interviews.

As far as the findings are concerned, it was found that SDH subtitles work better than ST subtitles for Chinese DHH viewers. They support better comprehension than ST subtitles, were similarly effective in terms of engagement and enjoyment. These findings can be situated within the broader context of limited SDH provision in China and offer important implications for future SDH reception studies in China. Finally, given the individual differences in participants' subtitle preferences, this study calls for more flexible and tailored SDH design that can accommodate diverse viewer preferences, rather than relying on a one-size-fits-all approach.

8.4 Limitations

If the contributions outlined above are one side of the coin, the limitations discussed below are the other. Among the possible limitations of this thesis, three are worth acknowledging in particular.

Sample limitations

Phase I, the interview study, is limited by its relatively small sample size, as it included only nine interviewees from three organisations. While these organisations are relevant to the aims of this research, the views of these participants cannot be taken to reflect the industry as a whole, and the findings are better understood as exploratory. The sample also did not include professional practitioners from major platforms such as Netflix. Their practices are likely to differ from what we saw in the three organisations involved in this study, and this absence is worth noting as a limitation.

Phase II, the reception study, also has some limitations in relation to the sample. All DHH participants were recruited from one special education school, and they were of a similar age and shared broadly comparable educational backgrounds, which limited the diversity of the sample. This may limit the extent to which the findings can be generalised to the broader Chinese DHH population. Additionally, grouping was based on clinically measured degrees of hearing loss, but the pre-task questionnaire indicated that some participants used hearing aids or cochlear implants. Because these devices can improve auditory access to varying extents, the reported degree of hearing loss did not necessarily reflect participants' functional hearing conditions. This limitation should be considered when interpreting the findings.

Methodological limitations

A further limitation relates to the research method and design in both phases. In Phase I, the study was exclusively based on interviews, in which nine subtitlers were asked to reflect on their work after the fact. However, their descriptions probably do not always correspond to what actually happens during the subtitling process. If interview data could be combined with records of the subtitling process in real time, this might give us a more direct view of how SDH subtitles are produced in practice.

In Phase II, the eye-tracking experiment took place in a controlled setting, not in natural viewing conditions, and this may have had an impact on participants' viewing behaviour

and, to some extent, the results obtained. A further point relates to the eye-tracking device itself. The Tobii Pro Nano used in this study had a 60 Hz sampling rate, which limited the granularity of the gaze data and made it less reliable for detecting high-velocity saccadic activity. Another point worth noting is that, during the experiment, group allocation was adjusted continuously. After each participant completed the eye-tracking task, their tracking ratio was checked immediately, and any data falling below 80% were excluded. The remaining participants were then reassigned where necessary. Because this was an ongoing process, it was not feasible to keep the two groups exactly equal in size. In the end, there were 17 participants in the SDH group and 15 in the ST group. Even though the difference is modest, the groups were not perfectly balanced.

The limited scope of AV content

Finally, the scope of the AV material used in Phase II also needs to be acknowledged as a limitation. As Szarkowska et al. (2024a: 2) address, the broader applicability of research findings can be affected by a range of factors, of which AV genre is only one. As far as this thesis is concerned, only an animated film was used in the reception study. However, different types of AV content may pose different processing demands for viewers, and may also differ in the way non-verbal audio information is distributed. In the case of the animated film used here, sound effects were more prominent than music, and this could have influenced the results. Therefore, caution is needed in extending the findings to other types of AV content.

8.5 Future work

In view of the limitations discussed above, several directions for future research seem worth pursuing in order to further explore the issues addressed in this thesis and, more broadly, SDH production and reception.

First, given the small number of subtitlers interviewed in Phase I, future research could involve a larger number of participants and cover a broader range of sectors. Since the present study did not include practitioners from major streaming platforms, it would be

worthwhile to include them in order to see whether the findings presented here also hold in different production contexts.

A second direction for future research would be to explore in more detail how SDH subtitlers work in practice. As noted in the limitations (Section 8.4), the Phase I interview study was entirely based on subtitlers' own reflections on their work, and these reflections cannot fully reflect what happens during the subtitling process itself. This means that our understanding of their role and work is likely to be partial. Given this, future research could make use of methods that allow the subtitling process to be examined more directly, such as eye-tracking, mouse tracking, and keystroke logging, which have been used in previous studies to explore subtitling practices (e.g., Orrego-Carmona et al., 2018).

Another area that would benefit from further research is SDH reception. In the present study, the sampling rate of the eye-tracking device used in Phase II was relatively low, which meant that some eye-tracking measures may not have been sufficiently reliable and therefore were not included in the analysis. Thus, future research could consider using a higher sampling rate in order to improve the precision of the data. In addition, given the issue of data loss in eye-tracking experiments, it is also necessary to recruit a larger number of participants. Pilot studies could also help refine the experimental design and make the study more reliable overall.

A further direction for future research would be to expand the scope of SDH reception studies. Future research could include various AV materials in order to further assess the robustness of the findings presented here. One possible approach would be to compare two or more videos from clearly different genres, especially where they differ in the type and density of non-verbal audio information. It would be valuable to know how different AV materials influence viewers' reception of SDH. Future studies could also include a broader range of participants, so as to explore how SDH subtitles are experienced by a wider audience within the Chinese DHH population beyond students alone.

Finally, even with the present study, we still know relatively little about SDH for Chinese DHH audiences, and there is still much to explore. This calls for further research into a broader range of subtitle variables among Chinese DHH viewers, since parameters such as subtitling speed and subtitle layout have not yet been systematically tested in this context. It is expected that such empirical research will contribute to advancing SDH production in China, where there is still a long way to go.

References

A

ACMA (Australian Communications and Media Authority) (2013) *The Broadcasting Services (Television Captioning) Standard 2013.* Available at: https://www.legislation.gov.au/Details/F2016C00225

ACMA (Australian Communications and Media Authority) (2023) *The Broadcasting Services (Television Captioning) Standard 2023.* Available at: https://www.legislation.gov.au/F2023L01210/latest/text

ACMA (Australian Communications and Media Authority) (2025) *Captioning rules on TV.* Available at: https://www.acma.gov.au/captioning-rules-tv

AENOR (Asociación Española de Normalización y Certificación) (2012) *Norma UNE 153010. Subtitulado para personas sordas y personas con discapacidad auditiva. Subtitulado a través del teletexto.* Madrid: AENOR.

Agrawal, C. and Peiris, R. (2021) 'I See What You're Saying: A Literature Review of Eye Tracking Research in Communication of Deaf or Hard of Hearing Users', in *the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, USA, pp. 1–13.

Agulló, B. and Matamala, A. (2019) 'Subtitling for the deaf and hard-of-hearing in immersive environments: Results from a focus group', *Journal of Specialised Translation*, 32, pp. 217–235.

Aickin, M. and Gensler, H. (1996) 'Adjusting for Multiple Testing When Reporting Research Results: The Bonferroni vs Holm Methods', *American Journal of Public Health*, 86(5), pp. 726–728.

Aleksandrowicz, P. (2019) 'The reading comprehension skill of d/deaf and hard-of-hearing Poles and its importance for media accessibility: A pilot study', *Journal of Audiovisual Translation*, 2(1), pp. 26–52.

Aleksandrowicz, P. (2020) 'Can subtitles for the deaf and hard-of-hearing convey the emotions of film music? A reception study', *Perspectives*, 28(1), pp. 58–72.

Atkinson, R. and Flint, J. (2001) 'Accessing hidden and hard-to-reach populations: Snowball research strategies', *Social Research Update*, 33.

Arnáiz-Uzquiza, V. (2010) 'The need of a deep analysis of data', in Matamala, A. and Orero, P. (eds.) *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang, pp. 163–174.

Arnáiz-Uzquiza, V. (2012) *Subtitling for the deaf and the hard-of-hearing: Some parameters and their evaluation*. Doctoral thesis. Universitat Autònoma de Barcelona.

Arnáiz-Uzquiza, V. (2015) 'Eye tracking Tests in Spain', in Romero-Fresco, P. (ed.) *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern, Switzerland: Peter Lang, pp. 263–287.

B

Bañós, R. (2017) 'Audiovisual translation', in Bedijs, K. and Maaß, C. (eds.) *Manual of Romance Languages in the Media*. Berlin, Germany: De Gruyter, pp. 471–488.

Bartoll, E. (2008) *Paramètres per a una taxonomia de la subtitulació*. Doctoral thesis. Universitat Pompeu Fabra.

Bartoll, E. and Martínez Tejerina, A. (2010) 'The positioning of subtitles for the deaf and hard of hearing', in Matamala, A. and Orero, P. (eds.) *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang, pp. 70–80.

BBC (British Broadcasting Corporation) (2026) *Subtitle Guidelines* (Version 1.2.5). Available at: https://www.bbc.co.uk/accessibility/forproducts/guides/subtitles/

Benecke, B. (2004) 'Audio-description', *Meta*, 49(1), pp. 78–80.

Boeije, H. (2010) *Analysis in qualitative research*. Los Angeles, CA: Sage.

Böhm, A. (2004) 'Theoretical coding: Text analysis in grounded theory', in Flick, U., Kardoff, E. von and Steinke, I. (eds.) *A Companion to Qualitative Research*. London: SAGE, pp. 270–275.

Bolaños García-Escribano, A. (2025) *Practices, Education and Technology in Audiovisual Translation*. Abingdon: Routledge.

Birks, M. and Mills, J. (2015) *Grounded theory: A practical guide*. 2nd edn. Los Angeles, CA: SAGE.

Black, S. (2022) ‘Could integrated subtitles benefit young viewers? Children’s reception of standard and integrated subtitles: A mixed methods approach using eye tracking’, *Perspectives*, 30(3), pp. 1–17.

Bogucki, Ł. (2004) ‘The constraint of relevance in subtitling’, *The Journal of Specialised Translation*, 1, pp. 69–85.

Bogucki, Ł. (2020) *A Relevance-Theoretic Approach to Decision-Making in Subtitling*. Cham: Palgrave Macmillan.

Braun, S. (2018) ‘The importance of being relevant? A cognitive-pragmatic framework for conceptualising audiovisual translation’, in Gambier, Y. and Ramos Pinto, S. (eds.) *Audiovisual Translation: Theoretical and Methodological Challenges*. Amsterdam: John Benjamins, pp. 121–132.

Bryant, A. and Charmaz, K. (2007) *The SAGE Handbook of Grounded Theory*. Los Angeles: Sage.

Burnham, D., Leigh, G., Noble, W., Jones, C., Grebennikov, M., Leonid, T. and Varley, A. (2008) ‘Parameters in Television Captioning for Deaf and Hard-of-Hearing Adults: Effects of Caption Rate Versus Text Reduction on Comprehension’, *Journal of Deaf Studies and Deaf Education*, 13(3), pp. 391–404.

C

CAB (Canadian Association of Broadcasters) (2012) *Closed captioning standards and protocol for Canadian English language television programming services*. Available at: https://assets.corusent.com/wp-content/uploads/2021/10/Closed_Captioning_Standards_Protocol.pdf

Canadian Department of Justice (1991) *The Broadcasting Act*. Available at: https://laws-lois.justice.gc.ca/PDF/B-9.01.pdf

Carroll, M. and Mueller, F. (2022) ‘The evolution of audiovisual translation in Australia’, in Wakabayashi, J. and O’Hagan, M. (eds.) *Translating and Interpreting in Australia and New Zealand.* Abingdon: Routledge, pp. 171–190.

Casas-Tost, H. and Rovira-Esteva, S. (2018) ‘A quest for effective and inclusive design of Chinese characters in subtitling’, *International Journal of Asian Language Processing*, 28(1), pp. 31–47.

CDPF (China Disabled Persons' Federation) (2021) *The Number of People with Disabilities in China by the End of 2010*. Available at: https://www.cdpf.org.cn/zwgk/zccx/cjrgk/15e9ac67d7124f3fb4a23b7e2ac739aa.htm

Cepak, R. (2021) *Subtitling for the Deaf and the Hard of Hearing in Spain: A study on Media Accessibility, Accessibility Studies and the Standard UNE 153010:2012*. Doctoral thesis. Universitat Pompeu Fabra.

Cepak, R. and Corrius, M. (2023) 'Moving forward to improve SDH: the needs of Spanish audiences', Perspectives, 31(4), pp. 745–761.

Charmaz, K. (1990) 'Discovering "chronic illness": Using grounded theory', *Social Science and Medicine*, 30(11), pp. 1161–1172.

Charmaz, K. (2006) *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. 1st edn. London: Sage Publications.

Charmaz, K. (2014) *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. 2nd edn. London: Sage Publications.

Chartrand, M.S. (2008) 'What Helen Keller knew; what popular thought overlooks', *Hearing Review*, 15(5), pp. 18–28.

Chaume, F. (2013) 'The turn of audiovisual translation. New audiences and new technologies', *Translation Spaces*, 2, pp. 105–123.

Chaume, F. (2018) 'Is audiovisual translation putting the concept of translation up against the ropes?', *Journal of Specialised Translation,* 30, pp. 84–104.

Chaume, F. (2020) 'Dubbing', in Bogucki, Ł. and Deckert, M. (eds.) *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. Cham: Springer International Publishing, pp. 103–132.

Chen, Y. and Wang, W. (2019) 'Semiotic analysis of viewers' reception of Chinese subtitles: A relevance theory perspective', *The Journal of Specialised Translation*, 32, pp. 194–216.

Chen, Z. (2024) *Subtitling as Multimodal Representation: A Corpus-based Experimental Approach to Text-image Relations*. Doctoral thesis. Hong Kong Polytechnic University.

Chiaro, D. (2009) 'Issues in audiovisual translation', in Munday, J. (ed.) *The Routledge Companion to Translation Studies*. London: Routledge, pp. 141–165.

Chiaro, D. (2013) 'Audiovisual translation', in Chapelle, C.A. (ed.) *The Encyclopedia of Applied Linguistics*. Oxford: Blackwell Publishing Ltd, pp. 1–5.

CNMC (Comisión Nacional de los Mercados y la Competencia) (2026) *La accesibilidad audiovisual avanza en España impulsada por nuevas cuotas legales y el uso de la inteligencia artificial*. Available at: https://www.cnmc.es/prensa/inf-accesibilidad-audiovisual-20260116

Congreso de los Diputados (2010) *Ley General de la Comunicación Audiovisual*. Available at: https://www.boe.es/buscar/pdf/2010/BOE-A-2010-5292-consolidado.pdf

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Coimisiún na Meán (2024) *Access Rules for Television Broadcasting Services*. Available at: https://www.cnam.ie/industry-and-professionals/codes-legislation/codes-and-rules/access-rules/

Cooke, M. (2008) *A History of Film Music*. Cambridge: Cambridge University Press.

Corbin, J. and Strauss, A. (2015) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 4th edn. Los Angeles, CA: Sage.

CRTC (Canadian Radio-television and Telecommunications Commission) (2007) *Broadcasting public notice CRTC 2007-54*. Available at: https://crtc.gc.ca/eng/archive/2007/pb2007-54.htm

CRTC (Canadian Radio-television and Telecommunications Commission) (2011) *Quality standards for French-language closed captioning*. Available at: https://crtc.gc.ca/eng/archive/2011/2011-741.pdf

CRTC (Canadian Radio-television and Telecommunications Commission) (2012) *Quality standards for English-language closed captioning*. Available at: https://crtc.gc.ca/eng/archive/2012/2012-362.pdf

CRTC (Canadian Radio-television and Telecommunications Commission) (2015) *Broadcasting regulatory policy CRTC 2015-104*. Available at: https://crtc.gc.ca/eng/archive/2015/2015-104.htm

Creswell, J.W. and Plano Clark, V.L. (2018) *Designing and Conducting Mixed Methods Research*. 3rd edn. Los Angeles: Sage.

D

Datta, P., Jakubowicz, P., Vogler, C. and Kushalnagar, R. (2020) 'Readability of punctuation in automatic subtitles', in Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M. and Peňáz, P. (eds.) *Computers Helping People with Special Needs.* Cham: Springer, pp. 195–201.

Dawadi, S., Shrestha, S. and Giri, R.A. (2021) 'Mixed-methods research: A discussion on its types, challenges, and criticisms', *Journal of Practical Studies in Education*, 2(2), pp. 25–36.

DCMP (Described and Captioned Media Program) (2019) *Captioning Key*. Available at: https://dcmp.org/learn/captioningkey

Delabastita, D. (1989) 'Translation and mass-communication: Film and TV translation as evidence of cultural dynamics', *Babel*, 35(4), pp. 193–218.

De Seriis, L. (2006) 'Il Servizio Sottotitoli RAI: televideo per i non udenti', *InTRAlinea*, Special Issue: Respeaking.

Díaz-Cintas, J. and Remael, A. (2007) *Audiovisual Translation: Subtitling*. London: Routledge.

Díaz-Cintas, J. (2008) 'Teaching and learning to subtitle in an academic environment', in Díaz-Cintas, J. (ed.) *The Didactics of Audiovisual Translation*. Amsterdam: John Benjamins, pp. 89–105.

Díaz-Cintas, J. and Anderman, G. (2009) 'Introduction', in Díaz-Cintas, J. and Anderman, G. (eds.) *Audiovisual Translation: Language Transfer on Screen*. Basingstoke: Palgrave Macmillan, pp. 1–17.

Díaz-Cintas, J. and Orero, P. (2010) 'Dubbing and voice-over', in Gambier, Y. and van Doorslaer, L. (eds.) *Handbook of Translation Studies*. Amsterdam: John Benjamins Publishing Company, pp. 441–445.

Díaz-Cintas, J. (2013) 'Subtitling: theory, practice and research', in Millán, C. and Bartrina, F. (eds.) *The Routledge Handbook of Translation Studies*. London: Routledge, pp. 273–289.

Díaz-Cintas, J. (2020) 'The name and nature of subtitling', in Bogucki, Ł. and Deckert, M. (eds.) *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. London, UK: Palgrave Macmillan, pp. 149–171.

Díaz-Cintas, J. and Szarkowska, A. (2020) 'Introduction: Experimental research in audiovisual translation–cognition, reception, production', *The Journal of Specialised Translation*, 33, pp. 3–16.

Díaz-Cintas, J. and Remael, A. (2021) *Subtitling: Concepts and Practices*. London and New York: Routledge, Taylor & Francis Group.

Dobie, R.A. and Van Hemel, S. (2005) *Hearing Loss: Determining Eligibility for Social Security Benefits*. Washington, DC: National Academies Press.

Doherty, S. (2012) *Investigating the effects of controlled language on the reading and comprehension of machine translated texts: a mixed-methods approach*. Doctoral thesis. Dublin City University.

Doherty, S. and Kruger, J.-L. (2018) 'The development of eye tracking in empirical research on subtitling and captioning', in Dwyer, T., Perkins, C., Redmond, S. and Sita, J. (eds.) *Seeing into Screens: Eye Tracking and the Moving Image*. London: Bloomsbury, pp. 46–64.

Duraj, K. and Szarkowska, A. (2025) 'Beyond traditional subtitles: How emojis and non-standard typography in subtitles boost engagement on TikTok', *Journal of Audiovisual Translation*, 8(1), pp. 1–28.

D'Ydewalle, G., Praet, C., Verfaillie, K. and van Rensbergen, J. (1991) 'Watching subtitled television: Automatic reading behavior', *Communication Research*, 18, pp. 650–666.

E

Eichstaedt, K.E., Kovatch, K. and Maroof, D.A. (2013) 'A less conservative method to adjust for familywise error rate in neuropsychological research: The Holm's sequential Bonferroni procedure', *NeuroRehabilitation*, 32(3), pp. 693–696.

Emerson, R.W. (2018) 'MANOVA (Multivariate Analysis of Variance): An expanded form of the ANOVA (Analysis of Variance)', *Journal of Visual Impairment & Blindness*, 112(1), pp. 125–126.

ERC (Entidade Reguladora para a Comunicação Social) (2025) *Plano Plurianual de Acessibilidade das Emissões Televisivas (2026–2029)*. Available at: https://www.erc.pt/document.php?id=YTlkYTk5ZGYtY2NhYi00NjQ3LTllYzgtNzE0Nzg4MDY3NzVl

Eugeni, C. (2015) ‘Long questionnaire in Italy’, in Romero-Fresco, P. (ed.) *The Reception of Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang, pp. 75–94.

F

FCC (Federal Communications Commission) (2010). *Twenty-First Century Communications and Video Accessibility Act.* Available at: https://www.fcc.gov/cvaa

FCC (Federal Communications Commission) (2025) *Media Bureau announces compliance date for closed captioning display settings requirements*. Available at: https://www.fcc.gov/consumer-governmental-affairs/media-bureau-announces-compliance-date-closed-captioning-display-settings-requirements

Field, A. (2024) *Discovering Statistics Using IBM SPSS Statistics*. 6th edn. London: Sage Publications.

Finstad, K. (2010) ‘Response interpolation and scale sensitivity: evidence against 5-point scales’, *Journal of Usability Studies*, 5(3), pp. 104–110.

Fox, W. (2018) *Can integrated titles improve the viewing experience?: Investigating the impact of subtitling on the reception and enjoyment of film using eye tracking and questionnaire data*. Cambridge: Language Science Press.

Fresno, N. and Sepielak, K. (2022) ‘Subtitling speed in media accessibility research: some methodological considerations’, *Perspectives*, 30(3), pp. 415–431.

G

Gambier, Y. and Gottlieb, H. (2001) ‘Multimedia, multilingualia: Multiple challenges’, in Gambier, Y. and Gottlieb, H. (eds.) *(Multi)Media Translation – Concepts, Practices, and Research.* Amsterdam and Philadelphia: John Benjamins, pp. viii–xx.

Gambier, Y. (2003) ‘Screen transadaptation: perception and reception’, *The Translator*, 9(2), pp. 171–189.

Gambier, Y. (2013) ‘The position of audiovisual translation studies’, in Millán, C. and Bartrina, F. (eds.) *The Routledge Handbook of Translation Studies*. London: Routledge, pp. 44–59.

Gambier, Y. (2026) ‘From a microscopic to a telescopic view of audiovisual translation’, in Yang, F., Wang, Y. and Li, D. (eds.) *Audiovisual Translation Theory, Practice and Technology*. Singapore: Springer, pp. 9–32.

Galletta, A. and Cross, W.E. (2013) *Mastering the Semi-Structured Interview and Beyond: From Research Design to Analysis and Publication*. New York: New York University Press.

Gernsbacher, M.A. (2015) ‘Video captions benefit everyone’, *Policy Insights from the Behavioral and Brain Sciences*, 2(1), pp. 195–202.

Glaser, B.G. (1978) *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory*. Mill Valley, CA: Sociology Press.

Glaser, B.G. and Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing.

Gottlieb, H. (2001) ‘Subtitling’, in Baker, M. (ed.) *Routledge Encyclopedia of Translation Studies*. London: Routledge, pp. 245–248.

Greco, G.M. and Jankowska, A. (2020) ‘Media accessibility within and beyond audiovisual translation’, in Bogucki, Ł. and Deckert, M. (eds.) *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. Cham, Switzerland: Palgrave Macmillan, pp. 57–82.

Greene, J.C. (2007) *Mixed Methods in Social Inquiry*. San Francisco, CA: Jossey-Bass.

Gürkan, A. (2019) *Subtitling for the Deaf and the Hard-of-Hearing: A Reception Study in the Turkish Context*. Doctoral thesis. University College London.

Gürkan, A. (2021) ‘Comparison of subtitling for the deaf and hard-of-hearing guidelines implemented across countries’, *Karamanoğlu Mehmetbey Üniversitesi Uluslararası Filoloji ve Çeviribilim Dergisi*, 3(2), pp. 249–278.

Guerberof-Arenas, A. and Toral, A. (2020) 'The impact of post-editing and machine translation on creativity and reading experience', *Translation Spaces*, 9(2), pp. 255–282.

Guerberof-Arenas, A., Moorkens, J. and Orrego-Carmona, D. (2024) '"A Spanish version of EastEnders": a reception study of a telenovela subtitled using MT', *The Journal of Specialised Translation*, 41, pp. 230–254.

Gutt, E.-A. (1991) *Translation and Relevance: Cognition and Context*. Oxford: Blackwell.

H

Hao, Y. and Li, P. (2020) 'Employment legal framework for persons with disabilities in China: Effectiveness and reasons', *International Journal of Environmental Research and Public Health*, 17(14), pp. 1–20.

Higgs, C. (2006) 'Subtitles for the deaf and the hard of hearing on TV: legislation and practice in the UK', *inTRAlinea*, Special Issue: Respeaking.

Holton, J. (2010) 'The coding process and its challenges', *The Grounded Theory Review*, 9, pp. 21–40.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. and van de Weijer, J. (2011) *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford: Oxford University Press.

Horiuchi, T. (2022) 'Efforts to disseminate broadcasting for persons with visual and hearing impairments', *Disability Information Resources (DINF)*. Available at: https://www.dinf.ne.jp/d/4/250.html

Howitt, D. and Cramer, D. (2014) *Introduction to Statistics in Psychology*. 5th edn. Harlow: Pearson Education Limited.

Hyönä, J., Tommola, J. and Alaja, A.-M. (1995) 'Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks', *Quarterly Journal of Experimental Psychology*, 48A(3), pp. 598–612.

I

ISO/IEC. (2018) *ISO/IEC 20071-23:2018 Information technology — User interface component accessibility — Part 23: Visual presentation of audio information, including captions and subtitles.* Geneva: International Organization for Standardization and International Electrotechnical Commission. Available at: https://www.iso.org/standard/70722.html

ITC (Independent Television Commission) (1999) *ITC Guidelines on Standards for Subtitling*. Available at: http://sign-dialog.de/wp-content/richtlinien_199902_england.pdf

Ivarsson, J. and Carroll, M. (1998) *Subtitling*. Simrishamn: TransEdit HB.

Ivarsson, J. (2002) 'Subtitling through the ages: A technical history of subtitles in Europe', *Language International*, 14(2), pp. 6–10.

J

Jakobson, R. (1959) 'On linguistic aspects of translation', in Brower, R.A. (ed.) *On Translation*. Cambridge, MA: Harvard University Press, pp. 232–239.

Jankowska, A. (2019) 'Accessibility mainstreaming and beyond: senior citizens as secondary users of audio subtitles in cinemas', *International Journal of Language, Translation and Intercultural Communication*, 8, pp. 28–47.

Jelinek Lewis, M. S. and Jackson, D. W. (2001) 'Television literacy: Comprehension of program content using closed captions for the deaf', *Journal of Deaf Studies and Deaf Education,* 6(1), pp. 43–53.

Jiang, Q. and Doherty, S. (2024) 'An eye-tracking study on the cognitive processing, reception, and perception of swearword translation in films', *Translation Spaces*, 13, pp. 293–329.

Johnson, R.B. and Onwuegbuzie, A.J. (2004) 'Mixed methods research: A research paradigm whose time has come', *Educational Researcher*, 33(7), pp. 14–26.

Johnson, R.B., Onwuegbuzie, A.J. and Turner, L.A. (2007) 'Toward a definition of mixed methods research', *Journal of Mixed Methods Research*, 1(2), pp. 112–133.

K

Kalinak, K. (2015) *Sound: Dialogue, Music, and Effects*. New Brunswick, NJ: Rutgers University Press.

Kemper, E.A., Stringfield, S. and Teddlie, C. (2003) 'Mixed methods sampling strategies in social science research', in Tashakkori, A. and Teddlie, C. (eds.) *Handbook of Mixed Methods in Social and Behavioral Research*. London/Thousand Oaks, CA: Sage Publications, pp. 273–296.

KMCC (Korea Media and Communications Commission) (2009) *Supporting Policy for the Broadcasting Underprivileged People*. Available at: https://www.kmcc.go.kr/user.do;jsessionid=0jVLjpWamDfdN9Dx2PC6yDtbq-CD811GY_yzcJZ9.servlet-aihgcldhome10?mode=view&page=E04010000&dc=E04010000&boardId=1058&cp=40&boardSeq=15654

KMCC (Korea Media and Communications Commission) (2026) *Guidelines for Providing Broadcasting Services for People with Disabilities*. Available at: https://www.kmcc.go.kr/user.do;jsessionid=1tcwF7gOu0ufpYWkDvaf1x73PRcComkGSByE3NNe.servlet-aihgcldhome20?mode=view&page=A02030700&dc=&boardId=1099&cp=1&nop=10&boardSeq=67928

Khoshsaligheh, M., Ahadi Farkoush, N. and Eriss Azadeh, F. (2024) 'Rendition of non-verbal acoustic elements for film audiences with hearing impairments', *Babel*, 70(6), pp. 759–782.

Kim, M. and Kang, H. (2016) 'Implement closed captioning systems for the deaf', *Journal of Korea Game Society*, 16(1), pp. 103–110.

Kovačič, I. (1994) 'Relevance as a factor in subtitling reductions', in Dollerup, C. and Lindegaard, A. (eds.) *Teaching Translation and Interpreting 2*. Amsterdam and Philadelphia: John Benjamins, pp. 245–251.

KRRiT (Krajowa Rada Radiofonii i Telewizji) (2026) *Dostępność audycji telewizyjnych dla osób z niepełnosprawnościami wzroku lub słuchu – raport KRRiT*. Available at: https://www.gov.pl/web/krrit/dostepnosc-audycji-telewizyjnych-dla-osob-z-niepelnosprawnosciami-wzroku-lub-sluchu--raport-krrit

Kruger, J.-L., Hefer, E. and Gordon, M. (2013) 'Measuring the impact of subtitles on cognitive load: Eye tracking and dynamic audiovisual texts', in *Proceedings of the 2013 Conference on Eye Tracking South Africa, ETSA 2013*, pp. 62–66.

Kruger, J.-L., Hefer, E. and Matthew, G. (2014) 'Attention distribution and cognitive load in a subtitled academic lecture: L1 vs. L2', *Journal of Eye Movement Research*, 7(5), pp. 1–15.

Kruger, J.-L., Szarkowska, A. and Krejtz, I. (2015) 'Subtitles on the moving image: an overview of eye tracking studies', *Refractory: A Journal of Entertainment Media*, 25, pp. 1–14.

Kruger, J.-L. and Doherty, S. (2016) 'Measuring cognitive load in the presence of educational video: Towards a multimodal methodology', *Australasian Journal of Educational Technology*, 32(6), pp. 19–31.

Kruger, J.-L., Soto-Sanfiel, M.T., Doherty, S. and Ibrahim, R. (2016) 'Towards a cognitive audiovisual translatology: Subtitles and embodied cognition', in Muñoz Martín, R. (ed.) *Reembedding Translation Process Research*. London, UK: John Benjamins, pp. 171–194.

Kruger, J.-L., Doherty, S., Fox, W. and De Lissa, P. (2018) 'Multimodal measurement of cognitive load during subtitle processing: same-language subtitles for foreign-language viewers', in Lacruz, I. and Jääskeläinen, R. (eds.) *Innovation and Expansion in Translation Process Research*. Amsterdam: John Benjamins Publishing Company, pp. 267–294.

Kruger, J.-L. and Liao, S. (2022) 'Establishing a theoretical framework for AVT research: the importance of cognitive models', *Translation Spaces*, 11(1), pp. 12–37.

Kuckartz, U. and Rädiker, S. (2019) *Analyzing Qualitative Data with MAXQDA: Text, Audio, and Video*. Cham: Springer International Publishing.

Kuder, G.F. and Richardson, M.W. (1937) 'The theory of the estimation of test reliability', *Psychometrika*, 2(3), pp. 151–160.

Kuo, S.-Y. (2014) *Quality in Subtitling. Theory and Professional Reality*. Doctoral thesis. Imperial College London.

Kuo, S.-Y. (2018) 'Subtitling quality beyond the linguistic dimension', in Shei, C. and Gao, Z.-M. (eds.) *The Routledge Handbook of Chinese Translation*. Abingdon: Routledge, pp. 415–431.

Kvale, S. and Brinkmann, S. (2015) *InterViews: Learning the Craft of Qualitative Research Interviewing*. 3rd edn. Thousand Oaks, CA: Sage Publications.

L

Law, M. and Kruger, H. (2018) 'How editors read: an eye-tracking study of the effects of professional editorial experience and task instruction on reading behaviour', in Lacruz, I. and Jääskeläinen, R. (eds.) *Innovation and Expansion in Translation Process Research*. Amsterdam: John Benjamins Publishing Company, pp. 241–265.

Larsen, P. (2005) *Film Music*. London: Reaktion.

Leeson, L. and Vermeerbergen, M. (2010) 'Sign language interpreting', in van Doorslaer, L. and Gambier, Y. (eds.) *Handbook of Translation Studies*. Amsterdam and Philadelphia: John Benjamins, pp. 324–328.

Leung, H.C.D. (2018) *Audio description of audiovisual programmes for the visually impaired in Hong Kong*. Doctoral thesis. University College London.

Li, D. (2013) *Hear&See: Audio-visual media for all*. Hangzhou: Zhejiang University Press.

Li, D. and Looms, P.O. (2015) ‘Television accessibility in China’, in Biswas, P., Duarte, C., Langdon, P. and Almeida, L. (eds.) *A Multimodal End-2-End Approach to Accessible Computing*. London: Springer, pp. 261–273.

Li, D. and Xiong, M. (2017) ‘After accessibility: Research and reflections on barrier-free communication in new media’, *Zhejiang Academic Journal*, 6, pp. 199–206.

Liao, S., Kruger, J.-L. and Doherty, S. (2020) ‘The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension’, *The Journal of Specialised Translation*, 33, pp. 70–98.

Liao, S., Yu, L., Kruger, J.-L. and Reichle, E.D. (2022) ‘The impact of audio on the reading of intralingual versus interlingual subtitles: Evidence from eye movements’, *Applied Psycholinguistics*, 43(1), pp. 237–269.

Liu, Y., Casas-Tost, H. and Rovira-Esteva, S. (2022) ‘Made in China versus Made in Spain: A corpus-based study comparing AD in Chinese and Spanish’, *The Journal of Specialised Translation*, 38, pp. 180–206.

Liu, Y., Casas-Tost, H. and Rovira-Esteva, S. (2023) ‘AD in China at the crossroads: a set of guidelines for localising AD scripts from Spanish into Chinese’, *Perspectives*, 32(5), pp. 815–832.

Liu, Q., Jia, H., Pang, C.L. and Jankowska, A. (2025) ‘Audio describers in Mainland China’, *Między Oryginałem a Przekładem*, 3(69), pp. 53–77.

Longari, G. (2022) *It’s not a hearing thing. It’s a cultural thing: A hybrid SDH proposal for post-fansubbing audiences*. Master’s thesis. Ca’ Foscari University of Venice.

Lorenzo, L. (2010) ‘Subtitling for deaf and hard of hearing children in Spain: A case study’, in Matamala, A. and Orero, P. (eds.) *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang, pp. 115–138.

Luyken, G.-M., Herbst, T., Langham-Brown, J., Reid, H. and Spinhof, H. (1991) *Overcoming Language Barriers in Television: Dubbing and Subtitling for the European Audience.* Manchester: European Institute for the Media.

M

Matamala, A. and Orero, P. (eds) (2010) *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Wien: Peter Lang.

Martínez-Sirés, P. (2016) 'Making the inaccessible accessible: a panorama of audiovisual translation, accessibility and audio description in Japan', *Asiédémica: Revista Universitaria de Estudios sobre Asia Oriental*, 7, pp. 24–50.

Marvasti, A.B. (2004) *Qualitative Research in Sociology: An Introduction*. London: SAGE.

Mellinger, C.D. and Hanson, T.A. (2017) *Quantitative Research Methods in Translation and Interpreting Studies*. London: Routledge.

MIC (Ministry of Internal Affairs and Communications) (2007) 視聴覚障害者向け放送普及行政の指針. Available at: https://www.soumu.go.jp/main_content/000126735.pdf

MIC (Ministry of Internal Affairs and Communications) (2012) 視聴覚障害者向け放送普及行政の指針. Available at: https://www8.cao.go.jp/shougai/suishin/seisaku_iinkai/k_37/pdf/s2-1.pdf

MIC (Ministry of Internal Affairs and Communications) (2018) 放送分野における情報アクセシビリティに関する指針. Available at: https://www.soumu.go.jp/main_content/000531258.pdf

MIC (Ministry of Internal Affairs and Communications) (2023) 放送分野における情報アクセシビリティに関する指針. Available at: https://www.soumu.go.jp/main_content/000906717.pdf

MIC (Ministry of Internal Affairs and Communications) (2024) 令和 5 年度の字幕放送等の実績. Available at: https://www.soumu.go.jp/menu_news/s-news/01ryutsu09_02000346.html

Miller, C. (2007) 'Access symbols for use with video content and information and communications technology devices', in Díaz-Cintas, J., Orero, P.C. and Remael, A. (eds.) *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam: Rodopi, pp. 53–69.

Miquel-Iriarte, M. (2017) *The reception of subtitling for the deaf and hard of hearing: viewers' hearing and communication profile & subtitling speed of exposure*. Doctoral thesis. Universitat Autònoma de Barcelona.

Mishra, P., Pandey, C.M., Singh, U., Gupta, A., Sahu, C. and Keshri, A. (2019) 'Descriptive statistics and normality tests for statistical data', *Annals of Cardiac Anaesthesia*, 22(1), pp. 67–72.

Mliczak, R. (2015) 'Signing and subtitling on Polish television: A case of (in)accessibility', in Bañós-Piñero, R. and Díaz-Cintas, J. (eds.) *Audiovisual Translation in a Global Context: Mapping an Ever-changing Landscape*. Basingstoke: Palgrave Macmillan, pp. 203–224.

Mliczak, R. (2019) *A Diachronic Study of Subtitling for the Deaf and the Hard-of-Hearing in Poland*. Doctoral thesis. University College London.

Moorkens, J. (2012) *Measuring Consistency in Translation Memories: A Mixed-Methods Case Study*. Doctoral thesis. Dublin City University.

Morettini, A. (2012) 'Profiling deaf and hard-of-hearing users of subtitles for the deaf and hard-of-hearing in Italy: A questionnaire-based study', *MonTI. Monografías de Traducción e Interpretación*, 4, pp. 321–348.

Morettini, A. (2014) 'Legislation on audiovisual and media accessibility in Italy and beyond', *inTRAlinea*, Special Issue: Across Screens Across Boundaries.

Muller, T. (2015) *A User-centred Study of the Norms for Subtitling for the Deaf and Hard-of-Hearing on French Television.* Doctoral thesis. Universitat Autònoma de Barcelona.

Muñoz, C. (2017) 'The role of age and proficiency in subtitle reading. An eye-tracking study', *System*, 67, pp. 77–86.

N

Napier, J. (2009) *Signed Language Interpreter Education and Training: A World Survey*. Washington, DC: Gallaudet University Press.

Netflix (2025) *Chinese (Simplified) Timed Text Style Guide*. Available at: https://partnerhelp.netflixstudios.com/hc/en-us/articles/215986007-Chinese-Simplified-Timed-Text-Style-Guide

Netflix (2025) *Chinese (Traditional) Timed Text Style Guide*. Available at: https://partnerhelp.netflixstudios.com/hc/en-us/articles/215994807-Chinese-Traditional-Timed-Text-Style-Guide

Neves, J. (2005) *Audiovisual translation: Subtitling for the Deaf and the Hard-of-Hearing*. Doctoral thesis. Roehampton University.

Neves, J. (2007a) 'Of pride and prejudice: The divide between subtitling and sign language interpreting on television', in Leeson, L. and Turner, G. (eds.) *The Sign Language Translator & Interpreter*, 1(2), pp. 251–274.

Neves, J. (2007b) *Guia de Legendagem para Surdos: Vozes Que Se Vêem*. Leiria: Instituto Politécnico de Leiria and Universidade de Aveiro.

Neves, J. (2008a) 'Ten fallacies about subtitling for the d/Deaf and the hard of hearing', *Journal of Specialised Translation*, 10, pp. 128–143.

Neves, J. (2008b) 'Training in subtitling for the d/Deaf and the hard-of-hearing', in Díaz-Cintas, J. (ed.) *The Didactics of Audiovisual Translation*. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 171–189.

Neves, J. (2009) ‘Interlingual subtitling for the deaf and hard-of-hearing’, in Díaz-Cintas, J. and Anderman, G. (eds.) *Audiovisual Translation: Language Transfer on Screen*. Basingstoke: Palgrave Macmillan, pp. 151–169.

Neves, J. (2012) ‘Audiovisual translation for accessible media in Portugal’, in Pym, A. and Assis Rosa, A. (eds.) *New Directions in Translation Studies* [Special issue], *Anglo-Saxónica*, 3(3), pp. 365–383.

Neves, J. (2019) ‘Subtitling for deaf and hard of hearing audiences – moving forward’, in Pérez-González, L. (ed.) *The Routledge Handbook of Audiovisual Translation*. London and New York: Routledge, pp. 82–95.

Neves, J. (2025) ‘Communication accessibility: 10 answers to 9 questions’, *Journal of Audiovisual Translation, 8*(2), pp. 1–21.

Neuhäuser, M. and Ruxton, G.D. (2025) ‘The choice between Pearson's χ2 test and Fisher's exact test for 2 × 2 tables’, *Pharmaceutical Statistics*, 24, e70012.

Nikolić, K. (2018) ‘Reception studies in audiovisual translation: interlingual subtitling’, in Di Giovanni, E. and Gambier, Y. (eds.) *Reception Studies and Audiovisual Translation.* Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 179–197.

NRTA (National Radio and Television Administration) (2013) *Specification of Closed Caption System for Digital Television*. Available at: https://www.nrta.gov.cn/module/download/downfile.jsp?spm=chekydwncf.0.0.1.gBQjZF&classid=0&filename=af5c85d528284b3d95c20ac53dea7984.pdf

NRTA (National Radio and Television Administration) (2020) *Guiding Opinions on Strengthening the Construction of the Radio and Television Public Service System*. Available at: http://www.nrta.gov.cn/art/2020/1/6/art_2062_53546.html

O

O'Brien, S. (2006) 'Eye tracking and translation memory matches', *Perspectives: Studies in Translatology*, 14(3), pp. 185–205.

Ofcom (Office of Communications) (2021) *Ofcom's Guidelines on the Provision of Television Access Services*. Available at: https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-2-6-weeks/178126-review-of-access-services-code/associated-documents/provision-of-tv-access-services-guideline.pdf?v=369712

Ofcom (Office of Communications) (2023) *Television channels required to provide access services in 2023*. Available at: https://www.ofcom.org.uk/siteassets/resources/documents/tv-radio-and-on-demand/broadcast-guidance/list-of-channels-2023.pdf?v=328179

Ofcom (Office of Communications) (2025) *Television and on-demand programme services: Access services report – January to December 2024*. Available at: https://www.ofcom.org.uk/tv-radio-and-on-demand/accessibility/television-and-on-demand-programme-services-access-services-report--january-to-december-2024

Ogea Pozo, M. del M. and Talaván, N. (2024) 'Teaching languages for social and cooperation purposes: Using didactic media accessibility in foreign language education', in Bolaños García-Escribano, A. and Oakhill, M. (eds.) *Inclusion, Diversity and Innovation in Translation Education*. London, UK: UCL Press, pp. 69–91.

O'Hagan, M. and Sasamoto, R. (2016) 'Crazy Japanese subtitle? Shedding light on the impact of impact captions', in Hansen-Schirra, S. and Grucza, S. (eds.) *Eye Tracking and Applied Linguistics. Translation and Multilingual Natural Language Processing*. Berlin, Germany: Language Science Press, pp. 31–57.

Oomori, K., Shitara, A., Minagawa, T., Sarcar, S. and Ochiai, Y. (2020) 'A preliminary study on understanding voice-only online meetings using emoji-based captioning for deaf or

hard of hearing users', in *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*, Virtual Event, Greece.

Orero, P., Doherty, S., Kruger, J.-L., Matamala, A., Pedersen, J., Perego, E., Romero-Fresco, P., Rovira-Esteva, S., Soler-Vilageliu, O. and Szarkowska, A. (2018) 'Conducting experimental research in audiovisual translation (AVT): A position paper', *The Journal of Specialised Translation*, 30, pp. 105–126.

Orero, P., Montagud, M., Mata, J., Torres, E. and Matamala, A. (2020) 'Audio subtitles or spoken subtitles/captions: an ecological media accessibility service', *Professional Communication and Translation Studies.* (ISSN 2065-099X).

Orrego-Carmona, D. (2013) 'Avance de la traducción audiovisual: desde los inicios hasta la era digital', *Mutatis Mutandis: Revista Latinoamericana de Traducción*, 6(2), pp. 297–320.

Orrego-Carmona, D. (2015) *The reception of (non)professional subtitling*. Doctoral thesis. Universitat Rovira i Virgili.

Orrego-Carmona, D. (2016) 'A reception study on non-professional subtitling: Do audiences notice any difference?', *Across Languages and Cultures*, 17(2), pp. 163–181.

Orrego-Carmona, D., Dutka, Ł. and Szarkowska, A. (2018) 'Using translation process research to explore the creation of subtitles: an eye tracking study comparing professional and trainee subtitlers', *The Journal of Specialised Translation*, 30, pp. 150–180.

O'Sullivan, C. (2016) 'Imagined spectators: The importance of policy for audiovisual translation research', *Target. International Journal for Translation Studies*, 28(2), pp. 261–275.

P

Paivio, A. (1979) *Imagery and Verbal Processes.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Pallant, J. (2005) *SPSS Survival Manual: A Step by Step Guide to Data Analysing Using SPSS for Windows (Version 12)*. Maidenhead: Open University Press.

Parliament of Australia (1992) *The Broadcasting Services Act 1992*. Available at: https://www.legislation.gov.au/C2004A04401/2020-12-18/text

Parliament of the United Kingdom (1990) *The Broadcasting Act 1990*. Available at: https://www.legislation.gov.uk/ukpga/1990/42/contents

Parliament of the United Kingdom (2003) *The Communications Act 2003*. Available at: https://www.legislation.gov.uk/ukpga/2003/21/section/303

Pedersen, J. (2017) 'The FAR model: Assessing quality in interlingual subtitling', *The Journal of Specialised Translation*, 28, pp. 210–229.

People's Government of Chongqing (2021) '*What is Disability Identification Card?*' Available at: https://www.cqyz.gov.cn/zwgk_229/cjzxwd/dshxqgwh_9124/202112/t20211228_10241225.html

Perego, E., Del Missier, F., Porta, M. and Mosconi, M. (2010) 'The cognitive effectiveness of subtitle processing', *Media Psychology*, 13(3), pp. 243–272.

Perego, E., Del Missier, F. and Bottiroli, S. (2015) 'Dubbing versus subtitling in young and older adults: Cognitive and evaluative aspects', *Perspectives: Studies in Translatology*, 23(1), pp. 1–21.

Perego, E., Orrego-Carmona, D. and Bottiroli, S. (2016) 'An empirical take on the dubbing vs. subtitling debate: An eye movement study', *Lingue e Linguaggi*, 19, pp. 255–274.

Perego, E., Del Missier, F. and Stragà, M. (2018) 'Dubbing vs. subtitling: Complexity matters', *Target*, 30(1), pp. 137–157.

Perego, E. and Pacinotti, R. (2020) 'Audiovisual translation through the ages', in Bogucki, Ł. and Deckert, M. (eds.) *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan, pp. 33–56.

Pereira, A.M. (2005) 'El subtitulado para sordos: estado de la cuestión en España', *Quaderns. Revista de Traducción*, 12, pp. 161–172.

Pereira, A.M. and Lorenzo, L. (2005) ‘Evaluamos la norma UNE 153010: Subtitulado para personas sordas y personas con discapacidad auditiva. Subtitulado a través de teletexto’, *Puentes*, 6, pp. 21–26.

Pereira, A. (2010) ‘Proposal of criteria for producing subtitles for deaf and hard of hearing adult audiences’, in Matamala, A. and Orero, P. (eds.) *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang, pp. 88–102.

Pintado Gutiérrez, L. (2025) *The challenges of audio description in Ireland: Bridging education, users and industry (ADESI)*. Project report. Dublin City University and Irish Research Council.

Preston, C.C. and Colman, A.M. (2000) ‘Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences’, *Acta Psychologica*, 104(1), pp. 1–15.

Q

R

Rayner, K. (1998) ‘Eye movements in reading and information processing: 20 years of research’, *Psychological Bulletin*, 124(3), pp. 372–422.

Remael, A. (2007) ‘Sampling subtitling for the deaf and the hard-of-hearing in Europe’, in Díaz-Cintas, J., Remael, A. and Orero, P. (eds.) *Media for All*. Amsterdam: Rodopi, pp. 23–52.

Remael, A. and Vercauteren, G. (2010) ‘The translation of recorded audio description from English into Dutch’, *Perspectives*, 18(3), pp. 155–171.

Robert, I.S. and Remael, A. (2016) ‘Quality control in the subtitling industry: An exploratory survey’, *Meta*, 61(3), pp. 578–605.

Robson, G.Y. (2004) *The Closed Captioning Handbook*. Burlington: Focal Press.

Romero-Fresco, P. (2009) 'More haste less speed: Edited versus verbatim respeoken subtitles', *Vial-Vigo International Journal of Applied Linguistics*, 6(6), pp. 109–133.

Romero-Fresco, P. (2010) 'D'Artagnan and the Seven Musketeers: SUBSORDIG travels to Europe', in Matamala, A. and Orero, P. (eds.) *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang, pp. 175–189.

Romero-Fresco, P. (ed.) (2015) *The Reception of Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang.

Romero-Fresco, P. and Chaume, F. (2022) 'Creativity in audiovisual translation and media accessibility', *Journal of Specialised Translation*, 38, pp. 75–101.

RTÉ (Raidió Teilifís Éireann) (1987) *RTÉ Launches Teletext Service 1987*. Available at: https://www.rte.ie/archives/2017/0613/882414-rte-launch-aertel/

RTÉ (Raidió Teilifís Éireann) (1995) *Closed Captions 1995*. Available at: https://www.rte.ie/archives/2025/1121/1545195-live-subtitles/

RTÉ (Raidió Teilifís Éireann) (2025) *RTÉ marks 50 years of providing access to news for the Deaf and Hard of Hearing*. Available at: https://about.rte.ie/2025/01/06/rte-marks-50-years-of-providing-access-to-news-for-the-deaf-and-hard-of-hearing/

S

Sala Robert, È. (2016) *Creative Subtitles: Subtitling for ALL*. Doctoral thesis. Universitat Pompeu Fabra.

Sala Robert, È. (2022) 'Rendering tone and mood in creative subtitles for deaf and hard-of-hearing: A proposal', *Między Oryginałem a Przekładem*, 1(55), pp. 73–91.

Saldanha, G. and O'Brien, S. (2014) *Research Methodologies in Translation Studies*. London: Routledge.

Saldaña, J. (2016) *The Coding Manual for Qualitative Researchers*. 3rd edn. Los Angeles, CA: SAGE.

Sanders, M.S. and McCormick, E.J. (1993) *Human Factors in Engineering and Design*. New York: McGraw-Hill.

Sasamoto, R., Doherty, S. and O'Hagan, M. (2021) 'The 'hookability' of multimodal impact captions: A mixed-methods exploratory study of Japanese TV viewers', *Translation, Cognition and Behavior*, 4(2), pp. 253–280.

Sasamoto, R. (2024) *Relevance and Text-on-Screen in Audiovisual Translation: The Pragmatics of Creative Subtitling*. London: Routledge.

Sheth, T. (2017) 'Non-verbal communication: A significant aspect of proficient occupation', *OSR Journal of Humanities and Social Science*, 22(11), pp. 69–72.

Sigauke, I. and Swansi, K. (2020) '"Classic" grounded theory: An extraction of Saldana's coding heuristics', *International Forum*, 23(1), pp. 5–23.

Sikkema, E.C. (2020) *Teleops for language learning: Japanese language learners' perceptions of authentic Japanese variety shows and implications for their use in the classroom*. Doctoral thesis. Dublin City University.

Silva, B.B., Orrego-Carmona, D. and Szarkowska, A. (2022) 'Using linear mixed models to analyze data from eye-tracking research on subtitling', *Translation Spaces*, 11(1), pp. 60–88.

Silver, C. and Lewins, A. (2014) *Using Software in Qualitative Research: A Step-by-Step Guide*. 2nd edn. Thousand Oaks, CA: SAGE.

Snyder, J. (2022) 'Audio description in the United States', in Taylor, C. and Perego, E. (eds.) *The Routledge Handbook of Audio Description*. London: Routledge, pp. 531–543.

Song, J.-G. (2003) 'A study on policy measures to expand broadcasting accessibility for people with disabilities (장애인 방송접근권 확대를 위한 정책방안 연구)', *Broadcasting Research (방송연구)*, pp. 147–178.

Sperber, D. and Wilson, D. ([1986] 1995) *Relevance: Communication and Cognition*. Oxford: Blackwell.

Standing Committee of the National People's Congress (2008) *Law of the People's Republic of China on Protection of Disabled Persons*. Available at: http://www.npc.gov.cn/zgrdw/englishnpc/Law/2011-02/14/content_1620414.htm

Standing Committee of the National People's Congress (2023) *Barrier-free Environment Development Law of the People's Republic of China*. Available at: https://www.gov.cn/yaowen/liebiao/202306/content_6888910.htm

Star, S.L. (2007) '"Living grounded theory: Cognitive and emotional forms of pragmatism"', in Bryant, A. and Charmaz, K. (eds.) *Handbook of Grounded Theory*. London: Sage, pp. 75–93.

Strauss, A.L. (1987) *Qualitative Analysis for Social Scientists*. Cambridge: Cambridge University Press.

Strauss, A.L. and Corbin, J. (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks, CA: Sage.

Szarkowska, A. (2010) 'Accessibility to the media by hearing impaired audiences in Poland: Problems, paradoxes, perspectives', in Díaz-Cintas, J., Matamala, A. and Neves, J. (eds.) *New Insights into Audiovisual Translation and Media Accessibility*. Amsterdam and New York: Rodopi, pp. 139–158.

Szarkowska, A., Krejtz, I., Kłyszejko, Z. and Wieczorek, A. (2011) 'Verbatim, standard, or edited? Reading patterns of different captioning styles among deaf, hard of hearing, and hearing viewers', *American Annals of the Deaf*, 156(4), pp. 363–378.

Szarkowska, A. (2013) ‘Towards interlingual subtitling for the deaf and the hard of hearing’, *Perspectives*, 21(1), pp. 68–81.

Szarkowska, A., Żbikowska, J. and Krejtz, I. (2013) ‘Subtitling for the deaf and hard of hearing in multilingual films’, *International Journal of Multilingualism*, 10(3), pp. 292–312.

Szarkowska, A., Pietrulewicz, J. and Jankowska, A. (2015) ‘Long questionnaire in Poland’, in Romero-Fresco, P. (ed.) *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern, Switzerland: Peter Lang, pp. 45–74.

Szarkowska, A., Krejtz, I., Pilipczuk, O., Dutka, Ł. and Kruger, J.-L. (2016) ‘The effects of text editing and subtitle presentation rate on the comprehension and reading patterns of interlingual and intralingual subtitles among deaf, hard of hearing and hearing viewers’, *Across Languages and Cultures*, 17(2), pp. 183–204.

Szarkowska, A. and Gerber-Morón, O. (2018a) ‘Viewers can keep up with fast subtitles: Evidence from eye movements’, *PLOS ONE*, 13(6), e0199331.

Szarkowska, A. and Gerber-Morón, O. (2018b) ‘Line breaks in subtitling: An eye tracking study on viewer preferences’, *Journal of Eye Movement Research*, 11(3), pp. 1–22.

Szarkowska, A. and Gerber-Morón, O. (2019) ‘Two or three lines: a mixed-methods study on subtitle processing and preferences’, *Perspectives*, 27(1), pp. 144–164.

Szarkowska, A. (2020) ‘Subtitling for the deaf and the hard of hearing’, in Bogucki, Ł. and Deckert, M. (eds.) *Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan, pp. 249–268.

Szarkowska, A., Silva, B. and Orrego-Carmona, D. (2021) ‘Effects of subtitle speed on proportional reading time: Re-analysing subtitle reading data with mixed effects models’, *Translation, Cognition & Behavior*, 4(2), pp. 305–330.

Szarkowska, A., Ragni, V., Orrego-Carmona, D., Black, S., Skriba, S., Kruger, J.-L., Krejtz, K. and Silva, B. (2024a) 'The impact of video and subtitle speed on subtitle reading: An eye-tracking replication study', *Journal of Audiovisual Translation*, 7(1), pp. 1–23.

Szarkowska, A., Ragni, V., Skriba, S., Black, S., Orrego-Carmona, D. and Kruger, J.-L. (2024b) 'Watching subtitled videos with the sound off affects viewers' comprehension, cognitive load, immersion, enjoyment, and gaze patterns: A mixed-methods eye-tracking study', *PLOS ONE*, 19(10), e0306251.

Szarkowska, A. and Jankowska, A. (2025) *Introducing Audiovisual Translation*. Abingdon: Taylor & Francis Group.

Szkriba, S. (2022) 'Accessible cinema for older adults: can an app for blind viewers benefit the sighted population?', *Journal of Audiovisual Translation*, 5(2), pp. 134–151.

T

Talaván, N. (2020) 'The didactic value of AVT in foreign language education', in Bogucki, Ł. and Deckert, M. (eds.) *Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan, pp. 567–591.

Tamayo, A. (2015) *Estudio descriptivo y experimental de la subtitulación en TV para niños sordos. Una propuesta alternativa*. Doctoral thesis. Universitat Jaume I.

Tamayo, A. (2016) 'Formal aspects in SDH for children in Spanish television: a descriptive study', *Estudios de Traducción*, 6, pp. 109–128.

Tamayo, A. (2017a) 'Signifying codes of audiovisual products: Implications in subtitling for the d/Deaf and the hard of hearing', *inTRAlinea*, Special Issue: Building Bridges between Film Studies and Translation Studies.

Tamayo, A. (2017b) 'Subtitling for d/Deaf and hard of hearing children: Past, present and future', *CoMe* II (1), pp. 74–90.

Tamayo, A., Lozano, J. de los R. and Martí Ferriol, J. L. (2021) 'The reception of subtitling for the deaf and hard-of-hearing in Spanish TV news programs', *Comunicación y Sociedad*, 18, e7875.

Taylor, C. (2020) 'Multimodality and intersemiotic translation', in Bogucki, Ł. and Deckert, M. (eds.) *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. London, UK: Palgrave Macmillan, pp. 83–99.

Teixeira, C.S.C. and O'Brien, S. (2018) 'Overcoming methodological challenges of eye tracking in the translation workplace', in Walker, C. and Federici, F.M. (eds.) *Eye Tracking and Multidisciplinary Studies on Translation*. Amsterdam: John Benjamins Publishing Company, pp. 33-54.

The Central People's Government of the People's Republic of China (2006) *Practical Evaluation Criteria for Persons with Disabilities in China*. Available at: https://www.gov.cn/ztzl/gacjr/content_459939.htm

The Oireachtas of Ireland (2009) *Broadcasting Act 2009*. Available at: https://www.irishstatutebook.ie/eli/2009/act/18/enacted/en/print.html

Titford, C. (1982) 'Sub-titling: constrained translation', *Lebende Sprachen*, 27(3), pp. 113–116.

Tobii Technology (2021) *Tobii Pro Nano User Manual*. v. 1.181. Available at: http://andrewd.ces.clemson.edu/courses/cpsc881/manuals/Tobii/Tobii-Pro-Lab-User-Manual.pdf

Tobii Technology (2023) *Tobii Pro Nano User Manual*. v. 2.5. Available at: https://device.report/m/d3fcc26184cc939b9f8185fd95985bf1dd3c7350aaecb4a4d565c6e412b7105c.pdf

Toledo, G. (2018) 'Subtitles for the deaf and hard-of-hearing: Comparing legislation and official orientation for SDH in Brazil and in other countries', *Transletters. International Journal of Translation and Interpreting*, 1, pp. 143–166.

Tor-Carroggio, I. (2020) *Audio Description in China: Past, Present and Future*. Doctoral thesis. Universitat Autònoma de Barcelona.

Tsaousi, A. (2015). 'Making sound accessible: The labelling of sound effects in subtitling for the deaf and hard-of-hearing', *Hermēneus*, 17, pp. 233–252.

U

United States Congress (1990) *Television Decoder Circuitry Act of 1990.* Available at: https://www.congress.gov/bill/101st-congress/house-bill/4267

United States Congress (1996) *Telecommunications Act of 1996*. Available at: https://www.congress.gov/104/plaws/publ104/PLAW-104publ104.pdf

Utray, F., Pereira, A.M. and Orero, P. (2009) 'The present and future of audio description and subtitling for the deaf and hard of hearing in Spain', *Meta*, 54(2), pp. 248–263.

V

Vanderplank, R. (1988) 'The value of teletext subtitles in language learning', *ELT Journal*, 42(4), pp. 272–281.

Vanderplank, R. (2016) *Captioned Media in Foreign Language Learning and Teaching: Subtitles for the Deaf and Hard-of-Hearing as Tools for Language Learning*. London: Palgrave Macmillan.

Verbit (2022) 'SDH subtitles vs closed captions'. Available at: https://verbit.ai/captioning/the-differences-between-subtitles-closed-captions-and-sdh/

W

Wang, J., O'Brien, S. and Sasamoto, R. (2025) 'Investigating the impact of personalized headnotes for English-speaking audiences in viewing Chinese comedies', *Perspectives*, pp. 1–20.

Wilson, D. and Sperber, D. (2004) 'Relevance theory', in Horn, L.R. and Ward, G. (eds.) *The Handbook of Pragmatics*. Oxford: Blackwell, pp. 607–632.

Wu, Y., Guerberof Arenas, A. and Lowie, W. (2026) 'Subtitles and accessibility: Insights from China's deaf community', *Journal of Audiovisual Translation*, 9(1), pp. 1–24.

X

Xiao, X. and Li, F. (2013) 'Sign language interpreting on Chinese TV: a survey on user perspectives', *Perspectives*, 21(1), pp. 100–116.

Xiao, X., Chen, X. and Palmer, J.L. (2015) 'Chinese Deaf viewers' comprehension of sign language interpreting on television: An experimental study', *Interpreting*, 17(1), pp. 91–117.

Y

Yang, H., Shen, W., Wang, Y. and Liu, B. (2022) 'Difficulties and solutions of the accessibility of subtitles for the deaf and hard of hearing in China', *Journal of Lanzhou Jiaotong University*, pp. 156–168.

Yang, H., Shen, W., Ma, X., Liu, B. and Wang, Y. (2025) 'Research on availability and reception of intralingual subtitles by the deaf and hard of hearing in China', *SAGE Open*, 15(3).

Yang, R. and Wang, W.S.Y. (2018) 'Categorical perception of Chinese characters by simplified and traditional Chinese readers', *Reading and Writing*, 31, pp. 1133–1154.

Ye, Z. and Jie, H. (2023) 'Enhancing cinema experience for visually impaired persons in China: A case study on Shenzhen', *Journal of Chinese Film Studies*, 3(2), pp. 221–240.

Yoon, J.-O. and Kim, M. (2011) ‘The effects of captions on deaf students’ content comprehension, cognitive load, and motivation in online learning’, *American Annals of the Deaf*, 156(3), pp. 283–289.

Youens, W. (2022) ‘Changing media accessibility landscapes in New Zealand’, in Wakabayashi, J. and O’Hagan, M. (eds.) *Translating and Interpreting in Australia and New Zealand*. Abingdon: Routledge, pp. 191–207.

Z

Zabalbeascoa, P. (2008) ‘The nature of the audiovisual text and its parameters’, in Díaz-Cintas, J. (ed.) *The Didactics of Audiovisual Translation*. Amsterdam: John Benjamins, pp. 21–37.

Zabalbeascoa, P. (2025) *Audiovisual Translation*. 1st edn. London: Routledge.

Zárate, S. (2014) *Subtitling for Deaf Children: Granting Accessibility to Audiovisual Programmes in an Educational Way*. Doctoral thesis. University College London.

Zárate, S. and Eliahoo, J. (2014) ‘Word recognition and content comprehension of subtitles for television by deaf children’, *The Journal of Specialised Translation*, 21, pp. 133–152.

Zárate, S. (2021) *Captioning and Subtitling for d/Deaf and Hard of Hearing Audiences*. London: UCL Press.

Zdenek, S. (2015) *Reading Sounds: Closed-Captioned Media and Popular Culture*. Chicago: University of Chicago Press.

Zhou, R., Hentschel, J. and Kumar, N. (2017) ‘Goodbye text, hello emoji: Mobile communication on WeChat in China’, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 748–759.

Appendices

Appendices for Phase I

Appendix A: Plain Language Statement

Appendix A1: Plain Language Statement (English)

Introduction

This study, titled "Subtitling Practices for the Deaf and Hard of Hearing in Ireland", is part of the doctoral research carried out by Rongyu Wang (rongyu.wang7@mail.dcu.ie) at the School of Applied Language and Intercultural Studies, Dublin City University. This research is supervised by Dr. Ryoko Sasamoto and Dr. Joss Moorkens, and funded by Dublin City University's Faculty of Humanities and Social Sciences PhD Scholarship Programme.

What is this research about?

This research sets out to understand how professional subtitlers make decisions in producing subtitling for the deaf and hard of hearing (SDH). It aims to discuss current SDH practices, challenges, and considerations in this field.

Why is this research being conducted?

By understanding the experiences and perspectives of professional subtitlers, the findings will help researchers have a better understanding of SDH practices and may contribute to the future development of more inclusive SDH standards and practices.

Why have I been invited to take part?

You have been invited to participate because you are a professional subtitler with experience in SDH production.

What will happen if I decide to take part in this research study?

If you agree to participate, you will be invited to take part in an individual interview. During the interview, you will be asked to share your experiences, views, and professional practices relating to SDH. With your permission, the interview will be audio-recorded for transcription and analysis. The interview is expected to last no longer than one hour.

How will my data be used?

The data collected, including interview recordings and transcripts, will be used solely for academic research purposes. The findings will contribute to a PhD thesis and may be included in future academic publications and presentations.

How will my privacy be protected?

Only the researcher and the supervisors will have access to the data collected for this study. All data will be stored securely and anonymised in order to protect your identity. Any findings that are published will be reported in a way that does not reveal your identity. This study complies with the General Data Protection Regulation (GDPR), and all personal data will be processed in accordance with GDPR requirements.

What are the benefits of taking part in this research study?

By participating in this research, you will potentially contribute to improving accessibility and SDH practices. In appreciation of your time and contribution, you will receive a One4all gift card or an equivalent voucher.

What are the risks of taking part in this study?

The risks involved in participating in this study are extremely low. If you feel uncomfortable discussing certain aspects of your work or experiences during the interview, you are encouraged to share only what you feel comfortable discussing.

How will I find out what happens with this project?

Regarding the findings of this research, updates may be obtained by contacting the researcher via email at rongyu.wang7@mail.dcu.ie. This may include a PhD thesis or publications related to the study.

Can I change my mind at any stage and withdraw from this study?

Yes, participation in this study is entirely voluntary. You may withdraw at any time without providing a reason.

----- Contact details for further information -----

If you have concerns about this study and wish to contact an independent person, please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9.

Tel: 01-7008000; E-mail: rec@dcu.ie

Appendix A2: Plain Language Statement (Chinese)

介绍

本研究题为《爱尔兰背景下的听障字幕制作实践》，是由都柏林城市大学应用语言与跨文化研究学院王容宇（rongyu.wang7@mail.dcu.ie）开展的博士研究项目的一部分。本研究由 Ryoko Sasamoto 博士和 Joss Moorkens 博士共同指导，并获得都柏林城市大学人文与社会科学学院博士奖学金资助。

本研究的内容是什么？

本研究旨在了解专业字幕制作人员在制作听障字幕过程中如何进行决策，并探讨当前听障字幕实践中的做法、挑战与相关内容。

为什么要开展这项研究？

通过了解专业字幕制作人员的经验与观点，本研究希望可以更深入理解听障字幕实践，并为未来制定更加包容的听障字幕标准与实践提供参考。

为什么邀请我参加？

您受邀参加本研究，是因为您具有听障字幕制作相关经验，并从事专业字幕工作。

如果我同意参加，需要做什么？

如果您同意参加本研究，您将受邀参与一次一对一访谈。在访谈过程中，您可以分享与听障字幕相关的制作经验、看法及专业见解。在获得您的同意后，访谈将被录音，用于后续的文字转录和分析。预计访谈时长不超过一小时。

数据将如何使用？

收集的数据（包括访谈录音和转录文本）将仅用于学术研究目的。研究结果将用于撰写博士论文，并可能被纳入未来的学术发表和学术会议报告中。

我的隐私如何得到保护？

只有研究者及其导师有权访问本研究收集的数据。所有数据都将被安全存储并进行匿名化处理，以保护您的个人信息。任何公开发表的研究结果均不会透露您的身份。本研究遵守《通用数据保护条例》，所有个人数据将按照该条例的要求进行处理。

参与本研究有什么好处？

通过参与本研究，您将有机会为加强无障碍服务及听障字幕实践作出贡献。为感谢您的时间与参与，您将获得一张 One4all 的礼品卡或代金券。

参加本研究是否存在风险？

参与本研究的风险极低。如果您在访谈过程中对讨论某些工作内容或个人经历感到不适，您可以选择只分享自己愿意讨论的内容。

我如何了解研究结果？

关于本研究的成果，您可以通过电子邮件（rongyu.wang7@mail.dcu.ie）联系研究人员获取。相关成果可能包括博士论文或与本研究相关的学术文章。

我可以在任何时候退出研究吗？

可以。参加本研究完全出于自愿。您可在任何时候选择退出，且无需说明任何理由。

----- **其他联系方式** -----

如您对本研究有任何疑问，并希望联系独立于研究团队的人员，请联系：

都柏林城市大学研究伦理委员会秘书处

地址： The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9.

电话：01-7008000

邮箱：rec@dcu.ie

Appendix B: Informed Consent Form

Appendix B1: Informed Consent Form (English)

Research title: Subtitling Practices for the Deaf and Hard-of-Hearing in Ireland

Principal investigators: Rongyu Wang, Dr. Ryoko Sasamoto, Dr. Joss Moorkens

Email: rongyu.wang7@mail.dcu.ie

Institution: School of Applied Language and Intercultural Studies, Dublin City University

- I have read and understood the Plain Language Statement about this research project.
- I voluntarily agree to participate in the study and consent to being interviewed by the researchers, during which I will answer questions regarding subtitling practices.
- I consent to the audio recording and transcription of my interview.
- I understand that all information I provide for this study will be securely stored. I am also aware that the confidentiality of the information provided is subject to GDPR regulations.
- I understand that I have the right to request access to DORAS and publications related to this study in order to obtain the research findings.
- I am aware of the potential risks and benefits of this research study.
- I understand that I can withdraw from this study at any time, and I am not obligated to provide a reason for withdrawal.

By signing this form, I confirm that I have read and understood the information provided. My questions and concerns have been addressed by the researchers, and I have received a copy of this consent form. Therefore, I consent to take part in this research project.

Participant's Name: _____

Participant's Signature: _____

Witness: _____

Date: _____

Appendix B2: Informed Consent Form (Chinese)

研究项目： 爱尔兰背景下的听障字幕制作实践

研究人员：王容宇、Ryoko Sasamoto 博士、Joss Moorkens 博士

电子邮件：rongyu.wang7@mail.dcu.ie

研究机构：都柏林城市大学应用语言与跨文化研究学院

- 我已阅读并理解本研究《研究说明书》中所提供的信息。
- 我可以就本研究提出问题并进行讨论。
- 我所有的问题均已得到满意的答复。
- 我自愿同意参与本研究，并知道我将参加由研究人员开展的访谈。
- 我同意对访谈进行录音，并同意对访谈内容进行转录。
- 我明白本研究所收集的所有信息将被安全保存，并进行匿名化处理。我也清楚，所提供的数据按照《通用数据保护条例》进行处理。
- 我知悉我有权查阅与本研究相关的博士论文及公开发表成果，以获取研究结果。
- 我知悉本研究可能存在的潜在风险与利益。
- 我知道我可以在任何时候退出本研究，且无需说明理由。

签署此表格即表示我确认已阅读并知悉所提供的信息。我的疑问和顾虑已得到研究人员的解答，并且我已收到本同意书的副本。因此，我同意参与本研究项目。

参与者姓名： _____

参与者签名： _____

见证人： _____

日期： _____

Appendix C: Ethical Approval



Faculty of Humanities & Social Sciences
DUBLIN CITY UNIVERSITY
23 April 2024

CONFIRMATION OF RESEARCH ETHICS APPROVAL FOR A PROJECT

Application Reference:	**DCU-FHSS-2024-037**
Project Title:	**Subtitling Practices for the Deaf and Hard-of-Hearing in Ireland**
Project contact(s):	**rongyu.wang7@mail.dcu.ie**

Let this letter certify that the proposed project identified above has been reviewed by the *Humanities & Social Sciences Faculty Research Ethics Committee* (F-REC) and has been approved as a low-risk project. The application was found to that comply with university requirements and best practices for research ethics, and with GDPR guidelines and requirements where personal data is processed in the project.

A copy of the application, including appended documents related to participant consent, is archived under the reference above. Queries about this project's approval may be directed to the F-REC Chair.

Dr Dónal Mulligan
donal.mulligan@dcu.ie

Chair, Faculty Research Ethics Committee
Faculty of Humanities & Social Sciences
Dublin City University

Dámh na nDaonnachtaí agus na nEolaíochtaí Sóisialta
Ollscoil Chathair Bhaile Átha Cliath

Appendix D: Three Rounds of Coding

The three rounds of coding are available via the following link:

https://github.com/wryuer/PhD-thesis_SDH-subtitlers-interview-study_Three-rounds-of-coding

Appendices for Phase II

Appendix E: Plain Language Statement

Appendix E1: Plain Language Statement (English)

Introduction

This study, titled "Reading Sounds: Subtitling for the Deaf and Hard of Hearing in China", is part of the doctoral research carried out by Rongyu Wang (rongyu.wang7@mail.dcu.ie) at the School of Applied Language and Intercultural Studies, Dublin City University. This research is supervised by Dr. Ryoko Sasamoto and Dr. Joss Moorkens, and funded by Dublin City University's Faculty of Humanities and Social Sciences PhD Scholarship Programme.

What is this research about?

This research explores how deaf and hard-of-hearing individuals in China process videos with regular subtitles and subtitles designed for the deaf and hard of hearing.

Why is this research being conducted?

By analysing how participants interact with different subtitle formats, the findings will help researchers develop more inclusive subtitling practices and enhance the media experience for this community.

Why have I been invited to take part?

You have been invited to participate because you are a native Chinese speaker, are 18 years or older, and identify as deaf or hard of hearing.

What will happen if I decide to take part in this research study?

If you agree to participate, you will first complete a short questionnaire to provide background information. Then, you will watch a subtitled video while an eye tracker records your eye movements. After watching the video, you will complete a comprehension test related to the video's content. The entire process is expected to take approximately 40 minutes.

How will my data be used?

The data collected, including questionnaire responses and eye-tracking data, will be used solely for academic research purposes to analyse how participants engage with subtitled videos. The findings will contribute to a PhD thesis and may be included in future academic publications and presentations.

How will my privacy be protected?

This study complies with the General Data Protection Regulation (GDPR). Only the research team will have access to the collected data. All data will be securely stored and anonymised to protect your identity. Any published results will be presented in a way that ensures your privacy is maintained.

What are the benefits of taking part in this research study?

By participating in this research, you will potentially contribute to making audiovisual content more accessible to the hearing-impaired community in China. In addition, you will receive a gift as a token of appreciation for your time and effort.

What are the risks of taking part in this study?

This study presents minimal risks. The eye-tracking technology used is non-invasive and does not cause harm. If at any time you feel discomfort or fatigue, you may take a break or withdraw from the study without any consequences.

How will I find out what happens with this project?

If you would like to receive a summary of the study's findings or any relevant updates, you are welcome to contact the researcher by email: rongyu.wang7@mail.dcu.ie.

Can I change my mind at any stage and withdraw from this study?

Yes, participation in this study is entirely voluntary. You may withdraw at any time without providing a reason.

----- Contact details for further information -----

If you have concerns about this study and wish to contact an independent person, please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9.

Tel: 01-7008000; E-mail: rec@dcu.ie

Appendix E2: Plain Language Statement (Chinese)

介绍

本研究题为《看见声音：中国背景下的无障碍听障字幕研究》，是都柏林城市大学应用语言与跨文化研究学院王容宇（rongyu.wang7@mail.dcu.ie）进行的博士研究项目的一部分。研究由 Ryoko Sasamoto 博士和 Joss Moorkens 博士共同指导，并获得都柏林城市大学人文与社会科学学院博士奖学金资助。

本研究的内容是什么？

本研究旨在探讨中国听障人士在观看带有普通字幕与听障字幕的视听内容时的观看过程及其理解方式。

为什么要开展这项研究？

通过分析参与者与不同类型字幕之间的互动方式，本研究希望帮助研究人员更深入地理解听障观众的实际需求，从而推动更具包容性的听障字幕制作，提升听障人士的观看体验。

为什么邀请我参加？

您受邀参加本研究，是因为您满足以下条件：母语为中文，年满 18 周岁，且为聋人或重听人士。

如果我同意参加，需要做什么？

如果您同意参加本研究，您将首先填写一份简短的问卷，以提供基本背景信息。随后，您将观看一段带有字幕的视频。在观看过程中，眼动仪将记录您的眼动情况。观看结束后，您需要完成一份与视频内容相关的理解测试。整个过程大约需要 40 分钟。

数据将如何使用？

本研究收集的数据（包括问卷和眼动数据）将仅用于学术研究目的，主要用于分析参与者如何与字幕进行互动。研究结果将用于撰写博士论文，并可能在未来的学术文章和学术会议中发表。

我的隐私如何得到保护？

本研究遵守《通用数据保护条例》。只有研究团队成员可以访问所收集的数据。所有数据将被安全存储并做匿名化处理，以保护您的个人信息。在任何发表的研究结果中，均不会包含可识别个人身份的信息。

参与本研究有什么好处？

通过参与本研究，您将为改善中国听障人士在视听内容方面的无障碍体验作出贡献。为感谢您的参与，您还将获得一份小礼物。

参加本研究是否存在风险？

本研究风险极低。所使用的眼动技术为非侵入式，不会对您造成任何伤害。如在研究过程中感到不适，您可以随时休息或退出研究，且不会产生任何不良后果。

我如何了解研究结果？

如果您希望获得研究结果的摘要或任何相关的信息，欢迎通过电子邮件联系研究人员：rongyu.wang7@mail.dcu.ie。

我可以在任何时候退出研究吗？

可以。参与本研究完全出于自愿。您可以在任何时候选择退出，且无需说明任何理由。

----- **其他联系方式** -----

如您对本研究有任何疑问，并希望联系独立于研究团队的人员，请联系：

都柏林城市大学研究伦理委员会秘书处

地址：The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9.

电话：01-7008000

邮箱：rec@dcu.ie

Appendix F: Informed Consent Form

Appendix F1: Informed Consent Form (English)

Research title: Reading Sounds: Subtitling for the Deaf and Hard of Hearing in China

Principal investigators: Rongyu Wang, Dr. Ryoko Sasamoto, Dr. Joss Moorkens

Email: rongyu.wang7@mail.dcu.ie

Institution: School of Applied Language and Intercultural Studies, Dublin City University

- I have read the Plain Language Statement and understand the information provided.
- I have had the opportunity to ask questions and discuss this study.
- have received satisfactory answers to all my questions.
- I voluntarily agree to participate in this study and understand that I will be required to complete two questionnaires and watch a video.
- I consent to the collection and analysis of my questionnaire responses and eye-tracking data.
- I understand that all information I provide for this study will be securely stored and disposed of appropriately. I am also aware that the confidentiality of the information provided is subject to GDPR regulations.
- I understand that I have the right to request access to the thesis and publications related to this study in order to obtain the research findings.
- I am aware of the potential risks and benefits of this research study.
- I understand that I can withdraw from this study at any time and that I am not obligated to provide a reason for my withdrawal.

By signing this form, I confirm that I have read and understood the information provided. My questions and concerns have been addressed by the researchers, and I have received a copy of this consent form. Therefore, I consent to take part in this research project.

Participant's Name: _____

Participant's Signature: _____

Witness: _____

Date: _____

Appendix F2: Informed Consent Form (Chinese)

研究项目：看见声音：中国背景下的无障碍听障字幕研究

研究人员：王容宇、Ryoko Sasamoto 博士、Joss Moorkens 博士

电子邮件：rongyu.wang7@mail.dcu.ie

研究机构：都柏林城市大学应用语言与跨文化研究学院

- 我已阅读并理解本研究的《研究说明书》中所提供的信息。
- 我可以就本研究提出问题并进行讨论。
- 我所有的问题均已得到满意的答复。
- 我自愿同意参与本研究，并我知悉将需要完成两份问卷并观看一段视频。
- 我同意收集并分析我的问卷回答及眼动数据。
- 我知道我为本研究提供的所有信息将被安全存储并妥善处置。我也清楚，所提供的数据按照《通用数据保护条例》进行处理。
- 我知悉我有权查阅与本研究相关的论文及出版物，以获取研究结果。
- 我知悉本研究可能存在的风险与利益。
- 我知道我可以随时退出本研究，并无需说明退出理由。

签署此表格即表示我确认已阅读并知悉所提供的信息。我的疑问和顾虑已得到研究人员的解答，并且我已收到本同意书的副本。因此，我同意参与本研究项目。

参与者姓名： _____

参与者签名： _____

见证人： _____

日期： _____

Appendix G: Ethical Approval

Ollscoil Chathair Bhaile Átha Cliath
Dublin City University

DCU

Rongyu Wang
School of Applied Language and Intercultural Studies

16th April 2025

REC Reference: **DCUREC/2025/043 (Notification Review)**

Proposal Title: **Reading Sounds: Subtitling for the Deaf and Hard of Hearing in China**

Applicant(s): Rongyu Wang, Dr Joss Moorkens

Dear Colleague(s),

Thank you for your application to DCU Research Ethics Committee (REC). Further to its review by the committee and resting on the assumption of information accuracy and completeness, DCU REC is pleased to issue ethical approval for this research project. Please include reference to this approval in all materials used to recruit research participants.

Researchers are responsible for ensuring that the research project to which this ethical approval refers is carried out as specifically described in the application form. Should modifications to the research project be required at a later stage, researchers must submit a research amendment application form to REC for approval, prior to the implementation of modifications.

Please note that it is the responsibility of the PI to ensure that any other DCU compliance requirements relevant to the research project, such as those related to data protection, insurance, health and safety, or legal issues, are fully met in advance of initiating the project.

As part of DCU REC's ongoing monitoring process, a research progress report may be required. DCU REC will request this report from the PI as appropriate.

DCU REC wishes you every success in your research.

Yours sincerely,

DCU
Research & Innovation

Taighde & Nuálaíocht Tacaíocht
Ollscoil Chathair Bhaile Átha Cliath,
Baile Átha Cliath, Éire

Research & Innovation Support
Dublin City University,
Dublin 9, Ireland

T +353 1 700 8000
F +353 1 700 8002
E research@dcu.ie
www.dcu.ie

Dr Melrona Kirrane
Chairperson
DCU Research Ethics Committee

Note: Please retain this approval letter for future publication purposes. Research students should include this letter as a thesis appendix.

Appendix H: Pre-task Questionnaire

Appendix H1: Pre-task Questionnaire (English)

- **Basic information**

1. Gender:

A. Male

B. Female

2. Age: _____

3. Eyesight:

A. Normal vision (no need for glasses or contact lenses)

B. Myopia (below 5.0 dioptres)

C. Myopia (5.0 dioptres or above)

D. Other (please specify): _____

4. What is your hearing condition?

(Please refer to the last digit of the disability identification card)

A. Level 1 – Complete hearing loss: almost no hearing of any sound; relies entirely on hearing aids or cochlear implants.

B. Level 2 – Profound hearing loss: requires high-volume hearing aids or other devices; has extremely limited auditory comprehension; daily communication is very difficult.

C. Level 3 – Severe hearing loss: can understand normal speech with hearing aids in quiet environments, but has difficulty communicating in noisy settings.

D. Level 4 – Moderate hearing loss: can hear normal conversation, but has difficulty perceiving soft or distant sounds.

5. When did your hearing impairment begin?

A. At birth

B. Before age 2

C. Between ages 2–6

D. Between ages 7–12

E. Between ages 13–18

F. Between ages 19–29

G. After age 30

6. Do you currently use hearing aids or cochlear implants?

A. No

B. Yes, hearing aids

C. Yes, cochlear implants

D. Yes (please specify the type of device): _____

7. What is your primary mode of communication in daily life?

A. Sign language

B. Spoken language (with hearing aids/cochlear implants or lip reading)

C. A combination of sign language and spoken language

D. Text-based communication (e.g., typing, writing, or speech-to-text software)

E. Other (please specify): _____

- **Viewing experience**

8. How many hours do you watch videos per day on average?

A. Less than 1 hour

B. 1–2 hours

C. 2–4 hours

D. More than 4 hours

9. What types of video content do you watch most frequently?

A. News

B. Movies

C. TV series

D. Animation

E. Sports

F. Vlogs

G. Reality shows and variety shows

H. Other (please specify): _____

- **Subtitle preferences**

10. Do you find it difficult to read subtitles when watching videos?

A. Yes

B. No

C. Sometimes, but it does not affect me much

D. Sometimes, and I feel it affects me to some extent

11. Do you decide which videos to watch based on whether they have subtitles?

A. I only watch videos with subtitles

B. I tend to choose videos with subtitles when available

C. It doesn't matter. I watch both videos with and without subtitles

12. When watching videos, which accessibility method do you prefer?

A. Subtitles only

B. Sign language only

C. Both subtitles and sign language

D. Neither subtitles nor sign language

13. Have you ever used subtitles designed specifically for deaf and hard-of-hearing audiences (i.e., subtitles that include descriptions of sounds)?

A. Yes

B. No

14. Do you think subtitles designed specifically for deaf and hard-of-hearing audiences are helpful?

A. Yes

B. No

C. It depends on the situation

15. Besides dialogue, what additional types of information would you like subtitles to include? (Multiple choice)

A. Speaker identification (e.g., character names)

B. Descriptions of tone of voice

C. Descriptions of important sound effects

D. Descriptions of background music

E. Descriptions of silence or quiet moments

F. All of the above

Appendix H2: Pre-task Questionnaire (Chinese)

■ 基本信息

1. 您的性别:

A. 男性

B. 女性

2. 您的年龄: _____

3. 您的视力状况：

A. 不近视（不需要佩戴眼镜）

B. 近视（度数在 500 度以下）

C. 近视（度数在 500 度以上，含 500 度）

D. 其他（请说明）_____

4. 您的听力状况：（请参照《中华人民共和国残疾人证》，选择证号上的最后一位数字）

A. 数字是 1：一级（极重度听力损失）基本无法听到任何声音，必须依靠助听设备或人工耳蜗辅助。

B. 数字是 2：二级（重度听力损失）需要极高音量或助听设备辅助，听觉理解极为有限，日常交流非常困难。

C. 数字是 3：三级（中度听力损失）在安静环境中通过助听设备可以听到普通对话，但在嘈杂环境中交流困难。

D. 数字是 4：四级（轻度听力损失）可以听到普通对话，但对细小声音、远距离声音的感知有困难。

5. 您是从什么时候开始有听力障碍的？

A. 出生时

B. 2 岁前

C. 2–6 岁

D. 7–12 岁

E. 13–18 岁

F. 19–29 岁

G. 30 岁以后

6. 您目前是否使用助听设备？

A. 否

B. 是，使用助听器

C. 是，使用人工耳蜗

D. 是（请说明使用哪种助听设备）_____

7. 您日常交流主要使用哪种方式？

A. 手语

B. 口语（佩戴助听器/人工耳蜗或读唇语）

C. 手语和口语结合

D. 文字交流（如打字、写字、语音转文字软件等）

E. 其他（请说明）_____

■ 观看习惯

8. 您平均每天观看视频的时间是多少？

A. 少于 1 小时

B. 1–2 小时

C. 2–4 小时

D. 超过 4 小时

9. 您最常观看哪些类型的视频内容？

A. 新闻

B. 电影

C. 电视剧

D. 动画片

E. 体育赛事

F. Vlog（视频博客）

G. 真人秀和综艺节目

H. 其他（请说明）_____

■ **字幕偏好**

10. 在观看视频时，您是否觉得阅读字幕有困难？

A. 是

B. 否

C. 有时会，但影响不大

D. 有时会，觉得有一定影响

11. 字幕是否会影响您对视频的选择？

A. 只看带字幕的视频（没有字幕的不看）

B. 有字幕的视频会优先选择

C. 无所谓，有字幕和没字幕的视频都会看

12. 观看视频时，您更倾向于哪种无障碍方式？

A. 仅字幕

B. 仅手语

C. 字幕和手语同时存在

D. 既不使用字幕，也不使用手语

13. 您是否使用过专为听障人士设计的字幕（即带有声音描述的字幕）？

A. 是

B. 否

14. 您是否认为专为听障人士设计的字幕有帮助？

A. 是

B. 否

C. 视情况而定

15. 您希望字幕中除了对话外，还包含哪些额外信息？（可多选）

A. 角色人名的标注

B. 角色语气的描述

C. 重要音效的描述

D. 背景音乐的描述

E. 对静音或沉默的描述

F. 以上全部

Appendix I: Post-task Questionnaire

Appendix I1: Post-task Questionnaire (English)

- **Comprehension text**

1. What did the Crow do to the Boar in the scene?



A. Touch the Boar's nose

B. Ask for the Boar's name

C. Pull out the Boar's tooth

D. Slap the Boar very hard

2. Why did the Crow shout "Ow! What are you doing?"



A. He wants to sing

B. He is annoyed

C. He is very happy

D. He got hit

3. What did the Boar do at the end after making the bamboo flute?



A. Smelled it with his nose

B. Put it on his head

C. Wiped it clean with his hand

D. Blew into it

4. In this scene, who said the line “Ow! Bro! My bristles are going to fall off” and what was their tone?



A. The Crow, feeling shocked

B. The Boar, feeling pain

C. The Bear, feeling sad

D. The Bull, feeling happy

5. Who are the two characters having this conversation in the scene?



A. The Boss and the Storyteller

B. The landlord couple

C. The Bear and the Wolf

D. The Boar and the Crow

6. What was the Boar’s tone when he said, “Is Sun Wukong also a monster”?



A. Angry

B. Upset

C. Confused

D. Happy

7. In this scene, how did the Boar's mother know that his bottle was empty?



A. She looked at it

B. She shook it and heard the sound

C. The Boar told her

D. She shocked the bottle but heard nothing

8. When the Boar asked the Bear if the Crow could come back, how did Bear respond?



A. Crow will definitely come back

B. Crow cannot come back

C. Crow might not come back

D. Crow has already come back

9. When the Wolf was looking for the Boar outside the cave, how did his footsteps change?



A. Getting farther and farther

B. Getting faster and faster

C. Staying the same

D. Stopping for a moment, then continuing

10. Why did the Boar lie down on the ground at the end of the story?



A. He collapsed from running

B. He was waiting for his friends to return

C. He was struck by Sun Wukong and was knocked unconscious

D. He was too exhausted from patrolling and lay down to rest

- **Narrative engagement and enjoyment**

11. I felt completely immersed in the story while watching the video.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neither agree or disagree
E. Somewhat disagree
F. Disagree
G. Strongly disagree

12. My attention was rarely distracted by other things during the viewing.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neither agree or disagree
E. Somewhat disagree
F. Disagree
G. Strongly disagree

13. After the video ended, I kept thinking about its plot or characters.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neither agree or disagree
E. Somewhat disagree
F. Disagree
G. Strongly disagree

14. When the main character was happy, I felt happy too; when they were in trouble, I felt worried or sad.

A. Strongly agree
B. Agree
C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

15. I was emotionally moved by the ending of the story in this video.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

16. After watching the video, I had characters that I particularly liked and others that I disliked.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

17. The story made me reflect on my own experiences or feelings.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

18. I was eager to find out what would happen next in the story.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

19. I found the video interesting, and it made me feel relaxed and happy.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

20. The subtitles improved my overall viewing experience.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

21. The subtitles in the video helped me better understand the story.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

22. I would recommend this type of subtitled video to my friends or family.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

23. If there is a similar video in the future, I would choose to watch the version with this type of subtitles again.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

24. I was satisfied with how the video content, sound, and subtitles worked together.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

25. Overall, I am very satisfied with this viewing experience.

A. Strongly agree

B. Agree

C. Somewhat agree

D. Neither agree or disagree

E. Somewhat disagree

F. Disagree

G. Strongly disagree

Appendix I2: Post-task Questionnaire (Chinese)

■ 理解测试

1. 画面中，乌鸦怪想对小猪妖做什么？



A. 摸小猪妖的鼻子

B. 喊小猪妖的名字

C. 拔小猪妖的牙齿

D. 打小猪妖一巴掌

2. 乌鸦怪喊“啊 你干嘛”是因为什么？



A. 想唱歌

B. 很难过

C. 很开心

D. 被吓到了

3. 小猪妖做完竹笛后，最后做了什么？



A. 用鼻子闻

B. 放在头上

C. 用手擦干净

D. 吹了一口气

4. “大哥，毛都没了”这句话是谁说的，带有什么语气？



A. 乌鸦怪说的，他很吃惊

B. 小猪妖说的，他觉得好痛

C. 熊獾头说的，他很伤心

D. 大黑牛说的，他很开心

5. 画面中，这段对话是谁和谁在说话？



A.老板和说书人

B. 屋里的男主人和女主人

C. 熊教头和狼大人

D. 小猪妖和乌鸦怪

6. 小猪妖说“孙大圣也是妖怪”时，他是什么心情？



A. 生气

B. 难过

C. 疑惑

D. 开心

7. 小猪妖的妈妈怎么知道葫芦是空的？



A. 看了一眼

B. 闻了闻葫芦

C. 小猪妖告诉她的

D. 摇了摇葫芦，没听到声音

8. 小猪妖问熊教头“乌鸦怪还能回来吗”，熊教头是如何回答的？

A. 乌鸦怪一定会回来

B. 乌鸦怪不能回来了

C. 乌鸦怪不一定能回来

D. 乌鸦怪已经回来了



9. 狼大人在找小猪妖的时候，脚步声是什么样的？

A. 越走越远

B. 越走越快

C. 一直一样

D. 走着走着停了一下，又继续走



10. 最后，小猪妖为什么躺在地上？

A. 跑步摔倒了

B. 等朋友回来

C. 被孙悟空打中，晕倒在地

D. 巡逻太累了，躺在地上休息



二．沉浸感与愉悦感

■ 观影体验

11.我在观看过程中感到完全沉浸在故事的情节中。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

12. 在观看过程中，我的注意力很少被其他事情分散。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

13. 视频结束后，我还在想里面的故事或人物。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

14. 当视频里的人物遇到危险和困难时，我会担心或难过。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

15. 我被这个视频最后的结局感动了。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

16. 看完视频后，我有特别喜欢的角色，也有不喜欢的角色。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

17. 这个故事让我回想到自己之前的经历和感受。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

18. 我特别想知道这个故事接下来会如何发展。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

19. 我觉得这个视频很有意思，让我觉得愉快和放松。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

20. 视频中的字幕提升了我的观影体验。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

21. 视频中的字幕，帮我更好地看懂这个故事。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

22. 我愿意向朋友或家人推荐这种类型字幕的视频。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

23. 如果下次还有类似的视频，我还会选择观看带有这种字幕形式的版本。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

24. 我对视频内容、声音和字幕之间的配合感到满意。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

25. 总体而言，我对这次观看体验非常满意。

A. 非常同意
B. 同意
C. 有点同意
D. 既不同意也不完全同意
E. 有点不同意
F. 不同意
G. 非常不同意

Appendix J: Eye-tracking Ratios of Participants

Participant	SDH group/ST group	Eye-tracking ratio
01	SDH group	94%
02	**SDH group**	**73%**
03	SDH group	88%
04	SDH group	91%
05	SDH group	97%
06	SDH group	93%
07	SDH group	93%
08	SDH group	97%
09	SDH group	94%
10	**ST group**	**55%**
11	ST group	94%
12	**ST group**	**79%**
13	ST group	97%
14	ST group	83%
15	**ST group**	**78%**
16	**ST group**	**77%**
17	ST group	92%
18	SDH group	82%
19	SDH group	96%
20	SDH group	97%

21	SDH group	97%
22	**SDH group**	**33%**
23	SDH group	93%
24	SDH group	90%
25	ST group	98%
26	ST group	99%
27	ST group	92%
28	ST group	98%
29	ST group	95%
30	SDH group	98%
31	SDH group	95%
32	SDH group	99%
33	ST group	89%
34	ST group	98%
35	ST group	97%
36	ST group	93%
37	ST group	97%
38	ST group	89%

Note:

As explained in Section 5.5.3.3, participants with an eye-tracking ratio below 80% were excluded from the final analysis. These excluded cases are shown in bold in the table. As a result, data from 32 participants were retained for analysis.